SSRG International Journal of Electronics and Communication Engineering
ISSN: 2348-8549/ https://doi.org/10.14445/23488549/IJECE-V12I4P120

Volume 12 Issue 4, 204-214, April 2025


*Review Article*

# Application of Deep Learning Algorithms in Lung Sound Classification: A Systematic Review Since 2015

Zakaria Neili[1], Kenneth Sundaraj[2]

[1]*Department of Electronics, University Badji Mokhtar, Annaba, Algeria.*
[2]*Fakulti Teknologi dan Kejuruteraan Elektronik dan Komputer, Universiti Teknikal Malaysia Melaka, Melaka, Malaysia.*

[2]*Corresponding Author : kenneth@utem.edu.my*

Received: 15 February 2025 Revised: 17 March 2025 Accepted: 18 April 2025 Published: 29 April 2025

***Abstract*** *- The article systematically explored the application of deep learning for lung sound classification in three popular scientific databases – PubMed, ScienceDirect and IEEE Xplore, for articles published between 2015 and 2024. Using specific keywords combined with deep learning terms, we identified 1428 articles. Based on their titles, abstracts and content, 33 articles were deemed relevant and selected for review. The article's thorough analysis revealed that deep learning algorithms have outperformed traditional machine learning techniques in lung sound classification.*

***Keywords*** *- Classification systems, Deep Learning, Lung sounds, Machine Learning, Systematic review.*

## 1. Introduction

Auscultation is a method used by physicians to evaluate pulmonary conditions by listening to patients' breathing sounds through a stethoscope. During this process, the physician listens to the patient's chest and diagnoses based on the patient's medical history. Pulmonary diseases can generally be categorized into two main types based on sound characteristics: normal and adventitious sounds. However, distinguishing these sounds often requires precision that depends on the physician's training. An automatic breathing sound classification system could, therefore, be a valuable alternative.

Traditional machine learning algorithms for detecting anomalies in respiratory signals involve two steps: feature extraction and pattern classification. These methods have been used to differentiate between continuous adventitious sounds like wheezes and discontinuous sounds like crackles. Digital stethoscopes are commonly used for lung sound acquisition and save audio data for further analysis. Wheezes and crackles are clinically significant symptoms of pulmonary diseases. Wheezes, high-pitched continuous sounds, occur in the 200 Hz to 2 kHz frequency range and last longer than 250 ms. Crackles, shorter discontinuous sounds, are detected during inspiratory and expiratory cycles and last less than 100 ms. Normal lung sounds range from 37.5 Hz to 1 kHz.

Pulmonary diseases are the third leading cause of death globally, with acute lower respiratory tract infection, asthma, lung cancer, tuberculosis, and Chronic Obstructive Pulmonary Disease (COPD) being the most prevalent, according to the World Health Organization (WHO). Recent research has focused on developing automated systems for classifying abnormal lung sounds using feature extraction techniques such as spectrograms, entropy, and Mel-Frequency Cepstral Coefficients (MFCC). Machine learning algorithms, including Hidden Markov Models (HMM), Gaussian Mixture Models (GMM), and logistic regression, have been employed for this purpose.

However, these methods face significant challenges. Most systems use binary classification to distinguish between crackles and wheezes, which is inadequate for multi-class classification. Additionally, the reliance on handcrafted features and the lack of freely available datasets hinder the evaluation and comparison of proposed algorithms. These limitations reduce the applicability of these methods in real-world scenarios.

Deep learning has emerged as a powerful solution, particularly when large datasets are available. Unlike traditional machine learning, deep learning does not require handcrafted feature extraction, as the model learns features directly from the data during training, as shown in Figure 1. This has made deep learning prominent in fields such as biomedical engineering.

Machine Learning (ML) and Deep Learning (DL) methods have been used for lung sound classification. However, to the best of our knowledge, few literature reviews have explored the application of machine learning algorithms in lung sound classification. Furthermore, the use of deep learning for lung sound classification has not been thoroughly studied, and a related research agenda has not been defined. Therefore, this systematic review aims to fill this gap by exploring the recent literature describing the use of DL to classify lung sounds. In this context, we identified several future research directions that could guide researchers when conducting future work. These include exploring the most promising research goals, investigating the effectiveness of transfer learning, and utilizing large datasets in research. Using both 1D raw audio data and 2D time-frequency images in lung sound classification using deep learning algorithms presents significant opportunities for improving the early detection and diagnosis of respiratory diseases. By identifying the most effective techniques and evaluating the results of various studies, this review could help streamline the field, enabling researchers to build upon previous work and develop more robust and effective models for lung sound classification. Ultimately, this could lead to improved patient outcomes and a reduced burden of respiratory diseases on global health.

