

A Comparative Study of Neural Networks and Support Vector Machines for Neurological Disordered Voice Classification

K. Uma Rani

Dept. of Biomedical Engineering
B.I.E.T
Davangere, India

Mallikarjun S. Holli

Dept. of Electronics and Instrumentation Engineering
University B.D.T. college of Engineering
Davangere, India

Abstract— Acoustical voice characteristics in neurological diseases might provide useful biomarkers for the detection of diseased voices. This paper presents a method for automatic detection of neurological disordered voices like Parkinson's disease; PD, cerebellar demyelination and stroke using the time domain features. The 16 features extracted in time domain were given to multilayer neural network and trained to classify whether the voice was neurological disordered or normal subject voice. Later the same features were given to the support vector machine; SVM for training and to classify as neurological disordered or normal subject voice for a comparative study. There are no risks involved in capturing and analysis of voice signals as it is noninvasive by nature and in carefully controlled circumstances, it can provide a large amount of meaningful data. The data collected in the present work consist of 263 sustained vowel phonations ; /ah/, among them 165 phonations are from 69 numbers of patients suffering from different neurological diseases and 98 phonations from 56 numbers of controlled subjects including both male and female subjects. 200 phonations were used to train the network and 63 phonations for testing. The best classification accuracy achieved for test data using multilayer neural network was 81.68% and with SVM was 86.11%. Hence, the time domain parameters with SVM classifier will give better classification of normal and neurological disordered voices.

Keywords— Time domain; neural network; support vector machine; neurological disorders; phonation; voice; sustained vowel; classification.

I. INTRODUCTION

Acoustic measures of vocal function are routinely used in the assessments of disordered voice. There are a great range of diseases that causes changes in the voice. They could appear as a modification of the excitation morphology, like the distribution of mass on vocal fold is increased and produce irregular vibration pattern. These are classified as organic pathologies such as nodules, polyps, cysts and edemas. Voice disorders can also be caused by other pathologies which are provoked by neurodegenerative diseases like Parkinson's disease (PD), stroke and cerebellar demyelination. The methods used by the medical community to evaluate the speech production system to diagnose pathologies are direct ones, which require inspection of vocal folds using laryngoscopic techniques such as fiberscope, which causes discomfort to the patient. The other method is

subjective, in which voice quality is evaluated by a doctor's audition. The measures obtained from recorded voice data allow in quantifying the voice quality by objective method [1], and such methods are very appealing due to their noninvasive nature. The extraction of measures from sustained vowel samples is common because of its simpler acoustic structure. Hence, in the present work sustained vowel /ah/ is used as voice sample [2].

There are extensive studies on speech measurement and analysis of voice in normal and PD [3], [4], [5]. In the standard tests, speech signals are recorded using a microphone and subsequently analyzed using quantitative measurement techniques to detect certain characteristics of speech signals. The traditional measurement methods to characterize the speech signal includes F_0 (the fundamental frequency or *pitch* of vocal oscillation), absolute sound pressure level (indicating the relative loudness of speech), *jitter* (the extent of variation in speech F_0 from one vocal cycle to other), *shimmer* (the extent of variation in speech amplitude from cycle to cycle), and *noise-to-harmonics ratios* (the amplitude of noise relative to tonal components in the speech) [6]. The earlier studies have shown variations in all these measurements for comparison of healthy controls to PD patients, indicating that these could be useful measures in assessing the extent of disorder in voice [1], [7]. Most of the earlier studies on speech analysis in neurological diseases have concentrated only on the PD subjects [8], [9], [10], [11]. In the present work an attempt has been made to extend these measures and analysis to other neurological disorders like stroke and cerebellar demyelination along with study on PD.

The use of classifier systems in medical diagnosis is increasing gradually. Recent advances in the field of artificial intelligence have led to the emergence of expert systems and decision support systems (DSS) for medical applications. However, expert systems and different artificial intelligence techniques for classification have the potential of being good supportive tools for the medical field. Classification systems can help in increasing accuracy and reliability of diagnosis and minimizing possible errors, as well as making the diagnosis more time efficient [12].

