Detection of Lung Condition from Respiratory Sounds using a Hybrid Deep Learning

CH. Swetha[1], Dr. A. Obulesu[2], K. Pranav Sreekar[3], K. Snigdha[4] , D. Dinesh[5] , T. Srikar[6]

Assistant Professor[1], Associate Professor[2], Student[3,4,5,6]

Vidya Jyothi Institute of Technology, Hyderabad, India

Abstract - Respiratory diseases such as asthma, pneumonia, bronchitis, and chronic obstructive pulmonary disease (COPD) require early and accurate diagnosis to reduce health risks. Conventional lung sound analysis through manual auscultation is subjective and dependent on clinical expertise, while respiratory sounds are complex and often affected by noise. These factors reduce the reliability of traditional diagnostic methods.

The paper presents a deep learning-based hybrid system for automatic lung disease detection using respiratory sounds. The signals are preprocessed and converted into Log-Mel spectrograms and MFCCs for feature extraction. A hybrid model combining ResNet-50 and ResNet-34 is used to learn discriminative spatial and frequency-domain features, where features from both networks are integrated within a single framework to improve classification performance.

The system performs multi-class classification of different lung conditions. Experimental evaluation on publicly available respiratory sound datasets shows consistent performance under varying recording conditions. The results suggest that the proposed approach can assist clinicians in early and non-invasive lung disease detection.

Keywords: Lung sound analysis, Respiratory disease classification, Hybrid deep learning, ResNet-50, ResNet-34, Log-Mel spectrogram.

I. INTRODUCTION

Respiratory disorders including asthma, pneumonia, bronchitis, chronic obstructive pulmonary disease (COPD), and COVID-19 represent a major challenge to global healthcare systems. These conditions differ in origin and progression but commonly affect airflow and gas exchange within the lungs [35], [24].

Asthma is a chronic airway disorder characterized by inflammation and intermittent airflow obstruction. Pneumonia is an infectious condition causing inflammation of the alveoli and impaired oxygen exchange. Bronchitis involves inflammation of the bronchial lining, leading to persistent

cough and mucus production, while chronic bronchitis contributes to long-term respiratory damage. COPD is a progressive disease marked by irreversible airflow limitation, often associated with smoking and environmental exposure. COVID-19, caused by SARS-CoV-2, primarily affects the respiratory system and ranges from mild symptoms to severe respiratory failure.

Respiratory sounds are non-stationary and sensitive to noise, recording conditions, and patient variability. Conventional diagnostic tools such as X-rays and CT scans are effective but costly and less accessible. These limitations have encouraged the development of automated and non-invasive lung disease detection using respiratory sound analysis [3].

Earlier studies used handcrafted features such as MFCCs with classifiers like SVM and KNN [1], [36], but these approaches showed limited generalization. Deep learning models, particularly CNNs, improved performance by learning features from spectrogram representations [29], [30], though many methods remain complex or disease-specific.

Motivated by these challenges, this work proposes a hybrid deep learning framework for automated lung disease classification using respiratory sound recordings. The system utilizes spectrogram-based feature extraction and a hybrid model combining ResNet-50 and ResNet-34 to perform reliable multi-class classification under varying recording conditions.

II. LITERATURE SURVEY

Automated respiratory sound analysis has gained attention for enabling early and non-invasive diagnosis of pulmonary diseases. Traditional auscultation is subjective and varies across clinicians, motivating the use of computational approaches for consistent lung sound analysis. Several studies have applied signal processing and deep learning techniques to address this challenge.

Huang et al. reviewed deep learning–based lung sound analysis for intelligent stethoscopes, outlining stages such as preprocessing, feature extraction, model development, and evaluation. Their work highlighted the effectiveness of spectrogram-based inputs with CNN architectures while noting challenges such as noise sensitivity and class imbalance [10]. Hybrid architectures have also been explored to improve performance. Petmezas et al. proposed a CNN–LSTM model to capture spatial and temporal characteristics, reporting improved sensitivity compared to standalone CNN models [7]. Performance studies further indicated that Mel-spectrogram features generally outperform traditional handcrafted features [17].

Recent works focused on improving robustness and efficiency.

