Detection of Parkinson's disease by using Voice Analysis

Ms. Ahilya V. Salunkhe[1] Prof. Ms. Pallavi S. Deshpande[2]
[1,2]Department of Electronics & Telecommunication Engineering
[1,2]Smt. Kashibai Navale College of Engineering, Vadgaon Bk. Pune, India

Abstract— In this paper, we tend to gift associate degree assessment of the sensible worth of existing ancient and nonstandard measures for discriminating healthy individuals from individuals with Parkinson's sickness (PD) by detection speech defect. We tend to introduce a brand new measure of speech defect, pitch amount entropy (PPE), that is powerful to several uncontrollable contradictory effects together with creaky acoustic environments and normal, healthy variations in voice frequency in our project uses feed forward neural network classifier to extend the classification performance with having high sensitivity, specificity and accuracy. Finally, we discover that nonstandard ways are best ready to separate healthy from metal subjects. The chosen nonstandard ways are sturdy to several uncontrollable variations in acoustic setting and individual subjects, and are so likeminded to tele monitoring applications.

Key words: Biomedical Measurements, Nervous System, Speech Analysis, Telemedicine

I. Introduction

NEUROLOGICAL disorders, together with Parkinson's malady (PD), Alzheimer's, and brain disorder, deeply have an effect on the lives of patients and their families. PD affects over one thousand people in North America alone. Moreover, associate degree aging population suggests that this range is predicted to rise as studies counsel chop-chop increasing prevalence rates when the age of sixty . In addition to exaggerated social isolation, the monetary burden of metallic element is important and is calculable to rise within the future currently, there's no cure, though medication is obtainable providing vital alleviation of symptoms, particularly at the first stages of the malady .Most people with Parkinson's (PWP) malady can thus be well obsessed on clinical intervention.

Parkinson could be a disorder and happens owing to lack of dopamine neurons. These dopamine neurons manage all body movements. Parkinson patients have issue in doing all daily routine activities, and even have disturbed vocal fold movements. Victimization voice analysis malady is diagnosed remotely at an early stage with a lot of reliability and economic way.

Diagnosis of Parkinson malady is extremely tough and no diagnostic science laboratory tests are accessible. Medical specialty tests and brain scans are done to diagnose it. These strategies are terribly costly and want high level of experience. Some physical identification can even be done however patients are needed to be ascertained for an extended time and this identification provide results once nearly eightieth of dopamine gets terminated. Concerning seventieth of individuals with Parkinson shows tremor that's most distinguished in hands and fingers. Stiffness within the muscles, slowness of movements and lack of coordination whereas doing daily routine activities also are vital signs of Parkinson. An explicit reason for the death of dopamine isn't renowned. Genetic issue is one amongst the reason for this malady.15% of the patients have their case history. Internal and external toxins scale back the dopamine production. Free radicals also are answerable for the death of dopamine as age will increase the possibilities of occurring this malady additionally will increase. Voice of the person shows changes at an earlier stage, therefore identification of Parkinson exploitation voice analysis may be done at an earlier stage. Reduced in voice level by approx ten sound unit, whispering, breathiness, tremors, shifting to higher tones are some voice characteristics visible in metallic element voice. This methodology (is extremely) reliable and of very radical low price. Methodology is totally computerized and no medical professionals are needed. Because the metallic element patients have problem in clinical visits, during this voice analysis methodology no clinical visits are needed. This methodology may be done telephonically, therefore the telediagnosis of the malady may be done by voice analysis with terribly less prices and efforts. Voice analysis for identification of malady isn't solely restricted to Parkinson however it may be used for several alternative diseases. Voice nodules, Reinke‟ oedema, respiratory disorder can even be diagnosed exploitation this methodology. Numerous classifiers are utilized in such kind of identification. With the assistance of classifiers accuracy and dependability of identification will increase.

II. Objectives

The Objective of this project listed below

- The calculation of features.
- The preprocessing of features.
- The application of a classification technique to all possible subsets of features for the discrimination of healthy from disordered subjects, selecting the subset that produces the best classification performance.
- To archive more accuracy of system.
- To make system more flexible and robust.

