Diagnosis of Parkinson's Disease in Human Using Voice Signals

Hamid Karimi Rouzbahani [1,2], Mohammad Reza Daliri [1,2]*

1. Biomedical Engineering Department Iran University of Science and Technology (IUST), Tehran, Iran.
2. Iran Neural Technology Research Center,ran University of Science and Technology (IUST), Tehran, Iran.

Article info:
Received: 12 January 2011
First Revision: 19 January 2011
Accepted: 10 February 2011

Key Words:
Classification,
Dysarthria,
Feature selection,
Evaluation,
Parkinson's Disease (PD).

A B S T R A C T

Introduction: A full investigation into the features extracted from voice signals of people with and without Parkinson's disease was performed. A total of 31 people with and without the disease participated in the data collection phase. In this study subjects' voice signals were used to let computer decide whether the person is suffering from the disease or they are not.

Methods: Their voice signals were recorded and processed. The relevant features were then extracted. Features were fed to different classifiers so as to be let them decide whether the subjects have the disease or not. Three different classifiers were used in order to rule out any doubt about the validity of classification performance on the given data.

Results: The use of a variety of feature selection methods resulted in a good performance for the diagnosis of Parkinson. The classifiers' performances were compared with one another and showed that the best performance was obtained with a correct rate of 0.9382 when using the KNN classifier.

Discussion: Results reveal that the use of proposed feature selection method results in a desirable precision for the diagnosis of Parkinson's disease (PD). The performances were assessed from different points of view, providing different aspects of the diagnosis, from which the physicians are able to choose one with higher accuracy in the diagnosis.

1. Introduction

Parkinson's disease (PD), is a disorder of the nervous system that affects muscle control. Marked by trembling of the arms and legs, muscular rigidity, and poor balance, Parkinson's disease is slowly progressive, worsening over time. Eventually symptoms may cause problems with walking or talking and, in some people, difficulty thinking. Physicians do not know how to cure Parkinson's disease, but drug therapy or surgery may alleviate some of the most troubling symptoms. The National Parkinson Foundation based in Miami, Florida, estimates that 1.5 million people in the United States are affected with Parkinson's disease (Lang et al., 1998), although estimates are difficult to make because symptoms of the disease are often mistaken for the normal effects of aging or are attributed to other diseases. Parkinson's disease occurs in people all over the world, with the incidence in men slightly higher than in women. People most commonly develop Parkinson's disease around the age of 60, and the incidence rises with age (Van Den Eeden et al., 2003). However, at least 10 percent of cases occur in people under age 40, and a rare form of the disease affects teenagers.

**** Corresponding Author:***
Mohammad Reza Daliri, PhD
Biomedical Engineering Department and Iran Neural Technology Research Center, Iran University of Science and Technology (IUST), Tehran 16846-13114, Iran. Tel: +98-21-73225738. Fax:+98-21-73225777
Email: daliri@iust.ac.ir

Diagnosing Parkinson's disease may be difficult, particularly in the early stages of the disease when symptoms resemble other medical conditions, and misdiagnosis occurs occasionally. No single laboratory test can diagnose the disease. Blood tests, brain imaging techniques such as magnetic resonance image (MRI), positron emission tomography (PET scan), and single photon emission computed tomography (SPECT), may be used to help doctors exclude other medical conditions, such as stroke or brain tumors, that produce symptoms similar to those of Parkinson's disease. Thus, there is no unique way of diagnosis among physicians and their diagnoses are mostly based on trial and error which is not desirable. Amongst others, one of the methods for the diagnosis of voice disorders which is commonly used by clinicians and vocal therapists is the use of acoustic tools that record the changes in pressure at lips or inside the vocal tract. Recently, upon signal processing a group of experts (Little et al., 2009) found some features in the voices of the people with Parkinson's disease that can be used as discriminatory measures to differentiate those who have the disease from those who do not. As mentioned before, the disease causes the muscles to shake and vibrate unintentionally and that it starts from smaller ones (Ho et al., 1998; Logemann et al., 1978). The very onset places in which the disorder starts to appear is the patient's vocal cords, so as of then the patient's voice signal will have some extra oscillations and other abnormal deviations from the expected signal that can be extracted if it is processed. This is obvious when the patient decides to speak and cannot produce the correct vocal sounds. This vocal impairment is called dysarthria. The problem they experience when trying to normally articulate speech is named dysarthria. Symptoms of dysarthria include reduced loudness, breathiness (which is a kind of noise mounted on voice signal), decreased energy in higher parts of harmonic spectrum and exaggerated vocal tremor (Rahn et al., 2007).