In the present work a system has been developed using artificial neural networks (ANNs) and support vector machines (SVMs) to distinguish neurological disordered subjects from normal subjects using the above said time domain measures extracted from the voice samples. A comparison of these two classifiers using these features for their performance is evaluated.

II. MATERIALS AND METHODS

A. Data Collection

The present work consists of 263 phonations of sustained vowel /ah/. Among them 165 phonations were collected from 44 male subjects (62.72 ± 8.0 yrs) and 25 female subjects (65.19 ± 8.8 yrs), who are found to be suffering from one or the other neurological disorder like PD, cerebellar demyelination and stroke. Remaining 98 phonations were from 56 normal subjects, who were selected among the age and gender-matched healthy persons who were not complaining of any voice problems. The data were collected from Outpatient Wing, Department of Neurology, J.S.S. Hospital, Mysore.

Voice signals are recorded as per the earlier standards through a microphone at a sampling frequency of 44,100 Hz using a 16-bit sound card in a laptop computer with a Pentium processor [13], [14]. The microphone to mouth distance was at 5 cm and the subjects were asked to phonate the vowels /ah/ for at least 3 sec at a comfortable level. Further, a steady portion of the signal of 2 sec duration was selected for the acoustic analysis. All the recordings were done in mono-channel mode and saved in WAVE format on the hard disk and acoustic analysis were done on these recordings [15], [16].

B. Acoustic Parameter

The time domain features in our study include three measures on fundamental frequency, fundamental frequency perturbations, five measures on *jitter*, amplitude perturbations, six measures on *shimmer*, and two measures on signal to noise ratios (*harmonics to noise ratio*) [17]. All 16 acoustic features are summarized in Table I. A free shareware program 'PRAAT' was used for extracting all the 16 features from the voice samples [15]. The duration between two successive openings or closures of the vocal folds defines a vocal fold cycle, where the vocal fold oscillation pattern (vocal fold opening and closure) is typically considered nearly periodic in healthy voices. That is, the intervals of time where the vocal folds are apart or in collision remain almost equal between successive cycles. This speech oscillation interval is called pitch period or fundamental frequency, whereas in voice pathologies this pattern may be severely affected. In addition, a common manifestation of vocal impairment is incomplete vocal fold closure, resulting in excessive breathiness (noise). This imbalanced vocal fold movement also results in turbulent noise and the appearance of vortices in the airflow from the lungs as shown in Fig. 1. In general, people with voice disorders cannot elicit steady phonations.

C. Statistical analysis

Differences in group means for the features tabulated in Table I was evaluated using two-sample Student's *t-test*. The significant differences were found between the mean values of neurological patients and controlled subjects with *p* value as indicated, jitter (%) ($p < 0.005$), jitter (Abs), RAP, PPQ, DDP, ($p < 0.01$), Shimmer, Shimmer (dB), Shimmer: APQ3, Shimmer: APQ5, Shimmer: APQ11, Shimmer: DDA ($p < 0.001$), NHR, HNR ($p < 0.05$) except for F_0 ($p = 0.585$), Flo ($p = 0.760$) and Fhi ($p < 0.480$) significant differences were not found [18].

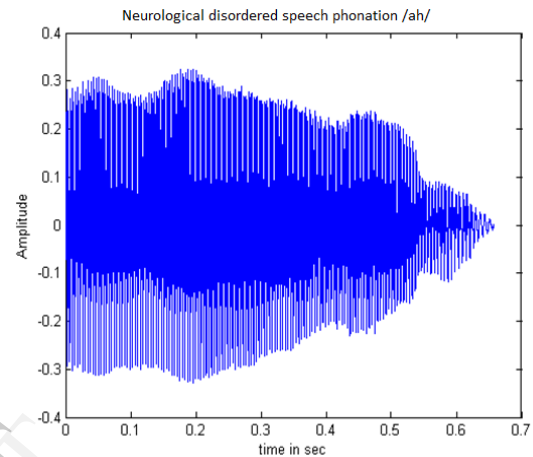


Fig1. Typical sustained phonation /ah/ of a neurological disordered subject. (PD)

