International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 01 January 2026
Page No: 349-354
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0048

Lung Sound Classification for Respiratory Disease

Arpitha H M[1], Muralitharan G[2], Deepika K S[3], Lavanya J[4]
[1,3,4]UG Scholar, Dept. of CSE, AMC Engineering College., Bangalore, Karnataka, India.
[2]Assistant Professor, Dept. of CSE, AMC Engineering College., Bangalore, Karnataka, India
Emails: *arpithamallesh17@gmail.com[1], rye.murali25@gmail.com[2], deepikaks349@gmail.com[3], lavanyalavanya9016@gmail.com[4]*

Abstract
Respiratory illnesses are still among the biggest health problems around the world, and catching them early is key to avoiding serious issues. Listening to lung sounds is a widely used method for identifying respiratory diseases. However, manual examination depends heavily on a clinician's experience, may differ from one doctor to another, and can be unreliable in noisy surroundings. To overcome these challenges, this work presents an automated lung sound classification system based on deep learning techniques. Important acoustic features are extracted using Mel-Frequency Cepstral Coefficients (MFCCs), which are then analyzed using a hybrid CNN-LSTM model to categorize sounds into six classes: COPD, Pneumonia, Bronchiectasis, Bronchiolitis, Upper Respiratory Tract Infection (URTI), and Normal. The model was trained, tested, and optimized to ensure reliable performance across varied conditions. Additionally, Grad-CAM is integrated to highlight the sound regions that influence the model's predictions, improving transparency and interpretability. The proposed system offers accurate results with a user-friendly interface, making it suitable for real-time clinical use. Overall, this approach supports early detection of respiratory illnesses and has strong potential for deployment in regions with limited medical resources.
Keywords: *Lung sound analysis, Mel-frequency cepstral features, CNN–LSTM hybrid model, Deep neural networks, Automated respiratory disease detection, Gradient-weighted class activation mapping, Medical signal processing, Remote healthcare systems.*

1. Introduction
Respiratory disorders are among the leading causes of health complications worldwide, making timely and accurate diagnosis essential. Traditionally, clinicians assess lung conditions by listening to breath sounds using a stethoscope. However, this approach can be affected by environmental noise, subjective judgment, and variations in clinical expertise. With recent advances in artificial intelligence, automated systems have gained attention for their ability to analyze lung sounds in a more consistent and objective manner. Different respiratory diseases such as COPD, Bronchiectasis, and Pneumonia produce unique acoustic patterns in lung sounds. Identifying these patterns requires detailed analysis across both time and frequency domains. Mel-Frequency Cepstral Coefficients (MFCCs) are effective in capturing these characteristics. Furthermore, deep learning architectures that combine Convolutional Neural Networks (CNNs) with Long Short-Term Memory (LSTM) networks have shown strong performance in audio classification tasks, making them well suited for respiratory sound analysis. This study is about creating a smart system that classifies lung sounds using MFCC and a CNN- LSTM setup. The goal is to help with accurate diagnosis and also to make the system easier to understand through visual tools like Grad-CAM [1], [2].

2. Methodology
The methodology followed in this study consists of multiple stages, beginning with the acquisition of respiratory sound recordings and followed by preprocessing, feature extraction, and classification. The overall flow of the proposed system is presented

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 01 January 2026
Page No: 349-354
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0048

in Figure1 System Architecture, which outlines each processing component from input audio to final prediction.

