

SpectroCough: Real-Time Respiratory Disease Screening Using Cough Audio Analysis

Hansa S. Borse¹; Harshal M. Jadhav²; Aryan A. Gaikwad³;
Mayank R. Lahare⁴; Abhishek S. Parmar⁵

¹Professor

^{1,2,3,4,5}Department of Artificial Intelligence and Data Science Pune Vidyarthi Griha's
College of Engineering Nashik, India

Publication Date: 2026/06/06

Abstract: Respiratory diseases such as COVID-19, pneumonia, asthma, bronchitis, and chronic obstructive pulmonary disease (COPD) require early diagnosis for effective treatment and healthcare support. Traditional diagnostic procedures are often expensive, time-consuming, and inaccessible in remote or resource-limited regions. This paper presents SpectroCough, an AI-powered respiratory disease screening framework that performs cough sound analysis using hybrid acoustic feature fusion and deep learning techniques. The proposed system supports both microphone-based cough recordings and stethoscopic respiratory recordings for flexible healthcare deployment. The framework performs audio preprocessing, Mel-spectrogram generation, MFCC extraction, handcrafted acoustic feature extraction, and fusion-based respiratory disease classification using convolutional neural networks and dense neural layers. Additional fake and non-cough detection modules improve system reliability by filtering invalid respiratory sounds such as sneezing, breathing, and artificial cough patterns. Experimental evaluation demonstrated real-time respiratory analysis with prediction latency below 2 seconds. The cough-based model achieved approximately 79.21% accuracy, while the stethoscopic model achieved approximately 80.77% accuracy with strong ROC-AUC separability across respiratory conditions. The proposed framework demonstrates the feasibility of practical mobile healthcare deployment for telemedicine, remote respiratory monitoring, and preliminary clinical screening support.

Keywords: Respiratory Disease Detection, Cough Analysis, Deep Learning, MFCC, Mel Spectrogram, Mobile Healthcare, Audio Classification.

How to Cite: Hansa S. Borse; Harshal M. Jadhav; Aryan A. Gaikwad; Mayank R. Lahare; Abhishek S. Parmar (2026) SpectroCough: Real-Time Respiratory Disease Screening Using Cough Audio Analysis. *International Journal of Innovative Science and Research Technology*, 11(5), 3394-3399. <https://doi.org/10.38124/ijisrt/26may1910>

I. INTRODUCTION

Respiratory diseases remain among the leading causes of global morbidity and mortality. Conditions such as COVID-19, pneumonia, asthma, bronchitis, and chronic obstructive pulmonary disease (COPD) require early diagnosis for effective treatment and reduced healthcare burden. However, conventional diagnostic procedures including laboratory testing, imaging scans, and clinical examinations are often expensive, time-consuming, and inaccessible in underserved regions.

Recent advancements in machine learning and audio signal processing have enabled respiratory cough sounds to serve as non-invasive diagnostic biomarkers. Variations in temporal, spectral, and acoustic characteristics across respiratory conditions provide a foundation for automated disease classification systems.

This work presents SpectroCough, a real-time respiratory screening framework integrating mobile deployment, backend inference systems, hybrid feature fusion, and intelligent respiratory disease classification. The system supports both microphone-based and stethoscopic respiratory recordings for flexible healthcare deployment.

➤ *The Primary Contributions of this Work Include:*

- Development of a complete end-to-end respiratory screening framework
- Support for both microphone and stethoscopic respiratory recordings
- Hybrid fusion of Mel-spectrogram, MFCC, and acoustic features
- Multi-class respiratory disease classification
- Fake and non-cough detection for improved robustness
- Real-time confidence scoring and respiratory report generation

- Mobile healthcare deployment for remote respiratory screening

II. RELATED WORK

Several studies have investigated respiratory disease detection using cough audio analysis. Mel-frequency cepstral coefficients (MFCCs), spectrogram-based representations, and convolutional neural networks have demonstrated promising classification performance across respiratory datasets.

Huang and Mushi developed a parallel-stream deep convolutional architecture using multiple spectrogram representations for cough classification. Hamdi et al. proposed a hybrid CNN-LSTM framework combined with spectral augmentation for COVID-19 detection.