III. Literature Survey

The paper by Saloni, R.K. Sharma., and A. K. Gupta works occurring this paper show that data processing have nice potential in illness detection for the advancement of medical field. Data processing is largely a tool for changing the data into some terribly helpful info. Data processing provides ways in which to extract info rework and gift the info during a helpful format. In this paper, they need used the feature dataset of Parkinson illness. Feature choice and classification is employed to classify healthy and pathological information sets. For feature choice a correlation filter is employed. Fuzzy C means that agglomeration and pattern recognition is applied on elect options for classifying traditional speakers and metallic element speakers Support vector machines builds a model victimisation set of coaching examples, every marked to its class so used for classification.

The paper by Max A. Little make a case for the sensible value of existing ancient and nonstandard measures for discriminating healthy folks from folks with Parkinson's disease(PD) by sleuthing speech defect. He introduce a brand new live of speech defect, pitch amount entropy (PPE), that is strong to several uncontrollable contradictory effects as well as strident acoustic environments and traditional, healthy variations in voice frequency. He collected sustained phonations from thirty one folks, 23 with PD. He then designated 10 extremely unrelated measures, associated an complete search of all potential combos of those measures finds four that together cause overall correct classification performance of ninety one.4%, employing a kernel support vector machine. lastly, He notice that nonstandard strategies together with ancient harmonics-to-noise ratios area unit best ready to separate healthy from metallic element subjects. the chosen nonstandard strategies area unit sturdy to several uncontrollable variations in acoustic setting and individual subjects, and area unit so similar temperament to telemonitoring applications.

The paper by Mohammed Shahbakhi, Danial Taheri Far make a case for a brand new algorithmic rule for designation of Parkinson's malady supported voice analysis. Within the beginning, genetic algorithmic rule (GA) is undertaken for choosing optimized features from all extracted options. Afterward a network supported support vector machine (SVM) is employed for classification between healthy and folks with Parkinson. The dataset of this analysis consists of a variety of medicine voice signals from thirty one individuals, twenty three with Parkinson's malady and eight healthy individuals. The topics were asked to pronounce letter "A" for three seconds. Twenty two linear and non-linear options were extracted from the signals that fourteen options were supported F0 (fundamental frequency or pitch), jitter, shimmer and noise to harmonics magnitude relation, that ar main factors in voice signal. As a result of dynamic in these factors is noticeable for the individuals with Pd, optimized options were elite among them. Of the varied numbers of optimized options, the information classification was investigated. Results show that the classification accuracy p.c of 94.50 per four optimized options, the accuracy p.c of 93.66 per seven optimized options and also the accuracy p.c of 94.22 per nine optimized options, may well be achieved. It may be discovered that the most effective classification accuracy is also achieved victimization Fhi (Hz), Fho (Hz), noise (RAP) and shimmer (APQ5).

IV. Experimental Setup

The data for this study contains a hundred sustained vowel phonations from some male and female subjects, of that sixty two were diagnosed with metallic element. The time since diagnoses ranged from zero to twenty-eight years, and also the ages of the topics ranged from forty six to eighty five years. Averages of six phonations were recorded from every subject, starting from one to thirty six s long. Fig.1 shows plots of 2 of those speech signals. The phonations were recorded in associate degree Industrial Acoustics Company (IAC) sound-treated booth employing a head-mounted mike (AKG C420) positioned at eight cm from the lips. The mike was label employing a category one sound-level meter (B&K 2238) placed thirty cm from the speaker. The voice signals were recorded directly on pc victimization computerized Speech Laboratory (CSL) 4300B hardware (Kay Elemetrics), sampled at forty four.1 kHz, with 16-bit resolution.

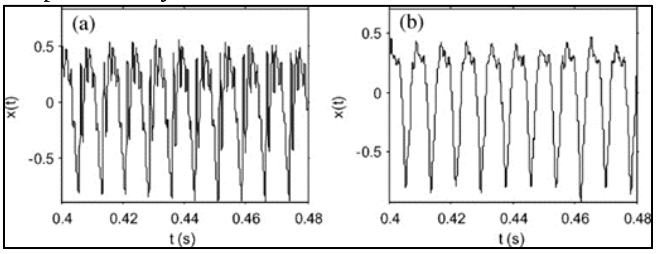

Fig. 1: Two selected examples of speech signals. (a) Healthy. (b) Subject with PD. The horizontal axis is time in seconds and the vertical axis is signal amplitude (no units).

