© 2018 IJSRST | Volume 4 | Issue 5 | Print ISSN: 2395-6011 | Online ISSN: 2395-602X
Themed Section: Science and Technology

A Review On Parkinson's Disease Diagnosis Through Speech

Amreen Saifer[1], Dr. S. M. Ali[2]
[1]Student, Electronics and Communication, RTMNU, ACET, Nagpur, Maharashtra, India
[2]Associate Professor , Electronics and Communication, RTMNU, ACET, Nagpur, Maharashtra, India

ABSTRACT

This paper gives detailed description of Parkinson's disease (PD) and systematic literature review on Parkinson's disease severity assessment methods based on speech impairment. Parkinson's disease is the most common disease of motor system degeneration that occurs when the dopamine-producing cells are damaged in substantia nigra. EEG, gait and speech are the various signals used to detect PD, these signals was also been investigated. Since approximately 90 percent of the people with PD suffer from speech disorders, speech analysis is considered as the most common technique for diagnosing. Researchers proposes various algorithm for diagnosing of Parkinson's disease based on voice analysis. Viz. Support vector machine (SVM), Genetic Algorithm, Artificial Neural Network.

Keywords: Parkinson's disease; Speech Analysis; Genetic Algorithm; Support Vector Machine.

I. INTRODUCTION

Parkinson's disease is the second most common neurodegenerative disease, affecting 6.2 million people globally. It is most common in the elderly, though 5-10% of diagnoses reflect early-onset Parkinson's disease. For some patients, symptoms may begin as early as age 20. Early warning signs of Parkinson's disease usually involve impairment of movement, including uncontrollable shaking, rigidity, difficulty walking, and unsteady gait. As the disease progresses, cognitive and behavioral complications may arise, commonly leading to dementia. Effective treatment and early diagnosis for Parkinson's disease are hindered by a lack of quantifiable biomarkers and objective measures of disease progression. As there are no diagnostic lab tests available for Parkinson's disease, the current gold-standard for diagnosis relies on an in-clinic neurological test and brain scans to rule out other neurological causes of symptoms. This process is extremely costly and requires a high level of expertise, placing stress on existing medical infrastructure. With improving life expectancies in developing countries and an aging population in many developed countries, early and accurate diagnosis of Parkinson's disease will undoubtedly pose an increasing challenge for current healthcare systems. Researchers have proposed a new model for diagnosis of Parkinson's based on speech recognition and machine learning technologies. Introduced by Dr. Max Little, Chairman of Parkinson's Voice Initiative, this approach requires only a single sound recording of sustained phonation (saying "aaah") from patients. Voice processing tools subsequently analyze the sound recordings and compare them to a database of recordings of Parkinson's patients and non-Parkinson's patients that serve as a control. The algorithm developed by this research team is able to detect specific variations in sound vibrations linked to vocal tremors, breathlessness, and weakness. By detecting such voice changes that are indicative of neurological degeneration, the algorithm is able to generate accurate diagnoses and predict disease progression based on the presence and severity of such degenerative symptoms.

II. METHODOLOGY

The algorithm used for detecting normal and Parkinson's patients consists of the following steps:

1) Data Acquisition,
2) Windowing,
3) Preprocessing,
4) Feature Extraction,
5) Classification and
6) An optional decision scheme.

A. Data Acquisition

For the data acquisition step, the patients were tested by asking patient to say something like "aaah" or saying some vowels in natural manner while being recorded.

B. Windowing

In windowing process the signal was divided into consecutive windows of 20ms length with 50% overlaps. This setup was considered to be suitable for real time analysis applications of speech, while, on the other hand it contains sufficient data on the phonetic level.

C. Preprocessing

During the preprocessing stage, the signal was offset to be set to a mean of zero i.e. "removing the DC", the amplitude of the signal obtained from the windows were then normalized and then filtered to remove frequencies not in the speech range.