In the field of vocal diagnosis there are a handful of methods for recording and sampling voices depending on what objectives we have (Baken et al., 2000; Dejonckere et al., 2001). One of the most famous approaches is running speech in which the subject is supposed to read a sentence out and their voice is recorded simultaneously. The sentence is a standard one and forces the subject to pronounce different sounds so as to expose their impairments. Other method is vowel (or sustained) phonation in which the subject is supposed to pronounce a vowel sound for as long as possible while having their voices recorded. Although the first method might seem more realistic, second method is sufficient to show the desired discriminatory criteria. Normal and disordered vowel sounds have a large range of behavior including nearly periodic or regular vibration, aperiodic or irregular vibration and sounds with no apparent vibration at all. All these types of signal suffer from some sort of breathiness of noise. In general vocal problems exhibit two characteristic phenomena: increased vibrating aperiodicity and increased breathiness (Michaelis et al., 1998). These categorizations are necessary for the analysis of speech signals. Because the mentioned noise is remarkable in analysis we cannot use classical methods (specially linear and deterministic methods like Fourier transform or spectrum analyzing) and there is a need for some new methods of stochastic signal processing (Carding et al., 2004).

Recently, some new measurement methods have been proposed which are based on nonlinear dynamical systems theory (Kantz et al., 1999; Little et al., 2007). This is in accordance with the nonlinear dynamics of voice production system which consists of vocal cords, pharynx, larynx and so on, each of which is a nonlinear system. Therefore, any changes in produced voice signal caused by any disease can change the dynamics of the whole system. These changes can be detected and the desired features extracted. According to the fact that randomness and noise are integral parts of the produced voice, as the new methods of processing, recurrence pitch entropy density (RPDE) and detrended fluctuation analysis (DFA) are applied to the voice signals as tools for showing the ability to detect general voice disorders (Little et al., 2007). Another important measure, that has been found invariant to the acoustic environment and the gender of the subject, is pitch period entropy (PPE) that is proposed by (Little et al., 2009), and used in this study.

These novel characteristics together with classical characteristics which are extracted using conventional methods make a sum of 22 discriminatory features that can differentiate a person with Parkinson from a healthy one using some classification method (Little et al., 2009). The measurements are provided by (Little et al., 2007). A classification method is applied on this data in (Little et al., 2009), revealed that these features are valid to be used in telemedicine for those who suffer from the disease. The result was a 0.914 of correct rate using the SVM classifier. Although relatively good, in the same study there are some features omitted, only one classifier used and no specific method mentioned for the feature selection phase thus ending up to the need for evaluation of all possible combinations among all the feature set. This will be burdensome and bring about lack of justification for the choice made. In addition, there is no comparison done on the dataset based on the

kind of classifier used. In this paper we will clarify the reasons for which our features are chosen based on the most state of the art criteria available in the literature while simultaneously comparing some widely used classifiers' behaviors that eventually result in a good performance for each of them.