V. System Architecture

A. *Feature Calculation Stage*

The feature calculation stage involves the application of a representative selection of traditional and nonstandard measurement methods to all the speech signal

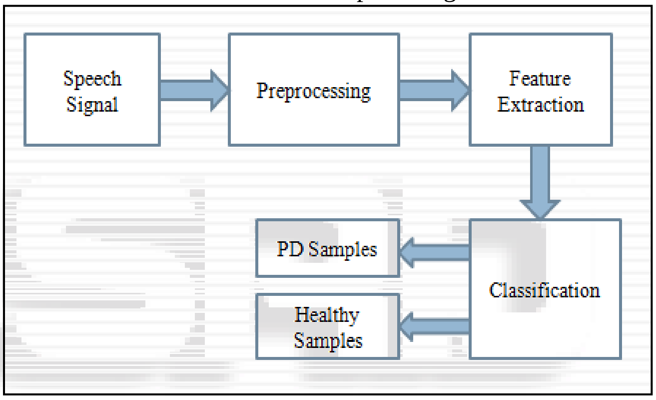

Fig. 2: Block Diagram of Proposed System

1) *HNR*

The noise - to- harmonics (and harmonics-to-noise) ratios are derived from the signal-to-noise estimates from the autocorrelation of each cycle.

2) *RPDE*

The repetition amount density entropy (RPDE) quantifies the extent to that dynamics within the reconstructed space once time-delay embedding is thought of as strictly periodic, i.e., repetition precisely [8]. A continual signal returns to constant purpose within the space once an exact length of your time, known as the repetition amount T. It's been shown that the deviation from cyclicity evaluated by the entropy H of the distribution of those repetition periods P(T) is a smart indicator of general voice disorders, as general voice pathologies cause impairment within the ability to sustain regular vibration of the vocal folds [8]. Dividing through by the entropy of the uniform distribution normalizes the RPDE values (Hnorm) to the vary[0, 1].

3) *DFA*

DFA may be a measure of the extent of the random self-similarity of the noise within the speech signal. The noise in speech is generally generated by turbulent flow of air through the vocal folds. Such turbulent processes are characterised by an applied mathematics scaling exponent α on a spread of physical scales, which manifests in measured aspects of the dynamics as well as acoustic pressure fields. In some voice disorders, incomplete vocal folds closure ends up in changes during this turbulent "breath" noise, thus the characteristics of the self-similarity of the noise within the speech signal is therefore an indicator of defect of speech [8]. It's found that for general voice disorders, the scaling exponent is larger for dysphonic than healthy subjects [8]. The DFA formula calculates the extent of amplitude variation F (L) of the speech signal over a spread of your time scales L, and therefore the self-similarity of the speech signal is quantified by the slope α of a line on a log–log plot of L versus F(L).A simple nonlinear transformation then normalizes these slope values (α norm) to the vary [0, 1] [8].

4) PPE

All healthy voices exhibit natural pitch (F0) variation characterized by sleek sound & microtremor. speakers with naturally high-pitched voices can have a lot of larger sound and microtremor than those with low-pitched voices, once these variations square measure measured on associate degree cardinal number (in hertz) scale. Therefore, measurements of abnormal speech pitch variation got to take under consideration these 2 necessary effects: healthy, sleek sound and microtremor, and therefore the exponent nature of speech eat utterance (and perception). These observations counsel that a lot of relevant scale on that to assess abnormal variations in speech pitch is that the perceptually relevant, exponent (tonal) scale, instead of absolutely the frequency scale. To implement these 2 insights algorithmically, we have a tendency to initial get the pitch sequence of the phonations and convert to the exponent half step scale p(t), wherever p is that the half step pitch at time t. we have a tendency to next analyze the roughness of variations during this sequence over and higher than any healthy, sleek variations, by initial removing linear temporal correlations during this half step sequence with a regular linear change of color filter (coefficients of that square measure calculable exploitation linear prediction by the variance technique to provide the relative half step variation sequence r(t). This filtering imp actively flattens the spectrum of the half step statistic and removes the effect of the mean half step (which depends on the individual preferences and gender). Later on, we have a tendency to construct a distinct likelihood distribution of incidence of relative half step variations P(r). Finally, we have a tendency to calculate the entropy of this likelihood distribution that then characterizes the extent of (non-Gaussian) fluctuations within the sequence of relative half step pitch amount variations. A rise during this entropy live higher reflects the variations over and higher than natural healthy variations in pitch discovered in healthy utterance.