The CycleGuardian framework introduced a lightweight deep model integrating clustering and contrastive learning for better discrimination between normal and abnormal sounds [16]. Feature fusion and ensemble approaches have also been investigated to enhance classification, though they often increase complexity and tuning requirements [12], [43].

Overall, existing studies confirm that deep learning models using spectrogram representations achieve superior results compared to traditional methods. However, issues related to generalization and robustness remain. Motivated by these findings, this work adopts a hybrid deep learning approach based on ResNet-50 and ResNet-34 to develop a reliable multi-disease lung sound classification framework using publicly available respiratory sound datasets.

III. DATA COLLECTION AND PREPROCESSING

Respiratory sound recordings used in this study are obtained from publicly available datasets, namely the ICBHI respiratory sound database and the Coswara dataset. These datasets include recordings from healthy subjects as well as patients diagnosed with respiratory diseases such as asthma, pneumonia, bronchitis, and chronic obstructive pulmonary disease (COPD). The diversity in recording environments and subject conditions makes them suitable for evaluating automated lung sound classification systems [3], [10].

Method	Features	Model	Limitation
Performance Evaluation	MFCC, Spectrogram	CNN	No unified model
Intelligent Stethoscope	Spectrogram	CNN / Hybrid	Review only
Feature Fusion CNN	Spectrogram	Multi-CNN	High complexity
Cycle Guardian	Grouped spectrogram	Lightweight CNN	Limited classes
Proposed System	Log-Mel, MFCC	CNN ensemble	Balanced approach

Table1: Comparative Analysis of Existing Works and Proposed System

The acquired recordings vary in duration, sampling frequency, and quality. To maintain consistency, all audio files are converted to a common format and resampled to a uniform sampling rate. This reduces device-based variability and supports efficient processing.

Preprocessing is applied to enhance signal quality before feature extraction. Silent and low-energy segments are removed, and amplitude normalization is performed to reduce signal variation. Basic noise suppression techniques are used to limit background interference while preserving important

frequency components [2], [20], [48].

After preprocessing, the audio signals are segmented into short overlapping frames to capture time–frequency characteristics. The segmented respiratory sounds are then forwarded to the feature extraction stage for further analysis.

IV. METHODOLOGY

The proposed methodology focuses on automated lung disease detection from respiratory sound recordings using a hybrid deep learning framework. The overall workflow includes data acquisition, audio preprocessing, feature extraction, deep feature learning, hybrid-based classification, and performance evaluation. The system follows commonly adopted pipelines in lung sound analysis with a simplified and unified model structure.
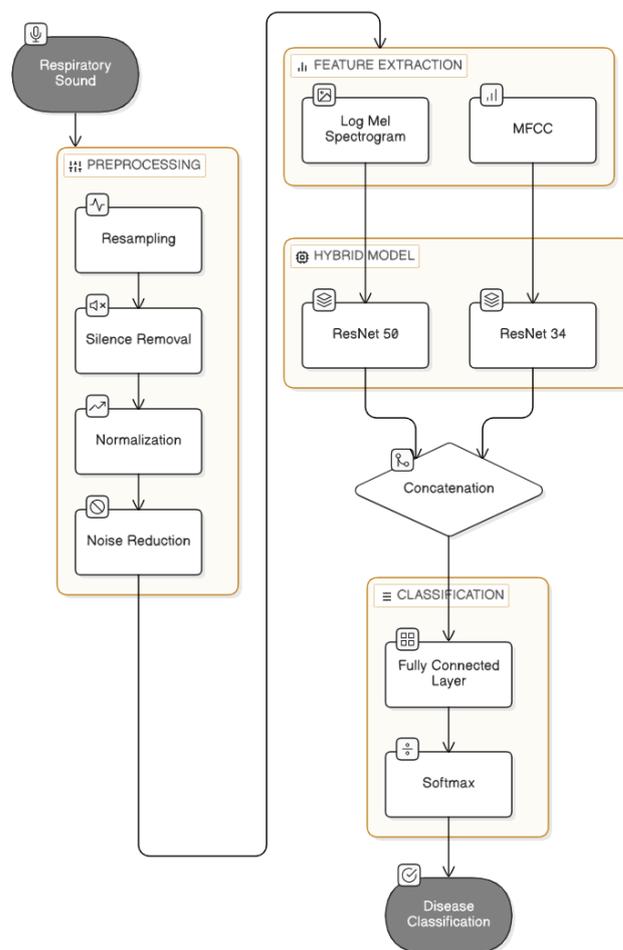

Fig 1: Architecture Diagram

Respiratory sound data are collected from publicly available datasets such as ICBHI and Coswara, which are widely used benchmarks for lung sound classification. These datasets include recordings from healthy subjects as well as patients with various respiratory conditions and support evaluation under diverse recording environments [3], [10].

