IoMT-Enabled Noninvasive Lungs Disease Detection and Classification Using Deep Learning-Based Analysis of Lungs Sounds

Muhammad Sajid[1], Wareesa Sharif[2], Ghulam Gilanie[3], Maryam Mazher[4],
Khurshid Iqbal[5], Muhammad Afzaal Akhtar[6], Muhammad Muddassar[7], Abdul Rehman[8]
Faculty of Computing, The Islamia University of Bahawalpur, Bahawalpur, Pakistan[1, 2, 3, 6, 8]
Department of Computer Science, COMSATS University Islamabad, Islamabad, Pakistan[4]
Department of Computer Science, Virtual University of Pakistan, Islamabad, Pakistan[5, 7]

***Abstract*—Noninvasive and accurate methods for diagnosing respiratory diseases are essential to improving healthcare consequences. The Internet of Medical Things (IoMT) is critical in driving developments in this field. This work presents an IoMT-enabled approach for lung disease detection and classification, using deep learning techniques to analyze lung sounds. The proposed approach uses three datasets: the Respiratory Sound, the Coronahack Respiratory Sound, and the Coswara Sound. Traditional machine learning models, including the Extra Tree Classifier and AdaBoost Classifier, are used to benchmark performance. The Extra Tree Classifier achieved 94.12%, 95.23%, and 94.21% across the datasets, while the AdaBoost Classifier showed improvements with 95.42%, 96.33%, and 94.76%. The proposed deep neural network (DNN) achieves accuracies of 98.92%, 99.33%, and 99.36% for the same datasets. This study explores the transformative potential of the Internet of Medical Things (IoMT) in augmenting diagnostic precision and advancing the field of respiratory healthcare.**

***Keywords*—Deep learning; respiratory sound; coronahack respiratory sound and coswara sound; IoMT**

I. Introduction

According to epidemiological statistics on respiratory disorders published by the World Health Organization (WHO), 210 million people worldwide suffer from chronic obstructive pulmonary disease (COPD), and 30 million people have asthma. Studies show that between 15 and 25 million people in India have asthma [1]. Physicians often use the noninvasive, low-cost lung auscultatory technique to assess the state of the lungs [2]. The noises the lungs make while air passes through them during breathing are known as lung sounds [3].

It is critical for recognizing lung diseases because it provides precise lung function data. Aberrant and accidental lung sounds are the two general classifications for lung sounds [4]. The Vesicular, Bronchial, Broncho-Vesicular, and Tracheal lung sounds are common [5]. Accidental lung sounds can be classified as constant or intermittent, depending on their duration and persistence [6]. Early identification and close observation of pneumonia are essential for adequate medical care [7]. Lung inspection is a standard clinical procedure for diagnosing respiratory disorders [8]. It involves hearing the sounds of an individual's lungs with a stethoscope. Usually, these noises are classified as either abnormal or normal [9]. The frequent unusual noises audible over characteristic lung sounds are crackles, wheezes, and squawks; they commonly exist in a lung condition [10].

Common lung sounds and cyclical patterns represent air passage during breathing. Pulmonary illnesses characterized by persistent, incomplete reversible airflow obstruction and normal breathing [11]. Auscultation using a listening instrument is only a qualitative diagnostic tool, even though it provides direct information [12]. However, the results of auscultation evaluation are inadequate due to several factors, such as inter and intra-observer inconsistency, bias errors in distinguishing fine sound structures, and frequency reduction [13]. Lung sound-based diagnosis is accurate and free of subjectivity errors due to the application of computer-based automated approaches and developments in lung sound recording techniques [14]. Computer-based lung sound assessment allows for a comprehensive assessment of lung sound features through visual representations, recording evaluations, suppression of noise contaminants, and evaluation of changes in lung sound action [15]. The sounds generated by air passing through the tracheobronchial tree are sounds produced by the respiratory system [16].

A. *Common Lungs Disease*

Asthma is a chronic respiratory condition characterized by inflammation and narrowing of the airways, leading to recurring episodes of wheezing, shortness of breath, chest tightness, and coughing [17]. It affects people of all ages but often starts in childhood and can persist into adulthood[18]. New opportunities for the early identification and categorization of lung diseases related to asthma are created by the combination of deep learning algorithms and heart sound analysis [19]. These algorithms can identify intricate patterns in cardiac sound data, which makes it possible to create precise and effective diagnostic models [20].

The respiratory ailment known as bronchitis is typified by discomfort in the bronchial tubes, the airways that supply oxygen to the lungs [21]. Two primary types of bronchitis can be distinguished: acute and chronic. Pneumonia is a dangerous respiratory system infection that affects the lungs [22]. Numerous pathogens, such as bacteria, viruses, fungi, or parasites, cause it. All ages are susceptible to pneumonia, but young children, aged people, and those with compromised immune systems are most at risk.

The symptoms of Chronic Obstructive Pulmonary Disease (COPD) frequently include wheezes and reduced breath sounds, which indicate airway narrowing and blockage. These auditory characteristics are analyzed by deep learning models to consistently identify COPD trends in IoMT-enabled, noninvasive detection. These devices capture precise auditory cues, allowing for early and accurate COPD monitoring and classification, facilitating proactive treatment and intervention.

Crackles in lung sounds, a sign of respiratory disorders like pneumonia or bronchitis, are detected using IoMT-enabled, noninvasive lung disease detection. Deep learning models analyze these acoustic patterns, providing accurate classification and automated assessment of potential lung abnormalities. Wheezes, high-pitched sounds during exhalation, indicate airway obstructions in conditions like asthma or COPD. IoMT-enabled lung disease detection uses deep learning algorithms to improve diagnostic accuracy, distinguishing between obstructive and restrictive respiratory disorders and facilitating timely medical interventions.

B. Research Objectives and Motivation

1) Develop a noninvasive system for detecting and classifying lung diseases using lung sound data collected via IoMT (Internet of Medical Things) devices.

2) Implement deep learning techniques, including an Extra Tree Classifier, an AdaBoost Classifier, and a Deep Neural Network, to classify lung disease from sound data.

3) Analyze the effectiveness of enabling monitoring, correct diagnosis, and efficient data processing for remote healthcare applications.