Benaliouche et al. investigated transfer learning approaches for respiratory disease classification and demonstrated strong performance using MFCC-based representations. Pahar et al. explored tuberculosis cough analysis under real-world noisy clinical conditions.

Although previous studies demonstrated promising classification performance, many lacked practical deployment capabilities such as real-time inference, mobile healthcare integration, multi-input respiratory analysis, fake cough filtering, and user-oriented diagnostic reporting systems. The proposed SpectroCough framework extends existing research by integrating these deployment-oriented capabilities into a unified respiratory screening platform.

III. PROPOSED METHODOLOGY

➤ System Overview

The SpectroCough framework is designed as a complete respiratory screening pipeline integrating audio acquisition, preprocessing, hybrid feature extraction, fusion-based classification, and real-time prediction generation.

The implemented architecture supports both microphone-based and stethoscopic respiratory recordings for deployment across mobile and clinical healthcare environments.

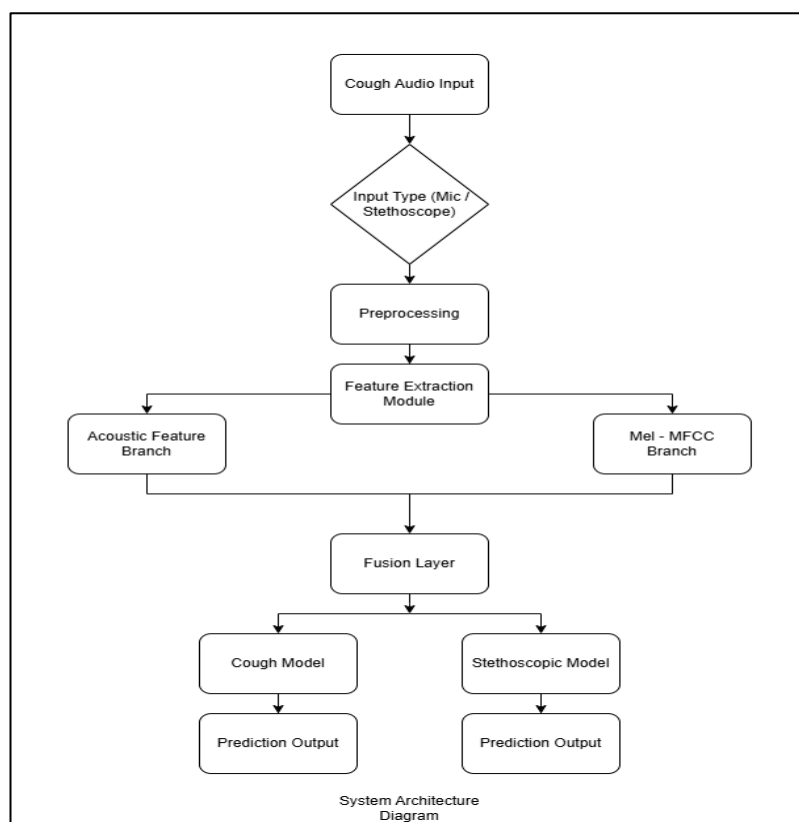


Fig 1 System Architecture Diagram

Fig. 1. Overall architecture of the implemented SpectroCough framework showing dual-input respiratory acquisition, preprocessing, hybrid feature extraction, fusion-based classification, and prediction generation.

➤ Audio Acquisition and Preprocessing

The framework accepts both microphone-based cough recordings and stethoscopic respiratory recordings. The

acquired respiratory audio undergoes multiple preprocessing stages including voice activity detection, silence removal, noise filtering, loudness normalization, and signal standardization. These operations improve signal consistency, suppress environmental interference, and enhance robustness during model inference.

➤ *Feature Extraction*

The implemented framework utilizes a dual-branch hybrid feature extraction pipeline.

• *Spectrogram Branch:*

This branch generates Mel-spectrogram and MFCC representations from respiratory audio signals. Convolutional neural network layers extract deep temporal and spectral representations from these inputs.

• *Acoustic Feature Branch:*

The acoustic feature branch extracts handcrafted descriptors including MFCC statistical means, chroma features, RMS energy, spectral centroid values, and Mel-band statistical representations. These features capture frequency distribution, acoustic energy variation, and respiratory sound dynamics across disease conditions.