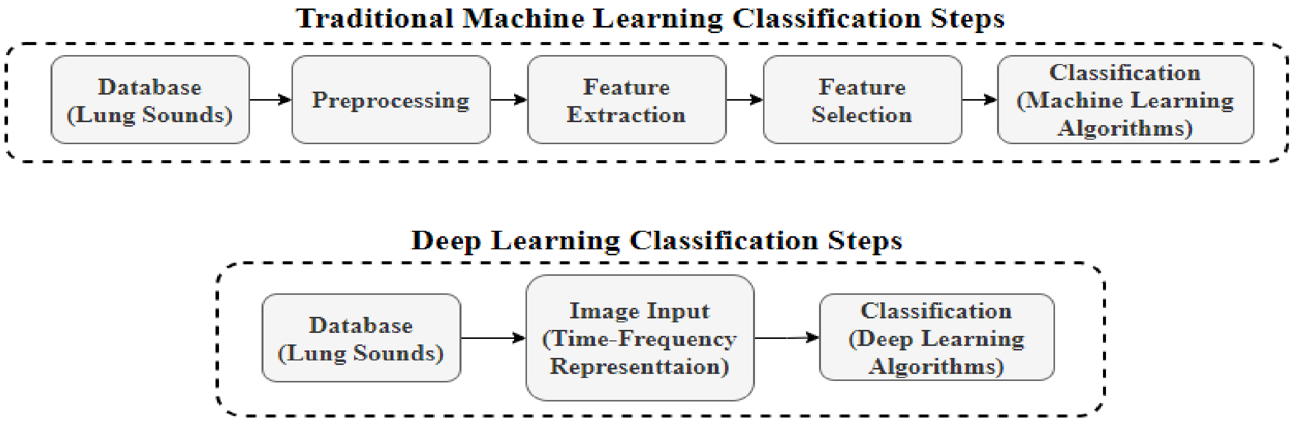


**Fig. 1 Differences among traditional classification methods based on machine learning algorithms and deep learning algorithms**

The main objectives of this review paper and its novelty which set it apart from other similar works in the literature are the following – 1) We evaluate various techniques for classifying lung sounds, not images of the lung (such as x-rays); 2) Lung sound classification is typically processed as either a one-dimensional audio signal or a two-dimensional image (time-frequency representation), 3) Our focus is specifically on lung sounds related to breathing, and we do not consider other sounds from the lungs, such as coughing, and 4) Our review specifically examines deep learning techniques for lung sound classification, rather than other machine learning approaches. The remainder of this manuscript is organized as follows – Section 2 provides a detailed description of the characteristics of lung sounds, the stethoscope used for recording them, and the available databases. Section 3 describes our search methodology, the process we used for selecting studies, and the data analysis presented in our review. Section 4 outlines the results of our study. We then discuss our findings and report on future work in Section 5. Finally, in Section 6, we present our conclusion.

## 2. Background on Lung Sounds

During the phases of respiration (inspiration and expiration), lung sounds are produced by airflow within the pulmonary system and can be heard over the chest wall. These sounds, characterized by their non-linear and non-stationary nature, pose challenges for clinicians in identifying abnormal and adventitious respiratory sounds, thereby complicating accurate diagnoses. Table 1 delineates the characteristics and types of lung sounds, correlating each type of respiratory pathology with specific breath sounds [39]. Lung sounds exhibit dominant frequencies ranging from 150 to 2000 Hz, while heart sounds predominantly fall below 150 Hz, enabling their differentiation. A variety of stethoscopes, including the Welch Allyn Meditron Master Elite Electronic Stethoscope, the 3M Littmann 3200 Electronic Stethoscope, the AKG C417L Microphone, and the 3M Littmann Classic II SE Stethoscope, are commercially available and employed for recording lung sounds. These devices are adept at capturing lung sound recordings and filtering out heart sounds, facilitating the processing and analysis of computerized lung signals. The development of the Computerized Respiratory Sound Analysis (CORSA) [37] standard has established guidelines for sensor placement, enhancing the quality of recorded patient sound data. The advent of electronic stethoscopes has significantly improved the collection of high-fidelity lung sound data from both healthy and diseased individuals, leading to the creation of extensive databases. These datasets are crucial for developing automated classification systems utilizing deep learning algorithms to identify pulmonary diseases.

Currently, two prominent databases are widely utilized in research, especially for deep learning applications – the International Conference on Biomedical Health Informatics (ICBHI) database [34], which is freely accessible, and the Respiration Acoustics Laboratory Environment (R.A.L.E.) repository [38], available commercially. Despite the availability of the R.A.L.E. database, recent studies predominantly employ the ICBHI dataset due to its public accessibility. The ICBHI database is notably unbalanced and contains a high noise level, presenting a challenging dataset for validating deep learning models under difficult conditions. In 2021, Fraiwan et al. [35] contributed lung sound data recorded from the chest wall to a publicly accessible database, further enriching resources available for respiratory sound analysis research. This comprehensive review underscores the importance of high-quality lung sound recordings and the role of electronic stethoscopes and standardized databases in advancing pulmonary diagnostics through deep learning algorithms.

**Table 1. Characteristics of lung sounds**

| Lung sound types | Sub-types of lung sound | Symptom | Dominant frequency range | Duration | Related Disease | Sensor location | Breath sound |
|---|---|---|---|---|---|---|---|
| Normal | Tracheal | / | / | / | / | / | High Pitch |
| | Vesicular | / | / | / | / | / | Low Pitch |
| | Bronchial | / | / | / | / | / | High Pitch |
| | Bronchovesicular | / | / | / | / | / | Low Pitch |
| Abnormal | Decreased normal lung sound | / | / | / | / | / | / |
| | Bronchial sounds | / | / | / | / | / | / |
| Adventitious | Continuous | Wheeze | 400 Hz or more | > 250 ms | Asthma | Over the lungs/trachea and most of the chest wall | High Pitch |
| | | Rhonchi | 200 Hz or less | > 250 ms | COPD | Over the lungs/trachea and most of the chest wall | Low Pitch |
| | Discontinuous | Crackles | 200 to 2000 Hz | Coarse Crackles < 100 ms<br>Fine Crackles < 100 ms | Pneumonia<br>Pulmonary fibrosis<br>CHF<br>IPF | Anterior and posterior chest wall for coarse crackles<br>Posterior lung base for fine crackles | Low Pitch for coarse crackles<br>High Pitch for fine crackles |

## 3. Materials and Methods

We systematically conducted a comprehensive literature search of the PubMed, ScienceDirect, and IEEE Xplore databases to identify journal articles on lung sound classification using deep learning published between 2015 and 2023. Conferences, books, letters, and clinical reports were excluded from this review. We screened the results based on predefined inclusion and exclusion criteria.

