



e-ISSN: 2147-8228

www.dergipark.org.tr/ijamec

Volume 08  
Issue 01

March, 2020

**Research Article****The Role of Dysphonia and Voice Recordings in Diagnosis of Parkinson's Disease****Gökçen ÇETİNEL<sup>a,\*</sup> , Elif SARICA DAROL<sup>b</sup> , Alhasan ALKHATIB<sup>c</sup>** <sup>a</sup> *Electrical and Electronics Eng., Sakarya University – Turkey*<sup>b</sup> *Neurology Department, Education and Research Hospital, Sakarya University - Turkey*<sup>c</sup> *Electrical and Electronics Eng., Sakarya University - Turkey*

## ARTICLE INFO

*Article history:*

Received 23 January 2020

Accepted 24 March 2020

*Keywords:*

Ensemble gentle boost

Jitter

Logistic regression

Parkinson's disease

Pitch

Shimmer

Weighted k-NN

## ABSTRACT

Parkinsonism is a syndrome that occurs as a combination of six cardinal signs; resting tremor, rigidity, bradykinesia, loss of postural reflex, flexion posture and freezing (motor block). Parkinson disease occurs with the loss of brain cells which are generating dopamine. The most important primary motor symptoms of Parkinson's disease are shaking of hands, slowness of movement, and speech changes. Sound changes are not recognized at the early stages of the disease while it becomes evident at the progressive stages. However, speech changes can be detected with some acoustic parameters. This study aims to detect Parkinson's disease by using voice recordings. In this study, 342 voice recordings that belong to 174 healthy subjects and 168 Parkinson's disease patients are used. 21 features are extracted from each voice record. The classification of subjects as healthy or with Parkinson disease is achieved by using logistic regression, k-nearest neighboring and ensemble gentle boost techniques. Furthermore, ten-fold and leave-one-out cross validation techniques are applied to improve the performance and reliability of the classifier. Sensitivity, specificity, maximum and average accuracy values are calculated to evaluate the success of the system. The obtained results show that the proposed system can be utilized by the neurologists to diagnose Parkinson's disease at its early stages.

This is an open access article under the CC BY-SA 4.0 license.  
(<https://creativecommons.org/licenses/by-sa/4.0/>)

**1. Introduction**

Parkinson's disease (PD) is a neurodegenerative disorder that leads to motor and cognitive dysfunctions. The most common form of parkinsonism is Idiopathic Parkinson's Disease (IPD), defined by James Parkinson in 1817 and diagnosed by the presence of at least three of the six cardinal signs (resting tremor, bradykinesia, rigidity) with asymmetric beginning [1,2].

IPD is the second most common neurodegenerative disease after Alzheimer's disease over the age of 60 years [3]. It is a multifactorial (e.g., genetic and environmental factors) disease and the probability of its occurrence increases by aging. Pathophysiology responsible for the occurrence of motor symptoms is; the degeneration of dopaminergic neurons in the pars compact of Substantia Nigra (SN) and Striatum, decreasing in dopamine release as a result [4]. However, many researchers believe that

Lewy bodies play a role in the pathophysiology of IPD. Braak reported that the accumulation of Lewy neurites in the caudal brainstem nuclei (e.g., the dorsal motor nucleus of the glossopharyngeus, vagus and anterior olfactory nucleus) began well before the loss of dopaminergic neurons in the SN [5,6]. This period is called premotor phase of PD and contains nonspecific findings such as joint pain, autonomic symptoms, depression, constipation, sleep problems and hoarseness [7]. That's why, Parkinson's disease has been described as a process with motor and non-motor symptoms.

In addition to cardinal motor findings, bulbar symptoms are often associated with PD. Bulbar symptoms (dysarthria, hypophonia, dysphagia, sialorrhea) are thought to be caused by oropharyngeal and laryngeal bradykinesia and rigidity [8]. It is reported that 90% of Parkinson's patients have such vocal problems in the early stages [9]. PD-related speech and voice disorder, often

\* Corresponding author. E-mail address: [gctinel@sakarya.edu.tr](mailto:gctinel@sakarya.edu.tr)  
DOI: 10.18100/ijamec.679038

referred to as hypokinetic dysarthria, is characterized by low volume (hypophonia), monotone voice and pitch (aprosodia), incorrectly pronounced consonants, pauses, and rapid speech propensity (tachyphemia) [10]. Low intensity sound in PD, partly due to the loss of muscle control and mass, is attributed to the shape of the bow of the vocal cords [11].

