❏   4265

Voice Assessments for Detecting Patients with Parkinson's Diseases in Different Stages

Elmehdi Benmalek, Jamal Elmhamdi, Abdelilah Jilbab
Department of Electrical Engineering, ENSET Rabat, Mohamed V University, Morocco

Article Info	ABSTRACT
	Recently, a wide range of speech signal processing algorithms (dysphonia measures) aiming to detect patients with Parkinson's disease (PD). So we have computed 19 dysphonia measures from sustained vowels collected from 375 voice samples from healthy and people suffer from PD. All the features are analysed and the more relevant ones are selected by the Principal component analysis (PCA) to classify the subjects in 4 classes according to the UPDRS (unified Parkinson's disease Rating Scale) score. We used k-folds cross validation method with (k=4) validation scheme; 75% for training and 25% for testing, along with the Support Vector Machines (SVM) with its different types of kernels. The best result obtained was 92.5% using the PCA and the linear SVM.

Corresponding Author:

Elmehdi Benmalek,
Department of Electrical Engineering,
ENSET Rabat, Mohamed V University,
Rabat, Morocco.
Email: elmehdi.benmalek@um5s.net.ma

1. INTRODUCTION

Parkinson's disease (PD) is a neurodegenerative disorder that results from the death of dopaminergic cells in the substantia nigra which is a basal ganglia structure located in the midbrain. Such neurological diseases profoundly affect the patients' quality of life and their families [1]. Age is one of the most important risk factor which explain that PD is generally seen in people over the age of 50. Diagnosis of PD is very difficult we use neurological tests and brain scans to diagnose it. These methods are very expensive and need high level of expertise.

Since most of the people with PD suffer from speech disorders [2], [3], it could be considered as the most reasonable way for detection of PD [4]. The range of symptoms present in speech disorders includes reduced loudness, increased vocal tremor, and breathiness. Vocal disorders do not appear abruptly, they are the result of a slow evolution whose early stages may be unnoticed. Voice assessments has proven to be an effective tool for PD detection, for this purpose, the processing of the quality of speech, and the identification of the causes of its degradation in the context of PD based on phonological and acoustic cues have become one of the main interest of clinicians and speech pathologists.

Among the most interesting recent works are those concerned with class of neurodegenerative diseases such as PD, multiple sclerosis among other, that affect motor, cognitive capabilities, and patient's speech [5], [6]. There are recent studies using machine learning tools such as Support Vector Machine (SVM) classifier, Gaussian radial basis kernel functions, regression, neural networks, DMneural and decision tree [7], [8], and acoustic measurements (features) of dysphonia for the detection of voice disorders, these include fundamental frequency or pitch of vocal oscillation (F0); Jitter which is the cycle-to-cycle variation of fundamental frequency; Shimmer that represents the extent of variation in speech amplitude from cycle to cycle; measures of noise-to-harmonics ratio components in the voice; the Nonlinear dynamical complexity

and fundamental frequency variation and Signal fractal scaling exponent [1], [4], [9]. Studies have shown variations in all these measurements in people with PD [10]. All these studies has been performed for binary classification, so for an early diagnosis of PD, multiclass classification based on severity of symptoms has been achieved with different classifiers using the Local Learning-Based Feature Selection feature selection algorithm and the cepstral analysis [11], [12],

In this study, we want to distinguish PD patients on different stages of symptoms' severity from healthy control using these acoustic measurements. So we aimed to discriminate 375 subjects on 4 groups; 55 healthy control, 178 in early 118 in intermediate and 24 in advanced stage according to the UPDRS scores. Each participant was invited to pronounce the sustained vowel /a/ and hold it at comfortable level, from each voice sample we have extracted 19 acoustic features, to reduce the number of these acoustic features and get only the most relevant ones, we applied the principal component analysis, and for classification we used k-folds cross validation method along with the SVM classifiers with its different kernels.

2.   RESEARCH METHOD
2.1. Dataset
The dataset collected in this study belong to The Patient Voice Analysis (PVA) dataset [8], [13], it contains voice recordings of voice phonations self-reported symptom assessment PDRS (Parkinson's Disease Rating Scale) and demographic information about the callers. Each row in the dataset corresponded to one report from a Parkinson's patient and the dysphonia measurements are represented in the columns. There are 375 users total (repeated and useless records are removed). All participants were asked to record the sustained vowel "a" hold as long as possible at a comfortable level. They also provided the following information; age, gender, age of diagnosis, years since first symptom, if they are on treatment or not, with (mean 62.17 years old, maximum 84 and minimum 34, standard deviation: 8.370254, variance: 69.88011, popular standard deviation: 8.359432, variance popular: 67.9286).