An initial selection process involved checking research titles to identify studies relevant to lung sound classification using deep learning. This step helped remove unreliable and duplicate articles. Research articles with unclear methodologies were also excluded. As depicted in Figure 2, our search strategy identified 1428 articles. After individual eligibility screening, 1335 articles were excluded because they were unrelated, duplicates, or lacked sufficient information.

Additionally, 66 articles were removed after reviewing their titles and abstracts, as they did not align with the scope of our review. These excluded articles either evaluated pulmonary pathologies without including lung sound classification, discussed respiratory sounds without performing lung sound analysis, did not evaluate pulmonary pathologies or lung sound classification, or were review articles. Further screening removed studies with unclear methodologies or those not focusing on the biomedical field and not using deep learning. The final stage identified 27 original research articles for this systematic review. An additional 6 articles, identified from other resources using similar eligibility criteria, provided supporting information on lung sound classification using deep learning.

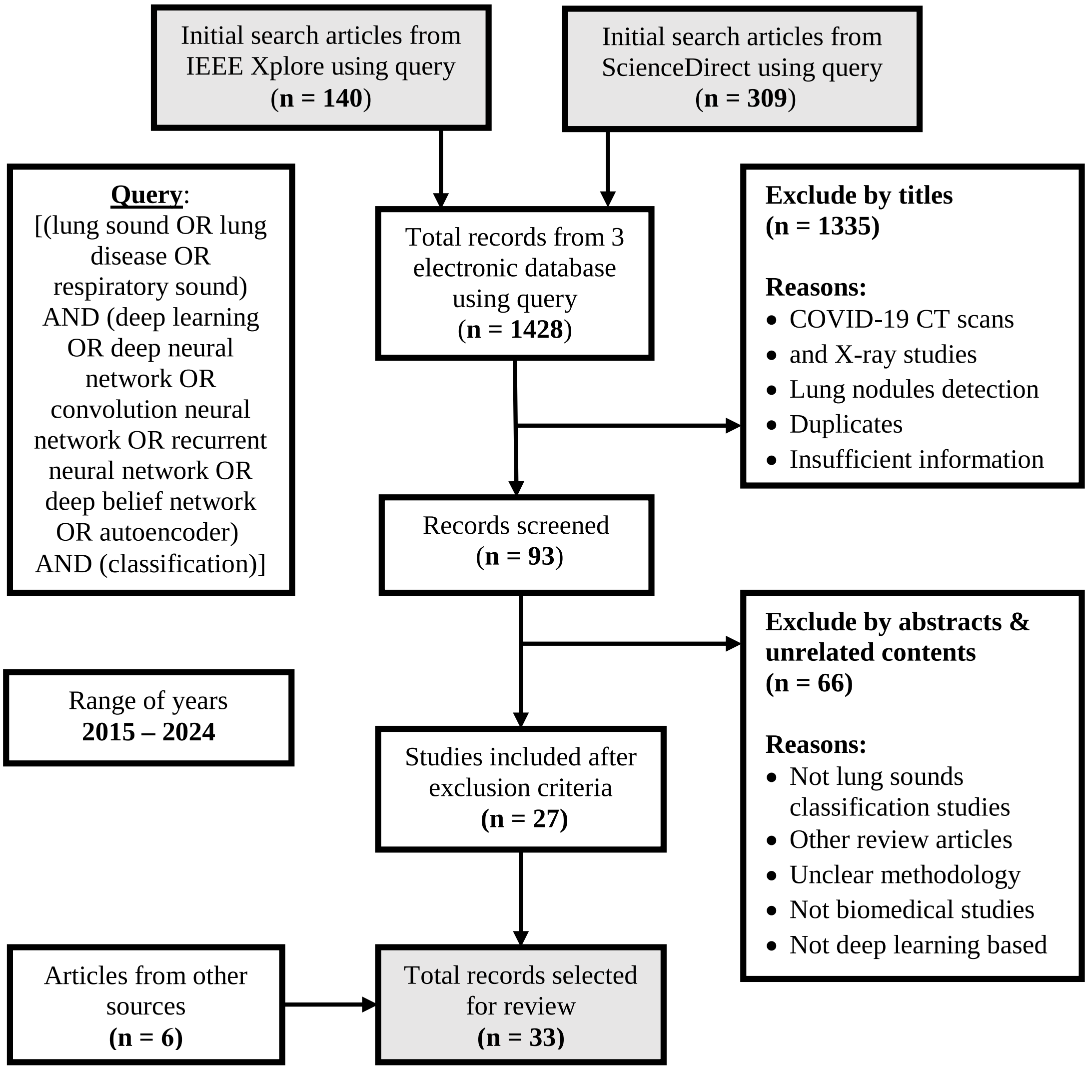


**Fig. 2 Summary of the search process on lung sound classification from the literature**

## 4. Results

The retrieved articles were sourced from three distinct databases, with the majority being published within the last three years. This trend highlights the increasing significance and interest in applying deep learning to lung sound classification. The reviewed papers, as summarized in Table 2, can be broadly categorized based on the type of input data used for deep learning networks: 1D input data (raw audio signals) and 2D input data (time-frequency images).

### *4.1. Raw Audio Signals (1D Input Data)*

Fraiwan et al. in [1] explored the use of deep learning architectures, specifically Convolutional Neural Networks (CNNs) and Bidirectional Long Short-Term Memory units (BDLSTM), to recognize pulmonary diseases from respiratory sounds. Their study utilized data from 213 patients, with 103 patients' data collected by the authors and 110 patients' data from the ICBHI database. To eliminate noise and artefacts, the dataset underwent preprocessing steps such as z-score normalization, displacement artefact removal, and wavelet smoothing. Evaluation metrics included accuracy, F1-score, Cohen's kappa, sensitivity, specificity, and precision, with the CNN + BDLSTM model achieving an accuracy of 99.62% and a precision of 98.85%. Despite these high-performance metrics, the study noted that preprocessing, although common in traditional machine learning, might not be suitable for clinical applications where real-time, raw data analysis is required, such as telemedicine during the COVID-19 pandemic.