2. Methods

2.1. Data Collection

2.1.1. Experimental Setup

31 people, form both genders, participated in the speech recording experiments (Little et al., 2009). They were aware of experimental conditions and their consents had already been collected. Their age ranged from 46 to 85 years. Some more detailed information on participants' condition is gathered in Table 1. The time since their disease diagnosis ranged from 0 to 28 years. It should be noted that we, as researchers, are blind to the diagnosis and all participants were diagnosed as with or without the disease according to experienced physicians. Averagely there were 6 voice signals recorded from each of them. There gathered 195 measurements altogether from the subjects. Among them, there were 23 with the disease and the rest 8 were healthy. There was a microphone (AKG C420) fitted on their heads, 8cm from their mouths, recording their voices while they pronounced a specific vowel sound for as long as they could. Calibration of microphone was done using a Class 1 sound level meter (B&K 2238) 30cm distant from the speaker. The phonation signal was sent to the

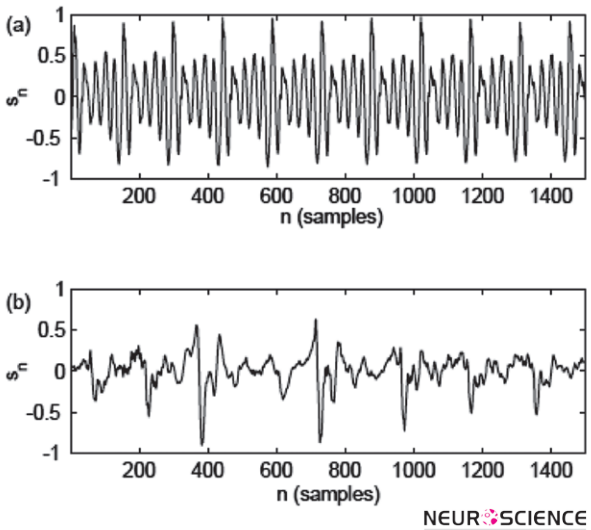


Figure 1. Two selected examples of speech signals: (a) healthy, (b) subject with PD. The horizontal axis is time in seconds, the vertical axis is signal amplitude (no units) (figure from (Little et al., 2007)).

Table 1. List of subjects with sex, age, Parkinson's stage and number of years since diagnosis.

Subject code	Sex	Age	Stage (H&Y)	Years since diagnosis
S01	M	78	3.0	0
S34	F	79	2.5	¼
S44	M	67	1.5	1
S20	M	70	3.0	1
S24	M	73	2.5	1
S26	F	53	2.0	1½
S08	F	48	2.0	2
S39	M	64	2.0	2
S33	M	68	2.0	3
S32	M	50	1.0	4
S02	M	60	2.0	4
S22	M	60	1.5	4½
S37	M	76	1.0	5
S21	M	81	1.5	5
S04	M	70	2.5	5½
S19	M	73	1.0	7
S35	F	85	4.0	7
S05	F	72	3.0	8
S18	M	61	2.5	11
S16	M	62	2.5	14
S27	M	72	2.5	15
S25	F	74	3.0	23
S06	F	63	2.5	28
S10(healthy)	F	46	n/a	n/a
S07(healthy)	F	48	n/a	n/a
S13(healthy)	M	61	n/a	n/a
S43(healthy)	M	62	n/a	n/a
S17(healthy)	F	64	n/a	n/a
S42(healthy)	F	66	n/a	n/a
S50(healthy)	F	66	n/a	n/a
S49(healthy)	M	69	n/a	n/a

Note: "H&Y" refers to the Hoehn and Yahr PD stage, where higher values indicate greater level of disability (Hoehn et al., (1967)).

Table 2. Extracted features.

No.	Feature	Explanation
1	MDVP: Fo(Hz)	Average vocal fundamental frequency
2	MDVP:Fhi(Hz)	Maximum vocal fundamental frequency
3	MDVP:Flo(Hz)	Minimum vocal fundamental frequency
4	MDVP:Jitter(%)	A measure of variation in fundamental frequency
5	MDVP:Jitter(Abs)	*
6	MDVP:RAP	*
7	MDVP:PPQ	*
8	Jitter:DDP	*
9	MDVP:Shimmer	A measure of variation in amplitude
10	MDVP: Shimmer(dB)	*
11	Shimmer:APQ3	*
12	Shimmer:APQ5	*
13	MDVP:APQ	*
14	Shimmer: DDA	*
15	NHR	A measure of ratio of noise to tonal components in the voice
16	HNR	*
17	RPDE	A nonlinear dynamical complexity measure
18	D2	*
19	DFA	Signal fractal scaling exponent
20	spread1	A nonlinear measure of fundamental frequency variation
21	spread2	*
22	PPE	*