Among 375 persons for which the data were recorded, we classify 55 subjects as healthy, 178 in early stage, 118 in intermediate stage, and 24 as advanced stage based on UPDRS scores. Voice recordings and the pre-processing are not sufficient in the assessment of voice disorders. Therefore, it is essential to devise and describe voice samples using a set of acoustic features, which are represented as a feature vector used for speech analysis.

2.2. Feature extraction
In this dataset, 19 linear and non-linear features were extracted. Table 1 contains all the features and a brief descriptions [14]. 16 features are based on four factors: F0 (fundamental frequency or pitch), several measures of variation in fundamental frequency and amplitude and measures of ratio of noise to tonal components in the voice, these measurements are the most important factors of the voice signal.

Table 1. Features Extracted

Feature number	Features	Description
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	Jitter (%)	Several measures of variation in fundamental frequency
5	Jitter (Abs)	
6	MDVP: RAP	
7	MDVP: PPQ	
8	Jitter: DDP	
9	Shimmer	Several measures of variation in amplitude
10	Shimmer (dB)	
11	Shimmer: APQ3	
12	Shimmer: APQ5	
13	MDVP: APQ	
14	Shimmer: DDA	
15	NHR	Two measures of ratio of noise to tonal components in
16	HNR	the voice
17	RPDE	Nonlinear dynamical complexity measures
18	DFA	Signal fractal scaling exponent
19	PPE	Nonlinear measure of fundamental frequency variation

Jitter (%): Expressed as a percentage, this is the average absolute difference between consecutive periods of fundamental frequency, divided by the average period

$$Jitter(\%) = \frac{\frac{1}{N}\sum_{i=1}^{N-1}|T_i - T_{i-1}|}{\frac{1}{N}\sum_{i=1}^{N}T_i} \tag{1}$$

Where T_i is the period of fundamental frequencies of window number "i" and N is the total number of windows. Jitter (ABS): Jitter absolute is the cycle-to-cycle variation of fundamental frequency, i.e. the average absolute difference between consecutive periods, expressed as:

$$Jitter(ABS) = \frac{1}{N-1}\sum_{i=1}^{N-1}|T_i - T_{i-1}| \tag{2}$$

Where T_i is the extracted F0 period lengths, and N are is the number of extracted F0 periods. Jitter (RAP): it is defined as the Relative Average Perturbation, the average absolute difference between a period and the average of it and its two neighbours, divided by the average period.

Jitter (PPQ) represents the Period Perturbation Quotient, defined as the average absolute difference between a period and the average of it and its four closest neighbors, divided by the average period [15],[16]. Shimmer: This is the average absolute difference between the amplitudes of consecutive periods, divided by the average amplitude

$$Shimmer = \frac{\frac{1}{N-1}\sum_{i=1}^{N-1}|A_i - A_{i-1}|}{\frac{1}{N}\sum_{i=1}^{N}A_i} \tag{3}$$

Shimmer (APQ5): It is defined as the five-point Amplitude Perturbation Quotient, the average absolute difference between the amplitude of a period and the average of the amplitudes of it and its four closest neighbours, divided by the average amplitude. HNR: Harmonics to Noise Ratio, NHR: Noise to Harmonics Ratio.

Recurrence Periodicity Density Entropy (RPDE) is based on the notion of recurrence [17], which can be seen as a generalization of periodicity [18]. This measure addresses the ability of the vocal folds to sustain stable vocal fold oscillation, quantifying the deviations from exact periodicity. Pitch Period Entropy (PPE) measures the impaired control of stable pitch during sustained phonations [1], a symptom common to people with PD [19]. Detrended Fluctuation Analysis (DFA) is a scaling analysis method used to quantify long range power-law autocorrelations in signals which are non-stationary, thus overcoming some of the problems of scaling analysis techniques which are only suitable for stationary signals [18],[20].

