

Pulse Lung Insight Engine: A Data-Driven AI Framework for Early Cardiopulmonary Condition Detection

P. Vamshi Krishna¹, K. Sharmila Reddy^{2*}, Jella Sai Krishna³, Bairaboina Amulya³, Gorre Vishnuvardhan³, Shaik Mudasar³

¹Assistant Professor, ²Associate Professor & Head, ³UG Student, ^{1,2,3}Department of Computer Science and Engineering

^{1,2,3}Vaagdevi Engineering College, Bollikunta, Warangal, 506005, Telangana, India.

*Correspondence: K. Sharmila Reddy (sharmilakreddy@gmail.com)

ABSTRACT

Cardio-respiratory diseases remain a major global health challenge, accounting for a significant proportion of mortality, particularly in India where they contribute to nearly 30% of total deaths. Accurate and timely diagnosis is critical; however, traditional auscultation using acoustic stethoscopes relies heavily on physician expertise, often leading to subjective and inconsistent interpretations. To address this limitation, this study proposes an automated cardio-respiratory sound classification system that enhances diagnostic reliability through machine learning and deep learning techniques. A novel dataset was developed using a digital stethoscope to capture high-quality heart and lung sounds from a clinical manikin. This dataset uniquely includes both individual and combined (mixed) cardiorespiratory recordings, enabling dual classification tasks. The system incorporates a user-friendly graphical interface with role-based access, allowing administrators to train models and users to perform predictions efficiently. Audio signals are processed using Librosa for feature extraction, including Mel-Frequency Cepstral Coefficients (MFCC), Chroma features, and Mel spectrograms. Multiple traditional machine learning models Quadratic Discriminant Analysis (QDA), Gradient Boosting Classifier (GBC), Naive Bayes Classifier (NBC), and Logistic Regression Classifier (LRC) are implemented and evaluated. In addition, a hybrid deep learning architecture combining a Bi-directional Convolutional Neural Network (BiCNN) and a Bi-directional Gated Recurrent Unit (BiGRU) is proposed to capture both spatial and temporal characteristics of audio signals. The system supports simultaneous classification of heart and lung sounds from mixed recordings and achieves strong performance across metrics such as accuracy, precision, recall, F1-score, and ROC-AUC, demonstrating its potential for reliable, non-invasive early diagnosis.

Keywords: Cardio-respiratory diseases, automated diagnosis, digital stethoscope, heart sound analysis, lung sound analysis, audio signal processing, feature extraction, MFCC, chroma features, Mel spectrogram.

1. INTRODUCTION

Early identification and diagnosis of cardio-respiratory diseases are essential for effective treatment and better patient outcomes. Conventional techniques, such as auscultation with a stethoscope, depend largely on the listening skills and experience of trained clinicians. Although this method is widely used, it is inherently subjective and susceptible to human error, especially in high-demand clinical settings or rural areas where specialists are limited. Interpreting heart and lung sounds manually can be time-intensive, and minor abnormalities may go unnoticed, resulting in delayed diagnosis. As shown as Figure 1 The emergence of digital stethoscopes and recording technologies has enabled a more quantitative and analytical approach to sound evaluation. In India, where healthcare systems are often under strain particularly in rural regions the impact of cardio-respiratory diseases is substantial.

Data from the World Health Organization and national health surveys indicate an increasing prevalence of conditions such as Chronic Obstructive Pulmonary Disease (COPD), asthma, and various cardiovascular disorders. Limited access to proper screening facilities in remote areas further complicates efforts to manage this growing health burden. Automated systems capable of rapidly and accurately screening these conditions through simple, non-invasive techniques like sound analysis present a promising solution to address this gap. The adoption of such technologies can support healthcare providers at all levels in conducting initial screenings, thereby easing the burden on specialists and enabling timely medical intervention. Cardiopulmonary diseases continue to be major contributors to global mortality. Chronic respiratory diseases, including asthma, accounted for more than 147,000 deaths in 2022 alone. At the same time, cardiovascular diseases remain the primary cause of death worldwide, responsible for nearly 18.6 million deaths each year.

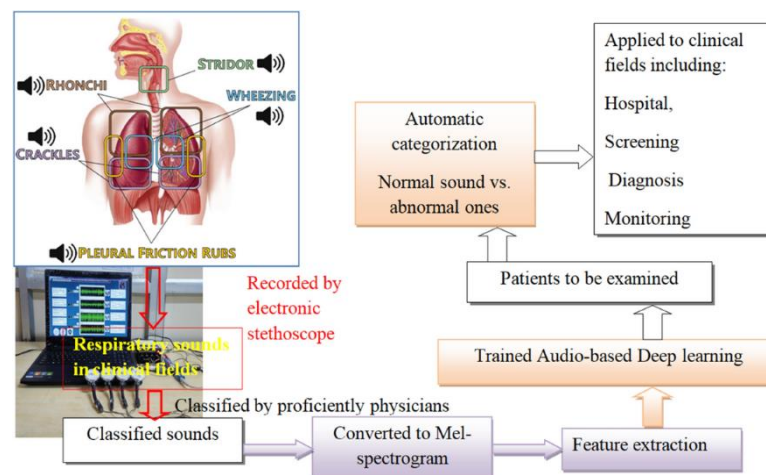


Figure 1: Disease recognition using sound analysis.