TABLE I
DESCRIPTION OF ACOUSTIC FEATURES

FEATURE	DESCRIPTION
F_0 (Hz)	Mean pitch
Flo(Hz)	Minimum pitch
Fhi (Hz)	Maximum pitch
Jitter (%)	Fundamental frequency perturbation (%)
Jitter (Abs)	Fundamental frequency perturbation (absolute)
RAP	Relative Amplitude Perturbation
PPQ	Five-point Period Perturbation Quotient
DDP	Average absolute difference of differences between cycles, divided by the average period
Shimmer	Shimmer Local amplitude perturbation
Shimmer (dB)	Local amplitude perturbation (decibels)
Shimmer: APQ3	Three point Amplitude Perturbation Quotient
Shimmer: APQ5	Five point Amplitude Perturbation Quotient
Shimmer: APQ11	11-point Amplitude Perturbation Quotient
Shimmer: DDA	Average absolute difference between consecutive differences between the amplitudes of consecutive periods
NHR	Noise-to-Harmonics Ratio
HNR	Harmonics-to-Noise Ratio

D. Classifier Models

i. Multilayer Perceptron (MLP)

ANN may have multiple layers of neurons and the architecture may either be feedback or feed forward structure. Depending upon the complexity of the class boundaries to be distinguished the layer of the network is decided. In our study

we have considered a three layer MLP as shown in Fig.2. A three-layer MLP network consist of an input layer which does not perform any processing and the number of neurons are usually set to the number of input parameters, a hidden layer is connected to all neurons in the next layer by weighted connections. These neurons compute weighted sums of their *inputs and add a threshold. The resulting sums are used to calculate the activity of the neurons by applying a sigmoid activation function. The output of each neuron can be calculated by means of*

$$H_k = f\left(\sum_{i=1}^{N_h} w'_{ik} \cdot I_i + \xi_i\right) \text{ and}$$

$$O_k = f\left(\sum_{k=1}^{N_o} W''_{kj} \cdot h_k + \theta_k\right) \quad (1)$$

Where I_i are the input features, θ_k and ξ_i are the thresholds, w'_{ik} are the weights associated to the hidden layer, W''_{kj} are the weights associated to the output layer, O_k are the net outputs, N_h is the number of neurons in the hidden layer, N_o is the number of neurons in the output layer and $f(\cdot)$ is the sigmoidal function [1].

The sigmoidal function is given as

$$\phi(u) = \tanh(u) = \frac{2}{1+e^{-2u}} - 1 \quad (2)$$

The operation of the network consists of a forward pass in which the outputs and the error at the output units are calculated. The error at the hidden nodes is calculated by back-propagating the error at the output units through the weights (backward pass), and finally the weights are adjusted using the back-propagated errors. For each data pair to be learned a forward and a backward pass is performed. These steps are repeated over and over again until a given function of the error arrives to a pre-established level. Within such architecture, the hidden layer learns to recode (or to provide a representation for) the input-output pairs. The MLP architecture is more powerful than single-layer networks [19]. In the present study, input layer consist of 16 numbers of input neurons which is equal to the number of features extracted, single hidden layer where the number of hidden neurons is a parameter to be adjusted during the training phase. The output layer has a single neuron to make the final decision about the presence (1) or absence (0) of the disordered voice. Choosing the net size is a critical problem, hence, numerous experimentations were done to arrive at the number of hidden units that allows good generalization capability as 40 units.

Training is carried out using the backpropagation algorithm with adaptive momentum and learning rate over 100 epochs. At this point the sum means squared error will be used as a measurement to control training termination. Weights are randomly initialized.

Support Vector Machine (SVM) is a new approach to classification standards and has recently attracted great interest in the scientific community, specifically in the areas of machine classification, regression and learning. SVM

theory was first introduced by Vapnik [18]. SVM is a useful technique for data classification. A classification task usually involves with training and testing data which consist of some data instances. Each instance in the training set contains one target values and several attributes.