• *YAMNet Embedding Extraction:*

For microphone-based cough analysis, the implemented framework additionally incorporates YAMNet-based embedding extraction to enhance respiratory sound representation learning. The pretrained YAMNet model generates high-level semantic audio embeddings from cough recordings, enabling the system to capture richer temporal and acoustic characteristics associated with respiratory conditions. These deep audio embeddings improve robustness under noisy real-world recording environments and contribute to improved classification performance for microphone-based respiratory screening.

➤ *Feature Fusion and Classification*

The extracted spectrogram features and acoustic features are fused using feature concatenation. The fused feature vector is passed through dense neural network layers followed by softmax classification for respiratory disease prediction.

The models were trained using categorical cross-entropy loss and optimized using the Adam optimizer.

➤ *Classification Modules*

The framework contains three classification modules including a stethoscopic respiratory classifier, a microphone cough classifier, and a fake/non-cough detection classifier.

The stethoscopic model predicts asthma, COPD, pneumonia, bronchitis, and healthy respiratory conditions. The microphone-based cough model predicts COVID-19 and healthy cough conditions. The fake/non-cough detection framework filters sneezing, breathing, laughing, and artificial cough patterns to improve deployment reliability.

IV. IMPLEMENTATION DETAILS

➤ *Technology Stack*

The implemented framework utilizes Python with TensorFlow and Keras for model development, Librosa for respiratory audio processing, Flutter for mobile deployment,

Streamlit for web-based testing, and PostgreSQL database integration.

➤ *Training Configuration*

The models were trained using the Adam optimizer with categorical cross-entropy loss, batch size of 32, learning rate scheduling, and early stopping regularization. Experimental evaluation was conducted using an Intel Core i5 system with 16 GB RAM and NVIDIA GPU acceleration.

V. EXPERIMENTAL RESULTS AND EVALUATION

➤ *Performance Metrics*

The system was evaluated using accuracy, precision, recall, F1-score, confusion matrix analysis, and ROC-AUC evaluation.

Table 1 Comparative Performance Metrics

Metric	Microphone-based Cough Model	Stethoscopic Model
Accuracy	79.21%	80.77%
Precision	80%	78.27%
Recall	79%	80.53%
F1-score	79%	79.04%

➤ *Confusion Matrix Analysis*

The confusion matrix demonstrates strong classification performance across respiratory categories, with most predictions concentrated along the diagonal. COPD and asthma classes achieved strong separability, while limited overlap was observed between bronchial and pneumonia classes due to similarity in respiratory acoustic characteristics.

The confusion matrix further indicates that the majority of respiratory predictions remain concentrated along the principal diagonal, demonstrating strong classification consistency across multiple respiratory categories. COPD achieved the strongest prediction stability with minimal inter-class confusion, while asthma and pneumonia exhibited limited overlap due to similarities in respiratory acoustic patterns. The healthy respiratory class also demonstrated reliable discrimination capability, validating the effectiveness of the proposed hybrid feature fusion architecture under practical respiratory screening conditions.

The obtained confusion matrix characteristics demonstrate that the proposed framework is capable of distinguishing multiple respiratory conditions with strong overall robustness. The integration of spectrogram-based deep representations and handcrafted acoustic descriptors improves class separability across complex respiratory patterns. These findings validate the suitability of the implemented framework for real-time respiratory screening applications in telemedicine and mobile healthcare environments.

