iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

RESPIRATORY SOUND CLASSIFICATION USING MACHINE LEARNING

Dheeraj KJ, Meghana R, Devika N, Syeda Mohammedi Kauser Azmi
Students, Department of Information Science & Engineering, Bangalore Institute of Technology, Bengaluru-560004
Prof. Prameela R,
Assistant Professor, Department of Information Science & Engineering, Bangalore Institute of Technology, Bengaluru-560004.

ABSTRACT
Respiratory diseases are a major global health challenge, ranking among the leading cause for mortality and disability worldwide. Conditions such as chronic obstructive pulmonary disease (COPD), asthma, tuberculosis, and lung cancer contribute significantly to global mortality, with COPD alone responsible for an estimated 3.91 million deaths annually. The burden is particularly severe in low-income regions with limited healthcare infrastructure, where socioeconomic factors and a shortage of medical professionals further hinder timely and accurate diagnoses. To tackle these issues, innovations in Artificial Intelligence provide effective strategies to minimize misdiagnoses and facilitate precise treatments. A cost-efficient and user-friendly algorithm was created to analyze respiratory sounds, ensuring compatibility with various devices. The methodology integrates machine learning approaches, utilizing Gammatone Cepstral Coefficients (GTCC) within a Convolutional Neural Network (CNN) and a CNN-LSTM hybrid model incorporating GTCC and STFC features. Four datasets were prepared for classifying respiratory audio, covering healthy versus pathological classification, rale/rhonchus/normal sound classification, singular respiratory sound type classification, and audio type classification with all sound types.

Keywords:
Respiratory sounds, Machine Learning, health challenges, lungs

I. INTRODUCTION

Respiratory illnesses are a primary cause of mortality and disability around the globe, particularly in the most disadvantaged areas. Essential risk factors include aging, tobacco use, environmental pollution, and body weight. Long-term respiratory disorders, such as Chronic Obstructive Pulmonary Disease (COPD) and asthma, present considerable public health challenges. In 2017, COPD was responsible for roughly 3.91 million fatalities, ranking as the third most prevalent cause of death worldwide. Asthma, the most widespread chronic disease in children, affects nearly 334 million people worldwide. Other critical respiratory ailments include pneumonia, tuberculosis (TB), and lung cancer. Pneumonia ranks among the leading causes of death in children under five, whereas tuberculosis (TB) is the most lethal infectious disease, responsible for 1.4 million fatalities each year. Lung cancer, the most fatal form of cancer, leads to 1.6 million deaths annually. Collectively, respiratory diseases significantly affect societal, economic, and healthcare systems. Economic hardship plays a crucial role in increasing mortality and disability rates, particularly in the most impoverished areas. Addressing lung diseases is critical in the healthcare industry, driving extensive research to enable early detection and timely treatment. As stated by the World Health Organization, numerous nations indicate fewer than one physician for every 1,000 individuals, falling short of the recommended ratio. This shortage places significant pressure on healthcare staff and increases the likelihood of errors. Thus, the development of reliable and automated diagnostic tools is essential to assist medical practitioners, save time, reduce errors.

iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

II. LITERATURE SURVEY

[1] Automatic Lung Sound Classification Using CNN. Authors: Ahmed A., Javed K., and Kaur S., March 2020

This research recommended utilizing Convolutional Neural Networks (CNN) for categorizing sounds of lungs such as wheezes and crackles. The original audio data was converted into spectrograms, which served as inputs for the CNN model. Data augmentation methods were employed to increase the dataset's diversity. CNNs exhibited excellent accuracy in identifying spectrogram patterns and proved resilient to slight data variations. The use of spectrograms enhanced feature representation. However, this approach required extensive preprocessing and a well-annotated dataset. Furthermore, the model's performance diminished when dealing with noisy audio samples or overlapping sounds.