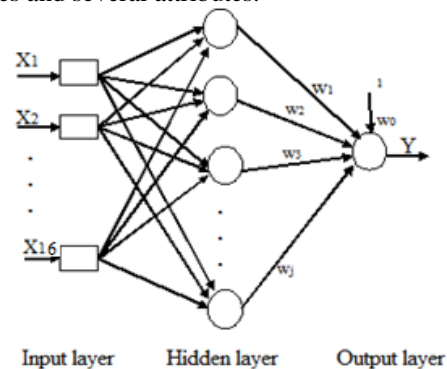


Fig. 2 A feedforward neural network with one hidden layer

ii. Support Vector Machine

The goal of SVM is to produce a model which predicts target value of data instances in the testing set which are given only the attributes [20], [21]. The classification in SVM is an example of supervised learning. A step in SVM classification involves identification of features which are intimately connected to the known classes. This is called feature selection or feature extraction. Feature selection and SVM classification together have a use even when prediction of unknown samples is not necessary. They can be used to identify key sets which are involved in whatever processes distinguish the classes. The SVM maps the input space to a high dimensional space. By calculating an optimal separating hyperplane in this new space, the SVM learns the border between areas belonging to both classes. The separating hyperplane is chosen to maximize separation distance between the closest training samples. SVM models were initially defined to classify linearly separable classes. An example of two linearly separable classes is shown in Fig. 3. The SVM may be equally used to separate non-linearly separable patterns. In these cases, object coordinates are mapped from the input space to a characteristic space using non-linear functions called characteristic function Φ . Since the characteristics space is high dimensional, it is not practical to directly use the function characteristic Φ to find the separation hyperplane. Rather, non-linear mapping induced by the characteristic function is calculated with the aid of special non-linear functions known as the *Kernel*. The *Kernel* has the advantage of operating in the input space where the solution to the classification problem is obtained by the weighted sum of the Kernel function evaluated by the support vectors. The SVM algorithm can construct a variety of learning machines by use of different kernel functions. Three kinds of kernel functions are usually used, they are as follows:

Polynomial kernel of degree d:

$$K(X, Y) = ((X \cdot Y) + 1)^d \quad (3)$$

Radial basis function with Gaussian kernel of width $C > 0$:

$$K(X, Y) = \exp \frac{-|X-Y|^2}{c} \quad (4)$$

Neural networks with tanh activation function:

$$K(X, Y) = \tan h (K(X, Y) + \mu) \quad (5)$$

Where the parameters K and μ are the gain and shift.

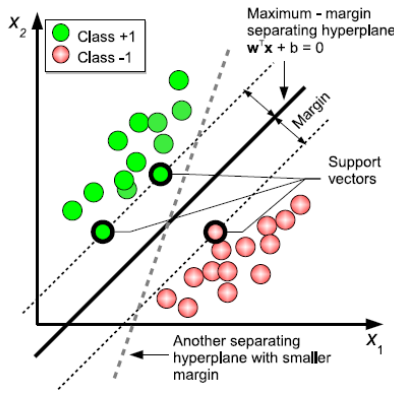


Fig.3. Basic principle of SVM classifier

In the present work SVM method has been used to classify the neurological disordered subject voice from the normal subject voice. SVM is a powerful machine learning tool which attempts to obtain a good separating hyper-plane between two classes in the higher dimensional space. The equation of the hyper-plane is [22]:

$$w^T x + b = 0 \quad (6)$$

Where ' w ' a weight, is vector and ' b ' is the bias. Nonlinearity is satisfied by mapping the input features x into higher dimensions using a function

$$\phi(x) : R^d \rightarrow R^d, \quad p > d \quad (7)$$

and hence the hyperplane becomes:

$$w^T \phi(x) + b = 0 \quad (8)$$

This leads to the following optimization problem:

$$\min_{\xi, w, b} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N \xi_i \quad (9)$$

