IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940



International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 11, November 2026

DOI: 10.17148/IJARCCE.2026.1511159

Deep Learning-Based Early COPD Detection and Severity Classification Using Respiratory Sound Analysis

Rajath Kiran A[1], Rithesh[2], Sanath K[3], Sheethal D Rai[4], Pramod Kumar P M[5]

Student, Department of Computer Science and Engineering, Vivekananda College of Engineering and Technology, Puttur, Karnataka, India[1-4]

Associate Professor (Guide), Department of Computer Science and Engineering, Vivekananda College of Engineering and Technology, Puttur, Karnataka, India[5]

Abstract: Chronic Obstructive Pulmonary Disease (COPD) is normally confirmed by spirometry, which needs calibrated equipment and trained staff that are scarce in rural and resource-limited settings. This paper presents a two-model deep learning framework for respiratory sound analysis on the public ICBHI 2017 database. Model 1 is a binary COPD screening network that passes a 128 × 157 log-Mel spectrogram through a frozen EfficientNetB0 backbone, a bidirectional LSTM and four-head self-attention. Model 2 is a four-stage residual CNN that assigns each annotated respiratory cycle to one of four coded acoustic categories, copd0 to copd3. The models are joined in a cascade in which Model 2 runs only when Model 1 predicts COPD. On a patient-level held-out test set of 1255 cycles, Model 1 reached 87.73 percent accuracy, 88.38 percent sensitivity, 83.80 percent specificity and an AUC of 0.9313. Model 2 reached 91.58 percent four-class accuracy and an ICBHI score of 92.56 percent. Because the binary test set is dominated by COPD cycles, Model 1 accuracy is only modestly above the 85.74 percent majority-class rate, and the four classes are acoustic labels rather than clinical severity grades. A React and Express web interface exposes the pipeline with an explicit research disclaimer.

Keywords: COPD detection; respiratory sound analysis; ICBHI 2017; log-Mel spectrogram; EfficientNetB0; BiLSTM; multi-head attention; residual CNN; deep learning

I. INTRODUCTION

Chronic Obstructive Pulmonary Disease (COPD) is a progressive, non-reversible respiratory condition marked by persistent airflow limitation, mainly caused by long-term exposure to tobacco smoke and biomass fuel. It is the third leading cause of death worldwide, accounts for over three million deaths a year and affects more than 300 million people. The disease advances silently and is often diagnosed late, when treatment options are already limited.

Spirometry is the established test for COPD and measures airflow obstruction through the FEV1/FVC ratio. It needs calibrated equipment and trained operators, which makes wide early screening impractical in rural, remote and resource-constrained settings. Respiratory sounds such as wheezes and crackles carry acoustic evidence of airway disease and can be captured with a digital stethoscope or a microphone. Deep learning can learn this evidence directly from spectrograms, and public datasets such as the ICBHI 2017 Respiratory Sound Database make such work reproducible.

This work builds a research aid for acoustic COPD screening. A binary model first decides whether a recording shows COPD-related evidence. A second model then assigns the recording to one of four coded ICBHI acoustic categories. The main contributions are: (1) an EfficientNetB0, bidirectional LSTM and multi-head attention network for binary COPD screening; (2) a compact four-stage residual CNN for four-class respiratory cycle classification; (3) a cascade that runs the second model only after a positive screen, served through a web interface that shows the decision path and a research disclaimer; and (4) an evaluation that reports the class imbalance of the test set and the meaning of the four labels, so that the headline figures are not over-read.

 This work is licensed under a Creative Commons Attribution 4.0 International License



International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 11, November 2026
DOI: 10.17148/IJARCCE.2026.1511159

II. LITERATURE REVIEW

A. CNN-Based Respiratory Sound Analysis

Deep learning on the ICBHI corpus began with spectrogram CNNs. Srivastava et al. [1] trained a CNN on MFCC and Melspectrogram features for COPD detection, using augmentation to cope with the small corpus, but treated each breathing cycle as a static image and produced only a binary output. Alghamdi et al. [2] varied MFCC coefficients, FFT window and hop length in a deep CNN and reached 97.4 percent validation accuracy, although results varied across patient groups. Naqvi and Choudhry [3] fused MFCC, chroma and spectral contrast features to separate COPD, pneumonia and healthy sounds with 99.70 percent accuracy under single-device conditions. Medina-Ramos et al. [4] showed that a custom CNN trained from scratch can outperform ImageNet-pretrained models for early COPD identification, while Surekha et al. [5] found that transfer learning raised sensitivity but suffered from the mismatch between natural images and spectrograms.