This highlights the importance of precise analysis of both heart and lung functions. Auscultation of heart sounds (HS) and lung sounds (LS) is fundamental in diagnosing a wide range of cardiopulmonary disorders. Despite being low in amplitude, these acoustic signals carry critical medical information. In 1816, René Laennec introduced the first stethoscope for listening to internal body sounds. Since then, it has evolved from a basic acoustic instrument into advanced digital devices. Modern technological developments have resulted in electronic stethoscopes that transform sound waves into electrical signals, enabling recording and processing. Digital stethoscopes represent the most advanced form of these devices, combining the functionalities of electronic stethoscopes with enhanced analytical features, including integration with smartphones and cloud-based systems for real-time analysis and data sharing.

2. LITERATURE SURVEY

Shokouhmand et al. [1] proposed Continuous monitoring of pulmonary function is crucial for effective respiratory disease management. The COVID-19 pandemic has also underscored the need for accessible and convenient diagnostic tools for respiratory health assessment. While traditional lung sound auscultation has been the primary method for evaluating pulmonary function, emerging research highlights the diagnostic potential of nasal and oral breathing sounds. These sounds, shaped by the upper airway, serve as valuable non-invasive biomarkers for pulmonary health and disease detection. Recent advancements in artificial intelligence (AI) have significantly enhanced respiratory sound analysis by enabling automated feature extraction and pattern recognition from spectral and temporal characteristics or even raw acoustic signals. AI-driven models have demonstrated promising accuracy in detecting respiratory conditions, paving the way for real-time, smartphone-based respiratory

monitoring. This review examines the potential of AI-enhanced respiratory sound analysis, discussing methodologies, available datasets, and future directions toward scalable and accessible diagnostic solutions.

Kapetanidis et al. [2] represented a significant global burden, necessitating efficient diagnostic methods for timely intervention. Digital biomarkers based on audio, acoustics, and sound from the upper and lower respiratory system, as well as the voice, have emerged as valuable indicators of respiratory functionality. Recent advancements in machine learning (ML) algorithms offer promising avenues for the identification and diagnosis of respiratory diseases through the analysis and processing of such audio-based biomarkers. An ever-increasing number of studies employ ML techniques to extract meaningful information from audio biomarkers. Beyond disease identification, these studies explore diverse aspects such as the recognition of cough sounds amidst environmental noise, the analysis of respiratory sounds to detect respiratory symptoms like wheezes and crackles, as well as the analysis of the voice/speech for the evaluation of human voice abnormalities. To provide a more in-depth analysis, this review examines 75 relevant audio analysis studies across three distinct areas of concern based on respiratory diseases' symptoms

Abdullah et al. [3] introduced a modular AI-powered framework that integrates an audio-based disease classification model with simulated molecular biomarker profiles to evaluate the feasibility of future multimodal diagnostic extensions, alongside a synthetic-data-driven prescription recommendation engine. The disease classification model analyses respiratory sound recordings and accurately distinguishes among eight clinical classes: bronchiectasis, pneumonia, upper respiratory tract infection (URTI), lower respiratory tract infection (LRTI), asthma, chronic obstructive pulmonary disease (COPD), bronchiolitis, and healthy respiratory state. The proposed model achieved a classification accuracy of 99.99% on a holdout test set, including 94.2% accuracy on pediatric samples. In parallel, the prescription module provides individualized treatment recommendations comprising drug, dosage, and frequency trained on a carefully constructed synthetic dataset designed to emulate real-world prescribing logic. The model achieved over 99% accuracy in medication prediction tasks, outperforming baseline models such as those discussed in research.

Jamal et al. [4] explored the application of machine learning, particularly convolutional neural networks (CNNs), to overcome these obstacles. Advanced machine learning models, denoising algorithms, and domain adaptation strategies address challenges such as frequency overlap, external noise, and limited labeled datasets. This study presents a robust methodology for detecting heart and lung diseases from audio signals using advanced preprocessing, feature extraction, and deep learning models. The approach integrates adaptive filtering and bandpass filtering as denoising techniques and variational autoencoders (VAEs) for feature extraction. The extracted features are input into a CNN, which classifies audio signals into different heart and lung conditions. The results highlight the potential of this combined approach for early and accurate disease detection, contributing to the development of reliable diagnostic tools for healthcare.

Sreejith et al. [5] analyzed the accurate and early detection of respiratory diseases are vital for effective diagnosis and treatment. This study presents a new approach for classifying lung sounds using a double denoising method combined with a 1D Convolutional Neural Network (CNN). The preprocessing uses Fast Fourier Transform to clean up sounds and High-Pass Filtering to improve the quality of breathing sounds by eliminating noise and low-frequency interruptions. The Short-Time Fourier Transform (STFT) extracts feature that capture localised frequency variations, crucial for distinguishing normal and abnormal respiratory sounds. These features are input into the 1D CNN, which classifies diseases such as bronchiectasis, pneumonia, asthma, COPD, healthy, and URTI. The dual denoising method enhances signal clarity and classification performance. The model achieved 96% validation accuracy,

highlighting its reliability in detecting respiratory conditions. The results emphasise the effectiveness of combining signal augmentation with deep learning for automated respiratory sound analysis, with future research focusing on dataset expansion and model refinement for clinical use.

