International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 07 , July - 2026

Intelligent Classification of Respiratory Diseases Using Machine Learning-Based Lung Sound Analysis

Tara V K[1], Varsha S[2]

[1]Assistant Professor, , Dept. of Computer Science and Engineering, City Engineering College, Bangalore, Karnataka, India-560062

[2]UG Student (IV Year), Computer Science and Engineering, City Engineering College, Bengaluru, Karnataka, India-560062

***Abstract* - Respiratory diseases such as asthma, chronic obstructive pulmonary disease (COPD), pneumonia, bronchiectasis, bronchiolitis and upper respiratory tract infection (URTI) remain among the leading causes of illness and death worldwide. Conventional diagnosis relies heavily on auscultation and spirometry, both of which depend on clinician experience and are susceptible to human error. This paper presents a machine learning-based system for automated respiratory disease classification using lung sound recordings. Audio signals are resampled, segmented into temporal snippets, denoised through DFT-based baseline wander removal and amplitude-normalized, before being converted into log-scaled Mel spectrograms and Gammatone Cepstral Coefficient (GTCC) features. A lightweight Inception-based deep learning model (RDsLINet) classifies the processed signals into seven categories — Asthma, Bronchiectasis, Bronchiolitis, COPD, Healthy, Pneumonia and URTI. Complementary experiments using a Convolutional Neural Network and a ResNet-50 transfer-learning model on chest X-ray images were also carried out for comparison. The proposed audio-based pipeline achieved a classification accuracy of up to 99.22%, while the image-based ResNet-50 model achieved 96.8% accuracy. The trained model is deployed through a lightweight web interface that allows a user to upload a lung sound recording and receive a real-time diagnostic prediction with a downloadable report. The results suggest that combining spectral audio features with a lightweight deep-learning architecture can provide a fast, non-invasive and scalable decision-support tool for respiratory-disease screening.**

Keywords - Respiratory Disease Classification; Lung Sound Analysis; Machine Learning; Deep Learning; Mel Spectrogram; Convolutional Neural Network; GTCC.

I. INTRODUCTION

Respiratory diseases are among the leading causes of death and disability globally, with the greatest burden falling on the poorest regions of the world. Chronic obstructive pulmonary disease (COPD) alone affects an estimated 65 million people and was responsible for 3.91 million deaths in 2017, accounting for roughly 7% of deaths worldwide. Asthma affects about 334 million people and remains the most common chronic disease of childhood. Pneumonia is a leading cause of death among children under five, tuberculosis affects over 10 million people annually, and lung cancer causes 1.6 million deaths each year. Collectively, more than one billion people suffer from acute or chronic respiratory conditions, and roughly four million people die prematurely from chronic respiratory disease every year.

Auscultation with a stethoscope remains the most widely used, low-cost, and non-invasive screening method, but it depends on the experience and hearing acuity of the examining physician and is vulnerable to ambient noise and instrument calibration issues. Advances in digital signal processing, machine learning and deep learning now make it possible to analyse respiratory sounds automatically, offering faster, more consistent and more objective screening. This paper proposes a machine-learning pipeline that converts raw lung-sound recordings into spectral features and classifies them using a lightweight deep-learning model, with the goal of supporting — not replacing — clinical decision-making.

II. LITERATURE SURVEY

Early work on automated lung-sound analysis relied on shallow-learning techniques such as Support Vector Machines (SVM), k-Nearest Neighbours (k-NN), Gaussian Mixture Models (GMM) and Artificial Neural Networks (ANN), typically applied to hand-crafted features such as Mel-Frequency Cepstral Coefficients (MFCC) or Gammatone Cepstral Coefficients (GTCC) [2], [6]–[8]. These approaches are computationally inexpensive and interpretable, and have reported classification accuracies ranging from roughly 85% to 96% depending on the dataset, the number of disease classes and the recording conditions used.

More recent studies have applied deep-learning architectures — including Convolutional Neural Networks (CNNs), hybrid CNN-RNN models, and residual networks — directly to spectrogram representations of lung sounds, removing the need for manual feature engineering. Reported accuracies for these deep models vary widely (roughly 40% to 99%) depending largely on dataset size, class imbalance and cross-validation strategy, since public respiratory-sound datasets such as ICBHI 2017 are relatively small and imbalanced. Across both shallow- and deep-learning studies,

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 07 , July - 2026

a recurring limitation is that most systems are trained and evaluated on a single, static modality (either audio or imaging), are computationally heavy, and are rarely validated for real-time or resource-constrained deployment. This motivates the lightweight, deployable pipeline proposed in this paper.