In the preprocessing stage, raw recordings are standardized to maintain consistency. Silence removal and amplitude normalization are applied to reduce signal variability, and basic noise suppression is used to minimize background disturbances while preserving relevant frequency components [2], [20], [48].

After preprocessing, the signals are transformed into discriminative representations. Log-Mel spectrograms and Mel-Frequency Cepstral Coefficients (MFCCs) are extracted to convert one-dimensional audio signals into two-dimensional time–frequency features, which effectively capture abnormal lung sound patterns [5], [29], [11].

For deep feature learning, a hybrid architecture combining ResNet-50 and ResNet-34 is employed. Transfer learning is adopted to improve convergence and reduce overfitting on limited medical audio data. Features extracted from both networks are integrated within a single framework to enhance classification reliability.

The framework performs multi-class classification to identify multiple lung diseases in a unified system. Performance evaluation is conducted using standard metrics such as accuracy, precision, recall, and F1-score [7], [17]. The methodology provides a non-invasive and reliable approach for early lung disease detection.

V. WORKING OF THE PROPOSED SYSTEM

The proposed system operates as a sequential pipeline designed to automatically analyze respiratory sound recordings and classify lung diseases using a hybrid deep learning approach. The working of the system begins with the acquisition of lung sound signals and proceeds through multiple processing stages to produce the final disease prediction.

Initially, respiratory sound recordings are collected from publicly available datasets and provided as input to the system. These raw audio signals pass through a preprocessing stage, where signal standardization ensures uniform sampling and format. Silence removal, amplitude normalization, and basic noise suppression are applied to enhance signal quality and reduce background interference [10], [20].

After preprocessing, the audio signals are transformed into time–frequency representations. Log-Mel spectrograms and Mel-Frequency Cepstral Coefficients (MFCCs) are extracted to capture spectral characteristics of lung sounds and highlight abnormal patterns such as wheezes and crackles [5], [11].

The extracted features are then fed into the deep learning module. A hybrid architecture combining ResNet-50 and ResNet-34 is employed to automatically learn discriminative spatial and frequency-domain features from the spectrogram images. Transfer learning is applied to improve learning efficiency and reduce overfitting. Each network extracts complementary features, which are integrated to generate class probability predictions [7], [29].
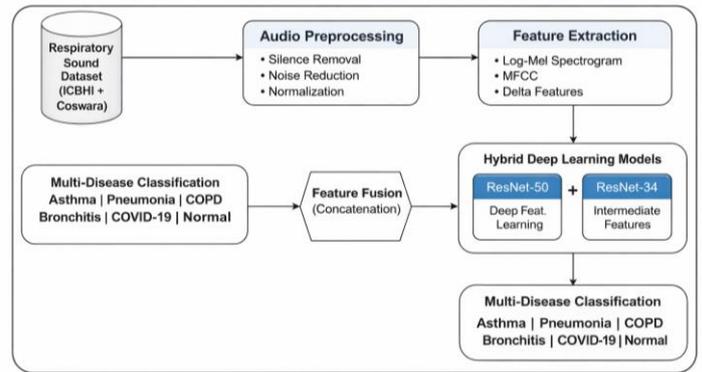

Fig 2: Work Flow

ResNet-50 is a deep residual network designed to capture high-level abstract features through residual connections that enable effective gradient flow across layers.
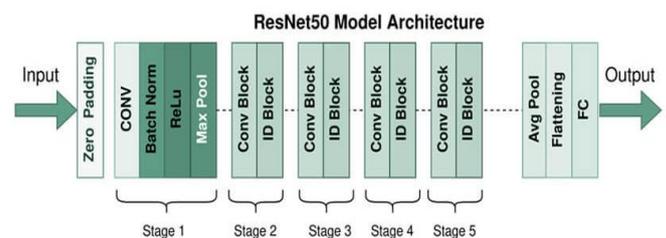

Fig 3: ResNet 50 Architecture

ResNet-34 is a moderately deep residual network that captures detailed and intermediate feature representations while maintaining computational efficiency.

