The Journal of Computational Science and Engineering.
ISSN: 2583-9055

DETECTION OF LARYNGEAL CANCER USING AUDIO PROCESSING

V.G.Lachake , Rushikesh D. Sonawane , Shrutika B. Pagare , Vedika S. Taji ,

Samruddhi V. Wakalkar

Department of Computer Engineering, Sanjivani K. B. P. Polytechnic, Kopargaon

Maharashtra State Board of Technical Education, Mumbai, Maharashtra, India

Keyword:

Audio analysis, Cancer detection, voice changes, Machine learning, Vocal patterns,

ABSTRACT

This project presents a groundbreaking method for early laryngeal cancer detection through an innovative blend of audio processing and cutting-edge machine learning algorithms. By scrutinizing voice recordings, this system aims to identify potential indicators of laryngeal abnormalities, potentially signifying the presence of cancer. Unlike traditional invasive diagnostic techniques, this approach offers a non-intrusive, accessible, and potentially real-time solution for detecting potential cases. The heart of this approach lies in the meticulous analysis of vocal characteristics, including pitch modulation, intensity fluctuations, resonance traits, and even nuanced pronunciation changes. These features are extracted from voice recordings, unveiling patterns that might point to laryngeal irregularities. Critical to this methodology is the integration of advanced machine learning models. These models are trained on a diverse dataset, encompassing both healthy voice samples and those from individuals with confirmed laryngeal cancer. Through this training process, the model discerns the subtle distinctions in vocal attributes between these two groups. When confronted with new, unlabeled voice recordings, the model is challenged to predict the likelihood of laryngeal abnormalities. Rigorous validation and accuracy assessment play an instrumental role in this endeavor. Comprehensive datasets, capturing a wide spectrum of voice characteristics, are employed to meticulously evaluate the model's predictive capabilities. Collaborations with medical experts further bolster the validation process, ensuring that the system's predictions align with expert medical opinions. The contributions of this project extend to various facets. Foremost, it strives to reshape the landscape of laryngeal cancer diagnosis. The prospect of early detection enabled by this system could lead to timely interventions, potentially altering outcomes and enhancing patient prognosis. Moreover, the non-invasive nature and real-time analysis potential offer a more accessible approach, bridging gaps in healthcare accessibility. From an economic perspective, the system's capacity to minimize the need for intricate diagnostic procedures could contribute to cost mitigation. In conclusion, this project introduces an avant-garde approach to laryngeal cancer detection, fusing audio processing and machine learning. This integration holds the promise of inaugurating a novel era of diagnostics that is accurate, attainable, and patient-centric. With ethical considerations governing patient data privacy and confidentiality,

Corresponding Author: Email: rushisonawane2514@gmail.com,

INTRODUCTION

Voice-based identity verification to voice pathology detection, are nowadays ubiquitous in our daily life. A significant attention has been paid to the science of voice pathology diagnostic and monitoring as it is offering solutions to companies seeking for efficiency enhancement with simultaneous cost saving, the market of speech technology forecast to be particularly promising in the next years. The voice disorders are caused due to defects in the speech organs, mental illness, hearing impairment, autism, paralysis, or multiple disabilities.

The main aim of this system is to help patients with pathological problems for monitoring their progress over the course of voice therapy. Currently, patients are required to routinely visit a specialist to follow up their progress. Furthermore, the existing ways to detect the voice pathology are subjective, invasive methods such as the direct inspection of the vocal folds and the observations of the vocal folds by endoscopic instruments are done. In addition, the present systems are usually based on information related to the vocal tract configuration, the airflow passing through the vocal folds, and called glottal flow. These techniques are expensive, risky, time consuming, discomfort to the patients and require costly resources, such as special light sources, endoscopic instruments and specialized video-camera equipment.

To avoid the above problems, a robust system is implemented to detect vocal fold pathology at an early stage from set of features like 12 Mel-Frequency Filter Bank Cepstral Coefficients (MFCC) and zero crossing rate (ZCR) derived from simple voice sample. The system helps the clinicians and speech therapists for early detection of vocal fold pathology and can improve the accuracy of the assessments using Saarbruecken Voice Database.

The proposed system can analyse voice source features in speech data for detection of glottal pathology using data mining techniques. SVM classifier is developed for various feature combinations to classify the glottal pathologic voice from normal voice. The system is implemented using 12 Mel-Frequency Filter Bank Cepstral Coefficients (MFCC) and zero crossing rate (ZCR).

