Computer-Aided Diagnosis of Speech Disorder Signal in Parkinson's Disease

Heba M. Afify[1], Basma Ahmed[2]

[1] Ph.D., Department of Bioelectronics Engineering, MTI University of Engineering, Cairo, Egypt
hebaaffify@yahoo.com

[2] Teaching Assistant, Department of Bioelectronics Engineering, MTI University of Engineering, Cairo, Egypt
eng_basma1990@yahoo.com

ABSTRACT

Computer-aided diagnosis (CAD) can be used as a decision support system by physicians in the diagnosis and treatment of disordered speech especially those who specialize in neurophysiology diseases. Parkinson's disease (PD) is a progressive disorder of the nervous system that affects movement. It develops gradually, sometimes starting with a barely noticeable tremor in speech. It has been found that 80% of persons with PD reported speech and voice disorders. Parkinson's disease symptoms worsen as the condition progresses over time. Therefore, Speech may become soft or slurred and these deficits in speech intelligibility impact on health status and quality of life. Different researchers are currently working in the analysis of speech signal of people with PD, including the study of different dimensions in speech such as phonation, articulation, prosody, and intelligibility. Here, we present the characteristics and features of normal speech and speech disorders in people with PD and the types of classification for implementation of the efficacy of treatment interventions. The results show that our classification algorithm using ANN is outperformed KNN and SVM. ANN is a practical and useful as a predictive tool for PD screening with a high degree of accuracy, approximately 96.1% of a correct detection rate (sensitivity 94.7%, and specificity 96.6%). Based on the high levels of accuracy obtained by our proposed algorithm, it can be used for enhancing the detection purpose to discriminate PD patients from healthy people. Our algorithm may be used by the clinicians as a tool to confirm their diagnosis.

Keywords

Computer-aided diagnosis; Parkinson's disease; Speech and voice disorders; Classification.

Academic Discipline And Sub-Disciplines

Biomedical Signal processing, Classification

SUBJECT CLASSIFICATION

Computer-aided diagnosis

INTRODUCTION

Many neurologic diseases [1] have been effected on speech production such as Parkinson's disease (PD), stroke [2], and multiple sclerosis [3]. People with Parkinson's develop problems with their speech and communication that can include: lack of volume, hoarse voice, lack of expression and indistinct speech [4]. Interestingly, there were differences concerning the perceptual speech ratings between male and female PD speakers which had been previously reported concerning some distinctive measures of acoustic speech analysis [5].

Although Parkinson's disease can't be cured, medications may markedly improve its symptoms. In occasional cases, early classification of speech abnormalities is critical for a speech-language therapist (SLT) to regulate certain regions of the brain and improve symptoms [6]. Schulz and Grant [7] conducted a review of the different treatment approaches for people with PD up to the time of their manuscript and examined the effects of these treatments on speech. Treatment methods reviewed included speech therapy, pharmacological intervention, and surgical procedures. A well-studied technique for improving vocal volume was the Lee Silverman Voice Treatment (LSVT). LSVT is an intensive program of therapy that aims to improve intelligible oral communication by teaching speakers to increase their vocal loudness using tasks designed to maximize voice and breathing functions [8].

Many methods for pathological voice classification have been proposed using this database. In general, pattern recognition of speech signal based on analysis of the effects on features of various data including normal and pathological voices [9]. The choice of the employed features is generally driven by the accuracy of classification, time needed for classification, and cost of performing classification.

There exists a wide range of database and algorithms of PD classification, a more recent algorithm is provided in [10]. In particular, many methods use personal and inaccessible databases, with a disparity in recording and patients conditions. The authors [11] presented pathological voice classification by using the MEEI-KayPENTAX database to achieve high classification accuracy. In [12] discussed the vowel metrics to differentiate disordered from non-disordered speakers and classified the dysarthria subtype. Toward this end, means testing (e.g. T-tests and analyzes of variance) and stepwise discriminant function analysis (DFA) were conducted. J.R. Orozco-Arroyave [13] presented a new database with speech recordings of 50 patients with PD and their respective healthy controls, matched by age and gender based on Spanish native speakers. A recent paper [14] concluded that classification algorithm using Naïve Bayes was outperformed KNN classifier and it was useful as a predictive tool for PD screening with a high degree of accuracy, approximately 93.3%.