To improve robustness and reliability, the outputs from ResNet-50 and ResNet-34 are combined within the hybrid framework. This integration reduces model-specific bias and enhances classification stability under varying recording conditions [16], [49].
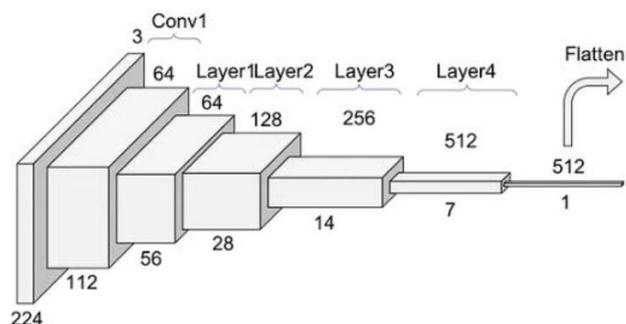

Fig 4 : ResNet 34 Architecture

The system performs multi-class classification to identify different lung conditions within a single framework. The predicted output corresponds to the detected respiratory condition of the input recording. This workflow provides an efficient and non-invasive approach for lung disease detection

and supports clinical decision-making.

VI. PERFORMANCE METRICS AND EVALUATION BASIS

The performance of the proposed lung disease classification system is evaluated using standard metrics commonly adopted in respiratory sound analysis and medical classification tasks [7], [17]. Since the dataset contains multiple disease classes and may exhibit class imbalance, accuracy alone is not sufficient to assess model performance. Therefore, additional evaluation metrics are considered to provide a comprehensive analysis.

Accuracy is used to measure the overall correctness of the classification system by computing the ratio of correctly predicted samples to the total number of samples. While accuracy provides a general indication of model performance, it may be misleading when class distributions are uneven [17].

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$

Precision is employed to evaluate the reliability of positive predictions by measuring the proportion of correctly predicted positive samples among all predicted positives. High precision indicates a lower false positive rate, which is important in medical diagnosis to avoid incorrect disease identification [22], [24].

$$Precision = \frac{TP}{TP+FP}$$

Recall, also known as sensitivity, measures the ability of the system to correctly identify actual positive cases. This metric is particularly significant in healthcare applications, as low recall may result in missed disease cases, leading to delayed diagnosis [7], [24].

$$Recall = \frac{True\ Positives\ (TP)}{True\ Positives\ (TP)+False\ Negatives\ (FN)}$$

The F1-score is used as a balanced metric that combines precision and recall into a single measure. It provides a more reliable evaluation of model performance when dealing with imbalanced datasets and reflects the trade-off between false positives and false negatives [17], [22].

$$F1\ score = \frac{2TP}{2TP+FP+FN}$$

The evaluation is conducted on publicly available respiratory sound datasets using a standard train–test split to ensure fair performance assessment. These metrics collectively provide a reliable basis for analyzing the effectiveness, robustness, and clinical applicability of the proposed system.

VII. RESULT AND DISCUSSION

The proposed hybrid deep learning system is evaluated on publicly available respiratory sound datasets to assess its effectiveness in classifying multiple lung diseases. The evaluation is conducted using standard performance metrics

including accuracy, precision, recall, and F1-score. The results demonstrate consistent and reliable performance across different respiratory conditions.
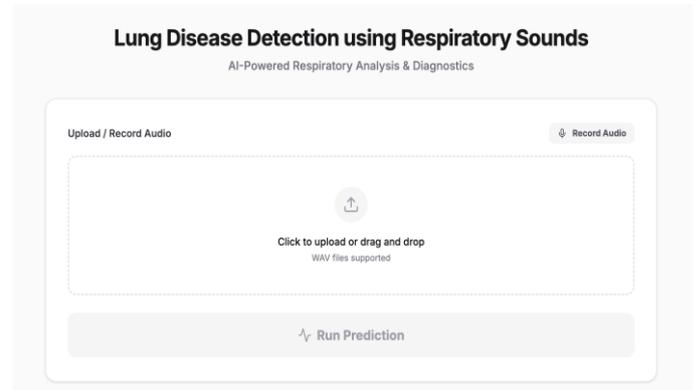

