IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

MACHINE LEARNING-DRIVEN PATTERN RECOGNITION FOR RESPIRATORY DISEASE DETECTION USING SPIROMETRY DATA AND HYBRID SVM-KNN-CNN MODELS

Shemla Asharaf
Ex-Assistant Professor
Department of Computer Science and Engineering
Government Engineering College, Palakkad
Street 600, Zone 74, Al khor, Qatar

Vaishaki VK
Assistant Professor
Department of Computer Science and Engineering
St.Thomas College Of Engineering and Technology, Kannur

Dr. Amitha I C
Associate Professor(HOD)
Department of Computer Science and Engineering
St.Thomas College Of Engineering and Technology, Kannur

Abstract *:* Respiratory illnesses, including chronic obstructive pulmonary disease (COPD) and other pulmonary conditions, are among the most widespread health challenges worldwide, highlighting the urgent need for advanced diagnostic systems. Early detection of COPD is crucial for establishing timely therapeutic strategies and personalized patient care. Epidemiological data on COPD and its severe impact on patient's health emphasize the need for advanced diagnostic tools. Traditional diagnostic methods may lack the precision required to detect early-stage COPD or to distinguish it from other respiratory diseases. This research aims to develop an adaptive and precise model for analyzing pulmonary audio data along with X-ray image to facilitate early COPD diagnosis. The goal is to enhance diagnostic accuracy by leveraging advanced machine learning techniques and feature extraction methods. In this study, Mel-Frequency Cepstral Coefficients (MFCCs) are used for feature extraction from pulmonary sound recordings. To address challenges related to dimensionality and computational complexity**,** Forward Feature Selection (FFS) is applied to identify the most relevant features. The classification approach synthesizes an SVM-KNN fusion model along with CNN-based models to reveal intricate patterns and boundaries in the data, making CNN useful for respiratory diagnostics. The COPD disease dataset serves as the foundation for all modeling and testing processes. X-ray images are collected and trained using CNN models. The implemented SVM-KNN-CNN fusion model in Python demonstrated exceptional performance, achieving an accuracy of 95% or greater. This level of accuracy highlights the model's capability to effectively differentiate between healthy and COPD-affected lungs. The developed framework significantly enhances the detection of COPD and the assessment of respiratory illness risk in patients. CNN-based models can significantly enhance COPD detection from lung X-rays, offering a non-invasive, fast, and cost-effective diagnostic tool**.** However, it should complement clinical tests like spirometry and physician expertise for accurate diagnosis.

I. INTRODUCTION

As of February 2025, the World Health Organization (WHO) continues to recognize Chronic Obstructive Pulmonary Disease (COPD) as a significant global health concern. COPD remains the fourth leading cause of death worldwide, responsible for approximately 3.5 million deaths annually, accounting for about 5% of all global deaths. Chronic obstructive pulmonary disease (COPD) refers to a group of progressive lung conditions, such as emphysema and chronic bronchitis, that make breathing increasingly difficult. These conditions hinder airflow in the lungs due to damage to lung tissue and inflammation of the small

airways or bronchi. Given the widespread prevalence and severe impact of COPD on life productivity, as well as the growing strain on healthcare systems, it is crucial to develop an accurate diagnostic tool to mitigate the effects of this disease on both individuals and society. Advanced machine learning (ML) and deep learning (DL) techniques have shown significant potential to transform the current diagnostic approaches in various medical fields, including COPD [3]. Traditional diagnostic methods for COPD, such as spirometry, imaging, and other tests, often result in high rates of false negatives and false positives, or have low sensitivity and specificity, which hinder early diagnosis and reduce treatment effectiveness. These limitations have existed since the pre-ML and DL era and can only be addressed by leveraging ML or DL methods. The key opportunity lies in the ability to apply complex algorithms to data, enhancing diagnostic accuracy.

Recent advancements in machine learning (ML) and deep learning (DL) techniques have been shown to enhance diagnostic accuracy and improve the overall efficiency of the diagnostic process. In deep learning, the model can autonomously learn from large datasets, identifying features and patterns without explicit programming. This is particularly valuable in medical imaging, where the output may not always be clear, and even subtle or ambiguous patterns can signal the presence of a disease in its early stages. In the context of COPD, extensive research has explored and explained how DL methods are applied to analyze medical images and pulmonary data, leading to improved diagnosis and disease classification [4].

Previous studies have highlighted the significant potential of machine learning (ML) and deep learning (DL) techniques in enhancing the diagnostic capabilities for COPD. Notably, models built using the COPD Gene database, which predominantly includes data from current and former smokers, have achieved an impressive area under the curve (AUC) of approximately 85.6% for identifying COPD. This success is attributed to the rich dataset, which captures various manifestations of COPD and the smoking histories of individuals. Additionally, recent advancements have introduced novel network models that specialize in identifying COPD in smokers who have undergone lung cancer screening. These methods have demonstrated even higher accuracy rates, exceeding 88%, in distinguishing COPD from other respiratory conditions, even when multiple illnesses like lung cancer are present. Overall, the results from these studies are promising and indicate substantial progress in the evaluation of COPD [5]. It is also important to address the limitations of current studies.

NEED OF THE STUDY

2.1 Incorporating new datasets

Many existing studies rely on datasets that may have limited patient diversity and may not fully reflect real-world scenarios. For example, the COPD Gene dataset primarily includes individuals with a smoking history or exposure to specific environmental factors, excluding non-smokers and those affected by other environmental influences. As a result, it is crucial for the machine learning (ML) and deep learning (DL) approaches discussed to be tested in more diverse and complex settings that better reflect real-world conditions [6]. This requires incorporating new datasets that represent a broader range of patient severities, different global regions, and various environmental conditions. For these models to be widely applicable, they must perform effectively across diverse populations, ensuring that they can provide generalizable, clinically relevant insights that improve COPD diagnosis and treatment for all patient groups.

2.2 Develop an accurate and dynamic model

The early changes in the auditory system associated with COPD, such as detection, diagnosis, and monitoring, are often difficult to identify using traditional spirometry and imaging techniques. The main goal of this research is to develop an accurate and dynamic model for analyzing pulmonary audio data to enable the early detection of COPD. By utilizing advanced ML techniques and feature extraction methods, the study aims to improve diagnostic precision and distinguish COPD from other respiratory disorders. This approach enhances patient management and treatment outcomes by improving diagnostic accuracy.

RESEARCH METHODOLOGY

The proposed methodology introduces several key innovations. First, Mel-frequency cepstral coefficients (MFCCs) are used as features in pulmonary sound recordings, which allows for capturing intricate details of the audio signals. This makes MFCCs valuable in detecting subtle patterns indicative of respiratory disorders. To enhance the model's predictive capability, we employ forward feature selection (FFS) to identify the most influential features, as not all features require significant computational resources. This approach helps reduce dimensionality and computational burden, while focusing on the features that are most likely to contribute to an accurate diagnosis in this context. In analyzing pulmonary data, we propose combining SVM, K-NN, and CNN techniques to enhance the classification results. Support Vector Machine (SVM) is particularly effective at identifying boundaries between non-linearly separable classes, making it ideal for distinguishing between complex categories. On the other hand, K-Nearest Neighbors (K-NN) excels at detecting intricate patterns and fine details within a specific class. By integrating these methods, along with CNN's ability to automatically extract hierarchical features from raw data, we aim to create a robust model capable of accurately classifying normal lungs from COPD-affected lungs.

