Tuijin Jishu/Journal of Propulsion Technology
ISSN: 1001-4055
Vol. 45 No. 3 (2024)

Enhancing Lung Sound Classification: A Review of Deep Learning Models with Transform and Spectral Features

[1]Ohmshankar S , [2]G.Sudhagar, [3]M.Hema
[1]*Electronics and communication Engineering Bharath Institute of Higher Education and Research* Tambaram, Chennai, India.
[2]*Information Technology Easwari Engineering College* Chennai

***Abstract*—**Accurate identification of lung sounds recorded by electronic stethoscopes is critical to the early diagnosis of respiratory disorders. In the last thirty years, ML(machine learning) methods have played a significant part in enhancing the accuracy of expert evaluations. This review article offers a thorough analysis of the developments in lung sound clas- sification, emphasizing a novel method that makes use of an Enhanced Long Short-Term Memory (ELSTM) model. Using a Savitzky-Golay filter to preprocess lung sound audio signals in order to eliminate noise is the first step. The preprocessed signals are then used to extract important characteristics, such as Stockwell Transform, spectral-based features, and Short-Time Fourier Transform (STFT). The multi-layered design of the ELSTM model makes use of these characteristics to enhance expressiveness, representation learning, and sequence learning. To improve classification accuracy, a hybrid loss function that combines Categorical Cross-Entropy (CCE) and Focal Loss (FL) is utilized to effectively forecast the error between actual and projected values. This work presents a comprehensive assessment of the literature on lung sound categorization, emphasizing the several approaches and machine learning models that have been investigated. It examines these strategies' advantages and disadvantages critically while highlighting the ELSTM model's contributions to solving contemporary problems. This study intends to lead future research and applications in the field of respiratory sound analysis by integrating prior research and offering a novel model. Ultimately, this will help in the timely identification and treatment of respiratory disorders.

***Index Terms*—**Lung sound, classification, ELSTM, Savitzky- Golay filter, and STFT.

I. Introduction

More than 500 million people worldwide suffer from respiratory disorders, making it one of the most common medical illnesses worldwide [1]. The majority of patients have difficulty identifying or comprehending the symptoms of chronic illnesses, which delays diagnosis. The most common technique used by doctors to assess patients' lung sounds in order to diagnose diseases is manual auscultation [2]. An experienced medical professional is required to auscultate the patient's lung sounds since poor equipment calibration and noisy surroundings, such as heartbeat and coughing noises, make disease diagnosis difficult [3]. There are three primary forms of lung sounds: wheeze, crackles, and rhonchi. When there's an obstruction in the airway, the patient experiences wheezing, which is a persistent, high-pitched, unnatural sound that comes from their lungs when they have diseases including viral pneumonia, COVID and pulmonary fibrosis [4], [5]. Dur- ing inhalation or exhalation, crackles are sudden, intermittent sounds associated with conditions such as chronic obstructive pulmonary disease (COPD) and asthma[4], [5]. Continuous, coarse, low-pitched noises resembling rattling or snoring are caused by secretions in the airways [5]. By employing these sound patterns during the auscultation procedure, medical pro- fessionals acquire their knowledge regarding the pertinent lung ailment. The healthcare industry has clearly benefited from AI in recent years, especially when it comes to the detection of diseases like cancer, respiratory issues, and neurological abnormalities.Researchers have shown considerable interest in utilizing deep learning (DL) for classifying lung sounds [6]. The study of disease-specific features is made easier by feature extraction using deep learning, a data driven technique that extracts unique features straight from unprocessed image or data [7]. Convolutional Neural Networks (CNNs) are ex- tremely ideal for tasks like recognition of object, segmentation of image, and classification because they can retrieve set of features from image data and learn to identify patterns. As a result, Convolutional Neural Networks (CNN) have been utilized to categorize spectrograms produced by the sounds of the lung [8]. In-depth assessments of current approaches and AI models for classifying lung sounds are provided in this work, along with a critical evaluation of their advantages and disadvantages.