Fig 5 : Taking input from the user

The use of spectrogram-based feature representations enables the hybrid ResNet-50 and ResNet-34 model to effectively capture discriminative acoustic patterns such as wheezes and crackles. The residual networks learn spatial and frequency-domain characteristics from these representations, contributing to accurate classification. Compared to traditional handcrafted feature-based methods, the proposed system shows improved robustness under varying recording conditions [6], [33].
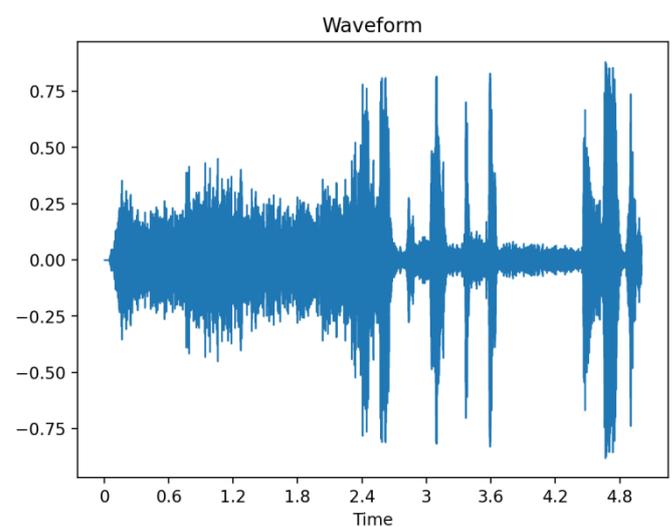

Fig 6:Time-domain waveform of a lung sound signal used for analysis.

The hybrid integration of ResNet-50 and ResNet-34 enhances classification reliability by combining deep and intermediate feature representations within a single framework. This reduces model-specific bias and improves generalization, particularly for classes with limited samples. The framework demonstrates a better balance between precision and recall, which is important for medical diagnosis [43], [49].

Temporal variations in respiratory sounds are effectively captured through the selected feature extraction and hybrid

learning strategy, enabling accurate multi-class classification. The results indicate competitive performance compared with existing approaches while maintaining moderate computational complexity [21], [32].
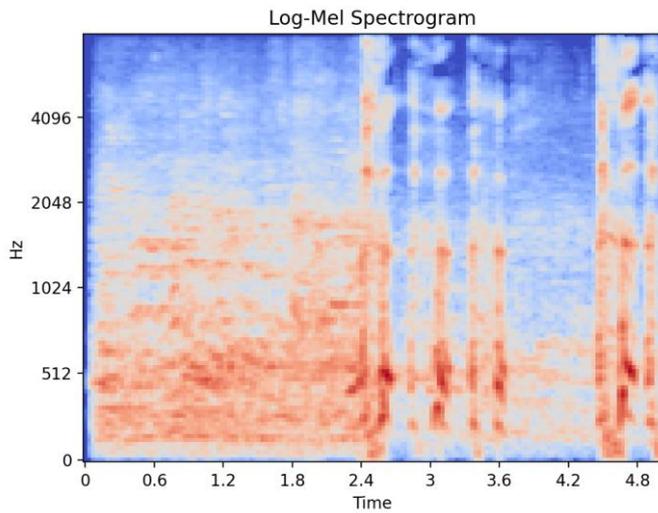

Fig 7:Log-Mel spectrogram representation of the lung sound signal used for feature extraction.

