Power System Technology
ISSN:1000-3673

Received: 16-09-2025 Revised: 05-10-2025 Accepted: 18-11-2025

Teleconsultation-Enabled Pulmonary Assessment Using a Multi-Task Deep Learning Framework

Preethi Kumar[1], Neha Kantikar[1], Srinivas C[1], Meenakshi N[1], Suma K V[1], Deepali B Koppad[1]

[1]*Department of Electronics and Communication, Ramaiah Institute of Technology*

Abstract: - Accurate assessment of pulmonary function is essential in healthcare for the timely diagnosis and management of respiratory conditions. Conventional studies have predominantly analyzed lung sounds and diseases as separate tasks. This study introduces a customized multi-task deep learning framework capable of performing simultaneous classification of lung sounds and lung diseases. The ICBHI 2017 Respiratory Sound Dataset was utilized for model development, wherein Mel Frequency Cepstral Coefficients (MFCCs) were extracted from lung sound recordings and concatenated with corresponding labels for training, validation, and testing. The proposed model achieved weighted average precision, recall, and F1-scores of 88%, 89%, and 86% for lung disease classification, and 61%, 65%, and 58% for lung sound classification, respectively. The overall classification accuracies were 89% for lung diseases and 65% for lung sounds. The trained model is integrated within a cloud-based health informatics framework that enables bidirectional data exchange via a dedicated mobile application. Patients can upload demographic and respiratory sound data to the cloud, where automated classification is performed, and results are made accessible to healthcare professionals. Clinicians can subsequently provide therapeutic recommendations through the same platform. The proposed system demonstrates the potential for real-time deployment in clinical and telehealth environments, promoting accessible and intelligent respiratory health monitoring. *Project Page:* https://github.com/preethik14/Pulmonary-Assessment-in-Health-Informatics-using-Deep-Learning

Keywords: *Multi-Task Learning, Teleconsulting App, Adventitious Sounds, IaaS Cloud Service, Cloud Platform.*

1. Introduction

Pulmonary assessment is a key component of healthcare that involves evaluating the functionality of the respiratory system through systematic data collection, physical examination, and pulmonary function tests [1]. The respiratory data obtained are analyzed by healthcare professionals to detect abnormalities and guide diagnosis. Despite their clinical utility, traditional pulmonary evaluation methods such as spirometry, imaging modalities (including computed tomography and chest X-ray), and auscultation face challenges such as inter-observer variability, limited accessibility, and the requirement for specialized expertise. These limitations underscore

Power System Technology
ISSN:1000-3673
Received: 16-09-2025 Revised: 05-10-2025 Accepted: 18-11-2025

the need for automated, scalable, and intelligent systems capable of improving diagnostic accuracy, supporting clinical decision-making, and enabling real-time monitoring.

Auscultation is a vital diagnostic procedure that allows clinicians to assess respiratory function by listening to lung and airway sounds using a stethoscope. Abnormal respiratory sounds, collectively known as adventitious sounds—such as crackles, wheezes, and rhonchi—are indicative of underlying pulmonary pathologies [2]. Accurate classification of these adventitious sounds and their association with specific respiratory diseases plays a crucial role in diagnosis and disease management.

With the rapid advancement of artificial intelligence (AI), deep learning (DL) techniques have emerged as powerful tools for analyzing complex biomedical signals. DL models can automatically extract intricate patterns from raw data, offering opportunities for automated pulmonary assessment and health informatics applications. Health informatics systems allow secure collection, storage, and management of patient data, while enabling remote health monitoring. Recent developments in transformer-based models, recurrent neural networks (RNNs), and convolutional neural networks (CNNs) have significantly enhanced pulmonary healthcare analytics. For instance, CNNs have demonstrated expert-level performance in detecting anomalies such as nodules, consolidation, and fibrosis from chest X-rays and CT images. Similarly, RNNs and their variants—such as long short-term memory (LSTM) and gated recurrent unit (GRU) networks—have been successfully employed for modeling time-series data like spirometric readings and respiratory sounds to predict dynamic respiratory patterns and detect pathological irregularities.

Health informatics has become a mainstream component of modern healthcare, offering patients self-monitoring capabilities through web-based digital tools [3]. Machine-learning (ML) techniques are particularly suited for interpreting complex data and answering clinically relevant questions [4]. Various DL frameworks have been explored for pulmonary disease diagnosis [5]. Mamalakis et al. developed a deep learning system to determine pathological ratios of lung diseases and map 3D anatomical lung models associated with pulmonary hypertension, thereby addressing prediction uncertainty [6]. Kocks et al. utilized electronic health record parameters from over 400,000 subjects to differentiate between asthma and chronic obstructive pulmonary disease (COPD) [7]. Ahmed et al. proposed an RNN-based model for diagnosing lung cancer using chest X-ray images [8], while Park et al. developed a semi-supervised learning model for early-stage lung cancer detection [9]. Similarly, Zafar et al. combined graphical deep features to classify multiple lung diseases such as COVID-19 and lung opacity [10], and Jameel et al. used X-ray chest images as inputs to Inception-ResNet and VGG16 models for pneumonia and COVID classification [11].

Hybrid and ensemble DL frameworks have also shown promise. Park et al. achieved a prediction accuracy of 92% and an F1-score of 81% by combining deep neural networks (DNNs) with traditional ML models [12]. Jiao et al. integrated multiple algorithms, including Genetic Algorithm (GA), Firefly Algorithm (FA), Particle Swarm Optimization (PSO), and Support Vector Machines (SVM), achieving effective diagnostic performance on the Cleveland Heart

Power System Technology
ISSN:1000-3673
Received: 16-09-2025 Revised: 05-10-2025 Accepted: 18-11-2025

Disease and Coswara cough sound datasets [13]. Gupta et al. employed differential evolution optimization to fine-tune the parameters of RF, AdaBoost, gradient boosting, LSTM, CNN, and RNN models [14]. The authors in [15] developed a model RespireNet based on CNN to perform classification of lung sounds with an accuracy of 77%. In [16] lung sound classification was performed using an ensemble of CNN. For the four lung sounds, an accuracy of 78% was achieved. Using CNN based methodologies, the authors in [17], discussed the results for lung disease namely COPD classification. The authors in [18] have used ResNet model as the base model to achieve multi class detection of lung sounds and respiratory diseases around 60% and 90% respectively.

The development of digital stethoscopes has allowed physicians to store and share respiratory sounds for research, education and consultation. Machine learning can enable automated analysis of these sounds and in turn can result in the development of intelligent stethoscopes [19]. The authors of [20] designed an intelligent system for predicting respiratory diseases. Combining digital stethoscopes with ML/DL techniques provides completely automated analysis of lung sounds.

Remote health monitoring using mobile phone app and cloud has been gaining momentum and popularity [21,22,23]. In a study by Joo et al, spirometric data were transmitted through a gateway to a cloud repository site. A smartphone game-based assessment of pulmonary function has shown good intraclass correlation coefficients with values from a spirometer for FVC and FEV1 of > 0.90 [24]. In a similar study on pediatric patients with cystic fibrosis and asthma, Kruizinga et al reported that FEV1 and FVC measured by a smartphone-connected spirometer were similar to those measured values from a conventional spirometer [25]. Cox et al mentioned that mobile apps can be a convenient telehealth intervention for pulmonary rehabilitation, especially if the subjects are socially or geographically isolated, are difficult to transport due to disease comorbidities or are engaged in full-time work [26]. Holland et al also opine that web-based rehabilitation, telerehabilitation and home-based rehabilitation have good clinical outcomes and hence are recommended for rehabilitation which needs minimal resources [27]. A disease diagnosis model for diabetes and heart disease was designed by Mansour et al by combing Ai and Internet of Things. Crow Search Optimization algorithm-based Cascaded Long Short-Term Memory (CSO-CLSTM) was employed for classification. The isolation forest (iForest) technique was used to remove outliers [28].

Despite significant progress, existing studies have largely focused on classifying either lung sounds or lung diseases independently. To the best of our knowledge, there are no existing works that perform simultaneous classification of both using a unified deep learning framework [29,30]. Addressing this gap, the present study introduces LuCoNet (Lung Convolution Network)—a customized multi-task deep learning model designed to classify both lung sounds and associated lung diseases concurrently. The model is deployed via cloud infrastructure (IaaS), enabling scalable computation, data storage, and remote accessibility. Complementing this, a mobile application named Svaastra (Lung Care App) facilitates patient-side data collection and real-time monitoring, allowing patients to upload demographic and sound data, while clinicians can access diagnostic results via a secure portal.

