

Classification of Respiratory Disease Using Lung Sound

Mohd Anas Shaikh¹, Mohammad Ali Shaikh², Mohammad Asim³, Saniya Chandiwala⁴, Irfan Jamkhandikar⁵

Department of Computer Engineering
Anjuman-I-Islam's Kalsekar Technical Campus
Navi Mumbai, India

ABSTRACT

Respiratory diseases such as asthma, chronic obstructive pulmonary disease, pneumonia, and bronchitis are major causes of morbidity and mortality worldwide. Early and accurate diagnosis plays a crucial role in effective treatment and management. This paper presents a machine learning based approach for the classification of respiratory diseases using lung sound recordings. Lung sounds are collected through electronic stethoscopes and preprocessed to remove background noise and artifacts. Relevant acoustic features are extracted from the filtered signals and used to train multiple classification models. The performance of different algorithms is evaluated using standard metrics such as accuracy, precision, recall, and F1 score. Experimental results demonstrate that the proposed method can effectively distinguish between normal and abnormal lung sounds, as well as classify specific respiratory conditions with high reliability. The study highlights the potential of computer aided diagnostic systems to support clinicians in real time respiratory disease screening, especially in resource limited settings.

Keywords:—respiratory disease classification, lung sound analysis, machine learning, feature extraction, computer aided diagnosis.

I. INTRODUCTION

Respiratory diseases like asthma, COPD, pneumonia, and bronchitis affect millions worldwide and require early diagnosis for effective treatment. Traditional methods such as X-rays, CT scans, and manual auscultation can be time-consuming, subjective, and inaccessible in many areas. AI-based analysis of lung sounds offers a faster, more objective, and cost-effective approach, supporting doctors and improving patient care, especially in resource-limited settings

Motivation: With the rapid rise in respiratory diseases caused by increasing air pollution, smoking habits, occupational hazards, and post-COVID respiratory complications, there is a growing demand for efficient and affordable healthcare solutions. Traditional diagnostic methods such as manual auscultation require significant medical expertise and may sometimes lead to subjective interpretations

Therefore, Identify applicable funding agency here. If none, delete this. developing an automated and intelligent system that can analyze lung sounds and assist medical professionals in identifying respiratory abnormalities at an early stage has become extremely important. Such

systems can improve diagnostic accuracy, reduce workload of healthcare professionals, and provide accessible healthcare support in rural and resource-limited areas.

Build an intelligent and automated system capable of detecting and classifying respiratory diseases by analyzing lung sound recordings collected through digital stethoscopes or audio sensors.

Develop a machine learning and deep learning based framework that can efficiently extract meaningful audio features from lung sound signals and accurately identify abnormal respiratory patterns.

Improve the speed, reliability, and consistency of respiratory disease diagnosis by reducing dependence on manual auscultation and subjective clinical judgment.

Assist doctors and healthcare professionals with a decision-support tool that can provide preliminary screening results for conditions such as COPD, pneumonia, bronchiectasis, bronchiolitis, and upper respiratory tract infections.

Promote early detection and timely medical intervention by providing an accessible artificial intelligence based diagnostic support system that can be used in hospitals, clinics, and remote healthcare environments

II. LITERATURE SURVEY

Survey of Existing Solutions: Deep learning approaches have been widely explored for lung sound classification, including CNNs on spectrograms, hybrid CNN+RNN models, and feature fusion methods. Most studies utilize publicly available datasets, such as the ICBHI 2017 Respiratory Sound Database, while some rely on smaller private hospital datasets. Techniques like Mixture of Experts, Transfer Learning, and Artificial Noise Addition have been investigated to improve classification performance

Analysis of Issues in Existing Solutions: While deep learning models achieve promising results, many face challenges when deployed in diverse real-world environments. Common issues include reduced performance with varying recording devices, background noise, and differences in patient populations. Additionally, existing systems often lack lightweight designs suitable for real-time applications or mobile integration.

