IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Hybrid Cnn-Lstm Model For Respiratory Condition Detection

[1]Santumon S D, [2]Anupama Jagadeesan, [3]Ankith Mahesh, [4]Akshara P A, [5]Bilha C Bhiju,

[1]Assisstant Professor, [2]Student, [3]Student, [4]Student, [5]Student

[1] Electronics and Communication Engineering,

[1] Vidya Academy of Science and Technology Thrissur, Kerala, India

Abstract: Respiratory diseases, including COPD, asthma, pneumonia, and other lung disorders, remain a significant global health concern, requiring early and accurate diagnosis for effective management. Traditional diagnostic methods, such as stethoscope auscultation and imaging techniques, often rely on subjective assessments and may not always provide real- time or automated analysis. As a non-invasive alternative, lung sound analysis has gained attention for its potential in detecting respiratory conditions efficiently. In this study, a hybrid Convolutional Neural Network–Long Short-Term Memory model is developed to effectively capture both spatial and temporal features of lung sounds. Mel-Frequency Cepstral Coefficients are extracted as input features to enhance the model's ability to differentiate between normal and pathological respiratory sounds. The CNN component identifies spatial patterns, while the LSTM network captures sequential dependencies within the audio signals, improving classification accuracy. To further strengthen prediction performance, patient metadata such as age, gender, medical history, environmental exposure, and reported symptoms are incorporated into the model, allowing for more context-aware and personalized diagnosis. The system is designed to function efficiently with real-time lung sound recordings, enabling early diagnosis and remote monitoring. Extensive experiments will be conducted to evaluate the model's performance based on accuracy, sensitivity, and specificity. The proposed approach has the potential to support clinicians in automated screening, telemedicine applications, and large-scale respiratory health assessments, contributing to improved patient outcomes and healthcare accessibility.

Index Terms **-** Respiratory disease detection, lung sounds, deep learning, CNN-LSTM, Mel-Frequency Cepstral-Coefficients (MFCCs), spatial and temporal features, classification accuracy, real-time monitoring.

I. Introduction

Respiratory diseases such as COPD, asthma, and pneumonia continue to be major global health concerns, necessitating early and accurate diagnosis for effective treatment. Beyond respiratory illnesses, other health concerns such as tuberculosis, lung cancer, and interstitial lung diseases also pose serious threats, often leading to complications if not detected early. The increasing prevalence of environmental pollution, smoking, and occupational hazards has further contributed to the rise in lung-related disorders. Moreover, infectious diseases like COVID-19 have highlighted the urgent need for efficient and scalable diagnostic tools to monitor respiratory health in real-time. Early detection of these conditions can significantly improve patient outcomes, reduce healthcare burdens, and enhance public health strategies. Recent advancements in deep learning have significantly improved the ability to analyze lung sounds for disease detection. Mel-Frequency Cepstral Coefficients are widely used to extract essential frequency features from lung sounds, while spectrogram conversion provides a visual representation for further analysis. By leveraging a CNN-LSTM model, the system effectively captures spatial features from spectrograms through CNNs and learns temporal dependencies using LSTMs, enhancing classification accuracy.

Fraiwan et al. [1] implemented a CNN-LSTM framework to recognize pulmonary diseases, demonstrating the effectiveness of combining spatial and temporal feature extraction. Yang and Zhang [4] proposed a CNN-based classification system, high- lighting the potential of convolutional networks in identifying abnormal lung sounds. Cruz and Pimenta [3] reviewed various machine learning techniques for lung sound classification, emphasizing the importance of high-quality datasets for model performance. Hirsch and Schneider [2] provided insights into acoustic analysis techniques, showing how spectrogram-based approaches can improve feature extraction.

Further advancements in AI-driven respiratory diagnostics include Wang and Li's [5] intelligent system that applies deep learning for lung disease detection. Gonza´lez and Farias [6] explored various applications of lung sound analysis, while Marin and Lo´pez [7] conducted a comparative study of machine learning methods, demonstrating their effectiveness in classification tasks.