Motivation: Respiratory health is safety-critical for human life, as effective diagnosis and treatment are essential to prevent severe consequences. Traditional methods are invasive and inaccessible, so there is a need for advancements in this domain. This work includes developing a noninvasive IoMT-enabled using deep learning for accurate and accessible lung disease detection, ensuring diagnosis, and supporting remote healthcare solutions.

II. Literature Review

Neural network model, lowering data leakage and memory utilization. CNN-LSTM layers, self-attention layers, dropout, Fully Linked (FC), and softmax layers comprise the model using the ICBHI 2017 dataset. The purpose of hyperparameter tuning is to reduce training failure. Self-attention is an independent layer that works with LSTM and CNN models [23]. According to experimental data, the suggested CNN+LSTM+Selfattention model performs better overall in terms of accuracy score than the CNN+LSTM+Hybrid CNN+LSTM, CNN+LSTM+Simple Attention, and CNN+GRU+Selfattention models. With a score of 57.02% for the initial train-test split, the model produces more dependable results.

A DNN is developed to diagnose interstitial lung diseases (ILD) in patients with connective tissue diseases (CTD), and preprocessing methods are evaluated on various lung sound data sets. The DNN offers remarkable accuracy on high-resolution CT scans, with an F1-score and an F2-score of 97% [24]. Since screening for ILD in patients with chronic autoimmune disorders is still a work in progress, this technique serves as an enabler for the early, safe, accurate, and affordable identification of CTD-ILD.

Augmentation techniques to resolve the imbalanced dataset problem. The model, which has two LSTM layers, five convolutional blocks, and no augmentation, achieves a remarkable F1 score of 0.9887 in 91 s per training epoch. Misclassifications accounted for just 3.05% of COVID-19 data and mostly happened in typical instances [25]. While the standard class showed recall and an F1 score, the pneumonia class showed exceptional precision. Deep Residual Network (DRN) uses a fractional water cycle swarm optimizer (Fr-WCSO-based DRN) to identify lung disorders from respiratory sound waves. The Fr-WCSO is a novel design that combines the Water Cycle Algorithm and Competitive Swarm Optimizer with Fractional Calculus and Water Cycle Swarm Optimizer (WCSO). To reduce overfitting problems, the system preprocesses respiratory input sound signals, identifies relevant features, and augments data [26]. DRN training and feature selection are then carried out using the Fr-WCSO algorithm.

Hybrid Interpretable Strategies with Ensemble Techniques (HISET) for respiratory sound classification. The first approach uses a GSSR technique, and the second uses a novel Realm Revamping Sparse Representation Classification (RR-SRC) technique, the third uses Distance Metric dependent Variational Mode Decomposition (DM-VMD) with Extreme Learning Machine (ELM) classification process, the fourth uses Harris Hawks Optimization with Scaling Factor based Pliable Differential Evolution (SFPDE), and the fifth uses Gray Wolf Optimization based Support Vector Classification (GWO-SVC) and Grasshopper Optimization Algorithm (GOA) based Sparse Autoencoder for dimensionality reduction techniques [27]. The ICBHI dataset is used to analyze the results, and the best results are obtained for the 2-class classification when Manhattan distance-based VMD-ELM is used. This method reported an accuracy of 95.39% for the 3-class classification, 90.61% for the 3-class classification, and 89.27% for the 4-class classification. The classification of pulmonary sounds obtained from patients with connective tissue illnesses using deep learning techniques [28].

Features such as Wavelet Entropy (WE) and wavelet packet energy (WPE) are extracted from the LS. Various classifiers, including Support Vector Machine (SVM), Decision Tree (DT), k-nearest Neighbor (KNN), and Discriminant Analysis (DA), are employed to classify healthy, COPD, and asthma cases using WE and WPE features. The proposed algorithm achieves a notable classification accuracy of 99.3% with the Decision Tree (DT) classifier, effectively distinguishing between healthy individuals and those with asthma or COPD based on LS[29]. Future work will validate this algorithm with real-time LS data from asthmatic and COPD patients.

Section I introduces the research work, focusing on using IoMT and deep learning for noninvasive lung disease diagnosis. Section II reviews the existing literature, highlighting the limitations of conventional methods and the potential of IoMT technologies. Section III discusses the background studies, describing foundational principles and relevant advancements in IoMT and lung sound analysis. Section IV outlines the materials and methods, including data acquisition and deep learning techniques. Section V presents the results and discussions, analyzing findings. Section VI concludes the study, summarizing contributions, implications, and recommendations for future work.

III. Background Studies

A. Internet of Medical Things

The Internet of Medical Things (IoMT) refers to a network of interconnected medical devices, software applications, and healthcare systems designed to collect, transmit, and analyze patient data in real time. IoMT advanced technologies, i.e., sensors, wearable devices, remote monitoring tools, and cloud computing, to provide continuous healthcare solutions. These systems enable personalized patient care, early disease detection, and effective chronic disease management by continuously tracking vital signs and other health parameters.

IoMT enhances patient outcomes by enabling remote consultations, reducing hospital visits, and facilitating proactive treatment through real-time alerts. It also streamlines healthcare workflows by integrating data from diverse sources, improving clinical decision-making. However, IoMT faces challenges, including data security, interoperability, and compliance with regulatory standards. Despite these hurdles, IoMT represents a transformative advancement in modern healthcare, driving a shift toward precision medicine and empowering patients to actively engage in health management.

B. Extra Tree Classifier

One method of group decision-tree education is called the Extra Trees Classifier. When splitting a tree node, the Extra Trees classifier strongly randomizes the choice of features and reduces points, producing an unpruned collection of decision trees and trees. Extra trees function by combining the output of several de-correlated decision trees into a forest, from which they derive the classification result determined by the bulk of the voting technique. Compared to conventional decision trees or even random forests, this model adds more randomness by constructing multiple decision trees using arbitrary portions of characteristics and splitting at random points. This volatility reduces overfitting and improves the model's ability to generalize. Features taken from the audio recordings, such as time-domain, frequency-domain, and time-frequency domain characteristics, are fed into the Extra Tree classifier in the context of lung sound analysis. The model uses the rich and varied feature set to differentiate between various lung sounds, including wheezes, crackles, and other pathological sounds connected to illnesses like pneumonia, COPD, or asthma, as well as typical breathing.

C. AdaBoost Classifier

The AdaBoost methodology is a method for improving the performance of a model by combining weak classifiers. It involves extracting relevant features from audio recordings, such as time-domain and frequency-domain features, and training a weak classifier, typically a decision tree with a single split. Classifier kj can express an opinion, denoted by kj(xi) when their proposal is considered a training example for classifier acquisition for a given input model xi. Taking into account the issue of splitting the learning vector gathering into two classes, kj(xi) only accepts two values, such as 0 or 1, respectively, as shown in equation Sign C(xi), the sign of the linear mixture of the weighted total of the sub-classifiers' opinions, determines the classifier K's ultimate decision.

$$C(x_i) = a_1 k_1(x_i) + a_2 k_2(x_i) + a_l k_l(x_i)$$

Where weights are denoted bya_1, a_2,..., a_l And sub-classifiers by k_1, k_2,..., k_l. To generate a set of subpar learners, the adaboost technique keeps track of weights across instruction data and continually adjusts them after each weak learning cycle. The weights of training instances that the current weak learner incorrectly classifies will be increased, while the weights of training instances that are correctly classified will be decreased.

D. Deep Neural Network

The components of deep learning networks developed in the two-stage model will be clarified in the following part on artificial neural networks. Artificial Neural Networks (ANNs) are dynamic models that can adapt their internal architectures to meet specific functional requirements, making them ideal for managing nonlinear type issues. The essential parts of an ANN are the links and nodes that make up it, each with an output and input for interaction with other nodes or the surroundings. Each neuron in the network applies an activation function to introduce non-linearity through weighted connections. A labeled dataset is used to train the network, which uses a loss function to minimize the error between predicted and actual class labels during training by adjusting the weights of the connections through backpropagation. The learning process is one of the core characteristics of ANN, as they can understand the connections that define the data by adapting its connection to the information structure that makes up its surroundings. Neurons can be arranged in any topological configuration based on the kind and volume of input data. The feed towards construction is used in designing the most widely used ANN, with an input layer typically consisting of a particular amount of neurons paired. The data is sent to the secret layer or layers operating within the ANN and the output layer are created specifically to address the issue and provide the solution. Each neuron in the layer below is linked to every other neuron, with a fixed number of inputs and weights. Measurements are crucial for operating the deep neural network as they can be learned parameters.

$$C(x_i) = a_1 k_1(x_i) + a_2 k_2(x_i) + \cdots + a_l k_l(x_i)$$

Weight values are randomly initialized to be near zero but not zero before acquiring starts. The values of the data are modified to new information during learning, and this modification will aid in determining the significance of inputs. The activation function translates the weighted average from one neuron into the afterwards neuron's stimulation. Numerous mechanisms for activation are described in this research. Two factors influence the selection of rectified linear activation units in hidden layers in this work: (1) their ease of computation; and (2) the possibility of deep neural network optimization because of their linear behaviour. After receiving input from hidden layer #2, the network's output layer transforms it into a binary (zero = unhealthy or one = healthy). The following equation is a representation of the sigmoid activation function:

$$\check{y} = \frac{1}{1 + e^{-z\prime}}$$

Where y̌ the neuron's results and z is the hidden layer #2 outputs.The average error was determined for each sample using the loss function with cross entropy. Here is a representation of the cross-entropy loss function in equation.

$$H\ (y, \check{y}) = -\sum_{i=1}^{n} y_i - \mathrm{Log}(\check{y}_i)$$

Where y̌ the network's output and y is is the real value. After every single propagation forward, the neural network searches for a set of heavy objects that minimizes the variance among the expected and actual values.

IV. Material and Methods

A. Dataset Descriptions

The audio files used in the thesis came from three distinct data sets. A variety of numbers of audio files from different datasets are included to create a balanced dataset. Table I Describes three Datasets used in this work.

1) Respiratory sound: The Respiratory Sound is a collection of 920 audio recordings from two research teams in different countries. Samples gathered at the Hospital Infante D. Pedro in Aveiro, Portugal, and the ESSUA Respiratory Research and Rehabilitation Laboratory by the School of Health Sciences, University of Aveiro research team. The second research team, from the Universities of Coimbra and Aristotle University of Thessaloniki, gathered respiratory sounds at the Papanikolaou General Hospital in Thessaloniki and the General Hospital of Imathia in Greece. Most of the database consists of audio, with samples from two hospitals in Portugal and Greece. The researchers analyzed the recordings using various instruments, including stethoscopes and microphones. They found 761 recordings suitable for evaluation, and 761 audio files were added to the model dataset without additional requirements [30]. The database includes 6898 breathing cycles from 126 patients, with 1864 having crackles, 886 having wheezes, and 506 having both.

2) Coronahack respiratory sound: The Coronahack Respiratory Sound includes respiration sound files from both COVID-19-affected and non-affected users. The file Corona-Hack-Respiratory-Sound-Metadata.csv includes additional disturbances and demographic data about the user. Audio recordings of patients with asthma or pneumonia were included in the dataset. Respiratory sound recordings from people with asthma or pneumonia were carefully selected from records that did not indicate probable COVID-19. The dataset contained these recordings [31].

3) Coswara sound: Coswara aims to develop a cost-effective method for diagnosing COVID-19 using speech, cough, and breath sounds. The study focuses on respiratory distress, a common symptom of the illness, and measures disease biomarkers in the acoustics of these noises. The project collects voice samples from healthy and sick individuals, examining nine categories of breathing, coughing, vowel phonation, and counting. Age, gender, location, current health status, and co-morbidities are collected. The dataset includes audio recordings from patients who have not contracted COVID-19 or recovered, tagged as "Asthma" or "Pneumonia." Thirty-eight healthy voice files with respiratory sound for at least 10 seconds were included in the dataset. The study includes 38 healthy, 58 asthmatic, and nine pneumonia voice recordings under specific conditions. The Coswara dataset is valuable for understanding and diagnosing COVID-19[32].