There is no definitive treatment and complete recovery in PD. The neurologists seek to improve the life quality of patients by performing dopamine replacement, symptomatic treatment and conventional deep brain stimulation. In recent years, scientists have turned to new neuroprotective treatment options to stop progression, but have not developed a clear curative method yet. Because of high treatment cost and disability of PD, non-invasive early diagnosis methods and machine learning techniques are gaining importance.

In recent years, several methods are suggested to early diagnose the PD. Some of these studies, that utilize voice recordings to detect the PD and related with the proposed study, can be summarized as follows: In [12], the authors aim to detect PD by taking 1208 voice recordings from 68 voluntaries. 28 features are extracted from each recording by using Praat software. Support vector machines (SVM) and cross-validation techniques are applied and the achieved accuracy percentage is 85%. In [13], 195 voice recordings are taken from 8 normal and 23 PD subject to determine the voice impairments that occur in PD subjects. The obtained accuracy is  $91.8\% \pm 2$ . In [14], 46 features are extracted from 100 voluntaries including 50 normal 50 PD subjects. Several classifiers are applied in this study and the highest accuracy is obtained by applying SVM as 85%. In the paper given in [15], it is aimed to detect early signs of PD through free-speech in uncontrolled background conditions. Random Forest (RF) and SVM methods are utilized to provide a reliable method to detect PD with a high accuracy percentage. The authors achieved 99% accuracy with RF and Leave-One Out (LOO) cross-validation techniques. In [16], the authors used tunable Q-factor wavelet transform in feature extraction step. They collected voice recordings of 252 subjects (188 PD and 64 healthy). The highest accuracy is 86% achieved by SVM classifier. In the study, a detailed analysis of signal processing techniques is performed for PD classification from voice recordings. Lahmiri and Shmuel consider the diagnosis of PD based on voice patterns. They aim to classify subjects as PD patients and healthy subjects. SVM with Bayesian optimization technique used for classification. The dataset contains voice recordings of 147 PD patients and 48 healthy control subjects. The maximum accuracy, sensitivity, and specificity values are 92.13%, 82.79%, and 95.27%, respectively [17]. Finally, in [18], two methods of vocal signal analysis have been proposed to evaluate dysarthria which is an anomalous condition in human speech and can be used to compare pathological

and healthy voices. They extracted acoustic and vowel metric features for 153 voice signals (60 PD, 54 multiple sclerosis (MS), and 39 healthy). The experimental results show that the extracted values can be considered reliable with good statistical significance to characterize PD and MS.

In the presented study, 342 voice samples are collected from 29 healthy subjects and 28 PD patients. 174 of 342 samples are recorded by the authors and the remaining 168 samples that belong to the PD patients are taken from UCI database. After the pre-processing step, 21 features are extracted from each recording. In this study, feature extraction step is performed without using Pratt software. Then, 10-fold and LOO cross validation methods are utilized to enhance the performance of the classifiers. In the classification phase three different classifiers are used. Logistic Regression (LR), Weighted k-NN (Wk-NN) and Ensemble Gentle Boost (EGB) techniques are performed to distinguish PD patients from healthy control subjects.

The paper is organized as follows. In Section 2, the main components of the proposed system are introduced. Section 3 gives the simulation results to show the performance of the system. Finally, Section 4 concludes the paper.

## 2. The Proposed System

In this section, the block diagram of the proposed system that is established to detect PD via voice recordings is illustrated. The main components of the system and utilized techniques are also described to better understand the study.

### 2.1. The Database

The database of the system is constructed from UCI 'Parkinson Speech Dataset with Multiple Types of Sound Recordings' database. This database includes 168 voice samples taken from 28 PD patients. The voice samples include sustained vowels 'a' and 'o'.

As mentioned before, the goal of the study is to distinguish the subjects as normal and with PD. So, the database of the study must also contain voice recordings of normal subjects. To construct a balanced database, 174 voice samples are taken from 29 normal subjects. Thus, the total number of voice samples is 342.