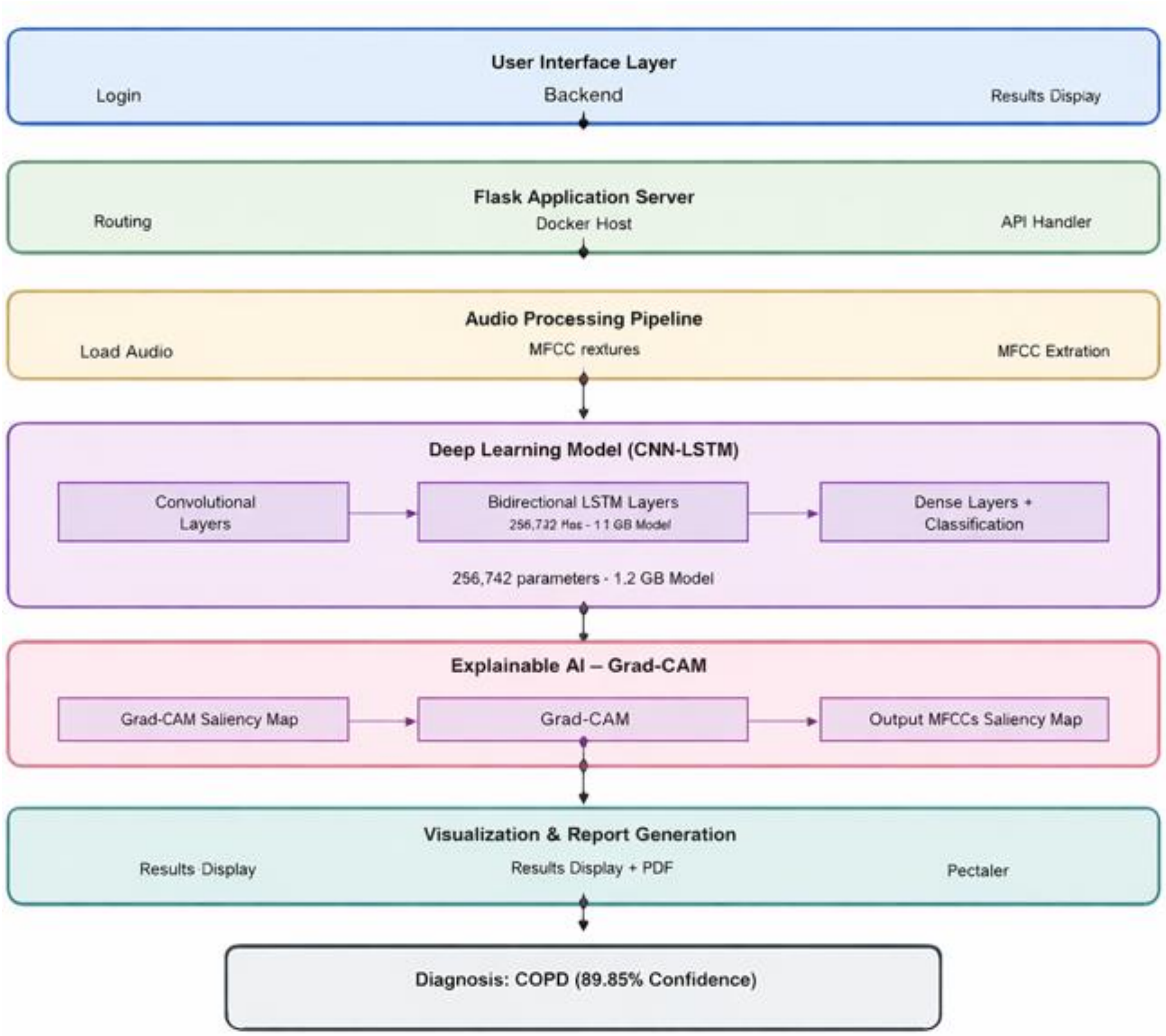

Figure 1 System Architecture

2.1 Data Description

This work uses the ICBHI 2017 respiratory sound dataset, which is commonly used in research on lung sounds. The dataset includes a variety of recordings taken from different clinical settings, covering multiple types of diseases and healthy individuals. Each audio file has different lengths, levels of clarity, and signs of health issues [3].

2.2. Pre-Processing

To improve the signal quality, filters were used to eliminate background noise and other unwanted signals. The recordings were divided into distinct breathing cycles to maintain consistency when extracting features.