2.3. Feature selection and validation

In most situations, we find ourselves with a number of variables which tends to exceed the number of observations. Dimensionality reduction process proceeds by applying a feature selection algorithm. In order to have a better representation of the data, redundant and useless information will be thus circumvented. The principal objectives of the reduction of dimension can be described by [21]. So to improve the task of classification and to aid the visualization and the comprehension of the data, we have to identify the more relevant features in order to reduce the storage of space necessary, minimize time consumption and CPU-expenditure.

However, the elimination of certain information can increase the classification error, considering this information can prove to be informative if they are used [22]. In this study we used the Principal Component Analysis (PCA), which considered the more recognized linear technique for dimensionality reduction, the PCA performs a linear mapping of the data to a lower-dimensional space in such a way that the variance of the data in the low-dimensional representation is maximized. Previous speech analysis has shown satisfactory results using this reduce dimensionality method [23].

After extracting all features and selecting the more relevant ones, we classify voice samples based on these features into four groups; Healthy cases, people with PD in early, intermediate and advanced stages. Subsequently, we built a matrix based on these parameters. The columns of the matrix represent the features and the rows represent the voice samples. In this study, we used k-folds cross validation method with (k=4) along with different kernel of the SVM classifier; Training and testing procedures are applied: 75% for training and 25% for testing. The dataset is divided into 4 subsets, each time, one of the 4 subsets is used as the test set and the other 3 subsets are put together to form a training set. Then the average error across all 4 trials is computed. The advantage of this method is that it matters less how the data gets divided, every data point gets to be in a test set exactly once, and gets to be in a training set 3 times.

3.   RESULTS AND ANALYSIS
3.1. Obtained results using linear kernel
The Table 2 represent the obtained results of classification using the linear SVM, and selecting the more relevant features by the PCA method, with 92.5% overall accuracy. For each class we have the ROC curve. In this model we have for:
1. The healthy control: we have 49 were correctly classified, 6 were misclassified and considered as early stage, with a percentage of 89% true positive rate;
2. The early stage class: we have 171 were correctly classified, 7 were misclassified (2 as healthy, and 5 as intermediate stage), with a percentage of true positive rate 96%;
3. The intermediate stage class: we have 113 were correctly classified, 5 were misclassified (4 as in early stage, and 1 as advanced stage) with a percentage of 96% true positive rate;
4. The advanced stage class: we have 14 were correctly classified, 10 were misclassified (as in intermediate stage), with a percentage of 58% true positive rate.