LITERATURE SURVEY

Title: “Survey on Signal Processing Based Pathological Voice Detection Techniques”
Author: R. Islam et al
Abstract: The motivation of the work is to address the need for non-invasive signal processing techniques to detect voice disability in the general population. This paper conducts a survey related to voice disability detection methods. The paper contains two main parts. In the first part, we present background information including causes of voice disability, current procedures and practices, voice features, and classifiers. In the second part, we present a comprehensive survey work on voice disability detection algorithms. The issues and challenges related to the selection of voice feature and classifier algorithms have been addressed at the end of this paper.[1]

Title: "Voice pathology detection based on the modified voice contour and SVM"

Author: Ali, Z et al.

Abstract: In this study, a novel method based on the voice intensity of a speech signal is used for automatic pathology detection with continuous speech. The proposed method determines the peaks from the speech signal to form a voice contour. The area under the voice contour allows us to discriminate between normal and disordered subjects.[2]

PROPOSED METHODOLOGY

Use an advanced algorithm to analyze vocal features in real-time, aiming to identify potential indicators of laryngeal cancer from voice recordings.
Design and implement a machine learning model that distinguishes between healthy and potentially cancerous vocal patterns with high accuracy.
Create a user-friendly interface that allows individuals to submit voice recordings for automated analysis, facilitating easy and early cancer detection.

I. MATHEMATICAL MODEL

Let us consider S be a Systems such that

S= {U,FT,P S,T,Ss,Ds}, where

- U= {U1, U2, U3. Un | 'U' is a Set of all Voters }

There may be number of users for making use of system. So this is the Infinite Set.

- I = {I | is the speech input of the system }

- FT= {F1, F2, F3. Fn | F are the features extracted }

FT is the feature extracted from speech input.

- T={ T is the technique used to process Input voice } P is the technique used to pre-process the input voice.

T is the technique used for detecting glottal differences from voice.

SS = {S REG, S LOGIN, S Feature Data| SS is a Set of Storage Service } STORAGE SERVER will provide four services like Registration, Login, and glottal parameters As this set also has finite attributes, so this is also Finite Set.

- DS = {Train Data| DS is a Set of data table for permanent storing of data on server }

EVENTS and ACTIVITIES:

- EVENT 1

User will make registration on SYSTEM & Storage Server. Let f(U) be a function of User Thus, f(U)→ {Ss}

- EVENT 2

Speech Input is given to the system.

Let f(U) be a function of User. Thus, f(U) → {I1,I2,I3. In } ε S

A] WORKING

Speech Input

The main aim of the proposed system is the extraction of glottal parameters from speech signal for the distinction between glottal pathological voice and normal voice. Firstly input speech data is given to the system which contain database of patients suffering from glottal and supra-glottal cancer and from normal persons.

Preprocessing: The speech data contain lot of noise. By using noise removal techniques or removing any other disturbances present in the data, it is preprocessed to get the fine-tuned data.

Feature Extraction:

The system can detect vocal fold pathology at an early stage from set of features like MFCC and ZCR, which are vocal tract parameters, in detection of the glottal Pathologies. a. MFCC (mel-frequency cepstrum (MFC)) The first step in any automatic speech recognition system is to extract features i.e., identify the components of the audio signal that are good for identifying the linguistic content and discarding all the other stuff which carries information like background noise, emotion. The Mel-Frequency Cepstral Coefficients (MFCC) features is the most used features in speaker recognition. It combines the advantages of the cepstrum analysis with a perceptual frequency scale based on critical bands. MFCC is based on Human hearing perceptions which cannot perceive frequencies over 1 Khz. In other words, in MFCC is based on known variation of the human ear's critical bandwidth with frequency. b. Zero Cross Rate (ZCR): The zero-crossing rate is the rate of sign-changes along a signal, i.e., the rate at which the signal changes from positive to zero to negative or from negative to zero to positive. The zero-crossing rate is strongly correlated with the spectral centroid, which can be computed using the MFCC spectrum, and is a measure for the high frequency content of a signal.

Classification:

Once the features extracted and normalized, we train a Support Vector Machine model. According to preliminary tests, the best results were achieved. SVM classifier is used for various feature combinations to classify the glottal pathologic voice from normal voice.