Survey of existing solution w.r.t. users’ requirements and improvements expected: Healthcare professionals expect reliable, explainable, and clinically validated tools, not just benchmark accuracy. Current systems often break down in noisy outpatient or ICU environments. Lightweight and mobile-friendly models are needed for integration with digital stethoscopes. Improvements should focus on usability, scalability, and trustworthiness in real healthcare contexts. Survey of Technologies Methodologies: Lung sound classification uses features like STFT, MFCC, Mel-spectrograms, Gamma-tone filters, and wavelet transforms, often converted to visual forms for deep learning. Models include CNNs, Depthwise Separable CNNs, ResNet, VGG16, hybrid CNN-RNNs, and lightweight networks. Advanced techniques like Teacher–Student frameworks, Mixture of Experts, ensemble fusion, and multi-task learning improve accuracy and reliability.

Gap Identification: Current lung sound classification models often perform well in controlled experiments but struggle in real-world scenarios. Most are trained on limited datasets, making it hard to handle variations in recording devices, patient demographics, and hospital environments. Few models are designed for mobile or low-power devices, and many lack clear explanations, which reduces their usefulness for doctors in daily practice

General Observations: AI techniques show strong potential for detecting respiratory diseases from lung sounds, achieving good accuracy in research studies. However, clinical adoption is limited, and models often fail in noisy or complex environments. There is a need for systems that are reliable, interpretable, and practical, supporting doctors in real-time decision-making and everyday healthcare settings.

LIMITATIONS OF EXISTING SYSTEM

- I. **Advantages / Results Achieved:** Deep learning models like CNNs, VGG16, ResNet, Inception, and BiLSTM have achieved high accuracy (85–99- Feature fusion (STFT + MFCC) and frameworks like Mixture of Experts (MoE) improve classification performance.
- II. Some studies developed lightweight or mobile-friendly models without significant loss of accuracy.
- III. Certain works have shown performance comparable or superior to medical trainees, demonstrating potential clinical utility.

Comparison of Existing vs Proposed System

Parameter	Existing Systems	Proposed System
Accuracy	High (85–99%)	87.90% with feature fusion
Real-world Performance	Poor in noise	Improved robustness
Feature Usage	Limited features	MFCC + Chroma + Mel
Deployment	Research-based	Practical use targeted
Explainability	Limited	Moderate
Scalability	Limited	Future scalable

Limitations:

1. Models are mostly tested on limited or specific datasets, restricting generalization.

2. Performance in noisy clinical environments is largely untested.
3. Few systems focus on lightweight, real-time deployment suitable for mobile/IoT devices.
4. Explainability and clinical trust are rarely addressed, reducing acceptance in healthcare.

III. SYSTEM DESIGN AND ARCHITECTURE

Dataset Inputs: The system utilizes raw respiratory audio recordings in .wav format along with corresponding clinical annotation files in .txt format. These annotation files contain precise timestamp information that marks the start and end of each respiratory cycle, including inhalation and exhalation phases. This pairing of audio data with clinical annotations ensures that the model receives accurately labeled respiratory events, which improves the reliability of the dataset and supports effective supervised learning

Audio Pre-processing: The raw audio signals undergo a comprehensive preprocessing stage before being provided to the model. Continuous recordings are first segmented into individual respiratory cycles using the timestamps from the annotation files. To maintain consistency across the dataset, padding and truncation are applied so that each audio sample maintains a uniform input length and fixed input shape. This step ensures compatibility with deep learning models and prevents dimensional mismatches during train

Parallel Feature Extraction: In order to capture diverse characteristics of respiratory sounds, the system performs parallel extraction of multiple audio features. The extracted features include Mel Frequency Cepstral Coefficients (MFCC), Chroma Short-Time Fourier Transform (Chroma STFT), and Mel Spectrogram. MFCC features represent perceptual sound characteristics similar to human auditory perception, Chroma STFT captures tonal and harmonic properties, and Mel Spectrograms provide a detailed time–frequency representation of the audio signal. Extracting these complementary features improves the model’s ability to identify subtle acoustic patterns.