In this paper, we discuss the classification algorithm between normal and pathological voices of PD by using KNN, SVM, and neural network. We concentrated on features for speech based human recognition as well as classification strategies. In addition, we evaluated a weighted mean accuracy of the classification methods to assess the overall performance of the PD dataset.

MATERIAL AND METHODS

Data Extraction

The present study included PD database by the National Center for Voice and Speech (NCVS) that comprises 31 people which composed of 23 with PD, and 10 healthy controls an extension of the database used in [15]. The 10 healthy controls consist of 4 males and 6 females, had an age range of 46 to 72 years with (mean ± standard deviation) 61 ± 8.6 years. The dataset was created by Max Little [16] of the University of Oxford, in collaboration with the NCVS, Denver, Colorado, who recorded the speech signals. The original study published the feature extraction methods for general voice disorders. Each column in the dataset is a particular voice measure, and each row corresponds one of 195 voice recording from these individuals ("name" column).

The main aim of the data is to discriminate healthy people from those with PD, according to "status" column which is set to 0 for healthy and 1 for PD. The data is in ASCII CSV format. The rows of the CSV file contain an instance corresponding to one voice recording. There are around six recordings per patient, the name of the patient is identified in the first column.

We divided our dataset to training data and testing data half by half. The training set is used for classifier design of known data while the testing set is used for performance assessment of unknown data. We used three methods for classification; artificial neural network (ANN) [17], K-nearest neighbor (K-NN) classifier [18], and support vector machines (SVM) [19] by using Matlab program to distinguish the differences between normal and abnormal speech.

Selection features from the speech signals

There are several types of features of a speech signal such as features that describing the phonation [20], features that describing the speech intensity, features that describing the speech quality, and non-linear dynamic features. We selected features that describing the phonation, which is the vibration of the vocal folds to create sound, such as fundamental frequency (F_0), and 5 types of jitter (e.g. Pitch Perturbation Quotient PPQ). Features that describing the speech intensity are 6 types of shimmer (e.g. Amplitude Perturbation Quotient APQ). Features that describing the speech quality are the harmonic-to-noise ratio (HNR), and the noise-to-harmonic ratio (NHR). The set of Non-linear dynamic features includes correlation dimension (D2), Recurrence Period Density Entropy (RPDE), Detrended Fluctuation Analysis (DFA) and Pitch Period Entropy (PPE). DFA is a method for determining the statistical self-affinity of a signal.

Most speech features based on Multidimensional Voice Program (MDVP), which is a computer program which analyzes various aspects of voice, can detect abnormal voice patterns of patients with upper airway pathology [21]. We introduced selection features of our dataset as shown in Table 1. All features for speech dataset serve as input to a set of statistical classifiers. The general algorithm that is applied in this work is depicted in Figure 1.
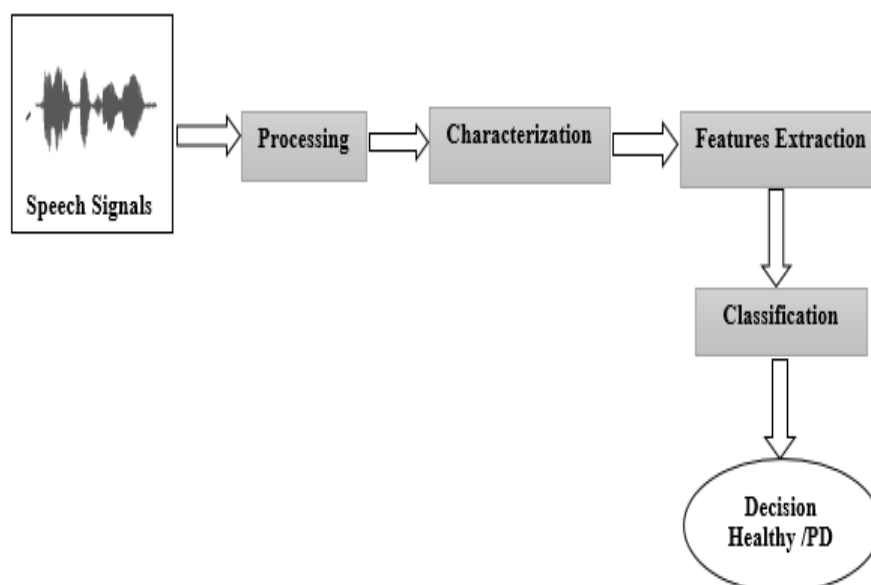