Power System Technology
ISSN:1000-3673
Received: 16-09-2025 Revised: 05-10-2025 Accepted: 18-11-2025

The paper outlines development in the following fields:

- Development of a cloud-integrated mobile platform for patients to upload demographic and respiratory sound data.
- Design of a customized multi-task deep learning model (LuCoNet) capable of simultaneous lung sound and disease classification.
- Deployment of a doctor's portal providing patient details and AI-generated classification reports for clinical decision support.

The remainder of this paper is organized as follows: Section II describes the methodology and system design of the proposed model and mobile application. Section III presents the experimental results and discussion. Section IV concludes the paper and outlines potential future directions.

2. Objectives

The main objectives of this study are as follows:

1. To design and develop a customized multi-task deep learning model (LuCoNet) capable of simultaneously classifying lung sounds and corresponding lung diseases from a single input source.
2. To integrate the proposed model with a cloud-based infrastructure (IaaS) for scalable computation, storage, and real-time accessibility of classification results.
3. To implement a mobile health application (Svaastra) that allows patients to record and upload respiratory sounds and demographic data to the cloud for automated analysis.
4. To develop a clinician-facing portal that provides patient details and AI-generated diagnostic reports, facilitating remote health monitoring and clinical decision-making.
5. To evaluate the performance of the proposed system using publicly available and real-world respiratory sound datasets, ensuring robustness, accuracy, and usability for clinical environments.

3. Methods

Fig. 1 presents an overall block diagram of the proposed method. There are two main aspects involved in this work: the design of AI-based model and a health informatics platform.

(i) **Design of AI-based model**: In this step, the lung sound dataset is procured, and the mel frequency cepstral coefficients (MFCC) are extracted. These coefficients were then used to train a deep learning model. The trained model was deployed onto the cloud which could be used for the classification of the test data.

(ii) **Health Informatics Platform**: The patients' demographic data as well as the lung sound are stored on the cloud data store via a mobile app. The app includes two portals, namely the patients portal and the doctors' portal. Both the portals have authenticated login and

Power System Technology
ISSN:1000-3673
Received: 16-09-2025 Revised: 05-10-2025 Accepted: 18-11-2025

bidirectional communication via the cloud. The patients' portal enables access to the patient's demographic data and upload lung sounds. The lung sound stored in the cloud was used by the trained deep learning model to identify any lung abnormalities. The Healthcare professionals can access lung abnormality labels along with patient details in the doctors' portal.

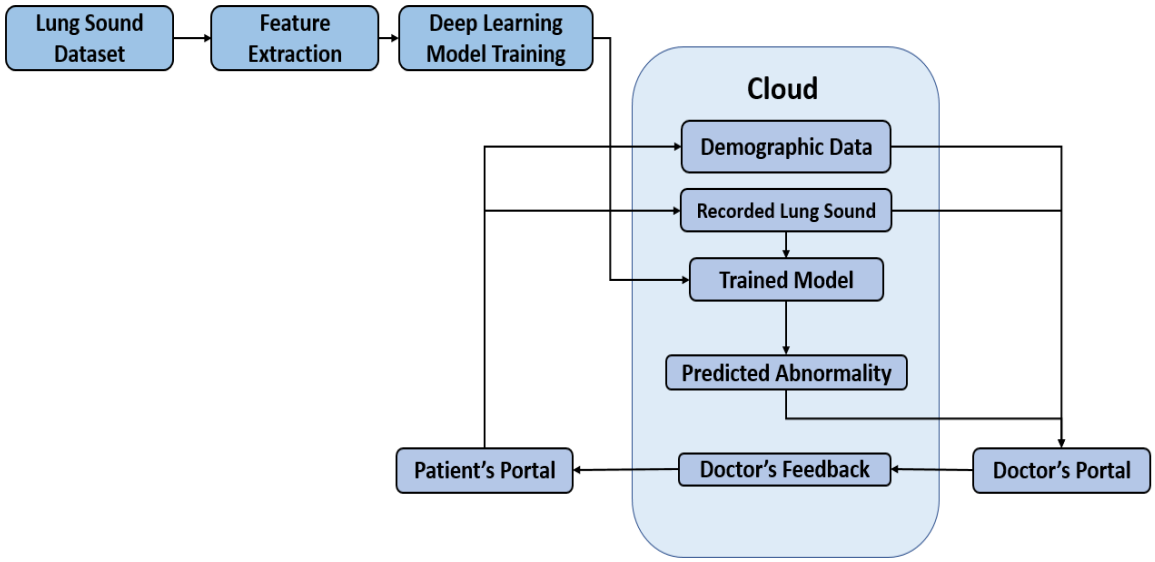

Fig. 1. Proposed Health Informatics System

The overview of the classification process is depicted in Fig. 2. The model is developed using the lung sound dataset, which consists of annotated respiratory audio signals that represent both normal and pathological circumstances (e.g., crackles, wheezes). Mel Frequency Cepstral Coefficients (MFCCs), which accurately represent the spectral characteristics of lung sounds in accordance with human auditory perception, are extracted from raw sound signals after they have been segmented. A condensed and discriminative representation of the acoustic signals is offered by the MFCC feature vectors, which are produced during brief time frames.

A novel deep neural network (DNN) architecture is engineered to process the acoustic signals. The network incorporates convolution layers for feature extraction. These convolution layers are followed by fully connected layers for nonlinear classification. Optimization of the architecture is achieved through hyper parameter tuning including layer depth, filter sizes, learning rate and regularization techniques. The model is trained under supervised learning paradigm using labeled MFCC features. Model performance is rigorously validated using stratified k-fold cross-validation to ensure robustness and reproducibility. Various metrics quantitatively assess the model before being deployed onto the cloud computing platform.

Power System Technology
ISSN:1000-3673
Received: 16-09-2025 Revised: 05-10-2025 Accepted: 18-11-2025

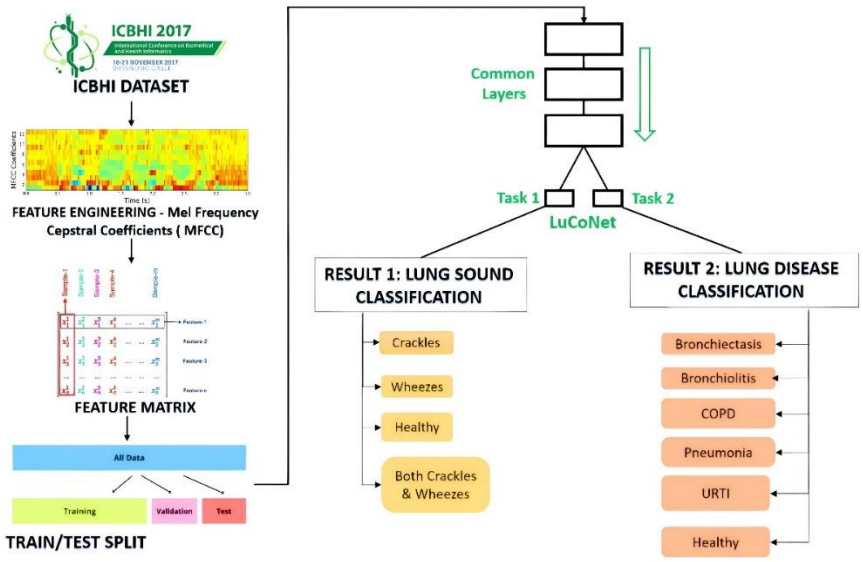

Fig.2 Classification Overview

3.1 Dataset

The process flow starts from the dataset collection to classify lung sounds and lung diseases. The ICBHI 2017 Respiratory Sound Dataset consisting of recordings totaling 5.5 hours, containing 6898 respiratory cycles was procured [31]. Within these recordings, there were instances of crackles and wheezes, with specific counts provided for each. The dataset is crucial for tasks such as the automatic classification of lung sounds, aiding in the diagnosis and treatment of respiratory diseases. Different lung diseases include COPD, bronchiectasis, bronchiolitis, URTI, pneumonia and Healthy. Different lung sounds included, crackles, wheezes, both crackles and wheezes along with healthy sets. The ICBHI data is split at a ratio of 80:20 for training and validation with testing. We have employed a simple hold-out validation method.