Overall, the experimental results validate the effectiveness of the proposed hybrid ResNet-based framework for automated lung disease detection. The balance between accuracy, robustness, and feasibility makes the system suitable for computer-aided diagnosis and potential deployment in remote healthcare environments [24].
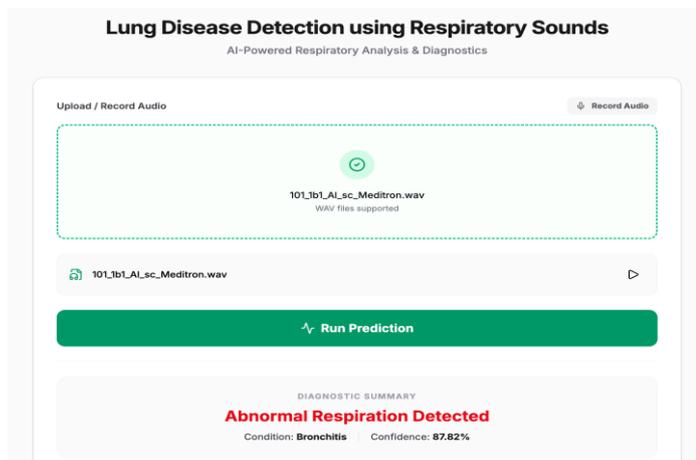

Fig 8:Output of the lung sound classification model indicating abnormal respiration with confidence score.

VIII. CONCLUSION

This paper presented a hybrid deep learning–based framework for automated lung disease detection using respiratory sound recordings. The proposed system integrates effective audio preprocessing, spectrogram-based feature extraction, and transfer learning–based ResNet-50 and ResNet-34 within a hybrid framework to achieve reliable multi-class lung disease

classification. The approach addresses the limitations of traditional auscultation by providing a consistent, non-invasive, and automated diagnostic solution.

Experimental evaluation on publicly available respiratory sound datasets demonstrates that the proposed system effectively captures discriminative acoustic patterns associated with different lung conditions. The hybrid integration of ResNet-50 and ResNet-34 improves robustness and reduces model-specific bias, resulting in balanced performance across multiple disease classes. It is also observed that datasets with minimal background noise and properly segmented respiratory cycles enable higher classification accuracy and more stable predictions, highlighting the importance of data quality in lung sound analysis.
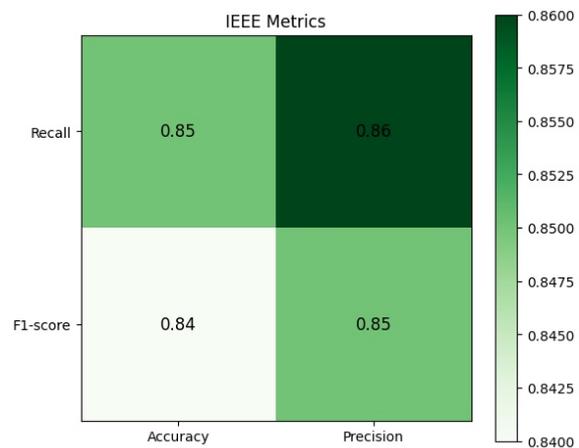

Fig 9: Evaluation metrics of the proposed hybrid lung sound classification model.

Overall, the proposed system shows potential as a computer-aided diagnostic tool to assist clinicians in early lung disease detection. Its moderate computational complexity and reliance on non-invasive data make it suitable for deployment in remote and resource-constrained healthcare environments.

IX. REFERENCES

[1] M. Bahoura and C. Pelletier, "New parameters for respiratory sound classification," in *Proc. IEEE Canadian Conf. Electrical and Computer Engineering (CCECE)*, vol. 3, 2003, pp. 1457–1460.

[2] D. Emmanouilidou, M. McCollum, D. Park, and M. Elhilali, "Computerized lung sound screening for pediatric auscultation in noisy field environments," *IEEE Transactions on Biomedical Engineering*, vol. 65, no. 7, pp. 1564–1574, 2018.

[3] B. M. Rocha, D. Pessoa, A. Marques, P. Carvalho, and R. P. Paiva, "A respiratory sound database for the development of automated classification," in *Precision Medicine Powered by pHealth and Connected Health*, Springer, 2017, pp. 33–37.

[4] Y. Ma, X. Xu, and Y. Li, "LungRN+NL: An improved adventitious lung sound classification using non-local block ResNet neural network with Mixup data augmentation," in *Proc. Interspeech*, 2020, pp. 2902–2906.

