Journal of Theoretical and Applied Information Technology
15th July 2025. Vol.103. No.13
© Little Lion Scientific
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

HYBRID CNN WITH ATTENTION-BASED FEATURE FUSION FOR LUNG DISEASE DETECTION FROM RESPIRATORY SOUNDS USING CLASS-BALANCED OPTIMIZATION

[1]Dr. P. GNANASUNDARI, [2] Dr. PRADIP RAM SELOKAR, [3]Dr. S. GOPINATH, [4]V.KIRUTHIKA, [5]B.SUGANTHI, [6]EPHIN M

[1]Associate Professor, School of Business, RV University, Bangalore, Karnataka
[2]Assistant Professor, Department of Electronics and Communication Engineering, Ramdeobaba University, Nagpur.
[3]Professor, Department of Electronics and Communication Engineering, Karpagam Institute of Technology, Coimbatore
[4]Assistant professor, Department of ECE, Sri Eshwar College of Engineering, Coimbatore
[5]Associate Professor, Department of ECE, RVS College of Engineering and Technology, Coimbatore
[6]Senior Faculty in the Department of Information Technology at Gloucestershire University, Achievers International Campus, UAE.
E-mail: [1]gnanasundarip@rvu.edu.in, [2]selokarpr@rknec.edu, [3]gopi.vasudev@gmail.com, [4]kiruthika.v@sece.ac.in, [5]suganth.b@gmail.com, [6]ephin0204@gmail.com

ABSTRACT

Lung disease detection using respiratory sounds involves analyzing audio recordings of breath sounds to identify abnormalities associated with conditions such as Chronic Obstructive Pulmonary Disease (COPD), pneumonia, and asthma. However, accurate detection remains challenging due to class imbalance in datasets and the difficulty of extracting discriminative features from complex audio signals. This study proposes a novel Hybrid Convolutional Neural Network with Attention-Based Feature Fusion (HCNN-AFF) to enhance detection accuracy. The model integrates spectrogram-based inputs with Mel-Frequency Cepstral Coefficients (MFCCs), effectively capturing both time-frequency and cepstral information for richer feature representation. To address class imbalance, a combination of data augmentation and a class-balanced loss function is employed, enabling improved learning from minority class samples. Furthermore, the model is optimized using the AdamW optimizer enhanced with a LookAhead mechanism (AWLA), promoting better convergence and generalization. Experimental results demonstrate that the proposed HCNN-AFF method outperforms existing approaches in respiratory sound analysis, achieving superior accuracy and robustness in lung disease detection.

Keywords: *Lung Disease, Respiratory Sounds, Machine Learning, Feature Extraction, Mel-Frequency Cepstral Coefficients, Early Diagnosis..*

1. INTRODUCTION

Lung disease detection by analyzing acoustic signals produced during breathing is a promising non-invasive approach for identifying abnormalities indicative of respiratory conditions such as asthma, Chronic Obstructive Pulmonary Disease (COPD), pneumonia, and bronchiolitis. These diseases often manifest through altered respiratory sounds—such as wheezes, crackles, and rhonchi—that can reflect underlying airway obstructions, lung tissue changes, or fluid secretions. As a result, respiratory sound analysis offers a cost-effective alternative to imaging methods like CT scans, which are expensive, less accessible in resource-constrained settings, and involve radiation exposure [1, 2].

Traditional signal processing techniques for analyzing respiratory sounds rely on handcrafted features such as Mel-Frequency Cepstral Coefficients (MFCCs), spectral features (e.g., spectral centroid, bandwidth, roll-off), and zero-crossing rate. MFCCs are widely used for capturing the timbral qualities of audio, while spectral features help characterize frequency-related attributes of the signal. These features have shown effectiveness in distinguishing certain respiratory conditions [3]. However, they suffer from several limitations in real-world applications. Respiratory sounds are often contaminated by background noise, and inter-patient variability (age, gender, disease severity) further complicates signal interpretation. Additionally, the quality of recordings can vary significantly depending on the device used. Manual feature engineering in these traditional methods is time-consuming, highly dependent on domain expertise, and not scalable for complex datasets with subtle distinctions between disease states [4, 5].

To overcome these challenges, **deep learning** has emerged as a powerful alternative. Unlike traditional methods, deep learning models can automatically learn hierarchical and abstract features from raw or minimally processed data. Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and hybrid architectures are increasingly applied to respiratory sound classification, as they excel at recognizing complex temporal and frequency patterns, even under noisy conditions. These models are capable of identifying pathological sounds such as wheezes and crackles, improving the reliability of automated diagnosis [6].

Despite their promise, existing deep learning models still face key limitations: inadequate fusion of complementary audio features, poor generalization to imbalanced datasets, and slow convergence or overfitting during training. To address these gaps, this study proposes a **Hybrid Convolutional Neural Network with Attention-Based Feature Fusion (HCNN-AFF)**, which combines the strength of CNN-based feature extraction with an attention mechanism that adaptively fuses spectrogram and MFCC-based features. This fusion enables the model to effectively capture both time-frequency and cepstral patterns critical for accurate classification. Furthermore, we integrate the **AdamW optimizer with LookAhead (AWLA)**, which enhances the training process by improving convergence speed and reducing the likelihood of overfitting.

The main contributions of this work are as follows:

1. **A novel hybrid architecture (HCNN-AFF)** that integrates CNN with attention-based feature fusion to enhance discriminative capability.
2. **Utilization of spectrogram and MFCC features**, capturing complementary aspects of respiratory sounds for robust representation.
3. **Incorporation of the AWLA optimizer**, which accelerates convergence and improves generalization by overcoming local minima and overfitting.
4. **Robustness to noisy and imbalanced datasets**, making the model suitable for deployment in real-world clinical environments.

