IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

RespiTech: Audio Analysis Using Deep Learning

[1]Aman Garasia, [2]Zainab Patel, [3]Shruti Shetty, [4]Tanvi Shah, [5]Mahendra Patil

[1]Student, [2]Student, [3]Student, [4]Student, [5]Professor

[1]Computer Engineering,

[1]Atharva College Of Engineering, Malad (W), India

Abstract:

The advancement of intelligent systems has significantly impacted the healthcare industry, enabling the development of automated diagnostic tools for early disease detection and monitoring. Respiratory diseases, particularly in children, are among the most common health concerns, requiring timely and accurate diagnosis. Traditional methods rely on clinical evaluations and laboratory tests, which can be time-consuming and resource-intensive. In this study, we introduce RESPITECH, a deep learning-based framework for automated cough sound classification. By leveraging audio signal processing techniques and machine learning, RESPITECH aims to differentiate between healthy and pathological coughs in individuals, with a focus on respiratory conditions such as asthma, upper respiratory tract infection (URTI), and lower respiratory tract infection (LRTI), COPD, Pneumonia, Bronchitis and also Healthy lungs.

Our approach employs a Convolutional Neural Network model (CNN) trained on Mel-Frequency Cepstral Coefficients (MFCCs) extracted from a carefully curated dataset of cough sounds labeled by clinicians. The CNN model is designed to capture temporal dependencies in cough sounds, enhancing classification performance. When trained to distinguish between healthy and pathological coughs, the model achieves an accuracy exceeding 84%, aligning closely with physician diagnoses. To improve classification performance for specific respiratory pathologies, multiple cough epochs per subject are aggregated, resulting in an overall accuracy surpassing 91% for detecting asthma, URTI, and LRTI. However, classification performance declines when distinguishing between four separate cough categories, as certain pathological coughs share overlapping acoustic characteristics, leading to misclassification.

A detailed analysis of the MFCC feature space through a longitudinal study indicates that pathological coughs, regardless of the underlying condition, exhibit similar acoustic patterns, making it challenging to differentiate between specific respiratory diseases using MFCCs alone. This suggests that while MFCCs provide valuable spectral information, additional features or multimodal data integration may be necessary for finer disease discrimination. Despite these challenges, RESPITECH demonstrates strong potential as an automated, non-invasive screening tool for respiratory disease monitoring. Future work will explore advanced feature extraction techniques, multimodal learning approaches, and real-world deployment strategies to enhance diagnostic accuracy and clinical applicability.

Keywords- Deep Learning, Neural Networks, Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM), Feature Extraction, Data Representation, Audio Classification.

I. INTRODUCTION

Recognizing meaningful patterns in medical data is essential for accurate diagnosis and effective disease classification, However, this task remains a significant challenge, particularly when working with large-scale, high-dimensional, and non-linear data. Traditional statistical methods and mathematical model often struggle to capture the inherent complexity and variability present in biomedical signals. [2] This has prompted the integration of more advanced computational techniques, particularly in the fields of artificial intelligence (AI) and machine learning (ML) which have demonstrated significant potential in medical diagnostics, enabling automated analysis and improved decision-making. RESPITECH is a deep learning-based system designed to classify respiratory sounds and assist in diagnosing pulmonary conditions. Unlike traditional rule-based expert systems, which suffer from high error rates, RESPITECH leverages AI-driven approaches to enhance the accuracy and reliability of respiratory disease classification.

Deep learning models, particularly Convolutional Neural Networks (CNNs), have been widely used in various domains, including image recognition, speech processing, and biomedical signal analysis. CNNs are highly effective in processing structured data due to their ability to learn spatial hierarchies of features. More recently, CNNs have been adapted for speech and audio analysis, leading to significant improvements in sound classification tasks. RESPITECH builds upon these advancements by incorporating Mel-Frequency Cepstral Coefficients (MFCCs) as input features and utilizing a Bidirectional Long Short-Term Memory (BiLSTM) network to analyze and classify respiratory sounds. While CNN layers focus on spatial feature extraction, BiLSTM layers are adept at modeling temporal dependencies, making the combined model highly effective in capturing both local and sequential patterns in respiratory audio signals. RESPITECH provides a robust framework for distinguishing between healthy and pathological lung sounds, including conditions such as asthma, upper respiratory tract infection (URTI), and lower respiratory tract infection (LRTI).