SYSTEM ARCHITECTURE

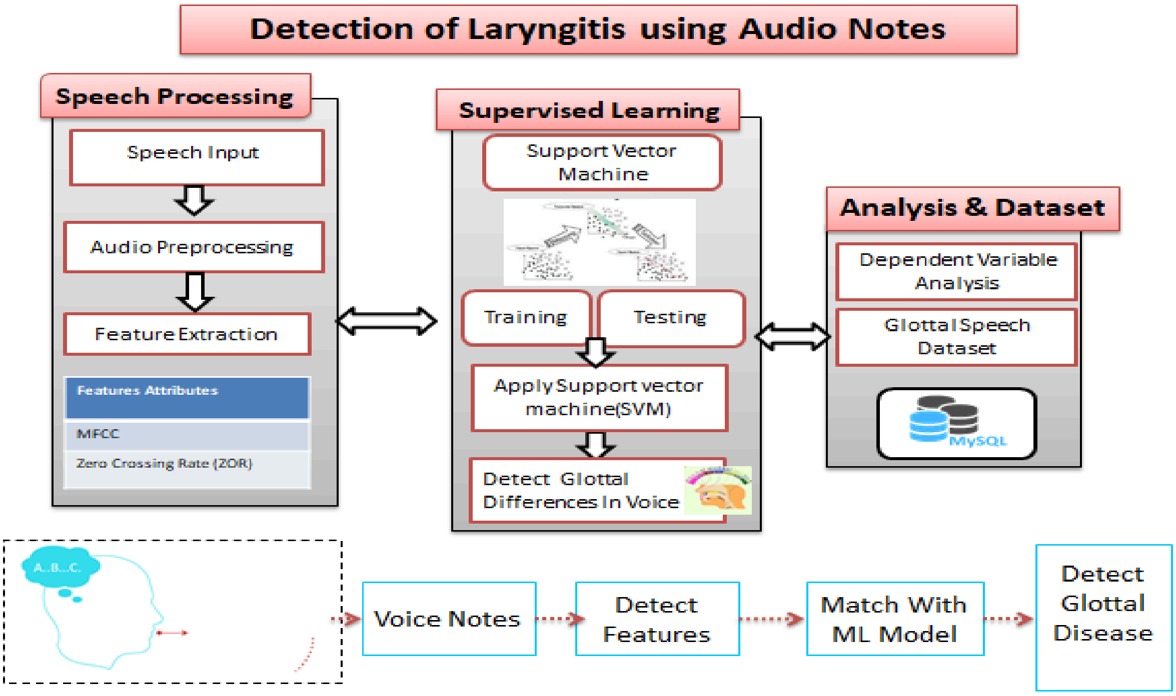

Fig 1. Block Diagram

FUTURE SCOPE

Future developments in audio processing applications for the identification of laryngeal cancer have great promise to transform early diagnosis and enhance patient outcomes. These apps can be extremely helpful in effectively and non-invasively screening, monitoring, and identifying laryngeal cancer by utilising developments in both technology and healthcare.

The improvement of algorithms for audio analysis is one of the main areas of progress. Researchers and developers can improve the app's capacity to identify small irregularities in voice that are linked to laryngeal cancer by continuously improving these algorithms. To find possible indicators of cancer, this entails examining a variety of aspects of voice recordings, including pitch, frequency, intensity, and temporal patterns. Adding machine learning techniques to the programme can help it learn from bigger datasets as these techniques advance, which will increase the app's accuracy and dependability over time.

Another interesting direction for future research and development is real-time monitoring capabilities. By analysing voice recordings in real-time, the software can notify users of any anomalies that might require additional research and give them quick feedback. When it comes to early detection, this real-time functionality is very helpful because it enables people to seek medical assistance as soon as they notice any worrying voice patterns.

Accessibility is a crucial aspect of any healthcare technology, and laryngeal cancer detection apps are no exception. Ensuring that these apps are user-friendly and accessible to individuals from diverse backgrounds is essential for maximizing their impact. This may involve designing intuitive user interfaces, providing clear instructions for recording voice samples, and offering support for multiple languages and dialects.