The research work introduced feature fusion with attention mechanism, unlike traditional CNN-based approaches.
The structure of the paper is described as follows. Section II gives the recent related work in lung disease detection. Section III explains the proposed work and its work flow in disease detection. Section IV discusses the results obtained by proposed feature extraction and classification techniques. The conclusion and future work is discussed in Section V.

2. RELATED WORK

Dianat et al. (2023) present a deep learning approach for classifying pulmonary sounds for diagnosing Interstitial Lung Diseases (ILD) secondary to connective tissue diseases. It uses CNN for the classification achieving promising results in identifying different lung conditions from recorded pulmonary sounds [8]. Karaarslan et al. (2024) explored the classification and characterization of respiratory diseases based on respiratory sounds using deep learning techniques. The feature extraction of relevant features from respiratory sounds is processed by machine learning models for different classifications of lung diseases [9]. Taloba and Matoog (2025) discussed machine learning-based pattern recognition to identify respiratory diseases through spirometry data. It proposed advanced pattern recognition techniques on lung function tests for classifying various respiratory conditions. This approach proved effective in detecting certain diseases based on spirometry data, but more data is needed for broader applicability [10]. Pessoa et al. (2024) presented an ensemble deep learning model designed for the estimation of dimensionless respiratory airflow using respiratory sound. The ensemble model made from various deep learning models improved the airflow estimation remarkably, even in noise [11].

Journal of Theoretical and Applied Information Technology
15th July 2025. Vol.103. No.13
© Little Lion Scientific
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Fava et al. (2024) examined pre-processing techniques to improve the classification of lung sounds using deep learning. Advanced signal processing methods applied to clean and prepare lung sound data improved classification performance. It established deep learning models' accuracy is dependent on the input data. However, the pre-processing strategy has high computation time [12]. Hassan et al. (2024), the authors presented EasyNet, a lightweight deep learning model for detecting lung diseases. EasyNet was designed to be computationally efficient and can work on devices with minimal computational resources, which would be ideal for portable diagnostic systems [13]. TaghiBeyglou et al.(2024) proposed TRespNET: A dual-route CNN tailored to identify pediatric adventitious respiratory sounds. It is designed to classify abnormal respiratory sounds in children, which enhances accuracy in pediatric classification [14]. Shen et al. (2024) proposed new feature extraction and data augmentation techniques for cough-based respiratory disease classification. It has the improved the feature extraction process and augmented the dataset by enhancing the model's robustness and performance. However, data augmentation was helpful but resulted in overfitting[15].

Sharan (2023) discussed a raw audio-based sound detection system utilizing SincNet and GRU-bidirectional structures., it is witnessed that by using deep learning techniques the cough sounds were identified without any processing into audio signals, which enhanced the sensitivity and specificity of detection systems for cough sounds [16]. Wang et al. (2024) proposed a lightweight network called LungNeXt for lung sound classification, with enhanced mel-spectrograms as input features. The LungNeXt is efficient and effective, giving an alternative to heavier models while achieving high accuracy in the classification of lung sounds [17]. Cansiz et al. (2024) applied the tunable Q-factor wavelet transform for lung signal decomposition and statistical feature extraction for better classification of lung diseases. This approach improved the accuracy of classification because it extracted more features from lung sounds, resulting in better performance in diagnosis [18]. Nitha and SS (2024) introduced the CEFNet framework for the detection of lung diseases and identification of the region of infection. This model primarily focused on the detection of infection regions in lung sounds and significantly improved the accuracy of classification of lung diseases [19]. Gupta et al. (2024) introduced DeepRespNet, a deep neural network designed for the classification of respiratory sounds. This model showed good performance in identifying various lung conditions; however, issues with noise and dataset diversity persisted [20, 21 & 22].

Despite advancements, existing deep learning models for lung disease detection struggle with class imbalance, noise, and generalization. This study addresses these issues by proposing HCNN-AFF, a hybrid model with attention-based feature fusion and AWLA optimization. Key research questions explore improvements in accuracy, robustness, and performance under real-world conditions.

Despite advancements in machine learning and signal processing, accurate detection of lung diseases from respiratory sounds remains a significant challenge. Clinical auscultation is subjective, and imaging tools like CT scans are costly and inaccessible in many regions. Traditional methods relying on handcrafted features struggle with noise, inter-patient variability, and data imbalance, often failing in real-world conditions. Deep learning offers potential but lacks robustness when faced with diverse, noisy, or underrepresented data. Therefore, there is a pressing need for a reliable, generalizable, and scalable solution that can effectively analyze respiratory sounds and support early, non-invasive diagnosis of lung diseases in diverse settings.

3. IMPLEMENTATION OF EFFICENT MULTIPATH LOCATION AWARE ROUTING PROTOCL

The proposed work combines Mel Spectrograms with MFCC to improve feature extraction in respiratory sound classification. Mel Spectrograms capture the time-frequency representation of audio signals and are effective in representing the frequency components of respiratory sounds. It provides features of interest through a logarithmic scale for lung disease detection. But Mel Spectrograms are good at catching the distribution of energy over time but will fail to capture fine-grained patterns of that sound.

Integrating a Hybrid Convolutional Neural Network with Attention-Based Feature Fusion (HCNN-AFF), combined with spectrogram and MFCC feature inputs and optimized using the AdamW optimizer with a LookAhead mechanism (AWLA), will significantly improve the accuracy, robustness, and generalizability of lung disease detection from respiratory sounds, particularly in noisy and class-imbalanced real-world datasets.