After database construction, the pre-processing step including noise reduction and silent part removing is performed for each voice sample. Noise reduction is performed by applying an average filter. By calculating discrete derivative, that represents the variation between consecutive samples, the salient part removing is implemented easily. In other words, the samples having minimum variations are discharged. The flowchart of the study is given in Figure 1.

**2.2. Feature Extraction**

The aim of applying the feature extraction step is to determine the parameters of each voice sample to diagnose PD from the voice recordings. The success of the feature extraction step directly affects the performance of the whole system. Furthermore, feature extraction decreases the computational complexity of the system. In the proposed study, 21 features are calculated for each voice sample and these features are given in Table 1.

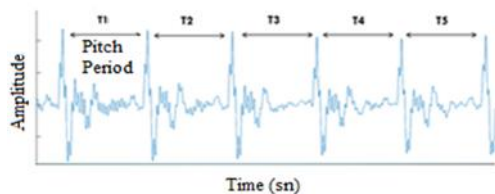
PD patients suffer from several speech impairments like dysphonia (defective use of the voice), hypophonia (reduced volume), monotone (reduced pitch range) and dysarthria (difficulty with the articulation of sounds or syllables). Among them, dysphonia is a commonly used speaking problem for the detection of PD. Dysphonia problem causes to reduced loudness, breathiness, roughness, decreased and exaggerated vocal tremor in voice. These impairments are the indicators of PD and can be detected by analyzing the voice samples. Thus, the features having given in Table 1 are calculated to detect the variations in voice samples that are originated from dysphonia.

**Table 1.** The extracted features for voice recordings.

Jitter	Shimmer	Pitch	Other
Jitter (Local)	Shimmer (Local)	Median Pitch	Autocorrelation (AC(T))
Jitter (Local-Absolute)	Shimmer (Local, dB)	Mean Pitch	Autocorrelation (AC(0))
Jitter (rap)	Shimmer (apq3)	Maximum Pitch	Harmonic-to-Noise Ratio (HNR)
Jitter (ppq5)	Shimmer (apq5)	Minimum Pitch	Noise-to-Harmonic Ratio (NHR)
Jitter (ddp)	Shimmer (apq11)	Standard Deviation Pitch	
	Shimmer (dda)	Pitch Range	



**Figure 1.** The flowchart of the study



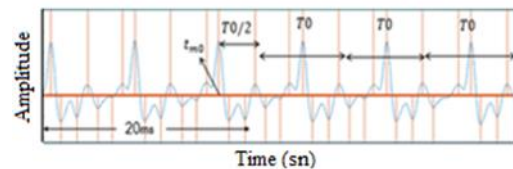
**Figure 2.** The voice pattern of a subject

At first, the pitch features are derived from the voice recordings. As discussed before, the voice recordings of the voluntaries include sustained vowels ‘a’ and ‘o’. In Figure 2, a sample voice recording of a subject is demonstrated. As can be seen from the figure, this is a periodic pattern and the distance between successive

maximum peak points is referred to as a pitch period. Even though the patterns of pitches are similar, their amplitudes and the duration between maximum amplitude values are different from each other. So, to obtain pitch features of the voice recordings T values should be written down correctly. Then the features given in Table 1 can be calculated easily. For this purpose, the following steps are applied:

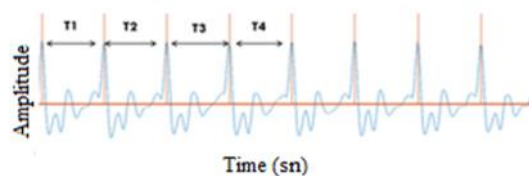
- i. Butterworth low-pass filter with 325 Hz cut-off frequency is used to eliminate the high-frequency components.
- ii. The highest amplitudes are detected.
- iii. The maximum point in the first 20 ms is found and the corresponding time point is pointed as  $t_{m0}$ .
- iv. Since the maximum point in every  $T_0$  interval ( $t_{m0} + \frac{T_0}{2}$ ) is obtained until the end of the signal. In this way, the T sequence is derived.

The procedure of building the T sequence is illustrated in Figure 3.