D. Feature Selection and Extraction

Special consideration was given for the selection of features. On the one hand, those features need to evaluate the degree of anomalous fluctuations in speech, but on the other hand, researcher wished to use standard features that are usually used in applications of speech analysis and recognition systems. Making the extraction process simpler, this also allows the researcher to easily integrate in existing systems. The ones chosen were as follows:

1) Pitch value and its power:

This feature represents the vibration rate of audio signals, which can be represented by the fundamental frequency and multiples thereof. The average pitch frequency time pattern, gain, and fluctuations change from one speaker to another. The values were calculated using an auto-correlation algorithm.

2) Short-time Energy:

The short-time energy (En) of speech signals reflects the amplitude variation, and is defined by the following equation:

$$E_n = \sum_{m=-\infty}^{\infty} x^2(m) \cdot h(n-m)$$

where h(n) is chosen to be a hamming window powered by 2. In voiced (periodic) speech the short time energy values are much higher than during the unvoiced speech.

3) Zero Crossing Rate (ZCR):

The zero-crossing rate of a short time window defined as a number of times the audio waveform changes its sign in the duration of the frame:

$$ZCR = \frac{1}{2}\sum_{n=1}^{N-1}\left|\operatorname{sgn}\left(x[n]\right) - \operatorname{sgn}\left(x[n-1]\right)\right|$$

where x(n) is the time domain signal for window t. This feature can indicate regarding the amount of noise in the speech signal, i.e. the periodicity of the signal.

4) Mean and Standard Deviation values of Zero Crossing Rates:

These values are computed using the statistics of the time intervals between consecutive zero crossings. Together with the ZCR feature these values can indicate about the speech abnormalities or 'noisiness' in different levels during the production of voiced and unvoiced sounds.

5) Mel Frequency Cepstral Coefficients (MFCC):

After computing the logarithm of the magnitude spectrum (computed by the Short Time Fourier Transform), and grouping the Discrete Fourier Transform (DFT) bins according to a Mel frequency scale (a logarithmic scale which approximates the response of the human auditory system), a discrete cosine transform is performed on the result. The first three coefficients (out of 26 coefficients) were used in this work.

E. Classification

For the classification stage, a Support Vector Machine (SVM) implementation of the algorithm was used with C-SVC (Support Vector Classification) kernel type and Radial Basis kernel function in order to do a 2-class classification of the data. SVM was chosen because it is known to perform well on generalization even with small amounts of data. Grid search methodology with a polynomial scale was applied in order to determine the free parameters ('C' parameter of the SVC and 'gamma' parameter of the RBF kernel function). Twenty four percent of data windows were randomly chosen without repetition from each patient (twelve percent for each of the two classes) for the training procedure each time. Then the selected data was oversampled in order to avoid learning on imbalanced data groups (due to the different number of patients in each group). After the best SVM parameters were found a cross validation procedure was applied in order to report on the results.

F. Decision scheme

The learning procedure implemented treats all windows equivalently without regard to their temporal origin. The researcher expect that further analysis will use the temporal arrangement which could be averaged in some sort of hierarchal fashion (a simple mechanism could be majority voting, but more complex methodologies are also possible).

SHAKIBI [34] propose a new algorithm for detection of PD based on genetic algorithm and SVM network. In first part, the strategy to select optimized features with genetic algorithm was described. In second part, SVM network and reasons that why it used for classification was explained.

III. GENETIC ALGORITHM:

Genetic Algorithm (GA) is an adaptive heuristic search algorithm premised on the evolutionary ideas of natural selection and genetic. It is one of the most influential methods in the process of data classification, which is effectively used to select optimized features. In genetic algorithm, the solution is called chromosome or string. This method requires a population of chromosomes (strings) representing a combination of features from the solution set, and requires a cost function (called an evaluation or fitness function). This function calculates the fitness of each chromosome. The algorithm manipulates a finite set of chromosomes (the population), based loosely on the mechanism of evolution. In each generation, chromosomes are subjected to certain operators, such as crossover, inversion and mutation, which are analogous to processes, which occur in natural reproduction. Crossover of two chromosomes produces a pair of offspring chromosomes, which are synthesis of the traits of their parents. Mutation of a chromosome produces a nearly identical chromosome with only local alternations of some regions of the chromosome. The optimization process is performed in cycles called generations. During each generation, a set of new chromosomes is created using crossover, inversion, mutation and other operators. Since the population size is fixed, only the best chromosomes were allowed to survive to the next cycle of reproduction. The crossover rate usually assumes to be of quite high value (on the order of 80%), while the mutation rate is small (typically 1% - 15%) for efficient search. The cycle repeats until the population "converges", that is all the solutions are reasonably the same and further exploration seems useless, or until the answer is "good enough".