In order to overcome this limitation, MFCCs which are typically used in speech and sound recognition are merged with Mel Spectrograms. The MFCC captures the perceptual characteristic of the sound by compressing the Mel Spectrogram into a set of coefficients that represent the short-term power spectrum. This combination enables the model to leverage the temporal and spectral details captured by the Mel Spectrogram while improving the representation of the underlying features of the audio via the MFCCs. It is expected that combining the two features will improve the ability of the model to classify respiratory sounds and be more effective for detecting and diagnosing lung diseases. In equation (!) the mel spectrogram converts an audio signal into a time frequency representation with a mel scale where $X(f,t)$ is the short time fourier transform magnitude of the audio signal and $H_{mel}(f,m)$ is the filter bank for mel scale.

$$S_{mel}(t,m) = \sum_{f} X(f,t).H_{mel}(f,m) \quad (1)$$

Equation (2) describes the logarithmic mel spectrogram using a discrete cosine transforms where M is the number of mel bins and $\log(S_{mel}(t,m))$ is the logarithmic mel spectrogram.

$$C_k = \sum_{m=1}^{M} \log(S_{mel}(t,m)).\cos(k.\frac{\pi(m-0.5)}{M}) \quad (2)$$

Equation (3) represent the feature extraction by applying the convolutional filters over the input where $W^{l}_{m,n}$ are the weights of the convolutional kernel, $F^{l}_{(i+m),(j+n)}$ is the feature map from previous layer , b^l is the bias term and σ is the activation function (softmax)

$$F^{l}_{i,j} = \sigma\left(\sum_{m,n} W^{l}_{m,n} \cdot F^{l}_{(i+m),(j+n)} + b^{l}\right) \quad (3)$$

In equation (4), the attention mechanism weights feature maps based on the relevance where $e_{i,j}$ is the attention score for position (i,j) and $A_{i,j}$ is the attention enhanced feature at (i,j).

$$\alpha_{i,j} = \frac{\exp(e_{i,j})}{\sum_{k,l}\exp(e_{k,l})}, A_{i,j} = \alpha_{i,j}.F_{i,j} \quad (4)$$

The AdamW with Lookahead optimizer updates the weights iteratively as given in equation (5) & (6) where θ_t are the model weights at iteration t, η is the learning rate, m_t and v_t are the fisrt and second moment estimates in AdamW. λ is the weight decay, α is the interpolation factor and θ_{fast} are the fast-moving weights in LookAhead.

$$\theta_{t+1} = \theta_t - \eta \cdot \left(\frac{m_t}{\sqrt{v_t}} + \lambda \cdot \theta_t\right) \quad (5)$$

$$\theta_{Lookahead} = \alpha.\theta_{t+1} + (1-\alpha) \cdot \theta_{fast} \quad (6)$$

The accuracy of the model is calculated using the equation (7).

$$Acc = \frac{True\ positive + True\ negative}{True\ positive + True\ negative + False\ positive + False\ negative} \quad (7)$$

Comparable approaches have also been used in other industries—such as speech emotion recognition in human-computer interaction and machine fault detection in manufacturing—where acoustic signals are analyzed using time-frequency representations and deep neural architectures. These cross-disciplinary parallels reinforce the validity and transferability of a deep learning-based experimental approach for solving complex, real-world acoustic classification problems.

The Hybrid Convolutional Neural Network is a deep learning architecture geared towards high accuracy and efficacy in extracting features and feature classification for tasks such as medical image processing and signal processing that require complexity in the processing. CNNs can very easily learn spatial hierarchies that extract local features in an image or signal. It is fail at highlighting essential parts or structures within the data to focus upon. To overcome this, the model incorporates the HCNN-AFF with an attention mechanism which enables the network to give various weights to different regions of input such that it would point towards more relevant features to make classifications.

HCNN-AFF allows the ability to augment the CNN to learn both the spatial dependencies of data but also temporal or contextual dependencies that could be pivotal to achieving accurate predictions.

The combination of attention-based features with CNN layers would enhance the model's capacity to focus on important parts of the input, which is specific textures or anomalies in medical signals. This hybrid approach integrates the computational efficiency and capabilities of local feature extraction through CNNs with the adaptability and focused learning provided by the attention mechanism.

The HCNN-AFF combined with the AdamW optimizer and LookAhead is an advanced deep learning architecture designed to improve model performance in complex tasks such as medical diagnosis. The HCNN-AFF uses CNNs for efficient local feature extraction and an attention mechanism to enhance focus on relevant parts of the input data. This combination helps the model learn spatial and contextual relationships within the data by identifying patterns, anomalies, or regions of interest.

