# Journal of Artificial Intelligence & Cloud Computing



### **Research Article**

## Deep Learning-Based Pneumonia Classification Based on Respiratory Sounds

#### Ahmed Ali Dawud<sup>1\*</sup> and Getamesay Haile<sup>2</sup>

<sup>1</sup>School of Biomedical Engineering, Jimma Institute of Technology, Jimma University, Jimma, Ethiopia

<sup>2</sup>Faculty of computing, Jimma Institute of Technology, Jimma University, Jimma, Ethiopia

#### ABSTRACT

A crucial aspect of pneumonic pathology is attributed and analyzed by respiratory sound, which also provides information on the patient's lung symptoms. Almost all health institutions worldwide acquire pneumonia as a common and therapeutically challenging diagnosis that can lead to a severe body response to an infection, critical illness, and respiratory failure. It is well acknowledged that clinical diagnosis and prognosis are insufficient for the proper assessment of the severity of the disease due to the complexity of its pathogenesis. Auscultation and chest X-ray (CXR) images are currently the primary methods for identifying and diagnosing pneumonia, but it can be difficult for a trained radiologist to identify pneumonia from a CXR image or for a pathologist to do so using a stethoscope because accurate interpretation of respiratory sounds requires significant clinical expertise.

We attempted to create an automated classification of respiratory sounds to detect pneumonia as well as a more complicated classification to distinguish between different forms of pathology by using a deep learning convolutional neural network to classify the six classes (healthy, COPD, URTI, bronchiectasis, pneumonia, and bronchiolitis) of sounds recorded in a clinical setting. In the proposed method for pneumonia detection and pathological respiratory sound classification, a pertained picture feature extractor of series, respiratory sound, and CNN classifier are all used. The CNN model had an average precision of 98.8%, with a split among the classes of 97.80%, 87.60%, 95.0%, 85%, 100%, and 100% for COPD, healthy, URTI, bronchiectasis, pneumonia, and bronchiolitis, respectively. It detected abnormal sounds with an accuracy of 94.2% and pneumonia with an accuracy of 100%. The performance results obtained suggest that CNN is a viable tool for detecting specific characteristics in respiratory data and is capable of accurately classifying pneumonia.

#### \*Corresponding author

Ahmed Ali Dawud, School of Biomedical Engineering, Jimma Institute of Technology, Jimma University, Ethiopia.

Received: May 19, 2023; Accepted: May 25, 2023; Published: May 31, 2023

**Keywords:** Convolutional Neural Network, Deep Learning, Lung Sounds, Pneumonia Detection, Pulmonary Diseases.

#### Introduction

When air flows during the process of breathing, lung sound is created, which refers to all respiratory sounds heard or recognized across the chest wall or within the chest [1]. Lung auscultation is used in clinical practice to identify pulmonary disorders, while respiratory sounds are the basic element of pneumonic pathologies, such as COVID-19, and pneumonia and offer symptomatic data about a patient's lung. Lung auscultation is the process of listening to a patient's lung sounds using a stethoscope to rule out serious illnesses and identify various pulmonary problems. As a result, for centuries, a stethoscope has been essential medical equipment for physicians to use in diagnosing lung problems. However, recognizing the refined distinctions among different lung sounds requires sufficient training and clinical experience. As COVID-19 sweeps the globe, lung auscultation remains vital for monitoring confirmed cases [2].

Remote automated auscultation methods may be essential in reducing infection risks among medical staff. As a result, artificial intelligence may be used to support clinicians in doing remote and accurate auscultation, which is becoming increasingly important. Recent research has established a number of lung sound categories. Lung sounds may be divided into two types: normal and abnormal [3]. Normal noise can be heard from the beginning of the inhalation phase to the beginning of the exhalation phase. The spectral properties of these normal sounds demonstrate that they have peaked with typical frequencies below 100 Hz, and the sound energy drops abruptly between 100 and 200 Hz [4]. Pathological sounds are sounds superimposed on normal a sound which is generated by respiratory disorder.