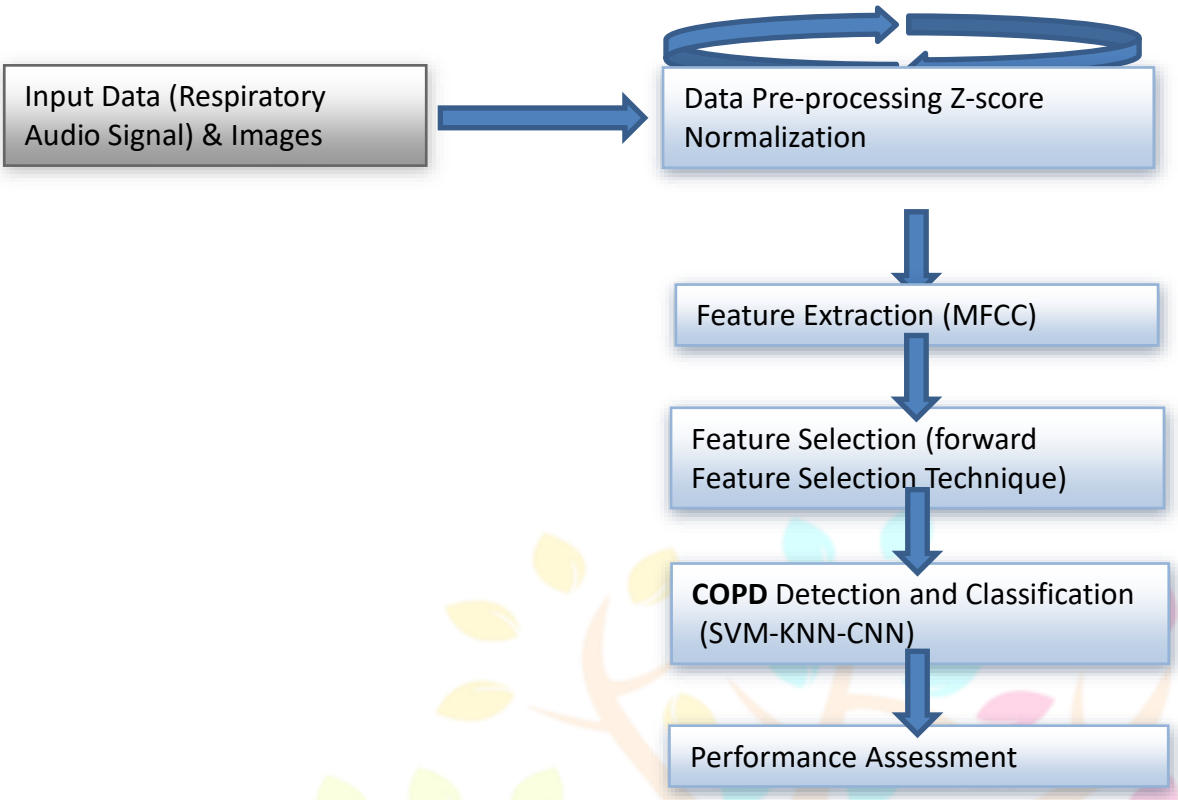

Fig.1. workflow of the suggested approach

3.1 Database assembly

The data utilized in this study was sourced from the open-access Kaggle database platform. The dataset comprises 920 annotated audio files, each lasting between 10 to 90 seconds, encompassing a diverse range of respiratory sounds. These recordings were collected from 126 individuals across various age groups, including pediatric, adult, and geriatric populations, thereby providing a representative demographic sample. Respiratory sounds are vital indicators of lung health and disease. The sounds generated during breathing are closely linked to airflow, lung tissue alterations, and secretion locations within the lungs. For instance, wheezing often signifies obstructive airway diseases such as asthma or chronic obstructive pulmonary disease (COPD). Alternative recording methods, including digital stethoscopes, can also capture these sounds. The Pulmonary Sound Database was developed by two research teams in Greece and Portugal. The database contains 920 annotated recordings, ranging from 10 to 90 seconds in duration, contributed by 126 participants. The dataset encompasses a total of 5.5 hours of recordings, comprising 6898 pulmonary cycles, of which 1864 exhibit crackles, 886 display wheezes, and 506 demonstrate both. The data includes both clear pulmonary sounds and loud recordings that simulate real-world environments.

3.1.1. Application of Z-Score normalization technique for data pre-processing

Z-score normalization is a valuable technique in audio signal analysis, ensuring that amplitude values are standardized around a mean of 0 and exhibit a standard deviation of approximately 1. This normalization is especially advantageous for audio sources with diverse scales and amplitudes, facilitating more accurate comparisons. Initially, the average (μ) and standard deviation (σ) of the amplitude values across the entire audio stream are calculated. Subsequently, z-score normalization is applied to standardize each amplitude value u_t in the audio signal at a specific time (t) was expressed in Eq(1).

$$Z_t = \frac{u_t - \mu}{\sigma} \quad (1)$$

Where: − Z_t represents the normalized amplitude value at time(t).
− u_t denotes the original amplitude value at time(t).
− (μ) is the mean of the audio signal′s amplitude values.
− (σ) is the standard deviation of the audio signal′s amplitude values.

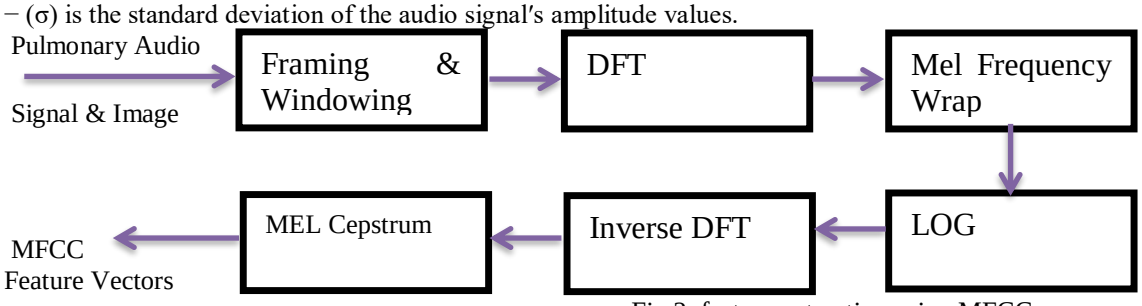

Fig.2. feature extraction using MFCC

The amplitude values are normalized through this equation, rendering them directly comparable and streamlining various audio processing tasks. By applying Z-score normalization, the analysis and modeling of audio data are facilitated, as the amplitude values of the audio signal are standardized to have a mean of zero and a standard deviation of approximately one, thereby enhancing their uniformity and consistency.

3.2. Employing Mel-frequency Cepstral coefficients to derive features

The Mel-Frequency Cepstral Coefficients (MFCC) technique is employed for feature extraction due to its unique ability to assess the spectral characteristics of sound with high precision, particularly in the context of respiratory sound analysis for COPD diagnosis. MFCCs are renowned for their efficacy in capturing brief power spectra and mimicking human auditory perception. Through a series of processing steps, including frame modulation, pre-emphasis, windowing, Fast Fourier Transform (FFT), and Mel filter bank application, the signal's frequency content can be accurately quantified. Subsequent logarithmic scaling and Discrete Cosine Transform (DCT) enable the de correlation of multiple coefficients, yielding a compact set of features [18]. This approach is particularly adept at detecting subtly nuanced audio patterns pertinent to respiratory health, leveraging solely the lower-order MFCC coefficients, which encapsulate essential information crucial to this analysis. The decision to employ MFCCs is guided by the technique's proven efficacy in diverse audio processing applications, especially in medical diagnostics, where the capture of acoustic features demands precision. Incorporating a diagram to supplement the explanation would be beneficial, illustrating the MFCC feature extraction process, from framing to FFT, Mel filtering, and subsequent stages, as well as its application in screening respiratory diseases, particularly Chronic Obstructive Pulmonary Disease (COPD)[19].

Mel-frequency Cepstral coefficients (MFCCs) are a valuable tool for identifying distinctive features of audio signals and provide a robust framework for characterizing the spectral properties of audio signals. MFCCs play a crucial role in precisely capturing the dynamic acoustic patterns inherent in respiratory sounds, particularly in the detection and initial evaluation of chronic obstructive pulmonary disease (COPD) and other respiratory conditions [19]. As a necessary step in extracting Mel-frequency Cepstral coefficients (MFCCs) from audio data, the audio signal must be divided into brief, overlapping segments, typically lasting 20-40 milliseconds. This fragmentation enables a temporal examination of the signal's spectral properties. A pre-emphasis filter is then applied to each segment to amplify high-frequency components, thereby diminishing the signal-to-noise ratio. Consequently, spectral leakage is minimized, and the signal is prepared for transformation via windowing techniques, such as Hanning or Hamming[21]. Ultimately, the Fast Fourier Transform (FFT) is applied to each windowed frame, facilitating the conversion of the signal from the time domain to the frequency domain, thereby enabling a comprehensive transformation. This process yields a power spectrum, which represents the amplitude of each frequency component present in the frame. The power spectrum is then processed through a Mel filter bank, which mimics the frequency resolution of the human auditory system. The overlapping triangular filters in this filter bank are non-uniformly spaced along the Mel-frequency scale, capturing the energy distribution across various frequency bands in a manner analogous to human hearing. Subsequently, a logarithmic transformation is applied, wherein the energy output of each filter is computed in a logarithmic format, simulating the logarithmic perception of loudness by humans. The final critical step involves applying a discrete cosine transform (DCT) to the logarithmic filter bank energies, ultimately yielding the cepstrum coefficients as expressed in Eq. (2) [22].

$$\mathbf{C^p = \sum_{u=1}^{T_f} G_u Cos\left[p(u = 0.5)\left(\frac{\pi}{T_f}\right)\right], p = 1, 2, 3, \ldots \ldots b} \qquad \mathbf{(2)}$$

Where, − $C^p = C^x$ = p-th MFCC Co-efficient
− T_f is the quantity of triangle filters present in the filter bank
− G_u represents the u-th filter coefficient′ slog energy output
− b represents the quantity of MFCC coefficients that want to compute

The Mel-frequency cepstral coefficients (MFCCs) are derived by applying a decorrelation process to the coefficients using the discrete cosine transform (DCT). A subset of these coefficients, typically the first 13, are retained for further analysis and are commonly utilized in sound analysis. The MFCC feature vectors extracted from the audio frames encompass their key spectral characteristics. These feature vectors can be fed into machine learning models, which can then leverage them to distinguish between the sounds of healthy breathing and those of a patient with chronic obstructive pulmonary disease (COPD). This methodology provides a non-invasive means of diagnosing respiratory diseases through data analysis, offering the potential to revolutionize diagnostic approaches and improve healthcare services by enhancing the accuracy of diagnoses and facilitating early intervention. MFCC-based feature extraction yields valuable insights into acoustic indicators of respiratory health, holding promise for the development of more accurate and efficient respiratory disease screening and COPD identification techniques. Feature Extraction using MFCC is shown in Fig. 2.

3.3. Feature selection using sequential forward selection method

The Forward Feature Selection technique is widely adopted in machine learning and data processing disciplines to pinpoint the most relevant and impactful features (variables) that contribute significantly to a particular task, such as regression analysis or classification. This iterative approach commences with a blank slate, incrementally incorporating features one by one, unlike alternative feature selection methods that evaluate all features collectively [23].

The following outlines the forward feature selection process:

3.3.1: **Initialization** - Begin by creating an empty feature subset, denoted as F_s={}. Set the optimal model performance score to zero, utilizing metrics such as accuracy or the F1-score.

3.3.2: **Feature Assessment** - For each feature u_x not currently part of the subset F_s, create a temporary feature subset, $\mathbf{F_s^{temp} = F_s + u_x}$. Utilize only the characteristics in $\mathbf{F_s^{temp}}$ to train and evaluate the machine learning model.

3.3.3: **Model Evaluation** - Assess the model's performance using a validation dataset and the selected evaluation metric. Compute the model's performance based on the features in $\mathbf{F_s^{temp}}$.