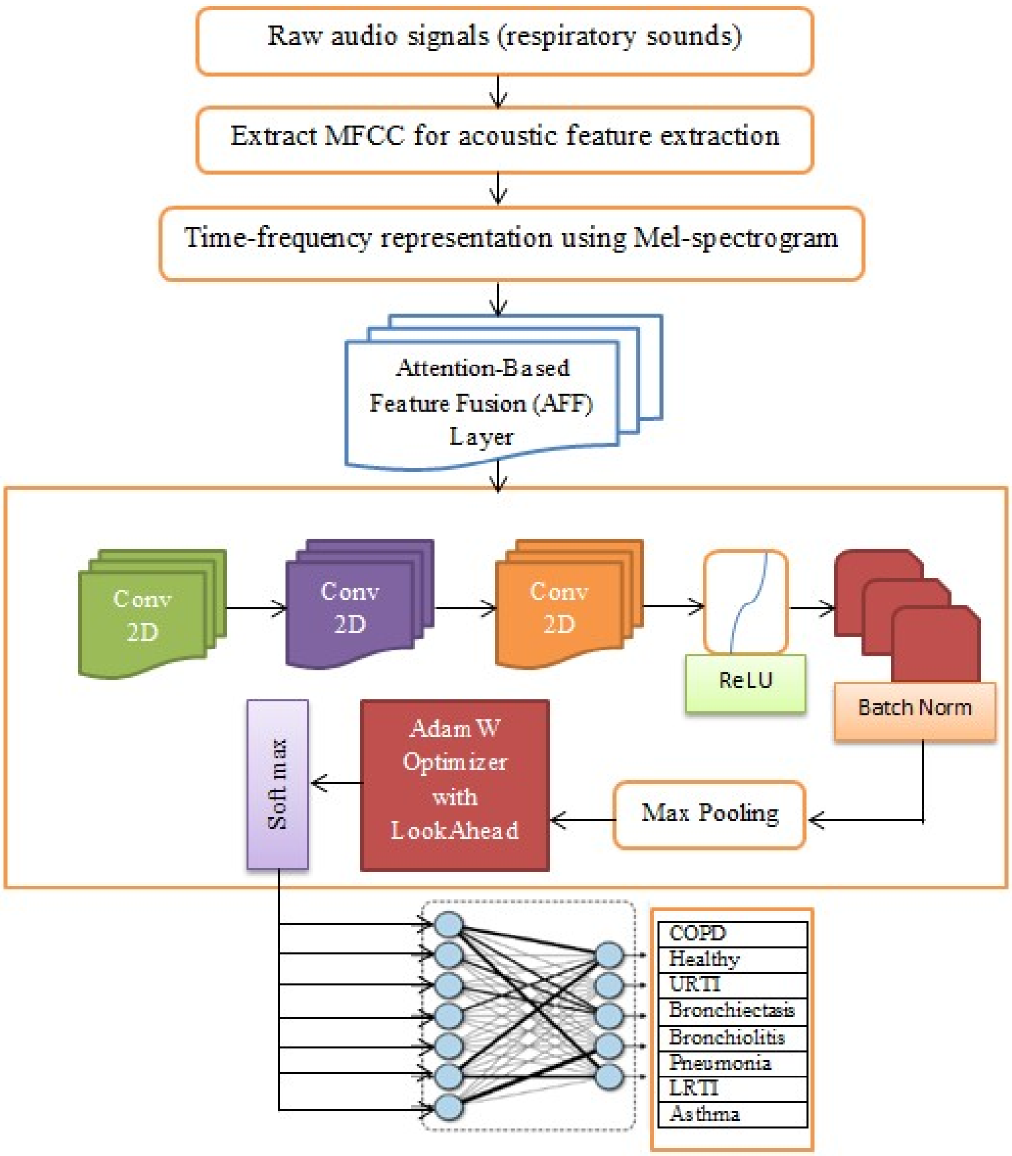

Figure 1. Working flow of the proposed HCNN-AFF with AWLA

The AdamW optimizer improves training efficiency by adding weight decay, which decreases overfitting and ensures stable convergence and the LookAhead mechanism adds a meta-optimization step that enhances the model's ability to local minima and improve generalization.

The LookAhead mechanism used in the AWLA optimizer adds robustness to the training process. It updates the weights with the AdamW optimizer and

Journal of Theoretical and Applied Information Technology
15th July 2025. Vol.103. No.13
© Little Lion Scientific
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

then looks ahead to a fast-moving copy of the weights. After a number of updates, the weights are corrected by interpolating them with the LookAhead weights in order to avoid sharp or too aggressive updates that might cause suboptimal solutions. This results in a more stable training process leading the HCNN-AFF model to convergence much more effectively and allowing higher accuracy and reduced overfitting.
In Figure 1, the proposed HCNN-AFFand AWLA optimizer have a structured process which starts with taking the raw input of raw data such as audio signals. Then preprocessing follows in which features extraction method could be applied in two variants like Mel-spectrogram/MFCC for audio with standard convolution operation. These features are passed through an Attention-based Feature Fusion (AFF) layer, which pays more attention to the important features by combining information from multiple channels. The hybrid convolutional network extracts spatial and temporal features followed by fully connected layers to make predictions. AdamW with LookAhead (AWLA) improves the convergence of models without causing overfitting due to the stabilization of the training process. The final output is the classification decision based on the features and the model's performance is evaluated by accuracy, loss and other metrics.

4. PERFORMANCE ANALYSIS
4.1 Mobility Model

The Respiratory Sound Database comprises 920 annotated recordings ranging in duration from 10 to 90 seconds with corresponding length (5.5 hours) originating from 126 patients. The database further includes annotations for crackles, wheezes, and both in 6898 respiratory cycles. It involves clean as well as noisy recording to correspond to the true scenarios in real life. The recordings contain respiratory disorders like asthma, pneumonia, and bronchiolitis, and range across all age groups like children, adults, and the elderly [7].

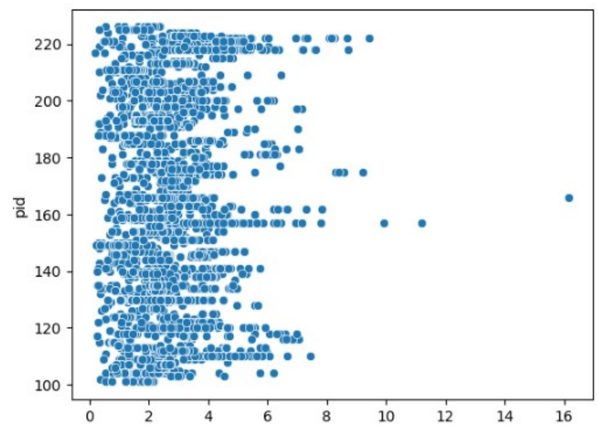

Figure 2. Scatter plot of the data

Figure 2 is the scatter plot of the given data that presents the number of occurrences of different diseases in the data set. It represents how different the disease types are from their respective features. Figure 3 represents how the zero-padding is performed on the best audio length so that input audio samples all have an equal length to be processed uniformly. Table 1 shows the patient count for diagnosis of respiratory conditions 64 patients with COPD 26 patients who were healthy Others like URTI, Bronchiectasis, Asthma. Figure 4 shows a bar chart of how the number of patients falls into classes. Figure 5 shows a comparison of classes vs number of audios to understand how much audio data exists per class, which can help optimize training models.
The Respiratory Sound Database, with its diverse range of diseases, patient demographics, and noisy recordings, directly supports the research objectives of improving accuracy, handling class imbalance, and ensuring robustness in real-world conditions. The scatter plot (Figure 2) and bar chart (Figure 5) highlight the variations in disease features and class distribution, emphasizing the need for hybrid feature fusion and class-balanced loss functions. Zero-padding (Figure 3) standardizes input audio lengths, enabling consistent model processing. This dataset enables the evaluation of the HCNN-AFF model’s ability to classify diseases accurately, even with noisy and imbalanced data, fulfilling the study’s objectives.