NEUROSCIENCE

Note: * means the same explanation as previous line.

computer using CSL 4300B hardware (KayElemetrics). The signal is then sampled at a rate of 44.1 kHz with 16 bit resolution. For the sake of brevity latter reference is recommended to the reader for more information. The 22 features shown in Table 2 are extracted, some from amplitude, some from the fundamental frequency and so on. Two example of speech signals are shown in Fig. 1, one from a healthy subject and the other from a patient. The differences are quite simple to recognize.

2.1.2. Feature Extraction

In calculations the second half of every signal is ignored, because they are greatly affected by spurious dysarthria caused by lack of lung pressure which is observed in those not suffering from PD as well. In this way the risk of misinterpretation is reduced and dignity of the diagnosis is retained.

Traditional measures were calculated using the software Praat (Boersma and Weenink, 2005, 2001). These measures rely on the application of successive windows of the signal, with the determination of the vibration frequency of vocal folds (F0 or pitch period) using peak-picking, and location in time of the beginning of each cycle of vibration of the vocal folds (pitch marks) (Boersma, 2005). The noise-to-harmonics and harmonics-to-noise measures are extracted using estimates of signal to noise by calculation of autocorrelation of each cycle.

Measures related to the perturbation of amplitude were extracted by analyzing vocal cycles. So was calculated shimmer. This was done by the investigation of maximum extent of amplitude sequence in each cycle. The difference appears, when averaging mentioned sequences, and this difference is used as a measure. The same method of averaging is used for the period perturbation measures and jitter. The only difference is that absolute differences between frequencies of each cycle are used here. The averaging is applied over a varying number of cycles which is then normalized by the overall average.

The correlation dimension (D2) is calculated by time delay embedding the signal to recreate the phase space of nonlinear dynamical system that is proposed to generate the speech signal (Kantz et al., 1999).

Ignoring the calculation method for this measure given in (Little et al., 2007), the recurrence period density entropy (RPDE) is a factor showing the non periodicity of the signal. This measure is a good evidence for general voice disorders. As we know, general voice pathologies perturb the normal activity of vocal folds, not allowing them produce regular vibration.

The next measure is named detrended fluctuation analysis (DFA). This measure shows the self-autocorrelation or self-similarity of noise in speech signal. The more self-similar this noise, the more problems may exist in voice production.

The noise heard in voice is mostly the product of turbulence of the air through the vocal folds. In dysphonic

disorders like incomplete vocal fold closure, there appear some self-similarity features of noise in speech signal which can be taken into consideration as measure for differentiating those who have the disease from those who do not.

A new measure is introduced in (Little et al., 2009). For the sake of awareness we have a glance at its calculation in this article. There is a normal variation in natural pitch (F0) which is observed in normal speech signal. This variation is characterized by smooth vibrato and microtremor (Schoentgen and Deguchteneere, 1995). The variation is distorted in people with PD. Of course, this fact is more remarkable during sustained phonation (Cnockaert et al., 2008), and that is one of the reasons for the choice of sustained phonation rather than running speech.

For the calculation of this measure, firstly, the pitch sequence of the phonation is obtained and converted to the logarithmic semitone scale, p(t), where p is the semitone pitch at time t. Then using a standard linear whitening filter, the linear temporal correlations in the semitone sequence are removed. This procedure will provide the analysis for the roughness of variations in the mentioned sequence to produce the relative semitone variation sequence r(t). Secondly, a discrete probability distribution of occurrence of relative semitone variations, P(r), is constructed. Finally, the entropy for this distribution is calculated which characterizes the extent of oscillations in the sequence of relative semitone pitch period variations. The increase in this entropy is corresponding to better variations over and above natural healthy variations in pitch observed in healthy speech production.