Figure 1: General Methodology

Table 1: List of the selection Dysphonia features

Dysphonia Features	Description	Speech subsystem
MDVP: Fo(Hz)	Average vocal fundamental frequency	Phonation
MDVP: Fhi(Hz)	Maximum vocal fundamental frequency	Phonation
MDVP:Flo(Hz)	Minimum vocal fundamental frequency	Phonation
MDVP:Jitter(%),MDVP:Jitter(Abs),MDVP:RAP,MDVP:PPQ,Jitter:DDP	Several measures of variation in fundamental frequency	Phonation
MDVP:Shimmer,MDVP:Shimmer(dB), Shimmer:APQ3,Shimmer:APQ5,MDVP:APQ,Shimmer:DDA	Several measures of variation in amplitude	Speech Intensity
NHR,HNR	Two measures of ratio of noise to tonal components in the voice status - Health status of the subject (one) - Parkinson's, (zero) – healthy	Speech Quality
RPDE,D2	Two nonlinear dynamical complexity measures	Non-linear dynamic
DFA	Signal fractal scaling exponent	Statistical self-affinity of a signal
spread1,spread2,PPE	Three nonlinear measures of fundamental frequency variation	Phonation

RESULTS

Classification of speech dataset may be a useful tool when diagnosing the speech disorders, as well as in evaluating response to treatment. We have applied 22 features extraction from our speech signal to three classification methods. These features were extracted to form the features vectors. In this work, the T-test [22] was used to test the significance of each feature to be used in classifying different speech types. Student's t-distribution or simply the t-distribution is a family of continuous probability distributions that arises when estimating the mean of a normally distributed population in situations where the sample size is small and population standard deviation is unknown. In ANN, we supposed that type of network was feed- forward and the number of hidden neurons were 15. In KNN, different numbers of K- nearest Neighbor were used and the best accuracy was given at k=1 as seen in Figure 2. The SVM is a powerful classification technique in the statistical learning theory field, which was developed as a binary classifier. We evaluated the classification performance of three classifiers configurations on the testing set. Sensitivity, specificity for testing and accuracy of each classifier are computed as in Table 2.
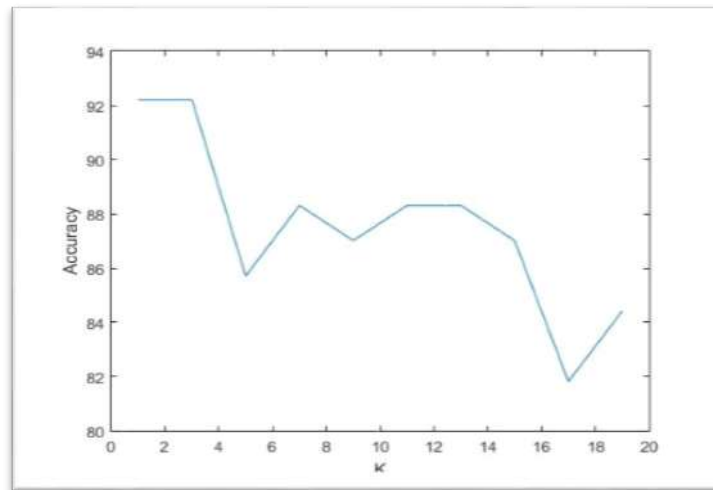
7119 | P a g e
J u l y   2 0 1 6
c o u n c i l   f o r   I n n o v a t i v e   R e s e a r c h
w w w . c i r w o r l d . c o m