B. Preprocessing

Preprocessing is essential for IoMT-enabled noninvasive lung disease classification and detection. Fig. 1 displays the proposed methodology for this work Preprocessing procedures are necessary to prepare the data for successful categorization since lung sound recordings frequently contain noise and unpredictability due to patient movements and natural influences. Special values in proportional variables are crucial for maintaining data integrity and optimizing model performance in lung disease detection and classification from lung sound analysis. 'Nan' values, which occur when numerator and denominator are zero, are removed in the first preprocessing step. Normalization or scaling techniques can be used to handle these special cases. When the denominator is zero, special values arise, such as positive infinity (+∞) or negative infinity (-∞), as shown in Eq. (1). These extreme values can significantly impact deep learning model performance during training.

$$X = \begin{cases} \frac{N}{D} & if\ D \neq 0 \\ +\infty & \text{if } N > 0 \text{ and } D = 0 \\ -\infty & \text{if } N < 0 \text{ and } D = 0 \end{cases} \tag{1}$$

TABLE I. Three Datasets used in this Work

Dataset Name	Audio Samples	Patients	Key Features	Source
Respiratory Sound	920 recordings, 761 used	126 patients, 6898 breathing cycles (1864 crackles, 886 wheezes, 506 both)	Breathing sounds recorded using stethoscopes and microphones	[30]
Coronahack Respiratory Sound	Varied, includes both COVID-19 and non-COVID-19 patients	Includes asthmatic, pneumonia, and COVID-19-negative patients	Demographic data, respiratory conditions, and sound disturbances included	[31]
Coswara Sound	38 healthy, 58 asthmatic, 9 pneumonia	Data collected includes age, gender, health status, and co-morbidities	Focus on speech, cough, and breath sounds for diagnosing respiratory distress	[32]

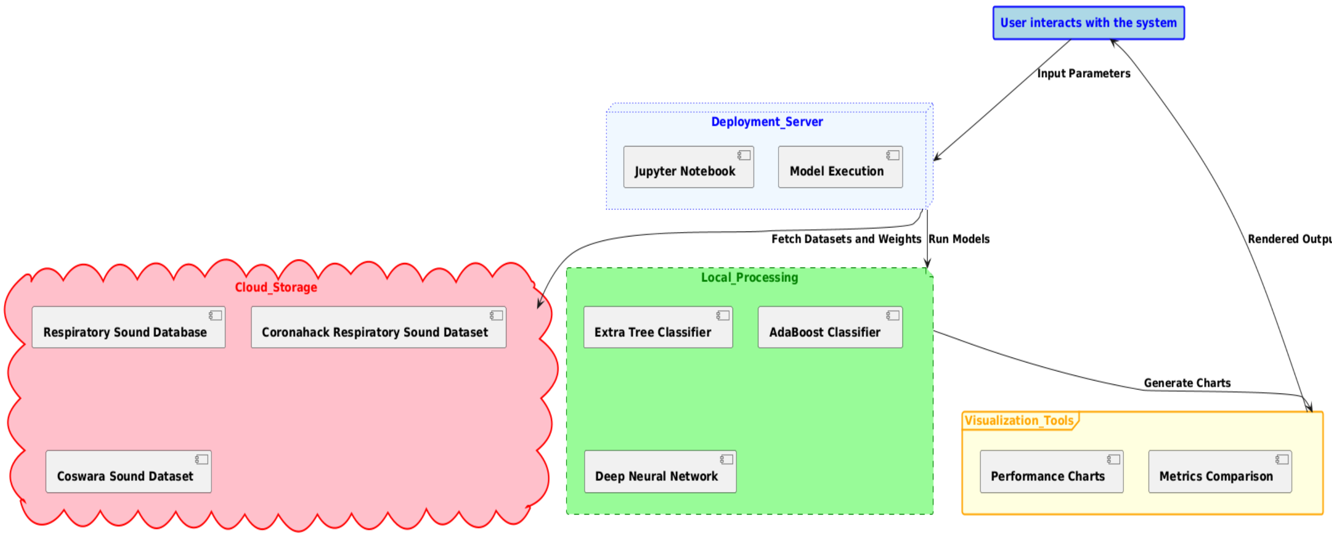

Fig. 1. Proposed methodology for this work.

The Weight of Evidence (woe) measures how well an organizing technique can discriminate between positive and negative results, or between 1 and 0. This method works for any issue where the binary variable is the target, even though it was initially created to create a predictive model for assessing credit default risk in the finance and credit sectors. The amount of evidence that either confirms or disproves a hypothesis is measured by its weight. The Weight of Evidence is determined as follows in Eq. (2):

$$WoE = \ln\left(\frac{Distribution\ of\ 1's}{Distribution\ of\ 0's}\right) * 100 \quad (2)$$

Feature elimination is crucial for optimizing model performance by removing unnecessary, redundant, or noisy features from the dataset using statistical testing, correlation analysis, and machine learning algorithms. Feature elimination techniques such as Recursive Feature Elimination (RFE) remove irrelevant or redundant features, reducing overfitting and computational load. The model's performance metric is calculated without feature removal, and the impact of eliminating features is assessed. The model's performance metric J is calculated without feature x_j, and the impact of elimination x_j is shown in Eq. (3).

$$J_{-x_j} = J\left(X\left\{x_j\right\}\right) \quad (3)$$

C. Feature Extraction

An individual measurable functionality or characteristic that defines a phenomenon is called a feature in machine learning. Useful algorithmic methods for classification rely on selecting independent, making distinctions, and useful characteristics. In this work, 1522 features from different categories were created for each sound recording as follows:

1) Time domain features: Three distinct groups are created from the signal's time series features: the audio recording's 0–1 s, 0–6 s, and 0–10 s segments. The relevant signal's increasing average series and accumulative series were calculated. The number of data points in the initial collection always equals the total moving average's term count, as shown in Eq. (4). C_k Is defined recursively as follows, where x1, x2,..., xn are the related respiration sound time series and C1, C2,..., Cn are the accumulated average with weights a series.

$$C_k = \frac{(x_k + (k-1) * x_{k-1})}{k} \quad (4)$$

2) Spectral feature: It is essential to differentiate between typical and abnormal respiratory conditions by capturing the frequency domain characteristics of lung sounds through features.

a) Time-frequency spectrogram statistical features: Mel-spectrogram, MFCC, Short-Term Fourier Transform, and Chroma In the earlier part, each respiration sound was subjected to a Short Term Fourier Transform to transform it into a time-frequency spectroscopy and extract features akin to those found in a spirometer. After computations, a long list of factors is created, which includes the statistical properties of the time-frequency spectra obtained for the 0–1, 0–6, and 0–10 s periods of each audio recording.

b) Power spectrogram statistical features: Power spectrogram statistical features provide insights into power distribution across different frequencies over time, which is essential for identifying and distinguishing various respiratory conditions, is described in Eq. (5).

$$f(x) = c_o + c_1 * x \quad (5)$$

Where c_1 Is the coefficient of the corresponding column and c_o is the constant term.

D. Feature Elimination Process

Feature elimination is crucial for optimizing model performance by removing unnecessary, redundant, or noisy features from the dataset using statistical testing, correlation analysis, and machine learning algorithms.

1) GINI Elimination: The Receiver Operating Characteristic (ROC) curve is a crucial tool in signal detection theory, used alongside the Neyman-Pearson method to visualize a classifier's efficacy. It is used for assessing and comparing the overall efficacy of testing or diagnostic procedures. The AUC index, a summary of the ROC curve, is often used in this assessment. The process is summarized as follows:

a) The training and test sets of the data set are split 80/20.

b) The single-variate regression procedure was applied to each variable's training set to determine the AUC of each variable.

c) Using the AUC values as a guide, the Gini coefficient for each variable was determined using the formula below Eq. (6):

$$GINI = (2 * AUC - 1) * 100 \quad (6)$$

E. Machine Learning Models

This section includes the details of two-stage machine/deep learning models and the working principles of applied algorithms. The Extra Tree Classifier and Ada Boost Classifier techniques with the most effective binary categorization were chosen as modeling algorithms using Python's open-source "pycaret" library.

Weight values are randomly initialized to be near zero but not zero before acquiring starts. The values of the data are modified to new information during learning, and this modification will aid in determining the significance of inputs. The activation function translates the weighted average from one neuron into the afterwards neuron's stimulation. Numerous mechanisms for activation are described in this research. Two factors influence the selection of rectified linear activation units in hidden layers in this work: (1) their ease of computation; and (2) the possibility of deep neural network optimization because of their linear behaviour. After receiving input from hidden layer #2, the network's output layer transforms it into a binary (zero = unhealthy or one = healthy). The following Eq. (7) is a representation of the sigmoid activation function:

$$\check{y} = \frac{1}{1 + e^{-z\prime}} \quad (7)$$

Where $\check{y}$ the neuron's results and z is the hidden layer #2 outputs. The average error was determined for each sample using the loss function with cross entropy. Here is a representation of the cross-entropy loss function in Eq. (8):

$$H(y, \check{y}) = -\sum_{i=1}^{n} y_i - \mathrm{Log}(\check{y}_i) \quad (8)$$

Where $\check{y}$ the network's output and y is the real value. After every single propagation forward, the neural network searches for a set of heavy objects that minimizes the variance between the expected and actual values.

Dropout is a regularization technique for deep neural networks that helps lessen learning when nerve cells are interconnected. It suggests that a subset of randomly chosen neurons from a particular layer may be removed during learning. As a result, during a specific forward or backward pass, the results of the eliminated nerve cells disappear. In the present study, every iteration saw the removal of 0.1% of the neurons in the relevant layer from the input and hidden layers.

F. Evaluation Measures

We utilize various assessment measures to evaluate the effectiveness of the proposed models. These measures provide insight into the models' accuracy, predictive power, and generalization capability. The percentage of accurately categorized cases out of all instances is known as accuracy, as shown in Eq. (9).

$$\text{Accuracy} = \frac{Number\ of\ correct\ prediction}{Total\ numer\ of\ prediction} * 100 \quad (9)$$

Other evaluation measures adopted for assessing the proposed models are specificity, sensitivity, and F1 score, shown in Eq. (10) and Eq. (11). The specificity and sensitivity formula are as follows:

$$Specificity = \frac{True\ positive}{True\ positive + False\ negative} \quad (10)$$

$$Sensitivity = \frac{True\ negative}{True\ negative + False\ positive} \quad (11)$$

True positives (TP) are positive in the test set and correctly labeled as positive by the classifier. True negatives (TN) are negative in the test set and correctly labeled as negative by the classifier. False positives (FP) are negative in the test set but incorrectly labeled as positive by the classifier. False negatives (FN) are positive in the test set but incorrectly labeled as negative by the classifier. Eq. (12) shows that the F1 score is the harmonic mean of precision and recall, providing a combined measure of precision and recall.

$$F1\ score = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (12)$$

Where precision and recall are calculated Eq. (13) and Eq. (14), respectively.

$$Precision = \frac{True\ positive}{True\ positive + False\ positive} \quad (13)$$

$$Recall = \frac{True\ positive}{True\ positive + False\ negative} \quad (14)$$

V. Discussion and Results

Spectrogram (top) and onset strength analysis (bottom) of lung sound data, critical for IoMT-enabled noninvasive lung disease detection. The spectrogram visualizes frequency (Hz) over time, with color intensity indicating sound energy. It highlights distinct acoustic patterns associated with respiratory cycles, facilitating feature extraction for disease classification. The onset strength graph below shows temporal variations in sound intensity, with detected onsets marked by red dashed lines, capturing significant events like wheezes or crackles. These features, analyzed using deep learning, improve the precision of lung sound classification, aiding in early and accurate detection of pulmonary diseases. Fig. 2 represents Spectro-Temporal Analysis for IoMT-Based Lung Disease Detection.