Khodabakhshi et al. [2] 2017 introduced a dynamic modelling method for lung sound characterization using an Adaptive Recurrent Neural Network (ARNN) based on fuzzy functions, evaluated with Recurrent Quantification Analysis (RQA). Their study classified healthy, COPD, and asthma conditions, achieving a 91% accuracy using the fuzzy functions ARNN with RQA features. However, the study was limited by a small sample size (27 COPD patients, 31 asthma patients, and 25 healthy individuals), highlighting the need for data augmentation and further validation before clinical application.

Altan et al. in [3, 4] quantified lung sounds in 3D space using deep learning models with data from the

RespiratoryDatabase@TR, which included 12 channels of lung sounds. Their approach requires validation with diverse datasets to confirm its clinical applicability.

These studies collectively emphasize that while preprocessing aids in classification accuracy, it contradicts the objective of deep learning to learn from raw, unprocessed data, which is essential for real-world clinical applications. Jayalakshmy et al. [5] examined using RNN-based stacked BiLSTM and Gammatone Cepstral Coefficients (GTCC) for lung sound classification. They used Empirical Mode Decomposition (EMD) to process respiratory signals from the ICBHI and R.A.L.E. databases. Their method, Stacked BiLSTM-IMF 3 with GTCC, achieved a specificity of 0.88 and a sensitivity of 0.78. The study used a 70/30 train/test split but lacked a validation set, which is crucial for model training and evaluation in deep learning.

A related study [6] evaluated three different deep learning models: CNN, LSTM, and a hybrid CNN-LSTM, applied to raw lung auscultation sounds without preprocessing. This study used data augmentation techniques to enhance model accuracy and utilized a large dataset from two online sources, classifying 11 categories of lung sounds, including ten diseases and healthy sounds. The findings demonstrated the potential for real-time classification of pulmonary diseases using deep learning on unprocessed data, a significant contribution to the field. However, further research is necessary to validate these methods across different datasets and real-time applications, which could ultimately provide valuable tools for clinicians in diagnosing and treating pulmonary diseases. In summary, these studies highlight the potential of deep learning models for pulmonary disease classification from respiratory sounds. Despite the high accuracy achieved through preprocessing, the ultimate goal is to develop models capable of handling raw data for real-time applications, especially in telemedicine and other clinical settings. Further research and validation across diverse datasets are essential to realize the full potential of these deep learning approaches in practical, clinical scenarios.

### *4.2. Time-Frequency Images (2D Input Data)*

Aykanat et al. [7] designed an electronic stethoscope, introducing a Convolutional Neural Network (CNN) instead of traditional machine learning techniques such as Support Vector Machines (SVM). They validated their CNN model using data collected with their device. However, this validation is insufficient for such systems; different CNN deep learning-based models should be tested with various datasets to accurately assess their device's effectiveness. Additionally, there is no information on the hyperparameters tuning for CNN configurations or the optimization algorithm used. The article also lacks details on positive predictive value, negative predictive value, the Area Under the Receiver Operating Curve (AUROC), and the confusion matrix. However, the authors did not compare their results with existing works to highlight the system's advantages.

Bardou et al. [8] compared CNN with traditional methods like SVM, kNN, and GMM for lung sound classification using a 70/30 train/test data partition. However, they did not mention any data validation. The dataset should be split into three sets: training, validation, and testing, using the validation data to stop model training and testing data for evaluation. The study did not use other texture features like LPQ and BISIF or time-frequency representations. Jácome et al. [9] demonstrated Breathing Phase Detection (BPD) using CNN-based deep learning from respiratory sound recordings. Their approach should be validated with larger databases like the ICBHI challenge database. The method used in the detection process was not applied to other time-frequency representations to test its effectiveness. Chen et al. [10] introduced a modified time-frequency representation called Stockwell transform (S-transform) for lung sound classification using ResNets and Optimized S-Transform (OST). The accuracy reached 98.79%, with sensitivity and specificity at 96.27% and 100%, respectively. However, the wheeze + crackles class was not included, and the model's generalization during the learning process was not described.

Shi et al. [11] improved lung sound recognition using a VGGish-BiGRU system, discussing the effects of heart sounds, different time-frequency analyses, transfer learning, and retraining methods. However, the study did not report results from different epochs for assessing model learning. Demir et al. [12] used 6898 respiratory cycles from the ICBHI 2017 database for lung sound classification with a CNN and SVM. Their first method achieved 65.5% accuracy, while the second method achieved 63.09%. The use of transfer learning and SVM for deep learning is not novel, and the main contributions were not clarified. Acharya et al. [13] presented a hybrid CNN-RNN model for classifying breathing cycles using the ICBHI dataset. They did not report hyperparameters such as batch size, optimization algorithm, learning rate values, or the number of epochs. They plan to develop an embedded implementation for telemedicine, which might require a lightweight CNN model.

Rocha et al. [14] used the ICBHI dataset to introduce an automatic lung sound classification method. They employed Linear Discriminant Analysis (LDA), SVM, boosted trees, and CNNs for model learning. The best results were achieved with a dual input CNN, but further discussion was not provided. García-Ordás et al. [15] addressed data imbalance using variational autoencoders for data augmentation with the ICBHI dataset. The accuracy was not reported, and further evaluation without preprocessing techniques was suggested. Shuvo et al. [16] presented a lightweight CNN architecture for classifying respiratory diseases using a hybrid scalogram time-frequency representation. The study achieved high accuracy but noted that long computation times could be a bottleneck for real-time implementation. Demir et al. [17] introduced a parallel pooling CNN-based model for lung sound classification. They achieved 71.15% accuracy but should report the number of epochs and use different time-frequency representations for further validation. Messner et al. [18] used a convolutional RNN for multi-channel lung sound classification. Data augmentation was necessary due to a small dataset, and a large number of data samples were needed for training and testing. Jayalakshmy et al. [19] used a pretrained CNN-based model AlexNet with EMD and scalogram representations for lung sound classification. They achieved 83.78% accuracy but should validate their system against the ICBHI database for a more reliable comparison.