Strategy for Selecting Optimized Features:

The process of running this algorithm in order to select the optimized pattern (feature) is explained below.

1. Calculate each pattern's entropy by using Equation (1) and output (target) vector's entropy by using Equation (2)

$$H(X) = -\sum_{i=1}^{n} p(x_i) \log p(x_i) \quad (1)$$

$$H(Y) = -\sum_{i=1}^{n} p(y_i) \log p(y_i) \quad (2)$$

where x is the vector of features and y is the vector of targets, p(x) and p(y) respectively density probability function of features and targets.

Measure mutual information between each pattern and every single output (target) via Equation (3)

$$I(X;Y) = H(X) - H(X|Y) = H(Y - H(Y|X))$$
$$= H(X) + H(Y) - H(X,Y) = H(X,Y) - H(X|Y) - H(Y|X) \quad (3)$$

In Equation (3), the patterns' entropy (H(X)), the target vector's entropy (H(Y)) and H(X,Y) are calculated by using Equation (4)

$$H(X,Y) = \sum_{i=1}^{N} \sum_{k=1}^{N} -\log\left(p(x_i, y_k)\right) p(x_i, y_k) \quad (4)$$

And

$$H(Y|X) = H(X,Y) - H(X) \quad (5)$$

And ultimately H(Y|X) is measured via Equation (6)

$$H(Y|X) = \sum_{x \in X} p(x) H(Y|X = x) = \sum_{x \in X} p(x) \sum_{y \in Y} p(y|x) \log\left(\frac{1}{p(y|x)}\right)$$
$$= -\sum_{x \in X} \sum_{y \in Y} P(x,y) \log(y|x) = -\sum_{x \in X, y \in Y} P(x,y) \log(y|x) \quad (6)$$

2. Initial population of genetic algorithm was produced randomly using 200 × n chromosomes, n was the number of features that need to be selected. Thus, each chromosome consists of n genes where the feature's number was placed randomly and it was possible for the feature number to be repeated randomly in a chromosome.

3. Measure the amount of relevance between patterns and targets for each chromosome using Equation (7)

$$V = \frac{1}{n} \sum_{i=1}^{n} I(X_i; Y_i) \quad (7)$$

where I is the mutual information between features and targets. The amount of redundancy among patterns and targets was measured for each chromosome using Equation (8)

$$P = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} I(X_i, Y_j) \quad (8)$$

4. The fitness value to each chromosome was assign via Equation (9)

$$\varphi = V - P \quad (9)$$

Purpose of the suggested genetic algorithm was to maximize the fitness function of Equation (9).

5. The chromosomes was rearranged according to the given fitness function.

6. Then select elite chromosomes as a parent.

7. Apply crossover and mutation and produce a new population. The chromosomes which can maximize the fitness function will remain and the rest will be removed and then

Steps 1 - 5 were repeated and this process continues as long as the changes in chromosomes fitness is less than 0.02 or the algorithm reaches the predetermined number of iterations which is supposed to be 80. Finally, the chromosome with the maximum fitness was chosen and the number of features in that chromosome was considered as selected features.

Support Vector Machine:

SVMs attempt to construct an optimal separating hyper plane in the feature space, between the two classes in the binary decision problem by maximizing a geometric margin between points from the two classes. In practical applications, data often cannot be linearly separated; in those cases, SVMs can use the kernel trick to transform the data into a higher dimensional space, and construct the separating hyper plane in that space. There was extensive research, beyond the scope of this study, on how to work with nonlinearly separable data. In general, this classifier requires the specification of some internal parameters, and SVMs are known to be particularly sensitive to the values of these parameters. Here, the researcher had used the LIBSVM implementation and followed the suggestions of the developers of that implementation:

They linearly scaled each of the input features to lie in the range [–1, 1], and used a Gaussian, radial basis function kernel. The determination of the optimal values of the kernel parameter γ and the penalty parameter C was decided using a grid search of possible values. They selected the pair (C, γ) that gave the lowest CV misclassification error. Specifically, they searched over the grid (C, γ) defined by the product of the sets

C = [2 5 , 2–13, . . . , 215], and

γ = [2–15 , 2–13, . . . , 23].