**Figure 3.** Determination of the T sequence

Note that, to apply the steps given above the  $T_0$  value should be known.  $T_0$  is the mean of the T sequence and  $F_0=1/T_0$  is the fundamental frequency. So, we need an algorithm to calculate fundamental frequency  $F_0$  without knowledge of T sequence. In the proposed study, to meet this requirement, zero-crossing algorithm is used. To enhance the accuracy of the zero-crossing algorithm the amplitude values corresponding to the T sequence elements are checked. If the amplitude value is 75% lower than the previous amplitude value, this value is deleted from the sequence. After all steps, the achieved T sequence for a voice sample is shown in Figure 4.



**Figure 4.** A sample T sequence on a voice recording

In several studies as reference [4], the Praat software is used to obtain the T sequence and then to calculate the features given in Table 1. The obtained values with the proposed method and Praat software are compared in Table 2. According to the comparison, the success of the improved zero-crossing method is verified. Hence, the features given in Table 1 can be calculated. The equations used to derive Jitter features are expressed as below:

$$\text{Jitter(Local)} = \frac{\frac{1}{N-1} \sum_{i=1}^{N-1} |T_i - T_{i-1}|}{\frac{1}{N} \sum_{i=1}^N T_i} * 100 \quad (1)$$

$$\text{Jitter(Local - Absolute)} = \frac{1}{N-1} \sum_{i=1}^{N-1} |T_i - T_{i-1}| \quad (2)$$

$$\text{Jitter(rap)} = \frac{\frac{1}{N-1} \sum_{i=1}^{N-1} |T_i - \frac{1}{3} \sum_{k=i-1}^{i+1} T_k|}{\frac{1}{N} \sum_{i=1}^N T_i} * 100 \quad (3)$$

$$\text{Jitter(ppq5)} = \frac{\frac{1}{N-1} \sum_{i=2}^{N-2} |T_i - \frac{1}{5} \sum_{k=i-2}^{i+2} T_k|}{\frac{1}{N} \sum_{i=1}^N T_i} * 100 \quad (4)$$

$$\text{Jitter (ddp)} = \frac{\sum_{i=1}^{2N-1} |(T_{i+1} - T_i) - (T_i - T_{i-1})|}{N - 2} \quad (5)$$

where T is the pitch periods sequence and N is the length of T sequence vector [19, 20]. Shimmer features that indicate the vocal tremor can be expressed as follows

$$\text{Sh(Local)} = \frac{\frac{1}{N-1} \sum_{i=1}^{N-1} |A_i - A_{i-1}|}{\frac{1}{N} \sum_{i=1}^N A_i} * 100 \quad (6)$$

$$\text{Sh(Local, dB)} = \frac{1}{N-1} \sum_{i=1}^{N-1} |20 * \log(\frac{A_{i+1}}{A_i})| \quad (7)$$

$$\text{Sh(apq3)} = \frac{\frac{1}{N-1} \sum_{i=1}^{N-1} |A_i - \frac{1}{3} \sum_{k=i-1}^{i+1} A_k|}{\frac{1}{N} \sum_{i=1}^N A_i} * 100 \quad (8)$$

$$\text{Sh(apq5)} = \frac{\frac{1}{N-1} \sum_{i=2}^{N-2} |A_i - \frac{1}{5} \sum_{k=i-2}^{i+2} A_k|}{\frac{1}{N} \sum_{i=1}^N A_i} * 100 \quad (9)$$

$$\text{Sh(apq11)} = \frac{\frac{1}{N-1} \sum_{i=5}^{N-5} |A_i - \frac{1}{11} \sum_{k=i-5}^{i+5} A_k|}{\frac{1}{N} \sum_{i=1}^N A_i} * 100 \quad (10)$$

$$\text{Sh(dda)} = \frac{\sum_{i=1}^{2N-1} |(A_{i+1} - A_i) - (A_i - A_{i-1})|}{N - 2} \quad (11)$$