[5] S. Shuvo, S. B. Ali, S. I. Swapnil, T. Hasan, and M. I. H. Bhuiyan, "A lightweight CNN model for detecting respiratory diseases from lung auscultation sounds using EMD–CWT-based hybrid scalogram," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 7, pp. 2595–2603, 2021.

[6] B. M. Rocha *et al.*, "Automated classification of adventitious respiratory sounds: A (un)solved problem?" *Sensors*, vol. 21, no. 1, p. 57, 2021.

[7] A. Petmezas *et al.*, "Automated lung sound classification using a hybrid CNN–LSTM network and focal loss function," *Sensors*, vol. 22, no. 3, p. 1232, 2022.

[8] S. Gupta, M. Agrawal, and D. Deepak, "Classification of auscultation sounds into objective spirometry findings using MVMD and 3D CNN," in *Proc. National Conf. Communications (NCC)*, IEEE, 2022, pp. 42–47.

[9] L. Brunese, F. Mercaldo, A. Reginelli, and A. Santone, "A neural network-based method for respiratory sound analysis and lung disease detection," *Applied Sciences*, vol. 12, no. 8, 2022.

[10] H. Huang *et al.*, "Deep learning-based lung sound analysis for intelligent stethoscope," *Military Medical Research*, vol. 10, no. 44, 2023.

[11] R. Roy and U. Satija, "A novel melspectrogram snippet representation learning framework for severity detection of chronic obstructive pulmonary disease," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, 2023.

[12] Y. Zhang, J. Zhang, J. Yuan, H. Huang, and Y. Zhang, "Research on lung sound classification model based on dual-channel CNN–LSTM algorithm," *Biomedical Signal Processing and Control*, vol. 86, 2024.

[13] F. Wang *et al.*, "OFGST-Swin: Swin transformer utilizing overlap fusion-based generalized S-transform for respiratory cycle classification," *IEEE Transactions on Instrumentation and Measurement*, 2024.

[14] A. Fraihi *et al.*, "Improving deep learning-based respiratory sound analysis with frequency selection and attention mechanism," *arXiv preprint*, 2025.

[15] Bikku *et al.*, "Deep learning-driven early diagnosis of respiratory diseases using lung sound analysis," *Scientific Reports*, Nature, 2025.

[16] Y. Chu, Q. Wang, E. Zhou, L. Fu, and Q. Liu, "CycleGuardian: A framework for automatic respiratory sound classification based on improved deep clustering and contrastive learning," *Complex & Intelligent Systems*, vol. 11, p. 200, 2025.

[17] Q. Zhang and J. Sun, "Performance evaluation of deep learning models for respiratory sound classification," *EURASIP Journal on Advances in Signal Processing*, 2024.

[18] A. Altan and Y. Kutlu, "Diagnosis of chronic obstructive pulmonary disease using deep learning," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 5, pp. 1344–1350, 2020.

[19] J. Li *et al.*, "LungAttn: Advanced lung sound classification using attention mechanism with dual TQWT and triple STFT spectrogram," *Physiological Measurement*, vol. 42, no. 10, 2021.

[20] S. Meng *et al.*, "Noise reduction and feature enhancement for respiratory sound analysis," *International Journal of Biological Sciences*, vol. 15, no. 9, pp. 1921–1930, 2019.

[21] M. Tariq *et al.*, "Multi-class respiratory disease recognition using spectrogram-based CNN with data augmentation," *Applied Sciences*, vol. 12, 2022.

[22] B. Wang and Y. Sun, "Deep learning approaches for respiratory sound classification: A comparative study," *EURASIP Journal on Advances in Signal Processing*, 2024.

[23] K. Jayalakshmy and G. F. Sudha, "Scalogram-based prediction model for respiratory disorders using optimized CNN," *Artificial Intelligence in Medicine*, vol. 103, 2020.

[24] P. Carvalho, R. P. Paiva, and I. Chouvarda, "Respiratory sound analysis and applications in computer-aided diagnosis," *IEEE Reviews in Biomedical Engineering*, vol. 15, pp. 180–195, 2023.