Table 2. Results Using Linear SVM

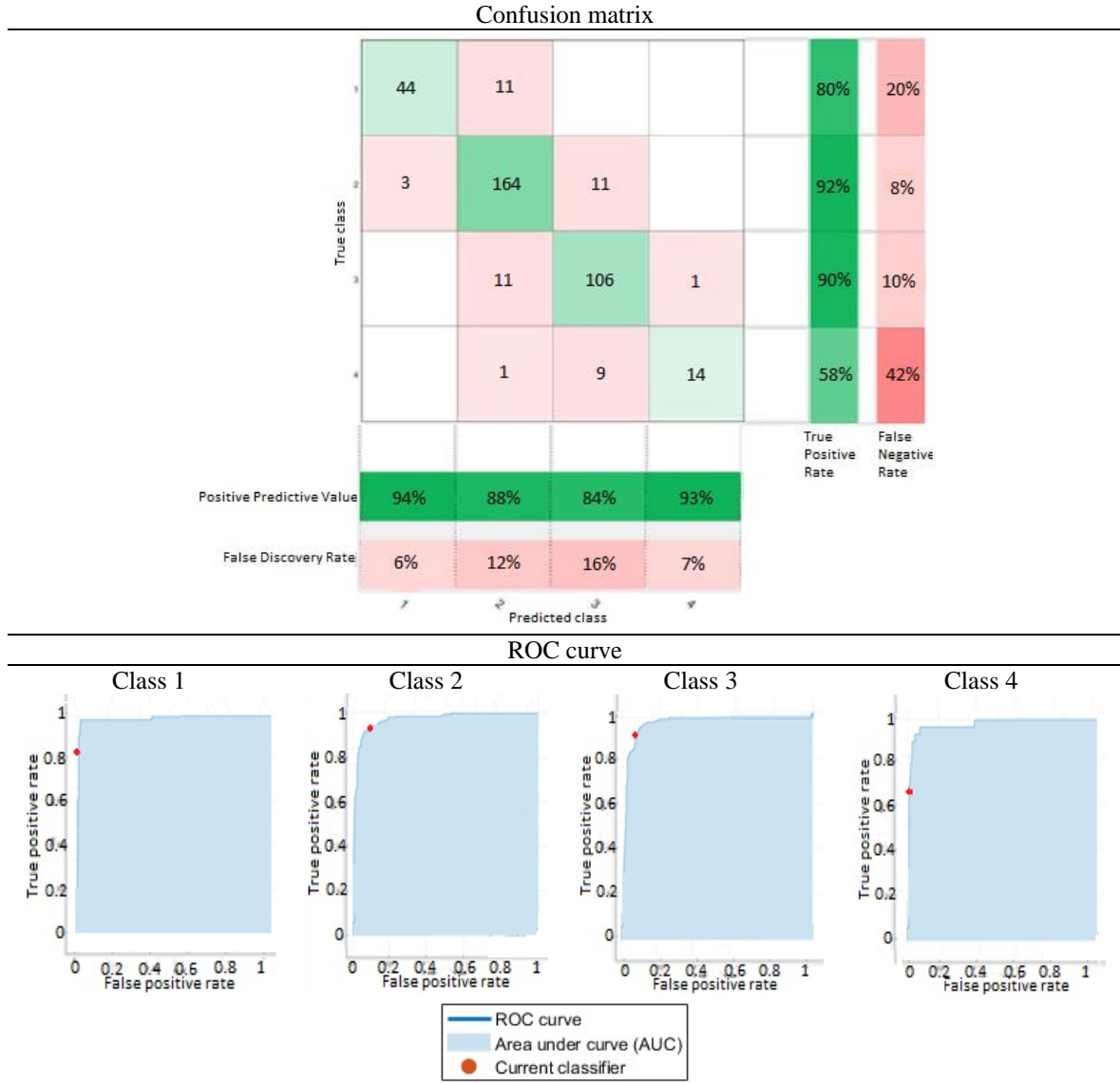Confusion matrix
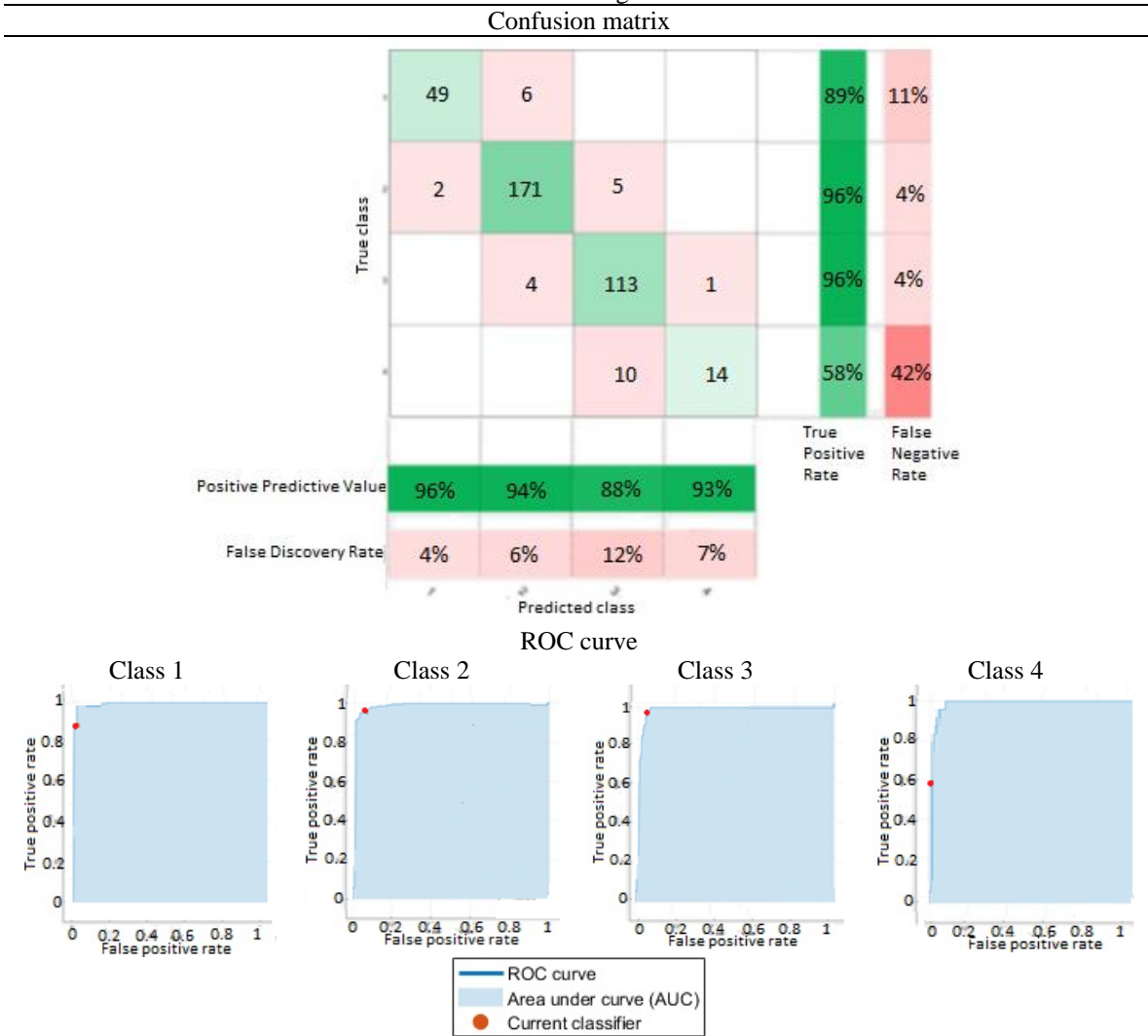
ROC curve