B. Temporal and Hybrid Models

Because breathing is a time-varying signal, later studies added recurrent layers. Fraiwan et al. [6] compared CNN and LSTM models for pulmonary sound recognition and found that LSTMs capture time-varying wheeze patterns better than static CNNs, and Demir et al. [7] studied CNN-based lung sound classification on the ICBHI data. Zhang et al. [8] applied LSTMs to Melspectrogram sequences and reported 98.82 percent accuracy, and Marepalli et al. [9] showed that lightweight CNN and LSTM models can run in low-resource settings. Sreejith et al. [10] proposed CALMNet, a TimeDistributed CNN-LSTM that reached 97.65 percent on multi-class lung disease classification. Reksohartono et al. [11] reported gradient instability in plain RNNs on long respiratory sequences, which supports the use of gated LSTMs. Islam et al. [12] applied WaveNet to raw waveforms with 95 percent accuracy at high computational cost, and Koishiyeva et al. [13] used multi-head attention to reach 93.8 percent while still confusing COPD with pneumonia.

C. Advanced Architectures and Severity Estimation

Recent work pushes accuracy with heavier models. Roy and Satija [14] reported 99.81 percent accuracy with a multi-head self-organized operational network, Roy et al. [15] fused three spectrogram types with Kolmogorov-Arnold networks, and Sait et al. [16] combined a CNN-ViT hybrid with Kolmogorov-Arnold networks for 97.62 percent accuracy at a high computational cost. Severity estimation is much rarer. Roy and Satija [17] used Melspectrogram snippets for four-class severity prediction, and Sahu et al. [18] proposed SOUNDNet for four-class COPD severity classification with 95.76 percent accuracy; neither aligned its classes with clinical severity grades. Altan et al. [19], [20] estimated severity from 12-channel lung sound recordings, which needs specialised hardware that primary care rarely has.

D. Other Approaches

Classical machine learning on acoustic features has also been used for severity prediction [21]. COPD Scope [22] links severity grading to clinical pathways using rule-based logic, and Haider et al. [23] and Taloba and Matoog [24] depend on spirometry data alongside or instead of sound. Further studies used deep belief networks on difference plots [25], a two-stage event and disease pipeline [26], k-NN with wrapper feature selection [27], PPG-derived respiration [28], CycleGAN-based augmentation [29] and multi-task learning [30].

The reviewed work leaves three gaps. Most systems give only a binary COPD output. Severity-oriented systems either need multi-channel hardware or do not align their classes with clinical grades. And few studies combine spatial feature extraction, temporal modelling and attention in a screening-then-staging cascade while reporting how class imbalance affects their accuracy figures.

III. METHODS AND MATERIALS

A. ICBHI 2017 Dataset and Labels

Both models use the ICBHI 2017 Respiratory Sound Database. Each recording has annotation files giving the start and end time of every breathing cycle together with its coded acoustic label. Patient identifiers are extracted from file names with regular expressions, and confirmed diagnoses are read from patient_diagnosis.txt. For Model 1 each cycle is labelled COPD (1) or NON-COPD (0). For Model 2 the annotations map to four coded classes, copd0, copd1, copd2 and copd3, which follow the ICBHI Normal, Crackle, Wheeze and Both categories; copd3 denotes cycles with both crackles and wheezes. These are sound-event labels. They are not spirometry-based severity grades.

B. Data Split

To prevent leakage, Model 1 splits the data by patient identifier and not by cycle. All cycles of one patient fall in exactly one of the training (70 percent), validation (15 percent) or test (15 percent) partitions. Class weights inversely proportional

 This work is licensed under a Creative Commons Attribution 4.0 International License

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 11, November 2026

DOI: 10.17148/IJARCCE.2026.1511159

to class frequency are computed to offset the imbalance between COPD and non-COPD cycles. Model 2 was trained and evaluated under its own split protocol, and its held-out partition contains 1033 cycles.