Furthermore, respiratory sounds are divided into two types: continuous and discontinuous. Wheeze and crackle were terms used to describe continuous and discontinuous noises, respectively. The word continuous denotes that the noises last for more than 250 milliseconds [5]. Narrowing of the airway diameter causes continuous noise [6]. The pitches of these continuous sounds are determined by two things. The bulk and suppleness of the airway walls are one factor, while the velocity of airflow is another [7]. The pitches of the continuous sounds match the power spectrum's prominent frequencies. Continuous adventitious sounds can clinically signify obstructive airway diseases such as asthma and chronic obstructive pulmonary disease (COPD) [8].

Discontinuous sounds (crackles) are nonmusical, brief, explosive sounds that typically last less than 20 milliseconds [9]. The sudden opening or closure of the airways, which are abnormally closed due to the lung's high elastic recoil pressure, causes those discontinuous noises. Depending on the airway diameter, the frequency of discontinuous noises ranges from 100 to 2000 Hz [10]. Furthermore, in individuals with pneumonia and interstitial lung illnesses, discontinuous sounds can be linked to disease progression and severity [9, 10].

This paper aims to propose signal analysis in conjunction with machine learning approaches that can provide better results in pneumonia detection for better health care services. Feature engineering of machine learning algorithms can be a useful way to differentiate between the same adventitious sounds of different pulmonic diseases. Thus, there is a need for an efficient, economic, convenient, and noninvasive diagnostic methodology for the identification of pulmonic disease.

#### **Related Work**

In recent years, many researchers have been working on the automatic classification of respiratory sounds by using different deep learning models. Using a deep learning model enhances the accuracy of respiratory sound classification. In this section, we review previous attempts made on respiratory sound classification models using deep learning which helps us to understand the basic features of pneumonia and other respiratory sounds.

In the author suggests using deep learning to classify crackles, wheezes, and rhonchi as respiratory sounds in the clinical setting [11]. A total of 1918 respiratory sounds captured in a clinical context were divided into categories by the author using deep learning convolutional neural networks (CNNs). They combined a pretrained image feature extractor of series, respiratory sound, and a CNN classifier to create their prediction model for respiratory sound classification. The authors conclude that their study provides good classification accuracy for respiratory sounds, but that pneumonia is not taken into account.

The author proposed combining CNN and LSTM to identify pulmonary diseases from lung sounds [12]. They made use of 103 patient datasets from locally recorded lung sounds using stethoscope from King Abdullah University Hospital, Jordan University of Science and Technology. They used bidirectional short-term memory units and convolutional neural networks as the two stages of their network design. They used Cohen's kappa, accuracy, sensitivity, specificity, precision, and F1-score as performance assessment measures, and their model was assessed using a k-fold cross-validation procedure of 10-fold. Their CNN + BDLSTM model classified patients according to lung illness with an average accuracy of 99.62% and a precision of 98.85%.

The authors proposed classifying lung sounds using CNN models that are depthwise separable and combine STFT and MFCC characteristics [13]. For the reduced DS-CNN (depthwise separable convolution neural network) model, the authors selected a total of three features: the short-time Fourier transformed (STFT), the Mel-frequency cepstral coefficient (MFCC) and the fused features of the two. The accuracy of DS-CNN models trained on either the two types of feature, according to the ir experiment, was 82.27% and 73.02%, respectively, and when both features were employed, it was 85.74%. Proposed using a convolutional neural network to classify respiration sounds [14]. To accomplish this, they created a visual representation of each audio sample employing the same techniques used to classify images with high precision, and they extracted resources with Mel Frequency Cepstral Coefficients (MFCCs) for respiratory disease classification.

From the related works, we noted that applying a deep learning model will enhance the performance of respiratory sound classification. Therefore, this research proposes to develop deep learning-based respiratory sound classification that can enhance the accuracy of existing respiratory sound classification models and pneumonia detection for better diagnosis of pneumonia.

#### **Materials and Methods**

This study's proposed approach for detecting pneumonia based on respiratory sound classification into predetermined classes consists of four steps. Figure 1 demonstrates signal acquisition, preprocessing, model development, and model evaluation.