ICBHI dataset addresses four different types of abnormal lung sounds, crackles, wheezes, a combination of both crackles and wheezes and healthy sound. Adventitious sounds are often of a higher frequency than the normal ones. Crackles have a frequency range between 100 and 200Hz [32]. Wheezes have a frequency of 400 Hz or sometimes more than it [33]. To ensure uniformity and consistency in our study we have opted for a standard sampling rate of 16,000 Hz. This standard sampling rate is preferred for higher resolution of lung sounds.

3.2 Feature Engineering using MFCC

MFCCs are compact representations of the spectrum of an audio signal. The logarithmic transformation in MFCC is a unique and essential step, aimed at mimicking how humans perceive sound intensity. The human auditory system perceives sound logarithmically rather than linearly, meaning we are more sensitive to changes in quieter sounds than louder ones. This is why MFCC applies a log function after computing the Mel-filtered energy [34]. Other methods like, wavelet transform, linear predictive coding and chroma features do not inherently apply a

Power System Technology
ISSN:1000-3673
Received: 16-09-2025 Revised: 05-10-2025 Accepted: 18-11-2025

logarithmic scale. MFCC's use of the log function is particularly beneficial for audio signals where human perception matters, such as speech recognition and lung sound analysis. This transformation emphasizes perceptually important information, making it highly effective for medical diagnostics where subtle sound variations are crucial.

The amplitude of the waveform was measured on the y-axis in the time series representations of the audio samples. The change in pressure surrounding the microphone or audio receiver that first detects the audio signal is typically used to calculate the amplitude. The network input for the model was the Mel spectrogram features of lung sounds.

1. **Pre-Emphasis** - Pre-emphasis is a pre-processing technique used in the signal processing field to make up for the high frequency of the signal that was suppressed during generation. This is achieved by using a high pass filter [35] shown in (1).

   $$y[n] = x[n] - \alpha * x[n-1] \quad (1)$$

   where y[n]: output of the pre-emphasis filter at time index n, x[n]: input signal at time index n, α: pre-emphasis coefficient (typically between 0.9 and 1.0)

2. **Signal Framing and Windowing**- The audio signal was framed with a frame size of 25ms. The signal cut down to a smaller frame passed through a Kaiser fast window. Each frame passes through a window to taper the boundaries of the signal in the frame as shown in (2).

   $$w(n) = I_0\left(\beta\sqrt{\frac{1-\left(n-\frac{M}{2}\right)}{I_0 B}}\right) \quad (2)$$

   w(n): value of the window function for sample n.

   β (beta): parameter that controls the trade-off between main lobe width and sidelobe attenuation. Higher beta values led to narrower main lobes but higher sidelobes.

   M: window length (total number of samples).

   $I_0(\beta)$: modified Bessel function of the first kind, zeroth order, evaluated at β.

3. **Discrete Fourier Transform** – Discrete Fourier Transform is used to compute the power spectrum of the signal [35]. The power spectrum is the power of the frequency components stored inside the signal depicted in (3).

   $$X(k) = \sum_{n=0}^{N-1} x(n) e^{\frac{-2\pi jnk}{N}}, k = 1, 2, 3, \ldots . N-1 \quad (3)$$

   where X(k): DFT coefficient at frequency bin k, x(n): time-domain signal at sample n, N: total number of samples in the frame.

4. **The Mel filter bank**- Mel filter bank is a set of 40 triangular filters. We multiply the power spectrum of each filterbank to determine the energies and then total the coefficients. After

Power System Technology
ISSN:1000-3673
Received: 16-09-2025 Revised: 05-10-2025 Accepted: 18-11-2025

this is completed, we are left with 40 numbers that represent the amount of energy in each filterbank.

$$Hm(k) = \{\frac{k - f(m-1)}{f(m) - f(m-1)} \qquad f(m-1) \le k < f(m)1$$

$$k = f(m)\frac{f(m+1)-k}{f(m+1)-f(m)} \qquad f(m) < k \le f(m+1) \qquad (4)$$

where Hm(k): response of the mth filter at frequency bin k, f(m): center frequency of the mth filter.

5. **Log-** A power spectrum is output by the Mel filter bank. However, the way sound is perceived by humans is logarithmic [36]. Consequently, we log the output of the power spectrum as in (5).

$$m = 2595\ (1 + \frac{f}{700}) \qquad (5)$$

where m = Mel scale value corresponding to frequency f (in Hz).

6. **Discrete Cosine Transform-** To choose the majority of accelerative coefficients or to isolate the relationship in log spectral magnitudes from the filter bank in the MFCC process [35, 36], the DCT is applied to the Mel filter bank as shown in (6).

$$X(k) = \sum_{n=0}^{N-1} x\,(n) cos\left(\frac{2\pi jnk}{N}\right), k = 1, 2, 3, \ldots . N - 1 \qquad (6)$$

where X(k): DCT coefficient at index k, x(n): log energy at time frame n.

The final input feature matrix of form 40x862 was generated for the audio file. Here, 40 represents the number of MFCC coefficients and 862 represents the total number of time frames for each audio file. We used a frame size of 25ms and a maximum padding length of 862 to maintain uniformity among all audio files and to prevent the occurrence of negative values in the matrix. Hence, the final size of the feature matrix consisting of the extracted cepstral coefficients is (40x862).

3.3 Multi-Task Learning

Multi-task learning (MTL) tackles multiple learning tasks within a single model. This approach leverages the shared underlying knowledge between tasks to improve efficiency and performance and allows the model to share initial layers that extract general features, while having separate final layers for specific predictions. According to Caruana [37], MTL aims to improve generalization by learning tasks in parallel and sharing low-dimensional representations. This is done by leveraging domain knowledge from related tasks. The fundamental tenet is that mastering one task aids in mastering others. This is accomplished by

Power System Technology
ISSN:1000-3673
Received: 16-09-2025 Revised: 05-10-2025 Accepted: 18-11-2025

learning all tasks simultaneously and using the knowledge associated with them to improve the mastery of each task independently. Equation 7 is the standard formulation for a traditional MTL algorithm [38]:

$$W = [w^1 w^2 \ldots . w^m] \sum_{m=1}^{M} L\ [X^m, y^m, w^m] + \lambda Reg(W) \qquad (7)$$

where M: number of tasks, X_m : input feature matrix for task m, ym : target output vector for task m, wm : weight vector (regression parameters) for task m, L: loss function for task m, Reg(W): regularization function on the weight matrix W, λ : regularization parameter.

The task involves an m-th task represented by input matrix X^m while y^m is the output vector for the m-th task. The regression parameters are represented by the Weight vector wm, mapping X^m to y^m. The scalars D, Nm and M denote the features for each input matrix, number of samples and the number of tasks, respectively. W is attained by concatenating the weight vectors in $\{w^m\}$. This is designed based on assumptions about task relationships and prior knowledge. The regularization parameter, λ, helps in controlling the balance between the regularizer and loss function.

In the proposed work, single-input multi-output (SIMO) multi-task learning shown in Fig. 3, was utilized. Using this technique, both lung sound and lung disease can be predicted from a single input which is the MFCC feature vector. SIMO multi-task learning is an appropriate method as features are shared among the common layers thereby making it suitable for the prediction of two different outputs simultaneously.

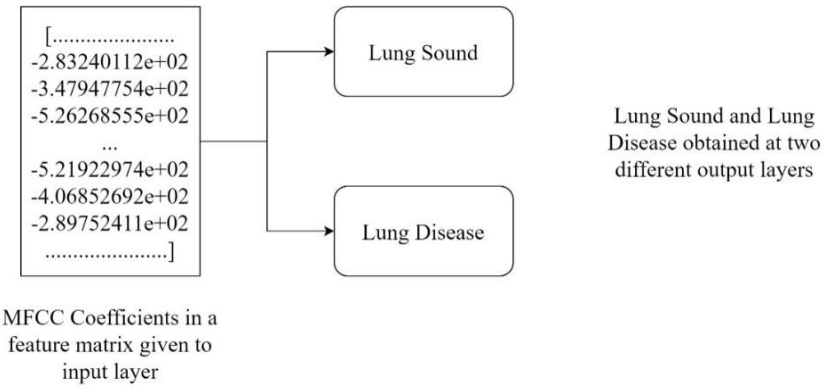

Fig. 3: Single-input multi-output (SIMO) multi-task learning

LuCoNet - Convolutional multi-tasking model

LuCoNet an acronym for Lung Convolution Network was custom built in this research work for pulmonary assessment. This model utilises the MFCC of the lung sounds and combines the annotation of both lung disease and lung sound, making it suitable for multi-task learning as mentioned in Section 3.3. The 2D convolutional layers are suitable for training the model. With 40 MFCC coefficients as rows and 862 padded lengths as column, the feature matrix was fed to

Power System Technology
ISSN:1000-3673
Received: 16-09-2025 Revised: 05-10-2025 Accepted: 18-11-2025

the model. The LuCoNet model consists of four convolutional blocks as shown in Fig. 4, where each block consists of a 2D depth-wise convolutional layer, 2D Max Pooling layer, a dropout layer and Batch Normalization layer. Each convolutional layer was fed with filters and an ReLU activation function. The number of filters in the convolutional layer of each convolutional block increases with a power of two [39].