3.3.4: **Feature Selection** - Update the best performance score and the optimal feature to include in the subset if the model performs better with $\mathbf{F_s^{temp}}$ than the current best model performance, as expressed in Eq. (3): $\mathbf{u_{best} = u_x}$.

Revise the optimal performance rating.
3.3.5: **Feature Addition** - If u_{best} is not None, indicating that adding an attribute enhanced the framework, incorporate the most effective feature, u_{best}, into the existing feature subset F_s, as expressed in Eq. (4): ($F_s = F_s + u_{best}$).
Terminate the loop if $\mathbf{u_{best}}$ = None, signifying that incorporating features did not improve the model.
3.3.6: Final Model Training - Train the final machine learning framework using the entire dataset and the selected attribute subset F_s.
3.3.7: Result - The optimal feature subset F_s, yielding the highest model performance based on the chosen evaluation metric, constitutes the end result of the Forward Feature Selection (FFS) procedure.
3.3.8: Termination Criterion - Define a termination point for the FFS process, such as identifying a point of diminishing returns on improved performance or reaching a predetermined number of features.

3.4. Developing a Hybrid SVM-KNN-CNN Model for Enhanced Respiratory Disease Classification and Diagnosis

3.4.1. Support vector machine (SVM) approach

Another potent supervised machine learning technique for tackling regression and classification problems is the Support Vector Machine (SVM)[24]. This approach operates by identifying the optimal hyper plane that can effectively categorize the data. The primary objective here is to find a hyper plane that maximally separates two classes, ensuring the largest possible margin. This margin refers to the distance between the two closest data points from different classes to the hyper plane, based on support vectors. Notably, SVM also facilitates nonlinear classification, enabling the solution of complex classification problems. To render nonlinearly separable data linearly separable, kernel functions are employed to transform the data into higher dimensions. For instance, in a binary classification scenario involving the categorization of patients into normal and COPD-afflicted groups, SVM is an ideal choice. Support Vector Machine (SVM) plays a vital role in binary classification, as it helps identify a decision boundary that effectively distinguishes between data points belonging to two classes: positive (patients with diseases like COPD) and negative (individuals without pre-existing conditions). The distance between this boundary and the nearest points from each class is referred to as the margin, and SVM's primary function is to maximize this distance. These crucial factors, known as support vectors, significantly influence classification outcomes. By maximizing the margin, SVM enhances the model's ability to handle new, unseen data, as new data points are likely to be closer to the correct side of the hyper plane. The decision boundary can be mathematically represented, as explained in Eq. (5).

$$\mathbf{w \cdot u - r = 0} \qquad \mathbf{(5)}$$

Where, − w symbolizes a weight vector that establishes the decision boundary's direction – r is a bias term that modifies the distance between the border and the origin.

Data points categorized as positive are situated on one side of the decision boundary, while negative points reside on the opposite side. This boundary confers upon SVM the ability to achieve high-precision class differentiation. The ultimate goal of the SVM optimization process is to determine the optimal values for w and r, thereby maximizing the separation of data points. Owing to SVM's distinctive capability to define decision boundaries, maximize margins, and effectively categorize data, it has emerged as a valuable tool in various applications, spanning from respiratory disease screening to general medical diagnostics. Furthermore, SVM can handle nonlinear data by leveraging kernel functions, including popular variants such as sigmoid, polynomial, radial basis function (RBF), and linear kernel. The context of the current situation and the attributes of the available data possess the potential to determine the most appropriate kernel function. Consequently, the Radial Basis Function (RBF) was frequently employed due to its ability to effectively model complex relationships within the data framework with flexibility. The kernel functions transform the data into a higher-dimensional space, enabling linear discriminants to distinguish between all data points.

3.4.2. K-Nearest Neighbors approach

K-Nearest Neighbors (KNN) is a fundamental and widely used supervised machine learning algorithm, applicable to both continuous and categorical data, and suitable for classification and regression tasks. The underlying concept of this method is rooted in the recognition of patterns and the notion that data points with similar attributes are equally likely to belong to the same category. The foundation of KNN is based on the Nearest Neighbor Principle, which states that the class of a data point can be predicted by examining the classes of its 'k-nearest' neighbors. The value of 'k' is a hyper parameter chosen by the user, and variations in the 'k' range significantly impact the algorithm's performance. As 'k' decreases, the number of local decision boundaries increases, whereas as 'k' increases, the decision boundaries become smoother, potentially introducing bias. Notably, KNN is also a prime example of instance-based learning, differing from other classification models that construct an explicit model during training. Instead, KNN stores all training data and, when presented with new data, selects the 'k' nearest data points from the training set to make predictions without learning a model.
In the classification context, KNN assigns an unlabeled instance to a specific category by identifying the classes of its nearest 'k' neighbors. Among these neighboring data points, the class with the highest frequency is selected to label the new instance. This classification approach is grounded in the assumption that data points with similar characteristics should be grouped together. To assess the differences or similarities between data points, KNN utilizes proximity metrics. Various distance measures can be employed, including the commonly used Euclidean distance, Manhattan distance, and Minkowski distance. The choice of distance metric depends on the characteristics of the data and the specific problem being addressed. For example, Euclidean distance is suitable for continuous attributes, while Hamming distance is more appropriate for categorical attributes. KNN involves calculating similarities, with proximity metrics playing a vital role in defining how distances are computed.

KNN algorithm and distance metrics
The KNN algorithm's functionality can be encapsulated in a sequence of steps:

3.4.2.1: Calculate, typically using a selected distance metric, the proximity between the target data point and all other data points in the dataset.
3.4.2.2: Based on the computed distances, select the 'k' most adjacent neighbors.
3.4.2.3: For classification tasks, determine the class labels of the k-nearest neighbors. The predicted class for the target data point is identified by examining the most frequently occurring class label among these neighbors.
3.4.2.4: For regression tasks, compute the mean (or weighted mean) of the target values for each k-nearest neighbor. This mean represents the forecasted target value for the data point.
It's essential to note that the characteristics of the data features will dictate the most suitable distance metric. For instance, when dealing with continuous data, Euclidean distance is a preferred option, as it can seamlessly integrate with real-valued attributes. In contrast, when working with categorical data, dissimilarity is often assessed using Hamming distance, which takes into account the number of differing categorical characteristics. Selecting the appropriate distance metric is crucial, as it influences the computation of data point proximity, ultimately impacting the algorithm's performance. Due to its high classification accuracy, KNN has proven to be a valuable tool in various domains, such as disease diagnosis and pattern recognition. In the context of KNN, when novel, unfamiliar data is encountered, the system assigns a class label based on the classes of the nearest neighbors. Specifically, the class label that appears most frequently among the nearest neighbors is assigned to the new data point, in accordance with the classification framework. The KNN methodology has been extensively applied to solve numerous problems, owing to its simplicity of implementation and high accuracy of classification.

3.4.3 Convolutional Neural Networks

Convolutional Neural Networks (CNNs) are a type of deep learning algorithm that's particularly well-suited for image and video processing. The main components are Convolutional Layers, Activation Functions, Pooling Layers, Flatten Layer and fully connected layers. The network takes an image as input. The convolutional layer applies filters to the image, detecting local patterns and features and the output is passed through an activation function, introducing non-linearity. The pooling layer down samples the output, reducing spatial dimensions. Above steps repeated, with each convolutional and pooling layer learning to detect more complex features and the output is flattened into a one-dimensional array. The fully connected layers classify the input image based on the extracted features.

3.5. Fusion of SVM, KNN, and CNN for Boosted Machine Learning Outcomes

Collect lung function test data (e.g., spirometry) and/or imaging data (e.g., chest X-rays or CT scans) from patients with COPD and healthy controls. Data is cleaned, preprocessed, and normalized. Extract relevant features from the preprocessed data, such as spirometry features and imaging features using SVM and KNN. Use convolutional layers to automatically extract features from the imaging data. Train an SVM model using the extracted features to classify patients as COPD or healthy. Train a KNN model using the extracted features to classify patients as COPD or healthy. Train a CNN model using the imaging data to classify patients as COPD or healthy. Combine the predictions from the SVM, KNN, and CNN models using techniques such as voting, weighted averaging and stacking etc. Create an ensemble model that integrates the predictions from the individual models. Evaluate the performance of the ensemble model using metrics such as accuracy, sensitivity, specificity, precision, recall, F1-score, and ROC-AUC and compare the performance of the ensemble model with the individual models (SVM, KNN, and CNN) to determine the effectiveness of the ensemble approach.

The strategic integration of SVM, KNN, and CNN offers a robust approach to enhancing the performance of machine learning models in automated screening for respiratory diseases and COPD. This synergy leverages the complementary strengths of each algorithm. SVM excels in classifying data points and defining decision boundaries, but may struggle with non-linear relationships or imbalanced classes. By utilizing SVM as the initial step for data categorization, the foundation for integration is established. SVM's ability to create a decision function provides a solid platform for the model. KNN is then incorporated to refine the categorization results, particularly when local patterns are crucial for prediction and neighboring effects enhance forecasting. Meanwhile, CNN's convolutional layers enable automatic feature extraction from imaging data, optimizing image processing tasks. Through this tactical collaboration, the hybrid model harnesses the local adaptability of KNN, the global optimization capabilities of SVM, and the efficient image processing of CNN. The SVM-KNN-CNN fusion model's workflow is a deliberate strategy to fully leverage the strengths of each algorithm.