Subject to:

$$\begin{aligned} y_i (w^T \phi(x_i) + b) &\geq 1 - \xi_i & i = 1 \dots N \\ \xi_i &\geq 0 & i = 1 \dots N \end{aligned} \quad (10)$$

C is constant determined by a cross validation process. The dual formulation of this problem is:

$$\max_{\lambda} \sum_{i=1}^N \lambda_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \lambda_i \lambda_j y_i y_j K(x_i, x_j) \quad (11)$$

Subject to:

$$\sum \lambda_i \lambda_i = 0 \quad 0 \leq \lambda_i \leq 1 \dots N \quad (12)$$

Here $\lambda_i, i = 1 \dots N$ are the Lagrange multipliers. The function $K(x_i, x_j) = \phi(x_i)^T \phi(x_j)$ is the kernel function. The kernel can be any one of the kernel discussed earlier. If the probability density functions of the feature vectors in both classes are known, there is a possibility of defining natural kernels derived from these distributions. The basic principle of an SVM classifier with the hyperplane and the support vectors is shown in Fig.3.

This linear classifier will not efficiently separate the pathological acoustic features from normal ones. Hence non linear SVM with non linear discriminant function is used, which can be written as

$$f(x) = \sum_{i=1}^N a_i t_i K(x, x_i) + d \quad (13)$$

Here $t_i \in \{+1, -1\}$ are the ideal output values. The support vectors x_i , their corresponding weights a_i and the bias term d , are determined from a training set using an optimization process. The kernel function $K(\dots)$ is designed so that it can be expressed as

$$K(x, x_i) = \phi(x)^T \phi(x_i) \quad (14)$$

where $\phi(x)$ is a mapping from the input space to kernel feature space of high dimensionality. The kernel function allows computing inner products of two vectors in the kernel feature space. In the present work a polynomial kernel of order 3 is used.

III. EXPERIMENTATION AND RESULTS

A. Statistical analysis

As discussed earlier in the statistical analysis from the student's t -test it is observed that the p value of all the features except the fundamental frequency components have significant differences in their mean values of neurological subjects and normal subject voices ($p < 0.05$), hence, showing the suitability of these features for assessment of vocal impairment. Further, the graphical results for some characteristics are illustrated for normal frame of vowel /ah/ and neurological frame of vowel /ah/. The scatter plots shown in Fig.4 illustrate the correlation of measurements relative to one another. As the measurements have low correlation they can be easily discriminated. Fig.5 shows the distribution of HNR and Shimmer measurement in box plots to show a comparison between normal and neurological subject voices. The boxes have lines at the lower quartile, median, and upper quartile values. The whiskers are lines extending from each end of the boxes to show the extent of the rest of data and '+' symbols mark the outlying points. If the line in the box plot does not overlap, we can conclude with 95% confidence that the true medians do differ, so medians are statistically different for normal and neurological disorder voices.

B. Classifier

In order to evaluate the performance of the classifier and to make comparisons, several measurements (TP, TN, FN, FP) and ratios (SE, SP, and E) were taken into account [1].

1. True negative (TN): The detector found no event (normal voice) when indeed none was present.
2. True positive (TP): The detector found an event (pathological voice) when one was present.