2.2. Feature Vector Preparation

2.2.1. Preprocessing

At the first step of feature preparation, an outlier removal method was tested trying to improve the efficiency of the classifiers. For this purpose, firstly, we chose a coefficient for the standard deviation based on which we could decide how big the span of the desired values of the observations was to be. Then, considering different values for the mentioned coefficient it was seen that the resulting performance did not improve, showing that reducing the number of observations in this special case, may be because of few number of observations, does not help us. So outlier removal was not considered a case of improvement anymore. Some of the results on outlier removal are provided in Table 3. Data normalization seemed to be necessary to be applied to the data at hand, because the span of variation was so much different from one feature to another, and this would have discriminate some features from some others in the classifier's view. We applied normalization using both linear and nonlinear formulations and according to the better performance resulted, the nonlinear one (softmax scaling with the free parameter r set to 0.5) was selected (Theodoridis and Koutroumbas, 2010). This method squashes data values nonlinearly in the interval [0,1].

Table 3. The effect of outlier removal on classification performance (SVM classifier).

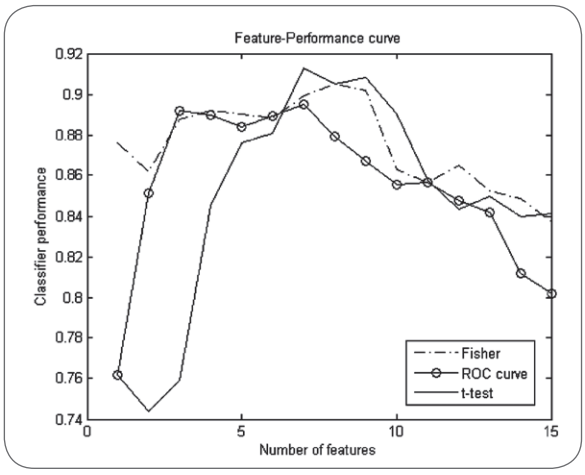
Coefficient Value	Number of Features Left	Classifier's Correct Rate
Original data	195	0.9144
1	146	0.9042
2	183	0.9099
3	194	0.9135

NEUROSCIENCE

2.2.2. Feature Selection

So far we had 22 features at hand that we are supposed to choose from, the best subset. The word 'subset' is used because it is desirable to reduce the number of features to a manageable size so as to firstly, reduce the complexity of classifier and secondly, to disregard the features that do not help us improving the performance of classifier if there is any. Firstly, for reaching a number that shows the optimum number of features which are considered informative enough, and that can lead the classifier towards its highest performance, we had to evaluate each feature individually. These features are supposed to be assessed based on some criteria which show the information that each of them contains. Because some of the features are extracted from the same parameters in the voice signals they showed a highly correlated dispersion in feature space. For the sake of improvement, one of every two features which showed a correlation rate over 95% was dismissed. The dismissed features are MDVP:Jitter(%), MDVP:RAP, MDVP:PPQ, MDVP:Shimmer, MDVP:Shimmer(dB), Shimmer:APQ3, Shimmer:APQ5 so there are 15 features left. According to the fact that there exist a few criteria in the literature that can prioritize the features independently we utilized in this study the most common ones which are Fisher's Discriminant Ratio (FDR), t-test and ROC curve. For more information on these criteria the reader is referred to (Theodoridis and Koutroumbas, 2010; Theodoridis and Koutroumbas, 2009; Duda, 2000).

Incrementing one by one the number of features that had already been prioritized and using them in a SVM classifier and simultaneously evaluating the performance of classifier it was revealed that the highest performance is achieved when using a combination of 7 most prior features, using the three classifiers at hand. It is confirmed by the resulted curves shown in Fig. 2 using the three mentioned criteria. As it is obvious when using 7 features two of the criteria (t-test and ROC curve) reach their maximum and the third one (Fisher) is almost in its peak so that the justification for the choice.



NEUROSCIENCE

Figure 2. Feature-performance curve. The three prevalent criteria have been used to evaluate the effect of number of features on classifier performance. SVM classifier with rbf kernel is used and Hold-out method is used for evaluation.