The  $A_i$  values are the maximum peak-to-peak values having shown in Figure 5. In addition to these features, four features derived in the time domain and given in the last column of Table 1. These features are related to autocorrelation of a signal. AC (0) is the autocorrelation coefficient at the origin consisting of all energy of the signal and AC (T) is the component of autocorrelation that belongs to the fundamental period of the signal. The difference between the energy of signal and the fundamental period of the signal is the noise energy. Harmonic-to-noise ratio (HNR) is the ratio between the periodic components (harmonic) and aperiodic components (noise) of the voice signal and gives information about the periodicity of the signal [21]. The HNR and NHR features are expressed as follows

$$\text{HNR} = 10 * \log(\frac{\text{AC}(T)}{\text{AC}(0) - \text{AC}(T)}) \quad (12)$$

$$\text{NHR} = \frac{1}{\text{HNR}} \quad (13)$$

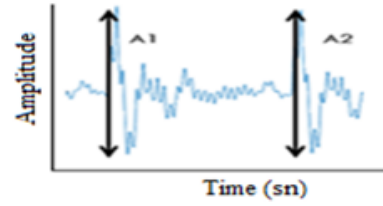


Figure 5. The Shimmer parameters on a voice sample

Table 2. The comparison of Praat software and proposed method.

Feature values for two subjects	Our algorithm	Praat software
Jitter (sec.) (local-absolute)	0.0000335	0.0000318
Jitter (local)(%)	0.00404	0.00392
Jitter rap(%)	0.00218	0.00208
Jitter ppq5(%)	0.00236	0.00228
Jitter (sec.) (local absolute)	0.0000228	0.0000217
Jitter (local)(%)	0.00461	0.00442
Jitter rap(%)	0.00270	0.00256
Jitter ppq5(%)	0.00312	0.00302

### 2.3. Classification

After feature extraction, the feature matrix with size  $342 \times 22$  is obtained. The number of rows is determined by the number of subjects and the column number is the sum of feature number and labels. Label “1” is assigned to normal subjects and “0” is used for the patients with PD. So, the performed classification task is a two-class classification problem and can be solved with available machine learning techniques.

Before applying classification techniques, cross validation methods are used to increase the performance of the system. In the k-fold cross validation process, the dataset is divided into k groups. Among them, (k-1) groups are used for training and the remaining group is the test group. After k experiment achieved, the test results are averaged and final performance is calculated. In the proposed study, 10-fold and LOO cross-validation techniques are carried out to improve the reliability of obtained classification results. The Logistic Regression (LR), Weighted-kNN (Wk-NN) and Ensemble Gentle Boost (EGB) classifiers are utilized for the given classification task.

LR is commonly preferred for linear classification problems. LR is a statistical method and can only be used for two-class classification tasks. The main difference between linear and logistic regression is that LR is used when the dependent variable is binary in nature. In contrast, linear regression is used when the dependent variable is continuous and the nature of the regression line is linear.

Wk-NN classifies the query according to the class labels

of its neighbors. In k-NN the weight of each neighbor is equal. However, in Wk-NN the impact of the nearest neighbor on the decision is the highest. In this study, eleven neighbors are chosen according to the Euclidean metric.

EGB is a community classifier in which the dataset is divided into subsets and for each subset different classification method is applied. The final decision is made according to the weighted average of the applied classification methods' results [22]. The weights of the classifiers are updated at each iteration.

### 3. Results and Discussions

The goal of the proposed study is to diagnose PD via voice recordings. The main idea behind the study is to detect speech impairments (dysphonia) that is resulted from PD. To recognize these impairments, the voice recordings of voluntaries are analyzed. 21 features are extracted from the jitter, shimmer and pitch patterns of the voice after a preprocessing step. The feature extraction step is followed by a cross validation technique to enhance the reliability of the system. Finally, three classifiers are performed to distinguish the subjects as "healthy" or "PD patient". The sensitivity, specificity, maximum and average accuracy values are calculated to evaluate the performance of the suggested system. In Table 3, the confusion matrix having constituted to calculate these performance metrics is given. In Table 3, the number of TP, TN, FP and FN subjects are also denoted for the classifier that provides the best classification performance. According to these values we can say that the accuracy, sensitivity and specificity percentages of the proposed system will be pretty high.