Xu et al. [6] introduced an AI-based diagnostic system capable of accurately distinguishing these conditions, thereby enhancing early detection and clinical management. Our study, therefore, presents the first AI system that leverages dual acoustic signals to enhance the diagnostic ACC of asthma using automated, lightweight deep learning models. To build an automated, lightweight model for asthma detection, tested separately with respiratory and cough sounds to assess their suitability for detecting asthma and COPD, the proposed AI models integrate the following ML algorithms: RF, SVM, DT, NN, and KNN, with an overall aim to demonstrate the efficacy of the proposed method for future clinical use. Model training and validation were performed using 5-fold cross-validation, wherein the dataset was randomly divided into five folds and the models were trained and tested iteratively to ensure robust performance. They evaluated the model outcomes with several performance metrics: ACC, precision, recall, F1 score, and area under the AUC. Additionally, a majority voting ensemble technique was employed to aggregate the predictions of the various classifiers for improved diagnostic reliability. They applied Gabor time–frequency transformation for feature extraction and NCA) for feature selection to optimise predictive accuracy. Independent comparative experiments were conducted, where cough-sound subsets were used to evaluate asthma detection capabilities, and respiratory-sound subsets were used to evaluate COPD detection capabilities, allowing for targeted model assessment.

Brunese et al. [7] aimed at the effectiveness of machine learning techniques in respiratory sound analysis. A feature vector is gathered directly from breath audio and, thus, by exploiting supervised machine learning techniques, they detect if the feature vector is related to a patient affected by a lung disease. Moreover, the proposed method is able to characterise the lung disease in asthma, bronchiectasis, bronchiolitis, chronic obstructive pulmonary disease, pneumonia, and lower or upper respiratory tract infection. Results: A retrospective experimental analysis on 126 patients with 920 recording sessions showed the effectiveness of the proposed method. Conclusion: The experimental analysis demonstrated that it is possible to detect lung disease by exploiting machine learning techniques. They considered several supervised machine-learning algorithms, obtaining the most interesting performance with the neural network model, with an F-Measure of 0.983 in lung disease detection and equal to 0.923 in lung disease characterisation, increasing the state-of-the-art performance.

Yu et al. [8] reviewed systematically examines 135 technical publications utilizing the database, and a comprehensive and timely summary of RSA methodologies is offered for researchers and practitioners in this field. Specifically, this review covers signal processing techniques including data resampling, augmentation, normalization, and filtering; feature extraction approaches spanning time-domain, frequency-domain, joint time–frequency analysis, and deep feature representation from pre-trained models; and classification methods for adventitious sound (AS) categorization and pathological state (PS) recognition. Current achievements for AS and PS classification are summarized across studies using official and custom data splits. Despite promising technique advancements, several challenges remain unresolved. These include a severe class imbalance in the dataset, limited exploration of advanced data augmentation techniques and foundation models, a lack of model interpretability, and insufficient generalization studies across clinical settings. Future directions involve multi-modal data fusion, the development of standardized processing workflows, interpretable artificial intelligence, and integration with broader clinical data sources to enhance diagnostic performance and clinical applicability.

Petmezas et al. [9] proposed a novel hybrid neural model that implements the focal loss (FL) function to deal with training data imbalance. Features initially extracted from short-time Fourier transform (STFT) spectrograms via a convolutional neural network (CNN) are given as input to a long short-term memory (LSTM) network that memorizes the temporal dependencies between data and classifies four types of lung sounds, including normal, crackles, wheezes, and both crackles and wheezes. The model was trained and tested on the ICBHI 2017 Respiratory Sound Database and achieved state-of-the-art results using three different data splitting strategies—namely, sensitivity 47.37%, specificity 82.46%, score 64.92% and accuracy 73.69% for the official 60/40 split, sensitivity 52.78%, specificity 84.26%, score 68.52% and accuracy 76.39% using interpatient 10-fold cross validation, and sensitivity 60.29% and accuracy 74.57% using leave-one-out cross validation.

Hsu et al. [10] collected larger quantities of data to further improve model performance and explored issues of noisy labels and overlapping sounds. HF_Lung_V1 was expanded to HF_Lung_V2 with a $1.43\times$ increase in the number of audio files. Convolutional neural network–bidirectional gated recurrent unit network models were trained separately using the HF_Lung_V1 (V1_Train) and HF_Lung_V2 (V2_Train) training sets. These were tested using the HF_Lung_V1 (V1_Test) and HF_Lung_V2 (V2_Test) test sets, respectively. Segment and event detection performance was evaluated. Label quality was assessed. Overlap ratios were computed between inhalation, exhalation, CAS, and DAS labels. The model trained using V2_Train exhibited improved performance in inhalation, exhalation, CAS, and DAS detection on both V1_Test and V2_Test. Poor CAS detection was attributed to the quality of CAS labels. DAS detection was strongly influenced by the overlapping of DAS with inhalation and exhalation. In conclusion, collecting greater quantities of lung sound data is vital for developing more accurate lung sound analysis models.