C. Audio Preprocessing

Raw variable-length recordings go through three steps. Cycles are sliced out using the annotation timestamps and resampled to 16 kHz, since respiratory sound content lies mostly below 2 kHz. Each cycle is then divided by its peak absolute value so that amplitudes lie in [−1, +1], which compensates for differences in recording-device gain. Finally each cycle is zero-padded or truncated to exactly 80,000 samples (5.0 s).

D. Log-Mel Feature Extraction

A 128-band Mel spectrogram is computed with an FFT size of 2048 and a hop length of 512, converted to decibels and min-max normalised to [0, 1]. A 5 s cycle gives 157 time frames, so Model 1 receives a (128, 157, 1) tensor. For the EfficientNetB0 backbone the tensor is rescaled to [0, 255], copied to three channels and resized to 224 × 224. Model 2 takes a 128 × 128 log-Mel input. Table I lists the main parameters.

TABLE I KEY PIPELINE CONFIGURATION PARAMETERS

Parameter	Value	Reason
Sample rate	16,000 Hz	Respiratory sound energy is mostly below 2 kHz
Clip length	5.0 s (80,000 samples)	Fixed network input size
Mel bands	128	Frequency resolution of the spectrogram
N_FFT / hop length	2048 / 512 samples	Time-frequency resolution trade-off
Time steps	157 frames	(80,000 // 512) + 1
Model 1 split	70 / 15 / 15 percent	By patient, to prevent leakage
Binary labels	COPD = 1, NON-COPD = 0	Binary target encoding
Decision threshold	0.5	Sigmoid output boundary

E. Model 1: EfficientNetB0-BiLSTM-Attention

Model 1 is implemented in Keras in five stages. (1) A frozen ImageNet-pretrained EfficientNetB0 extracts a (7, 7, 1280) feature map from the 224 × 224 × 3 input. (2) The map is permuted and reshaped into a sequence of 7 time steps of 8960 features, then projected by Dense(512) with batch normalisation and dropout. (3) A bidirectional LSTM with 128 units models forward and reverse breathing dynamics and outputs a (7, 256) sequence. (4) A four-head self-attention block with residual connections and layer normalisation emphasises discriminative time steps. (5) Concatenated global average and max pooling (512 dimensions) feeds Dense(128, ReLU), Dropout(0.4) and a Dense(1, Sigmoid) output. Fig. 1 shows the pipeline.

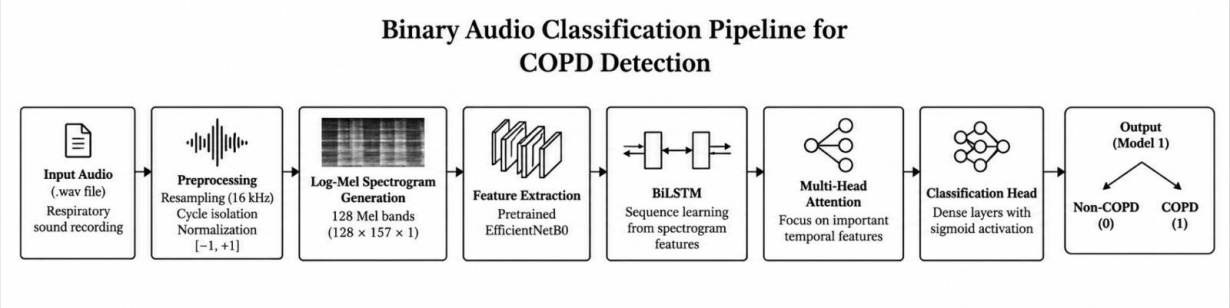


Fig. 1 Binary COPD screening pipeline of Model 1.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 11, November 2026

DOI: 10.17148/IJARCCE.2026.1511159

F. Model 2: Four-Stage Residual CNN

Model 2 is a PyTorch residual network, the RespiratoryCNN class, that follows the ResNet-18 layout. A stem of 7 × 7 convolution (64 channels, stride 2), batch normalisation, ReLU and 3 × 3 max pooling is followed by four stages of two residual blocks each, with 64, 128, 256 and 512 channels. The first block of stages 2 to 4 downsamples with stride 2, and every block has a shortcut connection and batch normalisation. Adaptive average pooling reduces the output to a 512-dimensional vector and a linear layer produces logits for copd0 to copd3.