The primary contributions of the paper are outlined below.

- This paper provides a comprehensive review of existing methodologies and machine learning models for lung sound classification. By critically analyzing the strengths and limitations of these approaches, it highlights the advancements and unique contributions of the ELSTM model. The synthesis of past research, combined with the introduction of the novel ELSTM model, aims to guide future studies and applications, ultimately contributing to the timely diagnosis and treatment of respiratory conditions.

- The research integrates advanced preprocessing tech- niques and feature extraction methods, such as the Savitzky-Golay filter, Short-Time Fourier Transform (STFT), and Stockwell Transform, along with spectral-based features, to enhance the quality of input data for the ELSTM model. Additionally, the employment of a hybrid loss function combining Categorical Cross- Entropy (CCE) and Focal Loss (FL) further refines the prediction accuracy, effectively reducing errors between actual and predicted values.

- The discussion includes potential future enhancements to further improve lung sound classification techniques. The synthesis of past research, combined with the introduction of the novel ELSTM model, aims to guide future studies and applications, ultimately contributing to the timely diagnosis and treatment of respiratory conditions.

The remaining sections of this paper are organized as follows: Section 2 presents basic information on lung sound classifi- cation and reviews state-of-the-art deep learning techniques in this area. The suggested model is explained in depth in Section 3, and its performance is assessed and analyzed in Section 4. Lastly, a summary of the study's conclusions may be found in Section 5.

II. Related Works

Through the use of a stethoscope to listen for ambient sounds, Rocha et al. [9] created categorization models for the diagnosis of chest diseases. Initially, they assembled 920 samples of lung sounds for various categories (e.g., COPD, Healthy, etc.) into a database. The challenge's second job was to identify the characteristics of the sounds and categorize them based on whether they were crackles, wheezy noises, or both. Finally, they used variables including MFCC,energy, spectral features, entropy, and coefficients of wavelet to con- duct possibility studies on machine learning(ML) techniques like SVM(support vector machine) and ANN(artificial neural networks).Using a lung data set, Dalal et al. [10] investi- gated several machine learning techniques for classifying lung sounds.Using the ICBHI challenge data set, Fateh et al. [11] suggested a CNN model that had already been trained to extract deep features. There are four categories in the data set: wheezes, normal, wheezes plus crackles, and crackles. Initially, they employed a visual method using spectrogram images produced by feature extraction from lung sounds. Next, they employed the Random Subspace Ensembles (RSE) approach to use the depth features as the input to the LDA(Linear Discriminant Analysis) classifier. Consequently, their model outperformed the current methods in terms of clas- sification accuracy, increasing it by 5Based on the ICBHI lung sound data collection, Demir et al. [12] suggested two CNN-based methods for classifying lung disorders. There are 6898 recordings across 4 classes in the data collection. First, they used the STFT(Short Time Fourier transform) method to turn the lung sounds into spectrogram images. A method for multi channel lung sound categorization using temporal, spectral and spatial data was published by Elmar et al. [13]. In order to categorize the sounds of lung gathered from 17 channel finding devices, they suggested utilizing spectrogram characteristics to train a CRNN(convolutional recurrent neural network). An F1-score of 92%was achieved by their CRNN model in the binary classification. Acharya et al. [14] used patient-specific data to retrain a deep RNN (recurrent neural network) method, which allowed them to get a score of 71.81% on four-class categorization. A deep residual network (ResNet) was trained by Chen et al. [15] to perform triple categorization of respiratory sounds with up to 100%, 96.27%, and 98.79% accuracy, sensitivity, and specificity, respectively. The suggested model reportedly fared better than CNN.

To train a lightweight neural network model, Shuvo et al. [16] utilized empirical mode decomposition (EMD) and continuous wavelet transform (CWT). Their lightweight CNN outperformed several larger networks and other modern lightweight models, achieving accuracy scores of 99.92% for three class chronic categorization and 99.70% for six class pathological categorization. Cinyo et al. [17] proposed a CNN architecture integrated with SVM(support vector machine) or softmax, showing that the VGG16-CNN-SVM model had the highest classification accuracy at 83%. Saraiva et al.