3.2. Obtained results using quadratic kernel
The Table 3 represent the obtained results of classification using the quadratic SVM and the PCA, with accuracy of 87.5%. For each class we have the ROC curve. In this model we have for:
1. The healthy control: we have 44 were correctly classified, 11 were misclassified (all as in early stage), with a percentage of 80% true positive rate;

2. The early stage class: we have 164 were correctly classified, 12 were misclassified (3 as healthy, and 11 as intermediate stage), with a percentage of true positive rate 92%;
3. The intermediate stage class: we have 106 were correctly classified, 12 were misclassified (11 as in early stage, and 1 as advanced stage) with a percentage of 90% true positive rate;
4. The advanced stage class: we have 14 were correctly classified, 10 were misclassified (1 as in early stage and 9 as in intermediate stage,), with a percentage of 58% true positive rate.

Table 3. Results using quadratic SVM

Confusion matrix

ROC curve

| Class 1 | Class 2 | Class 3 | Class 4 |

ROC curve
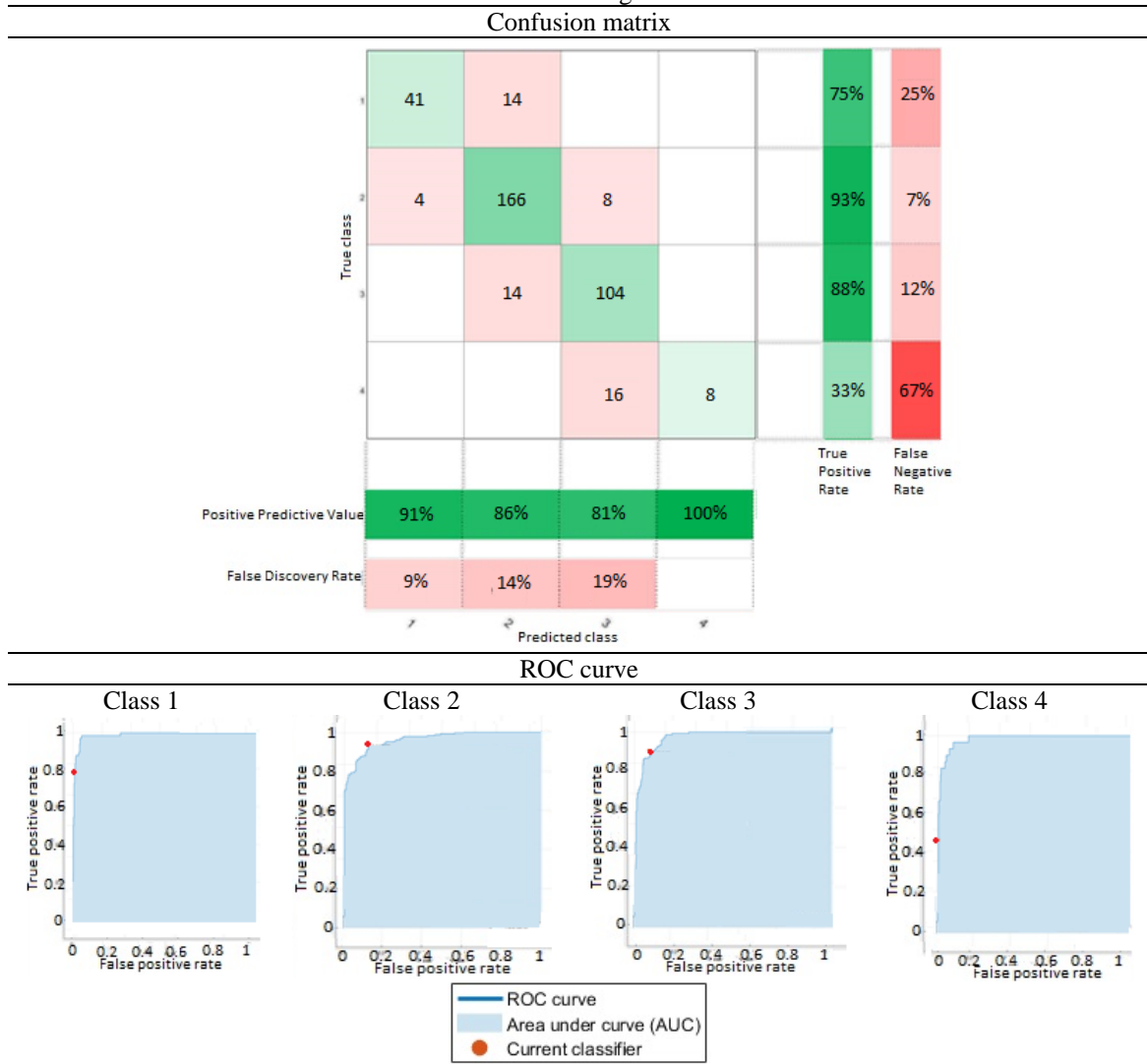Area under curve (AUC)
Current classifier

