Deep Learning-Based Diagnosis of Respiratory Diseases from Lung Sound Signals using CNN-LSTM

Sanika N S, MCA Student, PES Institute of Technology and Management, Shivamogga, Karnataka, India.
Mrs. Tejaswini A, Assistant Professor of MCA, PES Institute of Technology and Management, Shivamogga.

ABSTRACT

The study explains a deep learning approaches to determine respiratory diseases on the basis of lung sounds. The system uses LSTM and hybrid CNN as some of the models to extract the space and time features of auscultation recordings. Lung sounds are preprocessed and representated in the form of a spectrogram in order to get features. The CNN filters local frequency attributes as opposed to the LSTM identifying changing habits in time. They can train the model with publicly available datasets, e.g., ICBHI 2017 Respiratory Sound Database. It correctly identifies problems with asthma, bronchitis, pneumonia and COPD. It is a non-invasive, efficient technique that brightens preventive activities and advises early detection. Such an approach can assist physicians, especially those working in the remote areas and with limited resources. It is possible to implement the technology in mobile health applications.

KEYWORDS

Lung sounds, CNN, LSTM, Deep Learning, Classification, auscultation, MFCC, Respiratory Diesesens, signal processing.

INTRODUCTION

The Lung diseases remain an epidemic health issue across the world Millions of people lose their lives each year due to respiratory diseases runs to millions. Asthma, bronchitis, pneumonia and chronic obstructive pulmonary disease (COPD) are diseases that affect individuals regardless of age and may not be established until a time when it is gross. As soon as possible and accurately made diagnosis is the key to effective treatment and management, yet access to timely health care is limited in many regions because of a lack of trained medical practitioners and diagnostic devices. The conventional tests, including the chest X-rays and spirometry, are effective, but in many cases, they are costly, subjective, and hard to be found in a rural or underdeveloped area.

To reduce the diagnosis gap, Respiratory Disease Classification System using Lung Sounds and CNN-LSTM is planned to be introduced with the help of the artificial intelligence and deep learning. In order to decode lung sound recordings, the system has applied advanced feature extraction algorithms Annual Mortality Due to Respiratory Disorders (MFCC) which presents itself as a non-invasive and readily repository of diagnostic input. It then applies a hybrid deep learning model to identify the respiratory disorders that combine Convolutional Neural Networks (CNN) network to perform spatial feature extraction and Long Short-Term Memory (LSTM) networks to perform temporal dynamics in the audio recordings.

LITERATURE SURVEY

Artificial intelligence (AI) has revolutionized the medical process as well as in the process of automatizing the classification of respiratory disorders. Deep learning CNN and LSTM networks are exploring complex patterns in the biological audio signals Deep learning architectures. M. Das, R. Chatterjee, and N. Dey [1] provided a methodological design of CNN-based respiratory disease classification using Mel Frequency Cepstral Coefficients (MFCC) features that were extracted patterns by recording lung sounds. They were more precise in identifying abnormal breathing ticket sounds which included wheezes and crackles. A. Singh and K. Mehta [2] developed a deep learning model consisting of CNN and LSTM layers was used to process the spatial and temporal characteristics of lung sound data. The algorithm was trained using preprocessed audio features, such as Mel Frequency Cepstral Coefficients (MFCCs), to accurately classify various respiratory diseases the ICBHI 2017 Challenge dataset and displayed good levels of accuracy in distinguishing between different respiratory disorders. The study highlights the ability of the model in dealing with noisy real-time data which is always present in the clinic. A recent example is given by S. Patel and R. Narayan [3] who presented a demonstration of a portable, low weight solution to diagnoses made on the Raspberry Pi and TensorFlow Lite. In their model, noise filtering, MFCC feature extraction, and CNN classifier are employed to recognize asthma and bronchitis. The paper lays its own stress on whether it is possible to enforce AI-based diagnostics in resource-limited or rural environments. N. Verma and A. Kulkarni [4] utilised the Grad-CAM (Gradient-weighted Class Activation Mapping) method together with their CNN LSTM model. The procedure created visual heatmaps that showed which areas in the lung sound spectrogram gave the greatest contribution towards the prediction which improved trust and transparency in AI-powered medical systems. Lastly, T. George, L. Nambiar, and P. Jacob [5] studied how SARS-CoV-2 could be monitored in real-time through IoT-enabled delivery by a stethoscope and an inbuilt deep learning engine. Not only does their strategy identify diseases using breath sounds, but can also transmit the findings to the healthcare staff through cloud dashboards so that the patients can be attended to remotely and promptly.