G. Training Configuration

Model 1 uses two-stage transfer learning. In stage 1 (10 epochs) the backbone is frozen and the BiLSTM, attention and dense head are trained with Adam at a learning rate of 10^{-3}. In stage 2 (15 epochs) the top 30 percent of EfficientNetB0 is unfrozen and fine-tuned at 10^{-5} while batch-normalisation layers stay frozen. Validation AUC drives early stopping and checkpoint selection. Model 2 is trained with weighted cross-entropy and Adam at 10^{-4} with a batch size of 32, and the best checkpoint is saved. Table II summarises both configurations.

TABLE II TRAINING CONFIGURATION OF THE TWO MODELS

Item	Model 1 (COPD-EFF)	Model 2 (copd-mod2)
Framework	TensorFlow / Keras	PyTorch
Backbone	EfficientNetB0, BiLSTM, 4-head attention	4-stage ResNet-18 style CNN
Learning rate	10^{-3} (stage 1), 10^{-5} (stage 2)	10^{-4}
Dropout	0.4	0.3
Optimizer	Adam	Adam
Loss	Weighted binary cross-entropy	Weighted categorical cross-entropy
Class weights	Inverse class frequency	Inverse class frequency

H. Two-Stage Cascade and Web Application

At inference, Model 1 produces a sigmoid probability and the recording is labelled COPD when it reaches 0.5. A positive screen is passed to Model 2 for four-class labelling, whereas a negative screen marks Model 2 as bypassed, which saves computation and avoids producing an acoustic label after a negative screen. Fig. 2 shows the execution path and Fig. 3 the overall architecture.

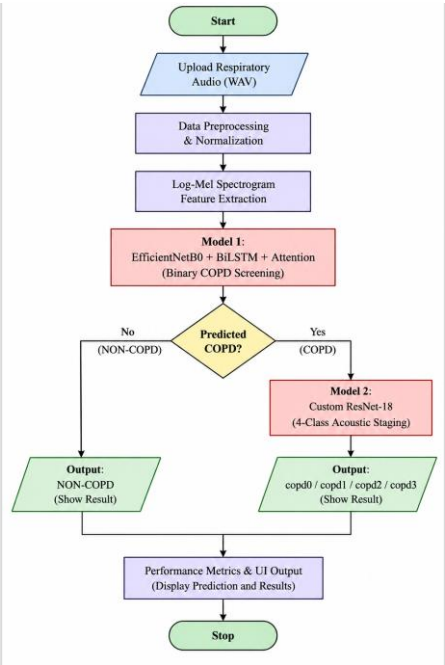


Fig. 2 Flowchart of the two-stage cascade.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 11, November 2026

DOI: 10.17148/IJARCCE.2026.1511159

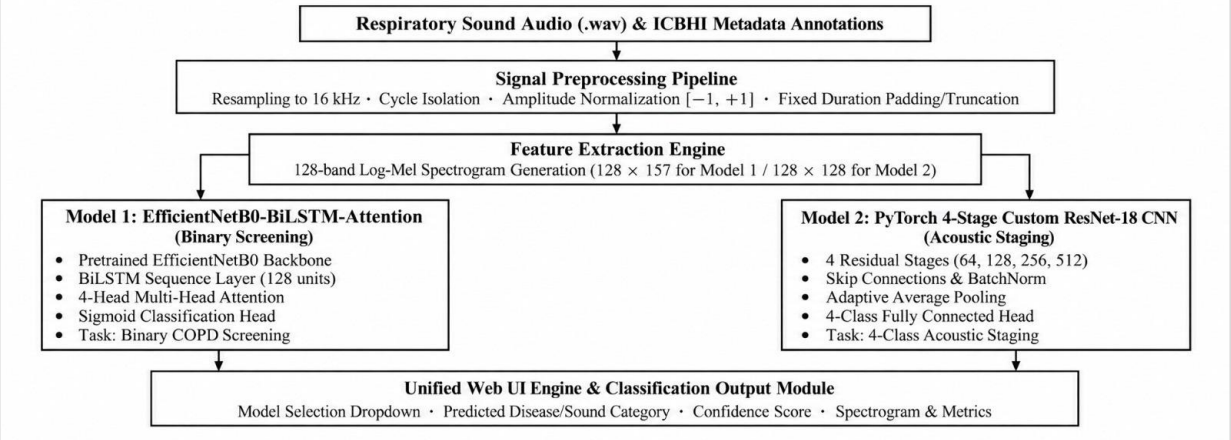


Fig. 3 System architecture of the dual-model platform.

