# Voice Disorders Recognition System using Artificial Neural Network

[1]Syed Akbar Ali, [2]Prof. Sanjay Ganar

[1]Assistant Professor,[2]Assistant Professor
[1] Department of Electronics and Telecommunication Engineering,
[1]Anjuman College of Engineering & Technology , Nagpur, India

***Abstract :*** Speech disorders are human disabilities widely present in young population but also adults may suffer from such disorders after some physical problems. In this context, the detection and further the correction of such disabilities may be handled by Automatic Speech Recognition (ASR) technology. The first works on the speech disorders detection began early in the 70s and seem to follow the same evolution as those on the ASR. Indeed, these early works were more based on the signal processing techniques. Progressively, systems dealing with speech disorders incorporate more ideas from ASR technology .Particularly, Artificial Neural Network, the state-of-the-art approaches in ASR systems, are used. systems that use ASR techniques to evaluate pronunciation of people who suffer from speech or voice impairments like balbuties and dysthria. investigate the existing systems and present the main innovation and some of the available resources. Than the output of this is given to feature extraction in which MFCC, LFCC,ZCR, Energy and LPC is present all this features give informative characteristic of patients and also in arithmetic values. Than the output of this is given to classifier which is a neural network, this classifier does all similar diseases patients

*IndexTerms* - **Pathological voices,MFCC,LFCC,,Voice Disorder, Artificial Neural Network.**

## I. INTRODUCTION

Speech disorders are human disabilities widely present in young population but also adults maysuffer from such disorders after some physical problems. So, the detection and further the correction of such disabilities may be handled by ASR technology. At the beginning, works were essentially based on signal processing techniques. In particular, these works were mainly based on the fundamental computation and the harmonics of the signal, and then the principal of the detection consists on looking for dissimilarities between the normal speech and the abnormal one. Recently, the need of the computer-aided speech therapy systems has increased. Such systems are getting more attention for researcherssince the number of persons suffering from speech impairment is great. The main purpose of these systems is to provide methods for improving the communication skills of person who suffer from disorder in speech or voice.

Among the available systems devoted to speech disorder, we would like first to present Vocaliza, a system which is developed in the context of the National Project TIN in Spain. Vocaliza (Vaquero et al., 2008a), is a SpeechTechnology-based application for computer-aided speech therapy in Spanish language. This software provides a user interface especially designed to be attractive even to the youngest users. It works on three level of language: phonological, semantic and syntactic. Each level was trained by a different method which was shown as a game, in order to attract young users. In fact, all games were based on ASR techniques. The goal here was to decide if the user has completed the game successfully.

The system includes speech synthesis to show how a word must be pronounced, speaker adaptation to estimate the acoustic models adapted to the user and utterance verification to evaluate user pronunciation. Most of Vocaliza functionalities are provided by different Human Languages Technologies (HLTs) like ASR which is the core module of Vocaliza application. Each game needs an ASR decoder to decode the user utterances, and to decide which word sequence has been pronounced.

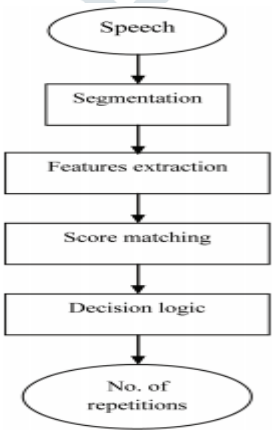


Figure 1. Flowchart of scheme

Hence ANN has been found to be more helpful than a traditional medical expert system in the diagnosis of diseases. In the system various needs to be designed.

There are various disease in voice of patients is affected. We will be choosing pathological & normal voice signal according to need of society from German voice disorder database. This German voice disorder database developed by Putzer contain healthy and pathological voice, where each one pronounced vowels [i, a, u] for 12 s in wav format.[12]. These files are sampled at 50 KHz..

In pre-processing chosen voice signals are passed through pre-emphasis filter which will remove noice and flatten voice signal spectrum thus helps in getting informative features of voice signal. Voice signals are non stationery so they are processed by making small frames of 20ms to 100 ms. Such frames will give more correct information. These framed voice signal is passed through window as window smoothens the signal.