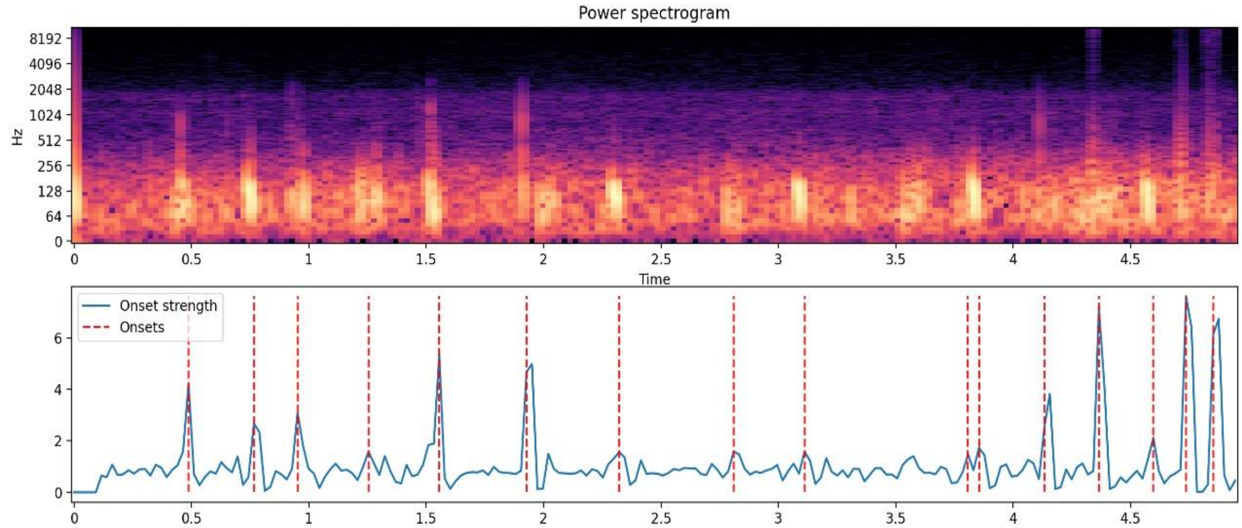

Fig. 2. Spectro-Temporal analysis for IoMT-Based lung disease detection.

Onset detection and energy analysis for lung sounds, crucial for IoMT-based noninvasive respiratory disease classification. The top graph visualizes onset strength with raw onsets (blue peaks) and backtracked onsets (red vertical lines). These onsets correspond to significant acoustic events, such as wheezes or crackles, indicative of lung abnormalities. By combining onset detection and energy analysis, these features enhance the ability of deep learning algorithms to accurately classify lung diseases, enabling early diagnosis through efficient feature extraction and temporal event mapping. This dual-layer analysis improves robustness in detecting subtle patterns in lung sounds, aiding real-time and remote healthcare applications. Fig. 3 shows Onset and Energy Analysis for Lung Sound Classification.

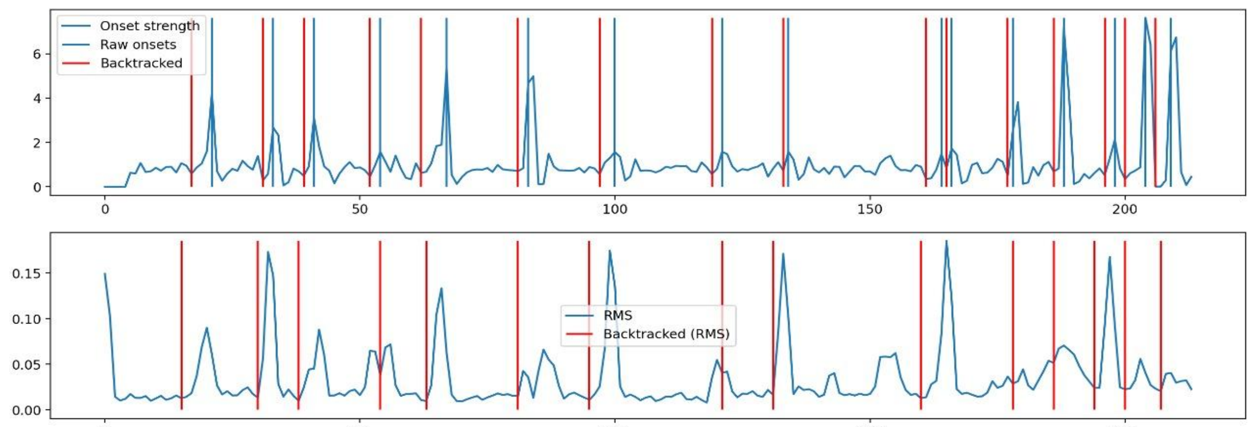

Fig. 3. Onset and energy analysis for lung sound classification.

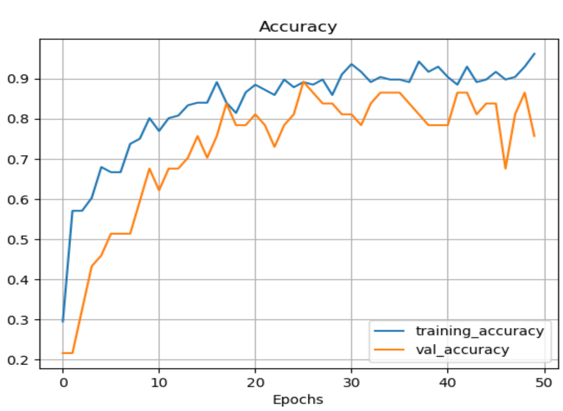

Fig. 4. Training history.

The Fig. 4 shows the training and validation accuracy trends over 50 epochs in a machine learning model. The x-axis represents epochs, and the y-axis represents accuracy values ranging from 0 to 1. The training accuracy (blue line) steadily improves, indicating the model's ability to fit the training data, with slight fluctuations towards the later epochs. The validation accuracy (orange line) increases initially, stabilizes, and exhibits minor oscillations, reflecting the model's generalization performance. The gap between training and validation accuracy suggests potential overfitting, as training accuracy surpasses validation accuracy in later epochs. This trend presents the need for optimization or regularization techniques.

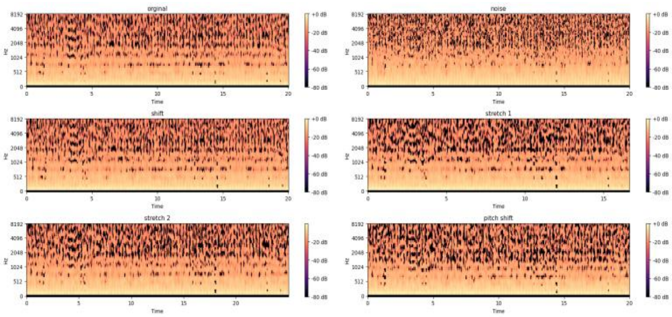

Fig. 5. Unaltered audio spectrogram.

The Fig. 5 show spectrograms visualizing feature extraction from an audio signal under different augmentations. The first plot (original) shows the unaltered spectrogram. Subsequent plots depict the effects of augmentations: noise addition, time shift, time stretching (two variations), and pitch shifting. These transformations simulate variability in audio datasets to improve machine learning model generalizability. The model can robustly learn features under varying conditions by augmenting the original signal, critical in tasks like speech recognition or environmental sound classification.

The results presents analysis of three machine learning models—AdaBoost Classifier, Extra Tree Classifier, and Deep Neural Network—evaluated on three datasets: the Respiratory Sound, Coronahack Respiratory Sound, and Coswara Sound. Each model's performance is measured using precision, recall, F1-score, and overall accuracy across diseases like crackles, wheezes, COVID-19, asthma, pneumonia, and healthy cases.

Table II, III and IV shows the performance of the models. Accuracy ranged from 94.12% to 95.23%, with consistent precision, recall, and F1-scores for all diseases, indicating robust yet moderate effectiveness. Table II displays the AdaBoost Classifier's assessment, which improved accuracy (95.42%–96.33%) and balanced precision-recall for detecting COVID-19 and other diseases, suggesting its superior predictive reliability compared to the ensemble approach. Table III evaluates a Deep Neural Network, showcasing the highest performance metrics, with accuracy surpassing 98% across all datasets. The network's precision, recall, and F1 scores consistently reached 0.99 for most diseases, demonstrating its efficacy in detecting subtle respiratory anomalies.

TABLE II. PERFORMANCE MEASURES OF EXTRA TREE CLASSIFIER

Dataset used for experiments	Diseases	Precision	Recall	F1-score	Overall Accuracy
Respiratory Sound	Crackles	0.95	0.96	0.95	**94.12%**
	Wheezes	0.93	0.94	0.96	
Coronahack respiratory sound	COVID-19	0.94	0.95	0.93	**95.23%**
	Healthy	0.96	0.93	0.92	
Coswara sound	Asthma	0.92	0.94	0.94	**94.21%**
	Pneumonia	0.94	0.93	0.94	
	Healthy	0.92	0.94	0.94	

TABLE III. PERFORMANCE MEASURES OF ADA BOOST CLASSIFIER

Dataset used for experiments	Diseases	Precision	Recall	F1-score	Accuracy
Respiratory Sound	Crackles	0.96	0.94	0.95	95.42%
	Wheezes	0.94	0.96	0.95	
Coronahack respiratory sound	COVID-19	0.97	0.96	0.97	96.33%
	Healthy	0.95	0.96	0.96	
Coswara sound	Asthma	0.96	0.95	0.94	94.76%
	Pneumonia	0.97	0.96	0.95	
	Healthy	0.95	0.95	0.96	

TABLE IV. PERFORMANCE MEASURES OF DEEP NEURAL NETWORK

Dataset used for experiments	Diseases	Precision	Recall	F1-score	Accuracy
Respiratory Sound	Crackles	0.99	0.98	0.98	98.92%
	Wheezes	0.99	0.99	0.99	
Coronahack respiratory sound	COVID-19	0.98	0.99	0.99	99.33%
	Healthy	0.99	0.99	0.99	
Coswara sound	Asthma	0.99	0.99	0.99	99.36%
	Pneumonia	0.98	0.99	0.99	
	Healthy	0.97	0.98	0.99	

VI. Conclusion and Future Work

This work presents an IoMT-based noninvasive approach to lungs disease detection and classification. The work uses an IoMT-enabled, noninvasive approach for lung disease detection and classification using Respiratory Sound, Coronahack Respiratory Sound, and Coswara Sound. Using machine learning models such as the Extra Trees classifier and AdaBoost classifier alongside a proposed deep learning model, this approach achieved impressive accuracy levels across various respiratory conditions. . The DNN achieves accuracy across all datasets, with 98.92% for the Respiratory Sound, 99.33% for the Coronahack Respiratory Sound, and 99.36% for the Coswara Sound. These results highlight the potential of deep learning models to support reliable and accurate respiratory health assessment in IoMT applications. Future work will enhance model robustness to handle diverse datasets and real-world variations and optimize the model for low-power IoMT devices to facilitate clinical deployment. Future work will optimize the proposed model for real-world applications, explore additional features such as multi-modal data for improved accuracy, and conduct large-scale evaluations across diverse network environments to assess generalizability.

Acknowledgment

The author would like to acknowledge the support of The Islamia University of Bahawalpur Pakistan, for providing the resources and guidance during the Ph.D. journey.

References

[1] Altan, G., Kutlu, Y., & Gökçen, A. (2020). Chronic obstructive pulmonary disease severity analysis using deep learning on multi-channel lung sounds. *Turkish Journal of Electrical Engineering and Computer Sciences, 28*(5), 2979-2996.

[2] Altan, G., Kutlu, Y., & Allahverdi, N. (2019). Deep learning on computerized analysis of chronic obstructive pulmonary disease. *IEEE Journal of Biomedical and Health Informatics, 24*(5), 1344-1350.

[3] Korenbaum, V., & Shiryaev, A. (2020). Features of sound conduction in human lungs in the 80–1000 Hz and 10–19 KHz frequency ranges. *Acoustical Physics, 66*, 548-558.

[4] Lee, S. H., Kim, Y.-S., Yeo, M.-K., Mahmood, M., Zavanelli, N., Chung, C., ... & Yeo, W.-H. (2022). Fully portable continuous real-time auscultation with a soft wearable stethoscope designed for automated disease diagnosis. *Science Advances, 8*(21), eabo5867.

[5] Tariq, Z., Shah, S. K., & Lee, Y. (2019). Lung disease classification using deep convolutional neural network. Paper presented at the *2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*.

[6] Au, Y. K., Muqeem, T., Fauveau, V. J., Cardenas, J. A., Geris, B. S., Hassen, G. W., & Glass, M. (2022). Continuous monitoring versus intermittent auscultation of wheezes in patients presenting with acute respiratory distress. *The Journal of Emergency Medicine, 63*(4), 582-591.

[7] Tong, L., Huang, S., Zheng, C., Zhang, Y., & Chen, Z. (2022). Refractory Mycoplasma pneumoniae pneumonia in children: Early recognition and management. *Journal of Clinical Medicine, 11*(10), 2824.