[18] recommended a CNN for quadruple classification and achieved a 74.3% accuracy. Tariq et al. [19] developed a feature-related fusion network using MFCCs, spectrograms and chromagrams to classify lung sounds into six categories, reaching an accuracy of 99.1%. Jayalakshmy et al. [20] used conditional generative adversarial networks and a CNN with scalograms as features, achieving 92.5% accuracy in four-class classification. Jasmine et al. [21, 22] introduced a novel deep learning system to identify lung disorders.

III. Materials and Methods

A. Data set

The ICBHI 2017 Challenge dataset, also known as the In- ternational Conference on Biomedical and Health Informatics 2017 Challenge dataset, is a comprehensive and widely used collection of respiratory sound recordings designed for re- search in respiratory disease detection and classification. This dataset is crucial for evaluating algorithms and models that identify various respiratory conditions based on lung sound analysis. The dataset comprises 920 recordings of respiratory sounds, totaling approximately 5.5 hours of audio, collected from 126 patients with a range of ages and clinical conditions.

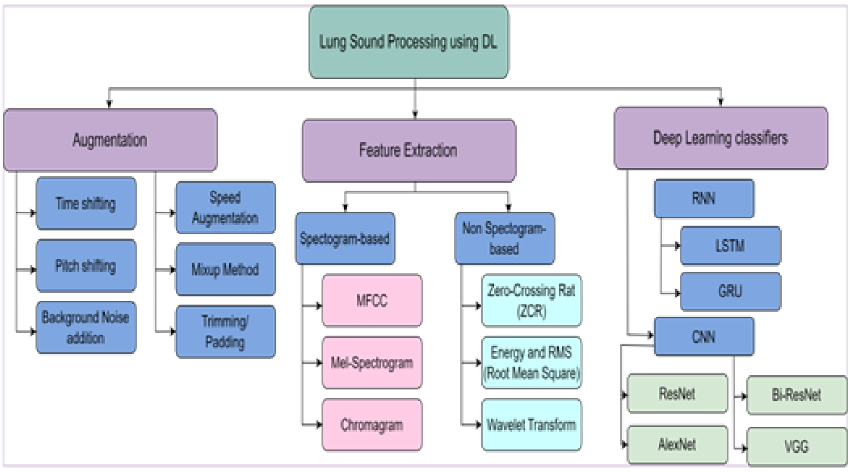

Fig. 1. Sound Processing Techniques

The recordings encompass different respiratory conditions, including healthy individuals, and patients with chronic ob- structive pulmonary disease (COPD), respiratory infections, bronchitis, pneumonia, and asthma. Each recording is anno- tated with detailed labels specifying the type of respiratory cycle (inspiration or expiration) and the presence of respiratory anomalies such as crackles and wheezes. The recordings are segmented into individual respiratory cycles, providing fine- grained labels that are useful for training and evaluating classification models. The primary goal of the ICBHI 2017 Challenge was to develop and benchmark machine learning and deep learning models for the automatic classification of respiratory sounds. This involves identifying respiratory cycles, detecting anomalies, and classifying conditions. The dataset is vital for advancing the field of respiratory sound analysis, offering a standardized benchmark for researchers to compare their models and methodologies, thereby fostering the development of robust and accurate diagnostic tools for respiratory diseases. Researchers use the dataset to train mod- els, validate techniques, and benchmark their results against those of other researchers, promoting the exchange of ideas and advancements in methodology.