The primary objective of RESPITECH is to develop a non-invasive, AI-powered diagnostic tool capable of assisting healthcare professionals in detecting respiratory conditions with high accuracy. Our study is based on a dataset of audio recordings collected by chest physicians, ensuring reliable ground truth for training and validation. The proposed system evaluates multiple cough epochs per subject, improving classification performance and reducing misdiagnosis. Additionally, RESPITECH explores the limitations of existing feature representations, particularly MFCCs, in differentiating between overlapping pathological conditions. Through this research, we aim to demonstrate the potential of RESPITECH as a scalable and efficient solution for respiratory sound analysis, ultimately contributing to AI-driven advancements in pulmonary disease detection and monitoring.

II. LITERATURE REVIEW

RESPITECH represents a transformative approach to AI-driven respiratory disease diagnosis. By automating sound analysis, improving classification accuracy, and enabling remote health monitoring, RESPITECH addresses the key limitations in existing diagnostic methods. This system has the potential to significantly reduce misdiagnoses, improve early detection rates, and enhance healthcare accessibility worldwide.

The study by **Rajkumar Palaniappan and K. Sundaraj**, **[1]** explores the use of Mel-Frequency Cepstral Coefficients (MFCCs) for feature extraction and Support Vector Machine (SVM) for classifying respiratory sounds into normal, airway obstruction, and parenchymal pathology. The model was trained using the RALE database and achieved an impressive mean classification accuracy of 90.77%, with 94.11% for normal, 92.31% for airway obstruction, and 88.00% for parenchymal pathology. The findings of this research emphasize the importance of MFCC-based feature extraction in distinguishing pathological lung sounds, which directly contributes to RESPITECH's signal processing techniques for automated respiratory sound classification.

In the paper **[2]** by **Ida Ayu Putu Ari Crisdayanti and Seong-Eun Kim** propose a lightweight Convolutional Neural Network (CNN) model for the automated classification of respiratory anomalies. Their approach utilizes multilevel feature fusion, allowing the CNN to extract both local and global acoustic features from respiratory sounds. By testing their model on public respiratory sound datasets, they demonstrate state-of-the-art accuracy in detecting abnormal lung sounds. The insights from this study directly influence RESPITECH's

CNN-based classification model, ensuring robust and scalable deep learning capabilities for real-time respiratory disorder detection.

Amoh and Odame's study **[4]** shows that even shallow neural networks, when trained on MFCCs, can classify cough types with moderate accuracy. Though simpler in structure, their approach validates the idea of using audio-based diagnostics, which RESPITECH advances with deeper networks and larger, clinically validated datasets.

Finally**, Laguarta et al. [6]** present a multi-task deep learning framework that uses cough, breathing, and voice data to detect COVID-19. Their work highlights the potential of multimodal learning for broader disease coverage, a strategy RESPITECH could adapt to expand its diagnostic utility beyond traditional pulmonary diseases.

By incorporating advanced signal processing techniques, deep learning architectures, and machine learning classifiers, RESPITECH aims to provide a highly accurate, scalable, and real-time AI-driven diagnostic tool for early detection of respiratory diseases. These research contributions form the technical foundation of RESPITECH, enabling efficient respiratory sound classification and automated clinical decision support.

III. PROPOSED SYSTEM

RESPITECH is a smart, AI-powered system developed to automatically analyze and classify respiratory sounds, with the goal of helping doctors detect lung-related diseases at an early stage. It combines modern deep learning techniques with audio processing to make diagnosis faster, more consistent, and more accessible specially in remote or under-resourced areas.The system starts by cleaning and preparing the recorded cough or breathing sounds to remove background noise. Then, it uses Mel-Frequency Cepstral Coefficients (MFCCs), a widely used method in audio analysis, to extract important features from the sound. These features help the system understand differences between normal and abnormal lung sounds.

RESPITECH uses a deep learning model that combines two powerful approaches: a Convolutional Neural Network (CNN), which identifies important sound patterns, and a Bidirectional Long Short-Term Memory (BiLSTM) network, which captures how those sound patterns change over time. [5] This combination allows the system to better recognize specific respiratory conditions like asthma, COPD, and pneumonia.To make the model more accurate and reliable in different real-world situations, various techniques are used during training—such as modifying the pitch and speed of cough sounds or adding background noise. These steps help the system learn to perform well even with varied or imperfect recordings.

RESPITECH is designed to be lightweight and fast, meaning it can work on devices like smartphones or tablets. This makes it a practical tool for mobile health apps or telemedicine platforms, where quick, on-the-go diagnosis is essential. It also includes features to estimate how confident the system is in its predictions and can show which parts of a sound were most important in making a decision—offering transparency and building trust for healthcare providers.