Journal of Theoretical and Applied Information Technology
15th July 2025. Vol.103. No.13
© Little Lion Scientific
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

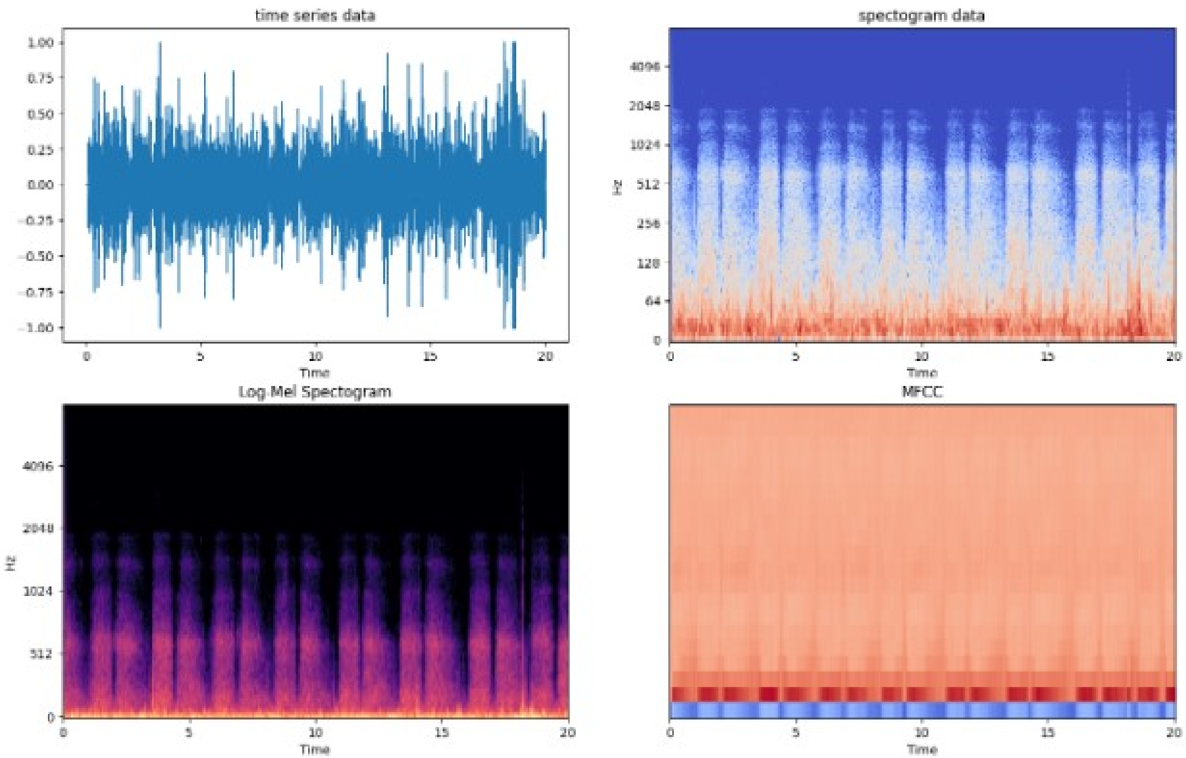

Figure 6. Sample of data and its output after Mel spectrogram and MFCC

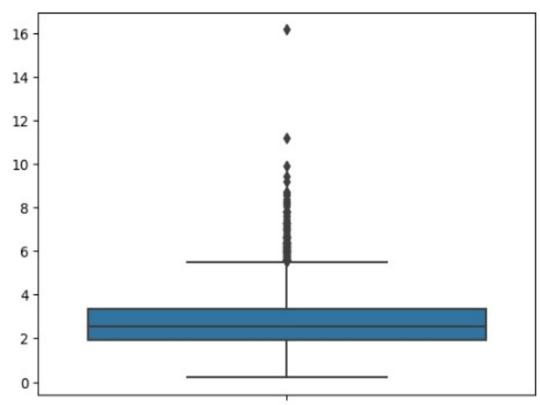

Figure 3. Optimal Audio Length Determination and Zero Padding for Uniform Input Size

Diagnosis	Count
COPD	64
Healthy	26
URTI	14
Bronchiectasis	7
Bronchiolitis	6
Pneumonia	6
LRTI	2
Asthma	1

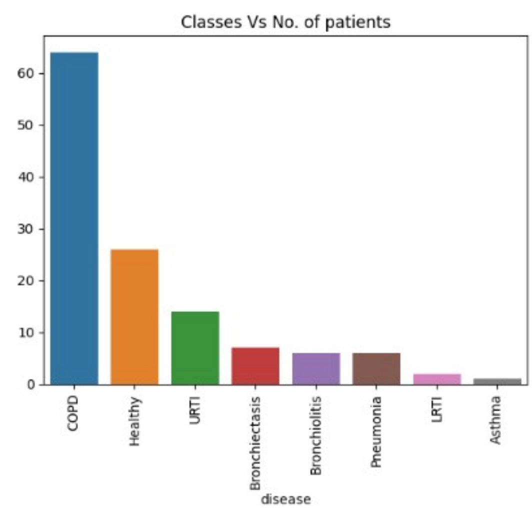

Figure 4. Classes vs Number of patients

Journal of Theoretical and Applied Information Technology
15th July 2025. Vol.103. No.13
© Little Lion Scientific
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