Now knowing the number of features required to attain the best performance it is time for choosing the best subset of the 7 features. If the exhaustive search is used there will be 6435 combinations for a combination of 7 out of 15 and they all need to be evaluated to find the best combination that leads to the best performance. It is proposed to use some criteria to prioritize the best subsets. The criteria used in this study are divergence, Bhattacharyya distance and scatter matrices which are the most prevalent in the literature of classification on this matter (Theodoridis and Koutroumbas, 2010). Generally speaking, each of these criteria somehow evaluates the difference that exists between features whether it is distance or dispersion or any other difference. We applied all three criteria on the 15 features and found the best subset of 7 using a Sequential Backward Selection (SBS) technique (Theodoridis and Koutroumbas, 2009). Thus in this way instead of having to do a collection of

Table 4. Chosen feature sets provided using different feature vector selection criteria.

Criterion	7 Chosen Features
Divergence	MDVP: Fo(Hz), MDVP: Fhi(Hz), DFA, spread1, spread2, D2, PPE
Bhattacharyya Distance	MDVP: Fo(Hz), MDVP: Fhi(Hz), MDVP:APQ,spread1,HNR, spread2, DFA
Scatter Matrices	MDVP: Fhi(Hz), spread1, DFA, Shimmer:DDA, RPDE, D2, PPE

NEUROSCIENCE

32767 evaluations for all combinations of 15 features we did a total search of 93 combinations. The calculation of this total search number is based on the formula for SBS searching technique (Duda et al., 2000).

The best combination is found according to some criterion and after that we did the evaluation of performance for each of the three criteria. The computational cost in (Little et al., 2009) is considered of importance any more, using this method. The results are shown in Table 4. For more information on these criteria and also SBS method the reader is referred to (Duda et al., 2000). According to the resulted 7 features chosen by the criteria and the fact that there are a total of 13 specific features mentioned by the three criteria, it can be concluded that the three criteria are highly correlated in result. We decided to choose the features proposed by divergence criterion (based on the best resulting performance of classifier) and will use them in the following section. They are MDVP: Fo(Hz), MDVP: Fhi(Hz), DFA, spread1, spread2, D2 and PPE.

3. Results

Now we have got the best subset of seven features according to divergence criterion. Our next goal is to understand how different classifiers would behave when encountering the chosen data and to compare their performance. Now we are sure that our subset is the best one among all existing subsets and we can focus on the classification part. We used the commercial software MATLAB[1] for all our computations and programs in this study. The adopted classifiers are Support Vector Machine (SVM), K-Nearest Neighbor (KNN) and some discrimination-function-based (DBF) classifiers. For the evaluation of classification there are different methods in the literature that each of them has some pros and some cons. K-fold cross-validation, HoldMout cross validation, Holdout method and Random subsampling

1. MATLAB is a registered trademark of The Math Works, Inc., Natick, MA.

methods are instances (Theodoridis and Koutroumbas, 2009).

We made use of the four above-mentioned methods to evaluate the performance of classifiers firstly to bring about a comparison case and secondly to cover all common methods of evaluation assuring that it provides enough evidence on the dignity of the calculations. There are some free parameters for these methods and for classifiers that have been chosen based on the best resulting performance for classifiers. They are K=3 for K-fold cross-validation, M=1 for Leave-M-out method and considering half of the samples for training and the rest for testing purpose in random sub-sampling. For SVM classifier we used RBF kernel function with the sigma value of 1, in KNN classifier K equals 4 and for DBF classifier we chose diag-quadratic discrimination function with the prior probabilities chosen according to number of samples in each class. According to the fact that the number of measurements are 147 and 48 for class1 (Patient) and class2 (Normal) respectively, prior probabilities are 0.674 and 0.326 respectively.

Table 5. Classification performance for SVM classifier.