In order to confirm the efficacy of laryngeal cancer detection applications in clinical settings, cooperation with medical specialists is essential. Through collaboration with otolaryngologists, oncologists, and other medical specialists, developers might get significant knowledge on the subtleties of diagnosing and treating laryngeal cancer. Furthermore, performing longitudinal studies and clinical trials might yield strong proof of the app's safety and effectiveness, which could eventually result in its broad use in healthcare settings.

The diagnostic capabilities of the software can be further improved by integrating multimodal data into its analysis. Apart from audio recordings, incorporating additional data formats, like textual or visual imagery, might offer further perspectives on the state of a patient. For instance, integrating audio analysis with imaging modalities such as magnetic resonance imaging (MRI) or laryngoscopy can provide a more thorough view of the larynx and adjacent structures, facilitating the identification and characterization of tumours.

Applications for the identification of laryngeal cancer are among the many healthcare technologies for which privacy and security are critical factors in development. Gaining user trust and adhering to legal standards necessitate the implementation of strong methods to protect sensitive user data, such as encryption, anonymization, and secure data storage.

Ultimately, there is a wide range of potential applications for audio processing software in the identification of laryngeal cancer. Developers may design effective solutions that promote early detection, enhance patient outcomes, and ultimately save lives by focusing on algorithm improvement, real-time monitoring, accessibility, collaboration with healthcare experts, multimodal data integration, privacy, and security.

RESULTS

The proposed system can help to detect glottal pathologies from the voice input. The system uses MFCC which are vocal tract parameters, in the detection of the glottal pathologies and ZCR for the detection. The SVM classifier is used to train the data. The classification of pathologies is carried out with the help of the SVM classification algorithm.
In our proposed system we used a large Saarbruecken Voice Database. Where accuracy and precision are calculated based on false positives images, i.e. which are items incorrectly labelled as belonging to the class and false negatives, which are items which were not labelled

as belonging to the positive class but should have been.

1. TP: Positive samples classified as positive.
2. TN: Negative samples classified as negative.
3. FP: Negative samples classified as positive.
4. FN: Positive samples classified as negative

$$Precision = \frac{TP}{(TP+FP)}$$

$$Recall = \frac{TP}{(TP+FN)}$$

For the mentioned classes the accuracy and precision are calculated by using the formula. The precision is the percentage of documents that are correctly classified as positive out of all the documents that are classified as positive. Where, TP, FP, and FN are truly positive, false positive and false negative images.

CONCLUSION

In summary, this project introduces a revolutionary approach to laryngeal cancer detection by amalgamating audio processing and machine learning. By scrutinizing vocal patterns, the system offers a non-intrusive and accessible means of early cancer detection. Its potential impact spans enhanced patient outcomes, improved diagnostic efficiency, and increased healthcare accessibility. While upholding ethical standards, the project's culmination signifies a pivotal stride toward merging technology and healthcare, promising a future where the power of voice becomes a formidable tool in the fight against cancer.

REFERENCES

1. YUANBO WU , CHANGWEI ZHOU, ZIQI FAN, DI WU, XIAOJUN ZHANG, AND ZHI TAO ,“Investigation and Evaluation of Glottal Flow Waveform for Voice Pathology Detection”,IEEE access, VOLUME 9, 2021.

2. Anis Ben Aicha,"Noninvasive Detection of Potentially Precancerous Lesions of Vocal Fold Based on Glottal Wave Signal and SVM Approaches “, International Conference on Knowledge Based and Intelligent Information and Engineering Systems, KES2018, 3-5 September 2018, Belgrade, Serbia.

3. Pranav S. Deshpande ; M. Sabarimalai Manikandan ,"Effective Glottal Instant Detection and Electroglottographic Parameter Extraction for Automated Voice Pathology Assessment", IEEE Journal of Biomedical and Health Informatics (Volume: 22 , Issue: 2, March 2018.

4. Korutla Sudhir Sai1, Polasi Phani Kumar,"Glottal Analysis UsingSpeech signals "International Journal for Research in Applied Science & Engineering Technology (IJRASET),2017.

5. Sudarsana Reddy Kadiri; Paavo Alku ,"Analysis and Detection of Pathological Voice Using Glottal Source Features", IEEE Journal of Selected Topics in Signal Processing (Volume: 14, Issue: 2, Feb. 2020).

6. Kadria Ezzine; Mondher Frikha ,"Investigation of glottal flow parameters for voice pathology detection on SVD and MEEI databases", 2018 4th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP).