Respiratory Disease Detection Using Lung Sound with CNN

Mrs.Sk.Nageena Jani[1],Mrs.J.Vidya[2] , M.Sneha[3*] , K.Jaya Shankar[4] ,N.NarendraBabu[5], K.Sathish[6]

[1,2]Assistant Professor, Department of CSE, Vignan's Lara Institute of Technology &Science, Vadlamudi, Guntur, Andhra Pradesh, India.
[3,4,5,6]Final Year, Department of CSE, Vignan's Lara Institute of Technology &Science, Vadlamudi, Guntur, Andhra Pradesh, India.

nageenavlits@gmail.com[1],vidyajyothula9@gmail.com[2],mulipirisneha@gmail.com[3*], kjshankar77@gmail.com[4], narendranaripeddi@gmail.com[5], kusathish79@gmail.com[6]

Abstract. Every year, respiratory disorders affect millions of people worldwide and pose a serious threat to public health. For treatment and therapy to be effective, a timely and accurate diagnosis is essential. In this work, we present a unique method that uses examination of lung sounds to improve the classification of pulmonary ailments. Our framework provides a comprehensive solution for autonomous respiratory audio analysis by utilizing cutting-edge deep learning techniques, namely a mixed convolutional neural network model that incorporates MFCC, or Mel-Frequency Cepstral Coefficients, chroma characteristicsand Mel-Spectrogram features. Our architecture focuses on speed optimization, ensuring rapid clinical application without sacrificing classification accuracy. It does this by using a lightweight combination of neural network models. By conducting extensive testing on several datasets, such as the chest wall lung sound database and the ICBHI 2017 competition database, our approach shows mastery in categorizing respiratory illnesses into several groups. Extensive assessment indicators, including recall, accuracy, precision, and F1 score, offer deep insights into our models' effectiveness. These results highlight how deep learning approaches may transform pulmonary diagnostics, giving medical personnel the essential tools they need to intervene quickly and improve patient outcomes in the process. With 95% accuracy, this model predicts respiratory illness.

Keywords: Respiratory disease classification, convolutional neural network, deep learning, MFCC, mel-spectrogram features.

1 Introduction

All around the world, respiratory disorders are among the main reasons for demise and impairment. The world's poorest regions exhibit the highest incidence of sickness. The researchers conclude that aging, body mass, smoking, and environmental pollution are all risk factors that have a significant impact. With 3.91

K. R. Madhavi et al. (eds.), *Proceedings of the International Conference on Computational Innovations and Emerging Trends (ICCIET 2024)*, Advances in Computer Science Research 112,
https://doi.org/10.2991/978-94-6463-471-6_84

million deaths, chronic obstructive lung disease was predicted to take the third position most prevalent leading cause of death worldwide in 2017. As such, it is a serious public health issue. Worldwide, an estimated 65 million people have chronic respiratory conditions. The number of deaths from chronic respiratory disorders rose from 3.32 million to 3.91 million, an 18% rise, from 1990 to 2017. Taking into account that 334 million individuals, or 14% of all children worldwide, are impacted.

Millions of people die each year from respiratory diseases like pneumonia, which also account for the majority of mortality among children under five. Over 10 million individuals are afflicted with tuberculosis (TB), the most common and fatal contagious illness, it results in 1.4 million deaths annually. The worst cancer, lung cancer, claims the lives of 1.6 million individuals a year. Chronic respiratory disorders are the cause of four million premature deaths globally. Respiratory illnesses account for five of the top 30 causes of death asthma ranks twenty-eight, COPD comes in third, tuberculosis comes in at number twelve, and lung, tracheal, and bronchial cancer come in at number six. Lower respiratory tract infections rank fourth. More than a billion individuals worldwide experience acute or long-term respiratory illnesses.

The sobering fact is that 4 million premature deaths worldwide are attributed to chronic respiratory diseases each year. It is especially dangerous for new born and young children. Nine million children death by under five years old from pneumonia every year, making it the primary reason for death. worldwide for this age group. Although breathing and respiratory health are frequently taken for granted, the lung is an essential organ that can be harmed by airborne infections. disorders affecting the respiratory system have a major influence on individuals' social, financial, and well-being lives. The most significant factor influencing death and disability rates was social deprivation, with the world's poorest places having the highest rates. In wealthy nations, lower death rates indicate better access to healthcare and advancements in medical research. Lung disease therapy is crucial in the medical field since lung disease is the leading cause of death globally. These features have led to a great deal of research into the early diagnosis and treatment of respiratory disorders. While it takes time and experience to effectively detect health risks associated with this information, statistics from the World Health Organization (WHO) indicate that in 45% of WHO Member States, there is fewer than one doctor per 1000 people, which is below the required ratio. When taking these numbers into consideration, errors can occur when a health professional who is already overbooked studies and diagnoses each patient individually. That's why it's critical to find innovative solutions to help physicians save time. Therefore, automated and dependable instruments can aid in the diagnosis of a greater number of patients and also assist specialists in reducing errors that may arise from overwork.

2 Literature Review

Fabio Martinelli et.al[1] proposed a deep learning network trained on Mel spectrogram characteristics was used to analyze respiratory audio recordings from patients with different respiratory illnesses. During testing, the prototype produced a correctness of 71.81%, a exactness of 0.798, and a recall of 0.849. The Grad-CAM technique was also used in the study to highlight spectrogram regions that are pertinent to disease identification, offering visual explanations for the model's predictions. With this method, the deep learning model should be easier for medical experts to understand and possibly more useful in diagnosing respiratory diseases in the real world.

Araya chatchaiwatkul et.al [2] suggested model achieves good accuracy in identifying and categorizing lung disorders from chest X-ray pictures by utilizing deep learning with VGG16 architecture. Its efficacy in accurately diagnosing diseases is demonstrated by its precise identification of disorders such as COVID-19, pneumonia, and normal lung states. Because of its accuracy, this model is a useful tool for improving patient care and medical diagnostics by aiding in the early detection and classification of a variety of lung disorders.

Sathwik Mangu et.al [3] discussed the importance of early respiratory disease diagnosis in the research, along with a unique approach that Using DS-CNNs, or depthwise separable convolutional neural networks to achieve an astounding 92% accuracy. The model is able to identify a number of chronic respiratory disorders by using spectral pictures and Mel Frequency Cepstral Coefficients (MFCCs). High precision, recall, and f1-score are shown in the results under various situations, highlighting the model's reliable performance. The study emphasizes how crucial it is to improve predictions for specific illnesses and how the suggested DS-CNN model has the ability to completely transform the identification of respiratory ailments.

Hazra Reetodeep et.al. [4] The primary objective of the proposed study was the automated analysis of respiratory sounds using convolutional neural networks (CNN). Although lung sounds like wheezes and crackles have been recognized in the past, the suggested method's accuracy in classifying respiratory diseases was higher than 92.39%. A range of normalization techniques were employed in the model design, in addition to convolutional, pooling, and fully connected layers. Performance analysis was conducted using a confusion matrix, and 500 training epochs were completed in a duration of 95 minutes. Compared to previous studies, our technique performed better and demonstrated potential for automated COPD detection in clinical settings.

Anupama H.S et.al [5] suggested methodology seeks to record lung sounds in real-time using digital stethoscopes in order to assist with the early determination of the respiratory conditions. Mel spectrograms and MFCC are extracted by the system by utilizing Audacity to convert raw audio files to the.wav format. After that, a Convolutional Neural Network (CNN) model is given these features in order to classify the data. The model architecture consists of layers for Adam optimization, ReLu and SoftMax activation functions, and feature extraction. Subsequent efforts will focus on optimizing noise cancellation for resilience, adding more audio characteristics, and adjusting hyper parameters. The ultimate objective is to create an intuitive application that streamlines the procedure and helps pulmonologists diagnose and treat patients more effectively. The suggested methodology seeks to record lung sounds in real-time using digital stethoscopes in order to diagnose respiratory diseases early.

Gaetan Chambres et.al[6] In order to predict respiratory cycle noises, such as crackles and wheezes, this study uses a machine learning model with boosted decisional trees. With an astounding 85% accuracy rate, a novel patient-level model has been built to distinguish between sick and healthy individuals. Classifying individual breathing cycles based on adventitious sound presence highlights the importance of sound analysis in the detection of heart and lung disorders. This method guarantees that ill patients are accurately identified, even if their cycles are primarily normal.

Ines Chouat et.al[7] suggested approach uses two CNN models—Mobile Net and VGGNet-16— to recognize and classify pneumonia in CT scans of the chest. These models were created via transfer learning, which made use of pretrained networks and five specially designed layers for feature extraction. The model design includes fully connected layers, convolution, pooling, and flattening. Photos were scaled to 224x224 pixels and pixel values were standardized between 0 and 1 using data normalization methodologies to enhance model training. To expand the amount of the dataset and improve model generalization, data augmentation was applied. In comparison to VGGNet-16, Mobile Net performed better, obtaining 82% accuracy, 83.5% recall, and an 82.5% F1 score. All things considered, the suggested model demonstrates how well deep learning algorithms can identify and categorize lung conditions from X-ray pictures.

Settipalli Naga Pranavi et.al[8] suggested approach seeks to use convolutional neural networks (CNNs) for the efficient and economical identification of lung conditions from chest CT and X-ray pictures, such as lung cancer and COVID-19. CNNs are capable of accurately classifying and analyzing these photos because of the large number of available data sets. The model helps physicians diagnose patients more quickly by taking pictures and processing the data to assess whether the lungs are healthy, have lung cancer, or are impacted by COVID-19. CNNs provide faster results and higher accuracy as compared to previous approaches. To improve disease diagnosis, the system makes use of frequency sub-bands from input photos and deep learning algorithms. Hospitals can benefit from this method since it is not only very accurate and simple to use, but it is also quite affordable.

3 Methodology

One of the main goals of the suggested methodology is to offer an automated algorithmic way for classifying lung sounds in various pathological stages. Proposing a lightweight deep learning architecture that can properly classify lung sounds while minimizing parameter sizes and computational complexity is another goal of the current research project.

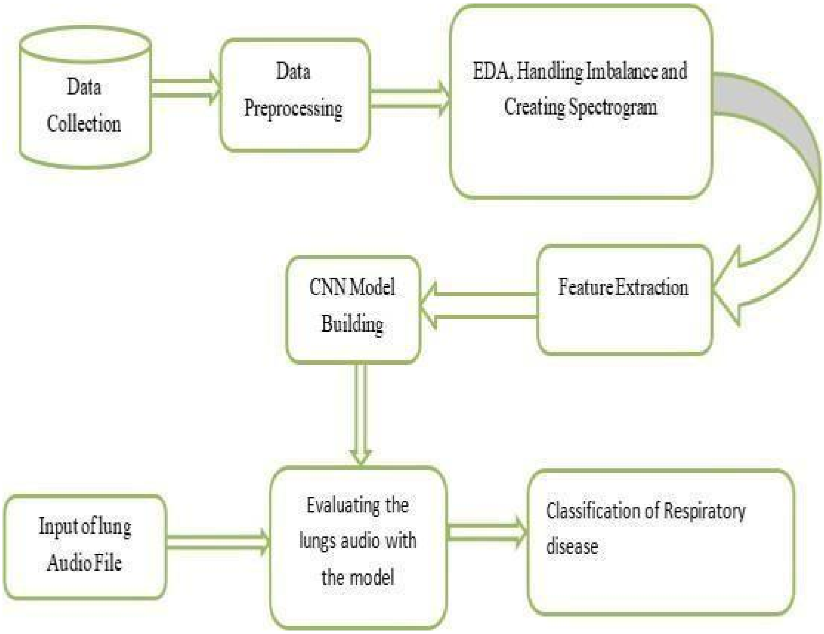

Fig. 1. Proposed System Architecture

The architecture of the suggested methodology, which we used to create a model to identify respiratory disorders, is shown in fig. 1 above.

Step 1: Data Collection: For respiratory the initial stage in illnessdetection is to collect data from the Respiratory Disease Classification Dataset, which was obtained via Kaggle. Even though we looked into additional datasets, following manualinspection, a large number of them had inadequate quality, missing important properties, or had extraneous information.

Step 2.Data preprocessing: The preprocessing steps performed onthe audio files can be summarized as follows:

- Loading the Audio Files
- Truncating the Audio
- Handling Multiple Cycles for the SamePatient
- Padding
- Saving the Processed Audio Files

By ensuring that the audio files have a consistent length and format, these pre treatment operations prepare them foradditional analysis and modeling work.

Step 3: Handling Imbalance and Creating Spectrogram &Feature Extraction: Analyzing patient audio recordings and demographic data is part of the entire procedure. First, demographic information is examined, including age, sex, andBMI. Next, gender distribution and disease frequencies are visualized. Libraries such as Librosa are used to load and process audio files, extracting features such as MFCCs for every file. To handle the data for machine learning operations,the files are sorted by patient ID and paired with the accompanying disease labels. To ensure that the distribution of disease classes is proportionate, the dataset is then split intotraining and validation sets. Ultimately, the audio files' features are taken out, encoded, and stored for training and validation. This creates the foundation for all upcoming machine learning operations. we have divided the dataset using an 80:20 proportion, meaning that 80% of the data is employed for training and 20% is used for testing.

Step 4: Model Bulding : The model is built using the training dataset with convolutional neural network. the model is described below:

i. Input Layers:

The input layers determine the structure of the input data and act as the model's foundation. The model accepts three different input layers that are tailored to different audio representations: MFCCs, Chroma features, and Mel- spectrogram features. Each input layer's shape corresponds to the dimensions of the corresponding feature data, which for MFCCs are (20, 259, 1), capturing coefficients, time steps, and channels. The model can handle many audio cues efficiently thanks to this customized technique, which is important for jobs like respiratory illness categorization

ii. Feature Extraction Networks:

a. MFCC Model: Mel-Frequency Cepstral Coefficients, or MFCCs, are a common tool for audio signal processing applications. There are several convolutional layers in the MFCC model. These layers extract spatial patterns at various scales by convolving over the input MFCC data. After every convolutional layer, batch normalization is used to speed up and stabilize training. ReLU activation functions give the network non-linearity, which makes it easier to identify intricate patterns in the MFCC data. The most notable features are retained when the spatial dimensions are down-sampled using max-pooling layers.

b. Chroma Model: Chroma characteristics are useful for tasks like chord recognition and music genre categorization because they show the energy distribution across various pitch classes. To extract hierarchical patterns from the Chroma data, Max-pooling layers, batch normalization, ReLU activation functions, and convolutional layers are used in this chroma model., just like the MFCC model.

c. Mel-spectrogram Model: Mel-spectrograms give audio signals a thorough spectro temporal representation. The architecture of the Mel-spectrogram model, which consists of convolutional layers activation functions for ReLU, max- pooling layers, batch normalization, is comparable to that of the MFCC and Chroma models. Together, these layers enable the input Mel-spectrogram data to be transformed into spectro temporal characteristics.

iii. Combine Network:

To create a single representation of the audio data, the outputs from the three feature extraction networks the MFCC, Chroma, and Mel-spectrogram models are combined as shown in the fig-2. By combining data from several audio feature categories, this combined representation enables the model to capture a variety of input audio signal characteristics. After concatenation, dropout layers are added to randomly deactivate a portion of the neurons during training. Regularization encourages the model to acquire more robust and generalizable characteristics, which helps prevent over fitting.

iv. Dense Layers:

A number of dense (completely linked) layers are passed through with the concatenated features. These tiers carry out additional processing on the amalgamated characteristics and derive elevated representations. By adding non-linearity to the thick layers, ReLU activation functions allow the model to learn intricate mappings between the input features and the output classes. Between thick layers, additional dropout layers are introduced to further regularize the model and reduce over fitting. Dropout layers enable the network to acquire redundant representations and keep it from becoming overly dependent on particular attributes by randomly dropping out connections during training.

v. Output Layer:

The output layer in this instance has eight neurons, each of which represents one of the eight potential kinds of respiratory illnesses. Soft-max activation function is applied in the output layer, converting the raw output scores into likelihoods for each class.

vi. Model Compilation:

The model is compiled with particular configurations prior to training. For optimization, the Nadam optimizer a variation of the Adam optimizer is selected. Because it works well for In multi-class classification problems, the sparse categorical cross-entropy loss function is utilized when the target labels consist of integers. Call backs like Reduce LROn Plateau and Early Stopping are used to monitor the training process and dynamically modify the learning rate. Reduce LROnPlateau lowers The rate at which learning occurs in the event that the reduction in validation plateaus, and Early Stopping terminates training in the event that the validation loss no longer improves after a predefined number of epochs.

vii. Training:

The fit approach takes the training data and their related labels and uses them to train the model. The model uses back propagation to iteratively update its parameters until it learns to match the relevant output classes with the input feature mapping. To evaluate the effectiveness of the model using unseen data and avoid over fitting, validation data is supplied. A set number of epochs are allocated for the training process to continue.

viii. Evalution:

Over the course of epochs, training and validation accuracy are tracked to evaluate the model's learning progress and identify any possible over fitting. It is easier to evaluate how successfully the model is learning and whether any adjustments must be made to improve its performance when the training and validation accuracy are visualized.

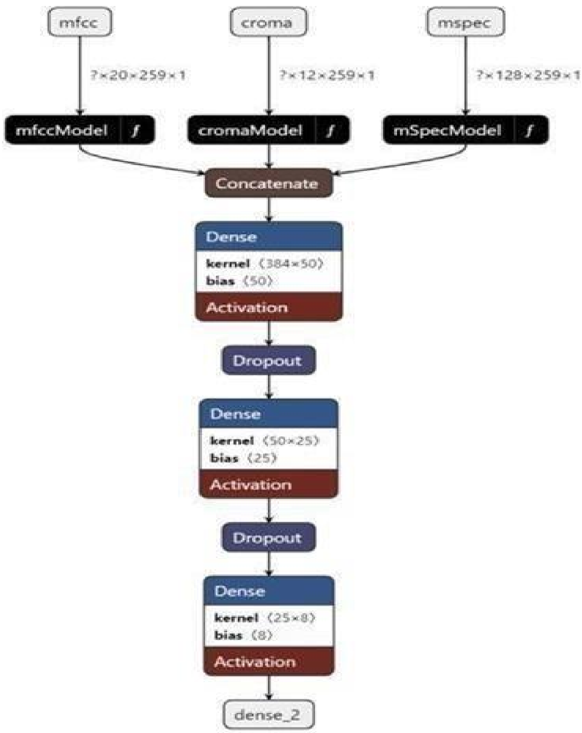

Fig. 2. The combined CNN architecture of the Model

Step 5: Evaluating the lungs audio with the model:

The model accurately predicts respiratory illnesses from the audio data, having been trained on audio variables such as MFCC, Chroma, and Mel-spectrogram. It correctly detects the most likely respiratory problems along with their associated probabilities by processing the audio file, extracting features, and putting them into the model. This gives important insights into potential medical conditions.

Step 6: Classification:

The model classifies the condition based on the percentage of the disease in our lungs after analysing the audio recording which is illustrated in fig-7. Seven distinct respiratory illnesses can be predicted and categorized by the model.

4 Results and Discussions

The fig-3 illustrate the dataset used in the model building.

	start	end	crackles	wheezes	pid	mode	filename	disease
0	0.036	0.579	0	0	101	sc	101_1b1_Al_sc_Meditron	URTI
1	0.579	2.450	0	0	101	sc	101_1b1_Al_sc_Meditron	URTI
2	2.450	3.893	0	0	101	sc	101_1b1_Al_sc_Meditron	URTI
3	3.893	5.793	0	0	101	sc	101_1b1_Al_sc_Meditron	URTI
4	5.793	7.521	0	0	101	sc	101_1b1_Al_sc_Meditron	URTI

Fig. 3. Dataset after merging the patient data and patient lungsound

The Fig-4 illustrate the overall training accuracy and validation accuracy obtained by the CNN model.

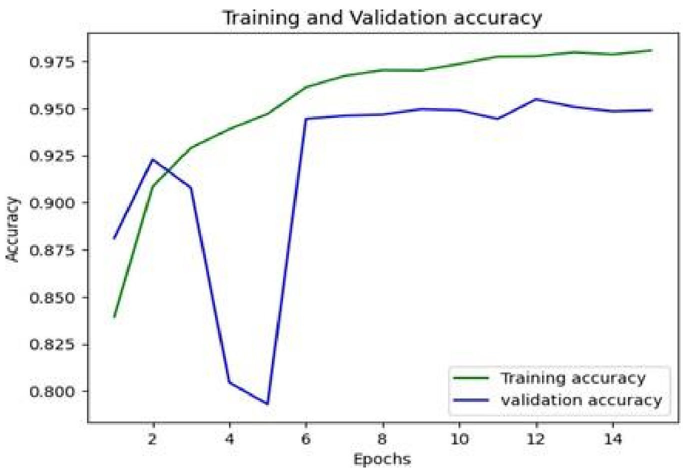

Fig. 4. Acurracy graph of the model

The fig-5 illustrate the training loss and the validation loss ofthe CNN model

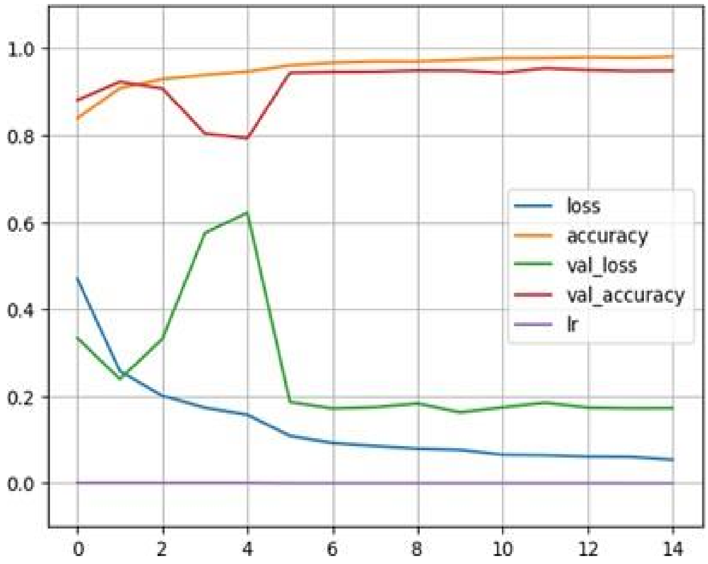

Fig. 5. Loss graph of the model

The Fig-6 describes the classification report of the CNN model.

Classification Report

	precision	recall	f1-score	support
Asthma	0.00	0.00	0.00	1
Bronchiectasis	1.00	0.96	0.98	26
Bronchiolitis	0.67	0.72	0.70	40
COPD	0.99	0.99	0.99	1437
Healthy	0.79	0.77	0.78	81
LRTI	0.00	0.00	0.00	8
Pneumonia	0.83	0.90	0.86	71
URTI	0.59	0.61	0.60	61
accuracy			0.95	1725
macro avg	0.61	0.62	0.61	1725
weighted avg	0.95	0.95	0.95	1725

Fig. 6. Classification Report

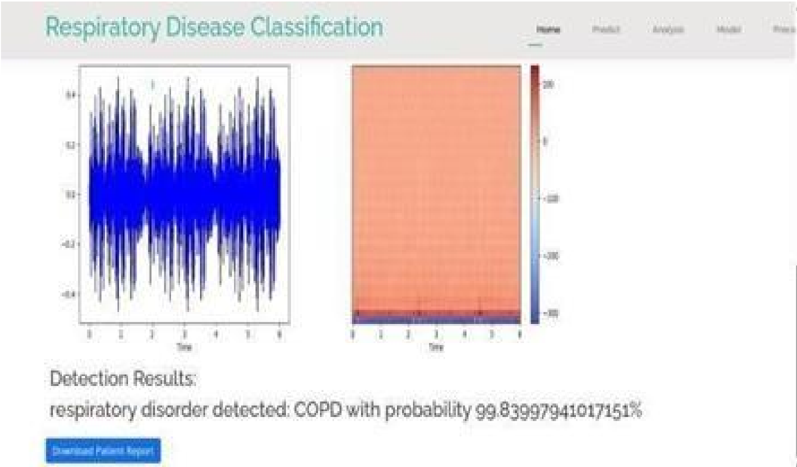

Fig. 7. Prediction Result

5 Conclusion and FurtherWork

Respiratory noises are difficult to identify based on the acoustic content. In our study, a method for differentiating between healthy and unhealthy lung sounds is described. This is a crucial step before determining the kind and severity of infection. We used a publicly accessible dataset for our trials, and the results show that the maximum accuracy of 95% outperforms the earlier models, which had accuracy of 92% and 83%.One potential way to get around the drawbacks of traditional auscultation is by automated adventitious sound detection or classification. Future research will focus on the following topic:To employ a bigger dataset and carry out more robustness tests when there is a higher percentage of noise.Attempts will also be made to separate breath sounds from heartbeat sounds and background noise in order to enable a more comprehensive investigation.To replicate lung sounds even more precisely, further acoustic approaches will be employed in addition to deep learning-based methods.To serve as a clinical tool for screening for pulmonary health and for distin- guishing between various lung disorders.Lastly, we will try to identify the kind and intensity of the illness based on the breath sounds.

References

1. "Respiratory Disease Detection through Spectogram Analysis with Explainable Deep Learning" , Fabio Martinelli, Mario Cesarelli, and Francesco Mercaldo, 2023, 8th International Conference on Smart and Sustainable Technologies (SpliTech),June 20-23, 2023. DOI: 10.23919/SpliTech58164.2023.10193020.
2. "Lung Disease Detection and Classification with Deep Learning Approach" Araya Chatchaiwatkul, Pasuk Phonsuphee, and Yurananatual Mangalmurti,2021 36th International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC), June 27-30, 2021. DOI: 10.1109/ITC- CSCC52171.2021.9501445.
3. "Respiratory Disease Detection using Depthwise Separable Convolutional Neural Networks",Sathwik Mangu, Raghu Indrakanti, and Srimanth Tenneti, 2022 ,Second International Conference on Next Generation Intelligent Systems (ICNGIS), July 29-31, 2022. DOI: 10.1109/ICNGIS54955.2022.10079793.
4. "Detecting Respiratory Diseases from Recorded Lung Sounds by 2D CNN" ,Reetodeep Hazra and Sudhan Majhi, 2020, 5th International Conference on Computing, Communication and Security (ICCCS), October 14-16, 2020. DOI: 10.1109/ICCCS49678.2020.9277101.
5. Madhavi, K. Reddy, K. Suneetha, K. Srujan Raju, Padmavathi Kora, Gudavalli Madhavi, and Suresh Kallam. "Detection of COVID 19 using X-ray Images with Fine-tuned Transfer Learning." Journal of Scientific and Industrial Research (2023): 241-248.
6. "Automatic Detection of Patient with Respiratory Diseases Using Lung Sound Analysis" ,Gaetan Chambres, Pierre Hanna, and Myriam Desainte-Catherine, 2018, International Conference on Content-Based Multimedia Indexing (CBMI), 04-06 September 2018. DOI: 10.1109/CBMI.2018.8516489.
7. "Lung Disease Detection in Chest X-ray Images Using Transfer Learning" , Ines Chouat, Amira Echtioui, Rafik Khemakhem, Wassim Zouch, Mohamed Ghorbel, and Ahmed Ben Hamida, 2022, 6th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP), 24-27 May 2022. DOI: 10.1109/ATSIP55956.2022.9805892.
8. "Lung Disease Detection using Deep Learning" ,Settipalli Naga Pranavi, Vandavasi Sreedevi, and K.V. Karthikeyan, 2023, International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI), 25-26 May 2023. DOI: 10.1109/ACCAI58221.2023.10199176.
9. "Lung Disease Detection using Deep Learning" , Syed Krar Haider Bukhari and Labiba Fahad,2022 ,17th International Conference on Emerging Technologies (ICET), November 29-30, 2022. DOI: 10.1109/ICET56601.2022.10004651.
10. "Diagnosis of Pneumonia from Chest X-ray Images using Deep Learning" , E. Ayan and H. M. Ünver, 2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT), pp. 1-5, 2019.
11. "Inter-observer Variation in the Interpretation of Chest Radiographs for Pneumonia in Community-Acquired Lower Respiratory Tract Infections",R. Hopstaken, T. Witbraad, J. Van Engelshoven, and G. Dinant, Clinical Radiology, vol. 59, no. 8, pp. 743-752, 2004.
12. Setio, A.A.A., Traverso, A., de Bel, T., Berens, M.S.N., van den Bogaard, C., Cerello, P., Chen, H., Dou, Q., Fantacci, M.E., Geurts, B., et al. (2017). "Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: the LUNA16 challenge," Med Image Anal, 42, pp. 1-13.

13. "Computerized Lung Sound Analysis as Diagnostic Aid for the Detection of Abnormal Lung Sounds: A Systematic Review and Meta- Analysis",Gurung A., Scrafford C.G., Tielsch J.M., Levine O.S., and Checkley W., Respiratory Medicinein 2011;23:1396–1403. DOI:10.1016/j.rmed.2011.05.007.
14. "Detection of Adventitious Lung Sounds using Entropy Features and a 2- D Threshold Setting", Liu X., Ser W., Zhang J., and Goh D.Y.T., Proceedings of the 2015 10th International Conference on Information, Communications and Signal Processing (ICICS), Singapore on 2–4 December 2015.
15. K.Venkateswara Rao, ―Prediction of Multiple Diseases Using Machine Learning & API‖, Journal of Engineering Sciences (JES), ISSN: 0377-9254, Vol-13, Issue-5, May 2022, Page No: 452-459.
16. K.Venkateswara Rao, ―Parkinson’s Disease Prediction and Classification Using Machine Learning Techniques‖, Journal of Optoelectronics Laser, ISSN : 1005-0086, Vol-41, Issue-5, 2022, Page No: 690-698.
17. K.Venkateswara Rao, ‖Prediction of Dengue Disease Cases by ML Techniques‖, International Journal of Data Science and Machine Learning (IJDSML), ISSN : 2692-5141, Vol-1 Issue-1, Sep 2020, Page No: 1-6.
18. K.Venkateswara Rao, ‖Disease Prediction and Diagnosis Implementing Fuzzy Neural Classifier based on IoT and Cloud‖, International Journal of Advanced Science and Technology (IJAST), ISSN : 2005-4238, Vol-29 Issue-5, May 2020, Page No: 737-745.