I S S N 2 2 7 7 - 3 0 6 1
V o l u m e 1 5 N u m b e r 9
I n t e r n a t i o n a l J o u r n a l o f C o m p u t e r s & T e c h n o l o g y

Figure 2: Different numbers of K in K-nearest neighbor (KNN) classifier

Table 2: Results for classification process using different classifiers

Classifier	Accuracy	Sensitivity	Specificity
ANN	96.1 %	94.7 %	96.6 %
KNN	92.2 %	78.9 %	96.5 %
SVM	76.6 %	78.9 %	75.8 %

The results confirm that the ANN classifier provides best results in the accuracy, sensitivity and specificity. On the other hand, SVM classifier shows the lowest specificity. KNN and SVM classifier show the lowest sensitivity. As can be seen, there is the noticeable increase in the value of the specificity of ANN compared to that of the performance ones.

The results show that our classification algorithm using ANN is outperformed KNN and SVM. ANN is a practical and useful as a predictive tool for PD screening with a high degree of accuracy, approximately 96.1% of a correct detection rate (sensitivity 94.7%, and specificity 96.6%). Also, we found that our results better than in [14] for recognition the speech signal of PD patient.

DISCUSSION

Some people with PD experience changes in cognition and language, which make it difficult to think quickly, to manage multiple tasks, to find words or to understand complex sentences. These voice changes may make a person's speech less precise and more difficult to understand. Some individuals with PD may speak more slowly. Others, perhaps 10 percent, accelerate their speech so much that they stumble over sounds, and seem to be stuttering. Therefore, CAD algorithms have been developed to help clinicians give an accurate diagnosis and to reduce the number of wrong decisions regarding the PD. In another word, it may be used by the clinicians as a tool to confirm their diagnosis.

Over the years, there has been an improvement in the detection algorithms but their performance is still not perfect. Another issue is extracting and selecting appropriate features that will give the best classification results. Furthermore, the choice of a classifier has a great influence on the final result and classifying abnormalities of speech is a difficult task even for expert clinicians. Further developments in each algorithm step are required to improve the overall performance of computer-aided detection and diagnosis algorithms.

Measuring Performance with ROC Curves

Receiver Operating Characteristic (ROC) curves are a common approach used to assess the performance of a classifier [23]. An ROC curve determines the number of true positive, true negatives, false positives, and false negatives and produces a summary statistic for performance, detailing model accuracy, sensitivity, and specificity. It achieves this by plotting the true positive rate against the false positive rate at different possible cut points. The area under the curve (AUC) value measures the discrimination, that is, the ability of the model to correctly classify the true positives and negatives.

I S S N   2 2 7 7 - 3 0 6 1
V o l u m e   1 5   N u m b e r   9
I n t e r n a t i o n a l   J o u r n a l   o f   C o m p u t e r s   &   T e c h n o l o g y

A perfect ROC curve (and therefore perfect test) would have an AUC value of 1, thus, the closer the curve follows the left-hand border and then the top border of the ROC Space, the more accurate the test will be. Figures 3, 4 and 5 show the ROC for testing set classification of ANN, KNN, and SVM, respectively. Therefore, results confirm that the ROC of ANN is a perfect one from the others.
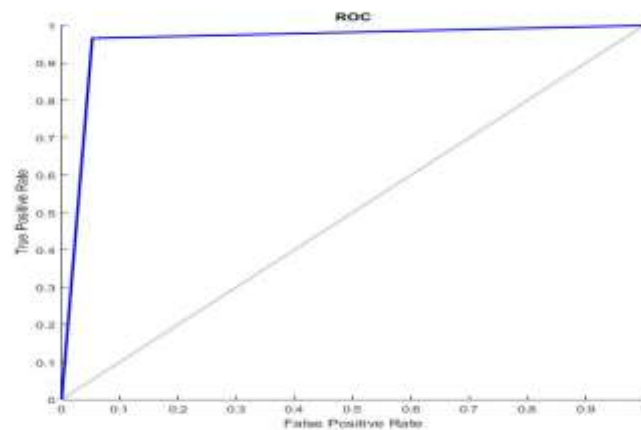