**Table 3.** Confusion matrix.

Confusion Matrix		Correct Classes	
		Positive	Negative
Predicted Classes	Positive	True Positive (TP) (152)	False Positive (FP) (12)
	Negative	False Negative (FN) (16)	True Negative (TN) (162)

Sensitivity is the ability of the decision system to

correctly identify those with the disease. Sensitivity (true positive rate) can be calculated as  $TP/TP+FN$ . If this value is close to its best value "one", the number of false negatives goes to zero. Specificity is the ability of the decision system to correctly identify those without disease. It is referred as to true negative rate and can be calculated as  $TN/TN+FP$ . If the number of false positives approaches to zero, the value of the specificity becomes closer to its highest value "one". Finally, accuracy is the proportion of correct decisions to all decisions. It can be expressed as  $(TP+TN)/(TP+TN+FP+FN)$ . A medical diagnosis systems having accuracy percentage higher than 80% can be used by the specialists for clinical experiments.

In our study, the achieved results are given in Table 4. According to the table, the LOO cross validation technique can increase the performance of the classifiers. The highest accuracy, specificity and, sensitivity percentages among the classifiers are 92.86%, 93.10% and 91.07%, respectively. Furthermore, the variance of accuracy is low. The low variance assures the stability of the system. Among all classifiers EGB performs the given task with higher accuracy, sensitivity and specificity percentages. Finally, in Table 5, the achieved accuracy, sensitivity and specificity values are compared to those of the existing similar studies. As can be seen from the table, the proposed study provides the highest sensitivity percentage among the considered studies. Furthermore, the accuracy and specificity values of the system are so high.

The performed simulations and achieved results show that the proposed system is a promising PD diagnosis system in which the speech impairments are detected by using signal processing and machine learning techniques. Hence, this system can be a decision-support system for the neurologists.

In future studies, it is planned to maximize the performance of the system by defining new features and using deep learning techniques.

### 4. Conclusion

PD is the second most common neurodegenerative disease after Alzheimer disease over the age of 60 years. PD degrades the life quality of the patients by causing shaking of hands, slowing the movement and disturbing speech. So, in recent years, the neurologists have turned to new neuroprotective treatment options to stop progression by early diagnosing the PD.

**Table 4.** The performances of the classifiers

	Logistic regression		Weighted k-NN		Ensemble Gentle Boost	
	10-fold	LOO	10-fold	LOO	10-fold	LOO
Mean Sensitivity (%)	84.52	85.12	83.69	83.33	89.76	91.07
Mean Specificity (%)	85.63	85.63	90.57	92.53	92.64	93.10
Maximum Accuracy (%)	85.96	85.38	87.72	88.01	92.86	92.11
Mean Accuracy (%)	85.08	85.38	87.19	88.01	91.22	92.11

**Table 5.** Comparison table.

	Average Accuracy (%)	Average Sensitivity (%)	Average Specificity (%)
Ref [12]	85.00	80.00	90.00
Ref [17]	92.13	82.79	95.27
Ref [23]	86.47	88.91	84.02
The Proposed	92.11	91.07	93.10

In this study, we aim to detect PD by using some acoustic parameters having derived from voice recordings. For this purpose, a database including 342 voice recordings (174 healthy and 168 PD) is constructed. After a preprocessing step, 21 features are extracted for each voice recordings. The classification of subjects as healthy or with Parkinson is achieved by using logistic regression, k-nearest neighboring and ensemble gentle boost techniques. Furthermore, ten-fold and leave-one-out cross validation techniques are applied to improve the performance and reliability of the classifier. Sensitivity, specificity, maximum and average accuracy values are calculated to evaluate the success of the system. The obtained results show that the proposed system can be utilized by the neurologists to diagnose Parkinson's disease at its early stages.