Building on the findings of previous research, this study introduces a deep learning-based framework for automated respiratory disease detection using lung sound analysis. The proposed system utilizes a CNN-LSTM architecture, where CNNs extract spatial features from spectrogram representations of lung sounds, while LSTM networks capture temporal dependencies, enabling the model to recognize patterns associated with different respiratory conditions.

To enhance feature extraction, the system applies MFCCs and spectrogram conversion, which allow for a detailed analysis of frequency components in lung sound recordings. The model is developed using Python, leveraging TensorFlow and Py Torch to ensure efficient training and implementation. These frameworks provide scalability, making the system adaptable for real-world applications, including clinical diagnosis and remote healthcare monitoring.

A key factor in improving the accuracy and reliability of the model is the availability of diverse and high-quality lung sound data. Expanding the dataset with recordings from different age groups, disease conditions, and environmental settings can enhance the model's ability to generalize across various patient populations. By continuously improving the dataset and optimizing the deep learning model, this approach aims to provide a non-invasive, automated, and scalable solution for early respiratory disease detection, contributing to more accessible and efficient healthcare diagnostics.

This section examines deep learning techniques for respiratory disease detection, emphasizing the CNN-LSTM architecture, MFCC-based feature extraction, and spectrogram analysis to effectively classify lung sounds. The remainder of the sections is structured as follows: Section II reviews related literature survey in lung disease detection using lung sound. Section III details the system architecture and methodology. Section IV presents results and analysis, and Section V concludes with future work and insights.

II. Proposed Model

This section presents a comprehensive analysis of the pro- posed model designed for respiratory disease detection using lung sound signals. The system integrates piezoelectric sensor technology with deep learning algorithms, enabling precise and non-invasive monitoring of pulmonary conditions.

Piezoelectric sensors are employed to record lung sounds due to their heightened sensitivity and immunity to ambient noise. These sensors transduce mechanical vibrations generated by breathing into electrical signals, making them particularly advantageous in clinical settings. Unlike conventional microphones that are susceptible to environmental interferences, piezoelectric sensors are better suited to capture internal acoustic signals, resulting in cleaner, high-fidelity lung sound recordings.

To compensate for the low amplitude of lung sounds, a signal amplification unit is included to enhance weak signals before processing. This ensures that subtle acoustic mark- ers such as wheezes, crackles, and rhonchi are retained for analysis. A digital filtering mechanism is also implemented to eliminate irrelevant noise sources, including heart sounds and background disturbances. This pre-processed signal is then converted into Mel-Frequency Cepstral Coefficients (MFCCs), which effectively represent the spectral properties of the signal by modeling how the human auditory system perceives sound. MFCCs are particularly useful for distinguishing frequency- based features associated with specific respiratory diseases.

The extracted MFCC features are input into a hybrid deep learning model composed of a Convolutional Neural Network (CNN) and a Long Short-Term Memory (LSTM) network. The CNN is tasked with identifying spatial patterns within spectrogram-like input features, while the LSTM layer captures temporal dynamics and sequential dependencies in breathing cycles. This combination ensures robust detection of various respiratory conditions, including asthma, chronic obstructive pulmonary disease (COPD), bronchitis, and pneumonia.

To further refine predictions, patient-specific metadata such as age, gender, medical history, and environmental exposure can be incorporated into the model as auxiliary input. By fusing this metadata with audio-derived features at the decision level or.

within a multimodal neural network architecture, the system can achieve personalized and context-aware disease prediction. For

instance, wheezing in a child may indicate a different pathology compared to a similar acoustic pattern in an elderly patient with

a history of smoking.

The model's performance can be further enhanced by integrating next-generation hardware components. Improved piezoelectric sensors with extended frequency response ranges and better calibration will allow for more accurate detection of subtle anomalies in lung sounds. Using sensor arrays placed at multiple thoracic locations can enable comprehensive mapping of respiratory sounds from different lung zones, improving diagnostic reliability.