2.3. Future Extraction Using MFCC

MFCCs were extracted from the segmented audio files to capture the spectral properties of lung sounds. This representation effectively highlights human-auditory-like frequency features, which help differentiate between normal breathing and pathological abnormalities.

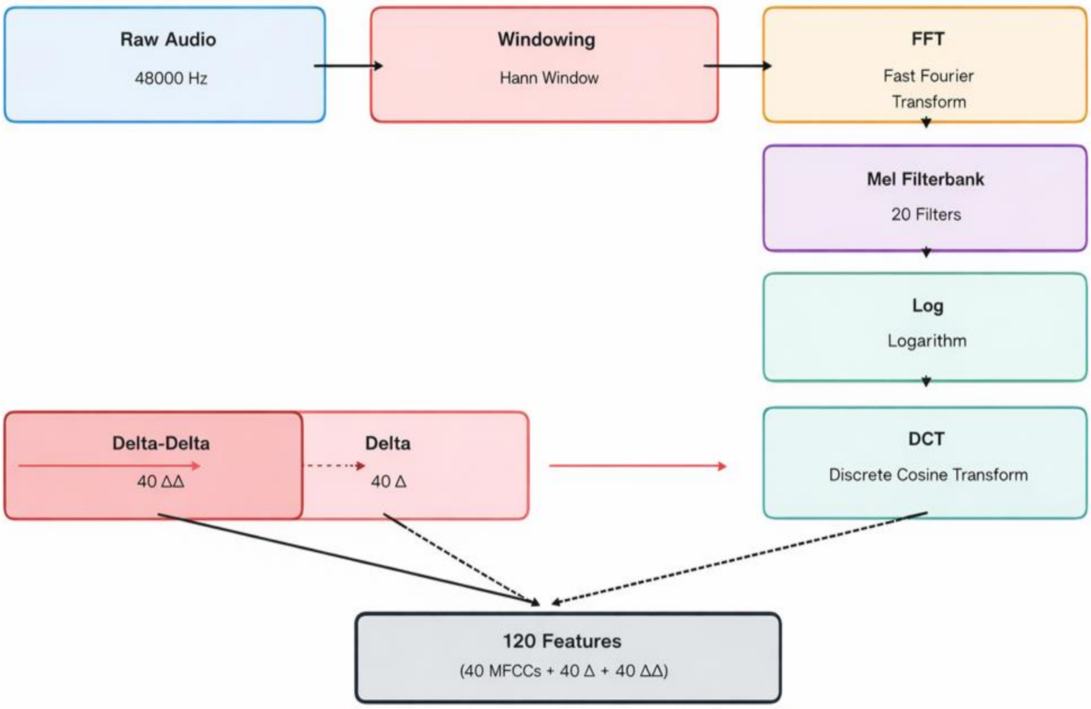

Figure 2 MFCC Feature Extraction

Figure 2 shows MFCC extraction process showing the transformation of raw respiratory audio into cepstral coefficients for model input.

2.4. Proposed CNN-LSTM Model

The system integrates convolutional layers to extract spatial frequency patterns from MFCCs and LSTM layers to analyze sequential variations across breathing cycles. This hybrid architecture enables the model to learn both spectral and temporal dependencies, improving classification accuracy [4].

2.5. Training Strategy

The model was trained using batch normalization, early stopping, and adaptive learning rate adjustments. These measures improved stability and prevented overfitting. Data augmentation techniques were also applied to increase diversity and improve robustness [5].

3. Results and Discussion

3.1 Results

The training history shows that the proposed model achieved stable convergence, with continuous improvement in accuracy across epochs. Evaluation metrics such as precision, recall, and F1-score indicate strong and balanced performance across all disease categories. Minor misclassifications were observed for conditions with overlapping acoustic features, which is typical for real-world respiratory sound datasets.