B. MFCC

MFCC is mostly useful tool in speech recognition process. It is very useful for cough analysis. MFCC involve the estimation of short time power spectra. Mapped to the MEL frequency scale and to compute cepstral coefficient.

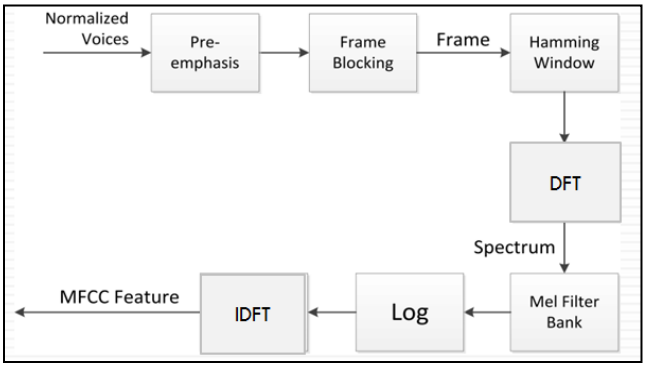

Fig. 3: Block diagram of MFCC

C. Feature Preparation & Classification Stage

Practical exploitation of the information in the measures calculated before requires us to construct feature vectors from these measures, which can then be subsequently used to discriminate healthy from PWP. FFNN classification performance is greatly enhanced by preprocessing of the values of each measure with an appropriate rescaling. Here, we scale each measure such that, overall signals, the measure occupies the numerical range [−1, 1].

In this step we automatically classified the Parkinson's sample signal and non-Parkinson's i.e. healthy sample signal.

In our work project we use the feed forward neural network classifier for the separation of the Parkinson's sample signal and non-Parkinson's sample signal. It is biologically inspired classification algorithm Feed forward neural network having the Capability to know the different types of the dysphonic sound. A feedforward neural network is an artificial neural network wherein connections between the units do not form a cycle. As such, it is different from recurrent neural networks.

The feedforward neural network was the first and simplest type of artificial neural network devised. In this network, the information moves in only one direction, forward, from the input nodes, through the hidden nodes (if any) and to the output nodes. There are no cycles or loops in the network

It having the advantages that it can be classifies the data using the linear decision boundaries.

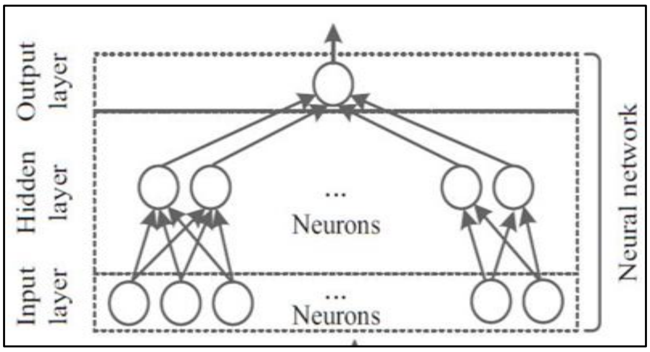

Fig. 4: Feed Forward Neural Network Classifier

VI. Implementation Results

A. Feature Calculation

The novel, nonstandard measures and harmonics-to-noise ratios are more evenly spread over a wider range of values.

Fig. 5 shows the results of calculative the RPDE and DFA values for a few designated speech signals. As are often seen, for healthy subjects, the repetition amount density P (T) shows one peak close to the time T at that the voice signal tends to repeat itself. for several PWP, however, the repetition periods square measure meet a large vary of values, that indicates that the vocal folds aren't periodical at regular intervals. This is often seemingly caused by impairment of the stable positioning of the intrinsic cartilaginous structure muscles (those that directly move the vocal folds), or adscititious cartilaginous structure muscles (connecting the cartilaginous structure and different structures), or by weakness within the production of stable flowing from the lungs.