## References

- [1] J. Parkinson, "An essay on the shaking palsy" London, Sherwood, Neely and Jones, 1817.
- [2] D. J. Gelb, E. Oliver, S. Gilman, "Diagnostic criteria for Parkinson's disease", *Arc. Neurology*, vol. 56, pp. 33-39, 1999.
- [3] Rijk M.C. *et al.*, "Prevalence of Parkinson's disease in Europe: A collaborative study of population-based cohorts", *Neurology*, vol. 54, pp. 21-23, 2000.
- [4] H. Ehringer, O. Hornykiewicz, "Distribution of noradrenalin and dopamine in the human brain and their behaviour in disease of extrapyramidal system", *Klin Wochenschr*, vol. 38, pp. 1236-1239, 1960.
- [5] H. Braak *et al.*, "Staging of brain pathology related to sporadic Parkinson's disease", *Neurobiology Aging*, vol. 24, pp. 197-211, 2003.
- [6] H. Braak *et al.*, "Stages in the development of Parkinson's disease-related pathology", *Cell Tissue Res.*, vol. 318, pp. 121-134, 2004.
- [7] H. Przuntek, "Early diagnosis in Parkinson's disease", *J. Neural Trans. Supp.*, vol. 38, pp. 105-114, 1992.
- [8] C. J. Hunker, J. H. Abbs, S. M. Barlow, "The relationship between parkinsonian rigidity and hypokinesia in the orofacial system: A quantitative analysis", *Neurology*, vol. 32, pp. 749-754, 1982.
- [9] S.B.O. Sullivan, T.J. Schmitz, G. Fulk, "Physical Rehabilitation" 5th ed., FA Davis Company, USA, 2013.
- [10] S. Fahn, J. Jankovic, M. Hallett, M. Jenner M, P. Anner, "Principles and practice of movement disorders", Churhill Livingstone, Chapter 2, Part 4, pp. 81, 2007.
- [11] G. M. Schulz, T. Peterson, C. M. Sapienza, "Voice and speech characteristics of persons with Parkinson's disease pre and post-pallidotomy surgery: Preliminary findings". *J Speech Lang. Hear Res.*, vol. 42, pp. 1176-1194, 1999.
- [12] B. Erdogdu Sakar, M. Isenkul, C.O. Sakar, A. Sertbas, F. Gurgun, S. Delil, H. Apaydin, and O. Kursun, "Collection and Analysis of a Parkinson Speech Dataset with Multiple Types of Sound Recordings", *IEEE Journal of Biomedical and Health Informatics*, vol. 17(4), 2013, pp. 828-834.
- [13] M.A. Little, P.E. McSharry, S.J. Roberts, D.A.E. Costello, and, I.M. Moroz, "Exploiting Nonlinear Recurrence and Fractal Scaling Properties for Voice Disorder Detection", *Biomedical Engineering Online*, vol. 6 (23), 2007, pp. 1-19.
- [14] Y. Campos-Roca, D. Montà, and C. Pérez, "Parkinson's Disease Detection Based on a Heterogeneous Acoustic Database", *EECCSS*, Barcelona, Spain, July 13-14, 2015.
- [15] D. Braga *et al.*, "Automatic detection of Parkinson's disease based on acoustic analysis of speech" *Eng. Applications of AI*, vol. 77, pp. 148-158, 2019.
- [16] C. O. Sakar *et al.*, "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 Journal*, vol. 74, pp. 255-263, 2019.
- [17] S. Lahmiri, A. Shmuel, "Detection of Parkinson's disease based on voice patterns ranking and optimized support vector machine", *Biomedical Signal Processing and Control*, vol.49, pp. 427-433, 2019.
- [18] P. Vizza *et al.*, "Methodologies of speech analysis for neurodegenerative diseases evaluation", *International Journal of Medical Informatics*, vol. 122, pp. 45-54, 2019.
- [19] J. Teixeira, "Algorithm for jitter and shimmer measurement in pathologic voices", *Procedia Computer Science*. vol.100, pp. 271-279, 2016.
- [20] P. Boersma and D. Weenink, "Praat: doing phonetics by computer", *Phonetic Sciences*, University of Amsterdam. <http://www.fon.hum.uva.nl/praat/>.
- [21] J. Teixeira and P. Fernandes, "Jitter, shimmer and HNR classification within gender, tones and vowels in healthy voices", *Procedia Technology*. vol. 16, pp. 1228-1237, 2014.
- [22] J. Han, "Data mining: Concepts and techniques", 3rd ed., Waltham, Mass.: Morgan Kaufmann Publishers, 2012.
- [23] L. Berus, S. Klancnik, M. Brezocnic, M. Ficko, "Classifying Parkinson's disease based on acoustic measures using artificial neural networks", *Sensors*. vol. 19, pp. 1-15, 2019.