[8] Mukhopadhyay, S., Mehrad, M., Dammert, P., Arrossi, A. V., Sarda, R., Brenner, D. S., ... & Ghobrial, M. (2020). Lung biopsy findings in severe pulmonary illness associated with E-cigarette use (vaping): A report of eight cases. *American Journal of Clinical Pathology, 153*(1), 30-39.

[9] Rajkumar, S., Sathesh, K., & Goyal, N. K. (2020). Neural network-based design and evaluation of performance metrics using adaptive line enhancer with adaptive algorithms for auscultation analysis. *Neural Computing and Applications, 32*, 15131-15153.

[10] Sathesh, K., Rajkumar, S., & Goyal, N. K. (2020). Least Mean Square (LMS)-based neural design and metric evaluation for auscultation signal separation. *Biomedical Signal Processing and Control, 59*, 101784.

[11] Sanchez-Perez, J. A., Berkebile, J. A., Nevius, B. N., Ozmen, G. C., Nichols, C. J., Ganti, V. G., ... & Wright, D. W. (2022). A wearable multimodal sensing system for tracking changes in pulmonary fluid status, lung sounds, and respiratory markers. *Sensors, 22*(3), 1130.

[12] Ferreira, H. d. M. G. (2021). Pulmonary auscultation using mobile devices: Feasibility study in respiratory diseases. *Universidade do Porto (Portugal)*.

[13] Razvadauskas, H., Vaičiukynas, E., Buškus, K., Arlauskas, L., Nowaczyk, S., Sadauskas, S., & Naudžiūnas, A. (2024). Exploring classical machine learning for identification of pathological lung auscultations. *Computers in Biology and Medicine, 168*, 107784.

[14] Sfayyih, A. H., Sulaiman, N., & Sabry, A. H. (2023). A review on lung disease recognition by acoustic signal analysis with deep learning networks. *Journal of Big Data, 10*(1), 101.

[15] Hsu, F.-S., Huang, S.-R., Huang, C.-W., Huang, C.-J., Cheng, Y.-R., Chen, C.-C., ... & Lai, Y.-C. (2021). Benchmarking of eight recurrent neural network variants for breath phase and adventitious sound detection on a self-developed open-access lung sound database—HF_Lung_V1. *PLOS One, 16*(7), e0254134.

[16] Waddingham, P. H., Mangual, J. O., Orini, M., Badie, N., Muthumala, A., Sporton, S., & Chow, A. W. (2023). Electrocardiographic imaging demonstrates electrical synchrony improvement by dynamic atrioventricular delays in patients with left bundle branch block and preserved atrioventricular conduction. *Europace, 25*(2), 536-545.

[17] Koefoed, H. J. L., Zwitserloot, A. M., Vonk, J. M., & Koppelman, G. H. (2021). Asthma, bronchial hyperresponsiveness, allergy, and lung function development until early adulthood: A systematic literature review. *Pediatric Allergy and Immunology, 32*(6), 1238-1254.

[18] Chang, A. B., Oppenheimer, J. J., Irwin, R. S., Adams, T. M., Altman, K. W., Azoulay, E., ... & Boulet, L.-P. (2020). Managing chronic cough as a symptom in children and management algorithms: CHEST guideline and expert panel report. *Chest, 158*(1), 303-329.

[19] 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*, e369.

[20] Kumar, N., & Kumar, D. (2021). Machine learning-based heart disease diagnosis using noninvasive methods: A review. Paper presented at the *Journal of Physics: Conference Series*.

[21] Deshmukh, R., Bandyopadhyay, N., Abed, S. N., Bandopadhyay, S., Pal, Y., & Deb, P. K. (2020). Strategies for pulmonary delivery of drugs. In *Drug Delivery Systems* (pp. 85-129): Elsevier.

[22] Oliva, J., & Terrier, O. (2021). Viral and bacterial co-infections in the lungs: Dangerous liaisons. *Viruses, 13*(9), 1725.

[23] Bhushan, P., Fahad, M. S., Agrawal, S., Kamesh, K. S. D., Tripathi, P., Mishra, P., ... Deepak, A. (2024). A Self-Attention Based Hybrid CNN-LSTM Architecture for Respiratory Sound Classification. *GMSARN International Journal, 18*(1), 54-61.

[24] Fava, A., Dianat, B., Bertacchini, A., Manfredi, A., Sebastiani, M., Modena, M., & Pancaldi, F. (2024). Preprocessing techniques to enhance the classification of lung sounds based on deep learning. *Biomedical Signal Processing and Control, 92*, 106009.

[25] Fachrel, J., Pravitasari, A. A., Yulita, I. N., Ardhisasmita, M. N., & Indrayatna, F. (2023). Enhancing an imbalanced lung disease x-ray image classification with the CNN-LSTM model. *Applied Sciences, 13*(14), 8227.

[26] Dar, J. A., Srivastava, K. K., & Mishra, A. (2023). Lung anomaly detection from respiratory sound database (sound signals). *Computers in Biology and Medicine, 164*, 107311.

[27] Prabhakar, S. K., & Won, D.-O. (2023). HISET: Hybrid interpretable strategies with ensemble techniques for respiratory sound classification. *Heliyon, 9*(8).

[28] Dianat, B., La Torraca, P., Manfredi, A., Cassone, G., Vacchi, C., Sebastiani, M., & Pancaldi, F. (2023). Classification of pulmonary sounds through deep learning for the diagnosis of interstitial lung diseases secondary to connective tissue diseases. *Computers in Biology and Medicine, 160*, 106928.

[29] Haider, N. S., & Behera, A. (2022). Computerized lung sound based classification of asthma and chronic obstructive pulmonary disease (COPD). *Biocybernetics and Biomedical Engineering, 42*(1), 42-59.

[30] Vbookshelf. (2017). *Respiratory sound database* [Dataset]. Kaggle. Retrievedfrom https://www.kaggle.com/datasets/vbookshelf/respiratory-sound-database

[31] Praveengovi. (2020). *Coronahack respiratory sound dataset* [Dataset]. Kaggle.Retrievedfrom https://www.kaggle.com/datasets/praveengovi/coronahack-respiratory-sound-dataset\

[32] Janashreeananthan. (2021). *Coswara* [Dataset]. Kaggle. Retrieved from https://www.kaggle.com/datasets/janashreeananthan/coswara