For many healthy subjects, the energy within the flowing of the lungs is well imparted to the movement of the vocal folds to get clear sustained phonations. Thus, the speech signal are drum sander, and this is often shown within the smaller DFA scaling exponent. However, several PWP square measure unable to take care of stable vocal cord vibration, and far a lot of of the flowing energy are transferred to turbulent acoustic noise generation mechanisms. Hence, the speech signal are rougher, and this could be seen in a rise within the DFA scaling exponent.

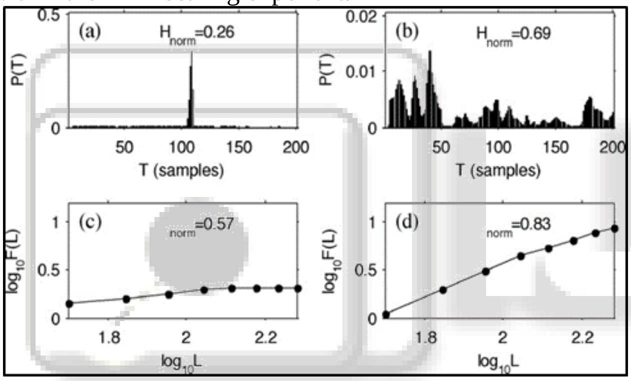

Fig. 5: RPDE and DFA Results for Healthy Subjects (Left Panels) And for Sub-Jects with Parkinson's (Right Panels). (A) And (B) Recurrence Period Density P(T) For Recurrence Times T. (C) And (D) Log–Log Plot Of Scaling Window Sizes L Against Fluctuation Amplitudes F(L)

Regarding the PPE measure (in Fig.6), we will see that healthy interval pitch sequences tend to be quite stable with signs of tiny, regular, swish sound, and micro tremor. Once removing this healthy variation with the lightening filter, the distribution of residuals shows a powerful peak at zero. The entropy of this distribution is correspondingly tiny. For PWP, however, the interval pitch sequence shows sizeable irregular variation; the colorless sequence is extraordinarily rough and also the distribution of residuals is cover a large vary of values. This can be picked up by the big entropy worth.

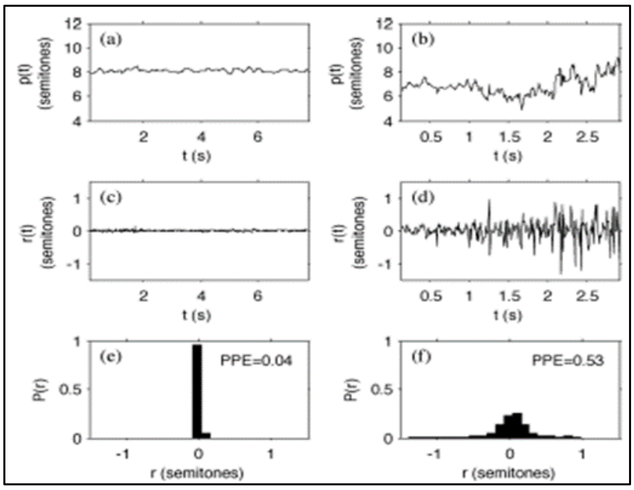

Fig. 6: Details of PPE Calculation. (A) And (B) Pitch Period P(T) In Semitones Relative To Note C3 On The Musical Scale. (C) And (D) Residual Of Pitch Period R(T) After Spectral Whitening Filter. (E) and (F) Probability Densities P(R) Of Residual Pitch Period R. PPE Value Is the Entropy of This Probability Density). Left Panels Are for a Healthy Subject, Right Panel Is for a Person with Parkinson's.

B. Feature Preparation & Classification

After preprocessing by range scaling, for classification purpose we uses the feed forward neural network classifier. In our study we collect the Parkinson's affected and healthy samples. i.e. total 100 samples out of which we use 50 for training phase and 50 for testing phase.

Correct classification of parkinson's infected person sample sound and normal person sample sound can be measured in terms of specificity, sensitivity and accuracy the following confusion matrix gives the output accuracy of proposed system.