The system is written in Python 3.9 with TensorFlow/Keras, PyTorch, Librosa, NumPy, SciPy and scikit-learn. The web application has a React and Vite front end, an Express API and a Python inference engine. A user uploads a .wav recording, previews the waveform and spectrogram, selects a model or the cascade, and views the predicted class, the confidence score and a research disclaimer. Inference needs no GPU.

IV. RESULTS AND DISCUSSION

A. Evaluation Protocol

Each model is evaluated on its own held-out test partition, which is not used for training, early stopping or checkpoint selection. Model 1 is assessed with accuracy, precision, sensitivity, specificity, F1-score, the confusion matrix and the area under the ROC curve (AUC). Model 2 is assessed with four-class accuracy, per-class recall and the ICBHI score reported by its evaluation script. Because the tasks and split protocols differ, the two sets of results are not a head-to-head ranking.

B. Model 1 on the Held-Out Cohort

Model 1 was tested on 1255 cycles: 179 NON-COPD and 1076 COPD. Fig. 4 shows the confusion matrix. It correctly identified 150 NON-COPD and 951 COPD cycles, with 29 false alarms and 125 missed COPD cycles. Table III lists the metrics.

TABLE III MODEL 1 PERFORMANCE ON THE HELD-OUT TEST COHORT

Metric	Value	Meaning
Accuracy	87.73%	(150 + 951) / 1255
Precision (COPD)	97.04%	Share of COPD predictions that are correct
Sensitivity (recall)	88.38%	COPD cycles correctly identified
Specificity	83.80%	NON-COPD cycles correctly identified
F1-score (COPD class)	92.51%	Harmonic mean of precision and recall
AUC	0.9313	Ranking quality across thresholds
Balanced accuracy	86.09%	Mean of sensitivity and specificity
Majority-class rate	85.74%	Accuracy of always predicting COPD

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 11, November 2026

DOI: 10.17148/IJARCCE.2026.1511159

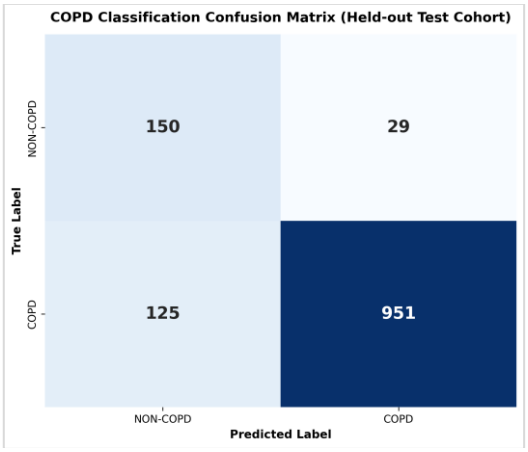


Fig. 4 Confusion matrix of Model 1 on the held-out cohort.

The test set is strongly imbalanced, with 85.74 percent COPD cycles, so a classifier that always predicted COPD would score 85.74 percent. The 87.73 percent accuracy is therefore only modestly above that baseline, and the balanced accuracy of 86.09 percent and the AUC of 0.9313 are better indicators of discrimination. The high COPD precision of 97.04 percent partly reflects the high prevalence of COPD cycles. For a screening aid the missed COPD cycles (125) outnumber the false alarms (29), which is the more serious error type.

C. Model 2 on Four-Class Respiratory Cycles

Model 2 was tested on 1033 held-out cycles. It classified 946 correctly, giving a four-class accuracy of 91.58 percent, and achieved an ICBHI score of 92.56 percent. Table IV gives per-class recall and Fig. 5 shows the confusion matrix. The large classes copd0 and copd1 reach about 94 percent recall, while the smaller classes copd2 and copd3 are harder at 81.20 and 84.21 percent. The macro-averaged recall is 88.31 percent. The errors arise mainly from low signal-to-noise ratio in ambulatory stethoscope recordings and from frequency overlap between subtle wheezes and background breath sounds.