Once the optimal parameter pair (C, γ) was determined, they trained and tested the classifier using these parameters.

Dataset:

The dataset was created by Max Little of the University of Oxford, in collaboration with the National Centre for Voice and Speech, Denver, Colorado, who recorded the speech signals. The first study published the feature extraction methods for general voice disorders. The data consists of 190 sustained vowel phonations from 35 male and female subjects, of which 23 were diagnosed with PD. The time since diagnoses ranged from 0 to 28 years, and the ages of the subjects ranged from 46 to 85 years (mean 65.8, standard deviation 9.8). Averages of six phonations were recorded from each subject, ranging from one to 35 seconds in length. The phonations were recorded in an IAC sound-treated booth using a head-mounted microphone (AKG C420) positioned at 9 cm from the lips. The voice signals were recorded directly to computer using CSL 4300B hardware (Kay Telemetries), sampled at 44.1 kHz, with 16 bit resolution. Although amplitude normalize action affects the calibration of the samples, the study was focused on measures insensitive to changes in absolute speech pressure level. Thus, to ensure robustness of the algorithms, all samples were digitally normalized in amplitude prior to calculation of the measures

Related Work:

In past years the detection of voice disorders with the help of machine learning turned into a hot topic. Various researchers have attempted to solve this problem by considering acoustic measurements of dysphonia as effective features to distinguish normal (control) from disordered cases [7, 8, 13,14].

Studies in this field can be divided into two main groups:

(1) those that attempt to find the most effective voice features and produce new datasets [8, 13, 15] and

(2) those that try to find more effective features from existing datasets and work on enhancing classification accuracy [14, 16–26].

Some studies focused on how to produce new datasets based on their research findings.

Little et al. in [8] focused to analyze the effectiveness of nonstandard measurements. Their work led to the introduction of a new dysphonia measurement named as PPE (pitch period entropy). In their research, they had gatherd sustained vowel “a” phonations from 35

subjects of which 25 were PD patients and they reached the classification accuracy of 91.4%.

In [13], Sakar et al. presented a dataset of 40 subjects including 20 PD. Each person was trained to say a set of 26 different disorder representative terms consisting of sustained vowels, words, numbers, and small sentences. They applied summarized leave-one-out (s-LOO) validation technique in which all the voice samples of each individual was summarized using central tendency and dispersion metrics such as median, mean, standard deviation, trimmed mean, interquartile range, and mean absolute deviation. Their approach obtained 77.5% of classification accuracy.

Tsanas et al. in [15] focused on monitoring the PD progression with the help of extracted features using signal processing techniques applied on a huge dataset of about 6000 voice samples from 42 patients with early-stage PD. They attempted to estimate the unified Parkinson's disease rating scale (UPDRS) using linear and nonlinear regression. Their results show the accuracy of about 7.5-point difference from clinical UPDRS estimations. These datasets are the main publicly available datasets of PD speech-based area of study. Other studies tried to improve the PD detection rate using the existing datasets.

Tsanas et al. in [14] discovered 135 dysphonia new measurements using an existing dataset consisting of 263 vowels "aaaa. . ." phonations from 45 cases by applying feature selection techniques. They had obtained 99% overall classification accuracy.

Sakar and Kursun [16] tried to assess the relevance and correlation between the features and PD score by applying mutual information-based selection algorithm with permutation test and feed the data with selected features ranked based on maximum-relevance-minimum-redundancy (mRMR) into an SVM classifier. They used leave-one subject-out (LOSO) as the cross validation technique of their model in order to avoid bias. In LOSO validation scheme, all the voice samples of an individual which is the candidate of being the testing sample will be left out from the rest of the data. Their approach gained 92.75% classification accuracy.

Neal and Shahbaba [17] presented a nonlinear model based on Dirichlet mixtures and obtained the classification accuracy of 87.8%.