	Prediction Outcome →	
Actual value ↓	True Positive	False Negative
	False positive	True Negative

Table 1: Confusion matrix

Where

- True positive (TP) = correct Parkinson disease samples classified
- True negative (TN) = correctly healthy samples classified.
- False negative (FN) = Parkinson samples classified as healthy
- False positive (FP) = healthy sample as Parkinson's sample

	Healthy persons	Parkinson's infected persons
Healthy persons	36	2

Parkinson's infected persons	1	61

Table 2: Final Result with Confusion Matrix

Calculations:

$$Sensitivity = \frac{TP}{TP + FN} * 100 = \frac{36}{38} * 100 = 94.73\%$$

$$Specificity = \frac{TN}{TN + FP} * 100 = \frac{61}{62} * 100 = 98.37\%$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} * 100 = \frac{97}{100} * 100 = \mathbf{97}\%$$

VII. Conclusion

In projected methodology it's simply attainable to classify the Parkinson's affected voice samples and healthy voice samples. Exploitation this methodology we tend to attain the accuracy of ninety seven, sensitivity of ninety four.73% and specificity of ninety eight.37% a vital note is that our results area unit supported broadband, uncompressed audio signals, and that we assume that future web information measure is enough that voice compression won't typically be needed. Future analysis might additional take a look at these findings by applying these measures to voice signals recorded in acoustic environments a lot of typical of sensible tele monitoring applications.

References

[1] Saloni, R.K. Sharma., and A. K. Gupta "Detection of Parkinson Disease Using Clinical Voice Data Mining." INTERNATIONAL JOURNAL OF CIRCUITS, SYSTEMS AND SIGNAL PROCESSING Volume 9, 2015

[2] J. I. Godino-Llorente and P. Gómez-Vilda "Automatic Detection of Voice Impairments by Means of Short-Term Cepstral Parameters and Neural Network Based Detectors", IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING, VOL. 51, NO. 2, FEBRUARY 2004.

[3] Saloni, R. K. Sharma, Anil K. Gupta "Artificial Voice Analysis for Telediagnosis of Parkinson Disease Using Artificial Neural Networks and Support Vector Machines," Department of Electronics and Communication Engineering, National Institute of Technology, Kurukshetra, India.

[4] J.R. Orozco-Arroyave, E.A. Belalcazar-Bolanos, J.D. Arias-Londono, "Characterization Methods for the Detection of Multiple Voice Disorders: Neurological, Functional, and Laryngeal Diseases." IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS,

[5] M. Little, P. McSharry, I. Moroz, and S. Roberts, "Nonlinear, biophysically-informed speech pathology detection," in Proc. ICASSP 2006. New York: IEEE Publishers, 2006.

[6] J. I. Godino-Llorente and P. Gomez-Vilda, "Automatic detection of voice impairments by means of short-term cepstral parameters and neural net-work based detectors," IEEE Trans. Biomed. Eng., vol. 51, no. 2, pp. 380– 384, Feb. 2004.

[7] S. Hadjitodorov, B. Boyanov, and B. Teston, "Laryngeal pathology detec-tion by means of class-specific neural maps," IEEE Trans. Inf. Technol. Biomed., vol. 4, no. 1, pp. 68–73, Mar. 2000.

[8] M. A. Little, P. E. McSharry, S. J. Roberts, D. A. Costello, and I. M. Moroz, "Exploiting nonlinear recurrence and fractal scaling properties for voice disorder detection," Biomed. Eng. Online., vol. 6, p. 23, 2007.

[9] Boyanov and S. Hadjitodorov, "Acoustic analysis of pathological voices," IEEE Eng. Med. Biol. Mag., vol. 16, no. 4, pp. 74–82, Jul./Aug. 1997.

[10] J. H. L. Hansen, L. Gavidia-Ceballos, and J. F. Kaiser, "A nonlinear operator-based speech feature analysis method with application to vocal fold pathology assessment," IEEE Trans. Biomed. Eng., vol. 45, no. 3, pp. 300–313, Mar. 1998.

[11] L. Cnockaert, J. Schoentgen, P. Auzou, C. Ozsancak, L. Defebvre, and Grenez, "Low-frequency vocal modulations in vowels produced by Parkinsonian subjects," Speech. Commun., vol. 50, pp. 288–300, 2008.