Voice Based Disease Identification System

Manasi Bendale[1], Jaya Soni[2], Garvit Chhugani[3], Nikhil Sahu[4], Nitin Choubey[5], Santosh Bothe[6]

[1,2,3,4,5,6]SVKM Narsee Monjee Institute of Management Studies, Shirpur, India
manasib311@gmail.com[1], jayasoni6287@gmail.com[2], garvitchhugani17@nmims.edu.in[3],
nikhil.sahu41@nmims.edu.in[4], nitin.choubey@nmims.edu[5], santosh.Bothe@nmims.edu[6]

Abstract: Human voice as well as the sound of the body is used as a clinical method to assess the health condition of an individual. The evaluation of the human voice has risen as a critical field of exploration. Speech analysis fundamentally involves the extraction of certain features from voice signals for generation of voice in alluring pertinence by using reasonable techniques. This paper brings up normal ailments that sway understanding voice patterns in proof for driving research that have affirmed voice modifications as demonstrative manifestations in their respective ailments and also the technique by which voice analysis can be done.

Keywords: Articulation points, voice analysis, pathological speech conditions, CNN, feature extraction

1. Introduction

Voice analytics uses a voice recognition tool to analyze and record an audio. Voice analytics program not only converts speech to text, but can also recognize the sentiments and intent of the speaker by interpreting audio signals. There is an abundance of research showing that a person's speech can be affected by multiple physical and mental health conditions. Also, during the speaking, there are 6300+ parameters which become active from which some set of the parameters are affected in each health condition. They could also make your voice creak or jitter so briefly that it's not detectable to the human ear. For example, speaking in a more nasal accent, elongating noises, slurring words or even noises that are not audible to the human ear. The below are the advantages of voice analytics, which will help us in detection and prevention of diseases:

i. Point of Care / Rapid Screening:It helps in achieving accurate real time results within minutes (rapidly)rather than hours and ensures that the patients receive the most effective and efficient treatment when and where it is needed with ease to use.
ii. Early Warning: It also helps in providing early warning to the users if there is a case of future emergency so that everyone should be safe and protected from the diseases.
iii. Disease Surveillance: It is an information which performs the gathering, evaluating, and interpretation of large data from a variety of sources.
iv. Preventive Care and Wellness: It helps in detecting or preventing serious diseases before they become crucial. This can lead to save us from future problems.

From our Voice production system to Throat, there are a total of 18 articulation points. Each point has distinct features. Whenever a healthy person communicates, the individual's speech characteristics are relatively usual. But when a person suffers from a particular health condition, certain parameters of the speaking voice of the spoken phrases impair them. The careful analytics of the speaking voice has a potential to map the underlying disease conditions. Voice analytics is used by the, speech pathologists to identify voice conditions by auditory perceptual criteria such as breathability, gruffness and harshness.

These diagnoses, on the other hand, depend on the knowledge of clinicians and require subjective recognition. It's easy to mix up pathologies with distinct symptoms with those that are often referred to as hoarse. Also when new technologies are used, for example, the same problem occurs when such speakers display a reflex motion in their supra-glottal cavity, resulting in incorrect judgments.

The problem remains in the absence of a doctor-to-patient recommendation system that alerts the doctor to the patient's condition and provides excellent care to the patient, since they have access to a doctor at all times. The problematic statement is building a model for disease prediction using voice recognition. Since the lack of a normalised text for sample collection is a significant issue, our project aims to make it language independent.
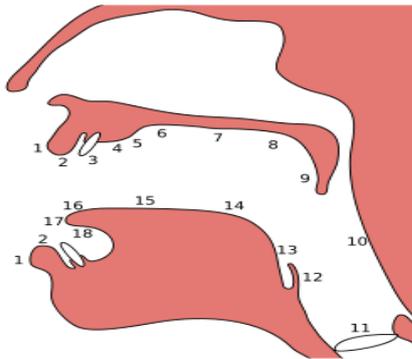
Figure1. Articulation Points

2. Background

The notion of Voice Based Disease Identification System came into existence when the person had to move from one place to another place for the detection of different health conditions and when they had to wait for a long time for the checkup. This led to the creation of major problems. In cases of emergencies, when a person is suffering from a heart attack, lung cancer, depression or kidney failure providing immediate care is difficult at times. But now as times have changed, the improvement of technology has made us able to make a Disease Identification System. This system will accept the voice sample of the patient. It functions by analyzing voice signals and extracting the features of sound. Henceforth it will provide the output of the disease detected of the patient which will help in early prediction of the disease.