Figure 1: Proposed Methodologies for the Detection of Pneumonia

#### Dataset

The 920 recordings that were obtained from 126 subjects made up the International Conference on Biomedical and Health Informatics (ICBHI) dataset, which was produced by two research teams (Greece and Portugal) [15]. A total of 5.5 hours of sound and 6898 breathing cycles were captured. To create a difficult dataset, high noise levels were used to mimic real-world conditions. In this study, respiratory sounds were taken from healthy subjects, people with chronic obstructive pulmonary disease (COPD), people with upper respiratory tract infections (URTI), people with pneumonia, and people with bronchiolitis.

#### Preprocessing

Lung sound recordings are vulnerable to acoustic disturbance from ambient noise, background conversation, electronic interference, or any stethoscope displacement, similar to any other biological signal that is electronically recorded [16]. Therefore, it is crucial to ensure the signals are treated and smoothed before any feature extraction steps are taken. The current study followed a number of preprocessing processes, including [12]:

- •1D wavelet smoothing
- Displacement artifact removal
- z-score normalization



**Figure 2:** Examples of lung sound signals from healthy individuals and patients with five types of respiratory diseases: a healthy, b pneumonia, c bronchiectasis, d URTI, e COPD and f bronchiolitis

#### **1D Wavelet Smoothing**

The wavelet transform (WT) has been extensively employed for nonstationary signal processing. Instead of complex sinusoids as in the Fourier transform (FT), the signal is divided into a collection of wavelets. The fundamental principle of WT is the translation and dilation of a signal into several functions using a mother wavelet (t)[17]. The mother wavelet is defined mathematically as:

Where t corresponds to the time instance, a is the dilation function, and b is the translation function. WT might be discrete or continuous. The continuous operation of the dilation and translation functions on the signal in continuous wavelet transform (CWT) raises the computational complexity. While using discretely sampled wavelets, the discrete wavelet transform (DWT) preserves both the frequency and the position information in time [18]. As a result, the DWT is more effective than the CWT for signal analysis. In this study, a level 4 soft thresholding with a db5 mother wavelet was chosen for the DWT to implement. The signal processing toolbox for MATLAB R2019a was used to carry out smoothing.

#### **Displacement Artifact Removal**

Any movement of the electronic stethoscope produces a lowfrequency wave-shaped signal in addition to the usable signal. Therefore, it is crucial to ensure that none of these forms inside the structure are being contaminated by the signals [19]. A local polynomial regression smoother (LOESS) function was used because it works quickly and effectively to eliminate these effects. In LOESS, the signal is fitted with a weighted leastsquares function, with larger weights for points that are closer to the fitted line and lower weights for farther away locations. It is specified as follows:

$$W(x) = (1 - |d|^3)^3 \dots (2)$$

Where d is the distance of each point to the fitted curve scaled to be between 0 and 1. A complete mathematical explanation of these two methods can be found in [20].

#### Z score Normalization

Each signal must be z-score normalized after the aforementioned signal preprocessing steps. The signal in the time domain will no longer have significant dynamic ranges between its corresponding values after z-score normalization. In other words, neither the larger nor the smaller trends in the signal predominate. Making a signal clear with no trend across time gives a better performance in deep learning algorithms [21]. As a result, the signal has the following mean value ( $\mu$ ) of 0 and standard deviation ( $\alpha$ ) of 1:

$$x = \frac{x - \mu}{\alpha}....(3)$$

#### **Training and Classification**

A combination of 1D convolutional neural networks has been used as the training and classification model (CNNs). With CNN,

characteristics about the signal's entire spatial dimensionality may be extracted. In contrast, LSTM collects the characteristics following changes in the temporal domain. Better results are often obtained when building a model to predict signals based on their spatial and temporal features by integrating the two networks [22]. Additionally, the model's effectiveness was assessed for both the merged network and while running independently as CNN.