By combining easy-to-use technology with powerful AI, RESPITECH provides a scalable, low-cost, and non-invasive solution for monitoring respiratory health. Its ability to work in real time and be used in remote areas makes it especially valuable in today's growing demand for accessible healthcare solutions.

3.1 Framework

RESPITECH is a deep learning-powered SaaS platform designed for automated classification of respiratory sounds to assist in the early detection of pulmonary diseases. The system leverages advanced machine learning frameworks and signal processing libraries to ensure efficient feature extraction, classification, and real-time processing. [8]

1. Convolutional Neural Network (CNN): RESPITECH utilizes a CNN-based deep learning model to classify respiratory sounds. CNNs effectively capture spatial and temporal patterns in spectrogram representations of lung sounds, enhancing the system's ability to differentiate between healthy and pathological cases.

2. Keras: Keras serves as the high-level deep learning API for designing, training, and fine-tuning CNN models. It simplifies the development process while enabling efficient GPU acceleration for training large-scale respiratory sound datasets.

3. TensorFlow: TensorFlow powers the backend computation for CNN training and inference. It enables optimized matrix operations, automatic differentiation, and distributed training, ensuring high-performance processing of respiratory audio data.

4. NumPy: NumPy provides essential mathematical and array operations for preprocessing respiratory sound data. It is used for handling multi-dimensional arrays, enabling efficient feature manipulation and dataset transformations.

5. Pandas: Pandas is utilized for data management and preprocessing, particularly in handling structured patient records, respiratory sound metadata, and classification results. It ensures seamless data integration between different stages of the pipeline.

6. Librosa: Librosa is the primary signal processing library for feature extraction from respiratory sound recordings. It is used to compute Mel-Frequency Cepstral Coefficients (MFCCs), spectrograms, and other acoustic features, providing the CNN model with rich input data for classification.

7. Google Colab: RESPITECH is developed and deployed in Google Colab, allowing cloud-based execution with GPU acceleration for deep learning model training and testing. Google Colab facilitates collaborative research, rapid prototyping, and scalability, ensuring efficient experimentation with large respiratory sound datasets.

3.2 Design Details

The RESPITECH design is centered around a secure, scalable, and AI-driven environment that facilitates real-time respiratory sound classification and diagnosis assistance.

1. User Roles and System Interaction

-Healthcare Professionals: Medical experts interact with the platform by securely uploading respiratory audio recordings, reviewing AI-generated classification results, and utilizing diagnostic suggestions. Access is granted via a secure login system, ensuring data protection and regulatory compliance.
-Patients: Patients contribute respiratory sound recordings through approved mobile applications or medical devices. These samples are automatically analyzed by the RESPITECH system to detect possible respiratory abnormalities.

2. Authentication and Security Framework

To ensure the confidentiality and integrity of sensitive medical data, RESPITECH incorporates a **multi-layered authentication system**:
-User Authentication: Multi-Factor Authentication (MFA) is implemented for both clinicians and patients, reinforcing secure access.
-Data Protection: Strong encryption protocols are used for storing and transmitting user credentials and medical records, aligned with HIPAA and other healthcare data protection regulations.

3. AI-Powered Sound Analysis

At the core of RESPITECH is a **deep learning pipeline** that enables accurate and timely analysis of respiratory sounds:
-Feature Extraction: Key audio features such as **Mel-Frequency Cepstral Coefficients (MFCCs)** and **Mel-Spectrograms** are extracted from user-submitted audio samples.
-Model Architecture: A **CNN-based classifier**, optionally enhanced with BiLSTM layers, is used to detect and differentiate between conditions like **COPD, asthma, LRTI, and URTI.**

-Preprocessing: Noise filtering and data augmentation methods are applied to improve robustness and accuracy under diverse recording conditions.

4. Cloud-Based Infrastructure

RESPITECH is hosted on a **cloud platform** to ensure scalability, performance, and reliability:
-Data Management: All audio recordings, patient metadata, and diagnostic results are securely stored in the cloud.
-Scalability & Availability: The system architecture supports auto-scaling and load balancing, ensuring uninterrupted service and seamless access to healthcare providers across different geographies.
-Compliance & Disaster Recovery: The infrastructure is designed to meet healthcare data compliance standards (e.g., HIPAA, GDPR) and includes backup mechanisms for disaster recovery.