3. Literature survey

This paper aims to identify the various health condition by analyzing the voice. Firstly the various parameters of voice were studied as proposed by Dixit et.al[**Error! Reference source not found.**], they adopted the methodology of linear predictive coding for analyzing the parameters. Some recommender systems were also studied for review as proposed by V.Shobana et. Al.[**Error! Reference source not found.**], they had made a personalized recommendation system for the prediction of health conditions using data analytics also another recommender system as presented by Anna Kasperczuk et. al.[**Error! Reference source not found.**] they made a recommender system for colon diseases and achieved a breakthrough accuracy of over 88%. The paper published by Olga Kaminska et. al.[**Error! Reference source not found.**], used fuzzy clustering and self organizing maps clustering techniques to select acoustic features of patients who are suffering from mania, depression, euthymia and mixed, they used recursive feature elimination and the two clustering techniques to achieve relevant results. They made an application through which they recorded calls of the patients which worked as their dataset, but they performed the study using about 15 patients only. They achieved the highest degree of agreement using Fuzzy C-means using a RFE set of 5 parameters reduced from a total of 86 parameters. As many as 39 new pathological voice parameters were introduced in this Jiri Mekyska et. al. [**Error! Reference source not found.**] paper, the importance of these features were tested in English, Czech and Spanish languages. They implemented SVM and Random Forest to detect pathological voices on 2 databases, namely MEEI database and PdA database. They achieved exceptional accuracy of 100% for MEEI database and accuracy of 82.1% was achieved for PdA database. This paper is exceedingly significant for researchers research as the new features described by this paper can be used to classify the pathological condition of the patient using voice. Paper by Kebin Wu et. al. [**Error! Reference source not found.**], uses a novel JOLLRRRR model which uses a fusion of 2 audio clips which has been proved by them as better than using a single audio clip. The only limitation faced by their work is parameter optimisation. In Daria Hemmerling [**Error! Reference source not found.**], a database of 1410 patients was used to classify the voice as pathological and healthy based on 28 parameters. They used random forest to achieve an astounding accuracy of 100%. This work is notable as researchers' will be training the model based on a random forest approach to classify the voice samples. The only demerit of this paper is instead of studying continuous speech, different intensities of /a/, /i/ and /u/ is included in the database. In paper by I. M. M. El Emarya et. al. [**Error! Reference source not found.**], Saarbrucken database was taken as the basis of research on developing a voice pathology detection system. They used MFCC, jitter, shimmer and GMM as classifiers. Feature selection can be improved in this work also the classes should have been more defined.

Then various papers were analyzed which detected different diseases based on voice analysis. The papers which analyzed cough detection Renard Xaviero Adhi Pramono et. al. [**Error! Reference source not found.**] used methodology as Three spectral features – widespread spectrum, low tone prominence, low harmonicity.

Windowing and pre-processing, STFT, frequency bands of B-HF, and B-01 were extracted, recording separation into training and test set using LOOCV scheme and henceforth achieved Sensitivity of 90.31%, specificity of 98.14%, and F1-score of 88.70%. Another paper which was based on cough detection was by Mateusz Solinski et. al. [**Error! Reference source not found.**], they achieved a phenomenal accuracy of 91% and they used classifiers as artificial neural networks to achieve this accuracy although they had some limitations as some curves were misclassified. In a paper by Hwan IngHee[**Error! Reference source not found.**], they did a proof of concept study for asthmatic and voluntary cough sounds and achieved an accuracy of 84% using MFCC, Gaussian Mixture Model–Universal Background Model (GMM–UBM). Research on Cough detection using audio samples using moment theory was done by Jesús Monge-Álvareza et. al. [**Error! Reference source not found.**], they have detected cough in noisy audio signals, which distinguishes this work from others, also this theory has proved to be more effective as voice analysis is based on the energy content in different frequency bands which is relatively easy to estimate. They have achieved sensitivity and specificity of around 90%. This method proved to be more precise than MFCC, LPCC, PNCC and SSCH. In paper presented by Yusuf A. Amrulloh et. al. [**Error! Reference source not found.**], cough segments were identified using pediatric sounds of patients by implementing artificial neural networks. This paper is noteworthy as sensitivity of 93% and specificity of 98% was achieved using ANN. However, they considered data of only 14 patients and the recording was a total of 35 hours.

The papers which analyzed diseases such as heart diseases and diabetes using voice - by Vishakha Pareek et. al. [**Error! Reference source not found.**], they achieved a moderate accuracy of 65%. For processing the voice signal they used CSL model 4500, it also contains MDVP, which analyzes and displays up to 22 voice parameters from a single voice analysis. In a paper by Divya Chitkara et. al. [**Error! Reference source not found.**], they used Time domain analysis for the extraction of parameters, and the acoustic analysis was done using MDVP for the detection of type 2 diabetes.