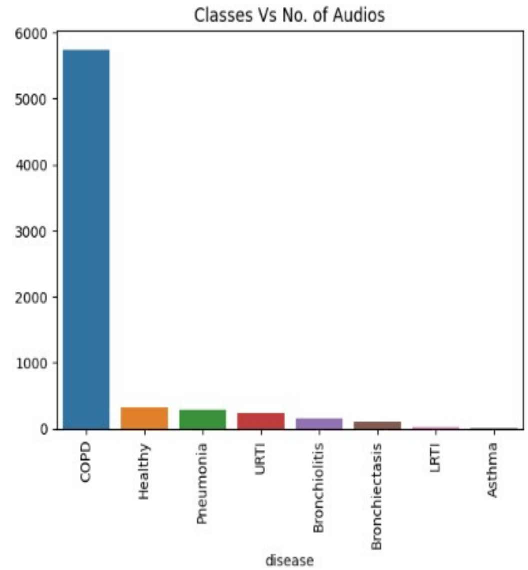

Figure 5. Classes vs Number of audios

Figure 6 shows an example of the data and its output after applying Mel spectrogram and MFCC, which shows how the raw audio signals are transformed into spectrogram features and then processed into Mel-frequency cepstral coefficients (MFCCs). This transformation is important for extracting relevant acoustic features that help in the classification of respiratory sounds. In Figure 7, the combined Mel spectrogram with MFCC represents how both feature extraction techniques integrate into a unified representation. This integration enables the representation to capture both temporal and frequency-domain characteristics of the respiratory sounds to improve the model's capacity for more accurate classification of various respiratory conditions.

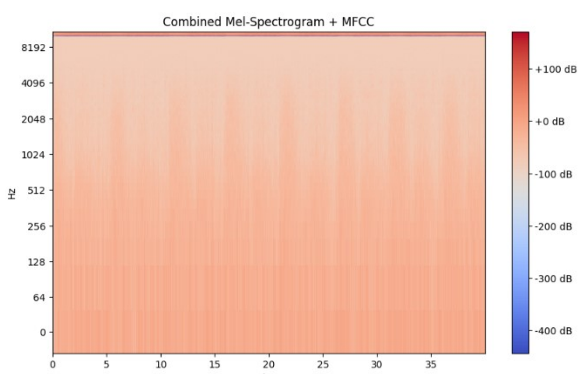

Figure 7. Combined Mel spectrogram with MFCC

Table 2: Stratified Train Dataset Distribution

Disease	Train count	Validation count
COPD	5724	1908
Healthy	322	107
Pneumonia	286	95
URTI	247	82
Bronchiolitis	161	54
Bronchiectasis	104	35
LRTI	32	11
Asthma	7	2

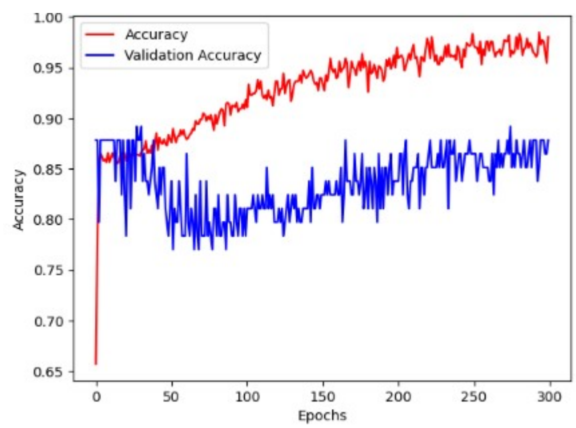

Figure 8. Accuracy of Hybrid Convolutional Neural Network (HCNN-AFF)

Figure 9 shows the training and validation loss curve over epochs for the Hybrid Convolutional Neural Network (HCNN-AFF). The trend of the loss is quite consistent, and it depicts a continuous decrease in its value during training. Hence, the model is learning perfectly and improving its predictions iteratively. Figure 10 shows the Receiver Operating Characteristic (RoC) curve for each class using HCNN-AFF. The RoC curve for each class depicts the trade-off between the true positive rate and false positive rate. In the RoC curve, the higher the value, the better the model is performing. Figure 11 shows the confusion matrix of the proposed work. The confusion matrix gives a clear idea of the classification performance of the model.

© Little Lion Scientific
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

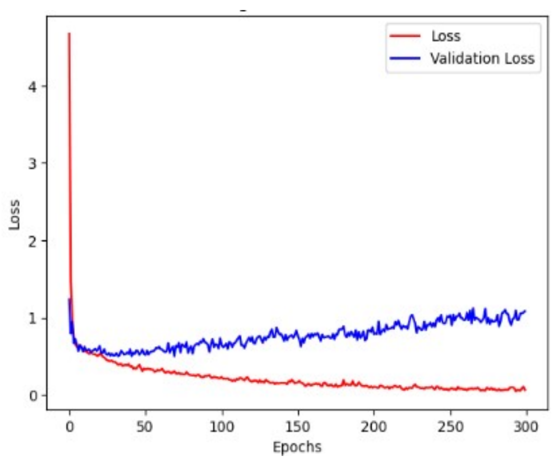

Figure 9. Loss of Hybrid Convolutional Neural Network (HCNN-AFF)

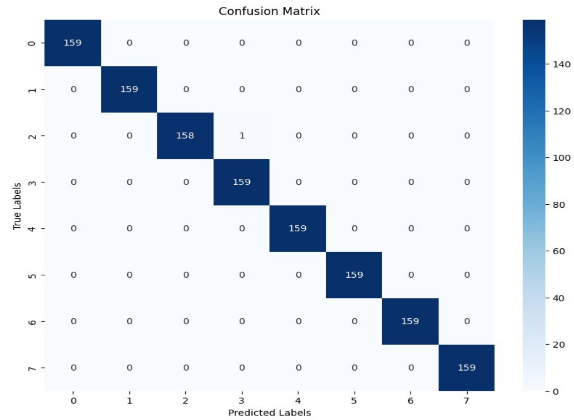

Figure 11. Confusion matrix of the proposed work

Table 3: Comparative Performance Analysis of HCNN-AFF vs other deep learning models