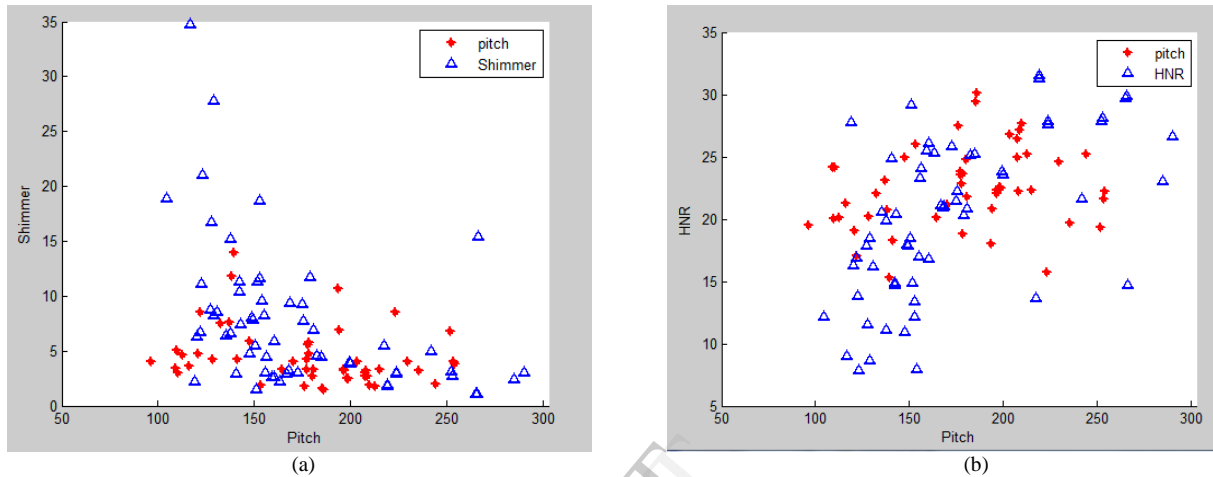


Fig. 4. Scatter plot of (a) pitch versus Shimmer and (b) pitch versus HNR showing low correlation.

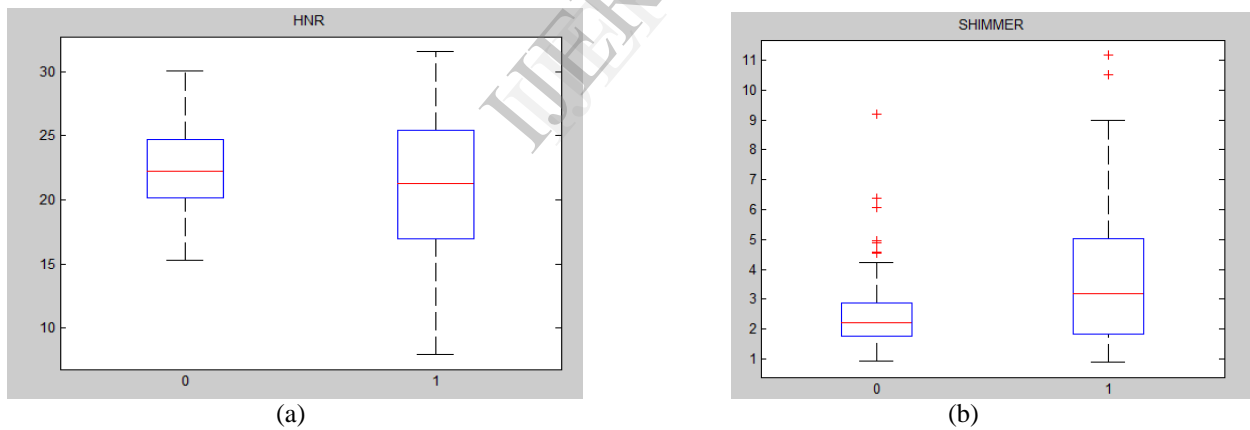


Fig. 5 Distribution of (a) HNR and (b) Shimmer for normal (0) and Neurological disorder voices

3. False negative (FN): The classifier missed an event, also called false rejection
4. False positive (FP): The detector found an event when none was present, also called as false acceptance.
5. Sensitivity (SE): Likelihood that an event will be detected given that it is present

$$SE = 100 \cdot \frac{TP}{TP + FN}$$

6. Specificity (SP): Likelihood that the absence of an event will be detected given that it is absent

$$SP = 100 \cdot \frac{TN}{TN + FP}$$

7. Efficiency (E): Likelihood that the classification is correct

$$E = 100 \cdot \frac{TN + TP}{TN + TP + FN + FP}$$

The confusion matrix for MLP is shown in Table II and for SVM in Table III. The ratios are tabulated in Table IV for comparison.