Features like MFCC,LFCC,Energy,ZCR,etc. may be extracted. Mel Frequency Cepstral Coefficients (MFCC)&Linear Frequency Cepstral Coefficients are most commonly used feature extraction method in speech recognition system. Pitch is nothing but frequency of voice signal & Formants are resonance in the voice spectrum. Selection of appropriate feature is important task which is also called principal component analysis.

## II. LITERATURE REVIEW

D. Razzouk, , et al., 2006] have suggested a Clinical decision support systems, a useful tools for assisting physicians to diagnose complex illnesses. Schizophrenia is a complex, heterogeneous and incapacitating mental disorder that should be detected as early as possible to avoid a most serious outcome. These artificial intelligence systems might be useful in the early detection of schizophrenia disorder. The results showed a relatively low rate of misclassification (18-34%) and a good performance by the diagnosis of schizophrenia, with an accuracy of 66-82%.

According to [Vahid Majidnezhad, Igor Kheidorov,2013] A new type of feature vector, based on wavelet packet decomposition and Mel-Frequency-Cepstral-Coefficients (MFCCs), is proposed. Also Principal Component Analysis (PCA) is used for feature reduction. An Artificial Neural Network is used as a classifier for evaluating the performance which yielded accuracy of 91.54% [J.Nayak, et al., 2005] classified the normal, abnormal and hyper function using continuous wavelet transform. The author has used three layered feed forward network with sigmoid activation function and trained the neural network for three classes giving efficiency in the range of 80-85% accuracy.

[Lofti Salhi, et al., 2010] have proposed a method to classify normal and pathological voice using multilayer neural network and using energy coefficients of wavelet transform with accuracy of classification 80 to 100 %.

According to [M.Hariharan, et al., 2010] time domain features based on energy variations are extracted from speech to form feature vector. They have used probabilistic neural network(PNN) for the classification of two class normal Pathological with classification accuracy of 96.42 % and 99.4 % respectively. [Jung-Won Lee, et al., 2013] proposes an efficient feature extraction method for automatic diagnosis systems to detect pathological subjects using continuous speech. Experimental results show that the projected method improves the classification error rate by 11.2% (relative) compared to the conventional method using HNR. [Karthikzeyan Umapathy, et al., 2005] have proposed joint time-frequency approach for classifying pathological voices using continuous speech signals that removes the need of segmentation. In his research speech signals were decomposed using an adaptive time-frequency transform algorithm. The classification was done by linear discriminant analysis .and have classified normal and pathological speech with the classification accuracy of 93.4%.

Research work of [Nicolas saenz-Lechon, et al., 2006] described some methodological concerns to be considered when designing systems for automatic detection of voice pathology, in order to enable comparisons to be made with previous future experiments. In their work tthey have classified two classes pathological and normal signal using 18 MFCC coefficients with percentage classification as 92.73 and 86.34 respectively. The system has accuracy of 89.6%. [Vanitha S., Bharthi.D, 2016] have classified parkinson disease form normal .The classifier used was support vector machine with linear kernel. According to them the accuracy of the classification depends on the voice data samples, voice features and their number. Six voice features are used and 99.6% accuracy is achieved.

According to [Abdullah Caliskan, et al, 2017] classified parkinson disease with accuracy of 93.79% . For this purpose, a deep neural network classifier, which contains a stacked autoencoder and a softmax classifier, is proposed. The several simulations are performed over two databases to demonstrate the effectiveness of the deep neural network classifier. The results of the proposed classifier are compared with the results of the state-of-art classification method. The experimental results and statistical analyses are showed that the deep neural network classifier is very efficient classifier for Parkinson disease diagnosis.

According to research work of [Alexandra Konig, et al., 2015] The classification accuracy of automatic audio analyses were as follows: between Healthy elderly control HCs and those with Mild cognitive impairment(MCI), 79% between HCs and those with Alzheimer disease(AD), 87% and between those with MCI and those with AD, 80% demonstrating its assessment utility..

## III. DETAILS OF DISEASES

Dysarthria : Dysarthria is a speech disorder caused by muscle weakness. Dysarthria is a motor speech disorder resulting from neurologic impairment affecting mainly the control and execution of movements related to speech production. Occurrence of dysarthria in adult age is commonly manifested as a consequence of degenerative disorder such as Parkinson's disease (PD), Huntington's disease (HD), multiple system atrophy (MSA), progressive supranuclear palsy (PSP) or cerebellar ataxia (CA). Interestingly, identification of specific deviant speech characteristics can provide important clues about the underlying pathophysiology and localization of neurological diseases. Speech may also serve as a valuable marker of disease onset or treatment efficacy. Therefore, the main aims of this doctoral thesis were

(a) to design the feasible algorithms, methodologies or measurements that would be sensitive and accurate enough to capture pathological changes in speech

(b) to objectively quantify the effect of neurological disorder on speech production and

(c)to relate the potentially observed speech changes to overall motor performance or medication doses in order to provide deeper insight into the patho physiology of speech disturbances.