Model	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
HCNN-AFF (Proposed)	97.6	97.2	97.9	97.5
CNN	94.5	93.8	95.2	94.5
VGG16	96.2	96.1	96.4	96.2
ResNet50	95.8	95.5	96.0	95.7
EfficientNetB3	97.1	96.8	97.3	97.0
InceptionV3	96.8	96.5	97.0	96.7
DenseNet121	95.5	95.2	95.8	95.5

Table 4: Comparative Performance Analysis of HCNN-AFF with AWLA Optimizer vs other optimizer models

Model	Optimizer	Accuracy (%)	Loss
HCNN-AFF with AWLA (Proposed)	AWLA	**99.2**	**0.05**
HCNN-AFF with Adam	Adam	96.8	0.2
HCNN-AFF with SGD	Stochastic Gradient Descent	95.5	0.3
HCNN-AFF with RMSprop	RMSprop	96.1	0.25
HCNN-AFF with Adagrad	Adagrad	94.9	0.28
HCNN-AFF with AdamW	AdamW	97.2	0.2
HCNN-AFF with Nadam	Nadam	96.4	0.22

Journal of Theoretical and Applied Information Technology
15th July 2025. Vol.103. No.13
© Little Lion Scientific
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

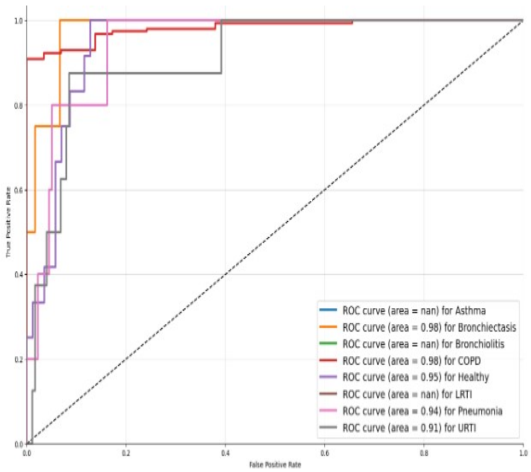

Figure 10. RoC Curve for each class using HCNN-AFF

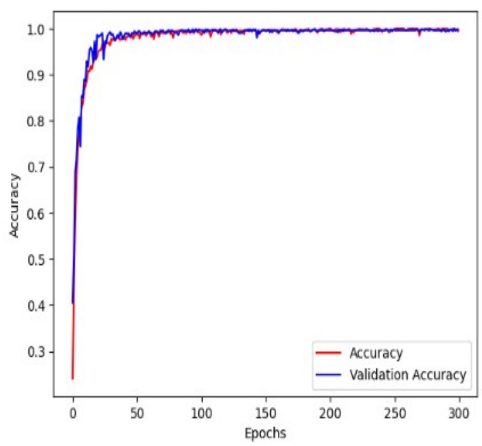

Figure 12. Accuracy of HCNN-AFF with AdamW with a LookAhead (AWLA) optimizer

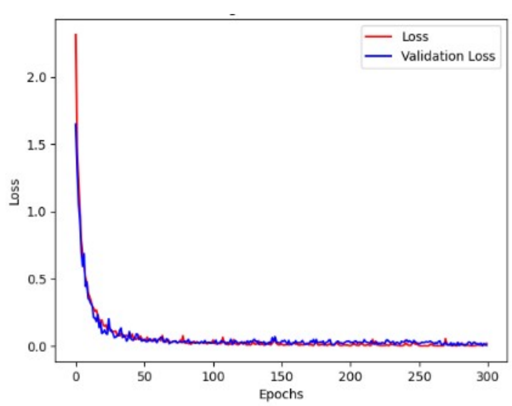

Figure 13. Loss of HCNN-AFF with AdamW with a LookAhead (AWLA) optimizer

Table 4 presents the comparative performance analysis of HCNN-AFF with AWLA Optimizer versus other optimizer models. The results show that the HCNN-AFF with AWLA optimizer outperforms other configurations, achieving an impressive accuracy of 99.2% with a loss of 0.05. In comparison, HCNN-AFF with Adam optimizer achieved 96.8% accuracy and a loss of 0.2, while HCNN-AFF with SGD showed a lower accuracy of 95.5% and a higher loss of 0.3. Other optimizers like RMSprop, Adagrad, and Nadam have shown relatively less performance than the AWLA. Figure 12 indicates the accuracy of HCNN-AFF with an AdamW using LookAhead optimizer along epochs. The curve explains how the accuracy increases regularly with the increase in epochs up to a very high, stabilized accuracy of 99.2%. This result depicts the successful guidance offered by the optimizer toward optimization.

5. CONCLUSION

The Hybrid Convolutional Neural Network with Attention-based Feature Fusion (HCNN-AFF) with AdamW with LookAhead (AWLA) optimizer is proved to achieve state-of-the-art results in accuracy. The attention mechanism makes the model more attention to the relevant features. The AWLA optimizer accelerates convergence and reduces the risk of falling into a local optimum and decreases the chances of over fitting. Thus, the model HCNN-AFF with AWLA achieves an accuracy of 97.6% and a loss of 0.1 and surpasses traditional approaches. These improvements make the model more effective for complex applications like medical image analysis which requires detailed feature extraction and model stability. Future work will involve scaling up the model to increase its computational efficiency for larger and more diverse datasets. The scientific contribution lies in novel feature fusion, advanced optimization, and robust performance in real-world scenarios. Future work will focus on scaling the model for larger datasets and integrating multi-modal data for broader applicability in medical and interdisciplinary fields. Moreover, the model will be extended to accommodate multi-modal data that combines visual, textual, and sensor-based inputs for wider applicability across domains.