B. Preprocessing

To ensure data quality, preprocessing procedures such as segmentation into respiratory cycles, signal normalization, and noise reduction using Savitzky-Golay filters are essen- tial. Lung sound characteristics are captured by the use of feature extraction techniques, such as time-domain statistics, frequency-domain transformations like Stockwell Transform and Short-Time Fourier Transform (STFT), and spectral-based features like Mel-Frequency Cepstral Coefficients (MFCCs). The ICBHI 2017 data set has made a substantial contribution to the advancement of lung sound classification methodolo- gies, despite obstacles such as class imbalance and variability in recording conditions. This has made it possible to de- velop models that improve the precision and dependability of respiratory disease diagnostics. Different techniques for classifying lung sounds are shown in Image 1. The ICBHI 2017 Challenge data set must be preprocessed in order for machine learning and analysis of the lung sound recordings to be done effectively. It guarantees that the data is uniform, clean, and appropriate for training models and extracting features.Fig.2 shows the lung sound classification model. The precise procedures for preparing the lung sound recordings are as follows:

Noise Reduction

Background noise can mask important respiratory signals in lung sound recordings. Techniques for reducing noise are used to improve the recordings' quality. The Savitzky-Golay filter is one efficient technique that smoothes audio signals by fitting successive polynomial functions to data points inside a moving window. This reduces high-frequency noise while maintaining the essential elements of lung sounds.The overall equations are shown in 1-5.

Segmentation

Lung sound recordings are continuous audio streams that need to be segmented into individual respiratory cycles (inspiration and expiration phases) for detailed analysis. This segmentation is crucial because different respiratory phases can exhibit distinct acoustic patterns.

Normalization

Normalization is used to provide consistency between record- ings by standardizing the amplitude of the audio signals. This step is crucial because differences in the environment and recording equipment can produce amplitude disparities in the signal. Among the normalization strategies are:

Min-Max Normalization: Scaling the amplitude of the audio signal to a fixed range, typically [0, 1] or [-1, 1].

Z-score Normalization: Standardizing the signal by subtract- ing the mean and dividing by the standard deviation of the signal amplitude.

Resampling: Resampling may be required if the recordings are collected at different sampling rates. Uniformity through- out the data collection is ensured by standardizing the sam- pling rate. Resampling all recordings to a common rate, such

44.1 kHz, is done to ensure consistency, for instance, if some are at 22 kHz and others at 44.1 kHz.

Feature Extraction

Lung sound categorization relies heavily on feature extraction, which converts the preprocessed audio signals into a set of features that machine learning models may utilize to find patterns and anticipate outcomes. The various feature types that can be retrieved from lung sound recordings are described in detail below:

Time-Domain Features

Time-domain features extract the essential statistical charac- teristics of lung sound recordings straight from the audio stream. The signal's core tendency is revealed by the mean, which represents the average amplitude. The variability and intensity of the signal are reflected in the variance, which quantifies the spread of amplitude values. The frequency content of the signal is shown by the zero-crossing rate (ZCR), which determines how frequently the signal crosses the zero amplitude line. When detecting pathological disorders, a higher ZCR can be a significant indicator of a higher frequency. These characteristics provide a basic description of lung sounds, which is necessary for more in-depth examination and categorization.

Frequency-Domain Features

The representation of the signal in the frequency domain is the source of frequency-domain characteristics. The spectral characteristics of the lung sounds are captured by these fea- tures.

Short-Time Fourier Transform (STFT)

A technique that computes the Fourier transform of each brief, overlapping section of the signal after splitting it into segments. This gives the signal a spectral representation that varies with time.

$$X(t,f) = \sum_{n=-\infty}^{\infty} x[n] \cdot w[n-t] \cdot e^{-j2\pi fn} \qquad (1)$$