Feature Concatenation: After extracting the different audio features, they are combined using a feature concatenation process. The MFCC, Chroma STFT, and Mel Spectrogram feature vectors are merged into a single unified feature vector. This fused representation integrates spectral, tonal, and perceptual information into a consolidated input format, enabling the model to learn from richer and more informative data.

Neural Network Classifier: The classification stage is implemented using a deep learning based neural network built with TensorFlow/Keras. A Sequential architecture is used to process the fused feature vectors and learn meaningful patterns associated with respiratory conditions. The network consists of multiple dense layers, activation functions, and optimization mechanisms that allow the model to effectively capture complex relationships between extracted features and disease labels.

Multi-Class Output: The final layer of the model performs multi-class classification, enabling the system to identify specific respiratory diseases. The trained model classifies respiratory sounds into categories such as Chronic Obstructive Pulmonary Disease (COPD), Pneumonia, Upper Respiratory Tract Infection (URTI), Bronchiectasis, Bronchiolitis, and Healthy.

Feature Extraction Techniques:-

Feature	Purpose	Benefit
MFCC	Captures spectral envelope	Good for audio recognition
Chroma STFT	Pitch and harmonic analysis	Identifies tonal patterns
Mel Spectrogram	Frequency on Mel scale	Mimics human hearing

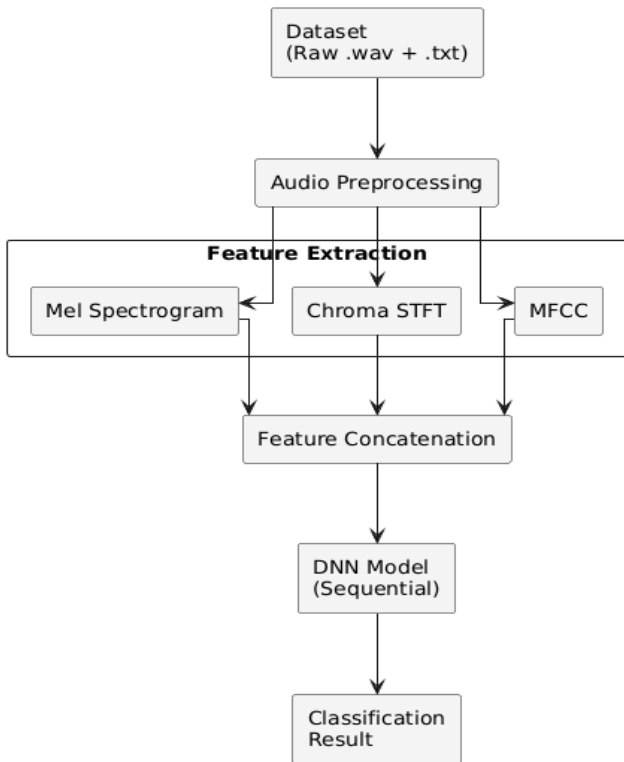


Fig. 1. Proposed system architecture for respiratory disease classification

IV. PROPOSED METHODOLOGY

Audio Preprocessing: Utilizing the ICBHI 2017 dataset raw recordings are segmented into individual respiratory cycles using clinical timestamps. These segments are padded and normalized via Librosa to ensure uniform input for the model.

Feature Extraction: Instead of a single feature, a parallel extraction method is used to capture a complete audio profile.

1. MFCC: Extracts the spectral envelope
2. Chroma STFT: Analyzes harmonic and pitch content.
3. mSpec: Provides frequency-domain data on the Mel scale.

Deep Neural Network (DNN) Classification: The fused features are fed into a Sequential Neural Network. The architecture uses multiple Dense layers and Dropout to accurately classify sounds into 6 categories: COPD, Pneumonia, URTI, Bronchiectasis, Bronchiolitis, and Healthy.