Some papers dealt with depression and Parkinson's diseases. In a paper by Yasin Ozkanca a MiraçGöksuÖztürk et. al. [**Error! Reference source not found.**], Depression Screening from Voice Samples of Patients Affected by Parkinson's Disease has been done and an accuracy of 77% was achieved using random forest algorithm, the only limitation of this paper was that it was conducted on a very small scale. In another paper by James R.Williamson et. al. [**Error! Reference source not found.**] depression severity was checked with the help of audio and video data of patients, they analyzed the HAM-D scores on basis of data and they used two datasets for conducting their study.On the WIBD dataset, the best prediction result is r = 0.63 and RMSE = 5.49(given true scores with range1-24). On the Mundt set the best result is r = 0.48 and RMSE = 5.99(given true scores with range 3-27). (Where RMSE - root mean squared error r - Spearman correlation). They used techniques like Dimensionality reduction, principal components analysis (PCA), Staircase regression, Gaussian Model, Cross-validation methodology for their paper.

Many papers were published for the detection of voice disorders. In a paper by Gaetan Chambres et. al.[**Error! Reference source not found.**] accuracy of 85% was achieved on their study of detection of patients with respiratory diseases using lung sound analysis. They used a monoclass and multiclass approach for achieving this accuracy. Another paper published by Everthon Silva Fonsecaa et. al. [**Error! Reference source not found.**] by using the concepts of signal energy (SE), zero-crossing rates (ZCRs) and signal entropy (SH), which provide a joint time-frequency-information map achieved a groundbreaking accuracy of 95%. The proposed approach classifies voice signals based on the discriminative paraconsistent machine (DPM), allowing for the application of paraconsistency to treat indefinitions and contradictions. In a paper by RimahAmami et. al. [**Error! Reference source not found.**], the methodology was to use Incremental DBSCAN-SVM, Support Vector Machines (SVM) classifier with a Radial Basis Function (RBF) Kernel for voice pathology detection and they achieved a great accuracy of 98% in their work. In a paper by Ghulam Muhammada et. al. [**Error! Reference source not found.**], Vocal tract irregularity features and support vector machine classifier was used on two datasets and 99.22% accuracy on the MEEI database and 94.7% accuracy on the SVD was achieved.

In a paper presented by Enam Biswas et. al. [**Error! Reference source not found.**], the Symptom-Based Disease Detection System was developed for the Bengali language which is also a limitation because they should've included other languages also, they secured accuracy of 63% using CNN. Another paper presented by TurkerTuncer et. al. [**Error! Reference source not found.**], They introduced an Algorithm MCMTTP for feature extraction, Nca based feature selection, they used 6 different classifiers and analyzed for 8 diff cases & got the highest accuracy using naive Bayes. (nb, KNN, CNN) for disease detection using voice. In the paper of Ms.Deepali S. Shrivastava et. al.[**Error! Reference source not found.**] linear predictive coding, Perception linear prediction, MFCC was used for speech processing. In a paper presented by Suman Deb et. al. [**Error! Reference source not found.**], Analysis and Classification of Cold Speech were done using MFCC, VMD,

SVM. They achieved an accuracy of 90%. In Nicholas Cummins et. al. [**Error! Reference source not found.**], SVM was used for Automatic Recognition of Speech Affected by Upper Respiratory Tract Infection, and accuracy of 70% was accomplished. Paper published by Ghulam Muhammada et. al.[**Error! Reference source not found.**], voice pathologies were detected on the MEEI database using interlaced derivative patterns on glottal source excitation, they achieved a phenomenal accuracy of 90.30%, IDP gave extremely useful results in comparison to other conventional methods like MFCC and MDVP but using MEEI database itself has some limitations like the data has been collected in different environments, so there might be a slight possibility that noise recorded consists on the background commotion. A paper by Hammami,L et. al. [**Error! Reference source not found.**] elaborates on using the wavelet domain to distinguish between normal and pathological voices, initially voice signals were separated by means of EMD-DWT, the two-stage analysis procedure, then vectors of features were extracted. Accuracy of 94.82% was achieved by their work however the reliability of interpolation near the end point is questionable while using EMD, and thus a selection process is necessary before decomposition. Classification of dysphonic voices was described precisely in the paper of João Paulo Teixeira et. al. [**Error! Reference source not found.**], artificial neural networks were used here to achieve an accuracy of 95%. This paper is significant for researchers paper as here jitter, shimmer and HNR parameters were extracted totalling to 81 parameters. The modified method of voice contour and SVM were used in Zulfiqar Ali et. al. [**Error! Reference source not found.**], to detect various voice pathologies, voice intensity was the main parameter used in this paper as dysphonic voices will have a small area covered in the MVC. Here, accuracy of 98.5% was achieved but they used a small set of samples and thus no proper deduction was made. In paper presented by Ghulam Muhammada et. al.[**Error! Reference source not found.**], low level MPEG-7 features were used to classify pathological voices and binary classification of pathologies, a remarkable 99.994% accuracy was achieved by them. SVM was used for classification and MEEI database was used which does not make use of continuous speech.