#### **Performance Evaluation**

Several indicators were incorporated in this work to investigate the classification confusion matrix and evaluate the performance of the proposed model. After tenfold cross-validation of the training/ classification method, the confusion matrix was constructed sequentially after each fold, and all evaluation metrics were derived from the total confusion matrix. These metrics are given by:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \dots (4)$$

$$F1 = \frac{2TP}{2TP + FP + FN}.$$
(7)

Where true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN).

#### **Results and Discussion**

Figure 3 only shows the preprocessing results for the 2.5 s segment for visual purposes. Figure 4 illustrates the outcomes of the metrics

used to assess how well the neural network model performed. As illustrated in part, the proposed method successfully eliminates any noise artifacts caused by nearby noise sources and extremely low frequencies while keeping any displacement or movement artifacts (b). After the signal's z score was normalized prior to the detection of pneumonia, the final signal utilized for training and classification is displayed in section (d).

#### Performance of the Proposed Model

MATLAB software R2019 a was used to implement the proposed method. The experiments were run on an Intel Core i7 computer with 32 GB of RAM. A 4 GB GPU and 8 GB display memory processing unit were used for the training procedure (VRAM). The entire training/classification scheme took approximately 15 minutes to complete because each fold was taught for 2 minutes. Under the aforementioned machine parameters for the detection of pneumonia and other respiratory illness classes, the prediction of the per-patient class took less than a second. Figure 4 shows a tenfold confusion matrix of the original vs predicted disorders for the CNN to evaluate its performance.

The CNN model has the highest average precision in the classification process, with a 98.85% split among the classes of 97.80%, 57.60%, 50.0%, 75%, 100%, and 100% for COPD, healthy, URTI, bronchiectasis, pneumonia, and bronchiolitis, respectively.

In comparison to previous models, the proposed approach exhibited the best levels of pneumonia prediction accuracy. The fact that each study used a different number of recordings to classify a different number of groups based on respiratory sounds is important to note. However, because pneumonia has a significant impact on many aspects of respiratory illness, our study primarily focused on distinguishing it from other respiratory problems.



**Figure 3:** Preprocessing of a sample lung sound signal (2.5 s segment): a Original Signal, b MODWT Wavelet Smoothing, c rLOESS Displacement Removal, d z Score Normalization

	precision	recall	f1-score	support
CODD	0.00	0.07	0.00	165
COPD	0.99	0.97	0.98	105
Healthy	0.80	0.57	0.67	7
URTI	0.67	0.50	0.57	4
Bronchiectasis	1.00	0.75	0.86	4
Pneumoina	0.33	1.00	0.50	3
Bronchiolitis	0.00	0.00	0.00	1
accuracy			0.93	184
macro avg	0.63	0.63	0.60	184
weighted avg	0.96	0.93	0.94	184

**Figure 4:** Results of the metrics used to evaluate the performance of neural the network shown in Figure 5.



**Figure 5:** The Confusion Matrix and Per-Class Precision Percentage of Pneumonia Detection in Respiratory Disease Recognition Using CNN

Figure 5 demonstrates the confusion matrices of the experiment test. With this, it is possible to show that the implementation had the best performance for the detection of pneumonia based on respiratory sound classification, as well as to analyze the training history for tests in Figures 6 and 7. The curves of training accuracy and validation accuracy, training loss and validation loss for the experiment are presented in Figures 6 and 7, while the confusion matrices that give the number of detected signals in the test are shown in Figure 5.



Figure 6: The Progress of Training Accuracy and Validation Accuracy for the First Experiment



**Figure 7:** The Progress of Training Loss and Validation Loss for the First Experiment.

Figures 6 and 7 demonstrate that the experiment's training accuracy performs better than its validation accuracy. This indicates that the model can be trained more effectively using the distribution of classes in the training. The method proposed in this article gives results above 94% in the classification of respiratory sounds using six classes and 100% for the detection of pneumonia among respiratory diseases available in the database used.