TABLE IV MODEL 2 PER-CLASS RECALL ON THE HELD-OUT TEST CYCLES

Class	Test cycles	Correct	Recall
copd0	545	512	93.94%
copd1	279	262	93.91%
copd2	133	108	81.20%
copd3 (both)	76	64	84.21%
Overall / macro	1033	946	91.58% / 88.31%

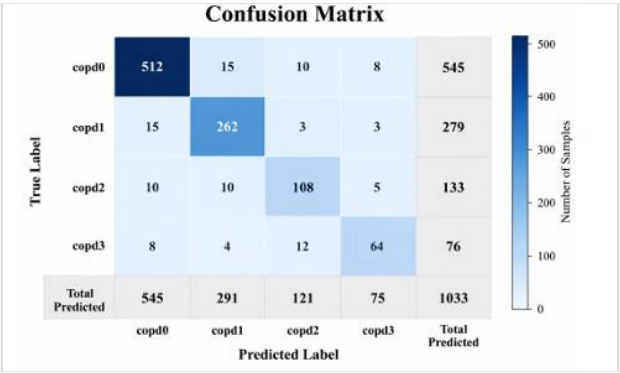


Fig. 5 Confusion matrix of Model 2 on the held-out four-class test cycles.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 11, November 2026

DOI: 10.17148/IJARCCE.2026.1511159

D. Prototype Interface

Fig. 6 shows the result view for a recording classified as non-COPD. It reports the predicted class, the confidence score and the Stage 1 probability bars, and marks Stage 2 as bypassed. The bypass message makes clear that the screening gate was negative and that the second model did not fail. The research disclaimer is displayed with every result.



Fig. 6 Cascade result view of the web interface for a non-COPD recording.

E. Discussion and Limitations

The results show that log-Mel spectrograms hold useful information for both binary screening and respiratory sound-event recognition. Several limits apply. First, the Model 1 test set is dominated by COPD cycles, so accuracy should be read with the balanced accuracy and AUC. Second, the four Model 2 classes come from ICBHI acoustic annotations and were not paired with spirometry or a clinical severity scale, so they must not be read as GOLD severity grades. Third, Model 2 was evaluated independently on annotated cycles, so the end-to-end cascade, in which Model 1 errors feed Model 2, has not been evaluated. Fourth, Model 2 used a split protocol different from the patient-level split of Model 1, and its figures may be optimistic if cycles of one patient appear in more than one partition. Fifth, all data come from one public database with a limited range of devices and patients, and there is no external or prospective validation. Finally, the confidence scores are not calibrated. Clinical utility and equivalence to spirometry remain unproven, and the system should be used only as a research aid.

V. CONCLUSION AND FUTURE SCOPE

This paper presented a two-model deep learning framework for respiratory sound analysis on the ICBHI 2017 database. An EfficientNetB0-BiLSTM-attention network screened for COPD with 87.73 percent accuracy, 88.38 percent sensitivity, 83.80 percent specificity and an AUC of 0.9313, and a four-stage residual CNN classified respiratory cycles into four acoustic categories with 91.58 percent accuracy. A web interface joins them in a cascade with an explicit research disclaimer.

Future work will collect patient-level severity labels paired with spirometry and train a clinically validated severity model, evaluate the full cascade end to end, and expand the data across institutions, regions and stethoscope types. Further directions include a prospective study against spirometry, explainability with attention visualisation and SHAP saliency maps, noise-robust preprocessing, calibration of the confidence scores, multi-modal fusion with health records, and quantised models for mobile or smart-stethoscope deployment.

ACKNOWLEDGMENT

The authors thank their guide, Prof. Pramod Kumar P M, Assistant Professor, Department of Computer Science and Engineering, VCET Puttur, for guidance throughout the project, and the department and management of the college for the facilities provided.

REFERENCES

[1]. A. Srivastava, S. Jain, R. Miranda, S. Patil, S. Pandya, and K. Kotecha, “Deep learning based respiratory sound analysis for detection of chronic obstructive pulmonary disease,” PeerJ Comput. Sci., vol. 7, p. e369, Feb. 2021.