V. RESULT:-

- The proposed system successfully demonstrates the transition from a theoretical concept to a practical and functional tool capable of identifying respiratory diseases with an accuracy of approximately 87.90%. By integrating multiple audio feature extraction techniques such as MFCC, Chroma STFT, and Mel Spectrograms, the system captures a wide range of acoustic characteristics present in lung sound signals.
- This multi-feature approach proves to be significantly more effective than relying on a single feature extraction technique, as it provides a richer and more informative representation of respiratory sounds.
- The developed Deep Learning model is capable of learning complex patterns from the combined feature set and accurately distinguishing between various respiratory conditions. In addition to achieving high classification accuracy, the model also demonstrates fast prediction speed, making it suitable for real-time or near real-time clinical applications. This efficiency is particularly valuable in healthcare environments where quick decision making can greatly assist doctors and medical practitioners.
- Furthermore, the proposed system highlights the potential of artificial intelligence in supporting medical diagnosis, especially in situations where expert medical resources may be limited. By providing an automated and objective analysis of lung sounds, the system can serve as a helpful decision support tool for identifying respiratory diseases such as COPD, Pneumonia, Bronchiectasis, Bronchiolitis, and other respiratory abnormalities. Overall, this research contributes to the development of intelligent healthcare technologies that aim to improve early detection, enhance diagnostic accuracy, and ultimately support better patient care and treatment planning.

REFERENCES :-

- [1]. G. Altan and G. Gur, 'CNN-MOE Based Framework for Classification of Respiratory Anomalies and Lung Disease Detection,' *Applied Acoustics*, vol. 187, p. 108502, 2022.
- [2]. II. Y. Ma, X. Xu, and Q. Yu, 'Efficiently Classifying Lung Sounds Through Depthwise Separable CNN Models With Fused STFT and MFCC Features,' *IEEE Access*, vol. 9, pp. 39774–39783, 2021.
- [3]. III. P. K. Jana, G. Saha, and D. K. Kole, 'Abnormal Respiratory Sounds Classification Using Deep CNN Through Artificial Noise Addition,' *Biomedical Signal Processing and Control*, vol. 68, p. 102573, 2021
- [4]. D. Perna and A. Tagarelli, 'Robust Respiratory Disease Classification Using Breathing Sounds (RRDCBS),' *Proc. IEEE Int. Conf. Bioinformatics and Biomedicine (BIBM)*, 2020, pp. 1478–1485
- [5]. V. J. Acharya and A. Basu, 'Respiratory Sound Classification for Crackles, Wheezes, and Rhonchi in the Clinical Field Using Deep Learning,' *IEEE Transactions on Biomedical Engineering*, vol. 67, no. 10, pp. 2894–2903, 2020.
- [6]. VI. M. Aykanat, O. Kilic, B. Kurt, and S. Sariyal, 'Recognition of Pulmonary Diseases from Lung Sounds Using CNN and LSTM,' *Proc. Signal Processing and Communications Applications Conference (SIU)*, 2017.
- [7]. VII. L. Brunese et al., 'A Neural Network-Based Method for Respiratory Sound Analysis and Lung Disease Detection,' *Applied Sciences*, vol. 10, no. 8, p. 3877, 2022
- [8]. VIII. L. Fraiwan et al., 'Automatic Identification of Respiratory Diseases from Lung Sound Signals Using Ensemble Classifiers,' 2020.
- [9]. IX. L. Fraiwan et al., 'Automatic Identification of Respiratory Diseases from Lung Sound Signals Using Ensemble Classifiers,' *Biocybernetics and Biomedical Engineering*, vol. 41, no. 1, pp. 1–14, 2021.
- [10]. X. Y. Kim et al., 'Respiratory Sound Classification for Crackles, Wheezes, and Rhonchi in the Clinical Field Using Deep Learning,' *Scientific Reports*, vol. 11, no. 1, 2021.
- [11]. XI. A. Sabry et al., 'Lung Disease Recognition Methods Using Audio-Based Analysis with Machine Learning,' 2024.
- [12]. XII. R. Zulfikar et al., 'Abnormal Respiratory Sounds Classification Using Deep CNN through Artificial Noise Addition,' *Frontiers in Medicine*, vol. 8, 2021.

c