From a signal processing perspective, the adoption of adaptive gain amplifiers and real-time digital filtering methods such as wavelet-based denoising and adaptive noise cancellation can significantly increase the signal-to-noise ratio, enhancing the clarity and diagnostic value of the recordings.

In the realm of feature extraction, future systems may integrate hybrid spectral analysis techniques, including Wavelet Transforms and Short-Time Fourier Transform (STFT), along- side MFCCs, to capture both time localized and frequency- specific patterns in lung sounds. Furthermore, deep autoencoders and unsupervised feature learning mechanisms could be employed to uncover hidden acoustic patterns related to rare or emerging respiratory diseases.

or practical deployment, leveraging Edge AI hardware, such as AI-accelerated microcontrollers and neuromorphic chips, would facilitate real-time inference on portable or wearable devices. These embedded systems can deliver diagnostic insights without reliance on cloud computing, offering a scalable solution for rural or resource-constrained healthcare environments.

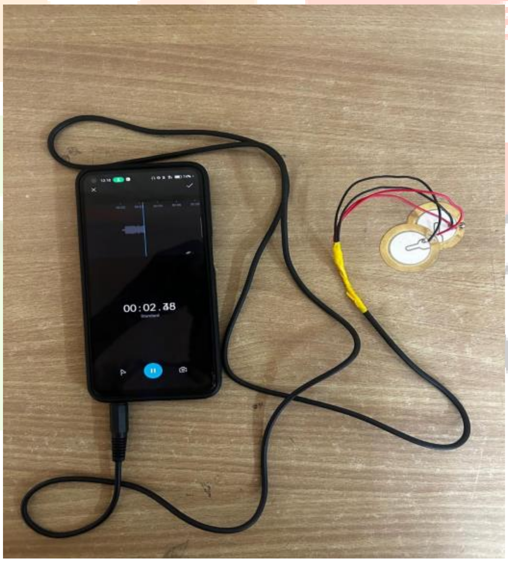

Figure 1:Hardware Model

Ultimately, the integration of multimodal data combining acoustic features, metadata, and physiological parameters within a unified AI framework will enable a fully automated, real-time respiratory health monitoring system. Such a system would offer high diagnostic accuracy, timely interventions, and enhanced accessibility for both clinical practitioners and patients managing chronic respiratory conditions at home.

III. Design And Methodology

The proposed approach for respiratory disease detection using lung sounds follows a systematic pipeline, integrating deep learning techniques to enhance classification and detection accuracy. The methodology comprises data preprocessing, feature extraction, model architecture design, and evaluation by integrating hardware-software.

A. System Architecture

The system architecture, illustrated in Figure 1, utilizes a piezoelectric sensor or a MEMS (Micro-Electro-Mechanical System) microphone to capture lung sounds, depending on the required sensitivity and noise reduction capabilities. If necessary, the recorded signals undergo amplification to enhance clarity and minimize background noise before further processing.

Once acquired, the lung sounds are transformed into spectrograms, which visually represent frequency variations over time. This conversion enables deep learning models to analyze both spatial and temporal characteristics of the sound. The CNN extracts spatial features from the spectrograms, identifying distinct patterns, while the LSTM processes temporal dependencies, capturing variations in respiratory cycles.

The extracted features are then compared against a pre- trained model, developed using a diverse dataset of normal and abnormal lung sounds. By analyzing these patterns, the system accurately classifies the input as either normal or indicative of a respiratory disease. This automated approach enhances early detection, offering a promising solution for clinical diagnosis and respiratory health monitoring.

Figure 2, illustrates the spectrogram obtained by plotting the magnitude of the frequency components over time. It pro- vides a time-frequency representation of lung sounds, making it useful for visualizing abnormal respiratory patterns. These extracted features are then fed into machine learning or deep learning models, such as CNN-LSTM, for classification. The effectiveness of disease detection improves with a diverse dataset, ensuring robust feature learning.