Das [18] performed a comparative study of neural networks (NN), DM neural, regression, and decision trees for PD diagnosis; their study resulted in classification performance of 92.9% based on NN.

Guo et al. [19] applied a combination of genetic programming and the expectation maximization (EM) and obtained a classification accuracy of 93.1%.

Luukka [20] proposed a method that used fuzzy entropy measures and similarity classifier and resulted in the mean accuracy of 85.03%.

Li et al. [21] found a fuzzy-based nonlinear transformation approach combined with SVM; their best classification accuracy was 93.47%.

Ozcift and Gulten [22] introduced classifier ensemble construction with a rotation forest approach which got classification accuracy of 87.13%.

Astrom and Koker [23] achieved the classification accuracy of 91.2% by using a parallel neural network model.

Polat [24] applied the fuzzy C-means clustering feature weighting together with the k-nearest neighbor classifier; their best obtained classification accuracy was 97.93%.

Chen et al. [25] proposed a model which combined PCA and the fuzzy k-nearest neighbor method; their classification approach achieved an accuracy of 96.07%.

Zuo et al. [26] used a diagnosis model based on particle swarm optimization (PSO) to strengthen the fuzzy k-nearest neighbor classifier which resulted in mean classification accuracy of 97.45%. In most of the studies, SVM was used as the base classifier to distinguish healthy subjects from Parkinson's patient [8, 14, 27] and the success of the diagnostic system was measured with ROC curves, AUC, and reporting True Positive and False Positive rates [28].

These datasets was grouped into two categories:
(1) those that contain the repetition of one term and (2) those that consist of different vocal terms.
The majority of datasets go to the first category. Hence, most of the studies on PD diagnosis were conducted on these datasets [14, 16–26]; however, 100% classification accuracy was still not obtained. The most popular and available datasets of this type are "Parkinson's Data Set" [7] and "Parkinson's Telemonitoring Data Set" [15], both dataset can be accessed from UCI Machine Learning Repository. The only dataset of the second category that was available in the form of processed data matrix was produced by Sakar et al. [13]. Less research has been done on this type of datasets; also, corresponding classification accuracies are not promising up to this time. The aim of this study is to show that this type of data collection can lead to high PD detection rates just by altering the classification strategy.

IV. CONCLUSION

An automated system for feature selection and classification of Parkinson's disease directly from original speech was developed using the techniques of Machine Learning. This system did not require any human intervention in the analysis. Besides the application itself, this method shows much promise for general machine perception of human conditions. Interestingly, this method shows that the deep human expertise in choice, selection and combination of speech signals can replaced by an automatic process on the auditory signal itself. While the features from the signal were, in fact, pre-chosen by the state of the art practice in general speech signal processing, it would be interesting in future work, to see if those features or replacements could be automatically discovered. We also expect that further pre-processing (e.g. removing "silent windows" or adding more global hierarchical windows) should be useful. As a further point, the general set-up of this work seems appropriate for application to a wide range of neurological diseases and states such as dementias, strokes and speech pathologies. Of course, one can foresee using such modules eventually in telemedicine systems.

Figure beside show Flow Diagram of the proposed scheme, starting from the signal acquisition phase through the preprocessing, to the feature extraction, training and classification modes followed by the decision scheme.
The full arrows indicate the trained system path and the dashed arrows indicate the machine learning training process.