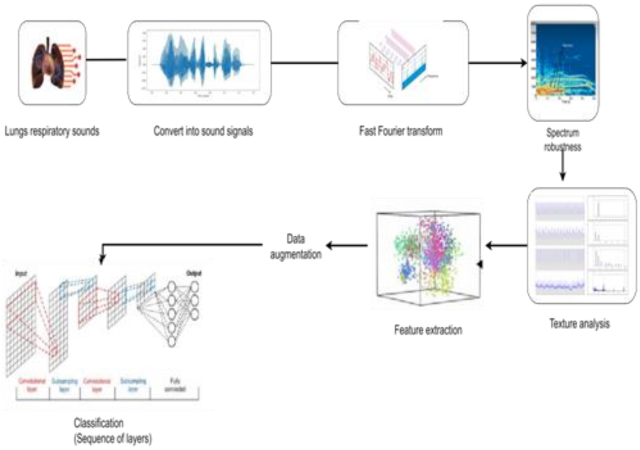

Fig. 2. Lung Sound classification model

where X(t,f) is the STFT of the signal x[n] at time t and frequency f, and w[n] is a window function that isolates the segment of the signal.

Stockwell Transform

It is sometimes referred to as the S-transform and offers a time- frequency representation of the signal by fusing wavelet and

which shows the location of the majority of the energy in the signal.

$$\text{Roll-off} = f_r \quad \text{such that} \quad \sum_{f=0}^{f_r} P(f) = 0.85 \sum_{f=0}^{f_s/2} P(f) \qquad (4)$$

$$\text{Centroid} = \frac{\sum_{f=0}^{f_s/2} f \cdot P(f)}{\sum_{f=0}^{f_s/2} P(f)} \qquad (5)$$

STFT components. It offers a multi-resolution analysis while reserving phase information.where S(t,f)is the S-transform of signal x(t) where P(f) is the power spectrum at frequency f.

C. Long Short-Term memory

$$S(t,f) = \int_{-\infty}^{\infty} x(\tau) \cdot \frac{|f|}{\sqrt{2\pi}} e^{-\frac{(t-\tau)^2 f^2}{2}} e^{-j2\pi f\tau} \, d\tau \qquad (2)$$

Spectral-Based Features

The spectrum representation of the signal is the source of spectral-based features, which offer comprehensive details on the energy distribution and frequency content.

Mel-Frequency Cepstral Coefficients (MFCCs)

These coefficients approximate the response of the human ear to various frequencies by representing the signal's short-term power spectrum on a mel scale. We first convert the signal to the frequency domain and compute the power spectrum before computing the MFCCs. After that, a Mel filter bank is applied to highlight perceptually significant frequencies. The dynamic range of the resulting Mel spectrum is compressed by taking its logarithm. Finally, we acquire MFCCs (main characteristics of the signal) by applying the DCT(Discrete Cosine Transform) on the log-Mel spectrum, which is useful for tasks like voice recognition.

$$\mathrm{MFCC}(k) = \sum_{n=1}^{N} \log(X[n]) \cdot \cos\left[\frac{\pi k}{N}\left(n - \frac{1}{2}\right)\right] \quad (3)$$

LSTM networks have demonstrated its ability to accurately simulate the temporal relationships present in audio data in the context of lung sound categorization. They handle lung sound data that has already been processed and is provided as a series of time- or frequency-domain feature vectors, each of which represents a distinct audio signal time frame. Input, forget, and output gates are used by LSTM networks to operate sequentially and record relevant temporal patterns and corre- lations, as seen in Figure 3. Because lung sounds frequently display patterns spanning numerous respiratory cycles, LSTMs are well-suited for studying lung sounds due to their ability to keep important information throughout extended periods. LSTM networks acquire hierarchical representations of these sequences through their multi-layered design, and their hidden states are particularly good at encapsulating the dynamic temporal features of lung sounds. These hidden states act as higher-order characteristics that distinguish between healthy

and unhealthy sounds. Subsequently, a fully connected layer receives the final hidden state or a combination of LSTM a measurement of the spectral distribution's skewness. where X[n] is the log-mel spectrum and k is the cepstral coefficient index. Spectral Roll-off: The frequency below which a given portion of the whole spectrum is present. It is layers, which is followed by a softmax layer for classification. With the use of this design, lung sound types can be precisely identified as each class (e.g., normal, crackle, wheeze) gives a probability output. The Enhanced LSTM (ELSTM) model is an enhancement over conventional LSTM networks that combines several changes to improve performance. It more thoroughly captures complex temporal connections and hierar chical patterns by stacking LSTM layers. By adding attention here P(f) is the power spectrum and fs is the sampling rate.