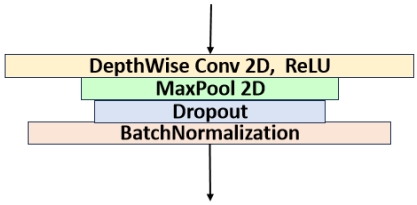

Fig.: 4 Individual Convolutional Block

Fig. 4 shows the architecture of Individual Convolutional Block

1. **Depth-wise Convolutional layer**: The function of the convolutional layer lies in its ability to extract features from input data. Depth-wise convolution is a type of convolution that can identify spatial correlations in the input data by applying a single convolutional filter per input channel, thereby minimizing the number of multiplications. This can result in faster training and inference times [7]. The ReLU activation function introduces non-linearity to the layer, enabling our network to learn the complex relationships between features.
2. **Max Pooling layer**: By choosing the maximum value inside each patch of the MFCC feature map, max pooling helps downsample MFCC feature maps by emphasizing the most notable features and removing less important data. This procedure helps mitigate overfitting in the model.
3. **Dropout layer**: uses diverse sets of neurons during each training iteration thus reducing the dependency on any specific set of neurons. This random deactivation prevents overfitting during training.
4. **Batch Normalization layer**: By normalizing the activations of a layer, Batch Normalization lessens the possibility of internal covariate shifts during training, which facilitates model optimization and enhances generalization capabilities.

The features extracted by the convolutional layers in this model were aggregated using Global Average Pooling and then sent to the dense layers which consisted of 256 units and a ReLU activation function, for final classification. This shared dense layer is then connected to the two branches of the output sound and disease, enabling it to acquire useful representations for tasks related to both sound and disease categorization.

Power System Technology
ISSN:1000-3673
Received: 16-09-2025 Revised: 05-10-2025 Accepted: 18-11-2025

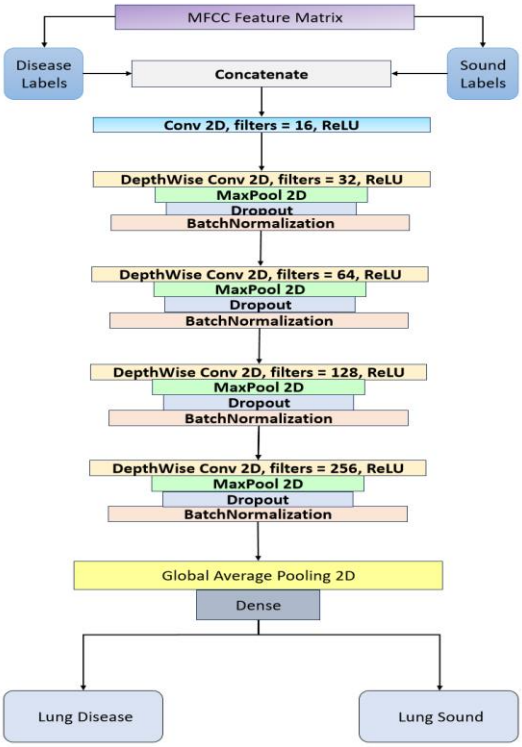

Fig. 5 Architecture of proposed LuCoNet model

The architecture of the LuCoNet model developed is depicted in Fig. 5. The convolutional block is designed to extract features from the depthwise convolutional layer, reduce spatial dimensions. reduce memory consumption through max pooling and dropout layers, and normalize the block using the batch normalization layer. Using a set of four blocks in a loop is not only efficient, but also improves the training time. Using a smaller number of blocks can make the model, simple and lightweight with a good computation time. As the number of blocks increased, the trainable parameters increased, thereby reducing efficiency. Table 1 presents the output feature shapes after each layer in the LuCoNet model.

Table: 1 Output Feature shape from each layer

Layer	**Output Feature shape**
Concatenate	(None, 40, 862, 2)
Convolutional 2D	(None, 20, 431, 16)
Convolution block 1	(None, 10, 216, 16)
Convolution block 2	(None, 5, 108, 16)

Power System Technology
ISSN:1000-3673
Received: 16-09-2025 Revised: 05-10-2025 Accepted: 18-11-2025

Convolution block 3	(None, 3, 54, 16)
Convolution block 4	(None, 2, 27, 16)
Global Avg. Pooling	(None, 16)
Dense	(None, 256)
Sound Output	(None, 4)
Disease Output	(None, 6)

The proposed model was carefully configured with the following hyperparameters:

1. ***Learning Rate***: Set to 0.001, the default value for the Adam optimizer. This learning rate provides a balanced trade-off between convergence speed and training stability, reducing the risk of overshooting minima during optimization.
2. ***Batch Size***: A batch size of 16 was chosen to balance computational efficiency and the stability of gradient updates. This size enables effective training with available resources while maintaining reliable gradient estimates.
3. ***Number of Epochs***: The model was trained for 20 epochs. This duration was sufficient to allow convergence without overfitting, given the dataset size (a small experimental dataset) and complexity.
4. ***Validation Split***: A validation split of 20% was used to monitor the model's generalization performance during training and prevent overfitting by early detection of performance degradation.
5. ***Dropout Rate***: A dropout rate of 0.2 was employed to mitigate overfitting by randomly deactivating neurons during training.
6. ***Optimizer***: The Adam optimizer was employed due to its adaptive learning rate capabilities and momentum, which improves convergence speed and robustness compared to traditional stochastic gradient descent.
7. ***Convolutional Filters***: The convolutional layers utilized increasing filter sizes of 16, 32, 64, 128, and 256 to enable hierarchical feature extraction, capturing both low-level and high-level patterns relevant to lung sound and disease patterns across the recorded file.
8. ***Kernel Size and Strides***: A kernel size of 3×3 with a stride of 2 was used in the initial convolutional layer to effectively capture local spatial features while downsampling the input for computational efficiency.
9. ***Activation Functions***: Rectified Linear Units (ReLU) were applied in hidden layers to introduce non-linearity and facilitate efficient gradient flow, while softmax activation was used in the output layers to generate probability distributions over the target classes.
10. ***Loss Functions***: Categorical cross-entropy loss was utilized for both output heads, suitable for multi-class classification tasks.
11. ***Input Shape***: The input dimensions were set to 40 × 862 × 1, corresponding to the MFCC feature representation of the audio data, thereby preserving both temporal and spectral characteristics for the convolutional network.

Power System Technology
ISSN:1000-3673
Received: 16-09-2025 Revised: 05-10-2025 Accepted: 18-11-2025

Significance of four convolutional blocks: Every block utilizes a certain number of filters which successively increases from filter size= 32 in the first convolutional block to filter size=256 in the fourth convolutional block. Using only four blocks has advantages such as lesser training time, fewer trainable parameters and simplification of the model. Adding more convolutional blocks can engender overfitting and the model undergoes biasing because it converges towards the class with a larger number of samples.

Table: 2 Significance of four convolutional blocks

No. of Convolutional blocks	**Sound Accuracy**	**Disease Accuracy**
1	60%	82%
2	62%	80%
3	64%	85%
4	65%	89%
5	63%	86%
6	63%	86%

Table 2 lists the accuracy obtained for lung sound and disease classification for a set of convolution blocks. As can be observed, increasing the number of blocks beyond four results in overfitting and a reduction in accuracy. Hence including only four convolutional blocks in the model offers better performance.

This trained model is deployed on a cloud platform to leverage the benefits of cloud computing infrastructure and services. The proposed work includes accessing these benefits through a mobile application for both patients and healthcare workers.

Svaastra - Mobile Application

Svaastra is an acronym for the Lung Care application, which is a mobile app designed in the proposed work. Central to the application architecture are portals for doctors and patients, that incorporate consumer-authenticated login functionality. This authentication technique, facilitated by a dual-motive window, calls for users to offer their credentials and identification. Within the patient portal, users can supply crucial demographic details, such as Name, Age, Gender, along with their diagnosis .wav files fostering green communication and diagnosis. Subsequently, the uploaded patient records, along with demographic statistics, are displayed on the doctor's portal, facilitating streamlined communication, and enhancing diagnostic accuracy.