TABLE II
CONFUSION MATRIX FOR MLP

Actual	Predicted	
	Positive	Negative
	Accuracy (%)	
Positive	TP = 84.5	FN = 15.5
Negative	FP = 21.1	TN = 78.9

TABLE III
CONFUSION MATRIX FOR SVM

Actual	Predicted	
	Positive	Negative
	Accuracy (%)	
Positive	TP = 83.3	FN = 16.7
Negative	FP = 11.1	TN = 88.9

It is observed from the Table II that the rate of identification of neurological disordered voice is more with 84.5%, whereas the misclassification of normal is more in MLP. In the case of SVM the identification of normal subject voice is 88.9% and hence showing a tendency towards misclassification of abnormal voice.

Table IV shows the best overall classification accuracy of SVM of 86.11% compared to the accuracy of MLP of 81.68%.

TABLE IV
CLASSIFICATION EFFICIENCY RESULTS WITH TRAINING AND TESTING SUBSETS FOR TWO DIFFERENT CLASSIFIERS

Classifier	Classifier's Parameter	Subset	Sensitivity	Specificity	Accuracy (%)
MLP	Hidden neurons = 40	Train	72.22	92.59	82.41
		Test	84.50	78.87	81.68
SVM	Polynomial kernel of order 3	Train	95.34	97.18	96.26
		Test	83.33	88.89	86.11

The sensitivity is more in case of MLP with 84.5%, (ability to identify neurological disordered voice) but the specificity is more in SVM with 88.89% (ability to identify normal voice).

IV CONCLUSION

Time domain parameters used for classification of normal voice from neurological disorder voice show a significant differences in their p value in all types of shimmers, jitters, NHR, and HNR which indicate that these features can be used as classifier inputs. A network using SVM for training is suggested. This network needs a shorter time to train compared to MLP network. The classification accuracy obtained from SVM classifier is 86.11% which is significantly better than MLP classifier with an accuracy of 81.68% with the ability of the network to identify normal voice more efficiently. Hence, the time domain parameters with SVM classifier will give better classification of normal

and pathological voices. Most of the approaches for acoustic analysis of pathological voices use time domain features like, jitter, shimmer, pitch period and extract these features require long duration of signal which is quite difficult to get from voice affected patients. The frequency domain analysis will require short duration data and also gives more information. In future work, to improve the classification accuracy, the experimentation could be done with spectral features as inputs for classifiers, and also by combining the different classifiers.

ACKNOWLEDGMENT

The authors are grateful to Dr. Harsha and Dr.Keshav, Neurological Department, J.S.S., Hospital, Mysore, for helping us to collect the voice data of neurological disordered patients.