[2] Classification of Adventitious Lung Sounds: Wheeze, Crackle Using Machine Learning Techniques. Authors: John Amose, Manimegalai P., Pon Selchiya R. (Date of Publication: July 8th, 2023)

SVM and CNN are frequently used and have shown high accuracy in the classification of lung sounds, demonstrating their effectiveness in distinguishing between different respiratory conditions. Models like SVM and CNN have demonstrated high accuracy and robustness in various studies, making them reliable tools for lung sound classification. The effectiveness of machine learning models is largely influenced by the quality and variety of the training dataset.

[3] A Novel Lightweight Inception Network for Respiratory Disease Classification Using Lung Sounds. Author: Arka Roy, Udit Satija, IEEE Vol. 72, 2023

This study introduces a novel lightweight inception network (RDLINet) for classifying lung sounds into seven respiratory disease categories. The methodology involves preprocessing lung sound signals (e.g., resampling, normalization) and converting them into mel spectrograms, which are subsequently categorized using the RDLINet architecture. The network employs multiscale filtering lightweight inception (MFLI) blocks, designed for resource-constrained devices, achieving an accuracy of 96.6% for seven-class classification and 94.0% for binary classification (healthy vs. asthma). Advantages include a lightweight architecture optimized for edge devices, high classification accuracy, and scalability for real-time clinical use. However, limitations include a reliance on annotated databases, challenges with generalization to unseen datasets, and the need for compatible microcontrollers for deployment in resource-constrained environments.

[4] Feature Extraction Using MFCC for Respiratory Disease Diagnosis. Authors: Patel R., Kumar N., and Sharma A., August 2018

This study concentrated on employing Mel-Frequency Cepstral Coefficients (MFCCs) for attribute extraction. MFCCs were computed for each lung sound recording, and the resulting attribute set served as input for classification models. This approach highlighted the potential of MFCCs in accurately distinguishing respiratory sounds, contributing to improved diagnostic systems.

III. EXISTING SYSTEM

The existing system for classifying respiratory sounds relies primarily on manual auscultation performed by healthcare professionals using stethoscopes. Traditional signal processing techniques are sometimes used to analyze audio sound recordings in order to detect characteristics like crackles and wheezes. These systems are frequently restricted to basic attribute extraction and categorization techniques, rendering them reliant on human skill. Furthermore, traditional machine learning models such as Support Vector Machines (SVMs) or Random Forests are occasionally utilized, but they necessitate considerable manual attribute extraction.

iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

Advantages:

1. The system automates the classification of respiratory sounds, reducing dependency on manual auscultation by healthcare professionals.
2. Advanced machine learning models like CNNs ensure high accuracy in identifying patterns, even in noisy datasets.
3. It is scalable and suitable for real-time applications, making it highly beneficial for large-scale or remote healthcare systems.

Disadvantages:

1. The system relies on high-quality annotated datasets, and performance can degrade with insufficient or poorly labeled data.
2. Computationally demanding models such as CNNs necessitate considerable processing capability, which may not be practical for resource-limited settings.
3. High Latency in Real-Time Noise in audio data can still affect performance, requiring robust preprocessing and augmentation techniques.

IV. PROBLEM STATEMENT

The manual auscultation and basic machine learning techniques used in the traditional methods of diagnosing respiratory diseases are prone to errors and inefficiencies. They rely heavily on expert intervention and often fail to analyze complex audio patterns from respiratory sounds accurately. Additionally, existing techniques lack scalability, making them unsuitable for real-time applications or handling large datasets with varying levels of noise. To address these limitations, an automated and scalable solution utilizing sophisticated machine learning techniques is required to improve diagnostic accuracy, reliability and efficiency.