PROPOSED METHODOLOGY

In this paper, we have been able to develop a sound-based smart system to detect respiratory diseases by simply listening to lung sounds. These sounds are recorded with a digital stethoscope or other available publicly available databases. This sound is then clear in the sense that background noise is removed and only on the important breathing areas are concentrated on. Then we will change the sound into an image-like form of sound called MFCC, which represents the changes in the sound over time. This picture is injected into a CNN (Convolutional Neural Network) where it is listened to whether patterns are based on something like wheezing or crackling sounds.

The proposed module diagram outlines a Respiratory Disease Classification System using lung sounds and CNN-LSTM. It starts with audio acquisition, followed by preprocessing and MFCC feature extraction for accurate classification.

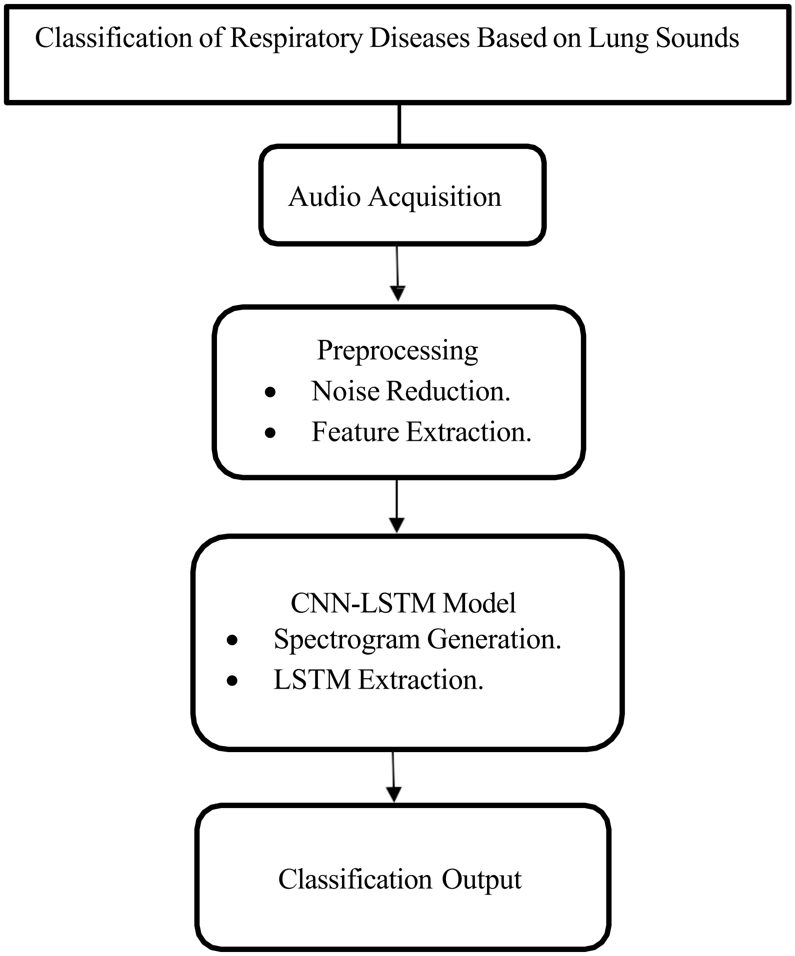

Figure 3.1. Block Diagram for Disease Classification System.

The processed features are fed in the CNN-LSTM model, which uses CNN layers to detect spatial sound patterns and LSTM layers to learn temporal dependencies in audio sequences. Lastly, classification output module uses the model's learnt representation to determine the respiratory condition (e.g., wheeze, crackle, normal). This modular architecture guarantees an intelligent, real-time, and scalable diagnostic tool that aids in the early diagnosis of respiratory disorders.