SVM	Correct Rate	Error Rate	Sensitivity	Specificity
Random Subs.	0.9175	0.0825	0.6400	0.9861
Holdout	0.9113	0.0887	0.6792	0.9877
K-fold	0.9128	0.0872	0.6875	0.9864
LeaveMout	0.9000	0.1000	0.8000	1
Average	0.9104	0.0896	0.7016	0.9900

NEUROSCIENCE

Results of classification are gathered in Table 5 to 7 for the mentioned classifiers respectively. For each of the classifiers correct rate, error rate, specificity and sensitivity are calculated. These criteria are defined in the following. By definition sensitivity relates to the test's ability to identify positive results. Here, if our test has high sensitivity then a negative result would suggest the absence of disease. For example, a sensitivity of 100% means that the test recognizes all actual positives – i.e. all sick people are recognized as being ill. So does the specificity for negative results.

Table 6. Classification performance for KNN classifier.

KNN	Correct Rate	Error Rate	Sensitivity	Specificity
Random Subs.	0.9533	0.0467	0.8333	1
Holdout	0.9165	0.0835	0.7917	0.9575
K-fold	0.9333	0.0667	0.8750	0.9524
LeaveMout	0.9500	0.0500	0.9000	1
Average	0.9382	0.0617	0.8500	0.9774

NEUROSCIENCE

A high rate of specificity shows the ability of our test to rule out the disease in the subject. Conclusively, the highest rate is desired for all these criteria but the error rate.

Table 7. Classification performance for DFB classifier.

DFB classifier	Correct Rate	Error Rate	Sensitivity	Specificity
Random Subs.	0.8031	0.1969	0.5833	0.8904
Holdout	0.7911	0.2089	0.6375	0.8416
K-fold	0.8103	0.1897	0.6458	0.8639
LeaveMout	0.8833	0.1167	0.8000	0.9667
Average	0.8220	0.1780	0.6666	0.8906

NEUROSCIENCE

As it can be seen in Table 5 SVM classifier has been successful in giving a great specificity and a relatively good correct rate. KNN classifier, however, achieved a better correct rate and a lower specificity at the same time. Discrimination-function-based classifier provides a lower performance in all aspects meaning a kind of failure for that classifier in this special case.

Table 8. Average classification performance for the three classifiers.

Classifier	Correct Rate	Error Rate	Sensitivity	Specificity
SVM	0.9104	0.0896	0.7016	0.9900
KNN	0.9382	0.0618	0.8500	0.9774
DFB	0.8220	0.1780	0.6666	0.8906

NEUROSCIENCE

Table 8 provides a comparison between the three classifiers from different views. As it is obvious, averagely the highest correct rate and sensitivity rate are resulted by the KNN classifier, which are 0.9382 and 0.85 respectively, and the highest specificity is provided by SVM classifier which is 0.99. This means that KNN classifier has the most ability to recognize positive cases of PD while the SVM classifier showed the most able to diagnose negative cases. Generally speaking, based on the correct rate which is the most general criterion for precision, KNN is the best classifier among the mentioned classifiers to differentiate between the people with PD and those without it.

4. Discussion

The main reason behind such studies as ours is to try to computerize the diagnosis process. The more computerized the diagnosis, the less expert intervention needed and so the less subjective diagnosis made. So the main goal of this study was to provide physicians with some diagnostic tool for Parkinson's disease, enabling them to make earlier diagnosis. This diagnosis would not be subjective as well. Subjective interpretation of a disease can be the result of factors such as experience, knowledge and even the mood of that physician at the time of diagnosis which is not desirable. It was tried to provide diagnosis through speech signals. By providing a basic concept for feature selection in this matter not only did our study reduced the computational costs (especially when coping with bigger databases) of previous works but also it improved the classification performance which is equivalent to a better diagnosis. The satisfactory performance of classification which is achieved means implicitly the self sufficiency of the method and the lack of need for physician's intervention. That is the desired result.

It is noted that the features of speech signals used in this study were extracted so as to expose some malfunctions of nervous system in PD. Different features can be extracted each of which might be representing some other problems. This piece of work can be used as a monitoring tool for telemedicine if used regularly by the physician at a remote surgery. In this way voice signals are checked for small changes during the time. This is our future goal to achieve.