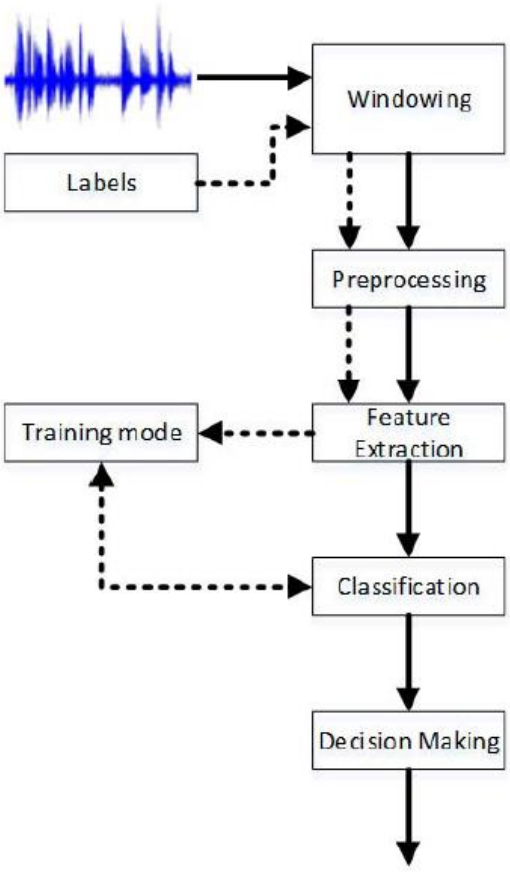

Figure 1. Flow Diagram Of Proposed Method.

V. REFERENCES

[1]. J.W.Langston,"Parkinson's disease: current and future challenges,"Neuro Toxicology,vol.23,no.4-5,pp.443-450,2002.

[2]. L.M.de Lau and M.M.Breteler,"Epidemiology of Parkinson's disease," The Lancet Neurology,vol.5,no.6,pp.525-535,2006.

[3]. M.C.de Rijk,L.J.Launer,K.Berger et al.,"Prevalence of Parkinson's disease in Europe: a collaborative study of population-based cohorts," Neurology,vol.54,no.11,supplement 5,pp.S21 S23,2000.

[4]. J.Jankovic,"Parkinson's disease: clinical features and diagnosis," Journal of Neurology,Neurosurgery and Psychiatry,vol.79,no.4,pp.368 376,2007.

[5]. A.K.Ho,R.Iansek,C.Marigliani,J.L.Bradshaw,and S.Gates,"Speech impairment in a large sample of patients with Parkinson's disease," Behavioral Neurology,vol.11,no.3,pp.131-137,1998.

[6]. J.A.Logemann,H.B.Fisher,B.Boshes,and E.R.Blonsky,"Frequency and co-occurrence of vocal tract dysfunctions in the speech of a large sample of Parkinson patients," Journal of Speech and Hearing Disorders,vol.43,no.1,pp.47-57,1978.

[7]. 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,article 23,2007.

[8]. M.A.Little,P.E.McSharry,E.J.Hunter,J.Spielman, and L.O.Ramig,"Suitability of dysphonia measurements for telemonitoring of Parkinson's disease," IEEE Transactions on Biomedical Engineering,vol.56,no.4,pp.1015-1022,2009.

[9]. J.Duffy,Motor Speech Disorders: Substrates,Differential Diagnosis,and Management,Mosby,St.Louis,Mo,USA,3rd edition,2012.

[10]. M.McNeil,Clinical Management of Sensorimotor Speech Disorders,Thieme,Stuttgart,Germany,2nd edition,2008.

[11]. R.J.Baken and R.F.Orlikoff,Clinical Measurement of Speech and Voice,Cengage Learning,2nd edition,1999.

[12]. P.H.Dejonckere,P.Bradley,P.Clemente et al.,"A basic protocol for functional assessment of voice pathology,especially for investigating the efficacy of (phonosurgical) treatments and evaluating new assessment techniques: guideline elaborated by the Committee on Phoniatrics of the European Laryngological Society (ELS)," European Archives of Oto-Rhino-Laryngology,vol.258,no.2,pp.77-82,2001.

[13]. B.E.Sakar,M.E.Isenkul,C.O.Sakar et al.,"Collection and analysis of a Parkinson speech dataset with multiple types of sound

recordings," IEEE Journal of Biomedical and Health Informatics,vol.17,no.4,pp.828-834,2013.

[14]. A.Tsanas,M.A.Little,P.E.McSharry,J.Spielman,and L.O.Ramig,"Novel speech signal processing algorithms for highaccuracy classification of Parkinsons disease," IEEE Transactions on Biomedical Engineering,vol.59,no.5,pp.1264-1271,2012.

[15]. A.Tsanas,M.A.Little,P.E.McSharry,and L.O.Ramig,"Accurate telemonitoring of Parkinson's disease progression by noninvasive speech tests," IEEE Transactions on Biomedical Engineering,vol.57,no.4,pp.884-893,2010.