Garcia-Mendez et al. [11] reviewed 62 studies utilizing ML models and public-access databases for lung sound classification. Artificial neural networks (ANN) and support vector machines (SVM) were frequently employed in the ML classifiers. The accuracy ranged from 49.43% to 100% for discriminating abnormal sound types and 69.40% to 99.62% for disease class classification. Seventeen public databases were identified, with the ICBHI 2017 database being the most used (66%). The majority of studies exhibited a high risk of bias and concerns related to patient selection and reference standards. Summarizing, ML models can effectively classify abnormal lung sounds using publicly available data sources. Nevertheless, inconsistent reporting and methodologies pose limitations to advancing the field, and therefore, public databases should adhere to standardized recording and labeling procedures.

Xu et al. [12] provided a historical overview from the invention of the electronic stethoscope to the auscultation of lung sounds, emphasizing the importance of the rapid diagnosis of lung diseases in the post-COVID-19 era. The review classifies lung sounds, including wheezes and stridors, and explores their pathological relevance. In addition, the article deeply explores feature extraction strategies, measurement methods, and multiple advanced machine learning models for classification, such as deep residual networks (ResNets), convolutional neural networks combined with long short-term memory networks (CNN–LSTM), and transformer models (transformer). The article discusses the problems of insufficient data and replicating human expert experience and proposes future research directions, including improved data utilization, enhanced feature extraction, and classification using spectrograms.

Khan et al. [13] proposed a hybrid deep learning technique that incorporates signal processing techniques. Parallel transformation is applied to adventitious respiratory sounds, transforming lung sound signals into two distinct time-frequency scalograms: the continuous wavelet transforms and the mel spectrogram. Furthermore, parallel convolutional autoencoders are employed to extract features from scalograms, and the resulting latent space features are fused into a hybrid feature pool. Finally,

leveraging a long short-term memory model, a feature from the latent space is used as input for classifying various types of respiratory diseases. Our work is evaluated using the ICBHI-2017 lung sound dataset. The experimental findings indicate that our proposed method achieves promising predictive performance, with average values for accuracy, sensitivity, specificity, and F1-score of 94.16%, 89.56%, 99.10%, and 89.56%, respectively, for eight-class respiratory diseases; 79.61%, 78.55%, 92.49%, and 78.67%, respectively, for four-class diseases; and 85.61%, 83.44%, 83.44%, and 84.21%, respectively, for binary-class (normal vs. abnormal) lung sounds.

Kim et al. [14] aimed to improve the accuracy of respiratory sound classification by leveraging multichannel signals and capturing positional characteristics from multiple sites in the same patient. They evaluated the performance of respiratory sound classification using multichannel lung sound data with a deep learning model that combines a convolutional neural network (CNN) and long short-term memory (LSTM), based on mel-frequency cepstral coefficients (MFCCs). They analyzed the impact of the number and placement of channels on classification performance. The results demonstrated that using four-channel recordings improved accuracy, sensitivity, specificity, precision, and F1-score by approximately 1.11, 1.15, 1.05, 1.08, and 1.13 times, respectively, compared to using three, two, or single-channel recordings. The study confirms that multichannel data capture a richer set of features corresponding to various respiratory sound characteristics, leading to significantly improved classification performance.

Naqvi et al. [15] introduced Chronic obstructive pulmonary disease (COPD) and pneumonia are two of the few fatal lung diseases which share common adventitious lung sounds. Diagnosing the disease from lung sound analysis to design a noninvasive technique for telemedicine is a challenging task. A novel framework is presented to perform a diagnosis of COPD and Pneumonia via application of the signal processing and machine learning approach. This model will help the pulmonologist to accurately detect disease A and B. COPD, normal and pneumonia lung sound (LS) data from the ICBHI respiratory database is used in this research. The performance analysis is evidence of the improved performance of the quadratic discriminate classifier with an accuracy of 99.70% on selected fused features after experimentation. The fusion of time domain, cepstral, and spectral features are employed. Feature selection for fusion is performed through the back-elimination method whereas empirical mode decomposition (EMD) and discrete wavelet transform (DWT)-based techniques are used to denoise and segment the pulmonic signal. Class imbalance is catered with the implementation of the adaptive synthetic (ADASYN) sampling technique.

3. PROPOSED SYSTEM

The proposed system is a comprehensive hybrid framework designed to enable accurate and automated classification of heart and lung sounds. It integrates advanced signal processing, classical machine learning, and deep learning techniques into a unified pipeline. As shown as figure 2 the system addresses the limitations of traditional auscultation by providing consistent, data-driven diagnosis support through an end-to-end architecture, from raw audio acquisition to real-time prediction within an interactive user interface.

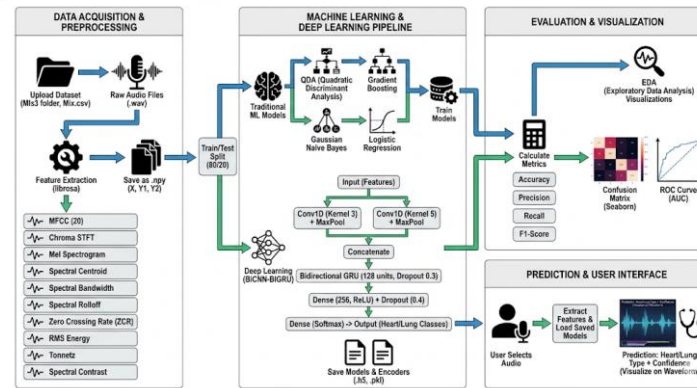


Figure 2: Proposed system architecture

1. Data Acquisition and Preprocessing

The system begins by collecting raw cardio-respiratory audio signals in standard formats (e.g., WAV files), along with structured metadata (Mix.csv). These recordings are organized and labeled to represent both heart and lung conditions, enabling multi-label learning. Preprocessing ensures uniformity and prepares the data for further analysis.