Quantitative results demonstrate that the optimized ensemble of KNN's localized learning, SVM's global optimization, and CNN's efficient image processing capabilities achieves superior classification accuracy. This fusion approach leverages the strengths of each algorithm, combining SVM's robust foundation, KNN's adaptability, and CNN's automated feature extraction. Consequently, the model exhibits enhanced resistance to noisy data, high accuracy, and powerful image processing capabilities. The integration of KNN's local context assessment and SVM's initial classification score enables the model to effectively address complex datasets and noisy data points. Moreover, the hybrid approach excels in handling skewed datasets, a common challenge in medical diagnostics. KNN's local operation mode ensures that underrepresented classes receive special consideration, fostering accurate classification. The effectiveness of the hybrid model is particularly pronounced when dealing with non-linear data characteristics. By combining KNN's local decision boundary detection and SVM's high-dimensional space operation, the model can capture intricate non-linear relationships within the data. Furthermore, the incorporation of CNN enables the model to automatically extract relevant features from imaging data, enhancing its performance in respiratory disease screening. The proposed hybrid model demonstrates its utility in identifying COPD and screening for respiratory diseases, showcasing its robustness, accuracy, and adaptability in responding to the unique characteristics of real-world medical data.

IV. RESULTS AND DISCUSSION

The effectiveness and potential benefits of the respiratory disease screening model in the healthcare sector have been demonstrated by its successful implementation in practical scenarios. By harnessing simulation tools, the proposed SVM-KNN-CNN fusion model can be comprehensively evaluated and directly compared to other prominent approaches in identifying respiratory disease risks. With a 95% accuracy rate in this rigorous assessment, the model exhibits remarkable precision. This outstanding outcome further reinforces its practicality in routine clinical practice by showcasing its ability to accurately diagnose and categorize respiratory diseases. Quantitative analysis, comprising mean, root mean square error (RMSE), and standard deviation (STD) measurements, was conducted to evaluate the predictive validity of the model for selected pathways. This analysis was performed using Python. Moreover, considering the goals of respiratory disease screening, pattern recognition, and COPD detection, careful consideration should be given to selecting the most suitable metrics for the SVM-KNN-CNN fusion technique.

4.1. Model performance metrics

Evaluating the effectiveness of machine learning and pattern recognition algorithms for respiratory disease assessment, COPD identification, and classification using SVM-KNN hybrid techniques relies heavily on model performance metrics. These metrics assess the model's ability and reliability to deliver accurate diagnoses and predictions, providing a comprehensive evaluation of its performance. The following key performance indicators are commonly used to measure model efficacy.

4.1.1 Accuracy

Accuracy represents the proportion of correct predictions made by the model out of all predictions [26]. It provides a comprehensive overview of the model's precision in categorizing cases, as formulated in Eq. (6).

$$\textbf{Accuracy} = \frac{\textbf{No.of Accurately Classified Predictions}}{\textbf{Total Number of Predictions}} \times \mathbf{100\%} \quad \textbf{(6)}$$

4.1.2 Precision

Precision gauges the model's correctness in detecting true positives, also known as the model's positive predictive accuracy. It calculates the ratio of actual positive predictions to the total number of positive forecasts made [27].Precision is crucial for minimizing false positive errors, as formulated in Eq. (7).

$$\textbf{Precision}(\mathbf{P}) = \frac{\mathbf{TP}}{\mathbf{TP+FP}} \quad \textbf{(7)}$$

4.1.3 Recall

Recall, also known as sensitivity or true positive rate, assesses the approach's effectiveness in accurately identifying all actual positive instances. It measures the proportion of genuine positives to the total number of actual positive cases [28]. When minimizing false negatives is the primary objective, recall becomes indispensable, as formulated in Eq. (8).

$$\mathbf{R} = \frac{\mathbf{TP}}{\mathbf{TP+FN}} \quad \textbf{(8)}$$

4.1.4 F1-Score

The F1-Score represents the balanced average of recall and precision, providing a comprehensive evaluation of both false positives and false negatives. It offers an unbiased assessment of the model's performance, striking a balance between precision and recall [29], as formulated in Eq. (9).

$$\mathbf{F1} = \mathbf{2} * \frac{\mathbf{P \cdot R}}{\mathbf{P+R}} \quad \textbf{(9)}$$

4.1.5 Mean

The mean serves as an indicator of the central value of a dataset, representing the average value of all data points, as calculated in Eq. (10).

$$\textbf{Mean}(\boldsymbol{\mu}) = \frac{\mathbf{1}}{\mathbf{N}} \sum_{\mathbf{i=1}}^{\mathbf{N}} \mathbf{x_i} \quad \textbf{(10)}$$

Where xi represents individual data points and N denotes the total count of data points.

4.1.6 Root Mean Squared Error (RMSE)

The Root Mean Squared Error (RMSE) measures the discrepancy between forecasted and actual values, with its calculation based on Eq. (11).

$$\textbf{RMSE} = \sqrt{\frac{\mathbf{1}}{\mathbf{N}} \sum_{\mathbf{I=1}}^{\mathbf{N}} (\mathbf{y_i} - \mathbf{y'_i})^2} \quad \textbf{(11)}$$

Where yi represents the true value, y′ denotes the forecasted value, and N signifies the overall number of data points or observations [30].

4.1.7 Standard Deviation

Standard Deviation (STD): A measure of data dispersion, quantifying the degree of variation within a dataset [28], calculated using Eq. (12).

$$\mathbf{STD}(\sigma) = \sqrt{\frac{1}{N-1}} \sum_{I=1}^{N} (x_i - \mu)^2 \quad \textbf{(12)}$$

Where xi represents each individual data point, μ denotes the average value of the dataset [30], and N signifies the total count of data points.

4.1.8 G-Mean (Geometric Mean)

The G-Mean provides a balanced evaluation of a classifier's performance across both positive and negative classes, making it particularly useful in scenarios with class imbalance. The calculation is based on the formula presented in Eq. (13):

$$\mathbf{G-Mean} = \sqrt{\mathbf{Sensitivity} \times \mathbf{Specificity}} \quad \textbf{(13)}$$

Where Sensitivity (True Positive Rate) and Specificity (True Negative Rate) are considered.

4.1.9 MCC (Matthews Correlation Coefficient)

The MCC provides a more nuanced evaluation of a binary classifier's performance, incorporating true positives, false positives, true negatives, and false negatives. It yields a score ranging from −1 to 1, where:
A score of +1 signifies flawless prediction, 0 implies performance equivalent to chance, and -1 indicates complete discord between predicted and observed outcomes. The calculation is based on the formula presented in Eq. (14):

$$\mathbf{MCC} = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP+FP) \times (TP+FN) \times (TN+FP) \times (TN+FN)}} \quad \textbf{(14)}$$

4.1.10 Specificity

Specificity refers to the percentage of genuine negative cases that are accurately classified as negatives (True Negative Rate). The calculation is based on the formula presented in Eq. (15):

$$\mathbf{Specificity} = \frac{TN}{TN+FP} \quad \textbf{(15)}$$

Where TP denotes True Positives, TN represents True Negatives, FP signifies False Positives, and FN indicates False Negatives. A confusion matrix provides a concise evaluation of a classification model's accuracy by juxtaposing actual and predicted class labels, as illustrated in Table 1.

Table 1
Confusion Matrix

	Predicted Normal	Predicted Disorder
Actual Normal	True Negative	False Positive
Actual Disorder	False Negative	True Positive

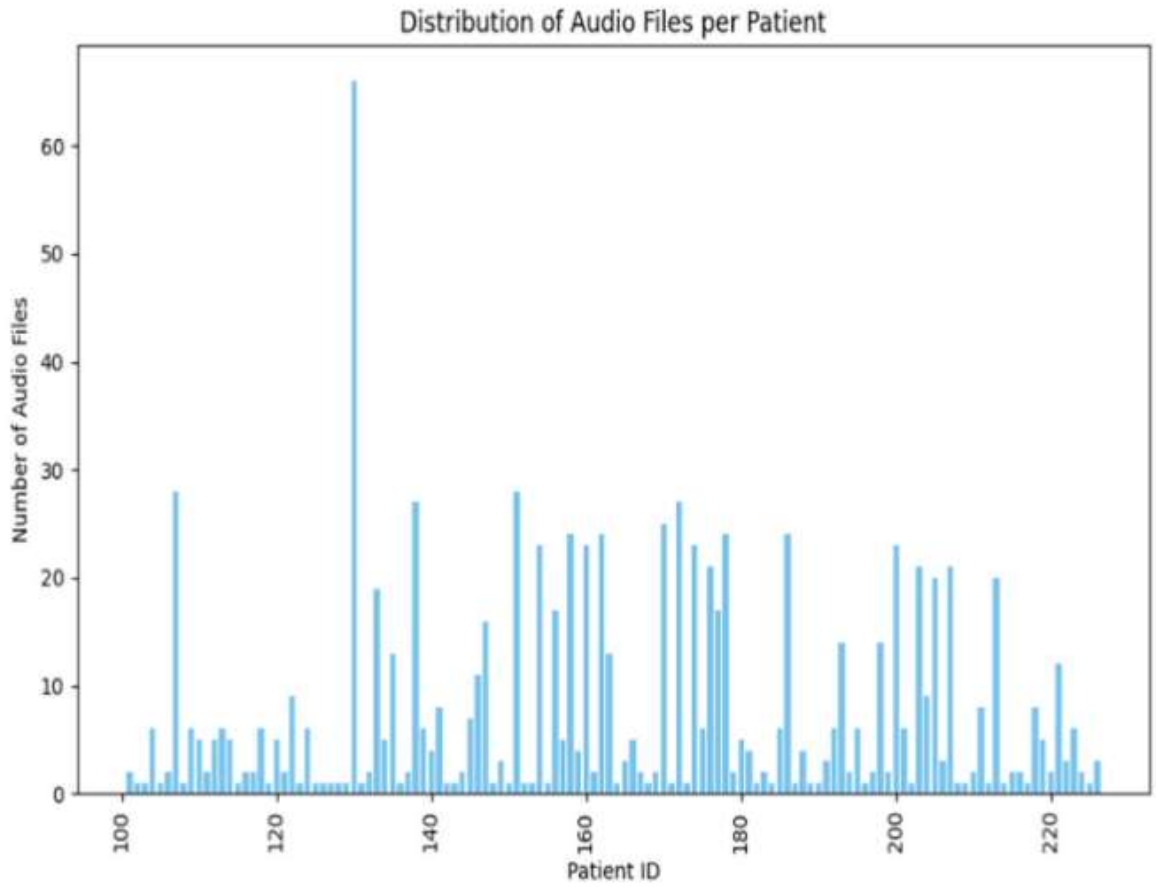

Fig. 3. Distribution of audio files per patient.