V. PROPOSED SYSTEM

The suggested system intends to classify respiratory sounds recordings such as wheezes and crackles using advanced machine learning algorithms. The system employs attribute extraction methods such as Mel-Frequency Cepstral Coefficients (MFCCs) and transforms audio recordings into spectrograms for enhanced feature representation. CNN are used for automated feature learning and classification. The system integrates noise reduction and augmentation techniques for improved model performance on diverse datasets. It is designed to provide a scalable, real-time, and automated diagnostic solution to overcome the limitations of manual and traditional ML approaches. Additionally, the project will explore the potential for real-time deployment in healthcare settings, making it a practical and accessible tool for enhancing the accuracy and efficiency of respiratory disease diagnosis worldwide.

iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

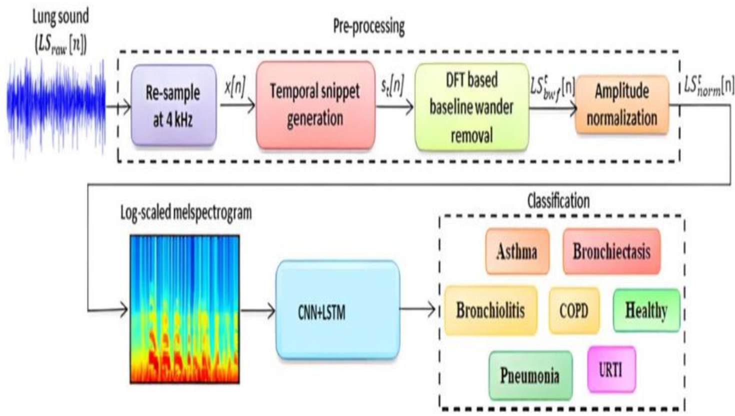

Figure 5.1 System Architecture

This figure illustrates the workflow for detecting and characterizing lung diseases using respiratory sound recordings. The process begins with capturing respiratory sounds, followed by feature extraction from these recordings. Before being used to detect lung diseases, the extracted data is preprocessed. If a disease is detected, the workflow proceeds to characterize the lung disease and further specifies the type of disease. If no disease is detected, the patient is classified as healthy.

The workflow concludes with the completion of disease characterization or confirmation of a healthy patient status. The figure outlines a diagnostic process for pulmonary diseases that utilizes electronic stethoscopes and machine learning. Initially, respiratory sounds from patients with pulmonary diseases are recorded using electronic stethoscopes. These recordings are collected in clinical settings, where pulmonologists classify them into categories such as normal, crackles, wheezes, and rhonchi. The classified sounds are then converted into Mel-spectrograms for feature extraction.

Datasets:
The respiratory sound data originates from lung airflow during breathing, collected using high-quality stethoscope systems such as the 3M Littmann Electronic Stethoscope and Meditron Master Elite. These sounds are recorded from six specific chest positions and include both normal and abnormal respiratory patterns, such as crackles and wheezes. The dataset utilized is the ICBHI (International Conference on Biomedical and Health Informatics) database, which comprises 920 audio tracks from 126 participants, making it one of the largest annotated collections of its kind. This dataset provides crucial annotations for respiratory cycles, identifying normal and problematic sounds and their combinations, which form the basis for building robust diagnostic tools.

Annotations:
Annotations in the dataset classify each respiratory cycle into four categories: normal, crackles, wheezes, and both crackles and wheezes. Of the 6898 cycles recorded, approximately 47% are healthy, while the rest display abnormal sound characteristics. This comprehensive labeling offers a standard for machine learning algorithms to distinguish between respiratory conditions and identify the severity of abnormalities through specific sound frequencies ranging from 100 Hz to 2500 Hz.
Challenges in Data Collection Recording respiratory sounds posed significant challenges, including the presence of noise such as stethoscope rubbing and background chatter, as well as the interference of

iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

heartbeats. These issues required meticulous preprocessing, including noise filtering and feature enhancement, to ensure the reliability of diagnostic models. Additionally, variability in patient conditions and recording environments further complicated data uniformity.

Preprocessing Techniques To enhance data usability, respiratory sound recordings were broken into smaller frames to facilitate analysis. Windowing functions, such as the Hamming window, were utilized to minimize spectral leakage during frequency range domain analysis. This preprocessing ensures that critical sound features remain intact while mitigating the impact of noise and other artifacts, ultimately improving the quality of input data for machine learning models.