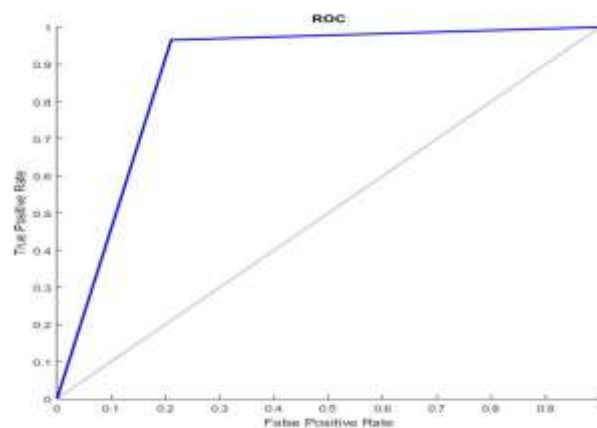Figure 3: Receiver Operating Characteristic (ROC) of artificial neural network (ANN)

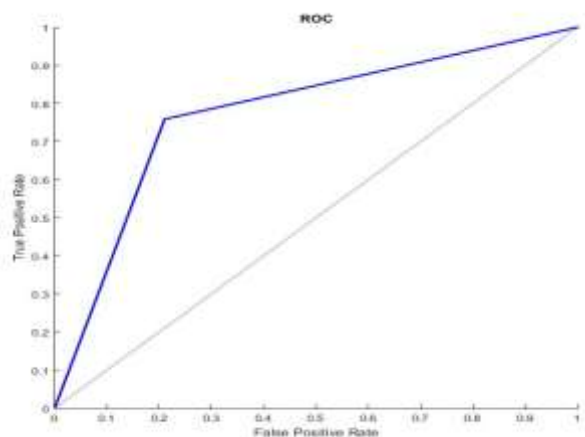Figure 4: Receiver Operating Characteristic (ROC) of K-nearest neighbor (KNN)

Figure 5: Receiver Operating Characteristic (ROC) of support vector machines (SVM)

CONCLUSION

Speech disturbance is a common symptom for individuals with PD. CAD may be defined as a diagnosis tool which takes into account the computer output as a second opinion. The purpose of CAD is to improve the diagnostic accuracy and the consistency of the speech signal interpretation. It should be of a particular value to enable the early detection of speech disturbance. Based on the high levels of accuracy obtained by our proposed algorithm, it can be used for enhancing the detection purpose to discriminate PD patients from healthy people.

Further work is required to establish appropriate methods of speech investigations which produce objective data, fulfill the demands of validity and reproducibility, time and cost effectiveness, and mirror bests the functional disability of patients.

REFERENCES

[1] Kim, Y.-J., Kent, R. D., & Weismer, G. An acoustic study of the relationships among neurologic disease, dysarthria type and severity of dysarthria Journal of Speech, Language, and Hearing Research. 2001; (54):417–429.

[2] Kim, Y.-J., Weismer, G., Kent, R. D., & Duffy, J. R. Statistical models of F2 slope in relation to severity of dysarthria. Folia Phoniatrica et Logopaedica. 2009; (61): 329–335.

[3] Rosen, K. M. Goozee, J. V., & Murdoch, B. E. Examining the effects of multiple sclerosis on speech production: Does phonetic structure matter?. Journal of Communication Disorders.2008; (41):49–69.

[4]. 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. Behavioural Neurology. 1998; 11(3):131–137.

[5] S. Skodda, W. Visser, and U. Schlegel. Gender-related patterns of dysprosody in Parkinson disease and correlation between speech variables and motor symptoms", Journal of Voice. 2011; 25(1):76–82.

[6] G. J. Canter.Speech characteristics of patients with Parkinson's disease. 3. Articulation, diadochokinesis, and over-all speech adequacy. The Journal of Speech and Hearing Disorders. 1965; (30): 217–224.