[16]. C.O.Sakar and O.Kursun,"Telediagnosis of Parkinson's disease using measurements of dysphonia," Journal of Medical Systems,vol.34,no.4,pp.591-599,2010.

[17]. B.Shahbaba and R.Neal,"Nonlinear models using dirichlet process mixtures," Journal of Machine Learning Research,vol.10,pp.1829-1850,2009.

[18]. R.Das,"A comparison of multiple classification methods for diagnosis of Parkinson disease," Expert Systems with Applications,vol.37,no.2,pp.1568-1572,2010.

[19]. P.-F.Guo,P.Bhattacharya,and N.Kharma,"Advances in detecting Parkinson's disease," in Medical Biometrics,vol.6165 of Lecture Notes in Computer Science,pp.306-314,Springer,Berlin,Germany,2010.

[20]. P.Luukka,"Feature selection using fuzzy entropy measures with similarity classifier," Expert Systems with Applications,vol.38,no.4,pp.4600-4607,2011.

[21]. D.-C.Li,C.-W.Liu,and S.C.Hu,"A fuzzy-based data transformation for feature extraction to increase classification performance with small medical data sets," Artificial Intelligence in Medicine,vol.52,no.1,pp.45-52,2011.

[22]. A.Ozcift and A.Gulten,"Classifier ensemble construction with rotation forest to improve medical diagnosis performance of machine learning algorithms," Computer Methods and Programs in Biomedicine,vol.104,no.3,pp.443-451,2011.

[23]. F.Astrom and R.Koker,"A parallel neural network approach to prediction of Parkinson's Disease," Expert Systems with Applications,vol.38,no.10,pp.12470-12474,2011.

[24]. K.Polat,"Classification of Parkinson's disease using feature weighting method on the basis of fuzzy C-means clustering,"International Journal of Systems Science,vol.43,no.4,pp.597- 609,2012.

[25]. H.-L.Chen,C.-C.Huang,X.-G.Yu et al.,"An efficient diagnosis system for detection of Parkinson's disease using fuzzy k-nearest neighbor approach," Expert Systems with Applications,vol.40,no.1,pp.263-271,2013.

[26]. W.-L.Zuo,Z.-Y.Wang,T.Liu,and H.-L.Chen,"Effective detection of Parkinson's disease using an adaptive fuzzy k nearest neighbor approach," Biomedical Signal Processing and Control,vol.8,no.4,pp.364-373,2013.

[27]. O.Kursun,E.Gumus,A.Sertbas,and O.V.Favorov,"Selection of vocal features for Parkinson's Disease diagnosis," International Journal of Data Mining and Bioinformatics,vol.6,no.2,pp.144-161,2012.

[28]. I.Bhattacharya and M.Bhatia,"SVM classification to distinguish Parkinson disease patients," in Proceedings of the 1st Amrita ACM-W Celebration on Women in Computing in India(A2CWiC '10),New Delhi,India,2010.

[29]. L.O.Ramig,S.Sapir,C.Fox,and S.Countryman,"Changes in vocal loudness following intensive voice treatment (LSVT_) in individuals with Parkinson's disease: a comparison with untreated patients and normal age-matched controls," Movement Disorders,vol.16,no.1,pp.79 83,2001.

[30]. J.Gibbons,Nonparametric Statistical Inference,Chapman & Hall/CRC,5th edition,2010.

[31]. M.Hollander and D.A.Wolfe,Nonparametric Statistical Methods,JohnWiley & Sons,New York,NY,USA,3rd edition,2013.
[32]. M.M.Kendall,Rank Correlation Methods,Griffin,London,UK.
[33]. Syed Mohammad Ali And P.T.Karule,"Design Of System For Classification Of Vocal Cord/Glottis Carcinoma Using ANN And Support Vector Machine," International Journal Of Computer Applications (0975-8887) Volume132- no.4,dec-15.
[34]. Mohammad Shahbakhi,Danial Taheri Far,Ehsan Tahami,"Speech Analysis for Diagnosis of Parkinson'sDisease Using Genetic Algorithm and Support Vector Machine," J.Biomedical Science and Engineering,2014,7,pp.147-156.