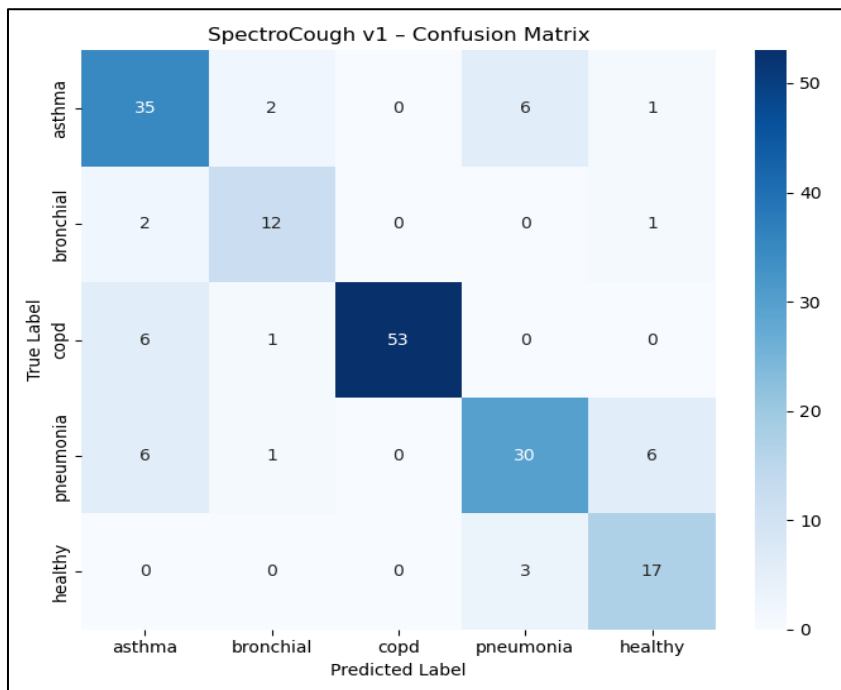


Fig 2 Confusion Matrix Analysis of the Proposed Stethoscopic Respiratory Classification Model Across Five Respiratory Conditions.

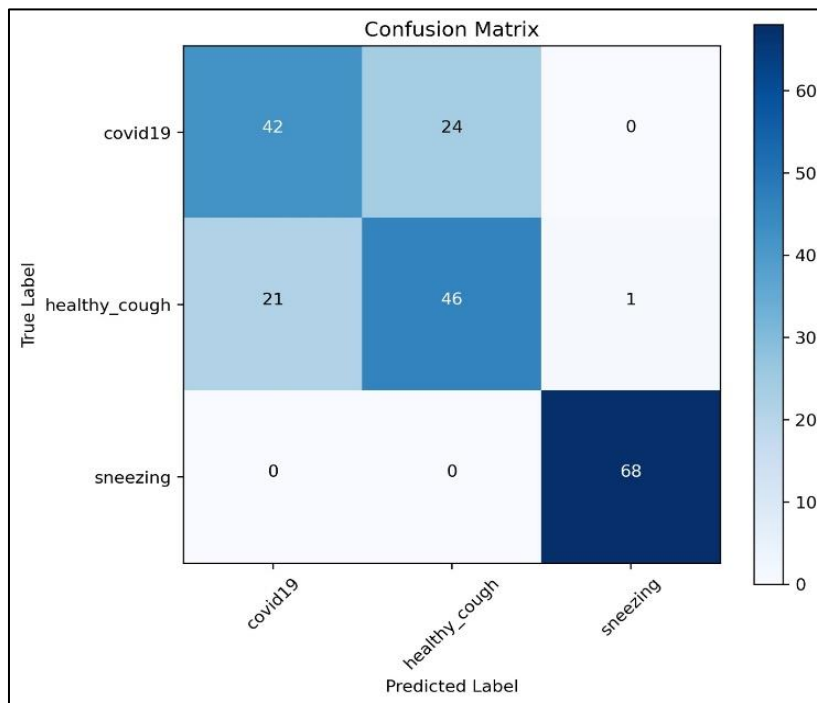


Fig 3 Confusion Matrix Analysis of the Proposed Microphone-Based Cough Classification Model.

➤ *ROC-AUC Evaluation*

The ROC curves demonstrate strong class separability across respiratory conditions. COPD achieved near-perfect classification performance with an AUC approaching 0.995. The macro-average AUC of approximately 0.965 indicates strong overall robustness of the proposed fusion-based architecture.

Most respiratory classes achieved AUC values above 0.93, demonstrating strong discriminative capability across

respiratory conditions. COPD exhibited the strongest class separability, while asthma and pneumonia also achieved high classification performance despite partial overlap in respiratory acoustic characteristics. The obtained ROC characteristics validate the effectiveness of the hybrid feature fusion approach for multi-class respiratory screening applications.