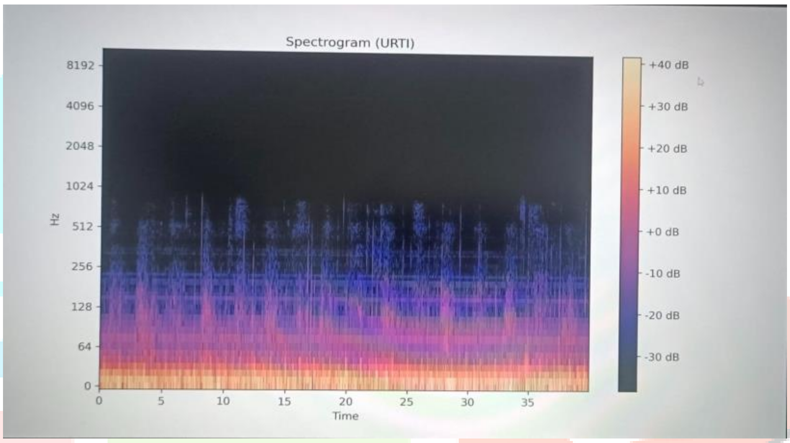

Figure 2. Spectrogram

B. Mathematical Design and Modelling

This section explores the methodologies behind the lung disease detection using lung sound, covering ML algorithms like cosine similarity MFCC feature extraction, dataset details and workflow.

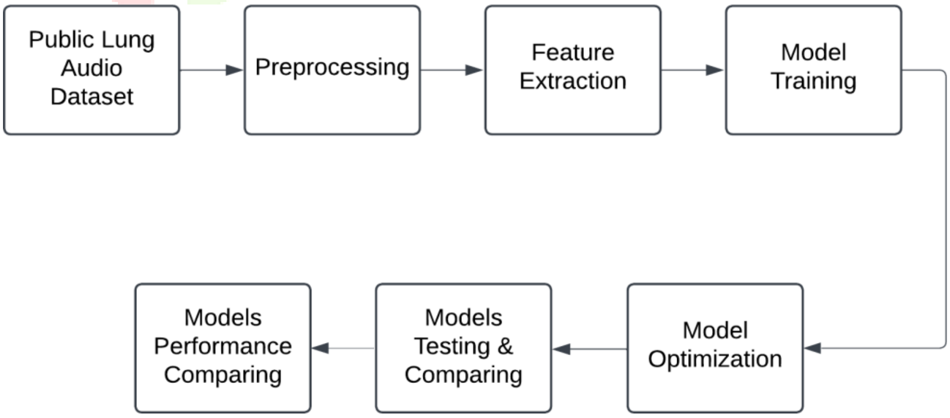

Figure 3. Block diagram of system

i. **Signal Acquisition and Preprocessing**: Lung sound signals (S(t)) are captured using a piezoelectric sensor, which converts mechanical vibrations into electrical signals:

$$V(t) = k\_p \cdot S(t) \qquad (1)$$

where:

- V(t) = Output voltage from the piezoelectric sensor
- k_p = Piezoelectric sensitivity constant
- S(t) = Lung sound signal in time domain

ii. **Mel-Frequency Cepstral Coefficients Calculation**: The Mel-Frequency Cepstral Coefficients are extracted from the lung sound signals to capture key acoustic features. The MFCC calculation follows these steps:

1. Compute the logarithm of the Mel-filtered signal:

$$\mathbf{S_m} = \sum_{\mathbf{k=1}}^{\mathbf{K}} \mathbf{X_k H_m(k)} \qquad (2)$$

Where:

- X_k is the magnitude of the Fourier-transformed signal
- H_m(k) is the Mel filter bank
- S_m is the Mel-scaled spectral energy

2. Apply the Discrete Cosine Transform (DCT) to obtain MFCCs:

$$C_n = \sum_{m=1}^{M} \log(S_m) \cdot \cos\left(\frac{n(m-0.5)\pi}{M}\right) \qquad (3)$$

Where:

- C_n is the n-th MFCC coefficient,
- M is the total number of Mel filters,
- S_m is the log energy of the Mel filter bank output

The resulting MFCCs provide a compact and discriminative representation of lung sounds, which are used as inputs to the classification model

iii. Cosine Similarity for Classification: Once the MFCC features are extracted, cosine similarity is used to compare new lung sound features with the trained dataset. Cosine similarity measures the angle between two feature vectors A and B as:

$$\text{Cosine similarity} = \frac{A \cdot B}{\|A\| \, \|B\|} \qquad (4)$$

Where,

- $A \cdot B$ is the dot product of the feature vectors.
- $\| A \|$ and $\| B \|$ are the Euclidean norms (magnitudes) of the vectors.

A similarity threshold T is set to determine if a given lung sound matches a known respiratory disease pattern:

$$\boldsymbol{S(A,B) \geq T \Rightarrow \{Detected\ Disease\}} \qquad \textbf{(5)}$$

The recorded signals undergo preprocessing, including noise reduction and normalization, to ensure clarity. Next, the pre- processed audio is converted into a spectrogram representation and Mel-Frequency Cepstral Coefficients are extracted to capture key sound features. These extracted features serve as inputs to a CNN-LSTM model, where Convolutional Neural Networks analyze spatial patterns in the spectrogram, and Long Short-Term Memory networks process temporal dependencies in the lung sounds.

The flowchart illustrates ,the workflow and trained model that compares the extracted features with stored disease patterns using cosine similarity, identifying abnormalities and classifying respiratory conditions. The system's accuracy improves with the inclusion of more diverse lung sound data, making it a scalable solution for real-time respiratory disease detection.

IV. Result And Analysis

The deep learning model was trained and tested using the ICBHI dataset, which consists of lung sound recordings from individuals with various respiratory conditions. Feature extraction was performed using MFCCs to capture essential acoustic patterns .Real-time data collection for the system is achieved using a piezoelectric crystal, which captures lung sounds and vibration into electrical signals. The collected sound data is transformed into a spectrogram representation, allowing the extraction of key features using Mel-Frequency Cepstral Coefficients. This processed data serves as input to the CNN-LSTM model, which classifies respiratory conditions based on trained patterns. Additionally, expanding the training data improved the model's ability to generalize, reducing misclassification rates and enhancing reliability. These results emphasize the effectiveness of deep learning for automated respiratory disease detection.