Fig. 3 highlights the varying frequency of recordings associated with each patient ID, with most patients having fewer than 20 audio files linked to their ID. However, notable exceptions exist, such as one patient with 51 audio files, significantly surpassing the others. This striking disparity suggests that while the majority of patients provide limited audio data, a select few offer more extensive and valuable data, which can be instrumental in conducting in-depth analyses for various medical treatment or research objectives—especially when audio recordings are utilized for disease monitoring or diagnostic purposes.

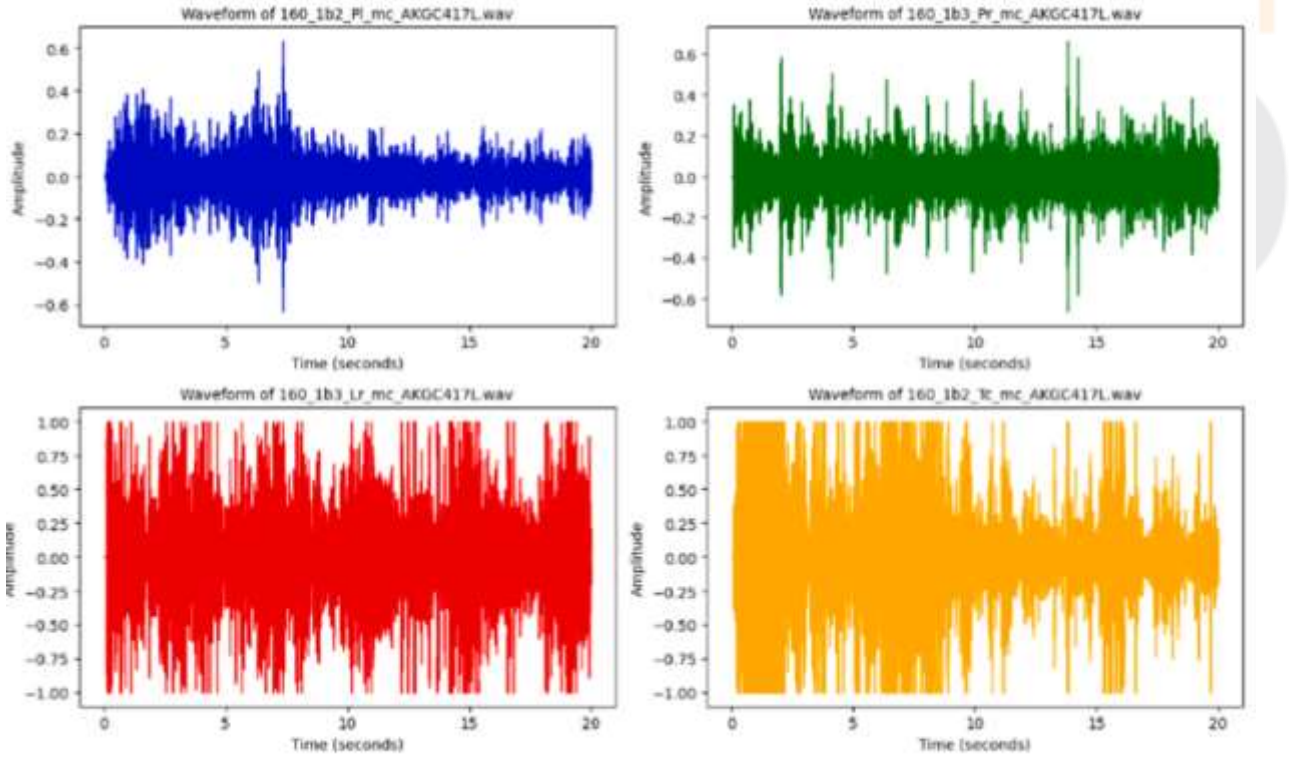

Fig. 4. Waveforms of four distinct audio signals.

Fig. 4 illustrates the temporal variations in the loudness of the sound produced during the experiment. The signals are labeled as follows: "160 162 Pl mc AKGC417Lwav" (flagged in blue), "160 163 Pr mc AKOC417Lwav" (flagged in green), "160 163 LF mc AKOCAITL Wav" (red), and "160 362 mc AKGC4171 wav" (detected in yellow). The X-axis represents time in seconds,

while the Y-axis indicates the amplitude or volume of the audio signal. The positive and negative signs show fluctuations in sound pressure. Comparing the blue and green waveforms reveals minimal distortion, with both waveforms displaying similar patterns and closely matching amplitudes, suggesting a potential relationship between the two audio sources. In contrast, the red waveform differs from the yellow one, indicating different sound characteristics or recording conditions. However, further information about the signals' nature, the audio source, the recording environment, and the analysis objectives is needed for a more comprehensive understanding.

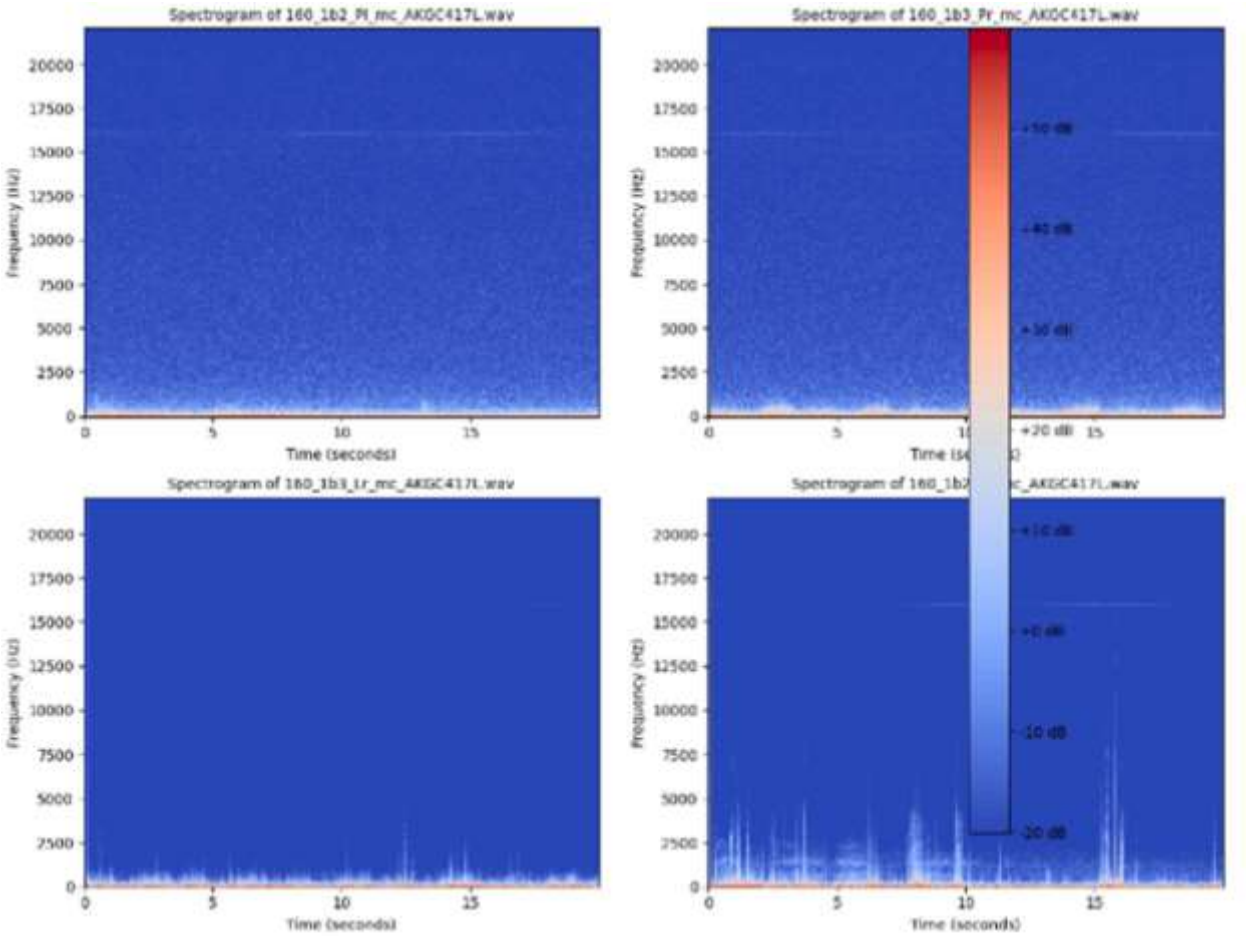

Fig. 5. spectrograms of four respiratory sound samples.

Fig. 5 presents spectrograms of four respiratory sound samples, displayed as frequency spectrograms, each showing the frequency components at different time intervals. All spectrograms primarily cover up to 20,000 Hz, which aligns with the frequency range typical of human vocalization and breathing. Each sample lasts approximately 15 seconds, with the spectrograms revealing varying spectral patterns. Spectrogram 1 is distinguished by high-frequency bands and the clear presence of harmonics, which may be related to respiration. Spectrogram 2 exhibits a high concentration in this frequency range, possibly indicating noise or a specific sound. Spectrogram 3 of the same signal highlights semi-low frequencies with intricate and well-defined harmonics, likely resulting from more resonant respiratory sounds. Spectrogram 4 focuses on low-frequency power, suggesting a potential respiratory condition. The amplitude is represented by warmer colors, indicating stronger signals, which contribute to the intensity and pattern of each sample. This analysis provides insights into the acoustic properties of respiratory sounds, offering valuable information for further research on respiratory health.