Stuttering (Balbuties) :Stuttering is a speech disorder characterized by repetition of sounds, syllables, or words; prolongation of sounds; and interruptions in speech known as blocks. An individual who stutters exactly knows what. he or she would like to say but has trouble producing a normal flow of speech. These speech disruptions may be accompanied by struggle behaviors, such as rapid eye blinks or tremors of the lips. Stuttering can make it difficult to communicate with other people, which often affects a person’s quality of life and interpersonal relationships. Stuttering can also negatively influence job performance and opportunities, and treatment can come at a high financial cost.

Symptoms of stuttering can vary significantly throughout a person's day. In general, speaking before a group or talking on the telephone may make a person's stuttering more severe, while singing, reading, or speaking in unison may temporarily reduce stuttering.

## IV. METHODOLOGY

Symptoms o Generally a Disease Detection includes the following components.
Speech waveform capture (analog to digital conversion) The a-to-d conversion is generally accomplished by digital signal processing hardware on the computer's sound card (a standard feature on most computers today). The typical sampling rate, 8000 samples per second, is adequate. The spoken voice is considered to be 300 to 3000 Hertz. A sampling rate 8000 gives a Nyquist frequency of 4000 Hertz, which should be adequate for a 3000 Hz voice signal.
Pre-emphasis filtering
Because speech has an overall spectral tilt of 5 to 12 dB per octave, a pre emphasis filter of the form 1 – 0.99 z-1 is normally used.This first order filter will compensate for the fact that the lower formants contain more energy than the higher.

4.1Feature Extraction
MFCC: For short time process, preprocess is again full applied for speech signals. Another MATLAB voice toolbox function "melbankm" is used for mel scale filter bank wrapping in For a chosen 256 points FFT, 32 filters are used in filterbank. Then Fast Fourier transformation is carried out use "fft" command and "log" for the logarithm. After the discrete cosine transform carries out, wipe out the first and the last frame.

Classification with Neural Network Enter "nnstart" to start the MATLAB GUI of neural network. It has several functions and choose pattern recognition tool then the general working to load input data and target data. Input data here which indicates the

e.