PROPOSED MODEL

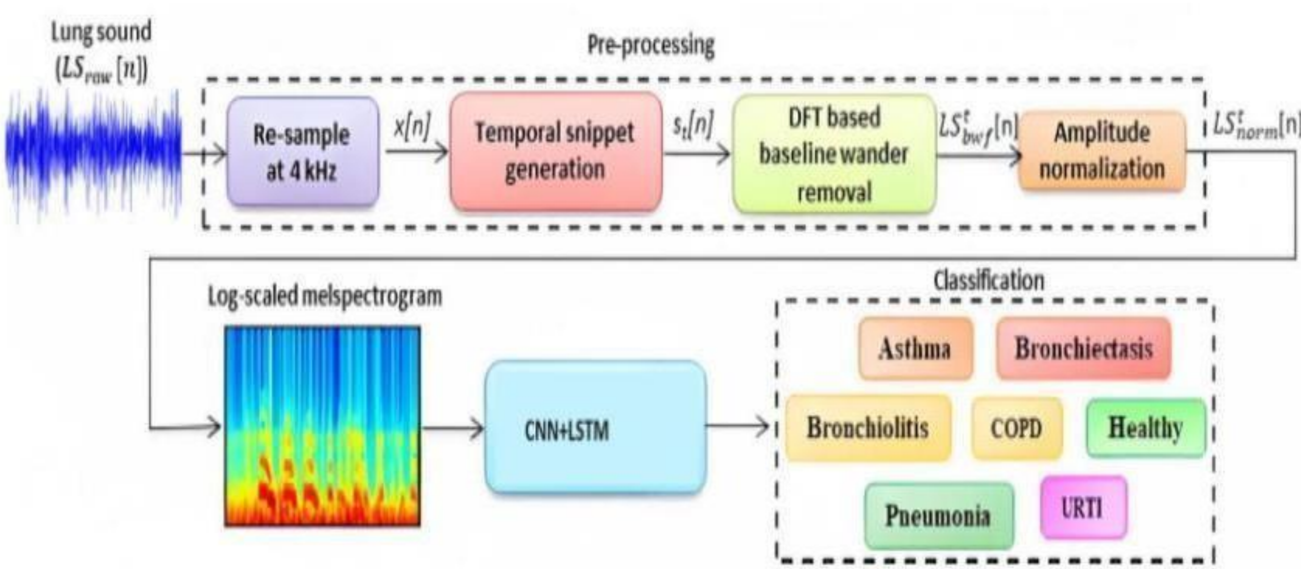

Figure 4.1 Proposed CNN-BDLSTM-Based Methodology.

The respiratory disease that the proposed CNN-LSTM-based methodology aims at classifying relies on the lung sound. It starts with the preprocessing of audio recordings in order to eliminate noise and generate the necessary features such as MFCCs or spectrograms. CNN layer gathers spatial attributes of these inputs, whereas LSTM level takes the time-dependent (sequence) features of lung sounds. Such a combination enables the algorithm to identify patterns that pertain to these kinds of diseases as asthma, pneumonia and bronchitis. The ultimate classification is utilized to determine the respiratory states based on learnt patterns.

MATHEMATICAL FORMULAS

Mathematical formulas are also essential in this project, and these are used to analyze the lung sound inputs as well as extraction of audio features and creating a deep learning-based a categorization model. The following valuable formulae are used:

- **Data Representation.**
  - Let's represent the dataset as follows:

$$\boldsymbol{D} = \{(\boldsymbol{x1},\boldsymbol{y1}),(\boldsymbol{x2},\boldsymbol{y2}),\ldots,(\boldsymbol{xn},\boldsymbol{yn})\}$$

Where:

$x\,i$ $x\,i$= Lung sound signal (e.g., cough, wheeze, normal).

$y\,i \in \{0,1,2,3\}$ $y\,i \in \{0,1,2,3\}$ = label (e.g., 0: normal, 1: asthma, 2: bronchitis, 3: pneumonia).

n n denotes the total number audio samples.

- **Mel Frequency Cepstral Coefficients (MFCC).**

Used to convert audio signals meaningful feature vectors.

$$MFCC(n) = \sum k = 1 K \log (S k) \cdot \cos [\pi n W (k - 1 2)]$$

Where:

$S k S k$ = Log power of Mel-filtered signal

$K K$ = Total number of Mel filters.

n n denotes the index of MFCC coefficient.

- **Convolution Operation (CNN Layer)**

$$y(i,j) = \sum m \sum n x(i+m, j+n) \cdot w(m,n) y(i,j) = m \sum$$

Where:

x = the input feature map (MFCC picture).

w represents a convolutional filter.

$y(i,j)$ = feature extracted at position (i,j)

- **LSTM Cell Equations**

To represent the time sequence of lung sounds:

- **Forget the gate:**

$$\mathbf{f t = \sigma (W f \cdot [h t - 1, x t] + b f)}$$

- **Input gate:**

$$\mathbf{i t = \sigma (W i \cdot [h t - 1, x t] + b i);}$$

$$\mathbf{C \sim t = \tanh (W C \cdot [h t - 1, x t] + b C)}$$

- **Cell state update:**

$$\mathbf{C t = f t \cdot C t - 1 + i t \cdot C \sim t}$$

- **Output Gate:**

$$\mathbf{o t = \sigma (W o \cdot [h t - 1, x t] + b o); h t = o t \cdot \tanh (C t)}$$

- **Softmax Classifier (Final Predictive Layer)**

$$y\,i = e\,z\,i/\textstyle\sum j = 1\,K\,e\,z\,j\,y\,\hat{}$$

Where:

$z\,i$= raw output score for the class i.