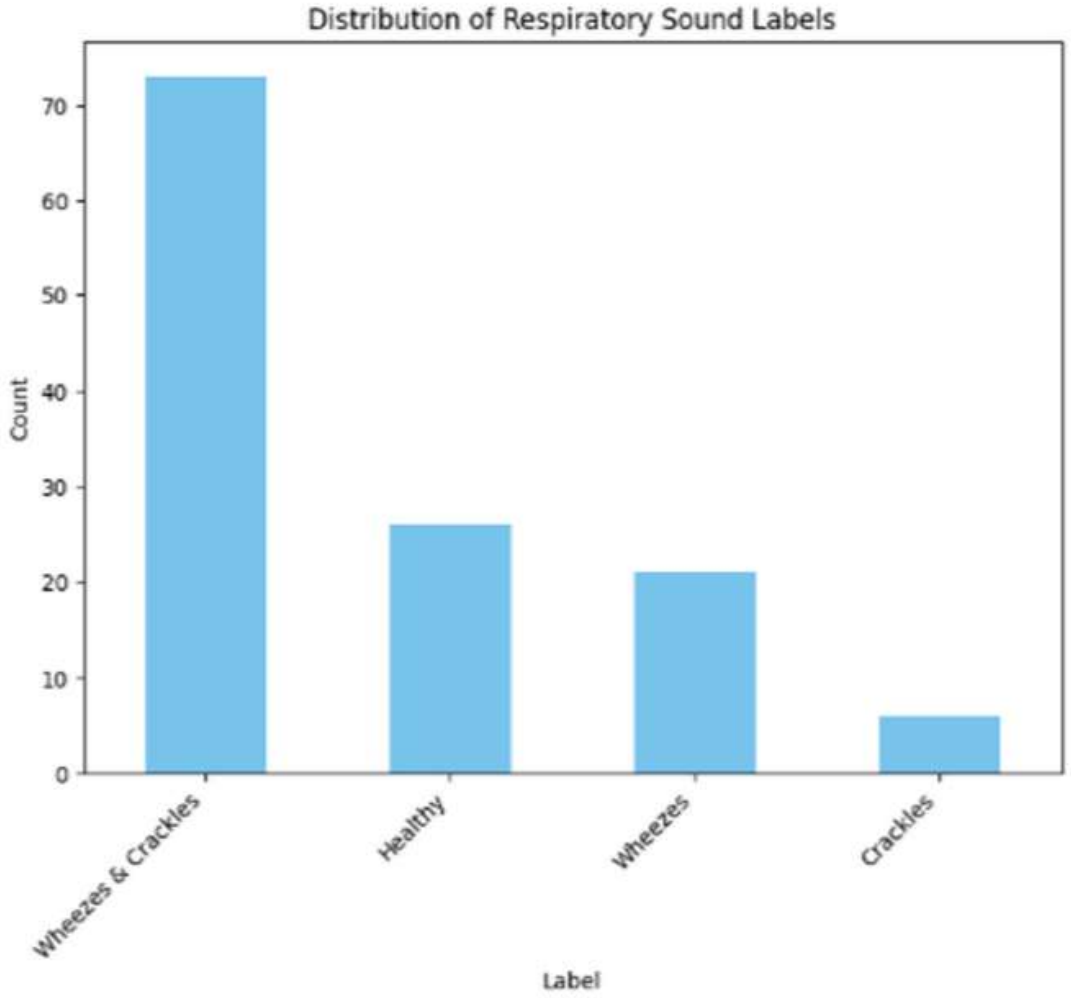

Fig. 6. Distribution of respiratory sound labels.

Fig. 6, titled "Distribution of Respiratory Sound Labels," displays the distribution of various respiratory sound labels. The x-axis represents the labels, which include “Wheezes & Crackles,” “Healthy,” “Wheezes,” and “Crackles,” while the y-axis shows the count of each label. The majority of the data falls under the “Wheezes & Crackles” label, followed by the “Healthy” label, with “Wheezes” and “Crackles” representing smaller proportions. These proportions suggest that most of the respiratory sounds in this dataset contain both wheezes and crackles, with fewer samples being categorized as healthy or containing only wheezes or crackles.

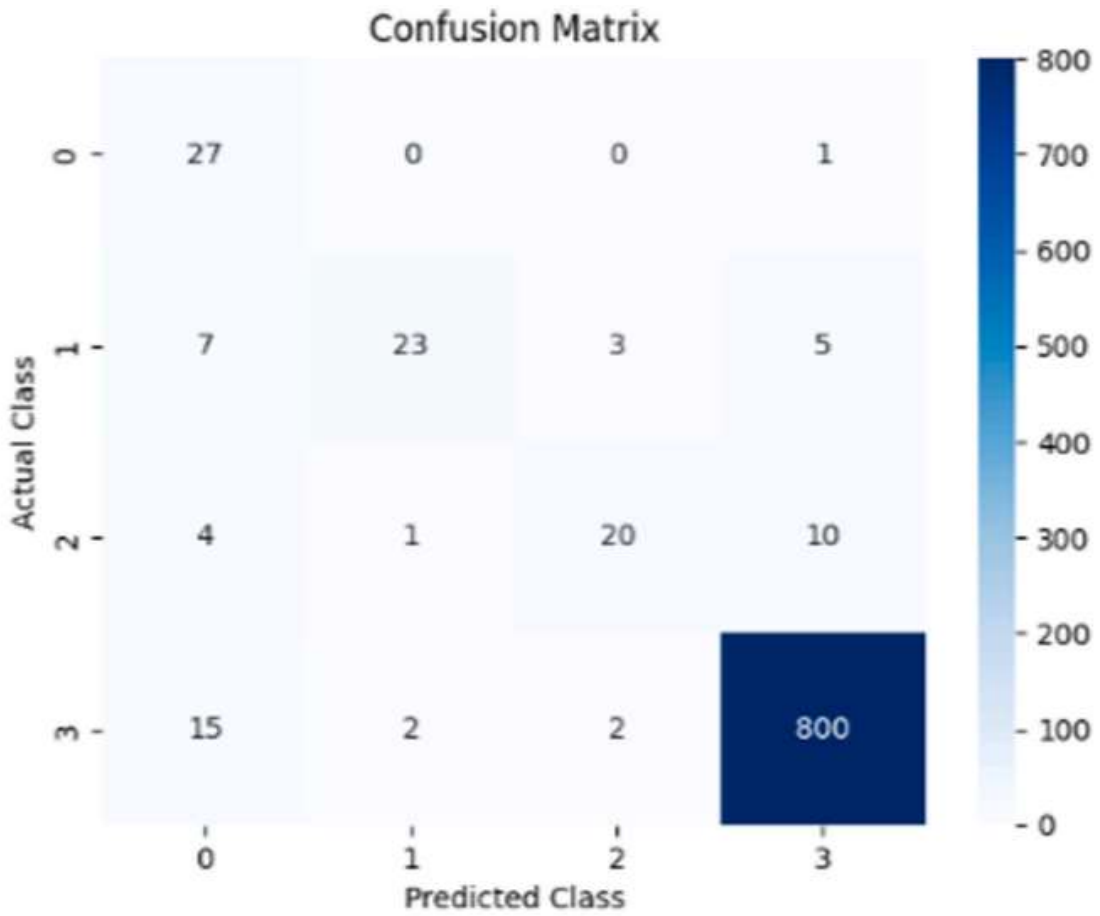

Fig. 7. Confusion matrix.

Fig. 7, illustrates a confusion matrix that assesses the accuracy of a respiratory sound classification model. The matrix is organized with actual classes on the vertical axis and predicted classes on the horizontal axis. The diagonal elements (27, 23, 20, 800) denote accurately predicted instances for each class, whereas non-diagonal elements signify incorrect classifications. For instance, 27 instances of class 0 (potentially normal sounds) were precisely categorized, whereas 3 instances of class 1 (probable crackles) were misidentified as class 2 (probable wheezes). This matrix showcases the model's ability to differentiate between various respiratory sound categories, with elevated diagonal values indicating exceptional performance and non-diagonal values revealing classification mistakes (Table 2).

Table 2
metrics for each fold

Fold	Accuracy	Precision	Recall	F1 Score	G-Mean	MCC	Specificity	RMSE
Fold 1	0.96	0.89	0.85	0.85	0.84	0.84	0.96	0.35
Fold 2	0.95	0.83	0.75	0.78	0.73	0.75	0.93	0.4
Fold 3	0.95	0.86	0.79	0.8	0.76	0.79	0.94	0.42
Fold 4	0.94	0.78	0.79	0.77	0.78	0.74	0.94	0.48
Fold 5	0.92	0.74	0.7	0.69	0.66	0.69	0.93	0.55

Table 3 illustrates the model's performance based on five-fold evaluation metrics, providing a comprehensive overview of its strengths and weaknesses across key performance indicators: Accuracy, Precision, Recall, F1 Score, G-Mean, MCC (Matthews Correlation Coefficient), Specificity, and RMSE (Root Mean Square Error). The results demonstrate consistent accuracy, ranging from 0.92 to 0.96, indicating reliable and robust performance. Although precision and recall vary across folds, fold 1 exhibits an impressive balance between the two. The F1 Score effectively balances precision and recall, while G-Mean and MCC, as measures of discriminatory power and overall predictive quality, exhibit a slight decline across successive folds. Specificity remains consistently high, indicating accurate identification of negative cases. Conversely, RMSE gradually increases with each fold, suggesting a rise in prediction error rates. This analysis reveals that the model performs well overall, but its precision, recall, and error rates vary across folds. A box plot illustrating the distribution of metrics across folds is presented in Fig. 8.