Power System Technology
ISSN:1000-3673
Received: 16-09-2025 Revised: 05-10-2025 Accepted: 18-11-2025

In addition, logout functionality is included in both portals to protect information privacy. Communication between the utility and cloud platform is installed through FirebaseTM SDK, leveraging HTTP protocols to ensure steady client-server communication [40]. It operates on a request-response model, in which the consumer sends a request to the server, which hosts the cloud platform, and the server responds with the requested records.

Because the cloud is used as an IaaS, the patient data, lung sounds, trained classification model, classified labels and doctors' feedback are stored. The LuCoNet model generates classified labels, on the cloud. The cloud facilitates communication between patients and healthcare personnel.

The overall workflow of the mobile device application is shown in Fig. 6. Upon receiving the HTTP request, the server methods the incoming statistics, which may additionally contain responsibilities, along with validation and authentication. The platform generates a unique identifier (ID) for each record access. This unique ID serves as a reference factor for retrieving and updating information during subsequent operations. Once the demographic information is efficiently processed and stored within the cloud, the server generates an HTTP reaction, confirming the receipt of statistics. This reaction/response includes the results after predication and the respective details of the patient as a mark of reaction on the doctor's portal.

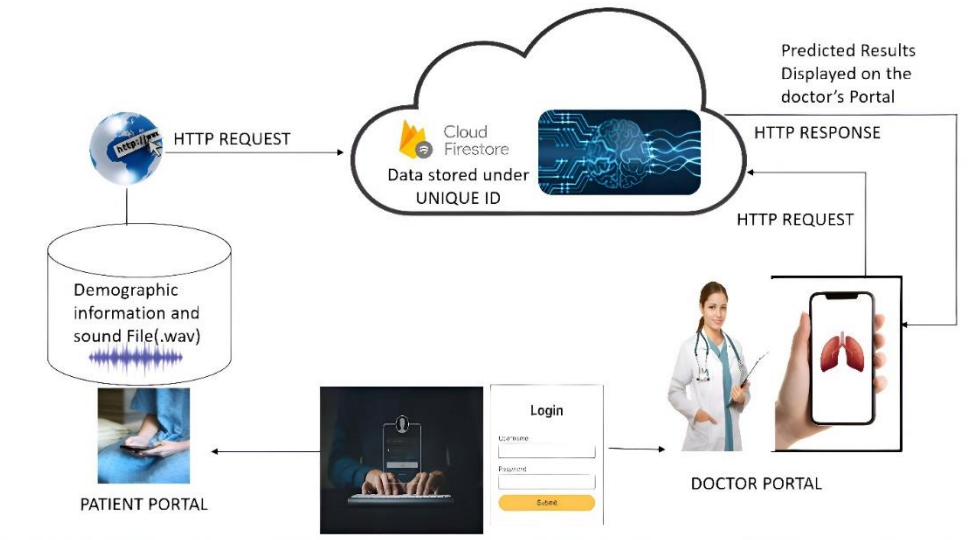

Fig. 6 Overview of Mobile Application

Svaastra is a secure health monitoring platform which can be conveniently used by all the stakeholders namely, the caretaker, patient and medical practitioner. After initial consultation the patient can have follow-up deliberation with the doctor via the app. Even without a vis-à-vis

Power System Technology
ISSN:1000-3673
Received: 16-09-2025 Revised: 05-10-2025 Accepted: 18-11-2025

interaction, teleconsulting can also be accomplished using the bi-directional channel provided by the app. The response of the patients over a specified period of time is stored by the app and a therapeutic study can be conducted. This can be used by the doctor to recommend further corrective regimen.

4. Results

4.1 Results from LuCoNet model

The model was trained on an NVIDIA RTX A6000 version 23.0.4 accelerated GPU. By simultaneously calculating the labels for both sound and disease, multi-task learning can clear the path. The detailed results are as follows:

Classification Report: Fig. 7 describes the classification results of LuCoNet model for both lung sound and lung diseases.

Classification Report - Sound:

	precision	recall	f1-score	support
Crackles	0.65	0.95	0.77	111
Healthy	0.00	0.00	0.00	10
Wheezes	0.72	0.27	0.39	49
Wheezes & Crackles	0.33	0.07	0.12	14
accuracy			0.65	184
macro avg	0.43	0.32	0.32	184
weighted avg	0.61	0.65	0.58	184

Classification Report - Disease:

	precision	recall	f1-score	support
Bronchiolitis	0.14	0.33	0.20	3
Bronchiectasis	0.00	0.00	0.00	3
COPD	0.93	0.99	0.96	159
Healthy	0.50	0.14	0.22	7
Pneumonia	1.00	0.14	0.25	7
URTI	0.50	0.40	0.44	5
accuracy			0.89	184
macro avg	0.51	0.34	0.35	184
weighted avg	0.88	0.89	0.86	184

Fig. 7: Classification Report – LuCoNet – Sound and Disease

This classification report offers a review of how well the model performed in categorizing sound recordings into various respiratory condition-related groups. 184 samples were used to test the model. The overall accuracy of the model for sound classification was 0.65 - indicating that 65% of the sound recordings could be accurately classified. The accuracy of the model for disease classification was 89%.

As observed from Fig. 7, the weighted average scores of precision, recall and F1-score for lung sound are 61%, 65% and 58% respectively. Similarly, the values for precision, recall and F1-score for lung diseases are 88%, 89% and 86% respectively. The sounds and diseases from the dataset namely, crackles and normal sounds along with COPD disease are richly trained. The model can be utilized for detecting COPD in patients and COPD being one of the crucial and highly growing lung disease [41]. Due to imbalance in the data distribution, the model has a reduced accuracy.

Accuracy and Loss: When training the model, the most crucial factors to consider are the accuracy and losses. The model was trained using various epochs and batch sizes under various settings. With 20 epochs and 32 batches, an accuracy of 65% for sound and 89% for illness were achieved. The accuracy and loss during validation on the test set are shown in Fig. 8 and Fig 9.

Power System Technology
ISSN:1000-3673
Received: 16-09-2025 Revised: 05-10-2025 Accepted: 18-11-2025

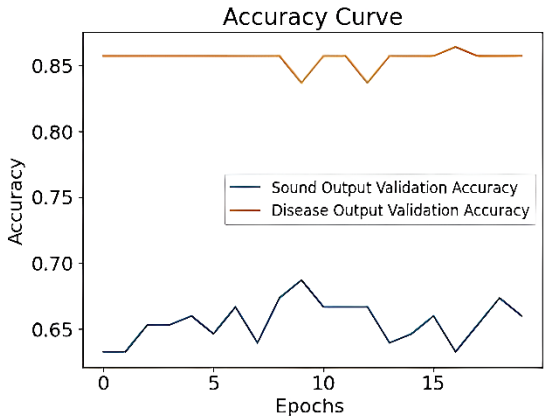

Fig.: 8 Validation Accuracy Curve for Sound & Disease

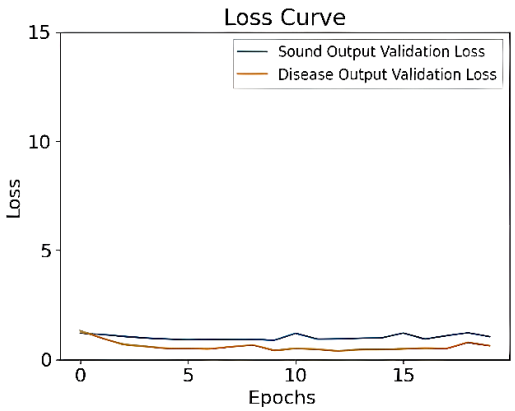

Fig.: 9 Validation Loss Curve for Sound & Disease

ROC Curves: The receiver operating characteristic (ROC) curve for lung sound is as shown in Fig. 10. The Area Under Curve (AUC) for the four types of lung sounds is seen at an average of 75.5%. The ROC for lung diseases is shown in Fig 10 and it can be observed that the AUC for the six lung diseases is an average of 93.6%. This indicates a very high probability of the model to predict the diseases correctly.