A remarkable paper published by Jordi Laguarta et. al.[**Error! Reference source not found.**], patients suffering from covid-19 were detected using the dataset of forced coughs by subjects. The paper was successful with 97.1% accuracy and most significantly 100% asymptomatic detection rate. The authors used the voice samples of 5320 patients, furthermore they used CNN and MFCC for their research. The dataset used was MIT Open Voice dataset for COVID-19 cough discrimination.

The researchers' will incorporate all the methodologies presented in the above-mentioned papers which have acquired accuracy of 85% or more and will try to improve the accuracy of researchers' models accordingly and also we will overcome the limitations posed by these papers so that researchers' model would be fairly accurate and has very few limitations.

The researchers' will be working on various datasets so that accuracy of researchers' models would be fairly high. Many voice features will be extracted from the datasets. Also, many models will be tested for the diseases and the one which gives the highest accuracy will be selected for the application. The application that we will deploy will make the users read a normalized text which will be language-independent, it will be based on a facility of read-and-speak and listen-and-speak so that even the illiterate people can make use of researchers' applications. The app will also be verified by a team of expert doctors who will validate researchers' models.

4. Clinical CONDITION CHARACTERIZED by the variations of the unnatural voice

Voice pathologists have identified sound patterns with signature abnormalities in patients of several medical problems such as Diabetes, Hypertension, Parkinson's, Autism Spectrum Disorders (PD), Alzheimer's disease, Dementia, cancer of the larynx, etc. In order to measure these voice changes corresponding to different medical conditions, scientists and medical experts are investigating the correlation for the design of voice recognition devices for the diagnosis and care in a controlled clinical setting of related health disorders. Such techniques of voice analysis allow physicians to track patients and scaling up the progress with continuing care. Numerous experiments have been performed to identify these acoustics, prosodic, emotional or lexical voice properties for the extraction of health details of the subjects. Some of the medical conditions are:

i. Voice Dysphonia associated with Asthma condition

Inflammation falls down to the vocal chords of the larynx in the nose passageways. Puffy, inflamed cord don't always vibrate correctly, making the voice's speech hoarse and thus impairing the voice's quality. Dysphonia is defined by any alteration of the voice qualitatively and/or quantitatively.

ii. Aphasia as evaluation symptom of Alzheimer

There seems to be no solution for the ailment and patient's issue declines as it advances, which unavoidably prompts demise. The common signs are: absence of memory, vulnerability, peevishness, hostility, language troubles, and emotional episodes. The most significant effect on essential psychological abilities related with Alzheimer's illness is aphasia, an absence of verbal communicational abilities, such as semantic processing breakdowns, superficial vocabulary and coarse vocabulary, challenges in word-finding that contribute to the degradation of spontaneous speaking.

iii. Vocal disorder in Parkinson's disease

It is possible to categorize Parkinson's disease through major differences invoice. Monotonous, diminished tone, failure to change tone, varying rate, quick surges of speech, uncertain consonants, failure to keep up continued vowel phonation and a hoarse and unpleasant voice are the trademark indications of Dysphonia in Parkinson's infection. Parkinson's patients are not able to generate the requisite loudness, pitch, voice modulations and rhythm patterns to express those emotions. Research findings also show that there are fewer words produced by people with Parkinson's (PWP), prolonged delays and abnormal speech rates.

iv. Correlating voice acoustics' in Depression patients

Study tests have connected the acoustic properties of the depressed and the self-destructive subjects by breaking down their words. Study at Georgia Tech and the Georgia Medical College tried patients with mental discouragement on the basis of speech qualities resulting from glottal depression waveform. Alpert et al. reported that acoustic parameters such as fluency and prosody of depression patients were significantly related to medical impressions.