[2]. N. S. Alghamdi, M. Zakariah, and H. Karamti, “A deep CNN-based acoustic model for the identification of lung diseases utilizing extracted MFCC features from respiratory sounds,” Multimedia Tools Appl., vol. 83, no. 35, pp. 82871–82903, Oct. 2024.

 This work is licensed under a Creative Commons Attribution 4.0 International License



International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 11, November 2026
DOI: 10.17148/IJARCCE.2026.1511159

[3]. S. Z. H. Naqvi and M. A. Choudhry, "An automated system for classification of chronic obstructive pulmonary disease and pneumonia patients using lung sound analysis," Sensors, vol. 20, no. 22, p. 6512, Nov. 2020.
[4]. C. Medina-Ramos, N. Sare-Vargas, W. Reategui-Romero, K. Paucar-Cuba, D. Carbonel-Olazabal, and J. Betetta-Gomez, "Utilizing deep learning algorithms for the prompt identification of chronic obstructive pulmonary disease," Eng. Technol. Appl. Sci. Res., vol. 15, no. 4, pp. 24940–24949, Aug. 2025.
[5]. V. Surekha, T. S. Kumar, and K. S. Rao, "COPD detection using respiratory sounds and transfer learning on CNN spectrograms," J. Biomed. Eng., vol. 40, no. 3, pp. 215–228, 2023.
[6]. M. Fraiwan, L. Fraiwan, M. Alkhodari, and O. Hassanin, "Recognition of pulmonary diseases from lung sounds using convolutional neural networks and long short-term memory," J. Ambient Intell. Humaniz. Comput., vol. 13, no. 10, pp. 4759–4771, 2022.
[7]. F. Demir, A. M. Ismael, and A. Sengur, "Classification of lung sounds with CNN model using parallel pooling structure," IEEE Access, vol. 8, pp. 105376–105383, 2020.
[8]. Y. Zhang, X. Liu, and J. Wang, "Pulmonary disease detection using LSTM neural networks on melspectrogram feature sequences," Expert Syst. Appl., vol. 213, p. 118943, 2023.
[9]. G. S. Marepalli, P. K. Kollu, and M. D. Inavolu, "Early detection of chronic obstructive pulmonary disease in respiratory audio signals using CNN and LSTM models," in Proc. 2024 IEEE Int. Conf. Contemp. Comput. Commun. (InC4), vol. 1, 2024, pp. 1–6.
[10]. R. Sreejith, R. K. Ramasamy, W.-N. Mohd-Isa, and J. Abdullah, "Enhanced lung disease classification using CALMNet: A hybrid CNN-LSTM-TimeDistributed model for respiratory sound analysis," IEEE Access, vol. 13, pp. 135053–135073, 2025.
[11]. C. C. Reksohartono, C. M. Lim, A. A. S. Gunawan, and J. J. Tedjasulaksana, "Deep learning for audio-based respiratory disease classification using clinical sound data," in Proc. 2025 Int. Conf. Smart Comput., IoT Mach. Learn. (SIML), 2025, pp. 1–6.
[12]. M. H. Islam, M. Rashed, N. K. Podder, M. I. Hossain, and A. K. Mondol, "Enhancing COPD detection with WaveNet: Assessing the efficacy of deep learning model," in Proc. 2025 IEEE Int. Conf. Quantum Photon., Artif. Intell., Netw. (QPAIN), 2025, pp. 1–6.
[13]. D. Koishiyeva, J. Kang, and A. Mukasheva, "Deep learning with multi-head attention for respiratory sound analysis," Comput. Methods Programs Biomed., vol. 260, p. 108543, 2025.
[14]. A. Roy and U. Satija, "A novel multi-head self-organized operational neural network architecture for chronic obstructive pulmonary disease detection using lung sounds," IEEE/ACM Trans. Audio, Speech, Lang. Process., vol. 32, pp. 2566–2575, 2024.
[15]. A. Roy, B. Gyanchandani, and S. K. Singh, "TriSpectraKAN: A novel approach for COPD detection via lung sound analysis," Sci. Rep., vol. 15, no. 1, p. 6296, Feb. 2025.