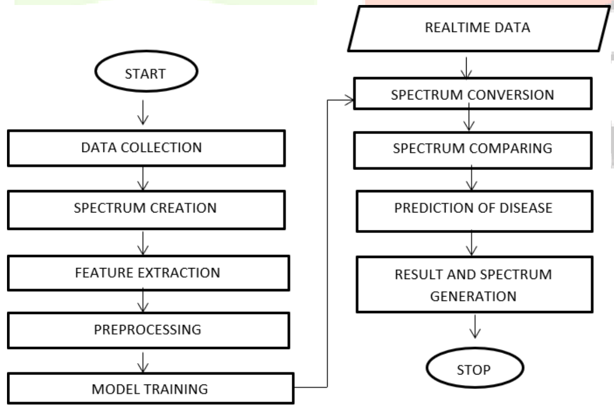

Figure 4. Workflow of system

A. Performance Matrices

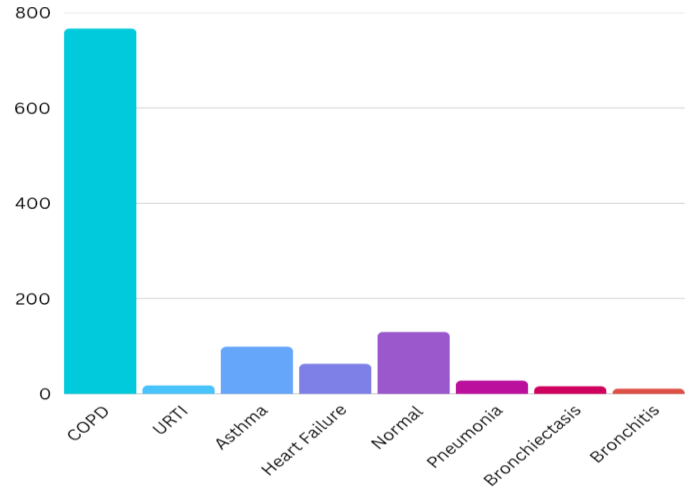

Figure 5. Graphical representation of disease classification in ICBHI 2017

The chart displays the prevalence of various respiratory and cardiovascular conditions. COPD is the most common condition, significantly outnumbering all others. Other conditions include Upper Respiratory Tract Infection , Asthma, Heart Failure, Pneumonia, Bronchiectasis, and Bronchitis, with ”Normal” cases also represented for comparison. The distribution highlights that COPD is the dominant condition among the cases, while others occur with much lower frequency.

B. Interface

The deep learning-based respiratory disease detection model was evaluated using a dataset of lung sound recordings, comparing its performance with traditional diagnostic methods. The system significantly improved efficiency, reducing analysis time while maintaining high diagnostic accuracy. The model achieved a average classification accuracy, surpassing conventional auscultation methods. Additionally, it minimized subjective variability in diagnosis, ensuring more consistent and objective evaluations. These findings demonstrate the potential of deep learning to enhance the speed, accuracy, and reliability of respiratory disease detection.

This approach has the potential to successfully develop an real time system for respiratory disease detection using lung sound analysis. The system integrates key components, including MFCC feature extraction, spectrogram conversion, and deep learning-based classification. Lung sounds are processed using a CNN-LSTM model, which effectively captures both spatial and temporal patterns for accurate disease identification. The system is implemented in Python using TensorFlow/Py Torch, ensuring efficiency and scalability. Additionally, incorporating a larger and more diverse dataset of lung sounds for training can further enhance model accuracy and robustness. This approach provides a non-invasive, auto- mated, and objective diagnostic solution, significantly reducing manual effort while improving efficiency and reliability in respiratory disease detection.

Figure 6. Interface model

V. CONCLUSION

This study presents a deep learning-based method for respiratory disease detection using lung sound recordings. The system uses piezoelectric sensors or MEMS microphones for accurate sound acquisition with minimal background noise .Key audio features are extracted using Mel-Frequency Cepstral Coefficients (MFCCs) and converted into spectrograms, allowing the model to learn both spatial and temporal patterns in the respiratory signals. To enhance prediction accuracy, the model also incorporates patient-specific metadata such as age, symptoms, and medical history. Additionally, integrating live data from wearable sensors or continuous monitoring enables real-time updates, improving early detection and personalized healthcare support.

The classification is performed using a CNN-LSTM architecture, where the CNN extracts spatial features, and the LSTM captures temporal dependencies in the lung sound sequences. The model is implemented in Python using frame- works like TensorFlow and Py Torch, ensuring efficient training and deployment.

Experimental results demonstrate the system's ability to accurately distinguish between normal and abnormal lung sounds, highlighting its potential for early and automated respiratory disease detection by incorporating more data in future. This method offers a non-invasive, cost-effective, and scalable solution that can be integrated into clinical applications and telemedicine for improved respiratory health monitoring.

REFERENCES

[1] M. Fraiwan, L. Fraiwan, M. Alkhodari, O. Hassanin, "Recognition of pulmonary diseases from lung sounds using convolutional neural net- works and long short-term memory", Journal of Ambient Intelligence and Humanized Computing (2022).

[2] Hirsch, M., Schneider, J. "Acoustic analysis of lung sounds for disease detection: A review." Journal of Medical Systems, 42(10), 1-12, 2018.

[3] Cruz, A., Pimenta, J. "Machine learning techniques for lung sound classification: A systematic review." Biomedical Signal Processing and Control, 58, 101840, 2020.

[4] Yang, Y., Zhang, H. "Development of a lung sound classification system based on convolutional neural networks." IEEE Access, 7, 162489-162497, 2019.