[25] G. Serbes, S. Ulukaya, and Y. P. Kahya, "An automated lung sound preprocessing and classification system based on spectral analysis," IEEE Journal of Biomedical and Health Informatics, vol. 23, no. 2, pp. 867–874, 2019.

[26] T. A. Mesquita et al., "Respiratory sound classification using deep neural networks," Biomedical Signal Processing and Control, vol. 64, 2021.

[27] A. Rizal, S. H. Salleh, and S. S. Hameed, "Lung sound analysis using time–frequency features and CNN," Procedia Computer Science, vol. 163, pp. 538–545, 2019.

[28] J. A. Charlton et al., "Breath sound analysis for disease detection using deep learning," IEEE Access, vol. 8, pp. 202012–202021, 2020.

[29] Y. Perna and A. Tagarelli, "Deep auscultation: Predicting respiratory pathologies using CNNs," Applied Soft Computing, vol. 95, 2020.

[30] M. Aykanat, Ö. Kılıç, B. Kurt, and S. Saryal, "Classification of lung sounds using convolutional neural networks," EURASIP Journal on Audio, Speech, and Music Processing, 2017.

[31] J. Pahar et al., "COVID-19 cough and lung sound analysis using deep learning," IEEE Transactions on Instrumentation and Measurement, vol. 70, 2021.

[32] A. Abbas et al., "Respiratory sound classification for disease detection using deep CNN," Neural Computing and Applications, vol. 34, pp. 11921–11935, 2022.

[33] F. Demir, A. Sengur, and V. Bajaj, "Convolutional neural networks for lung sound classification," Health Information Science and Systems, vol. 8, no. 1, 2020.

[34] Y. Ren et al., "Automatic detection of abnormal respiratory sounds using deep learning," IEEE Sensors Journal, vol. 20, no. 9, pp. 4973–4982, 2020.

[35] H. Pasterkamp, S. S. Kraman, and G. R. Wodicka, "Respiratory sounds: Advances beyond the stethoscope," American Journal of Respiratory and Critical Care Medicine, vol. 156, 1997.

[36] D. Gurung et al., "Respiratory sound classification using MFCC and CNN," International Journal of Biomedical Engineering and Technology, vol. 32, 2020.

[37] A. N. Rizal and J. Hidayat, "Hybrid CNN-RNN model for lung disease detection," Journal of Medical Systems, vol. 45, 2021.

[38] S. Abbas et al., "Transfer learning for respiratory sound classification," Computers in Biology and Medicine, vol. 129, 2021.

[39] R. Liu et al., "Attention-based deep learning model for lung sound classification," Biomedical Signal Processing and Control, vol. 68, 2021.

[40] A. Kochetov et al., "Respiratory disease recognition using spectrogram images and CNN," Pattern Recognition Letters, vol. 138, pp. 160–166, 2020.

[41] J. Imran et al., "AI-based analysis of lung sounds for early disease detection," IEEE Access, vol. 9, pp. 117667–117676, 2021.

[42] T. Nguyen et al., "Deep feature fusion for respiratory sound classification," Sensors, vol. 22, no. 6, 2022.

[43] M. T. Islam et al., "Ensemble deep learning for lung sound classification," Expert Systems with Applications, vol. 182, 2021.

[44] Y. Wang et al., "Multi-resolution spectrogram analysis for respiratory disease detection," IEEE Signal Processing Letters, vol. 28, 2021.

[45] P. Harshavardhan et al., "Deep learning-based lung sound classification using Log-Mel features," International Journal of Medical Informatics, vol. 158, 2022.

[46] K. Minami et al., "Explainable AI for respiratory sound classification," Artificial Intelligence in Medicine, vol. 134, 2022.

[47] J. Pons et al., "End-to-end learning for respiratory sound analysis," IEEE Transactions on Audio, Speech, and Language Processing, vol. 28, 2020.

[48] A. B. R. Ahmed et al., "Noise-robust lung sound classification using deep learning," Biomedical Engineering Letters, vol. 11, 2021.

[49] L. Yang et al., "Respiratory disease detection using ensemble CNN models," Healthcare Analytics, vol. 2, 2022.

[50] S. Mishra et al., "Automated lung disease diagnosis using deep spectrogram learning," Journal of Healthcare Engineering, 2023.