5. Techniques for voice analysis

The dynamic idea of a voice signal which differs in time is making it hard to assess. It likewise represents a significant test to researchers in making a productive voice analysis framework. That is the reason the vast majority of the voice diagnostic techniques so far have been intended to extricate time-varying highlights of the voice signal so as to improve the assessment, deterioration or adjustment of the signal. Speech signal includes voiced areas that identifies with periodicity and reach energy of the signal and with the non-incidental parts, unvoiced bits of the signal fuse. A modified signal or a large number of voice signals will be an ideal voice signal representation. Signals or a collection of parameters with respect to the principal signal with the ultimate objective that terrifically significant information can be found in a more obvious and ordered manner[1]. The idea is to design a language independent script to read so as to extract the spectrum from all the 18 points of articulation. Recording the voice with desired sampling rate and other attributes and the preprocessing for normalizing the voice. The next step is for voice parameter extraction and then to build a machine learning model to establish the correlations with health conditions and accordingly predict the health conditions.

1. Analysis of signal using short-timeframe window:

Short-time window analysis of voice signal depends on time-varying voice signal properties to be captured. Voice is evaluated in this approach for a limited time window interval for which signal properties remain unchanged. In order to ensure high precision and to study the effect of time on the chosen functions, the parameters which are extracted are measured numerous times from different time windows and the effects are then summed. The window function is defined by $W(n) \cdot S(n)$, where $S(n)$ is the voice signal, and $W(n)$ is the signal spectrum area of interest. The window's shape, time and scale depend on the characteristics to be evaluated in the desired application. It is desirable to take the window size as small as possible to reduce random signal noise.

Rectangular windows with large frequency resolution, similar weighing functions and high band leakage are the most widely used windows at the output, creating noise. Using a short window (1) and a large frequency resolution using a longer window, a better temporal resolution is achieved[**Error! Reference source not found.**]. Therefore, to represent the opening, an appropriate size must be selected to represent the exact harmonic structure. It is possible to represent the rectangular window as:

$W_R(n) = \{$ Ndkffsdfsdfs   1 if n=0,…..,R

     0, elsewhere

(1)

2. Time- domain features for voice analysis:

Analysis of the voice signal time space is categorizes as the adjustment of the voice signal into a progression of boundaries which infer an insignificant time change that can be effortlessly examined. Fig.2.1 depicts the time area analysis of speech as sufficiency versus time acknowledged utilizing the Fourier analysis technique. The framework is known as Zero-intersection Rate and Short-Time Auto connection for the time area. The Zero-intersection Rate procedure accumulates ghostly information. Short-Time Auto time signal correlation is the energy spectrum for inverse Fourier transform containing the periodicity information, harmonics and amplitude [1].
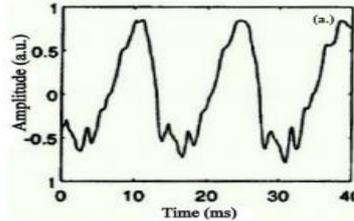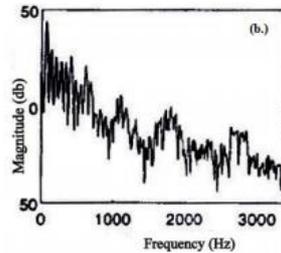
Figure 2.1. Time domain analysis for voice signal

Figure 2.2. Frequency domain analysis for voice signal

3. Frequency-domain features for voice analysis:

The frequency area is an energy part of the signal range, so the voice parts of the frequency space are more significant regarding data qualities than its stage or timing viewpoints. Fig. 2.2 shows frequency space, spoken to by energy versus frequency. Filter Bank Analysis and Short-Time Fourier Transform Analysis can be utilized to infer the frequency space boundaries. A lot of band pass channels are utilized in filter bank analysis to show the phantom dispersion of energy in the ghastly envelope [**Error! Reference source not found.**].

4. Linear predictive coding [LPC] technique:

A powerful and broadly acknowledged methods for accomplishing voice analysis is straight prescient coding. Utilizing the short request filter, LPC makes the short-time connection in the voice tests. The LPC is utilized to assess the music of the voice, the capacity of the vocal plot, the frequency and the signal transfer speed. The example's assessment of voice qualities depends on a direct blend of its past perceptions. The engineering of the Vocoder (an understanding plan where the spectrum of a source signal is weighted by the ghostly segments) is applied by the LPC. The extraction of ghostly envelopes accomplished by LPC analysis is extremely exact. In this manner, in fixed spectrum portrayal, LPC is valuable.

There are two methods to apply the LPC to the signal: least-square autocorrelation and least-square covariance. By analysing the speech signal in the time-limited window, the least-square autocorrelation method minimises the mean energy in the error signal in a sample frame. In the least-square covariance method, instead of the input speech signal S(n), the error signal e(n) is windowed. Prediction and identification of pitch is an advanced process of LPC research accomplished by the implementation of certain major differences in the spectrum. The residual signal would have long-term associations with the speech area that is spoken. Hence, in the second step of prediction, the residual signal is spectrally flattened. To fit 16 samples to 160 samples, the window size is wide enough. In the analysis of the voice signal, the pitch symbolises the basic frequency of a

signal. It is possible to determine the pitch either from the periodicity in the time domain or from the frequency domain's frequently spaced harmonics[**Error! Reference source not found.**].