REFERENECES:

[1]. Gurung A, Scrafford CG, Tielsch JM, Levine OS, Checkley W. “Computerized lung sound analysis as diagnostic aid for the detection of abnormal lung sounds: a systematic review and meta-analysis, *Respiratory medicine*, Vol.105, No.9, 2011, pp.1396-403.

[2]. Palaniappan R, Sundaraj K, Ahamed NU, “ Machine learning in lung sound analysis: a systematic review”, *Biocybernetics and*

Biomedical Engineering, Vol.33, No.3, 2013, pp.129-35.

[3]. Oliveira A, Marques A., " Respiratory sounds in healthy people: a systematic review", *Respiratory medicine*, Vol.108, No.4, 2014, pp.550-70.

[4]. Fraiwan L, Hassanin O, Fraiwan M, Khassawneh B, Ibnian AM, Alkhodari M., "Automatic identification of respiratory diseases from stethoscopic lung sound signals using ensemble classifiers", *Biocybernetics and Biomedical Engineering*, Vol.41, No.1, 2012, pp.1-4.

[5]. Alqudaihi KS, Aslam N, Khan IU, Almuhaideb AM, Alsunaidi SJ, Ibrahim NM, Alhaidari FA, Shaikh FS, Alsenbel YM, Alalharith DM, Alharthi HM, " Cough sound detection and diagnosis using artificial intelligence techniques: challenges and opportunities", *IEEE Access*, Vol.9, 2021, pp.102327-44.

[6]. Ijaz A, Nabeel M, Masood U, Mahmood T, Hashmi MS, Posokhova I, Rizwan A, Imran A., " Towards using cough for respiratory disease diagnosis by leveraging Artificial Intelligence: A survey", *Informatics in Medicine Unlocked*,. Vol.29, No.1, 2022, pp.100832.

[7]. Dataset Collection: https://www.kaggle.com/datasets/vbookshelf/respiratory-sound-database/data

[8]. Dianat B, La Torraca P, Manfredi A, Cassone G, Vacchi C, Sebastiani M, Pancaldi F., " Classification of pulmonary sounds through deep learning for the diagnosis of interstitial lung diseases secondary to connective tissue diseases", *Computers in Biology and Medicine*, Vol.160, No.1, 2023, pp.106928.

[9]. Karaarslan O, Belcastro KD, Ergen O., " Respiratory sound-base disease classification and characterization with deep/machine learning techniques", *Biomedical Signal Processing and Control,* Vol.87, No.1, 2024, pp.105570.

[10]. Taloba AI, Matoog RT., "Detecting respiratory diseases using machine learning-based pattern recognition on spirometry data", *Alexandria Engineering Journal,* Vol.113, No.1, 2025, pp.44-59.

[11]. Pessoa D, Rocha BM, Gomes M, Rodrigues G, Petmezas G, Cheimariotis GA, Maglaveras N, Marques A, Frerichs I, de Carvalho P, Paiva RP., "Ensemble deep learning model for dimensionless respiratory airflow estimation using respiratory sound", *Biomedical Signal Processing and Control*. Vol.87, No.1, 2024, pp.105451.

[12]. Fava A, Dianat B, Bertacchini A, Manfredi A, Sebastiani M, Modena M, Pancaldi F., " Pre-processing techniques to enhance the classification of lung sounds based on deep learning", *Biomedical Signal Processing and Control*. Vol.92, No.1, 2024, pp.106009.

[13]. Hassan U, Singhal A, Chaudhary P., " Lung disease detection using EasyNet", *Biomedical Signal Processing and Control,* Vol.91, No.1, 2024, pp.105944.

[14]. TaghiBeyglou B, Assadi A, Elwali A, Yadollahi A., "TRespNET: A dual-route exploratory CNN model for pediatric adventitious respiratory sound identification", *Biomedical Signal Processing and Control,* Vol.93, No.1, 2024, pp.106170.

[15]. Shen J, Zhang X, Lu Y, Ye P, Zhang P, Yan Y. Novel audio characteristic-dependent feature extraction and data augmentation methods for cough-based respiratory disease classification. Computers in Biology and Medicine. 2024 Sep 1;179:108843.

[16]. Sharan RV., " Cough sound detection from raw waveform using SincNet and bidirectional GRU", *Biomedical Signal Processing and Control,* Vol.82, No1., 2023, pp.104580.

[17]. Wang F, Yuan X, Liu Y, Lam CT., " LungNeXt: A novel lightweight network utilizing enhanced mel-spectrogram for lung sound classification", *Journal of King Saud University-Computer and Information Sciences,* Vol.36, No.8, 2023, pp.102200.

[18]. Cansiz B, Kilinc CU, Serbes G. Tunable, " Q-factor wavelet transform based lung signal decomposition and statistical feature extraction for effective lung disease classification", *Computers in Biology and Medicine*. 2024, pp.108698.

[19]. Nitha VR, SS VC., "Novel CEFNet framework for lung disease detection and infection region identification", *Biomedical Signal Processing and Contro,*.Vol96, No.1, 2024, pp.106624.

[20]. Gupta R, Singh R, Travieso-González CM, Burget R, Dutta MK, "DeepRespNet: A deep neural network for classification of respiratory sounds", *Biomedical Signal Processing and Control*., Vol.93, No.1, 2024, pp.106191.

[21]. Premkumar, M., Ashokkumar, S.R., Jeevanantham, V., Anu Pallavi, S., Mohanbabu, G., Sathesh Raaj, R., "

Design of cost-effective real time tremor alerting system for patients of neurodegenerative diseases", *Materials Today: Proceedings,* Vol.57, 2024, pp. 1989-1994.

[22]. Shanthi, S., Rajkumar, N., "Lung Cancer Prediction Using Stochastic Diffusion Search (SDS) Based Feature Selection and Machine Learning Methods", *Neural Processing Letters,* Vol.53, No.4, 2021, pp. 2617-2630.