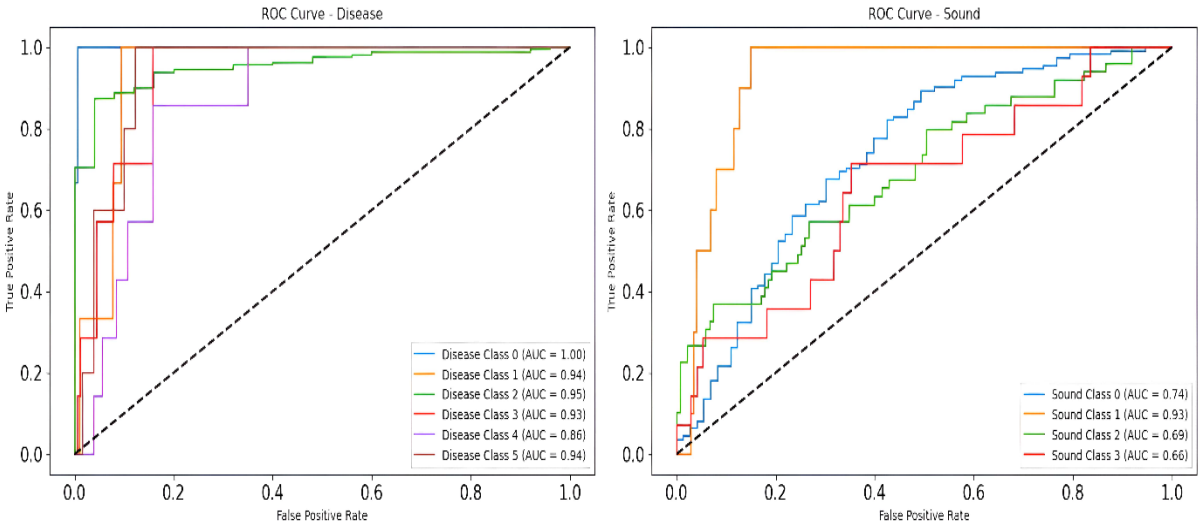

Fig 10. ROC Curve for Lung Sound and Lung Disease

Power System Technology
ISSN:1000-3673
Received: 16-09-2025 Revised: 05-10-2025 Accepted: 18-11-2025

4.2 Results from Svaastra

The mobile application provides access to both the patient and healthcare provider with an authenticated login. The patient could enter demographic details along with the lung sounds. The model deployed on the cloud classifies the lung abnormalities. The demographic details along with the classification results are visible to the doctor, as shown in Fig. 11. In turn, the doctor can offer recommendations to the patient through the app.

Fig.: 11 Patient details and predictions shown on Mobile Application

5. Discussion

Lung sounds play a pivotal role in assessing lung conditions. Acquisition of these lung sounds is usually performed using a simple non-invasive diagnostic instrument, stethoscope. Lung sounds have a frequency range of 50 to 2500 Hz. Abnormal sounds can reach a frequency of 4000 Hz. Owing to this special feature of lung sounds, a highly sensitive mic is required for the proper acquisition of lung sounds. An intelligent acquisition system requires microphones with high sensitivity. Two microphones, SEN0487 and SEN0526 have a frequency response range of 100 Hz to 8000 Hz. This frequency range is suitable, and the microphone can capture all ranges of adventitious sounds. The SEN0487 mic is compatible with the microcontroller - ARDUINO UNO R3, and SEN0526 is compatible with the Espressif ESP32 Wroom-32E module.

Power System Technology
ISSN:1000-3673
Received: 16-09-2025 Revised: 05-10-2025 Accepted: 18-11-2025

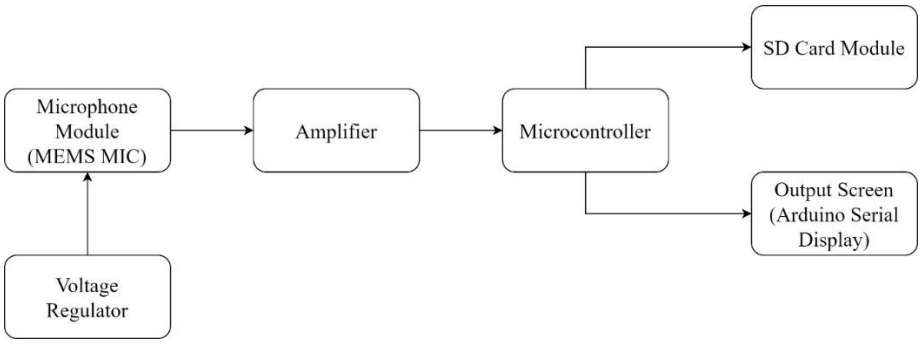

Fig. 12 Acquisition of real time audio signal

A data acquisition module was built using high sensitivity microphones interfaced with microcontrollers. The analog output of the microphones was assigned to the in-built ADCs of the microcontrollers. Fig. 12 shows the flow followed to obtain .wav file. The microphone has three basic elements: a voltage regulator to ensure consistent and reliable voltage, an MEMS mic module, which is a microphone module, and an amplifier to amplify the signal. An amplified signal was produced and fed to the microcontroller. The inbuilt ADC of the microcontroller takes continuous input and produces discrete samples at a sampling rate of 44.1KHz. These samples are stored in the external SD storage medium. The samples are stored on to a text file and all these stored values can be reconstructed using different audio libraries like SciPy which is capable of assembling the samples and forming a wav file. Reconstruction in this context refers to recalling amplitudes in order to combine them into a single audio file. The wav file is directly given as input to pre-processing, followed by predictions from the multi-tasking deep learning model.

Fig. 13 shows the signal reconstructed from the digital data. The audio was sampled using the ADCs of the microcontroller, and the obtained values were stored in text file that was later on reconstructed.

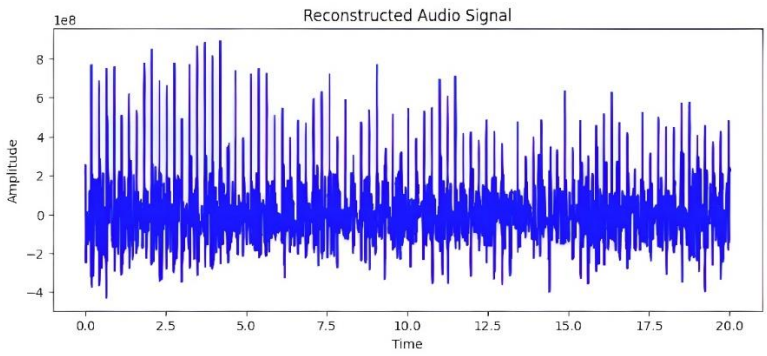

Fig. 13 Reconstructed audio signal from digital input

Power System Technology
ISSN:1000-3673
Received: 16-09-2025 Revised: 05-10-2025 Accepted: 18-11-2025

In this study a multi-tasking model capable of predicting lung sound and lung disease using a single LuCoNet model was developed. Previous studies have shown the prediction of lung sounds and lung diseases separately using different models and a detailed summary of previous related work using the ICBHI Database is shown in Table 3.

Table 3: Comparison of proposed work with existing literature

Study	Model	Accuracy (%)
Sound		
Serbes et al. [42]	SVM	49.86
Chambres et al. [43]	SVM	49.1
Minami et al. [44]	VGG-Net	52.79
Ma et al. [45]	Bi-ResNet	50.16
Acharya et al. [46]	CNN-RNN	66.31
Proposed work	LuCoNet	65
Disease		
Tariq et al. [47]	2D CNN	97
Do et al. [48]	CNN, MLP	94.1, 99.2
Demir et al. [49]	CNN+SVM	65.5
Hazra et al. [50]	2D CNN	92.39
Proposed Work	LuCoNet	89

As can be observed from Table 3, the LuCoNet model performance is on par with the existing models for the same dataset. The sound classification accuracy was approximately 65% and the diseases classification accuracy was approximately 89%.

The proposed customized model includes the idea of using multi-task learning to predict both lung sound and lung disease simultaneously [51]. This results in a reduction in the training time to 8.267 s in comparison to the previous experimentation of MTL using pre-trained CNN models, such as ResNet, MobileNet and DenseNet [51], as shown in Table 4. The proposed model has less trainable parameters (11,514) and the sound and disease classification accuracies were similar to the results in [51].