6. Data description

The training dataset will be made by using the voice samples of some of our college mates and some friends as well as some colleagues, they will be asked to read a small paragraph which helps us to predict the disease. The paragraph will be of over a duration of 1 minute. Henceforth, this dataset will be used for training the model so as the model is able to predict diseases. Further, for the testing dataset, we will request the doctors of associated hospitals for seeking the voice samples of their patients, so that our model achieves a good accuracy.

The dataset is investigated and vital information is chosen and the dataset is changed over into machine reasonable structure. Highlight extraction is the cycle to lessen the size of information to just take enlightening, non-excess and important information, in order to encourage ensuing learning and speculation step to gain better human understanding. The overall sickness forecast framework predicts possibility of essence of an ailment present in a patient based on their side effects. It will likewise prescribe vital careful steps needed to treat the anticipated infection. The framework will at first be taken care of information from various sources for example patients. When the framework preparing is done, our overall body infection forecast model will be prepared for utilizing.

There will be numerous information bases utilized inside this framework including:

1. Database for user details
2. Database for voice recordings
3. Database for user's diseases
4. Database for issues
5. Database for feedback
6. Database for record history

This will include assortment of clinical data relics from different sources like clinics, release slips of patients and from UCI store at that point reprocessing is applied on dataset which will eliminate all the superfluous information and concentrate significant highlights from information. The Disease Prediction model will be prepared on the dataset of ailments to do the forecast precisely and produce Confusion lattice. The prepared and tried forecast model will be conveyed in a genuine situation made by the human specialists and will be utilized for additional improvement in the strategy.

Precision v/s number of perception regarding illnesses forecast model will be estimated. This will help us in forecast of infection.

7. Limitations

After various assessments the following observations and assumption based on information can be made these results.

One drawback is that if we regularly reduce the study, then the frame shift reduces and hence the frame rate increases. This implies that we can process a lot more than sufficient voice analytic data, which will undoubtedly increase the complexity. Owing to the speculative nature of the speech signal, the second drawback of reducing the window length of voice processing analytics is that the spectral predictions will appear to become less prevalent. Based on an analytical assumption the voice recognition use lots of memory and that to very precise requirement of hardware, However, the voice analytic use the less memory as compared to the recognition technique. Voice analytic technology is not quintessential, and although comes with few limitations.

However, to get the best output from the voice analysis, we need a quiet environment, but as we all know about the system don't work properly if their is lot of background noice. The system will not differentiate between the your speech and other ambient noice as well as other natural disturbance which leads to transcript error. This can cause the problem while using voice analysis in the busy office or noice environment. Although if you use the microphone while recording your voice or using the noice cancellation headset can help the system to detect the voice properly and can get you the better analytical result.

8. Conclusion and future work

The inspiration driving this research was to communicate the need of a system for voice analysis syin the pre-evaluation of specific health condition. This paper had presented the different clinical conditions associated with unnatural voice and methodologies that are adopted for assessing characteristic varieties in voice examples of people. The papers read, confirm the accomplishment of the voice investigation frameworks in convenient conditions. However, all these techniques isn't proper for planning a wide range of dysphonia designs. Therefore, there is a need to develop a model for a standard voice assessment for clinical practice.