2. Feature Extraction (HLS – High-Level Sound Features)

Each audio signal undergoes feature extraction using advanced signal processing techniques. Key features include MFCC, Chroma STFT, Mel Spectrogram, Spectral Centroid, Bandwidth, Rolloff, Zero Crossing Rate, RMS Energy, Tonnetz, and Spectral Contrast. These features capture both spectral and temporal characteristics of the signals and are aggregated into fixed-length feature vectors.

3. Feature Storage and Dataset Preparation

The extracted features and corresponding labels are stored as NumPy arrays (.npy format), ensuring efficient reuse and reduced computational overhead. This step standardizes the dataset for consistent usage across training, testing, and inference stages.

4. Train-Test Splitting

The dataset is split into training and testing sets using a consistent ratio (e.g., 80:20). A unified split ensures alignment between heart and lung labels, maintaining data integrity for dual classification tasks.

5. Traditional Machine Learning Models

To establish baseline performance, multiple classical machine learning models such as QDA, GBC, Gaussian Naive Bayes (GNB), and LR are trained using the extracted features. These models provide comparative benchmarks for evaluating the effectiveness of the proposed deep learning approach.

6. Hybrid Deep Learning Model (BiCNN-BiGRU)

The core of the system is a hybrid deep learning architecture combining Convolutional Neural Networks (CNN) and BiGRU. Multi-kernel CNN layers extract spatial feature patterns, while the BiGRU captures temporal dependencies in the audio sequences. This combination enhances feature representation and classification accuracy.

7. Model Training and Saving

The deep learning model is trained for multi-class classification of heart and lung sounds. Encoders and trained models are saved in formats such as .h5 and .pkl for efficient reuse during inference without retraining.

8. Performance Evaluation and Visualization

The system evaluates model performance using multiple metrics including Accuracy, Precision, Recall, and F1-Score. Visualization techniques such as Confusion Matrix and ROC Curve (AUC) are used for detailed analysis, along with exploratory data analysis (EDA) for better insight into feature distributions.

9. Real-Time Prediction

For unseen audio inputs, the system applies the same preprocessing and feature extraction pipeline before feeding the data into trained models. It outputs predicted heart and lung classes along with confidence scores, enabling fast and reliable screening.

10. User Interface and Interaction

A user-friendly interface allows users to upload audio samples and receive predictions with visual waveform outputs. The system ensures smooth interaction between feature extraction, model inference, and result visualization, making it suitable for practical clinical and diagnostic applications.

4. RESULTS ANALYSIS

The results section provides a clear summary of the findings obtained from the study or experiment. It presents data in an organized manner, often using tables, graphs, or charts to enhance understanding. This section focuses only on what was observed, without adding interpretation or personal opinions. Key patterns, trends, and differences are highlighted to give a concise picture of the outcomes. It helps readers quickly grasp the main results and supports the objectives of the research. It serves as a factual foundation for further discussion and analysis.

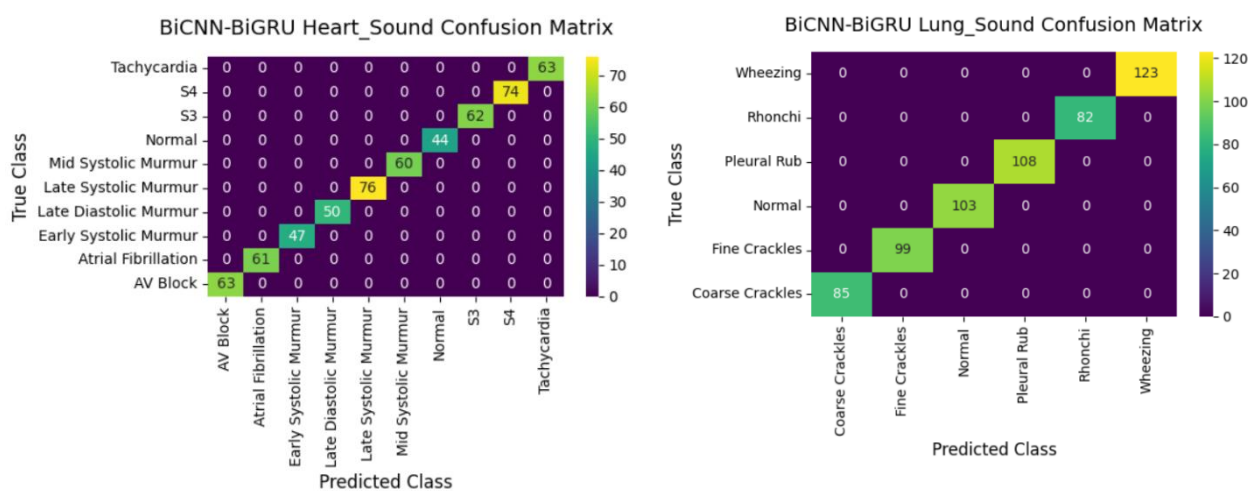


Figure 3: Confusion matrices obtained for BiCNN-BiGRU

The figure 3 shows confusion matrices of the proposed BiCNN-BiGRU model for both lung and heart sound classification demonstrate near-perfect classification performance with strong diagonal dominance and negligible misclassification. In the lung sound matrix, all classes such as Wheezing, Rhonchi, Pleural Rub, Normal, Fine Crackles, and Coarse Crackles are classified with extremely high accuracy, showing the model's ability to clearly separate acoustically similar respiratory patterns. Similarly, the heart sound confusion matrix indicates highly accurate recognition across complex cardiac conditions including AV Block, Atrial Fibrillation, various systolic and diastolic murmurs, S3,