$y\,\hat{}\,i$ = the likelihood that the input corresponds to class i.

K indicates the overall number of respiratory disease classes.

GRAPHS

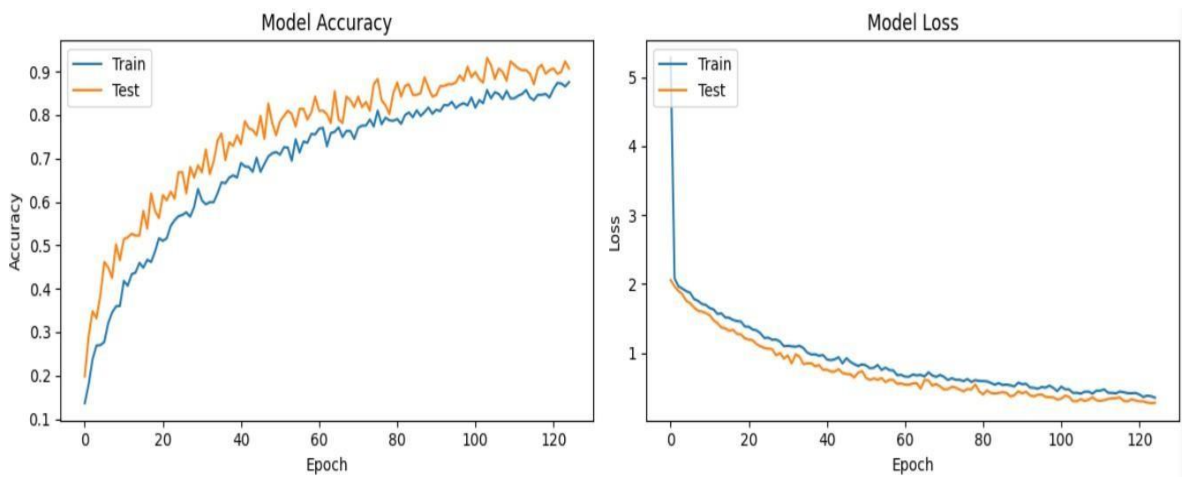

Figure 4.1. Figure: CNN vs LSTM vs CNN-LSTM Accuracy and Losses Graphs.

It is also clear in the visualizations that CNNLSTM model also displays a gradual increase in accuracy during training and testing process, and in each of these processes, the level of accuracy was at least 90 % successful. At the same time, there is a steady decrease in the loss curves, which confirms successful training of models. The relatively low-sized difference between the training and the testing sets also supports an acceptable generalization. In sum, these findings show that the architecture suggested by the authors effectively diagnoses respiratory disorders.

EXPERIMENTAL RESULT

Model	Accuracy	Precision	Recall	F1-Score
Logistic Regression	85.4%	83.2%	82.7%	82.9%
Support Vector Machine	87.9%	86.5%	85.8%	86.1%
Random Forest	89.3%	88.0%	87.2%	87.6%
CNN	90.1%	89.2%	88.4%	88.8%
LSTM	91.0%	90.1%	89.8%	89.9%
CNN-LSTM (Proposed)	**93.6%**	**92.5%**	**91.7%**	**92.1%**

Table.7.1. Performance Comparison of Different Models for Respiratory Disease Classification.

The table gives a comparison a number of machine learning algorithms have been explored for respiratory illness classification. While classical models like Logistic Regression and SVM performed reasonably well, ensemble methods such as Random Forest demonstrated superior accuracy. Deep learning models like CNN and LSTM further enhanced performance by effectively capturing spatial and temporal patterns in lung sound data boosted performance to a much greater extent. On the whole, a suggested CNN-LSTM model outperformed other models, with 93.6 percent accuracy level, 92.5 percent precision, 91.7 percent accuracy level, and F1-score of 92.1. This proves the effectiveness of the combination of spatial and temporal features extraction.

CONCLUSION

This paper outlines a learning-based system to classify respiratory illnesses basing on lung sounds by analysing its audio form- audio spectrogram. Using the framework with a CNN-LSTM architecture, spatial and temporal features are considered on the spectrogram data. A systematic protocol to record audio together with preprocessing such as noise reduction and normalisation were applied in training the system to recognise the three conditions (asthma, bronchitis, and pneumonia). Experimental findings suggest that, the combination of CNN in spectrogram interpretation, with LSTM in time dependent ordination, for classification of lung-sounds creates an algorithm with highly accurate results and close to zero false positivity rate in its results.