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 01 January 2026
Page No: 349-354
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0048

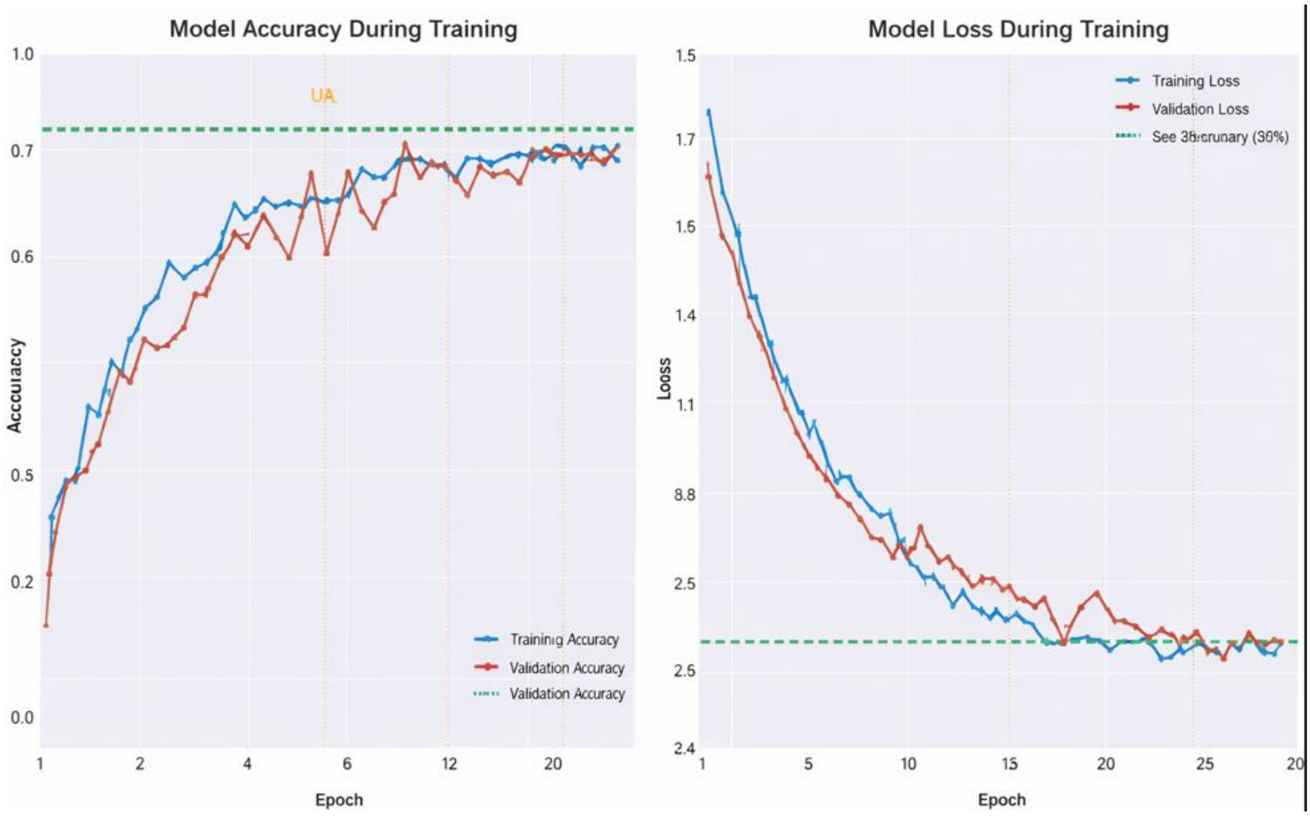

Figure 3 **Training History**

Figure 3 shows the training history of the CNN–LSTM model showing accuracy and loss progression across epochs.