Feature Extraction:

Linear Predictive Coefficients (LPC) and other important audio characteristics were taken out of every frame. By providing a mathematical depiction of the signal's behavior, these coefficients make it possible to spot patterns particular to particular respiratory disorders. A standardized dataset for model training and testing was then produced by normalizing and grading the features to ensure consistency across different clip lengths.

Classification Techniques:

The chapter details the use of Multilayer Perceptrons (MLPs) and feed-forward neural networks for classifying respiratory sounds. These models are structured with input, hidden, and output layers, leveraging activation functions to map input data to specific diagnoses. By employing algorithms like the MLP, the study demonstrates how non-linear relationships between features can be effectively captured to improve classification accuracy.

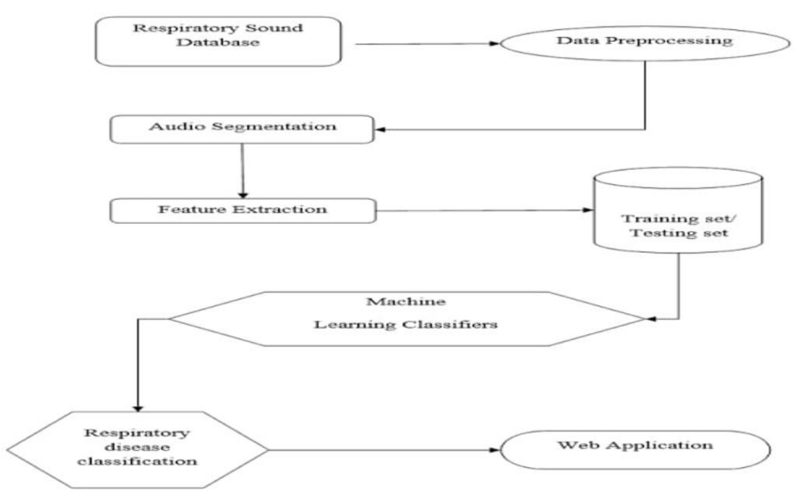

Figure 5.2 High Level Design

VI. ALGORITHMS

1. Mel-Frequency Cepstral Coefficients (MFCCs)

MFCC is a feature extraction technique designed to mimic the human auditory system by converting audio signals into the frequency domain. It filters out noise and unnecessary data while capturing important frequency components. Short frames make up the audio signal. The frequency spectrum is computed using a Fourier Transform. The Mel filter bank prioritizes frequencies that are audible to humans. Compact feature vectors are created using a discrete cosine transform (DCT) and logarithmic compression. Wheezes and

IJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

crackles are examples of respiratory sounds from which MFCCs are used to extract important features, lowering dimensionality while maintaining significant patterns.

2. Convolutional Neural Networks (CNNs)

CNN is a deep learning algorithm designed for image and grid-based data processing. It uses convolutional layers to automatically detect spatial patterns, making it ideal for processing spectrograms derived from audio signals. Convolutional layers extract local patterns using filters or kernels. Pooling layers reduce dimensionality, retaining essential features. Fully connected layers map learned features to output classes (e.g., wheezes, crackles). They eliminate the need for manual feature engineering by learning patterns directly from the data.

Data Augmentation Methods

Data augmentation involves artificially increasing the training dataset's size and diversity by applying transformations such as noise addition, time shifts, and pitch variations. Noise injection adds random background noise to mimic real-world conditions. Time shifting moves audio signals slightly forward or backward. Pitch shifting alters the frequency of the audio signal while retaining its essence. These techniques are used to make the model robust to variations in audio data, improving generalization to unseen datasets.