III. PROBLEM STATEMENT

Existing automated respiratory-diagnosis systems, while promising, exhibit several practical limitations:

- Limited use of sequential / time-series information such as continuous spirometry or monitoring data.
- Dependence on large, well-labelled datasets, with degraded performance on imbalanced disease classes.
- Poor generalisation across different recording devices, patient populations and clinical settings.
- Limited interpretability, which reduces clinician trust in automated predictions.
- High computational cost, making real-time deployment on portable or low-resource devices difficult.

IV. PROPOSED METHODOLOGY

A. Input Acquisition

Lung-sound recordings are acquired using a digital stethoscope or microphone and stored as the raw waveform LS[n], which forms the input to the pipeline.

B. Preprocessing

Each recording is resampled to a standard rate of 4 kHz to remove device-dependent variability. The continuous signal is then divided into fixed-length temporal snippets so that the model can focus on localised, disease-relevant patterns. Discrete Fourier Transform (DFT)-based baseline-wander removal eliminates low-frequency noise caused by sensor movement or ambient interference, and the resulting signal is amplitude-normalised to standardise volume across recordings and equipment.

C. Feature Extraction

The cleaned snippets are converted into log-scaled Mel spectrograms, which capture both time- and frequency-domain characteristics of the sound and emphasise low-energy components that may indicate subtle abnormalities. Gammatone Cepstral Coefficients (GTCC), which approximate human auditory perception, are extracted in parallel to provide a complementary, compact feature representation.

D. Classification

The extracted features are passed to RDsLINet, a lightweight Inception-style convolutional network designed specifically for this task. Its multi-branch convolutional blocks capture patterns at multiple scales while keeping the parameter count low enough for deployment on resource-constrained hardware. The network outputs a probability distribution over seven classes: Asthma, Bronchiectasis, Bronchiolitis, COPD, Healthy, Pneumonia and URTI.

E. Deployment

The trained model is served through a lightweight Flask-based web interface. A user uploads a lung-sound file, the backend performs preprocessing and feature extraction, and the model returns a predicted class with a confidence score and a downloadable report, enabling near real-time use in a clinical or screening setting.

V. SYSTEM ARCHITECTURE

The end-to-end pipeline can be summarized as six sequential stages, outlined in Table I.

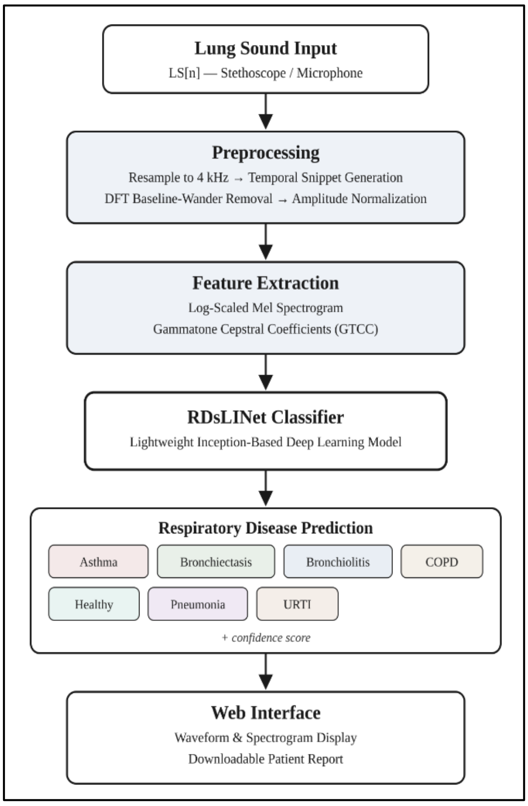

FIG I. OVERALL ARCHITECTURE OF THE PROPOSED RESPIRATORY DISEASE CLASSIFICATION FRAMEWORK.

VI. IMPLEMENTATION

The system was implemented in Python. Pandas and NumPy were used for data handling, Scikit-learn for classical baselines and evaluation utilities, and TensorFlow/Keras for building and training the CNN, RDsLINet and ResNet-50 models. OpenCV was used to pre-process the supplementary chest X-ray images used for comparison, and Matplotlib/Seaborn were used to visualise training curves and confusion matrices. The web interface was built with Flask, allowing a clinician to upload a recording and view the waveform, spectrogram and predicted class in the browser.

VII. RESULTS AND DISCUSSION

The models were evaluated using an 80/20 train-test split together with k-fold cross-validation, and assessed using accuracy, precision, recall, F1-score and ROC-AUC. Table II summarises the best results obtained for each approach evaluated in this work.