Power System Technology
ISSN:1000-3673
Received: 16-09-2025 Revised: 05-10-2025 Accepted: 18-11-2025

Table: 4 Comparison between previous models and proposed LuCoNet model

Model name	Trainable parameters	Training Time (s)	Sound Accuracy (%)	Disease Accuracy (%)
2D CNN	154,986	24.164	66	89
ResNet 50	25,642	9.846	66	86
DenseNet	40,714	36.786	60	86
MobileNet	397,834	36.892	74	91
LuCoNet	11,514	8.267	65	89

The ICBHI dataset is rich in COPD patients whereas very little data is present for other lung abnormalities. Almost 80% of the dataset focuses on COPD and crackles data. In future iterations, we aim to explore techniques such as class weighting, oversampling, or synthetic data generation to further enhance performance on underrepresented classes.

6. Conclusion and Future Scope

Although various research studies are available for lung sound analysis leading to detection of type of lung sound or classification of lung diseases, there is a lack of research integrating both in a user-friendly environment. The current work has developed a single Deep Learning model which can detect both lung sound abnormality as well as lung disease, simultaneously, in a convenient Mobile Application. In the proposed work, a health informatics system employing an interactive portal LuCApp for patients and healthcare personnel was developed. The computational backbone is a customized deep learning model, LuCoNet, which can concurrently predict lung sounds and lung diseases. The cloud service deployed here is the Infrastructure As A Service for the storage of data, functioning of the computational model, and communication between stake-holders. The prediction accuracy of LuCoNet model classification was 89% for lung diseases and 65% for lung sounds. The training time of this model was 8.267 s on an NVIDIA RTX A6000 version 23.0.4 accelerated GPU system. The interactive mobile app provides an authenticated login for both the patient and the doctor. The patient can upload their demographic data and lung sounds while the doctor can access this information along with the classification label provided by the model. Doctors can also communicate their professional recommendations to patients via this portal. Thus, this mobile app plays an important role in clinical diagnostics by facilitating remote monitoring, therapeutic study and teleconsulting.

One of the major bottlenecks for our study has been the lack of sufficient data. Further work involves enhancing the dataset to improve the prediction accuracy, especially for lung sounds.

Power System Technology
ISSN:1000-3673
Received: 16-09-2025 Revised: 05-10-2025 Accepted: 18-11-2025

This enhancement of the dataset may include different populations from various demographics and ethnicities, which can help in a comprehensive experimentation. Accordingly, the model must be re-trained periodically as and when new data are available. While the current model demonstrates strong performance using empirically determined hyperparameters, subsequent work could benefit from a systematic approach to hyperparameter optimization. Specifically, adjusting the learning rate, batch size, convolutional filter configurations, and the incorporation of dropout layers could further enhance the model's accuracy. Should sufficient data be available, use of computationally intensive models like transformers, may offer more precise predictions. As there is a lack of similar work, more experimentation for justifying the model is to be conducted.

7. Acknowledgement

The authors would like to thank all the doctors for their valuable support during this project. This work was supported by the MSRIT Seed Funding (MSRIT/PRIN/2023/1598).

Informed consent and patient details – No subjects were involved in this study. For validation purposes, only open access datasets were used as described in the paper.

Declaration of competing interest – The authors declare that they have NO competing interests in the execution of this work.

References

[1] Zander Williams, James H. Hull, Ben Moseley. (2023). Assessment of pulmonary function. Medicine, 51(10), 693-698.

[2] Zimmerman B, Williams D. (2024). Lung Sounds. StatPearls Publishing.

[3] Himes BE, Weitzman ER. (2016). Innovations in health information technologies for chronic pulmonary diseases. Respir Res.

[4] Talha Khan Burki. (2019). The Role of AI in diagnosing lung Diseases.The Lancet Respiratory Medicine, 7(12), 1015-1016.

[5] Sfayyih, A.H.; Sabry, A.H.; Jameel, S.M.; Sulaiman, N.; Raafat, S.M.; Humaidi, A.J.; Kubaiaisi, Y.M.A. (2023). Acoustic-Based Deep Learning Architectures for Lung Disease Diagnosis: A Comprehensive Overview. Diagnostics, 13, 1748. https://doi.org/10.3390/diagnostics13101748

[6] Mamalakis, M., Dwivedi, K., Sharkey, M., SamerAlabed, David Kiely & Andrew J. Swift. (2023). A transparent artificial intelligence framework to assess lung disease in pulmonary hypertension. Scientific Reports, 13, 3812.

[7] Kocks JWH, Cao H, Holzhauer B, Kaplan A, FitzGerald JM, Kostikas K, Price D, Reddel HK, Tsiligianni I, Vogelmeier CF, Bostel S, Mastoridis P. (2023). Diagnostic Performance of a Machine Learning Algorithm (Asthma/Chronic Obstructive Pulmonary Disease [COPD] Differentiation Classification) Tool Versus Primary Care Physicians and

Power System Technology
ISSN:1000-3673
Received: 16-09-2025 Revised: 05-10-2025 Accepted: 18-11-2025

Pulmonologists in Asthma, COPD, and Asthma/COPD Overlap. J Allergy ClinImmunolPract,11(5), 1463-1474.

[8] H. N. Ahmed, M. O. Ibrahim, and A. R. Suleiman, "A hybrid RNN-based deep learning model for lung cancer and COPD detection," Engineering, Technology & Applied Science Research (ETASR), vol. 15, no. 2, pp. 10255–10260, 2025.

[9] S. L. Park, T. J. Wang, and Y. S. Kim, "Lightweight advanced deep neural network (DNN) model for early-stage lung cancer detection," Diagnostics, vol. 14, no. 21, p. 2356, Nov. 2024.

[10] A. Zafar, M. S. Javed, and M. M. Khan, "DeepChestGNN: A comprehensive framework for enhanced lung disease identification through advanced graphical deep features," Sensors, vol. 24, no. 9, p. 2830, May 2024.

[11] M. K. Jameel and S. K. Jain, "A multiclass deep learning algorithm for healthy lung, COVID-19, and pneumonia disease detection from chest X-ray images," Discover Artificial Intelligence, vol. 4, no. 1, pp. 1–14, Jan. 2024.

[12] Dong Jin Park, Min Woo Park, Homin Lee, Young-Jin Kim, Yeongsic Kim, Young Hoon Park. (2021). Development of machine learning model for diagnostic disease prediction based on laboratory tests. Scientific Report. 11(7567).

[13] Jiao, N. (2023). An efficient disease prediction framework based on optimized machine learning models for a smart healthcare application. Multimed Tools Appl.

[14] Punit Gupta, FurqanRustam, Khadija Kanwal, WajdiAljedaani, Sultan Alfarhood, MejdlSafran, Imran Ashraf. (2024). Detecting Thyroid Disease Using Optimized Machine Learning Model Based on Differential Evolution. International Journal of Computational Intelligence Systems, 17(3).

[15] S. Gairola, F. Tom, N. Kwatra and M. Jain. (2021). RespireNet: A Deep Neural Network for Accurately Detecting Abnormal Lung Sounds in Limited Data Setting.43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC),527-530.

[16] T. Nguyen and F. Pernkopf, (2020). Lung Sound Classification Using Snapshot Ensemble of Convolutional Neural Networks. 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 760-763, doi: 10.1109/EMBC44109.2020.9176076.

[17] Srivastava A, Jain S, Miranda R, Patil S, Pandya S, Kotecha K. (2021). Deep learning based respiratory sound analysis for detection of chronic obstructive pulmonary disease. PeerJ Computer Science, 7.

[18] T. Nguyen and F. Pernkopf, (2022). Lung Sound Classification Using Co-Tuning and Stochastic Normalization. IEEE Transactions on Biomedical Engineering, 69(9), 2872-2882.

Power System Technology
ISSN:1000-3673
Received: 16-09-2025 Revised: 05-10-2025 Accepted: 18-11-2025

[19] SubasingheAkila, AbeywickramaAsira, DissanayakeHarsha, SachithaHerathIsuru, JayeratneRajitha, DinalankaraRandima, EdirisingheSajith, JayasooriyaVidura. (2022). Smart Stethoscope: Intelligent Respiratory Disease Prediction System.2nd International Conference on Advanced Research in Computing (ICARC), 242-247.

[20] Huang DM, Huang J, Qiao K, Zhong NS, Lu HZ, Wang WJ. (2023). Deep learning-based lung sound analysis for intelligent stethoscope. Mil Med Res, 10(1).doi: 10.1186/s40779-023-00479-3.

[21] Devos P., Debeer J., Ophals J., Petrovic M. (2019) Cognitive impairment screening using m-health: An android implementation of the mini-mental state examination (MMSE) using speech recognition. Eur. Geriatr. Med.

