JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue

JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)

An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Machine Learning based Voice Disease Detection System

[1]Sairaj P. More, [2]Prof. N. M. Wagdarikar, [3]Mohit P. Pawar, [4]Rutik P. Ramteke

Department of Electronics & Telecommunication engineering

Smt. Kashibai Navale College of Engineering Pune, India

[1]sairaj.ydbhc19@sinhgad.edu, [2] nmwagdarikar@sinhgad.edu, [3]mohit.aephc19@sinhgad.edu, [4]rutik.deahb19@sinhgad.edu

Abstract

The use of mobile devices in the healthcare sector is rapidly rising. Mobile technologies provide not only forms of communication for multimedia content (e.g., clinical audio-visual notes and medical records), but also potential solutions for people who want to detect, monitor, and treat their health concerns at any time and from any location. Mobile health systems can help to improve, speed up, and lower the cost of patient treatment. The usage of mobile technologies can benefit a variety of pathological disorders. This paper focuses on dysphonia, a change in voice quality that affects one out of every three people at some point in their lives. Voice disorders are quickly spreading, despite the fact that they are frequently overlooked. Mobile health systems can provide simple and quick support for vocal pathology detection. To realise a legitimate and precise mobile health system, an algorithm that discriminates between sick and healthy voices with greater precision must be identified. This paper's main contribution is to explore and compare the performance of several machine learning approaches beneficial for speech pathology identification. All studies are carried out on a set of voices drawn from the voice database. The collected results are graded based on their accuracy, sensitivity, specificity, and receiver operating characteristic region. They demonstrate that the machine learning method achieves the highest accuracy in detecting voice illnesses.

Keywords: Machine Learning, voice disease, classification.

Introduction

About 5% of the world's population suffers from vocal illnesses or signs of voice ailments to varying degrees. Voice sickness has an impact on human communication, life, and work quality. Voice pathology is caused by a variety of conditions, including significant organ damage, air pollution, smoking, dietary habits, and stress. Voice illness is more common in persons who must talk loudly and regularly for extended periods of time for professional reasons, such as teachers, singers, auctioneers, lawyers, and

actors. Laryngeal endoscopy is always used to assess the site of laryngeal lesions, and laryngeal electromyography is used to detect laryngeal muscle and vocal cord illnesses. These two clinical approaches of detecting voice disorder rely heavily on clinical experience and subjective judgement of clinicians, and they are also more expensive. There is an intuitive biological relationship between speech conditions and the acoustic characteristics of healthy and various types of diseased sounds, which also provides a physical explanation for the speech waveforms generated and transmitted by sound organs. The extraction and processing of patient speech signal parameters allows for objective and scientific evaluation of voice quality.

Literature Survey

Decades of research has proved that the application of machine learning method (SVM, GMM, ELM) in the classification of health and disease speech is effective and accurate. Researchers studied and applied various Deep learning architectures which include a 5-layer network, 5-layer CNN and RNN. In the experiment, 5-layer CNN got a satisfying result, which received 96% sensitivity and 18% specificity [1]. The sensitivity and specificity mentioned here are related to the recall and precision in the medical field. Sensitivity, figuratively speaking, is given a series of samples, how many proportions of patients can be judged; specificity is the ability to determine a certain disease. In [2] a Deep belief network has been proposed to extract feature with SVM to accomplish the classification. Also, some researchers built and improved a DNN model, which includes 5 hidden layers with 200 neurons. Balanced the sensitivity and specificity by adjusting parameters and achieved a result of 93.1% and 46% [3]. Traditional machine learning models (such as Support Vector Machine, Random Forest, K-Nearest Neighbor, and Gradient Boosting) have also achieved reliable results in classification [4][5]. For four-category problems, researchers proposed convolutional architectures combined with feed-forward neural networks, to classify four types of voice disorders and achieved 57% accuracy [6].

Proposed System

Almost all automatic speech recognition systems begin by identifying audio components that can be used for content identification or feature extraction. As illustrated in Fig. 1, the study work in this section of this paper is divided into three components.

The gathered audio signal is first preprocessed in wav format; next the features of the preprocessed audio signal are retrieved; and finally, an algorithm model for training and testing is established.

System Architecture:

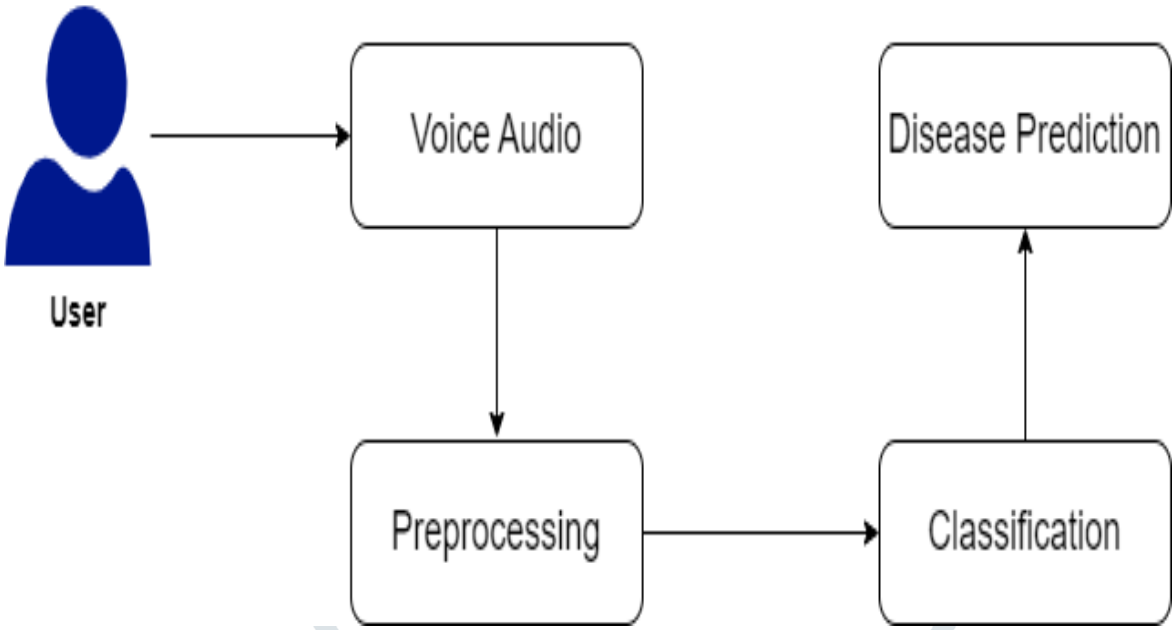
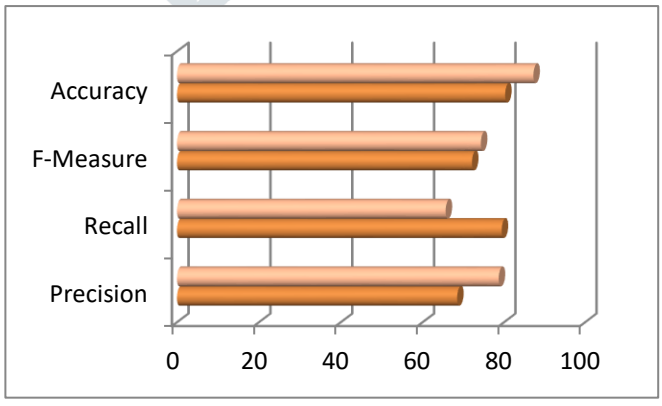


Figure 1. System Architecture

Result and Discussion

Experiments are done by a personal computer with a configuration: Intel (R) Core (TM) i3-2120 CPU @ 3.30GHz, 4GB memory, Windows 7. The application is web application used tool for design code in VS code and execute on Tomcat server.

We compared the proposed disease prediction accuracy on number of samples and show the result graphically. Let see the following graph and table shows the disease prediction accuracy result based on classification technique.



	Existing System	Proposed System
Precision	68.45	77.70
Recall	79.44	66.64
F-Measure	72.11	74.31
Accuracy	80.29	88.26

Figure 2 : Performance Evaluation of proposed system

Conclusion

We suggest a more accurate and rapid technique to classify four prevalent voice diseases in order to fulfil the clinical needs of Xiangya Hospital and increase the diagnostic efficiency of voice diseases. First, we trained the model using normal quality patient audio (8 phonemes) gathered in the clinical setting, then changed the parameters and examined the classification effect, which had an accuracy of 88% and a recall rate of 66%.

References

[1] Ramalingam, Archana, Sharat Kedari, and Chandrasekar Vuppalapati, "IEEE FEMH voice data challenge 2018," in 2018 IEEE International Conference on Big Data (Big Data). IEEE, pp. 5272-5276, 2018.

[2] Islam, Kazi Aminul, Daniel Perez, and Jiang Li, "A transfer learning approach for the 2018 FEMH voice data challenge." in 2018 IEEE International Conference on Big Data (Big Data). IEEE, pp. 5252-5257, 2018.

[3] Chuang, Zong-Ying, et al., "Dnn-based approach to detect and classify pathological voice," in 2018 IEEE international conference on big data (big data). IEEE, pp. 5238-5241, 2018.

[4] Pham, Minh, Jing Lin, and Yanjia Zhang, "Diagnosing voice disorder with machine learning," in 2018 IEEE International Conference on Big Data (Big Data). IEEE, pp. 5263-5266, 2018.

[5] Pishgar, Maryam, et al., "Pathological voice classification using melcepstrum vectors and support vector machine," arXiv preprint arXiv:1812.07729, 2018.

[6] Miliaresi, Ioanna, Kyriakos Poutos, and Aggelos Pikrakis, "Combining acoustic features and medical data in deep learning networks for voice pathology classification," in 2020 28th European Signal Processing Conference (EUSIPCO). IEEE, pp.1190-1194, 2021.