IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

## IN COMPUTER & COMMUNICATION ENGINEERING

Volume 12, Issue 7, July 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 7, July 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1207012 |

# Automatic Pathological Voice Detection using Audio Processing

**Vikram G. Lachake**

Lecturer, Sanjivani K.B.P. Polytechnic, Kopargaon, India

**ABSTRACT:** The prevention and early detection of voice diseases can provide effective help for the treatment and rehabilitation of patients. The pathological voice detection is done through signal processing and machine learning. The glottal flow waveform is estimated by glottal inverse filtering method (IAIF). The features extracted from glottal flow waveform used for to test and train classifier like Random forest classifier or Support Vector Machine (SVM) for voice pathological detection and normal voce.

The 10 f0ld cross validation used to get the best result. This method contributes better accuracy of pathological voice detection compared to other similar methods because the feature set used extracted from glottal flow waveform.

**KEYWORDS:** Glottal flow waveform, Glottal invers filtering, Random forest classifier, Support Vector Machine (SVM).

## I. INTRODUCTION

Speech is produced by exciting a time-varying vocal tract system that consists of various articulators (such as the tongue, jaw, lips) by a time-varying excitation signal. The main purpose of speech in communication is to convey a linguistic message. Apart from linguistic content, speech also contains rich information about the language and dialect as well as about the speaker's gender, age, emotions and state of health. This work studies pathological voice and compares it to normal voice using both analysis and detection (i.e., normal vs. pathological). Voice pathologies arise due to infections, physiological and psychogenic causes and due to vocal misuse that is prevalent in professions such as singers, teachers, and customer service representatives . Automatic detection of voice pathology enables an objective assessment and an early intervention for the diagnosis. A typical voice pathology detection system consists of two main stages: the first stage is the representation of the input acoustic speech signal (i.e., Speech is a type of signal produced by the human vocal organs (such as the lungs, vocal folds, nasal cavity and lips), and represents a certain practical significance, especially for social communication. Causes of voice diseases generally include misuse of pronunciation, improper use of voice, local infection or mental problems. Furthermore, the proportion of patients with pathological voices is higher for specific professions such as teachers, singers and lawyers . The prevention and early detection of voice diseases can provide effective help for the treatment and rehabilitation of patients. However, an laryngoscope is often used to diagnose the disease by directly observing the patient's larynx, which is often invasive and uncomfortable for the patient. Therefore, it is important and valuable to study the noninvasive and objective methods of pathological voice detection based on signal processing and machine learning. A typical voice pathology detection system is usually composed of two main procedures: (1) features extractionand (2) results classification. Hence, the key points that can effectively improve the performance of the detection system are the feature set used and the classifiers utilized. The Gaussian mixture model (GMM), hidden Markov model (HMM), support vector machine (SVM) and random forest (RF) methods have been widely used due to their outstanding performances, especially SVM . With the development of neural network technology, methods such as artificial neural networks (ANNs) , deep neural networks (DNNs) and convolutional neural networks (CNNs) have also been applied to the study of pathological voice [5].

## II. LITRATURE SURVEY

**YUANBO WU and CHANGWEI ZHOU [1] stated t**o investigate the effectiveness of the glottal flow waveform and improved the performance of voice pathology detection, a proposed pipeline system consisting of a feature extractor and a separate classifier was built in this paper. The acoustic features computed using the openSIMLE toolkit, the audio features calculated using the MPEG-7 standard and the glottal features obtained using the glottal flow estimated with the IAIF glottal inverse filtering method were extracted from raw speech utterances and the corresponding glottal source signals. The individual and combined feature sets were utilized as input parameters for the RF classifier to detect pathological voices.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 7, July 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1207012 |

Anis Ben Aicha [2] proposed method that takes advantage for the non-invasive process based on the recorded speech. First, they have extracted the source signal, namely the glottal flow signal from the acoustic speech using IAIF technique. Some descriptors are built from the time and frequency representations of the glottal flow signal. At the second stage, they have studied the relevance of used descriptors using boxplot and PCA analysis. As a result, they kept only those which are the most pertinent. Finally, an SVM module is used to classify and discriminate premalignant lesions from normal voices. Experimental results show the validity of the proposed idea with a rate of classification accuracy about 92%.