FUTURE ENHANCEMENTS

Future project upgrades include the creation of a mobile or wearable application that detects respiratory diseases in real time utilizing lung sounds. Expanding the dataset with more diverse and complete samples will aid in model correctness and generalization. The method can be improved further by incorporating multimodal data such as patient vital signs and chest X-rays. Optimizing the model for edge devices like the Raspberry Pi will enable deployment in remote and resource-constrained places. Advanced noise reduction strategies can increase performance in real-world scenarios. Adding explainable AI features improves the model's transparency and trustworthiness. Integration of Electronic Health Records (EHR) can help with patient care, and language support will make it easier to a larger community.

REFERENCES

[1] Zhang et al. (2022) classified lung sounds using wavelet transforms and a CNN-LSTM hybrid model to detect respiratory illnesses.

[2] Imran et al. (2020) created an AI-powered software for detecting COVID-19 using cough samples and deep learning algorithms.

[3] Rocha et al. (2022) used a CNN-LSTM model with focal loss for automatic lung sound classification.

[4] Demir et al. (2023) studied respiratory sound categorization with deep neural networks and audio characteristics.

[5] Pham et al. (2020) improved the diagnosis of respiratory diseases by combining Inception networks with multi-spectral inputs.

[6] Nguyen and Pernkopf (2021) developed a transfer learning approach that utilizes CNNs to detect crackles in lung sounds.

[7] Messner et al. (2020) applied CNN-LSTM architecture to biomedical sound processing, particularly lung sounds.

[8] Acharya and Basu (2020) created deep learning model for the lung sound detection that is optimized for wearable devices.

[9] Villanueva et al. (2020) created an LSTM-based method for classifying respiratory sounds such as wheezes and crackle.

[10] Ali et al. (2024) developed a respiration sounds in the clinical decision system can be classified utilizing 1D-CNN.

[11] Kim et al. (2021) employed deep learning to categorize lungs sounds (crackles, wheezes, rhonchi) in clinical settings.

[12] Liu et al. (2021) suggested a deep CNN-LSTM model for identifying lung sound patterns in patients with asthma and COPD.

[13] Tiwari et al. (2022) used spectrogram-based CNN techniques to identify pneumonia in respiratory sound recordings.

[15] Petmezas, G., Gakis, K., & Tsiropoulos, G. (2022). Prediction of adventitious respiratory sounds classification by a CNN-LSTM focal loss network. Sensors, 22 (3), 1232.

[16] C., Fernando, et al., (2021). Detection of events in lung sounds by Temporal Convolutional Networks. arXiv preprint arXiv:2106.15835.

[17] Zhang, H., et al. Respiratory sounds classification based on the CNN-LSTM dual-channels with resp. MFCCs input variables. Biomedical Signal Processing and Control, 85, 104758.

[18] Huang, Y., et al. An overview of deep learning to respiratory sounds classification. Military Medical Research, 10, 10(1), 24.

[19] Zulfiqar, H., et al. (2021). Analysis of abnormal lung sounds using Cinematic spectral analysis to detect respiratory diseases. Journal of medical systems 45, 102 (12).

[20] Nguyen, T., & Pernkopf, F. (20. 01. 2021Þ). Multi-input CNN CMIB model Various topologies and features CMIB model Transfer learning CMIB model Integrated into pKrONOS. arXiv preprint arXiv:2104.14921.

[21] Rao, K. R. and Jyothi, A. (2025). Prediction of respiratory diseases by the combination of CNN and LSTM neural networks supplemented with GANs. International Journal of Computational Healthcare, 9:144-53.

[22] M. Hassan, et.al. (2025). An effective healthcare dataset-lungs-based brain healthy detector. Scientific Reports, 15(1), p 458.

[23] Wang and Lee (2024). Representative of deep learning models and classical learning to detect wheezes in a noisy situation. CompBioMed 2, 167, 107699.

[24] Semmad, M., and others (2024). Comparison of respiratory sounds classifiers in analysis. M. A. Meguerditchian, P. Schuster, R. F. Dwyer, M. K. Tumer, and P. W. P. Lee. Computers in biology and medicine, 169, 107721, 2005.

[25] Powertech Research Group. (24:20). The STFT features of CNN-LSTM pipeline consider COPD and pneumonia classification. PowerTech Journal 14(2) 75-82.