Authors	Classes	Extracted features	Models	Performance
Aykanat et al. (2017) [23]	3: Normal pneumonia, rhonchus	MFCC/spectrograms	SVM/CNN	Accuracy: 80% Sensitivity:79.00% Specificity: N/A
Shi et al. (2019) [24]	3: Normal, asthma, pneumonia	Spectrograms	VGG-BDGRU	Accuracy: 87.41% Sensitivity: N/A Specificity: N/A
García-Ordás et al. (2020) [25]	6: Normal, asthma, pneumonia, BRON, COPD, respiratory tract infection	Spectrograms	CNN	Accuracy: N/A Sensitivity: 98.81% Specificity: 98.61%
Fraiwan et al. (2021) [12]	6: Normal, asthma, pneumonia, BRON, COPD. Heart failure	Spatial and temporal	(CNN + BDLSTM) CNN + BDLSTM	Accuracy: 99.62% Sensitivity: 98.43% Specificity: 99.69%
This study	Healthy, COPD, URTI, Bronchiectasis, Pneumonia and Bronchiolitis	Spatial	CNN	Accuracy: 99.9% Sensitivity: 98.5% Specificity: 99.4%

Table 1: Summary table of recent studies found in the literature for the use of deep learning approaches in lung sound classification

According to Table 1's results, it is obvious that the given method had higher classification accuracy than the studies that were examined, which might result in a more accurate diagnosis of pneumonia. The created algorithm can therefore identify pneumonia and classify the five types of respiratory sound signals completely automatically and with sufficient reliability.

#### Conclusion

Based on the research methodology, a convolutional neural network with a deep learning framework was created that initially integrates MATLAB software R2019a and processing algorithms such as 1D wavelet smoothing, displacement artifact removal, and z-score normalization for improved respiratory sound classification and pneumonia detection. This algorithm has an overall average accuracy of 94.2% for classifying respiratory sounds into various respiratory disorders, with 100% accuracy for predicting pneumonia. This study significantly improved the results of using a deep learning-trained model in clinical settings to help doctors make decisions on the diagnosis of pneumonia.

The performance results obtained suggest that CNN is a viable tool for detecting specific characteristics in respiratory data and is capable of accurately classifying pneumonia inside and outside of laboratory environments, which inspires and enables further research in the analysis of respiratory sounds. However, the classification for other respiratory diseases may be further improved by modifying both the preprocessing techniques and the training structure without affecting the performance of pneumonia detection. Future research will thus focus on extending the dataset's size to include more subjects and a wider spectrum of pathologies, such as COVID-19, to increase the proposed model's reliability.

#### References

- 1. A R A Sovijärvi, F Dalmasso, J Vanderschoot, L P Malmberg, G Righini, et al. (2000) Definition of terms for applications of respiratory sounds. Eur Respir Rev 10: 597-610.
- C Jiang, J Zhao, B Huang, J Zhu, J Yu (2021) A basic investigation into the optimization of cylindrical tubes used as acoustic stethoscopes for auscultation in COVID-19 diagnosis. J. Acoust Soc Am 66-69.
- H Pasterkamp, P L P Brand, M Everard, L Garcia-Marcos, H Melbye, et al. (2016) Towards the standardisation of lung sound nomenclature. Eur Respir J 47: 724-732.
- N Gavriely, M Nissan, A H E Rubin, D W Cugell (1995) Spectral characteristics of chest wall breath sounds in normal subjects. Thorax 50: 1292-1300.
- Chizu Habukawa, Naoto Ohgami, Naoki Matsumoto, Kenji Hashino, Kei Asai, et al. A wheeze recognition algorithm for practical implementation in children. PLoS One 15 1-12.
- C Jácome, J C Aviles-Solis, Å M Uhre, H Pasterkamp, H Melbye (2018) Adventitious and normal lung sounds in the general population: Comparison of standardized and spontaneous breathing. Respir Care 63: 1379-1387.
- Sovijarvi ARA, Malmberg LP, Charbonneau G, Vanderschoot J, Dalmasso F, et al. (2000) Characteristics of breath sounds and adventitious respiratory sounds. Eur Respir Rev 10: 591-596.
- 8. P Piirila, ARA Sovijarvi (1995) Crackles: Recording, analysis and clinical significance. European Respiratory Journal 8: 2139-2148.
- 9. P Piirila (1992) Changes in crackle characteristics during the clinical course of pneumonia. Chest 102: 176-183.
- 10. S Z H Naqvi, M A Choudhry (2020) An automated system for classification of chronic obstructive pulmonary disease

and pneumonia patients using lung sound analysis. Sensors (Switzerland) 20: 1-23.