Pham et al. [20] used various time-frequency representations for lung sound classification with a deep learning C-DNN model. They introduced a mixture of expert techniques to enhance performance but did not compare it with other well-established representations. Sharan et al. [21] used different time-frequency representations for CNNs to classify lung sounds in children with pertussis. They achieved 90.48% accuracy but should use data collected according to the CORSA standard and report confusion matrices. Jung et al. [22] investigated three features for a DS-CNN model-based lung sound classification, reporting accuracy for STFT and MFCC. However, they did not provide information on hyperparameters. Jayalakshmy et al. [23] used deep learning for respiratory signal classification with the ICBHI dataset, Thinklabs Lung Sounds Library, and R.A.L.E. respiratory data. They achieved high accuracy but should report 95% confidence intervals for performance metrics.

Kim et al. [24] developed an automatic breath sound classification system using CNN-based deep learning. They achieved 86.5% accuracy and an AUC of 0.93 but should clarify their research aim and visualize learned features on the Mel-spectrogram. In [25], the authors introduced a preprocessing technique using Variational Mode Decomposition (VMD) for denoising respiratory signals and generating gammatonegram images for classification using deep convolution neural networks. They achieved high performance but should compare VMD with other techniques like Empirical Mode Decomposition (EMD) and test other Intrinsic Mode Functions (IMFs). In [26], the authors used cochleograms to enhance CNN models in respiratory adventitious sound classification. However, they did not compare cochleograms with other time-frequency representations to validate effectiveness. The authors in [27] made a comparison with several time-frequency representations as input data. They showed that different standard CNN architectures produce different experimental results for the various types of input data representation. The authors used the ICBHI database as the input data source in that work.

The study in [28] utilized CNN and SVM for multi-class respiratory sound classification. The authors should compare their approach with techniques like gammatonegram and scalogram for a more comprehensive evaluation. The authors in [29] introduced a hybrid CNN-RNN model for classifying lung sounds into normal, crackle, wheeze, and both categories. They should evaluate the CNN model's performance before combining it with RNN and conduct further experiments for validation. In [30], normal and abnormal respiratory sounds were collected by specialists and classified using a dense, lightweight CNN-bidirectional GRU model. Future work should explore eXplainable Artificial Intelligence (XAI) to understand correlations between diseases and symptoms. In [31], a hybrid neural model with focal loss function addressed data imbalance, using CNN and LSTM networks for lung sound classification. The authors should compare their standard time-frequency representation with others like scalograms and gammatonegram. The work in [32] classified abnormal respiratory sounds using CNN and artificial noise addition. The authors used five different networks with various settings and achieved a best performance of 100%. The research in [33] proposed a new preprocessing technique and CNN architecture for lung sound diagnosis. They should isolate and analyse the effects of each stage separately before combining them to establish a baseline performance.