REFERENCES

- [1] J. I. Godino-Llorente and P. Gomez-Vilda, "Automatic detection of voice impairments by means of short-term cepstral parameters and neural network based detectors" *IEEE Trans. Biomed. Eng.* vol. 51, no.2, pp.380-384, Feb. 2004.
- [2] R.J.Baken, R.F.Orlikoff, "Clinical measurement of speech and voice", 2nd edition, singular Thomson learning, 2000.
- [3] Antoine Giovanni, Ping Yu, Joana Revis, Marie-Dominique Guarella, Bern a rd Teston, Maurice Ouaknine, "Objective dysphonia evaluation using the EVA® workstation. A review", *Fr ORL*, vol. 90, pp. 183-19, 2006.
- [4] M.A.Little, P.E.McSharry, S.J.Roberts, D.Costello, andI.M.Moroz, "Exploiting nonlinear recurrence and fractal scaling properties for voice disorder detection," *Biomedical Engineering [Online]*. vol. 6, no. 23, 2007.
- [5] B. T. Harel, M. S. Cannizaro, H. Cohen, N. Reilly, and P. J. Snyder, "Acoustic characteristic of Parkinsonian speech: A potential biomarker of early disease progression and treatment," *J. Neuro. Linguistics*, vol. 17, pp.439-453, 2004.
- [6] M. A. Little, P. E. McSharry,E.J.Hunter, Jennifer Spielman and Lorraine O.Ramig, "Suitability of Dysphonia Measurements for Telemonitoring of Parkinson's Disease", *IEEE Trans. Biomed. Eng.*, vol. 56, no. 4, pp. 1015-1022, April 2009
- [7] Petra Zwirner, Thomas Murry, Gayle E. Woodson, "Phonatory function of neurologically impaired patients", *Journal of Communication Disorders* vol. 24, no.4, pp.287-300, Aug. 1991.
- [8] D. A. Rahn, M. Chou, J. J. Jiang, and Y. Zhang, "Phonatory Impairment in Parkinson's disease: Evidence From Nonlinear Dynamic Analysis And perturbation Analysis," *J. Voice*, vol. 21, pp. 64-71, 2007.
- [9] A.Tsanas, Mas.A.Little, Patrick E. McSharry, J.Spielman, Lorraine O. Ramig, "Novel Speech Signal Processing Algorithms For High-Accuracy Classification Of Parkinson's Disease", *IEEE Tras. Of Bio. Med. Eng.*, vol57, no.5, pp.1264-1271, 2012.
- [10] Mehmet Fatih CAGLAR, Bayram CETISLI, Inayet Burcu TOPRAK, " Automatic Recognition of Parkinson's Disease from Sustained Phonation Tests Using ANN and Adaptive Neuro-Fuzzy Classifier", *J. Eng. Sci. and Design*, vol.1, no.2, pp.59-64, 2010.
- [11] Resul Das, "A Comparison of Multiple Classification Methods for Diagnosis of Parkinson Disease", *J. Expert system with applications*, vol. 37, pp. 1568-1572, 2010.
- [12] David Gil A, Magnus Johnson B, "Diagnosing Parkinson by using Artificial Neural Networks and Support Vector Machines", *Global Journal of Computer Science and Technology*, pp. 63-71, 2009.
- [13] Michael R. Chial, "Suggestions for Computer-based audio recording of speech samples for perceptual and acoustic analyses", *Phonology Project Technical Report*, no. 13, Oct. 2003.
- [14] Luis M. T. Jesus, Anna Barney, Ricardo Santos, Janine Caetano, Juliana Jorge, Pedro Sa Couto," Universidade de Aveiro's Voice

- Evaluation Protocol”, in *Proc. of Interspeech 2009, Brighton, UK, 7-10 Sept. 2009*, pp. 971-974.
- [15] P. Boersma, and D. Weenink, “Praat: doing phonetics by computer” (Version 5.2.2.1) [Computer program]. Retrieved from <http://www.praat.org/>, 2011.
- [16] Tripti Kapoor, R.K. Sharma, “Parkinson’s disease Diagnosis using Mel-frequency Cepstral Coefficients and Vector Quantization”, *Int. J. of Comp. Application.*, vol.14., no.3, pp.43-46, 2011.
- [17] Little1, Patrick E. McSharry, Lorraine O. Ramig, “Enhanced Classical Dysphonia Measures And Sparse Regression For Telemonitoring Of Parkinson’s Disease Progression”, in *proceedings of 35th International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Dallas, Texas, USA, March 14 - 19, 2010, pp.594-597.
- [18] Uma Rani K and Mallikarjun S. Holi, “Analysis of Speech Characteristics of Neurological Diseases and their Classification”, in *Proceedings of ICCNT 2012, Coimbatore, India, July 2012*, [proceedings in the form of CD].
- [19] Yegnanarayana B. “Artificial neural network”. New Delhi: *Prentice-Hall of India*; 1999.
- [20] Steve.R.Gunn, “Support Vector Machine for Classification and Regression”, Technical Report, School of Electronics and Computer Science University of Southampton, 1998.
- [21] P. Dhanalakshmi, S. Palanivel, V. Ramalingam, “Classification of audio signals using SVM and RBFNN,” *Expert Systems with Applications*, vol.36, no.3, pp. 6069-6075, 2009.
- [22] K.M Ravikumar, R.Rajagopal, H.C.Nagaraj, “An Approach for Objective Assessment of Stuttered Speech Using MFCC Features”, *DSP Journal*, vol. 9, no.1, pp19-24, June 2009.

IJERT