V.RESULT

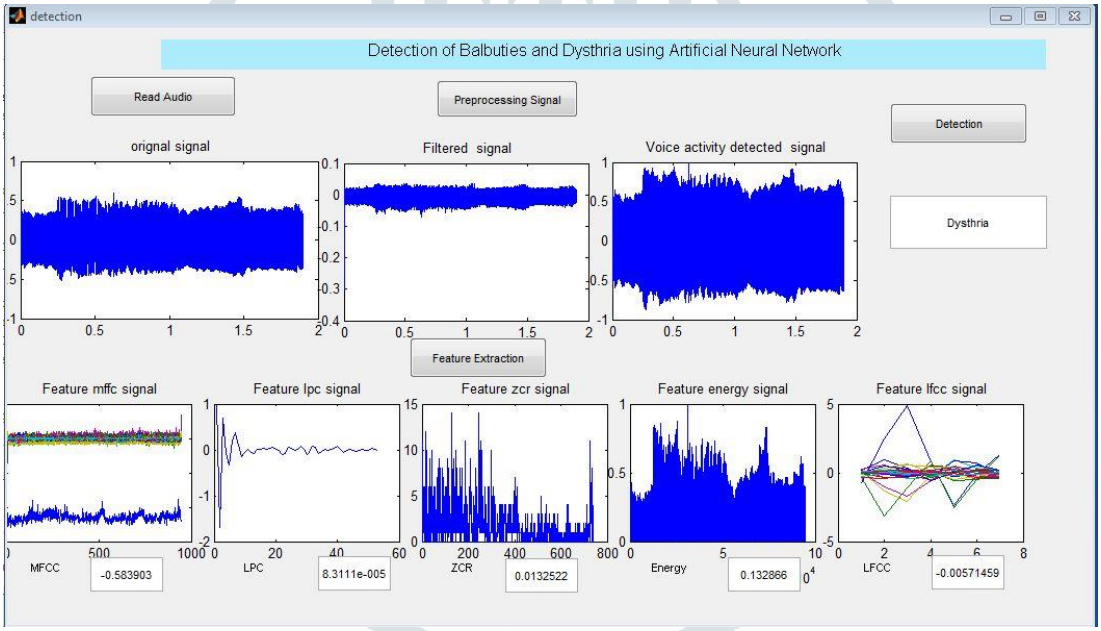


Figure 2. Final Result of Proposed Work

## V. CONCLUSION

The goal of this work is to conceive a tool of help to the clinicians in the Tunisian hospitals. This tool allow to follow up of patients who suffer from illness of vocal and neurological origin. We presented in this paper a material and software interface of numeric treatment of the patient's vocal signal based on neural networks. Result of the MNN classifier gives the correct classification. The classification rate is between 80% and 100%. We have demonstrated in this study, LPC, MFCC, Energy,LFCC,ZCR coefficients is useful for classification of normal and pathological speech data. At a preliminary level, the speech data is classified into two classes balbuties or dysthria . The MNN with BPA used as a classifier has been proved to be more efficient and more precise than the time-frequency analysis method. The MNN classifier represents a low cost, accurate, and automatic tool for pathological voice classification using feature . It is presented in this paper as diagnostic tools to aid the physician and clinician in the analysis of speech disease. Therefore, future work will be focused on the specific recognition of illness type that causes the speech pathology. This work has to be validated on a larger speech pathology database to increase the result reliability

**REFERENCES**

[1]Razzouk D, Mari J.J, Shirakawa I, Wainer J, and Sigulem D.,2006, "Decision support system for the diagnosis of schizophrenia disorders", Brazilian journal of medical and biological research, Volume 39,No.1, (Jan 2006),119-128.
[2]Majidnezhad V, Kheidorov I.,2013, "An ANN-based Method for Detecting Vocal Fold Pathology",International Journal of Computer Applications, Volume 62, No.7 ( 2013)1-4..
[3]Nayak J., Bhatt P.S, Acharya R., Aithal U. V., "Classification and analysis of speech abnormalities", www.sciencedirect, (2005), 319-327.
[4]Salhi, L., Mourad, T., Cherif, A., "Voice Disorders Identification Using Multilayer Neural Network", The International Arab Journal of Information Technology, Vol. 7, No.2, (April 2010), 177-185.
[5]Hariharan, M., Paulraj, M.P., Jaacob, S., "Time Domain Features and Probabilistic Neural Network For the Detection Of Vocal Fold Pathology", Malaysian journal Of Computer Science, Vol No.23, (2010),60-67.
[6]Lee J.W., Kang H.G., Lee Y., "Detecting pathological speech using local and global characteristics of harmonic-to-noise ratio", Proceedings of the IEEE, Signal and Information Processing Association Annual Summit and Conference (APSIPA) Asia-Pacific, (2013) 6010-6014.
[7]Umapat K., Krishnan S., Parsa V., Jamieson D. G., "Discrimination of Pathological Voices Using a Time-Frequency Approach", IEEE Transaction on Biomedical Engineering, Vol.52, No.3, (2005) 421-431.
[8]Saenz-.Lechon N., Godino-Llorente J. I., Osma-Ruiz V., P Gomez-Vilda., "Methodological issues in the development of automatic systems for voice pathology detection", Biomedical Signal Processing and Control, Vol.52, No.3, (2005) 120-128.
[9]Vanitha S., Bharthi D, "Computer aided diagnosis of Parkinson disease using speech signal", International Journal of Advances in Electronics and Computer Science, Vol 3, No 6, (2016)140-143.
[10]König A, Satt A, Sorin A, Hoory R, Toledo-Ronen O, Derreumaux A, Manera V, Verhey F, Aalten P, Robert
[11]PH, David R, " Diagnosis Assessment Disease monitoring", Automatic speech analysis for the assessment of patients with predementia and Alzheimer's disease, Vol. 17 No,1,(2015), 112-124.
[12]Manfred Putzer & Jacques Koreman ''A german database for a pattern for vocal fold vibration '' Phonus 3,
[13]Institute of Phonetics, University of the Saarland, 1997, 143-153