**Table 2. Summary of findings on lung sound classification from selected literature**

| REF | YEAR | DL TYPE | DATABASE | NO. OF CLASSES | 1D / 2D DATA | RESULTS |
|---|---|---|---|---|---|---|
| [1] | 2021 | 1D CNN + BDLSTM | Self-collected dataset + ICBHI dataset | 6 | 1D raw data | Accuracy 99.62%<br>Precision 98.85% |
| [2] | 2017 | Attractor RNN (ARNN) | Self-collected dataset | 3 | 1D raw data | Accuracy 91% |
| [3] | 2018 | Deep Belief Networks (DBN) | Self-collected dataset | 2 | 1D raw data | Accuracy 93.67%<br>Sensitivity 91%<br>Specificity 96.33% |
| [4] | 2019 | Deep Belief Networks (DBN) | Self-collected dataset | 2 | 1D raw data | Accuracy 95.84%<br>Sensitivity 93.34%<br>Specificity 93.65% |
| [5] | 2021 | RNN-based stacked | ICBHI dataset | 4 | 1D raw data | Sensitivity 0.78 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | BiLSTM | | | | Specificity 0.88 |
| [6] | 2022 | CNN–LSTM | KAUH + ICBHI | 8 | 1D raw data | 99.6%, 99.8%, 82.4%, and 99.4% for datasets 1, 2, 3, and 4 (w/o data augmentation)<br>100%, 99.8%, 98.0%, and 99.5% for datasets 1, 2, 3, and 4 (with data augmentation) |
| [7] | 2017 | CNN | Self-collected dataset | 78 | Spectrogram (STFT) | healthy vs. pathological 86%<br>rale, rhonchus, and normal 76%<br>singular respiratory 80%<br>all audio types 62% |
| [8] | 2018 | CNN | RALE database | 7 | Spectrogram (STFT) | Accuracy 95.56% |
| [9] | 2019 | CNN | Self-collected dataset | 2 | Spectrogram (STFT) | Sensitivity of 97%<br>Specificity of 84% |
| [10] | 2019 | ResNet | ICBHI Database | 3 | OST-spectrogram | Accuracy 98.79%<br>Sensitivity 96.27%<br>Specificity 100% |
| [11] | 2019 | VGGish-BiGRU | Self-collected dataset | 3 | Mel-spectrogram | Asthma 83.33%<br>Pneumonia 86.75%<br>Normal 91.94% |
| [12] | 2020 | VGG16 + SVM | ICBHI Database | 4 | Spectrogram (STFT) | Accuracy of 65.5% for the first Method Accuracy of 63.09% for the second method |
| [13] | 2020 | CNN-RNN | ICBHI Database | 4 | Spectrogram (STFT) | 66.31% of the 4 classes on overall data<br>71.81% of the 4 classes on patient-specific data |
| [14] | 2021 | CNN | ICBHI Database | 3 | Spectrogram<br>Mel-spectrogram | 3 class fixed durations tasks, best classifier accuracy 96.9%<br>3 class variable durations tasks, best classifier accuracy 81.8% |
| [15] | 2020 | CNN & Variational Autoencoders | ICBHI Database | 3, 6 | Mel-spectrogram | 0.993 F-Score in 3 class<br>0.990 F-Score in 6 class |
| [16] | 2020 | CNN | ICBHI Database | 6 | Scalogram (STFT) | 99.20% for ternary chronic classification<br>99.05% for six-class pathological classification |
| [17] | 2020 | CNN | ICBHI Database | 4 | Spectrogram (STFT) | Accuracy 71.15.% |
| [18] | 2020 | GRNNs & CNN | Self-collected dataset | 2 | Spectrogram (STFT) | F1-score 92% |
| [19] | 2020 | AlexNet | RALE database | 4 | Scalogram (STFT) | Accuracy 83.% |
| [20] | 2021 | CNN-MoE | ICBHI Database | 8 | 4 Spectrogram types | Sensitivity 99% |
| [21] | 2021 | 2D CNN | Dataset constructed from YouTube | 5 | Mel-spectrogram<br>Wavelet scalogram<br>Cochleogram | Accuracy 90.48% |
| [22] | 2021 | Depthwise Separable CNN (DS-CNN) | Self-collected dataset | 4 | Spectrogram (STFT) | Accuracy 85.74% |
| [23] | 2021 | cGAN | ICBHI dataset | 4 | Scalogram<br>Spectrogram | Accuracy 92.68% |
| [24] | 2021 | VGG16 & CNN | Self-collected dataset | 4 | Mel-spectrogram | Accuracy 86.50% |
| [25] | 2021 | CNN | Self-collected dataset | 3 | Gammatonegram | Accuracy 98.8%<br>Precision 97.7%<br>Sensitivity 100%<br>Specificity 97.6% |
| [26] | 2023 | CNN | ICBHI Database | 4 | Cochleogram | 85.1% in wheezes |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | | | | 73.8% of crackles |
| [27] | 2022 | CNN | ICBHI Database | 4 | Scalogram<br>Spectrogram<br>Mel-spectrogram<br>Gammatonegram | Gammatonegram and scalogram produced the best classification results |
| [28] | 2023 | CNN + SVM | Self-collected dataset | 3 | Spectrogram (STFT) | Accuracy 83% |
| [29] | 2021 | CNN + RNN | ICBHI Database | 4 | Spectrogram (STFT) | Sensitivity 0.63<br>Specificity 0.83<br>Average score 0.73<br>Harmonic score 0.72 |
| [30] | 2022 | CNN and BiGRU | Self-collected dataset | 6 | Mel-spectrogram | Accuracy of 92.3%<br>Sensitivity of 92.1%<br>Specificity of 98.5%<br>F1-score of 91.9% |
| [31] | 2022 | CNN-LSTM | ICBHI Database | 4 | Spectrogram (STFT) | 60/40 split - Sensitivity 47.37%, Specificity 82.46%, F1-Score 64.92%, Accuracy 73.69%<br>10 Fold CV - Sensitivity 52.78%, Specificity 84.26%, F1-Score 68.52%, Accuracy 76.39%<br>LOO CV - Sensitivity 60.29%, Accuracy 74.57% |
| [32] | 2021 | ANA + CNN | Datasets from multiple internet sources | 7 | Spectrum analysis | VGG-B1 = 0.95%, VGG-B3 = 0.95%, VGG Drop = 0.95%, VGG-V1 = 0.84%, VGG-V2 = 0.84%, AlexNet = 100%, InceptionNet = 0.95%, ResNet = 0.95%, LeNet5 = 0.89% |
| [33] | 2022 | CNN | ICBHI Database | 8 | Mel-spectrogram | Recall of 0.991, F1 score of 0.993 and precision of 0.994 |

## 5. Discussion

We systematically analysed 33 articles that explore the application of deep learning to lung sound classification. The studies reviewed not only contribute to the field of lung sound analysis but are also significant across various health domains. Deep learning, a prominent technique in artificial intelligence, is particularly prevalent in the biomedical field. The majority of studies included in this work were conducted within the last two years, indicating that technology-based deep learning is a recent research hotspot. Current lung sound applications, including eHealth embedded system diagnosis, require improvements, especially for diagnosing lung conditions. Digital lung sound recordings, widely collected by electronic stethoscope systems and uploaded to cloud repositories, are expected to see increased efficiency in healthcare applications due to the rapid development of the Internet and 5G networks, which will enhance data variability and volume, particularly in pulmonary care.

The increasing availability of data is well-suited for deep learning algorithms. Currently, we find two publicly available databases [34, 35]. These data, coupled with tools like TensorFlowLite, can be used to deploy deep learning applications on inference systems such as smartphones for real-time diagnosis and classification of pulmonary pathologies. To integrate deep learning technologies into healthcare, several initiatives in artificial intelligence aim to enhance model functionality and transparency. For instance, the Shapley Additive Explanations (SHAP) framework demonstrates how input data features contribute to the final output, which has been validated by many healthcare applications.