Model	Data	Accuracy
GTCC + shallow classifier	Lung audio	99.22%
RDsLINet	Mel spectrogram	High
ResNet-50 (transfer learning)	Chest X-ray	96.80%

TABLE II. COMPARISON OF MODEL PERFORMANCE

In a representative deployment test, the system correctly flagged a sample recording as COPD with a confidence of approximately 99.99%, and the web interface displayed the waveform, log-mel spectrogram and downloadable report within seconds of upload. These results indicate that a lightweight, audio-driven pipeline can achieve accuracy competitive with heavier image-based transfer-learning models, while remaining suitable for near real-time, low-resource deployment. As with all machine-learning-based diagnostic aids, the reported figures reflect performance on the specific datasets used and should be validated on larger, independent and multi-site datasets before any clinical use.

VIII. CONCLUSION AND FUTURE SCOPE

This paper presented a machine-learning pipeline for classifying respiratory diseases from lung-sound recordings, combining DFT-based preprocessing, log-Mel/GTCC feature extraction and a lightweight Inception-based classifier (RDsLINet), deployed through a simple web interface for near real-time screening. The audio-based approach achieved up to 99.22% accuracy, comparable to a ResNet-50 transfer-learning baseline trained on chest X-rays. Future work will focus on: (i) multimodal fusion of audio, imaging and clinical data; (ii) training on larger and more diverse datasets to improve generalization; (iii) handling class imbalance through oversampling or weighted losses; (iv) incorporating explainability techniques such as LIME or SHAP to improve clinician trust; and (v) integrating the system with electronic health record (EHR) platforms for real-world clinical deployment.

ACKNOWLEDGMENT

The authors thank the Department of Computer Science and Engineering, City Engineering College, Bengaluru, for the resources and support extended throughout this work.

REFERENCES

[1] T. Lallukka, A. Millear, A. Pain, M. Cortinovis, and G. Giussani, "GBD 2015 Mortality and Causes of Death Collaborators: Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015," Lancet, vol. 389, no. 10064, p. E1, 2017.

[2] I. Sen, M. Saraclar, and Y. P. Kahya, "A comparison of SVM and GMM-based classifier configurations for diagnostic classification of pulmonary sounds," IEEE Trans. Biomed. Eng., vol. 62, no. 7, pp. 1768–1776, 2015.

[3] M. Milicevic, I. Mazic, and M. Bonkovic, "Classification accuracy comparison of asthmatic wheezing sounds recorded under ideal and real-world conditions," in Proc. 15th Int. Conf. Artificial Intelligence, Knowledge Engineering and Databases (AIKED), Venice, 2016.

[4] B. M. Rocha, L. Mendes, I. Chouvarda, P. Carvalho, and R. P. Paiva, "Detection of cough and adventitious respiratory sounds in audio recordings by internal sound analysis," in Proc. Int. Conf. Biomedical and Health Informatics, Springer, 2017, pp. 51–55.

[5] R. B. Urquhart, J. McGhee, J. E. S. Macleod, S. W. Banham, and F. Moran, "The diagnostic value of pulmonary sounds: a preliminary study by computer-aided analysis," Comput. Biol. Med., vol. 11, no. 3, pp. 129–139, 1981.

[6] M. Bahoura and C. Pelletier, "Respiratory sounds classification using cepstral analysis and Gaussian mixture models," in Proc. 26th Annu. Int. Conf. IEEE EMBS, vol. 1, 2004, pp. 9–12.

[7] P. Mayorga, C. Druzgalski, R. L. Morelos, O. H. Gonzalez, and J. Vidales, "Acoustics based assessment of respiratory diseases using GMM classification," in Proc. Annu. Int. Conf. IEEE EMBS, 2010, pp. 6312–6316.

[8] R. Palaniappan, K. Sundaraj, and S. Sundaraj, "A comparative study of the SVM and k-NN machine learning algorithms for the diagnosis of respiratory pathologies using pulmonary acoustic signals," BMC Bioinformatics, vol. 15, no. 1, pp. 1–8, 2014.

[9] W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," Bull. Math. Biophys., vol. 5, no. 4, pp. 115–133, 1943.

[10] S. K. Chowdhury and A. K. Majumder, "Frequency analysis of adventitious lung sounds," J. Biomed. Eng., vol. 4, no. 4, pp. 305–312, 1982.

[11] Z. Korona and M. M. Kokar, "Lung sound recognition using model-theory-based feature selection and fusion," Appl. Signal Process., vol. 5, no. 3, p. 152, 1998.