[16]. A. R. W. Sait, A. K. Dutta, and M. A. Shaikh, "Optimized Kolmogorov–Arnold networks-driven chronic obstructive pulmonary disease detection model," IEEE Access, vol. 13, pp. 162947–162960, 2025.
[17]. A. Roy and U. Satija, "A novel melspectrogram snippet representation learning framework for severity detection of chronic obstructive pulmonary diseases," IEEE Trans. Instrum. Meas., vol. 72, pp. 1–11, 2023.
[18]. P. Sahu, S. Kumar, and A. K. Behera, "SOUNDNet: Leveraging deep learning for the severity classification of chronic obstructive pulmonary disease based on lung sound analysis," in Proc. 2024 IEEE Int. Conf. Electron., Comput. Commun. Technol., 2024, pp. 1–6.
[19]. G. Altan, Y. Kutlu, and A. Gokcen, "Chronic obstructive pulmonary disease severity analysis using deep learning on multi-channel lung sounds," Turkish J. Electr. Eng. Comput. Sci., vol. 28, no. 5, pp. 2979–2996, 2020.
[20]. G. Altan, Y. Kutlu, and N. Allahverdi, "Deep learning on computerized analysis of chronic obstructive pulmonary disease," IEEE J. Biomed. Health Inform., vol. 24, no. 5, pp. 1344–1350, May 2020.
[21]. M. Archana, S. Dash, H. R. Sindhu Shree, and V. Shree Deeksha, "Predicting the severity of pulmonary disease from respiratory sounds using ML algorithms," in Proc. 2025 6th Int. Conf. Mobile Comput. Sustain. Informat. (ICMCSI), 2025, pp. 1750–1755.
[22]. B. Lakshmipriya, S. Jayalakshmy, S. A. Ajagbe, and P. I. Ogie, "COPD scope: Smart clinical pathways for respiratory illness diagnosis and management," Ann. Data Sci., 2025.
[23]. N. S. Haider, B. K. Singh, R. Periyasamy, and A. K. Behera, "Respiratory sound based classification of chronic obstructive pulmonary disease: A risk stratification approach in machine learning paradigm," J. Med. Syst., vol. 43, no. 8, p. 255, Aug. 2019.
[24]. A. I. Taloba and R. T. Matoog, "Detecting respiratory diseases using machine learning-based pattern recognition on spirometry data," Alexandria Eng. J., vol. 113, pp. 44–59, Feb. 2025.
[25]. G. Altan, Y. Kutlu, A. Ö. Pekmezci, and S. Nural, "Deep learning with 3D-second order difference plot on respiratory sounds," Biomed. Signal Process. Control, vol. 45, pp. 58–69, 2018.
[26]. O. Karaarslan, K. D. Belcastro, and O. Ergen, "Respiratory sound-base disease classification and characterization with deep/machine learning techniques," Biomed. Signal Process. Control, vol. 87, p. 105570, Jan. 2024.

 This work is licensed under a Creative Commons Attribution 4.0 International License



International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 11, November 2026
DOI: 10.17148/IJARCCE.2026.1511159

[27]. T. M. R. Yenumula, J. P. Marin, and A. Kumar, "Multi-class respiratory sound classification using machine learning," in Proc. 2025 IEEE Int. Conf. Omni-layer Intell. Syst. (COINS), 2025, pp. 1–7.
[28]. A. Bose, S. Chowdhury, and J. K. Chandra, "PPG-derived respiration signal predictor-guided simple rule-based detection of COPD," IEEE Trans. Instrum. Meas., vol. 74, pp. 1–9, 2025.
[29]. K. Le Trung, P. Nguyen Anh, and T. T. Han, "A novel method in COPD diagnosing using respiratory signal generation based on CycleGAN and machine learning," Comput. Methods Biomech. Biomed. Eng., pp. 1–16, Mar. 2024.
[30]. D. Tran-Anh, N. H. Vu, K. Nguyen-Trong, and C. Pham, "Multi-task learning neural networks for breath sound detection and classification in pervasive healthcare," Pervasive Mobile Comput., vol. 86, p. 101685, 2022.

 This work is licensed under a Creative Commons Attribution 4.0 International License