Spirometry, a fundamental measure of pulmonary function, plays a crucial role in diagnosing conditions like asthma and COPD early. These conditions have similar symptoms, making differential diagnosis challenging. Traditionally, a spirometer, which uses a threshold for the forced air applied by the patient, helps determine the condition. Developing an automatic pulmonary disease classification system using deep learning on lung sound signals could offer an alternative strategy for earlier diagnosis and serve as a validated tool alongside spirometry. Current studies focus on diagnosing and classifying various pulmonary diseases, but no previous study has developed an automatic system to distinguish between asthma and COPD. For example, Trivedy et al. [36] developed an automatic disease classification system using a CNN for a smartphone-enabled spirometer, achieving an accuracy of 98.98%. Future research should focus on classifying lung sounds from asthma and COPD to aid doctors in early diagnosis, even those not well-trained.

Deep learning technologies have shown advantages over traditional machine learning algorithms. Designing a real-time lung sound classification system must consider real-world hospital settings, where background noise, patient movement, and conversation are common. The literature review reveals that researchers typically perform preprocessing steps, facilitating deep learning-based classification and improving accuracy.

However, deploying these algorithms on Single-Board Computers (SBCs) for real-time use may reduce accuracy in actual diagnoses. To address this, exploring the direct use of raw audio recordings with deep learning algorithms without preprocessing could reveal the true power of deep learning and aid in developing SBC-based hardware systems for hospital use.Nearly all reviewed studies (99%) utilized the ICBHI challenge database, employing preprocessing steps like slicing, resampling, and feature extraction. No study has examined classifying ICBHI audio data without these steps, indicating a need for further evaluation to ensure the robustness and effectiveness of deep learning methods. Researchers should address challenges such as noise and artefact inclusion in input data and feature selection difficulties.

The review also highlighted that transfer learning often used models pretrained on ImageNet, which may not be suitable for time-frequency representations of lung sounds. Future research should explore transfer learning using lung sound spectrogram data to improve accuracy and reliability, enhancing deep learning performance in clinical diagnosis and treatment. There is a critical need to collect and share diverse lung sound datasets, such as those for asthma, COPD, and lung cancer, to support related research. This would facilitate the development of accurate diagnostic tools such as and improve the reproducibility and comparability of research findings. Despite challenges like background noise and patient movement in real-world settings, deep learning models trained on large, diverse datasets can identify complex patterns in lung sounds. Future research should focus on developing robust deep learning models, possibly using generative models for data augmentation and techniques to handle missing or imbalanced data.

## 6. Conclusion

This manuscript systematically summarizes and reviews various deep-learning studies on lung sound classification. The relevant research has been evaluated based on the data, algorithms, and models utilized. The studies are assessed for using deep learning algorithms in lung sound classification, focusing on their capabilities to distinguish various time-frequency inputs. While applying deep learning in lung sound classification has shown promising performance, certain challenges and limitations remain unresolved.

The review's contributions are twofold – (1) it highlights the limitations of current studies and identifies future research opportunities, and (2) it provides a systematic and comprehensive review of deep learning studies on lung sound classification. Further research is needed to explore the potential of deep learning in enhancing the diagnosis and management of respiratory diseases in clinical settings. This review can serve as a valuable reference for future research in this field.