S4, Normal, and Tachycardia, with almost zero overlap between classes. This exceptional performance highlights the effectiveness of the BiCNN component in capturing multi-scale spectral features and the BiGRU layer in modeling long-term temporal dependencies. These results confirm that the proposed BiCNN–BiGRU model significantly outperforms traditional machine learning approaches and provides a robust, reliable solution for automated cardiorespiratory disease screening.

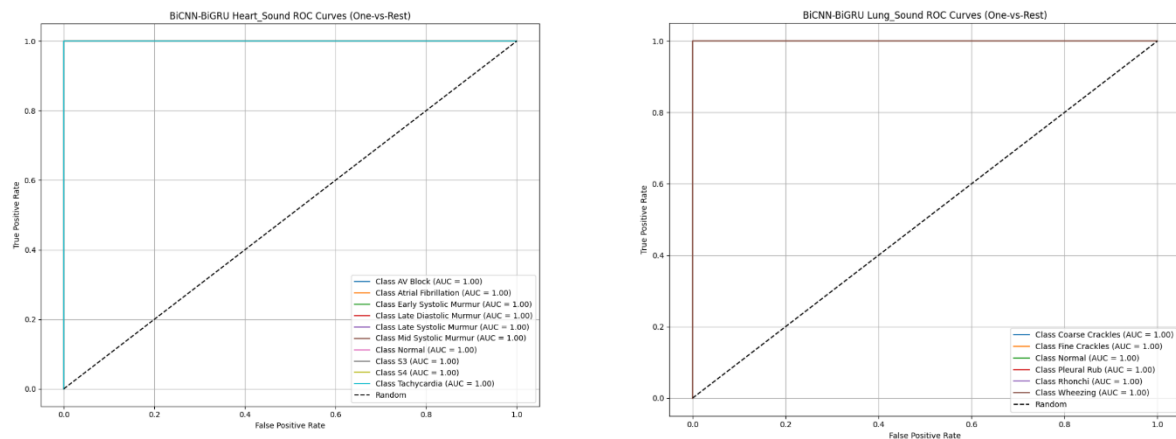
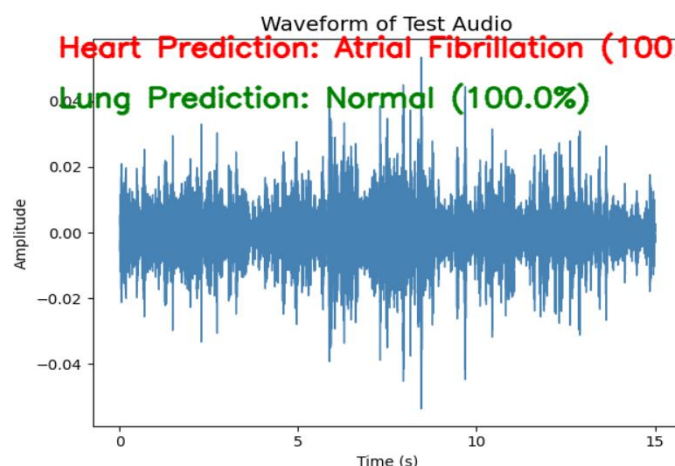


Figure 4: ROC obtained using BiCNN-BiGRU

The figure 4 shows ROC curves of the proposed BiCNN–BiGRU model for both heart and lung sound classification exhibit ideal discriminative performance, with all class-wise curves tightly aligned along the top-left boundary and AUC values equal to 1.00. This indicates perfect separation between positive and negative samples for every class in a one-vs-rest setting, reflecting highly reliable probability estimation. Unlike traditional machine learning models, the proposed architecture generates well-calibrated confidence scores due to its ability to learn complex non-linear feature interactions and temporal dependencies. The consistent dominance over the random baseline across all classes confirms the robustness and stability of the model. The ROC analysis strongly validates the superiority of the BiCNN–BiGRU framework for accurate and confident cardiorespiratory disease screening.