Spectral Centroid The "center of mass" of the spectrum,mechanisms, the model is better able to identify different lung sound properties by focusing attention on prominent portions

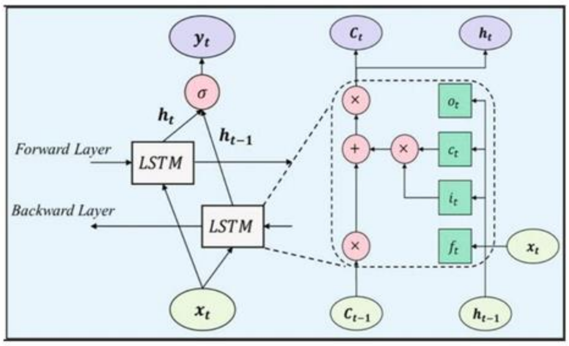

Fig. 3. LSTM model

of the input sequences. In addition, class imbalance is ad- dressed by using hybrid loss functions, which increase the sen- sitivity of the model to less common classes. One such hybrid loss function is the combination of Categorical Cross-Entropy (CCE) and Focal Loss (FL). Together, these improvements strengthen the ELSTM model's resilience and effectiveness in correctly identifying lung sounds, which greatly aids in the early identification and diagnosis of respiratory disorders.

Focal Loss Function

In lung sound classification using LSTM or ELSTM models, Focal Loss addresses the challenges of class imbalance, es- pecially when distinguishing minority classes like abnormal lung sounds (e.g., crackles or wheezes). After processing lung sound features through LSTM or ELSTM layers, the model's predictions are refined by a softmax layer, which produces class probabilities. Focal Loss modifies the standard Cross-Entropy Loss by incorporating a focusing parameter and a class balancing factor. This adjustment penalizes mis- classifications more effectively, particularly for minority class instances, by emphasizing harder-to-classify examples. Conse- quently, Focal Loss enhances the model's ability to learn and distinguish subtle features indicative of respiratory conditions, thereby improving diagnostic accuracy in lung sound analysis tasks.

$$FL = -\sum_{i=1}^{M} y_i (1-\hat{y}_i)^{\gamma} \log(\hat{y}_i), \quad \gamma \geq 0 \qquad (6)$$

Challenges

lung sounds, can bias model training and lead to poorer

perfor- mance on minority classes. Techniques like data

augmentation, class weighting, and specialized loss functions

such as Focal Loss are essential to mitigate this imbalance

and improve

model sensitivity to less frequent classes. Complexity and Variability of Lung Sounds: Lung sounds exhibit a wide range of variations influenced by factors such as patient age, body size, and underlying respiratory conditions. Capturing and interpreting these variations accurately requires models that can handle complex temporal dependencies and subtle acoustic patterns present in the audio signals. Model Interpretability: Deep learning models like LSTM and ELSTM, while effec- tive in learning from sequential data, can be challenging to interpret due to their black-box nature. Understanding how and why a model makes specific predictions, especially in medical diagnostics, is crucial for gaining trust and acceptance among healthcare professionals. Data Quality and Preprocessing: The quality of input data, including factors like noise, artifacts, and inconsistencies in recording conditions, can significantly impact model performance. Robust preprocessing techniques such as noise reduction, normalization, and feature extraction play a vital role in preparing the data for accurate classifi- cation. Scalability and Deployment: Deploying LSTM or EL- STM models in clinical settings requires consideration of scal- ability, computational efficiency, and real-time performance. Models must be optimized to handle large volumes of data efficiently while maintaining high accuracy and responsiveness in practical applications.