[22] Brittz M., Heinze B., Mahomed-Asmail F., Swanepoel D.W., Stoltz A. (2019). Monitoring Hearing in an Infectious Disease Clinic with mHealth Technologies. Journal of the American Academy of Audiology, 30(6), doi.org/10.3766/jaaa.17120.

[23] Pérez Y.I.V., Medlow S., Ho J., Steinbeck K. (2019). Mobile and Web-Based Apps That Support Self-Management and Transition in Young People with Chronic Illness: Systematic Review. Journal of medical Internet research, 21(11).

[24] Joo S., Lee K., Song C. (2018). A comparative study of smartphone game with spirometry for pulmonary function assessment in stroke patients. Biomed. Res. Int.

[25] Kruizinga M. D., Essers E., Stuurman F. E., Zhuparris A., van Eik N., Janssens H. M., et al. (2020). Technical validity and usability of a novel smartphone-connected spirometry device for pediatric patients with asthma and cystic fibrosis. Pediatr. Pulmonol, 55 2463–2470. Doi:10.1002/ppul.24932

[26] Cox NS, Dal Corso S, Hansen H, Narelle S Cox, Simone Dal Corso, Henrik Hansen, Christine F McDonald, Catherine J Hill, Paolo Zanaboni, Jennifer A Alison, Paul O'Halloran, Heather Macdonald, Anne E Holland. (2021). Telerehabilitation for chronic respiratory disease. Cochrane Database Syst Rev, 1(1):CD013040.

[27] Anne E Holland, Narelle S Cox, LinzyHouchen-Wolloff, Carolyn L Rochester, Chris Garvey, Richard ZuWallack, Linda Nici, Trina Limberg, Suzanne C Lareau, Barbara P Yawn, Mary Galwicki, Thierry Troosters, Michael Steiner, Richard Casaburi, Enrico Clini, Roger S Goldstein. (2021). Defining modern pulmonary rehabilitation. Ann Am Thorac Soc. 18(5):33929307.

[28] R. F. Mansour, A. E. Amraoui, I. Nouaouri, V. G. Díaz, D. Gupta and S. Kumar. (2021). Artificial Intelligence and Internet of Things Enabled Disease Diagnosis Model for Smart Healthcare Systems, IEEE Access, 9, 45137-45146.

[29] Fraiwan, M., Fraiwan, L., Alkhodari, M., O. Hassanin. (2022). Recognition of pulmonary diseases from lung sounds using convolutional neural networks and long short-term

Power System Technology
ISSN:1000-3673
Received: 16-09-2025 Revised: 05-10-2025 Accepted: 18-11-2025

memory. J Ambient Intell Human Comput 13, 4759–4771. https://doi.org/10.1007/s12652-021-03184-y

[30] Petmezas G, Cheimariotis GA, Stefanopoulos L, Rocha B, Paiva RP, Katsaggelos AK, Maglaveras N. (2022). Automated Lung Sound Classification Using a Hybrid CNN-LSTM Network and Focal Loss Function. Sensors (Basel), 22(3):1232. doi: 10.3390/s22031232. PMID: 35161977; PMCID: PMC8838187.

[31] B. Rocha, D. Filos, L. Mendes, I. Vogiatzis, E. Perantoni, E. Kaimakamis, P. Natsiavas, A. Oliveira, C. Jacome, A. Marques, R. P. Paiva, I. Chouvarda,Carvalho, N. Maglaveras . (2017). A respiratory sound database for the development of automated classification, Precision Medicine Powered by pHealth and Connected Health: ICBHI 2017, Thessaloniki, Greece, 18-21 November. Springer, 2018, pp. 33–37.

[32] Reichert S, Gass R, Brandt C, Andrès E. (2008). Analysis of respiratory sounds: state of the art. Clin Med CircRespiratPulm Med. 16(2), 45-58. doi: 10.4137/ccrpm.s530. PMID: 21157521; PMCID: PMC2990233.

[33] Meslier N, Charbonneau G, Racineux JL. Wheezes. (1995). EurRespir J. 8(11):1942-1948. doi: 10.1183/09031936.95.08111942. PMID: 8620967.

[34] Z. K. Abdul, A. K. Al-Talabani. (2022). Mel Frequency Cepstral Coefficient and its Applications: A Review, IEEE Access, 10, 122136-122158. doi: 10.1109/ACCESS.2022.3223444.

[35] John G. Proakis, Dimitris G. Manolakis. (1996). Digital signal processing (3rd ed.): Principles, Algorithms, and Applications. Prentice-Hall, Inc., USA.

[36] Z. K. Abdul, A. K. Al-Talabani. (2022). Mel Frequency Cepstral Coefficient and its Applications: A Review. IEEE Access, 10, 122136-122158.

[37] Caruana R. (1998). Multi-task learning. Learning to learn, Springer, 95–133.

[38] Thung, KH., Wee, CY. (2018). A brief review on multi-task learning. Multimedia Tools Applications 77, 29705–29725.

[39] Ahmed, Wafaa, Karim, Abdulamir. (2020). The Impact of Filter Size and Number of Filters on Classification Accuracy in CNN. Proceedings of the International Conference on Computer Science and Software Engineering (CSASE), Duhok, Iraq, 88-93. doi: 10.1109/CSASE48920.2020.9142089.

[40] Chougale, Pankaj, Yadav, Vaibhav, Gaikwad, Anil, BharatiVidyapeeth. (2022). Firebase -Overview and Usage. Journal of Engineering and Technology Management. 2582-5208.

[41] Brandsma CA, Van den Berge M, Hackett TL, Brusselle G, Timens W. (2020). Recent advances in chronic obstructive pulmonary disease pathogenesis: from disease mechanisms to precision medicine. J Pathol. 250(5): 624-635. doi: 10.1002/path.5364. Epub 2019 Dec 3. PMID: 31691283; PMCID: PMC7216938

[42] Serbes G., Ulukaya S., Kahya Y.P. (2017). An Automated Lung Sound Preprocessing and Classification System Based OnSpectral Analysis Methods. Proceedings of the

Power System Technology
ISSN:1000-3673
Received: 16-09-2025 Revised: 05-10-2025 Accepted: 18-11-2025

International Conference on Biomedical and Health Informatics. Springer; Singapore: 45–49.

[43] Chambres G., Hanna P., Desainte-Catherine M. (2018). Automatic Detection of Patient with Respiratory Diseases Using Lung Sound Analysis; Proceedings of the International Conference on Content-Based Multimedia Indexing (CBMI); La Rochelle, France. 4–6 September.

[44] Minami K., Lu H., Kim H., Mabu S., Hirano Y., Kido S. (2019). Automatic Classification of Large-Scale Respiratory Sound Dataset Based on Convolutional Neural Network; Proceedings of the 2019 19th International Conference on Control, Automation and Systems (ICCAS), 15–18 October, 804-807.

[45] Ma Y., Xu X., Yu Q., Zhang Y., Li Y., Zhao J., Wang G. (2019). LungBRN: A Smart Digital Stethoscope for Detecting Respiratory Disease Using bi-ResNet Deep Learning Algorithm; Proceedings of the IEEE Biomedical Circuits and Systems Conference (BioCAS), Nara, Japan. 17–19 October, 1-4.

[46] Acharya J., Basu A. (2020). Deep Neural Network for Respiratory Sound Classification in Wearable Devices Enabled by Patient Specific Model Tuning. IEEE Trans. Biomed. Circuits Syst. 14:535–544.

[47] Z. Tariq, S. K. Shah, Y. Lee. (2019). Lung Disease Classification using Deep Convolutional Neural Network. Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine (BIBM), San Diego, CA, USA, 732-735,

[48] Do QT, Lipatov K, Wang HY, Pickering BW, Herasevich V. (2021). Classification of Respiratory Conditions using Auscultation Sound. AnnuIntConf IEEE Eng Med Biol Soc. Nov,1942-1945.

[49] Demir F, Sengur A, Bajaj V. (2019). Convolutional neural networks based efficient approach for classification of lung diseases. Health InfSci Syst. Dec 23, 8(1):4.

[50] Hazra, Reetodeep, Majhi, Sudhan. (2020). Detecting Respiratory Diseases from Recorded Lung Sounds by 2D CNN. 5th International Conference on Computing, Communication and Security (ICCCS), Patna, India, 1-6, doi: 10.1109/ICCCS49678.2020.9277101.

[51] Suma K V, DeepaliKoppad, Preethi Kumar, Neha A Kantikar, Surabhi Ramesh. (2024). Multi-Task Learning for Lung sound & Lung disease classification. 2404.03908, arXiv.