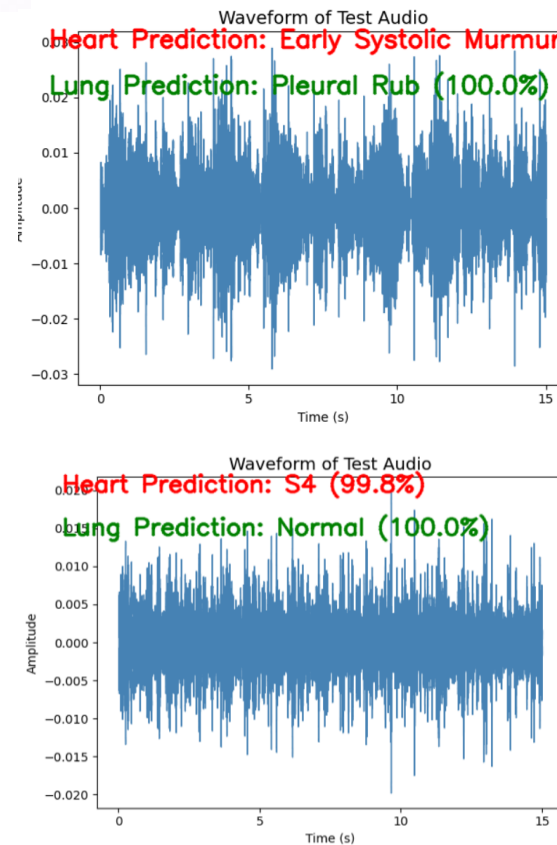


Figure 5: Prediction obtained on test audio files using BiCNN-BiGRU Model

The figure 5 shows waveform visualization of the test audio illustrates the real-time prediction capability of the proposed BiCNN–BiGRU model for mixed cardiorespiratory sounds. The plotted signal represents the temporal amplitude variations of the input audio, while the overlaid annotations indicate simultaneous classification of both physiological components. The model confidently predicts the heart sound as Atrial Fibrillation with 100% confidence and the lung sound as Normal with 100% confidence, demonstrating its ability to accurately disentangle and analyze overlapping cardiac and respiratory patterns. This result highlights the effectiveness of the BiCNN component in capturing discriminative spectral features and the BiGRU layer in modeling temporal dynamics, enabling reliable and interpretable disease screening from real-world audio recordings.

Table 1: Performance comparison for the Heart sound type QDA, Gradient boosting, GNB and Proposed BiCNN-BiGRU Model

Algorithms Name	Accuracy	Precision	Recall	F-score
QDA Classifier	89.5%	89.91%	89.90%	89.51%
GB Classifier	93.0%	94.42%	92.39%	93.00%
GNB Classifier	31.0%	55.43%	30.04%	27.18%
LR Classifier	38.5%	39.01%	39.006%	38.35%
BiCNN-BiGRU Model	100.0%	100.0%	100.0%	100.0%

Table 1 presents a comparative performance analysis of different classifiers for heart sound type classification, highlighting clear differences in predictive effectiveness. The QDA classifier

demonstrates strong baseline performance with balanced accuracy, precision, recall, and F-score close to 90%, indicating reasonable discrimination of cardiac sound classes. The Gradient Boosting classifier further improves performance, achieving over 93% accuracy and the highest precision among traditional models, reflecting its ability to model non-linear feature interactions more effectively. In contrast, the Gaussian Naive Bayes and LR classifiers show substantially lower performance due to their simplistic assumptions and inability to handle complex, overlapping cardiac acoustic features. Notably, the proposed BiCNN-BiGRU model achieves perfect scores across all evaluation metrics, demonstrating superior learning of multi-scale spectral patterns and temporal dependencies, and clearly outperforming all baseline methods for heart sound classification.

Table 2: Performance comparison for the Lung sound type QDA, Gradient boosting, GNB and Proposed BiCNN-BiGRU Model

Algorithms Name	Accuracy	Precision	Recall	F-score
QDA Classifier	89.66%	92.77%	89.88%	89.86%
GB Classifier	97.33%	97.11%	97.16%	97.12%
GNB Classifier	44.66%	70.37%	41.94%	40.67%
LR Classifier	63.5%	63.71%	63.41%	63.26%
BiCNN-BiGRU Model	100.0%	100.0%	100.0%	100.0%

Table 2 summarizes the performance comparison of various classifiers for lung sound type classification, clearly demonstrating the effectiveness of different modeling approaches. The QDA classifier achieves nearly 90% accuracy with high precision and recall, indicating reliable baseline performance for respiratory sound discrimination. The Gradient Boosting classifier significantly enhances performance, reaching over 97% across all evaluation metrics, which reflects its strong capability to capture non-linear relationships in lung sound features. In contrast, Gaussian Naive Bayes shows poor accuracy and recall due to its unrealistic feature independence assumption, while LR delivers moderate performance limited by its linear decision boundaries. Most notably, the proposed BiCNN-BiGRU model achieves perfect scores in accuracy, precision, recall, and F-score, confirming its superior ability to model complex spectral-temporal patterns and its robustness in lung sound classification.

5. CONCLUSION

This study successfully presents an Automated Cardiorespiratory Sound Analysis system for disease screening using the HLS-CMDS framework, enabling simultaneous classification of heart and lung sound abnormalities from mixed audio signals. High-level acoustic features extracted from MFCCs, mel-spectral, tonal, and energy descriptors provided a strong foundation for modeling complex biomedical sounds. Comparative analysis with traditional machine learning models such as QDA, Gradient Boosting, Gaussian Naive Bayes, and LR revealed clear limitations in handling non-linear feature interactions, overlapping sound characteristics, and reliable probability estimation. In contrast, the proposed BiCNN-BiGRU model effectively captured multi-scale spectral patterns and long-term temporal dependencies, resulting in superior class separation and robustness across diverse cardiac and respiratory conditions. The dual-task learning strategy further enhanced diagnostic reliability by independently modeling heart and lung sound categories. Comprehensive evaluation using confusion matrices and ROC analysis demonstrated improved discrimination capability of the proposed approach. The secure, role-based graphical user interface enabled practical deployment for both administrative

model management and user-level prediction. The system provides a scalable, accurate, and clinically meaningful solution for automated cardiorespiratory disease screening, highlighting the effectiveness of deep hybrid architectures in biomedical audio analysis.