D. Performance measures

To assess the discrimination performance, several indicators were employed, including sensitivity, accuracy, specificity, F1- score and precision. The equations are shown in (1) to (5). These metrics were crucial in evaluating the effectiveness of discrimination. Four performance indexes were calculated: false positive index (FP), true positive (TP), false negative index (FN), and true negative index (TN). A TP (True Positive) case signifies accurate identification of COVID-19, while an FP (False Positive) case indicates misidentification of COVID- 19 as viral pneumonia. Conversely, an FN (False Negative) case occurs when COVID-19 is correctly identified as viral pneumonia, and a TN (True Negative) case denotes correct identification of viral pneumonia as COVID-19.The equations are shown in 7-11.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$
(7)

$$Precision = \frac{TP}{TP + FP}$$
(8)

In the domain of lung sound classification using LSTM or ELSTM models, several challenges need to be tackled to achieve effective and accurate diagnosis of respiratory condi- tions. These challenges include: Data Imbalance: Imbalanced datasets, where certain classes (e.g., abnormal lung sounds like crackles or wheezes) are underrepresented compared to normal

$$sensitivity = \frac{TP}{TP + FN} \quad (9)$$

$$sensitivity = \frac{TN}{TN + FP} \quad (10)$$
(10)

$$F1 - Score = 2 * \frac{\text{Precision} * \text{sensitivity}}{\text{Precision} + \text{sensitivity}} \quad (11)$$
(

Conclusion

This paper reviews the literature on lung sound classification using deep learning (DL) techniques. The analysis of various studies highlights the field's dynamic evolution and the con- tinuous efforts to improve classification model accuracy and efficiency. DL has proven to be a promising and adaptable tool for diagnosing respiratory diseases, as evidenced by its related applications. Advancements in feature extraction, data augmentation, and model explainability contribute to the robustness and versatility of these models. Despite significant progress, future research should address challenges such as implementing real-world applications, enhancing model inter- pretability, and incorporating clinical data to further validate the practical utility of DL in lung sound analysis.

References

[1] J. B. Soriano, P. J. Kendrick, K. R. Paulson, V. Gupta, E. M. Abrams, R. Adedoyin, T. B. Adhikari, S. M. Advani, A. Agrawal, E. Ahmadian et al., "Prevalence and attributable health burden of chronic respiratory diseases, 1990–2017: a systematic analysis for the global burden of disease study 2017," The Lancet Respiratory Medicine, vol. 8, no. 6, pp. 585–596, 2020.

[2] Z. Tariq, S. K. Shah, and Y. Lee, "Feature-based fusion using cnn for lung and heart sound classification," Sensors, vol. 22, no. 4, p. 1521, 2022.

[3] L. Brunese, F. Mercaldo, A. Reginelli, and A. Santone, "A neural networkbased method for respiratory sound analysis and lung disease detection," Applied Sciences, vol. 12, no. 8, p. 3877, 2022.

[4] J. Acharya and A. 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.

[5] Y. Kim, Y. Hyon, S. S. Jung, S. Lee, G. Yoo, C. Chung, and T. Ha, "Respiratory sound classification for crackles, wheezes, and rhonchi in the clinical field using deep learning," Scientific reports, vol. 11, no. 1, p. 17186, 2021.

[6] L. Fraiwan, O. Hassanin, M. Fraiwan, B. Khassawneh, A. M. Ibnian,and M. Alkhodari,"Automatic identification of respiratory diseases from stethoscopic lung sound signals using ensemble classifiers," Biocyber- netics and Biomedical Engineering, vol. 41, no. 1, pp. 1–14, 2021.

[7] D.-M. Huang, J. Huang, K. Qiao, N.-S. Zhong, H.-Z. Lu, and W.- J. Wang, "Deep learning-based lung sound analysis for intelligent stethoscope," Military Medical Research, vol. 10, no. 1, p. 44, 2023.

[8] Hema.M "Audiovisual spatial-audio analysis by means of sound localization and imaging: A multimedia

healthcare framework in ab- dominal sound mapping," IEEE Transactions on Multimedia, vol. 18, no. 10,pp. 1969–1976, 2016.