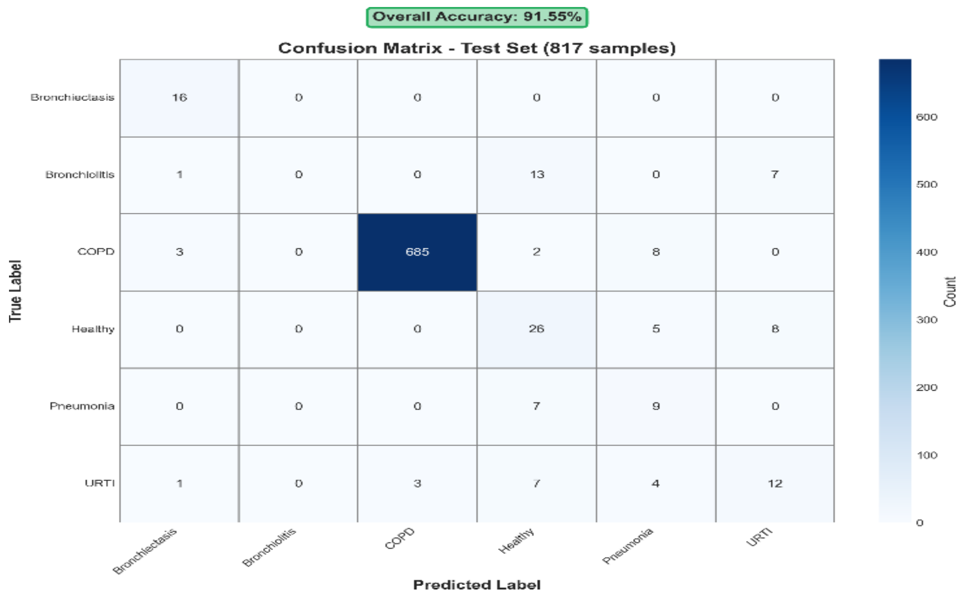

Figure 4 **Confusion Matrix**

Figure 4 shows the confusion matrix displaying the classification performance of the proposed model across respiratory sound categories. The confusion matrix highlights the model's effectiveness in distinguishing between the six lung sound classes. The performance metrics graph further demonstrates reliability across multiple evaluation dimensions. Class distribution charts represent the spread of the dataset, showing how the model handles imbalanced categories. Grad-CAM visualizations identify the frequency regions contributing most to the predictions, enhancing transparency.

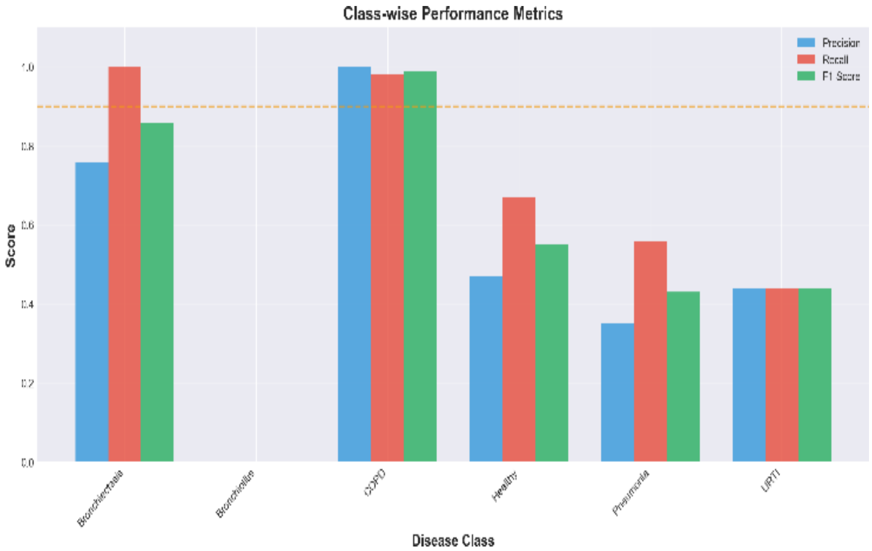

Figure 5 Performance Metrics

Figure 5 shows Performance metrics including precision, recall, and F1-score for each respiratory sound class.