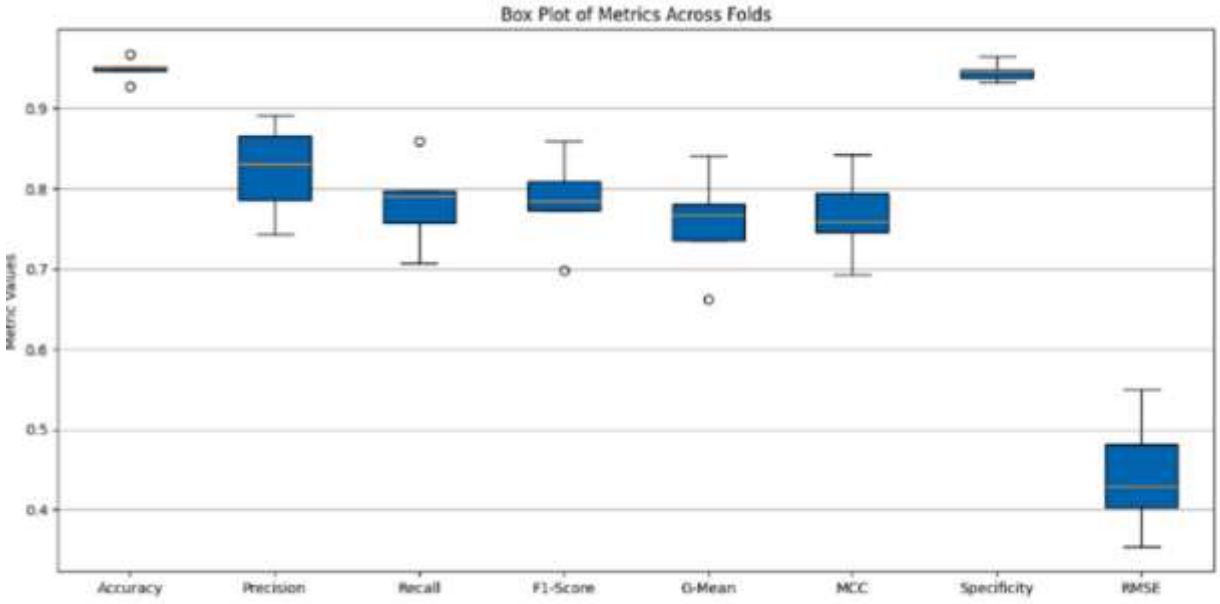

Fig. 8. Box plot of metrics across folds.

Table 3
Overall metrics for mean and standerd deviations and RMSE

Metric	Mean	STD	RMSE
Accuracy	**0.94**	**0.013**	**0.0508**
Precision	**0.82**	**0.053**	**0.1762**
Recall	**0.78**	**0.050**	**0.2173**
F1 Score	**0.78**	**0.052**	**0.2151**
G-Mean	**0.75**	**0.058**	**0.2423**
MCC	**0.76**	**0.049**	**0.2332**
Specificity	**0.94**	**0.010**	**0.05389**
RMSE	**0.44**	**0.067**	

Table 4 summarizes the overall performance metrics of the model, including the mean, standard deviation (STD), and root mean square error (RMSE) for each metric. The accuracy metric exhibits a high mean value (0.94) with minimal variability (STD: 0.013) and a low RMSE of 0.0508, indicating consistent and reliable performance. Precision (0.82) and recall (0.78) display moderate values with somewhat higher variability (STD: 0.053 and 0.050, respectively), while the RMSE suggests opportunities for improvement in predicting relevant instances. The F1 score (0.78) combines precision and recall, demonstrating similar performance and variability. G-Mean and MCC,which assess the balance between different classes, are slightly lower (0.75 and 0.76, respectively) but stable and consistent. Specificity (0.94) is high, reflecting accurate and reliable identification of negative cases. The RMSE for model prediction (0.44) exhibits the highest variability (STD: 0.067), indicating fluctuations in performance. A boxplot illustrating the overall metrics is presented in Fig. 9.

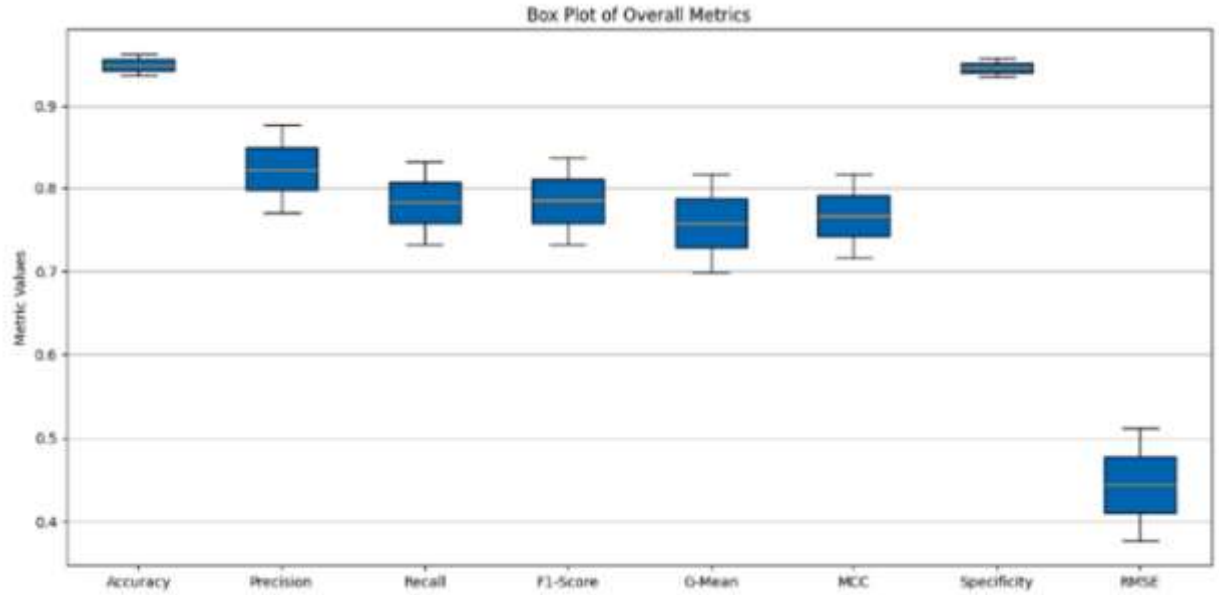

Fig. 9. Boxplot of overall metrics.

4.2. Accuracy comparison

The efficiency of various screening models can be evaluated by utilizing these accuracy metrics as a key benchmark. The most suitable model will depend on several factors, including the dataset's characteristics, the balance between false positives and false negatives, and specific clinical requirements. Beyond accuracy, this study examines additional performance indicators, such as precision, recall, and F1 score, to inform the selection of the most appropriate model for respiratory illness screening. Table 3 presents a comparative accuracy assessment using the proposed methodology.

A comprehensive comparison of various respiratory illness screening procedures, as presented in Table 4, is essential for determining the precision of these methods in detecting and classifying respiratory disorders. The proposed SVM-KNN-CNN fusion model demonstrates exceptional accuracy, achieving an impressive 95% or higher success rate, thereby surpassing the performance of other evaluated strategies and showcasing its remarkable capability for accurate disease diagnosis. This outstanding accuracy underscores the model's potential to significantly enhance the effectiveness of respiratory disease assessment, rendering it a notable breakthrough in healthcare. Although other methods, including CNN, LSTM-GRU-CNN, Naive Bayes, Boosted Decision Tree, SVM, XGBoost, ELM, and KNN, exhibited varying degrees of accuracy, the SVM-KNN-CNN model's superior performance highlights the crucial role of machine learning and pattern recognition in facilitating early and precise disease identification in the medical field. A graphical representation of the accuracy comparison evaluation is illustrated in Fig. 10.

Table 5
Accuracy Comparison

Techniques	**Accuracy**
LSTM-GRU-CNN [30]	75%
Naive Bayes [31]	72.8%
CNN [32]	79.3%
Boosted Decisional Tree [33]	85%
SVM and XGBoost [34]	89%
ELM and KNN [35]	90.71%
SVM-KNN [1]	94%
Proposed (SVM-KNN-CNN)	95% or higher

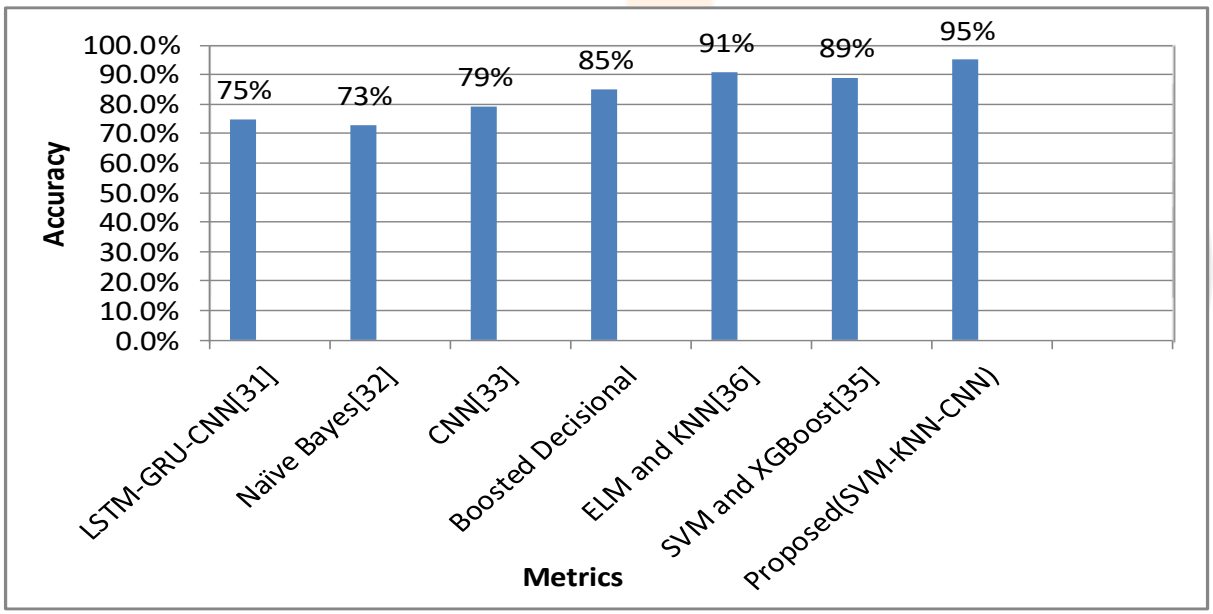

Fig.10. Accuracy comparison

References

[1]Ahmed I. Taloba a,R.T. Matoog, Detecting respiratory diseases using machine learning-based pattern recognition on spirometry data, Alexandria Engineering Journal 113(2025) 44-59

[2] M.A.A. Albadr, M. Ayob, S. Tiun, F.T. Al-Dhief, M.K. Hasan, Gray wolf optimization-extreme learning machine approach for diabetic retinopathy detection, Front. Public Health 10 (2022) 925901.