3.3. Obtained results using cubic kernel

The Table 4 represent the obtained results of classification using the cubic SVM, and selecting the more relevant by the PCA, with accuracy of 85.1%. For each class we have the ROC curve. In this model we have for:
1. The healthy control: we have 41 were correctly classified, 14 were misclassified (all as in early stage), with a percentage of 75% true positive rate;
2. The early stage class: we have 166 were correctly classified, 14 were misclassified (4 as healthy, and 8 as intermediate stage), with a percentage of true positive rate 93%;
3. The intermediate stage class: we have 104 were correctly classified, 14 were misclassified (all as in early stage) with a percentage of 88% true positive rate;
4. The advanced stage class: we have 8 were correctly classified, 16 were misclassified as in intermediate stage, with a percentage of 33% true positive rate.

Table 4. Results using cubic SVM

Confusion matrix

ROC curve

Class 1	Class 2	Class 3	Class 4

ROC curve
Area under curve (AUC)
Current classifier

From all previous results, it is seen that the maximum classification accuracy of 92.5% was achieved using the linear SVM. Compared with previous studies done, the proposed method give better results than the cepstral analysis approach (86.7%) [12], but this findings could be improved by using feature selection algorithm dedicated for multiclass classification and combinig the voice features with the cepstral analysis where a score of 96% has been achieved in [11], but the approach was more complex than the one proposed in this study. The results show also that the feature selection play critical role in classification optimization. And the misclassification is explained by the relative merits of the UPDRS scale for accurately determining the degree of disease progression. The purpose of this study is to show the effectiveness of using voice recording to classify people with Parkinson's disease by the severity of symptoms using only 19 features.

4. CONCLUSION

Clinicians and voice pathologists have become progressively watchful to any techniques, which might provide supplementary information to help them in the evaluation and the diagnosis of PD. In this paper, we presented new technique that can separate between healthy people and PD patients at different severity stages based on voice features. As a result, we achieved 92.5% of accuracy using linear SVM and the PCA. The results show also that the feature selection play critical role in classification optimization. And the misclassified samples are usually mingled with the nearest class, which clinically explained by the relative merits of the UPDRS scale for accurately determining the degree of disease progression. These results are very encouraging, in future works we consider to determine correlation between the voice

disorders and the symptoms, which will be of great help to the medicine and could also extended for other voice pathologies.