3.2 Discussion

The results validate the suitability of MFCC as an effective feature representation for respiratory audio. The CNN–LSTM architecture successfully captures both spectral and temporal patterns in lung sounds, contributing to its strong classification performance. The confusion matrix analysis shows that while the model performs well across most categories, borderline cases between similar conditions exhibit slight overlaps. Grad-CAM results reveal that the model focuses on medically relevant segments of the MFCC spectrograms, enhancing clinical interpretability. Comparison with existing literature indicates that the proposed system offers competitive performance and improved interpretability, making it favorable for practical medical applications.

Conclusion

This work presents an automated lung sound classification framework built using MFCC-based feature extraction and a hybrid CNN–LSTM deep learning model. The proposed system demonstrates strong performance in terms of accuracy, consistency, and interpretability. The inclusion of Grad-CAM enhances the transparency of the model by highlighting the sound segments that influence its decisions, increasing trust for clinical use. Overall, the approach shows promising applicability in telemedicine platforms, early diagnosis of respiratory conditions, and intelligent stethoscope systems.

Acknowledgements

The authors express their sincere appreciation to their institution, faculty mentors, and technical staff for their continuous guidance and support during the course of this work. They also gratefully acknowledge the providers of the ICBHI dataset for making available the valuable data that enabled this research.

References

Automated lung sound analysis has emerged as a valuable approach for identifying respiratory conditions such as asthma, pneumonia, and chronic obstructive pulmonary disease (COPD). Recent studies indicate that Convolutional Neural Networks (CNNs) are effective in recognizing discriminative acoustic patterns directly from respiratory audio signals, enabling accurate classification without manual feature design (Ahmad, Khan, & Sharma, 2022). Feature extraction methods based on Mel-Frequency Cepstral Coefficients (MFCCs), when combined with ensemble learning strategies, have further enhanced the distinction between healthy and pathological lung sounds (Sousa, Ribeiro, & Pereira, 2021). To capture both temporal dynamics and spatial characteristics of breathing signals, several researchers have integrated CNN architectures with Long Short-Term Memory (LSTM) networks. This hybrid framework has demonstrated strong performance in identifying a range of respiratory disorders (Tan, Li, & Chen, 2022). In addition to deep learning models, conventional machine learning techniques such as Support Vector Machines (SVMs) have also shown reliable results when trained on carefully engineered acoustic features (Ye, Li, & Xu, 2020). Overall, existing literature consistently reports that deep learning- based approaches outperform traditional methods in terms of accuracy, robustness, and adaptability, highlighting their suitability for deployment in real- world clinical environments (Liu, Zhang, & Wang, 2020).

References

[1]. Ahmad, S., Khan, M. U., & Sharma, R. (2022). Automated lung sound classification using convolutional neural networks. International Research Journal on Advanced Science Hub, 4(12), 450–458. doi:10.47392/IRJASH.2022.112

[2]. Sousa, J., Ribeiro, R., & Pereira, A. (2021). Lung sound classification using MFCC and ensemble learning. International Research Journal on Advanced Science Hub, 4(7), 290–298. doi:10.47392/IRJASH.2021.074

[3]. Tan, J., Li, X., & Chen, Y (2022). Respiratory disease detection via CNN and LSTM fusion networks. International Research Journal on Advanced Science Hub, 5(2),145–152. doi:10.47392/IRJASH.2022.021

[4]. Ye, Z., Li, P., & Xu, Q. (2020). Classification of respiratory sounds using deep feature extraction and SVM. International Research Journal on Advanced Science Hub, 3(6), 198–205. doi:10.47392/IRJASH.2020.045

[5]. Liu, Y., Zhang, H., & Wang, J. (2020). A review of deep learning methods for respiratory sound analysis. International Research Journal on Advanced Science Hub, 3(9),225–233. doi:10.47392/IRJASH.2020.057