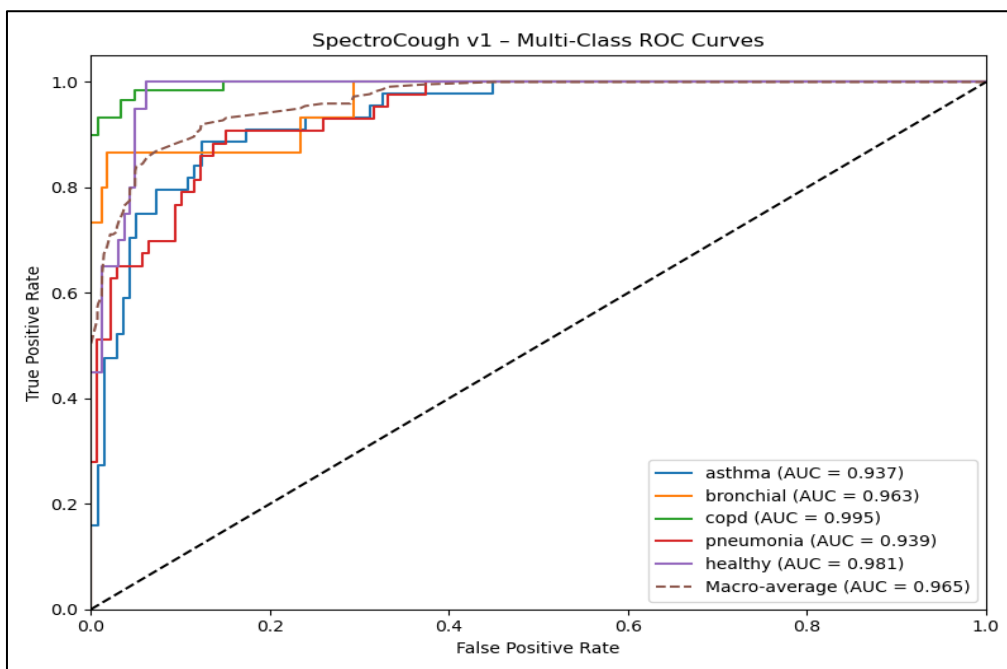


Fig 4 ROC-AUC Analysis of the Proposed Stethoscopic Respiratory Classification Framework.

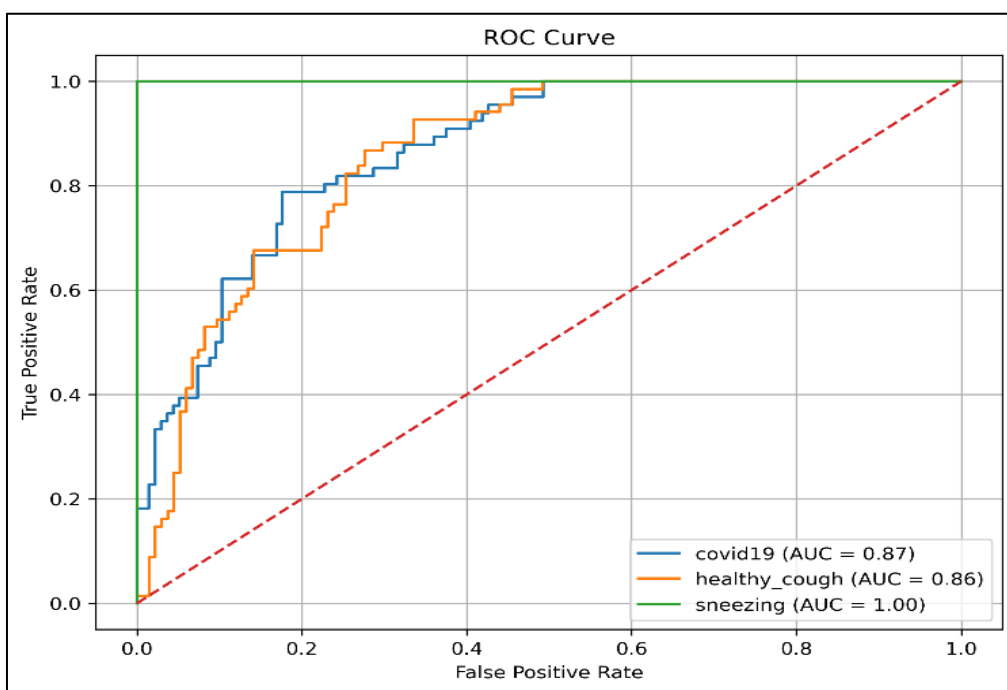


Fig 5 ROC-AUC Analysis of the Proposed Microphone-Based Cough Classification Model.