VII. RESULTS

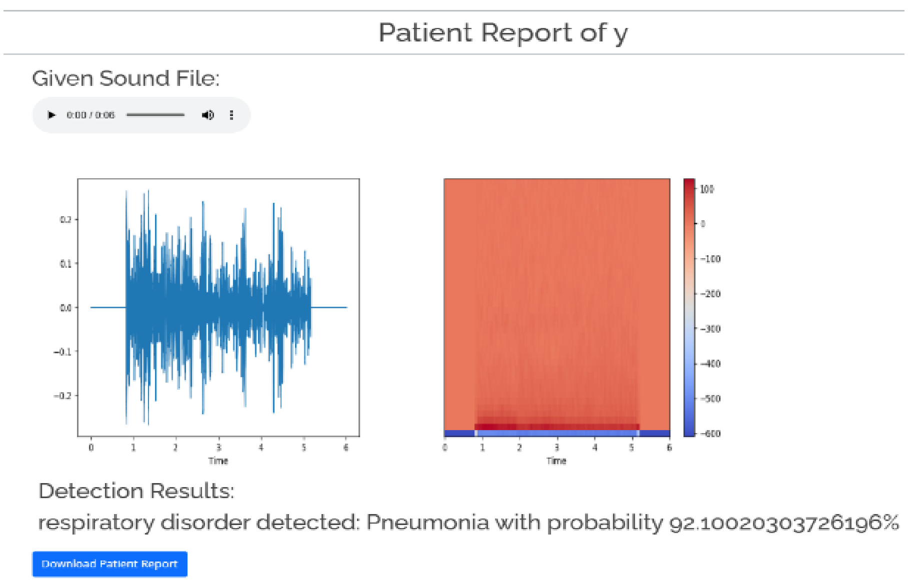

Figure 7.1 Result on desktop application

The interface of your Respiratory Sound Classification System is designed to enable seamless interaction for classifying respiratory sounds (e.g., wheezes, crackles, or normal sounds) using machine learning. The interface is divided into two main sections: Audio Upload & Classification and Result Display.

Audio Upload & Classification Section:

In this section, users can upload an audio file or record a new respiratory sound using a microphone. Once the audio is selected, the system processes it to extract features (such as MFCCs) and performs classification using CNN.

iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

Upload or Record Audio:
Users will click a button to upload a pre-recorded audio file of respiratory sounds (e.g., wheeze.wav, crackle.wav)

Preprocessing & Feature Extraction:
After the audio file is uploaded or recorded, the system applies preprocessing steps such as noise reduction, followed by MFCC feature extraction. These features are important for feeding the data into the convolutional neural network model for classification.

Classification:
The extracted features are processed by the trained CNN model, which classifies the respiratory sound as either Wheeze, Crackle, or Normal. The classification is based on the patterns the model learned during training, allowing it to recognize different types of respiratory sounds.

Result Display Section:
Once the classification is completed, the outcomes are shown in an intuitive interface that includes both textual and graphical representations. Class Prediction: The system shows the predicted class (e.g., Wheeze, Crackle, or Normal) along with the confidence score (e.g., 85%, indicating the model's certainty about the classification).

Spectrogram Visualization:
A graphical depiction of the spectrogram of the audio is presented. This enables users to observe how the audio information was converted into a visual format, aiding them in understanding the patterns that the model identified.

Graphical Output:
The outcomes are also represented graphically, showing a bar chart or pie chart displaying the percentage confidence for each possible class. This gives users a clear view of how confident the system is in its predictions.

Save or Export Results:
The user has the option to download the results as a PDF for further analysis or reporting.

VIII. CONCLUSION

“Respiratory Sound Classification Using Machine Learning” demonstrates the power of machine learning algorithms, particularly CNN, in identifying and classifying respiratory sounds such as wheezes, crackles, and normal sounds. The system leverages advanced feature extraction techniques, including Mel Frequency Cepstral Coefficients (MFCCs), to transform raw audio data into usable features for classification. By utilizing a CNN model, the project successfully classifies audio files into respective categories with high accuracy, providing a reliable tool for the early detection of respiratory ailments.

This system possesses considerable promise in healthcare applications, especially for distant observation of individuals with persistent respiratory illnesses like asthma and COPD. The ability to automatically detect and classify abnormal respiratory sounds can significantly reduce diagnostic time, allowing healthcare professionals to respond more promptly to critical cases. Moreover, it offers a non-invasive and cost-effective approach for monitoring respiratory health. The project’s modularity also enables easy adaptation to various healthcare environments, making it scalable and flexible for future applications.