## References

[1] M. Fraiwan et al., "Recognition of Pulmonary Diseases from Lung Sounds Using Convolutional Neural Networks and Long Short-Term Memory," *Journal of Ambient Intelligence and Humanized Computing*, vol. 13, pp. 4759-4771, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[2] Mohammad Bagher Khodabakhshi, and Mohammad Hassan Moradi, "The Attractor Recurrent Neural Network Based on Fuzzy Functions: An Effective Model for the Classification of Lung Abnormalities," *Computers in Biology and Medicine*, vol. 84, pp. 124-136, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[3] Gokhan Altan et al., "Deep Learning with 3D-Second Order Difference Plot on Respiratory Sounds," *Biomedical Signal Processing and Control*, vol. 45, pp. 58-69, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[4] Gokhan Altan, Yakup Kutlu, and Novruz Allahverdi, "Deep Learning on Computerized Analysis of Chronic Obstructive Pulmonary Disease," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 5, pp. 1344-1350, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[5] S. Jayalakshmy, and Gnanou Florence Sudha, "GTCC-Based BiLSTM Deep-Learning Framework for Respiratory Sound Classification Using Empirical Mode Decomposition," *Neural Computing and Applications*, vol. 33, pp. 17029-17040, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[6] Ali Mohammad Alqudah, Shoroq Qazan, and Yusra M. Obeidat, "Deep Learning Models for Detecting Respiratory Pathologies from Raw Lung Auscultation Sounds," *Soft Computing*, vol. 26, pp. 13405-13429, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[7] Murat Aykanat et al., "Classification of Lung Sounds Using Convolutional Neural Networks," *EURASIP Journal on Image and Video Processing*, vol. 2017, pp. 1-9, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[8] Dalal Bardou, Kun Zhang, and Sayed Mohammad Ahmad, "Lung Sounds Classification Using Convolutional Neural Networks," *Artificial Intelligence in Medicine*, vol. 88, pp. 58-69, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[9] Cristina Jácome et al., "Convolutional Neural Network for Breathing Phase Detection in Lung Sounds," *Sensors*, vol. 19, no. 8, pp. 1-10, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[10] Hai Chen et al., "Triple-Classification of Respiratory Sounds Using Optimized S-Transform and Deep Residual Networks," *IEEE Access*, vol. 7, pp. 32845-32852, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[11] Lukui Shi et al., "Lung Sound Recognition Algorithm Based on VGGish-BiGRU," *IEEE Access*, vol. 7, pp. 139438-139449, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[12] Fatih Demir, Abdulkadir Sengur, and Varun Bajaj, "Convolutional Neural Networks Based Efficient Approach for Classification of Lung Diseases," *Health Information Science and Systems*, vol. 8, pp. 1-8, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[13] Jyotibdha Acharya, and Arindam Basu, "Deep Neural Network for Respiratory Sound Classification in Wearable Devices Enabled by Patient Specific Model Tuning," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 14, no. 3, pp. 535-544, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[14] Bruno Machado Rocha et al., "Automatic Classification of Adventitious Respiratory Sounds: A (Un)Solved Problem?," *Sensors*, vol. 21, no. 1, pp. 1-19, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[15] María Teresa García-Ordás et al., "Detecting Respiratory Pathologies Using Convolutional Neural Networks and Variational Autoencoders for Unbalancing Data," *Sensors*, vol. 20, no. 4, pp. 1-16, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[16] Samiul Based Shuvo et al., "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, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[17] Fatih Demir, Aras Masood Ismael, and Abdulkadir Sengur, "Classification of Lung Sounds with CNN Model Using Parallel Pooling Structure," *IEEE Access*, vol. 8, pp. 105376-105383, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[18] Elmar Messner et al., "Multi-Channel Lung Sound Classification with Convolutional Recurrent Neural Networks," *Computers in Biology and Medicine*, vol. 122, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[19] S. Jayalakshmy, and Gnanou Florence Sudha, "Scalogram Based Prediction Model for Respiratory Disorders Using Optimized Convolutional Neural Networks," *Artificial Intelligence in Medicine*, vol. 103, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[20] Lam Pham et al., "CNN-MoE Based Framework for Classification of Respiratory Anomalies and Lung Disease Detection," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 8, pp. 2938-2947, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[21] Roneel V. Sharan et al., "Detecting Pertussis in the Pediatric Population Using Respiratory Sound Events and CNN," *Biomedical Signal Processing and Control*, vol. 68, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[22] Shing-Yun Jung et al., "Efficiently Classifying Lung Sounds through Depthwise Separable CNN Models with Fused STFT and MFCC Features," *Diagnostics*, vol. 11, no. 4, pp. 1-12, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[23] S. Jayalakshmy, and Gnanou Florence Sudha, "Conditional GAN Based Augmentation for Predictive Modeling of Respiratory Signals," *Computers in Biology and Medicine*, vol. 138, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[24] Yoonjoo Kim et al., "Respiratory Sound Classification for Crackles, Wheezes, and Rhonchi in the Clinical Field Using Deep Learning," *Scientific Reports*, vol. 11, no. 1, pp. 1-11, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[25] Sonia Gupta, Monika Agrawal, and Desh Deepak, "Gammatonegram Based Triple Classification of Lung Sounds Using Deep Convolutional Neural Network with Transfer Learning," *Biomedical Signal Processing and Control*, vol. 70, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[26] L.D. Mang et al., "Cochleogram-Based Adventitious Sounds Classification Using Convolutional Neural Networks," *Biomedical Signal Processing and Control*, vol. 82, pp. 1-13, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[27] Zakaria Neili, and Kenneth Sundaraj, "A comparative Study of the Spectrogram, Scalogram, Melspectrogram and Gammatonegram Time-Frequency Representations for the Classification of Lung Sounds using the ICBHI Database based on CNNs," *Biomedical Engineering*, vol. 67, no. 5, pp. 367-390, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[28] Funda Cinyol et al., "Incorporating Support vector Machine to the Classification of Respiratory Sounds by Convolutional Neural Network," *Biomedical Signal Processing and Control*, vol. 79, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[29] Naoki Asatani et al., "Classification of Respiratory Sounds Using Improved Convolutional Recurrent Neural Network," *Computers & Electrical Engineering*, vol. 94, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[30] Youngjin Choi et al., "Lightweight Skip Connections With Efficient Feature Stacking for Respiratory Sound Classification," *IEEE Access*, vol. 10, pp. 53027-53042, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[31] Georgios Petmezas et al., "Automated Lung Sound Classification Using a Hybrid CNN-LSTM Network and Focal Loss Function," *Sensors*, vol. 22, no. 3, pp. 1-13, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[32] Rizwana Zulfiqar et al., "Abnormal Respiratory Sounds Classification Using Deep CNN Through Artificial Noise Addition," *Frontiers in Medicine*, vol. 8, pp. 1-16, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[33] Saumya Borwankar et al., "Improvise Approach for Respiratory Pathologies Classification with Multilayer Convolutional Neural Networks," *Multimedia Tools and Applications*, vol. 81, pp. 39185-39205, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[34] B.M. Rocha et al., "A Respiratory Sound Database for the Development of Automated Classification," *Precision Medicine Powered by pHealth and Connected Health*, Thessaloniki, Greece, pp. 33-37, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[35] Mohammad Fraiwan et al., "A Dataset of Lung Sounds Recorded from the Chest Wall Using an Electronic Stethoscope," *Data in Brief*, vol. 35, pp. 1-6, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[36] Sudipto Trivedy et al., "Design and Development of Smartphone-Enabled Spirometer with a Disease Classification System Using Convolutional Neural Network," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 9, pp. 7125-7135, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[37] A.R.A. Sovijärvi, J. Vanderschoot, and J.E. Earis, "Standardization of Computerized Respiratory Sound Analysis," *European Respiratory Review*, vol. 10, no. 77, pp. 586-590, 2000. [Google Scholar]

[38] H. Pasterkamp, "RALE: A Computer-Assisted Instructional Package," *Respir Care*, vol. 35, 1990. [Google Scholar]

[39] Hans Pasterkamp, Steve S. Kraman, and George R. Wodicka, "Respiratory Sounds Advances Beyond the Stethoscope," *American Journal of Respiratory and Critical Care Medicine*, vol. 156, no. 3, pp. 974-987, 1997. [CrossRef] [Google Scholar] [Publisher Link]