ACKNOWLEDGEMENTS

REFERENCES

[1]   M. A. Little, *et al.*, "Suitability of dysphonia measurements for telemonitoring of Parkinson's disease," *Biomedical Engineering, IEEE Transactions on*, vol/issue: 56(4), pp. 1015- 1022, 2009.

[2]   A. Ho, *et al.*, "Speech impairment in a large sample of patients with Parkinson's disease," *Behavioral Neurology*, vol. 11, pp. 131–137, 1998.

[3]   J. A. Logemann, *et al.*, "Frequency and co-occurrence of vocal-tract dysfunctions in speech of a large sample of Parkinson patients," *Journal of Speech Hearing Disorder*, vol. 43, pp. 47–57, 1978.

[4]   D. A. Rahn, *et al.*, "Phonatory impairment in Parkinson's disease: Evidence from nonlinear dynamic analysis and perturbation analysis," *J. Voice,* vol. 21, pp. 64-71, 2007.

[5]   V. Parsa and D. G. Jamieson, "Interactions between speech coders and disordered speech," *Speech Communication*, vol/issue: 40(7), pp. 365–385, 2003.

[6]   S. B. Davis, "Acoustic characteristics of normal and pathological voices," *Speech and Language: Advances inBasic Research and Practice*, vol. 1, pp. 271–335, 1979.

[7]   R. Das, "A comparison of multiple classification methods for diagnosis of Parkinson disease," *Expert Systems with Applications*, vol/issue: 37(2), pp. 1568–1572, 2010.

[8]   A. Tsanas, *et al.*, "Novel speech signal processing algorithms for high-accuracy classification of Parkinson's disease," *IEEE Transactions on Biomedical Engineering*, vol/issue: 59(5), pp. 1264-1271, 2012.

[9]   M. A. Little, *et al.*, "Exploiting nonlinear recurrence and fractal scaling properties for voice disorder detection," *Biomed. Eng. Online*, 2007.

[10]   P. Zwirner, *et al.*, "Phonatory function of neurologically impaired patients," *Journal of communication disorders*, vol/issue: 24(4), pp. 287300, 1991.

[11]   E. Benmalek, *et al.*, "Multiclass classification of Parkinson's disease using different classifiers and LLBFS feature selection algorithm," *International Journal of Speech Technology*, vol/issue: 20(1), pp. 179-184, 2017.

[12]   E. Benmalek, *et al.*, "Multiclass classification of Parkinson's disease using cepstral analysis," *International Journal of Speech Technology*, pp. 1-11, 2017.

[13]   Patient Voice Analysis (PVA) Synapse ID: syn2321745 https://www.synapse.org,

[14]   V. J. Philippe, *et al.*, "Kernel Methods in Computational Biology," MIT Press, Cambridge, 2004.

[15]   M. Farrús, *et al.*, "Jitter and shimmer measurements for speaker recognition," *INTERSPEECH,* 2007.

[16]   R. A. Shirvan and E. Tahami, "Voice analysis for detecting Parkinson's disease using genetic algorithm and KNN classification method," *Biomedical Engineering (ICBME), 2011 18th Iranian Conference of. IEEE*, 2011.

[17]   H. Kantz and T. Schreiber, "Nonlinear time series analysis," Cambridge University Press, 2nd edition, 2004.

[18]   M. A. Little, *et al.*, "Exploiting Nonlinear Recurrence and Fractal Scaling Properties for Voice Disorder Detection," *Biomedical Engineering Online,* vol/issue: 6(23), 2007.

[19]   L. Cnockaert, *et al.*, "Low frequency vocal modulations in vowels produced by Parkinsonian subjects," *Speech Communication*, vol. 50, pp. 288-300, 2008.

[20]   Z. Chen, *et al.*, "Effect of nonstationarities on detrended fluctuation analysis," *Physical Review E*, vol/issue: 65(4), pp. 041107, 2002.

[21]   Guérif S., "Réduction de dimension en apprentissage numérique non supervisée," PhD thesis, Université Paris 13, pp. 420 148, 2006.

[22]   Ferchichi S. E., *et al.*, "Feature selection using an SVM learning machines," *Proceedings of the 422 3rd International Conference on Signals, Circuits and Systems (SCS 2009),* pp. 1-6, 2009.

[23]   A. Benba, *et al.*, "Voice assessments for detecting patients with Parkinson's diseases using PCA and NPCA," *International Journal of Speech Technology*, vol/issue: 19(4), pp. 743–75412, 2016.