[5] Wang, J., Li, Y. "An intelligent system for lung disease detection using respiratory sounds and deep learning." Artificial Intelligence in Medicine, 104, 101837, 2020.

[6] Gonza´lez, A., Farias, D. "Analysis of lung sounds: An overview of techniques and applications." Health Information Science and Systems, 9(1), 1-14, 2021.

[7] Marin, J., Lo´pez, A. "A comparative study of machine learning meth- ods for lung sound classification." Computer Methods and Programs in Biomedicine, 218, 106706, 2022.

[8] Forum of International Respiratory Societies. The Global Impact of Respiratory Disease—Second Edition. Sheffield, UK: European Respi- ratory Society; 2017.

[9] Gairola, S. et al. "Respirenet: A deep neural network for accurately detecting abnormal lung sounds in a limited data setting." In Pro- ceedings of the 2021 43rd Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC), Mexico City, Mexico, 1–5 November 2021, pp. 527–530.

[10] Ma, Y. et al. "Lung RN+NL: An Improved Adventitious Lung Sound Classification Using Non-Local Block ResNet Neural Network with Mixup Data Augmentation." In Proceedings of the Interspeech 2020, Shanghai, China, 25–29 October 2020.

[11] Pham, L.D. et al. "CNN-MoE Based Framework for Classification of Respiratory Anomalies and Lung Disease Detection." IEEE Journal of Biomedical and Health Informatics, 2021, 25, pp. 2938–2947.

[12] Nguyen, T. et al. "Lung Sound Classification Using Co-tuning and Stochastic Normalization." arXiv preprint arXiv:2108.01991, 2021.

[13] Nguyen, T. et al. "Lung Sound Classification Using Snapshot Ensemble of Convolutional Neural Networks." In Proceedings of the 42nd Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC), Montreal, QC, Canada, 20–24 July 2020. Rocha, B.M.; Pessoa, D.; Marques, A.; Carvalho, P.; Paiva, R.P. "Auto- matic Classification of Adventitious Respiratory Sounds: A (Un)Solved Problem?" Sensors 2020, 21, 57. [CrossRef] [PubMed]

[14] Padilla-Ortiz, A.L.; Ibarra, D.; Padilla, A. "Lung and Heart Sounds Analysis: State-of-the-Art and Future Trends." Crit. Rev. Biomed. Eng. 2018, 46, 33–52. [CrossRef] [PubMed]

[15] Watt, J.; Borhani, R.; Katsaggelos, A.K. "Machine Learning Refined: Foundations, Algorithms, and Applications," 2nd ed.; Cambridge Uni- versity Press: Cambridge, UK, 2020. [CrossRef]

[16] Pramono, R.X.A.; Bowyer, S.; Rodriguez-Villegas, E. "Automatic adventitious respiratory sound analysis: A systematic review." PLoS ONE 2017, 12, e0177926. [CrossRef] [PubMed]

[17] Rocha, B.M.; Filos, D.; Mendes, L.; Serbes, G.; Ulukaya, S.; Kahya, Y.P.; Jakovljevic, N.; Turukalo, T.L.; Vogiatzis, I.M.; Perantoni, E.; et al. ”An open access database for the evaluation of respiratory sound classification algorithms.” Physiol. Meas. 2019, 40, 035001. [CrossRef] [PubMed]

[18] Mazic´, I.; Bonkovic´, M.; Dzˇaja, B. ”Two-level coarse-to-fine classifi- cation algorithm for asthma wheezing recognition in children’s respi- ratory sounds.” Biomed. Signal Process. Control 2015, 21, 105–118. [CrossRef]

[19] Matsutake, S.; Yamashita, M.; Matsunaga, S. ”Abnormal-respiration detection by considering correlation of observation of adventitious sounds.” In Proceedings of the 23rd European Signal Processing Conference (EUSIPCO), Nice, France, 31 August–4 September 2015;pp. 634–638. [CrossRef]