[3] M.A.A. Albadr, S. Tiun, M. Ayob, F.T. Al-Dhief, Particle swarm optimization-based extreme learning machine for covid-19 detection, Cogn. Comput. 16 (4) (2024) 1858–1873.

[4] S. Aghalya, N.V. Rao, R.R. Roy, C. Srinivasan, G. Premi, Real-time monitoring and prediction of respiratory diseases using IoT and machine learning, in: in: Proceedings of the Second International Conference On Smart Technologies For Smart Nation (SmartTechCon), 2023, IEEE, 2023, pp. 635–640.
[5] Y. Feng, Y. Wang, C. Zeng, H. Mao, Artificial intelligence and machine learning in chronic airway diseases: focus on asthma and chronic obstructive pulmonary disease, Int. J. Med. Sci. 18 (13) (2021) 2871.
[6] S. Bhattacharjee, B. Saha, P. Bhattacharyya, S. Saha, Classification of obstructive and non-obstructive pulmonary diseases on the basis of spirometry using machine learning techniques, J. Comput. Sci. 63 (2022) 101768.
[7] G. Zhang, L. Luo, L. Zhang, Z. Liu, Research progress of respiratory disease and idiopathic pulmonary fibrosis based on artificial intelligence, Diagnostics 13 (3) (2023) 357.
[8] F.T. Al-Dhief, N.M.A. Latiff, N.N.N.A. Malik, M.M. Baki, N. Sabri, M.A.A. Albadr, Dysphonia detection based on voice signals using naive bayes classifier, in: in: Proceedings of the IEEE Sixth International Symposium on Telecommunication Technologies (ISTT), 2022, IEEE, 2022, pp. 56–61.
[9] M. Nahiduzzaman, et al., Parallel CNN-ELM: a multiclass classification of chest Xray images to identify seventeen lung diseases including COVID-19, Expert Syst. Appl. 229 (2023) 120528.
[10] J. Sun, X. Liao, and Y. Yan, Detection and Staging of Chronic Obstructive Pulmonary Disease Using a Computed Tomography- based Weakly Supervised Deep Learning Approach
[11] A. Celik, Diagnosis of the diseases using resampling methods with machine learning algorithms, Proc. Bulg. Acad. Sci. (2023) 1065–1076.
[12] A. Kaplan, et al., Artificial intelligence/machine learning in respiratory medicine and potential role in asthma and COPD diagnosis, J. Allergy Clin. Immunol. Pract. 9 (6) (2021) 2255–2261, https://doi.org/10.1016/j.jaip.2021.02.014.
[13] A. Srivastava, S. Jain, R. Miranda, S. Patil, S. Pandya, K. Kotecha, Deep learning based respiratory sound analysis for detection of chronic obstructive pulmonary disease, PeerJ Comput. Sci. 7 (2021) e369.
[14] O. Papadopoulou, I. Kallistratos, M. Trevlakis, A. Xalkia, M. Stavropoulou, and A. H. Papadopoulou, The Effectiveness of Respiratory Physiotherapy in Children with Severe Respiratory Problems: Literature Review, 2024.
[15] C.D. Visser, M.R. Faay, A. Ozdemir, ¨ H.-J. Guchelaar, M. Teichert, "Short-acting β2- agonists (SABA) overuse in asthma and patients' perceptions for this behavior,", Respir. Med. 231 (2024) 107723.
[16] A.-C. Paola Patricia, P.-M. Marlon Alberto, B.-M. Er-nesto, B.-A. Sharith Alejandra, B.-C. Camilo, R.- Fabian, Improving the Accuracy of Predictive Models in Imbalanced Lung Cancer Data. in: Proceedings of the International Conference on Swarm Intelligence, Springer, 2024, pp. 219–230.
[17]. Respiratory Sound Database. (Accessed 6November 2023)., 2023. (Online). 〈https://www.kaggle.com/datasets/vbookshelf/respiratory-sound-database.
[18] F.T. AL-Dhief, et al., Voice pathology detection using machine learning technique, in: in: Proceedings of the IEEE Fifth International Symposium on Telecommunication Technologies (ISTT), 2020, IEEE, 2020, pp. 99–104.
[19] M.A.A. Albadr, S. Tiun, M. Ayob, M. Mohammed, F.T. AL-Dhief, Mel-frequency cepstral coefficient features based on standard deviation and principal component analysis for language identification systems, Cogn. Comput. 13 (2021) 1136–1153.
[20] M.A.A. Albadr, et al., Fast learning network algorithm for voice pathology detection and classification, Multimed. Tools Appl. (2024) 1–32.
[21] F.T. AL-Dhief, N.M.A. Latiff, M.M. Baki, N.N.N.A. Malik, N. Sabri, M.A.A. Albadr, Voice pathology detection using support vector machine based on different number of voice signals, in: in: Proceedings of the Twenty Sixth IEEE Asia-Pacific Conference on Communications (APCC), 2021, IEEE, 2021, pp. 1–6.
[22]Mel-frequency cepstrum, Wikipedia. Oct. 25, 2023.(Accessed 6 November 2023). (Online). 〈https://en.wikipedia.org/w/index.php?title=Mel-frequency_cepstr um&oldid=1181829232〉.
[23]Feature Selection Techniques in Machine Learning- Javatpoint.(Accessed 6 November 2023), 2023. (Online). https://www.javatpoint.com/feature-selection -techniques-in-machine-learning.
[24] A. Saini, Guide on Support Vector Machine (SVM) Algorithm, Analytics Vidhya. (Accessed 6 November 2023), 2023. (Online). 〈https://www.analyticsvidhya.co m/blog/2021/10/support-vector-machinessvm-a-complete-guide-for-beginners/.
[25] T. Srivastava, A Complete Guide to K-Nearest Neighbors (Updated 2023)" Analytics Vidhya. (Accessed 6 November 2023), 2023. (Online). 〈https://www.ana lyticsvidhya.com/blog/2018/03/introduction-k-neighbours-algorithm-clustering/ .
[26] M.I.H. Alzawali, et al., Facial emotion images recognition based on binarized genetic algorithm-random forest, Baghdad Sci. J. 21 (2) (2024) 0780.
[27] F.T. AL-Dhief, et al., Voice pathology detection using decision tree classifier, in: in: Proceedings of the fourteenth international conference on information and communication technology convergence (ICTC), 2023, IEEE, 2023, pp. 36–41.
[28] M.A.A. Albadr, S. Tiun, M. Ayob, F.T. Al-Dhief, T.-A.N. Abdali, A.F. Abbas, Extreme learning machine for automatic language identification utilizing emotion speech data. in: Proceedings of the International Conference on Electrical, Communication, and Computer Engineering (ICECCE), IEEE, 2021, pp. 1–6.
[29] M.A.A. Albadr, F.T. AL-Dhief, L. Man, A. Arram, A.H. Abbas, R.Z. Homod, Online sequential extreme learning machine approach for breast cancer diagnosis, Neural Comput. Appl. (2024) 1–17.
[30] M.A.A. Albadr, M. Ayob, S. Tiun, F.T. Al-Dhief, A. Arram, S. Khalaf, Breast cancer diagnosis using the fast learning network algorithm, Front. Oncol. 13 (2023) 1150840.
[31]Respiratory Disease Classification, LSTM/GRU/CNN. (Accessed 6 November 2023),2023.(Online). 〈https://kaggle.com/code/pvnbalaramamurthy/respiratory-disea se-classification-lstm-gru-cnn〉.
[32] R.V. Sharan, H. Rahimi-Ardabili, Detecting acute respiratory diseases in the pediatric population using cough sound features and machine learning: a systematic review, Int. J. Med. Inf. 176 (2023) 105093.
[33] P. Podder, et al., Lddnet: a deep learning framework for the diagnosis of infectious lung diseases, Sensors 23 (1) (2023) 480.
[34] G. Chambres, P. Hanna, M. Desainte-Catherine, Automatic detection of patient with respiratory diseases using lung sound analysis. in: Proceedings of the International Conference on Content-Based Multimedia Indexing (CBMI), IEEE, 2018, pp. 1–6.

[35] V. Binson, M. Subramoniam, Y. Sunny, L. Mathew, Prediction of pulmonary diseases with electronic nose using SVM and XGBoost, IEEE Sens. J. 21 (18) (2021) 20886–20895.
[36] Z. Neili, M. Fezari, A. Redjati, ELM and K-nn machine learning in classification of Breath sounds signals, Int. J. Electr. Comput. Eng. 10 (4) (2020) 3528–3536.