Pranav S. Deshpande , M. Sabarimalai Manikandanv [3] presented an effective method based on the aVMD autocorrelation feature for automatically determining the glottal closure and opening instants from EGG signals including both voiced speech and non-speech portions contaminated with different kinds of artifacts and noise. The proposed method consists of five stages: the VMD-based baseline and HF noise removal, the EGG feature signal extraction using mode center frequency criterion, the zero crossing based glottal instant detection, the ACF feature based nonglottal instant removal and the EGG parameter extraction including pitch period, CQ, OQ, and SQ. The accuracy and robustness of the method is tested using noise-free and noisy EGG signals.

Sudarsana Reddy Kadiri; Paavo Alku [5] presented system in which glottal source features were studied in the analysis of normal and pathological voice, and these features were further used in voice pathology detection. The glottal source features were derived from three signals: from the glottal flows estimated with the QCP inverse filtering method, from the approximate source signals computed with the ZFF method and directly from acoustic voice signals. Analysis of features revealed that glottal source features help in discriminating normal voice from pathological voice. Detection experiments were carried out using two databases with individual glottal source feature sets and with a combination of features. Experiments showed that on their own the studied glottal source features provide better discrimination compared to spectral features such as MFCCs and PLPs features. Also, it was shown that complementary information exists among the different glottal source features. Further, the combination of the existing spectral features with the glottal source features resulted in improved detection performance, indicating the complementary nature of features.

## III. METHODS

Figure 1 describes the structure of the proposed voice pathology detection method. This traditional method is often used in the previous works, which usually consists of source signal preprocessing, feature extraction and classifiers. Most researchers focused on how to solve the problems in feature extraction and classifiers, such as multi-band analysis, dimensionality reduction , classification algorithm fusion. However, the importance of the source signals used is rarely systematically studied and analyzed. As a matter of fact, voice diseases usually directly affect the tissue structure of the vocal folds, such as asymmetry and unilateral paralysis. But in previous works, raw speech signals affected by vocal tract resonance and lip radiation were generally used as the source signals, rather than glottal flow waveforms which are directly related to vocal folds vibration. To systematically study the contribution of glottal flow waveforms as source signals to the improvement of pathological voice detection performance, raw speech utterance is used as a contrast to the glottal flow waveform that is estimated by using the glottal inverse filtering method of IAIF in this work. The features extracted from the glottal flow waveform are the acoustic features obtained using openSMILE (openSMILE-G), the audio features calculated by MPEG-7 standard (MPEG-7- G) and the glottal source features computed with time- and frequency-domain measurements. For raw speech signals from the voice database, openSMILE toolkit and MPEG-7 standard are used to calculate the two feature sets, called openSMILE-R and MEPG-7-R, respectively. With regard to the openSMILE and MPEG-7 features, the parameters settings in feature extraction have no differences with respect to raw voice and glottal signal, to investigate the effectiveness of the glottal information in voice pathology detection. Individual or combined features are used as input parameters to train and test the machine learning classifier RF. To obtain robust and reliable experimental results, 10-fold cross-validation is used . The input data set will be divided into 10 portions of equal size each time: 9 for training and 1 for testing. The total number of experiments carried out is 10. In the end, the mean value of the results of 10 experiments was taken as the final accuracy for pathological voice detection[1].

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 7, July 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1207012 |

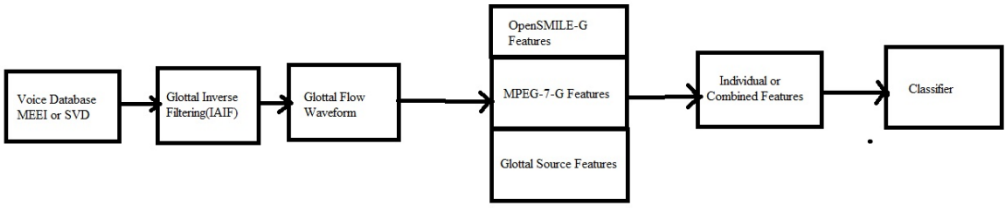


Figure 1

## IV. RESULT

Performance of the classifier was improved by using the combined feature set of OpenSMILE-G features, MPEG-7-G Features and Glottal source features.

## V. CONCLUSION

To improve the performance of voice pathological detection, the proposed system consisting of feature extraction from glottal flow waveform and the classifier like Support Vector Machine (SVM) or Random Forest classifier used to separate normal voice from pathological voice detection. The glottal features obtained using glottal flow estimated with IAIF glottal inverse filtering method extracted from raw speech and corresponding glottal source signal. The results also indicated feature sets extracted from glottal source signals leads to best classification performance which simply classify pathological voice detection from glottal waveform.

## 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).

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi crossref

निस्केयर NISCAIR

# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

## IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details