Overall, the project not only contributes to the healthcare sector by improving diagnostic efficiency but also serves as a foundation for further exploration in the domain of medical sound classification using machine learning.

iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

IX. FUTURE ENHANCEMENT

While the current implementation of the respiratory sound classification system demonstrates promising results, there are several potential enhancements that can further improve its accuracy, robustness, and applicability in real-world settings.

One important future enhancement would be the incorporation of more advanced deep learning frameworks, such as Recurrent Neural Networks (RNNs) or Long Short-Term Memory (LSTM) networks, which are particularly suited for sequential data like audio. This could potentially improve classification accuracy, especially for complex respiratory sounds that may have subtle characteristics. Additionally, integrating data augmentation techniques to simulate various environmental noise conditions would further enhance the model's robustness to real-world scenarios.

Another area for improvement is the inclusion of a broader dataset that includes a diverse range of respiratory sounds from different populations, including various age groups, genders, and ethnic backgrounds. This would guarantee that the model can handle a broad range of sound profiles and enhance its generalization ability.

Finally, implementing a real-time prediction system would significantly increase the practical value of the project, enabling its use in healthcare settings where timely diagnosis and monitoring are critical. This could be achieved by optimizing the processing pipeline to reduce latency and integrating the system with cloud-based healthcare platforms for remote monitoring of patients.

In conclusion, future advancements in algorithm development, data diversity, and real-time capabilities will enhance the project's overall impact and utility in the healthcare industry.

X. REFERENCES

[1] A. Ahmed, K. Javed, and S. Kaur, "Automatic Lung Sound Classification Using CNN," IEEE Transactions on Biomedical Engineering, March 2020.

[2] J. Amose, M. P., and P. S. R., "Classification of Adventitious Lung Sounds: Wheeze, Crackle using Machine Learning Techniques," ISSN: 2147-6799, July 8, 2023.

[3] A. Roy and U. Satija, "RDLINet: A Novel Lightweight Inception Network for Respiratory Disease Classification Using Lung Sounds," IEEE Transactions on Instrumentation and Measurement, vol. 72, 2023

[4] R. Patel, N. Kumar, and A. Sharma, "Feature Extraction Using MFCC for Respiratory Disease Diagnosis," Springer Journal of Signal Processing Systems, August 2018.

[5] S. K. Gupta, A. Kumar, and N. Singh, "Deep learning techniques for medical sound analysis," J. Med. Syst., vol. 43, no. 6, pp. 220-225, 2019.

[6] S. R. Gubbi and S. S. Devaraj, "A novel approach to classify respiratory sounds using deep convolutional neural networks," J. Healthcare Eng., vol. 2018, pp. 1-9, 2018.

[7] J. Xie, L. Zhang, and F. Jiang, "Deep learning for respiratory sound classification and its applications in healthcare," Sensors, vol. 19, no. 5, pp. 1056, 2019.

[8] P. Kumar, "Applications of machine learning in medical diagnostics," Int. J. Data Sci. Anal., vol. 6, no. 3, pp. 205-215, 2020.

iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

[9] S. Jayalakshmy, S. Rithika, and S. Rajasri, "Average Power Based Classification of Respiratory Sound Using SVM Classifier," IEEE Transactions on Biomedical Engineering, Aug. 2020.

[10] J. Amose, M. P., and P. S. R., "Classification of Adventitious Lung Sounds: Wheeze, Crackle Using Machine Learning Techniques," July 2023.

[11] J. P. Garcia-Mendez and A. Lal, "Machine Learning Automated Classification of Abnormal Lung Sounds Obtained from Public Databases: A Systematic Review," IEEE Transactions on Artificial Intelligence, Oct. 2023.

[12] B. S., K. P. N., and M. K., "Respiratory Lung Disease Classification Using Machine Learning Techniques," 2020.