[7] Schulz, G.M. & Grant, M.K. Effects of speech therapy and pharmacologic and surgical treatments on voice and speech in Parkinson's disease: A review of the literature. Journal of Communication Disorders. 2000; 33(1): 59-88.

[8] Sharkawi .A, Ramig L, Logemann JA, et al. Swallowing and voice effects of Lee Silverman Voice Treatment (LSVT): A pilot study. Journal of Neurology, Neurosurgery, and Psychiatry. 2002; 72: 31–36.

[9] Schwartz, R., et al. Speech Recognition in Multiple Languages and Domains: The 2003 BBN/LIMSI EARS System. Proc. ICASSP, 2004.

[10] Tsanas A. et al. Novel speech signal processing algorithms for high accuracy classification of Parkinson's disease. IEEE Trans. Biomedical Engineering. 2012; 59(5):1264–1271.

[11] Khalid Daoudi, Blaise Bertrac. On classification between normal and pathological voices using the MEEI-KayPENTAX database: Issues and consequences. INTERSPEECH, Jun 2014.

[12] Kaitlin L. Lansford, and Julie M. Liss, Vowel Acoustics in Dysarthria: Speech Disorder Diagnosis and Classification. Journal of Speech, Language, and Hearing Research. February 2014; 57:57–67.

[13] J.R. Orozco-Arroyave, J.D. Arias-Londo˜no, J.F. Vargas-Bonilla. New Spanish speech corpus database for the analysis of people suffering from Parkinson's disease. Ninth International Conference on Language Resources and Evaluation (LREC'14), At Reykjavik, Iceland, 2014.

[14] Yahia Alemami, Laiali Almazaydeh. Detection of Parkinson Disease through Voice Signal Features. Journal of American Science, 2014; 10(10).

[15] Parkinson's disease dataset: https://archive.ics.uci.edu/ml/datasets/Parkinsons

[16] M.A. Little, P.E. McSharry, E.J. Hunter, J. Spielman, L.O. Ramig. Suitability of Dysphonia Measurements for Telemonitoring of Parkinson's Disease. IEEE Trans. Biomedical Engineering. 2009; 56(4):1015-1022.

[17] Balcázar, José. Computational Power of Neural Networks: A Kolmogorov Complexity Characterization. Information Theory, IEEE Transactions. Jul 1997; 43(4): 1175–1183. Doi: 10.1109/18.605580.

[18] Altman, N. S. An introduction to kernel and nearest-neighbor nonparametric regression. The American Statistician. 1992; 46 (3):175–185.

[19] Meyer, D.,Leisch, F.,Hornik, K. The support vector machine under test. Neurocomputing. 2003; 55: 169.

[20] A. M. Goberman. Correlation between acoustic speech characteristics and non-speech motor performance in Parkinson disease. Med. Sci. Monit. 2005; 11:109–116.

[21] Zelcer S, Henri C, Tewfik TL, Mazer B. Multidimensional voice program analysis (MDVP) and the diagnosis of pediatric vocal cord dysfunction. Ann Allergy Asthma Immunol. 2002; 88(6):601-8.

[22] John A. Rice. Mathematical Statistics and Data Analysis. Third Edition, Duxbury Advanced; 2006.

[23] Fogarty, James; Baker, Ryan S., Hudson, Scott E. Case studies in the use of ROC curve analysis for sensor-based estimates in human computer interaction. ACM International Conference Proceeding Series, Proceedings of Graphics Interface; 2005.

Author' biography with Photo

Dr. Heba M.Afify had completed her M.Sc. & Ph.D. in Biomedical Engineering from Cairo University, Cairo, Egypt in 2007 & 2011 respectively. Her research area is Biomedical Signal Processing and currently working on Bioinformatics. She working as Lecturer in Bioelectronics at MTI University, Cairo, Egypt. She had authored seven research papers in reputed international & national journals.

Eng. Basma Ahmed had completed her B.Sc. in Biomedical Engineering from Helwan University, Cairo, Egypt in 2012 and she is currently Master Degree candidate. Her research area is Biomedical Signal Processing and currently working on Biomedical Image. She working as Teaching assistant in Bioelectronics at MTI University, Cairo, Egypt.