➤ *Comparison with Existing Methods*

Although some existing studies reported higher classification accuracy, direct comparison remains challenging because many previous approaches focused on binary classification tasks using highly controlled datasets.

The proposed framework addresses more complex multi-class respiratory classification under practical deployment conditions and real-world acoustic variability.

Table 2 Comparison with Existing Respiratory Classification Approaches

Ref	Method	Features Used	Classes	Accuracy
[1]	CNN	MFCC	Binary	94%
[2]	CNN-LSTM	Spectrogram	Binary	92%
[3]	AI Model	Acoustic	Multi-class	88%
Proposed	Hybrid (Fusion)	Mel+MFCC + Acoustic	Multi-class	~80%

➤ *System Efficiency*

The developed framework achieves real-time respiratory analysis with prediction latency below 2 seconds, making it suitable for telemedicine, remote respiratory monitoring, and preliminary clinical triage support.

VI. DISCUSSION

The experimental findings demonstrate that hybrid fusion of deep spectrogram representations and handcrafted acoustic descriptors improves respiratory classification robustness across multiple disease conditions. The integration of fake and non-cough filtering further enhances deployment reliability by reducing invalid respiratory predictions caused by environmental sounds and artificial cough patterns.

The proposed framework demonstrates practical applicability in telemedicine, remote respiratory monitoring, and preliminary clinical triage support. Although certain respiratory classes exhibit partial acoustic overlap, the overall classification performance and ROC-AUC analysis validate the effectiveness of the implemented fusion-based architecture for real-world respiratory screening applications.

VII. CONCLUSION AND FUTURE WORK

This work presented SpectroCough, a practical real-time respiratory screening framework integrating hybrid acoustic feature fusion, deep learning-based classification, and mobile healthcare deployment for respiratory disease analysis.

The proposed framework integrates Mel-spectrogram analysis, MFCC feature extraction, acoustic feature fusion, multi-class respiratory classification, fake cough detection, and mobile healthcare deployment into a unified respiratory screening platform.

Experimental evaluation demonstrated reliable classification performance with real-time inference latency below 2 seconds. The stethoscopic classification model achieved approximately 80.77% accuracy with strong ROC-AUC separability across respiratory conditions.

The obtained results validate the feasibility of practical mobile respiratory screening systems for telemedicine, remote monitoring, and accessible healthcare support.

Future research will focus on expanding dataset diversity, improving robustness under noisy environmental conditions, integrating explainable artificial intelligence techniques, and conducting broader clinical validation studies across diverse patient populations.

ACKNOWLEDGMENT

The authors acknowledge the support of Pune Vidyarthi Griha's College of Engineering & S. S. Dhamankar Institute of Management, Nashik

The authors also express gratitude to Prof. Hansa S. Borse for valuable guidance throughout the project.

REFERENCES

- [1]. P. Huang and R. Mushi, "Classification of cough sounds using spectrogram methods and a parallel-stream one-dimensional deep convolutional neural network," *IEEE Access*, vol. 10, pp. 97089–97100, 2022.
- [2]. S. Hamdi, M. Oussalah, A. Moussaoui, and M. Saidi, "Attention-based hybrid CNN-LSTM and spectral data augmentation for COVID-19 diagnosis from cough sound," *Journal of Intelligent Information Systems*, vol. 59, pp. 367–389, 2022.
- [3]. H. Benaliouche, H. Hafi, H. Bendjenna, and Z. Alshaikh, "Toward AI-driven cough sound analysis for respiratory disease diagnosis," *IEEE Access*, vol. 13, pp. 92554–92568, 2025.
- [4]. M. Pahar, M. Klopper, B. Reeve, R. Warren, G. Theron, and T. Niesler, "Automatic cough classification for tuberculosis screening in a real-world environment," *Physiological Measurement*, vol. 42, p. 105014, 2021.