References

1. Y. S. VM Dixit, "Voice parameter Analysis for disease detection," *IOSR Journal of Electronics and Communication Engineering,* vol. 9, no. 3, pp. 48-55, 201
2. O. Y. G. Ö. M. E. M. ·. A. D. ·. D. C. ·. H. G. R., "Depression Screening from Voice Sample of patients," *Depression Screening from Voice Samples of Patients Affected by Parkinson's Disease,* pp. 72-82, 12 June 2019.
3. R. S. Divya Chitkara, "Voice based Detection of type 2 diabetes," in *2016 2nd International Conference on Advances in Electrical, Electronics, Information, Communication and Bio-Informatics (AEEICB)*, Chennai, India, 2016.
4. A. K. D. Enam Biswas, "Symptom-Based Disease Detection System Using Convolution Neural Network," in *7th International Conference on Smart Computing & Communications (ICSCC)*, Sarawak, Malaysia, 2019.
5. R. K. S. Vishakha Pareek, "Coronary heart disease detection from voice analysis," in *IEEE Students' Conference on Electrical, Electronics and Computer Science (SCEECS)*, Bhopal, India, 2016.
6. T. Tuncer, S. Dogan, F. Özyurt, S. B. Belhaouari and H. Bensmail, "Novel Multi Center and Threshold Ternary Pattern Based Method for Disease Detection Method Using Voice," *IEEE Engineering in Medicine and Biology Society Section,* vol. 8, pp. 84532 - 84540, 2020
7. P. S. P. B. Ms.Deepali S. Shrivastava(P.G Student), "Voice Analysis Based Disease Identification Using Speech Processing,," *International Journal of Emerging Technologies and Innovative Research ,* vol. 3, no. 1, pp. 47-49, 2016.
8. O. B. B. O. A. K. D. H. O. O. H. T. ,. K. P. L. ,. S. S. T. ,. S. L. a. J. M. C. O. Hwan Ing Hee, "Development of Machine Learning for Asthmatic and Healthy Voluntary Cough Sounds: A Proof of Concept Study," pp. 1-14, 16 july 2019.
9. S. D. J. K. Suman Deb, "Analysis and Classification of Cold Speech using Variational Mode Decomposition," *IEEE Transactions on Affective Computing ,* vol. 11, no. 2, pp. 296 - 307, 2017.
10. M. S. A. K. B. S. Nicholas Cummins, "Automatic Recognition of Speech Affected by Upper Respiratory Tract Infection," in *39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Seogwipo, South Korea, 2017.
11. R. C. G. B. J. D. R. G. C. M. P. Everthon Silva Fonseca, "Acoustic investigation of speech pathologies based on the discriminative paraconsistent machine (DPM)," *Biomedical Signal Processing and Control,,* vol. 55, pp. 1746-8094, 20 August 2019.
12. A. S. Rimah Amami, "An incremental method combining density clustering and support vector machines for voice pathology detection," *Computers & Electrical Engineering,,* vol. 57, pp. 257-265, 2016.
13. M. Ł. Ł. K. Mateusz Soliński, "Automatic cough detection based on airflow signals for portable spirometry system," *Informatics in Medicine Unlocked,,* vol. 18, 2020.
14. D. Y. A. A. N. J. N. B. S. H. T. F. Q. James R. Williamson, "Tracking depression severity from audio and video based on speech articulatory coordination," *Computer Speech & Language,,* vol. 55, pp. 40-56, 2018.
15. G. A. M. A. Z. A. T. A. M. M. F. K. H. M. A. A.-n. Ghulam Muhammad, "Automatic voice pathology detection and classification using vocal tract area irregularity,," *Elviser,* vol. 36, no. 2, pp. 309-317, 2016.
16. S. A. I. ,. E. R.-V. Renard Xaviero Adhi Pramono, "Automatic Cough Detection in Acoustic Signal using Spectral Features," in *41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Berlin, Germany, 2019.
17. A. I. Renard Xaviero Adhi Pramono, "Automatic detection of patients with respiratory diseases using lung sound analysis," in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society.*, 2018.

18. N. K. V. Shobana, "A personalized recommendation engine for prediction of disorders using big data analytics," in *International Conference on Innovations in Green Energy and Healthcare Technologies (IGEHT)*, Coimbatore, India, 2017.

19. A. D. D. Anna Kasperczuk, Recommender system for diagnosis of colon diseases, Bialystok, Poland: IEEE, 2018.

20. G. &. T.-R. M. &. T. V. &. L. M. &. L.-R. B. &. Q. R. &. M. M. &. A. W. Guzmán, "A Collaborative Framework for Sensing Abnormal Heart Rate Based on a Recommender System: Semantic Recommender System for Healthcare," *Journal of Medical and Biological Engineering,* May 2018.

21. W. &. X. Z. &. Z. N. &. W. Y. Hong, "A Medical-History-Based Potential Disease Prediction Algorithm," *IEEE Access,* vol. 7, pp. 131094 - 131101, 11 september 2019.

22. M. &. S. H. &. A. I. &. A. J. Sabaee, "A Preliminary Acoustic Analysis of Laryngectomized speech in Adult," in *16th International Workshop on Acoustic Signal Enhancement (IWAENC)*, chicago, 2018.

23. G. &. A. M. &. A. Z. &. M. T. &. F. M. &. M. K. &. A.-n. A. &. B. M. Muhammad, "Voice Pathology Detection Using Interlaced Derivative Pattern On Glottal Source excitation," *Biomedical Signal Processing and Control,* vol. 31, pp. 156-164, 2018.

24. M. A. I. E. G. M. T. A. M. M. F. K. H. M. Zulfiqar Ali, "Voice pathology detection based on the modified voice contour and SVM," *Biologically Inspired Cognitive Architectures.,* vol. 15, pp. 10-18, 2016.