- 11. Yoonjoo Kim, YunKyong Hyon, Sung Soo Jung, Sunju Lee, Geon Yoo, et al. (2021) Respiratory sound classification for crackles, wheezes, and rhonchi in the clinical field using deep learning. Sci Rep 11: 1-11.
- M Fraiwan, L Fraiwan, M Alkhodari, O Hassanin (2021) Recognition of pulmonary diseases from lung sounds using convolutional neural networks and long short-term memory. J Ambient Intell Humaniz Comput 3: 4759-4771.
- S Y Jung, C H Liao, Y S Wu, S M Yuan, C T Sun (2021) Efficiently classifying lung sounds through depthwise separable cnn models with fused stft and mfcc features. Diagnostics 11: 732.
- 14. A A Saraiva, D B S Santos, A A Francisco, Jose Vigno Moura Sousa, N M Fonseca Ferreira (2020) Classification of respiratory sounds with convolutional neural network. BIOINFORMATICS 2020 - 11th International Conference on Bioinformatics Models, Methods and Algorithms, Proceedings; Part of 13th International Joint Conference on Biomedical Engineering Systems and Technologies BIOSTEC 2020 138-144.
- B M Rocha, D Filos, L Mendes, I Vogiatzis, E Perantoni, et al. (2018) A respiratory sound database for the development of automated classification. IFMBE Proceedings 2018 66: pp. 33-37.
- 16. J Huang, B D Sekar, R Guo, M Dong, X Hu (2016) Characterization of noise contaminations in realistic heart sound acquisition. Lecture Notes in Electrical Engineering 375: 491-501.
- 17. J P Martínez, R Almeida, S Olmos, A P Rocha, P Laguna (2004) A Wavelet-Based ECG Delineator Evaluation on Standard Databases. IEEE Trans Biomed Eng 51: 570-581.
- 18. A Mohsin, O Faust (2019) AUTOMATED CHARACTERIZATION of CARDIOVASCULAR DISEASES USING WAVELET TRANSFORM FEATURES EXTRACTED from ECG SIGNALS. J Mech Med Biol 19: 1940009.
- J Zheng, J Zhang, S Danioko, H Yao, H Guo, et al. (2020) A 12-lead electrocardiogram database for arrhythmia research covering more than 10,000 patients. Sci Data 7: 48.
- 20. David A Swayne, C Obimbo (2015) Automating the Smoothing of Time Series Data. J Environ Anal Toxicol 5.
- 21. Y Roy, H Banville, I Albuquerque, A Gramfort, T H Falk, et al. (2019) Deep learning-based electroencephalography analysis: A systematic review. Journal of Neural Engineering 16: 051001.
- 22. Z Zheng, Z Chen, F Hu, J Zhu, Q Tang, et al. (2020) An automatic diagnosis of arrhythmias using a combination of CNN and LSTM technology. Electron 9: 121.
- M Aykanat, Ö Kılıç, B Kurt, S Saryal (2017) Classification of lung sounds using convolutional neural networks. Eurasip J Image Video Process 2017.
- L Shi, K Du, C Zhang, H Ma, W Yan (2019) Lung Sound Recognition Algorithm Based on VGGish-BiGRU. IEEE Access 7: 139438-139449.
- 25. F Demir, A M Ismael, A Sengur (2020) Classification of Lung Sounds with CNN Model Using Parallel Pooling Structure. IEEE Access 8: 105376-105383.

**Copyright:** ©2023 Ahmed Ali Dawud. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.