[9] Rocha, B.; Filos, D.; Mendes, L.; Vogiatzis, I.; Perantoni, E.; Kaimakamis, E.; Natsiavas, P.; Oliveira, A.; Jácome, C.; Marques, A.; et al. A respiratory sound database for the development of automated classification. In Proceedings of the International Conference on Biomedical and Health Informatics, Thessaloniki, Greece, 18–21 November 2017; pp. 33–37.

[10] Bardou, D.; Zhang, K.; Ahmad, S.M. Lung sounds classification using convolutional neural networks. Artif. Intell. Med. 2018, 88, 58–69

[11] Demir, F.; Ismael, A.M.; Sengur, A. Classification of Lung Sounds with CNN Model Using Parallel Pooling Structure. IEEE Access 2020, 8, 105376–105383

[12] Demir, F.; Sengur, A.; Bajaj, V. Convolutional neural networks based efficient approach for classification of lung diseases. Health Inf. Sci. Syst. 2020.

[13] Messner, E.; Fediuk, M.; Swatek, P.; Scheidl, S.; Smolle-Jüttner, F.M.; Olschewski, H.; Pernkopf, F. Multi-channel lung sound classification with convolutional recurrent neural networks. Comput. Biol. Med. 2020, 122, 103831.

[14] J. Acharya, A. Basu, Deep neural network for respiratory sound classifcation in wearable devices enabled by patient specifc model tuning. IEEE Trans. Biomed. Circuits Syst. 14(3), 535–544 (2020). https://doi.org/10.1109/TBCAS.2020. 2981172

[15] H. Chen, X. Yuan, Z. Pei, M. Li, J. Li, Triple-classification of respiratory sounds using optimized s-transform and deep residual networks. IEEE Access 7, 32845–32852 (2019). https://doi.org/10.1109/ACCESS.2019.2903859.

[16] S.B. Shuvo, S.N. Ali, S.I. Swapnil, T. Hasan, M.I.H. Bhuiyan, A lightweight CNN model for detecting respiratory diseases from lung auscultation sounds using EMD-CWT-based hybrid scalogram. IEEE J. Biomed. Health Inform. 25(7), 2595–2603 (2021). https://doi.org/10.1109/JBHI.2020.3048006.

[17] F. Cinyol, U. Baysal, D. Köksal, E. Babaoğlu, S.S. Ulaşlı, Incorporating support vector machine to the classification of respiratory sounds by convolutional neural network. Biomed. Signal Process. Control 79, 104093 (2023). https://doi. org/10.1016/j.bspc.2022.104093.

[18] A.A. Saraiva, D.B.S. Santos, A.A. Francisco, J.V.M. Sousa, N.M.F. Ferreira, S. Soares, A. Valente, Classification of respiratory sounds with convolutional neural network, in Proceedings of the 13th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2020)—BIOINFORMATICS, INSTICC (SciTePress, Val- letta, 2020), pp. 138–144. https://doi.org/10.5220/0008965101380144.

[19] Z. Tariq, S.K. Shah, Y. Lee, Feature-based fusion using CNN for lung and heart sound classification. Sensors 22(4), 1521 (2022). https://doi.org/10.3390/s22041521.

[20] S.Jayalakshmy, G.F. Sudha, Conditional GAN based augmentation for predictive modeling of respiratory signals. Comput. Biol. Med. 138, 104930 (2021). https://doi.org/10.1016/j.compbiomed.2021.104930.

[21] S. G. Jasmine,et.al, "Leveraging Radiomics and Genetic Algorithms to Improve Lung Infection Diagnosis in X-Ray Images Using Machine Learning," in IEEE Access, vol. 12, pp. 47656-47671, 2024, doi: 10.1109/ACCESS.2024.3383781.

[22] Jasmine, S. G,et.al. (2023). Enhancing multiclass pneumonia classifica- tion with Machine Learning and textural features.Machine Graphics and Vision,32(3/4), 83–106. https://doi.org/10.22630/MGV.2023.32.3.5