25. P. O. F. N. A. João Paulo Teixeira, "Vocal Acoustic Analysis –Classification of Dysphonic Voices With Artificial Neural Networks," *Procedia Computer Science.,* vol. 121, pp. 19-26, 2017.

26. T. U. Mane, "Smart heart disease prediction system using Improved K-Means and ID3 on Big Data," in *International Conference on Data Management, Analytics and Innovation (ICDMAI)*, Pune,India, 2017.

27. E. J. P. G.-V. Z. S. I. R. I. E. M. K. M. M. J. B. A.-H. M. F.-Z. K. L.-d.-I. Jiri Mekyska, "Robust and complex approach of pathological speech signal analysis," *Neurocomputing,,* vol. 167, pp. 94-111, 2015.

28. A. Ed-Daoudy and K. Maalmi, "Real - Time machine learning for early detection of Heart Disease using big data approach.," in *nternational Conference on Wireless Technologies, Embedded and Intelligent Systems (WITS),* , Morocco, 2019.

29. N. K. S. Banu and S. Swamy, "Prediction of Heart Disease at early stage using Data Mining and Big Data Analytics: A Survey.," in *International Conference on Electrical, Electronics, Communication, Computer and Optimization Techniques (ICEECCOT),*, Mysuru, 2016..

30. A. &. M. G. &. A. M. &. A. Z. &. M. K. &. M. T. &. F. M. nasheri, "Pathological Voice Detection And Binary Classification Using MPEG-7 Audio features.," *Biomedical Signal Processing and Control,* vol. 11, pp. 1-9, 2014..

31. D. Z. G. L. Z. G. Kebin Wu, "Joint Learning For Voice Based Disease detection.," *Elvesier Pattern Recognition,,* vol. 87, pp. 130-139,, 2019.

32. L. M. R. C. J. H. P. C. a. R. P. P. B. M. Rocha, "Detection of Explosive Cough Events in Audio Recordings by Internal Sound Analysis,," in *IEEE Engineering in Medicine and Biology Society (EMBC)*, Seogwipo, 2017..

33. C. &. W. L. Alexander, "Big Data Analytics in Heart Attack Prediction,," *Journal of Nursing & Care,* vol. 6, 2017..

34. S. A. I. a. E. R.-V. R. X. Adhi Pramono, "Automatic Identification of Cough Events from Acoustic Signals," in *41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC),*, Berlin, Germany, 2019..

35. H. &. S. A. A. &. S. C. &. C. T. Mohammadi, "Automatic Discrimination Between coughAnd No-cough Accelerometry Signal artefacts.," *Biomedical Signal Processing and Control. ,* vol. 52, 2019.

36. C. H.-B. K. D. P. C.-d.-l.-H. Jesús Monge-Álvarez, "Audio-cough event detection based on moment theory,," *Applied Acoustics,,* vol. 135, pp. 124-135, 2018..

37. M. Berkay Yılmaz and M. A. E. Yacoubi, "Method of Pathology Detection by Speech Analysis: Survey,," in *International Conference on Computer Science and Engineering (UBMK)*, Antalya, , 2017.

38. S. C. A. &. N. J. Upadhya, "Multitaper perceptual linear prediction features of voice samples to discriminate healthy persons from early stage Parkinson diseased persons,," *Int J Speech Technol ,* vol. 21, p. 391–399 , 2017..

39. L. Z. Z. T. a. H. Z. X. Zhang, "Acoustic Characteristics of Normal and Pathological Voices Analysis and Recognition,," in *nternational Conference on Systems and Informatics (ICSAI)*, Shanghai, China, 2019..

Authors

Manasi Bendale
B. Tech [Information Technology], SVKM Narsee Monjee Institute of Management Studies, Shirpur, India, manasib311@gmail.com

Jaya Soni
B. Tech [Information Technology], SVKM Narsee Monjee Institute of Management Studies, Shirpur, India, jayasoni6287@gmail.com

Garvit Chhugani
B. Tech [Information Technology], SVKM Narsee Monjee Institute of Management Studies, Shirpur, India, garvit.chhugani17@nmims.edu.in

Nikhil Sahu
B. Tech [Information Technology], SVKM Narsee Monjee Institute of Management Studies, Shirpur, India, nikhil.sahu41@nmims.edu.in

Dr. Nitin Choubey
Professor & Head of Information Technology Department, SVKM Narsee Monjee Institute of Management Studies, Shirpur, India
nitin.choubey@nmims.edu

Dr. Santosh Bothe
Assistant Professor, Computer Engineering Department, SVKM Narsee Monjee Institute of Management Studies, Shirpur, India
santosh.bothe@nmims.edu