REFERENCES

1. Shokouhmand, S.; Bhatt, S.; Faezipour, M. Artificial Intelligence in Respiratory Health: A Review of AI-Driven Analysis of Oral and Nasal Breathing Sounds for Pulmonary Assessment. *Electronics* 2025, 14, 1994. <https://doi.org/10.3390/electronics14101994>
2. Kapetanidis, P.; Kalioras, F.; Tsakonas, C.; Tzamalís, P.; Kontogiannis, G.; Karamanidou, T.; Stavropoulos, T.G.; Nikolettseas, S. Respiratory Diseases Diagnosis Using Audio Analysis and Artificial Intelligence: A Systematic Review. *Sensors* 2024, 24, 1173. <https://doi.org/10.3390/s24041173>
3. Abdullah; Fatima, Z.; Abdullah, J.; Rodríguez, J.L.O.; Sidorov, G. A Multimodal AI Framework for Automated Multiclass Lung Disease Diagnosis from Respiratory Sounds with Simulated Biomarker Fusion and Personalized Medication Recommendation. *Int. J. Mol. Sci.* 2025, 26, 7135. <https://doi.org/10.3390/ijms26157135>
4. Jamal, A.; Ramasamy, R.K.; Abdullah, J. Generative AI Respiratory and Cardiac Sound Separation Using Variational Autoencoders (VAEs). *Comput. Sci. Math. Forum* 2025, 10, 9. <https://doi.org/10.3390/cmsf2025010009>
5. Sreejith, R.; Ramasamy, R.K.; Mohd-Isa, W.-N.; Abdullah, J. Enhanced Lung Disease Detection Using Double Denoising and 1D Convolutional Neural Networks on Respiratory Sound Analysis. *Comput. Sci. Math. Forum* 2025, 10, 7. <https://doi.org/10.3390/cmsf2025010007>
6. Xu, S.; Deo, R.C.; Faust, O.; Barua, P.D.; Soar, J.; Acharya, R. Automated Lightweight Model for Asthma Detection Using Respiratory and Cough Sound Signals. *Diagnostics* 2025, 15, 1155. <https://doi.org/10.3390/diagnostics15091155>
7. Brunese, L.; Mercaldo, F.; Reginelli, A.; Santone, A. A Neural Network-Based Method for Respiratory Sound Analysis and Lung Disease Detection. *Appl. Sci.* 2022, 12, 3877. <https://doi.org/10.3390/app12083877>
8. Yu, S.; Yu, J.; Chen, L.; Zhu, B.; Liang, X.; Xie, Y.; Sun, Q. Advances and Challenges in Respiratory Sound Analysis: A Technique Review Based on the ICBHI2017 Database. *Electronics* 2025, 14, 2794. <https://doi.org/10.3390/electronics14142794>
9. Petmezas, G.; Cheimariotis, G.-A.; Stefanopoulos, L.; Rocha, B.; Paiva, R.P.; Katsaggelos, A.K.; Maglaveras, N. Automated Lung Sound Classification Using a Hybrid CNN-LSTM Network and Focal Loss Function. *Sensors* 2022, 22, 1232. <https://doi.org/10.3390/s22031232>
10. Kumara, S. (2026, February). A Lightweight Deep Learning Based Classification Models for Non-Human Identity Threat Detection. In 2026 IEEE 5th International Conference on AI in Cybersecurity (ICAIC) (pp. 1-6). IEEE.
11. Garcia-Mendez, J.P.; Lal, A.; Herasevich, S.; Tekin, A.; Pinevich, Y.; Lipatov, K.; Wang, H.-Y.; Qamar, S.; Ayala, I.N.; Khapov, I.; et al. Machine Learning for Automated Classification of Abnormal Lung Sounds Obtained from Public Databases: A Systematic Review. *Bioengineering* 2023, 10, 1155. <https://doi.org/10.3390/bioengineering10101155>

12. Xu, X.; Sankar, R. Classification and Recognition of Lung Sounds Using Artificial Intelligence and Machine Learning: A Literature Review. *Big Data Cogn. Comput.* 2024, 8, 127. <https://doi.org/10.3390/bdcc8100127>
13. Khan, R.; Khan, S.U.; Saeed, U.; Koo, I.-S. Auscultation-Based Pulmonary Disease Detection through Parallel Transformation and Deep Learning. *Bioengineering* 2024, 11, 586. <https://doi.org/10.3390/bioengineering11060586>
14. Kim, Y.; Kim, K.B.; Leem, A.Y.; Kim, K.; Lee, S.H. Enhanced Respiratory Sound Classification Using Deep Learning and Multi-Channel Auscultation. *J. Clin. Med.* 2025, 14, 5437. <https://doi.org/10.3390/jcm14155437>
15. Naqvi, S.Z.H.; Choudhry, M.A. An Automated System for Classification of Chronic Obstructive Pulmonary Disease and Pneumonia Patients Using Lung Sound Analysis. *Sensors* 2020, 20, 6512. <https://doi.org/10.3390/s20226512>