References

Lang A.E., Lozano A. M. (1998). Parkinson's disease - First of two parts. New England Journal of Medicine 339:1044-1053.

Van Den Eeden S. K., Tanner C. M., Bernstein A. L., Fross R. D., Leimpeter A., Bloch D. A., Nelson L. M. (2003). Incidence of Parkinson's disease: Variation by age, gender, and Race/Ethnicity. American Journal of Epidemiology 157:1015-1022.

Singh N., Pillay V., Choonara Y. E. (2007). Advances in the treatment of Parkinson's disease. Progress in Neurobiology 81:29-44.

Ho A. K., Iansek R., Marigliani C., Bradshaw J. L., Gates S. (1998). Speech impairment in a large sample of patients with Parkinson's disease. Behavioural Neurology 11:131-137.

Logemann J. A., Fisher H. B., Boshes B., Blonsky E. R. (1978). Frequency and Co-Occurrence of Vocal-Tract Dysfunctions in Speech of a Large Sample of Parkinson Patients. Journal of Speech and Hearing Disorders 43:47-57.

Rahn D. A., Chou M., Jiang J. J., Zhang Y. (2007). Phonatory impairment in Parkinson's disease: Evidence from nonlinear dynamic analysis and perturbation analysis. Journal of Voice 21: 64-71.

Baken R. J., Orlikoff R. F. (2000). Clinical Measurement of Speech and Voice (2nd ed.). San Diego: Singular Thomson Learning.

Dejonckere P. H., Bradley P., Clemente P., Cornut G., Crevier-Buchman L., Friedrich G., Van De Heyning P., Remacle M., Woisard V. (2001). 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). Eur Arch Otorhinolaryngol 258:77-82.

Michaelis D., Frohlich M., Strube H.W. (1998). Selection and combination of acoustic features for the description of pathologic voices. Journal of the Acoustical Society of America 103(3):1628-1639.

Carding P.N., Stecn I.N., Webb A., Mackenzie K., Deary I.J., Wilson J.A. (2004). The reliability and sensitivity to change of acoustic measures of voice quality. Clinical Otolaryngology 29(5):538-544.

Kantz H., Schreiber T. (1999). Nonlinear Time Series Analysis. New York: Cambridge University Press.

Little M. A. (2007). Biomechanically Informed Nonlinear Speech Signal Processing. D.Phil. Thesis, University of Oxford, Oxford.

Little M.A., McSharry P.E., Hunter E.J., Ramig L.O. (2007). Exploiting Nonlinear Recurrence and Fractal Scaling Properties for Voice Disorder Detection. BioMedical Engineering Online, 6:23.

Little M.A., McSharry P.E., Hunter E.J., Ramig L.O. (2009). Suitability of dysarthria measurements for telemonitoring of Parkinson's Disease. IEEE Transactions on Biomedical Engineering, 56(4):1015-1022.

Boersma P., Weenink D. (2005). Praat: doing phonetics by computer (Version 4.3.14).

Hoehn M. M., Yahr M. D. (1967). Parkinsonism - Onset Progression and Mortality. Neurology 17 (5): 427-442.

Boersma P., Weenink D. (2001). Praat, a system for doing phonetics by computer. Glot International 5:341-345.

Boersma P. (1993). Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound. Proceedings of the Institute of Phonetic Sciences,17: University of Amsterdam.

Schoentgen J., Deguchteneere R. (1995). Time-Series Analysis of Jitter. Journal of Phonetics 23:189-201.

Cnockaert L., Schoentgen J., Auzou P., Ozsancak C., Defebvre L., Grenez F. (2008). Low-frequency vocal modulations in vowels produced by Parkinsonian subjects. Speech Communication 50:288-300.

Theodoridis S., Koutroumbas K. (2010). Introduction to pattern recognition: a MATLAB® approach. Academic Press Publications.

Theodoridis S., Koutroumbas K. (2009). Pattern recognition (4th ed.). Academic Press Publications.

Duda R.O., Hart P.E., Stork D.G. (2000). Pattern classification (2nd ed.). Wiley Publications.