5. Diagnostic Feedback and Continuous Learning

-Clinical Feedback Interface: After analysis, detailed diagnostic reports are presented to medical professionals, including **classification results, probability scores**, and suggested clinical insights.
-Model Improvement: The system adopts a **continual learning approach**, updating its models based on new and diverse respiratory sound samples, thereby enhancing performance over time and reducing false positives or negatives.

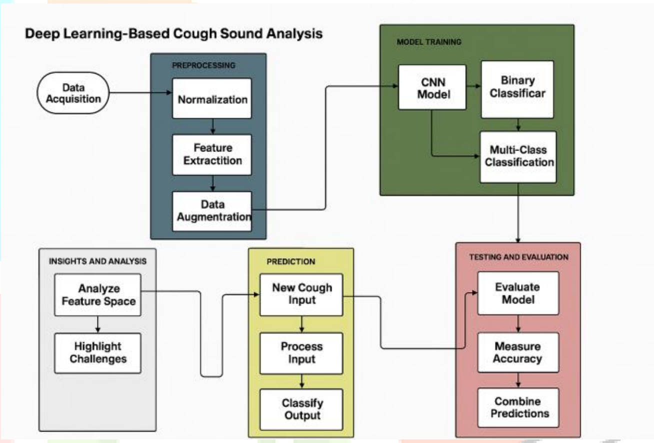

Fig. 1 Design Overview

IV. METHODOLOGY

An agile, iterative approach is used in the development of RESPITECH to guarantee adaptability, effectiveness, and quick deployment.

1. Data Acquisition: The dataset used in RESPITECH comprises lung sound recordings obtained from both healthy individuals and patients with respiratory dysfunctions such as COPD, asthma, lower respiratory tract infection (LRTI), and upper respiratory tract infection (URTI). The recordings include data from male and female subjects spanning all age groups, including children, adults, and senior citizens.

2. Data Preprocessing: Since the collected respiratory audio recordings were irregular in length and structure, preprocessing techniques were applied to standardize the data. Using the Librosa Python library, the raw audio files were:
-Trimmed or padded to a uniform length to ensure consistency across samples.
-Normalized to remove variations in amplitude caused by recording conditions. Converted to a consistent sampling rate to maintain feature uniformity.
-This preprocessing step ensures that the data is structured optimally for feature extraction and deep learning model training.

3. Feature Extraction: For feature extraction, Mel-Frequency Cepstral Coefficients (MFCCs) were used. MFCCs are widely used in speech and audio analysis because they represent the short-term power spectrum of sound. The MFCC extraction process involved:

-Applying the Fast Fourier Transform (FFT) to convert time-domain signals into the frequency domain.
-Mapping the frequency to the mel scale, which aligns better with human auditory perception.
-Computing the mel spectrogram, where the color dimension represents amplitude. Since MFCCs capture phonemes and sound variations based on vocal tract shape, they are highly effective for respiratory sound classification.

4. Data Augmentation: Due to class imbalance, particularly in the number of COPD versus non-COPD samples, audio augmentation techniques were applied to increase the diversity of non-COPD samples. Augmentation techniques included:
-Time-stretching: Adjusting the speed of audio samples without altering pitch.
-Pitch shifting: Modifying the frequency components to simulate natural variations.
-Adding noise: Introducing low-level background noise to improve model robustness. These augmentations help in preventing model overfitting and enhance generalization to real-world respiratory sound data.

5. CNN Model Architecture: The proposed Convolutional Neural Network (CNN) is designed using Keras with a TensorFlow back-end. The model follows a sequential architecture, consisting of:
-Input Layer: Accepts MFCC features as input.
-Convolutional Layers (Conv2D): Extracts spatial features from the input spectrogram.
-MaxPooling Layers: Reduces dimensionality and prevents overfitting.
-Dropout Layers: Regularization technique to improve generalization by randomly deactivating neurons.
-Fully Connected Dense Layer: Aggregates features for final classification. The CNN model effectively captures the patterns in respiratory sound signals, leading to high accuracy in classifying normal vs. pathological coughs.

6. Model Training and Evaluation
The CNN model was trained using:
-Cross-entropy loss function for multi-class classification.
Adam optimizer for adaptive learning rate adjustment.
-Stratified k-fold cross-validation to evaluate performance across different subsets of data. Performance metrics such as accuracy, precision, recall, and F1-score were used to assess the model's classification ability.

V. OUTPUT

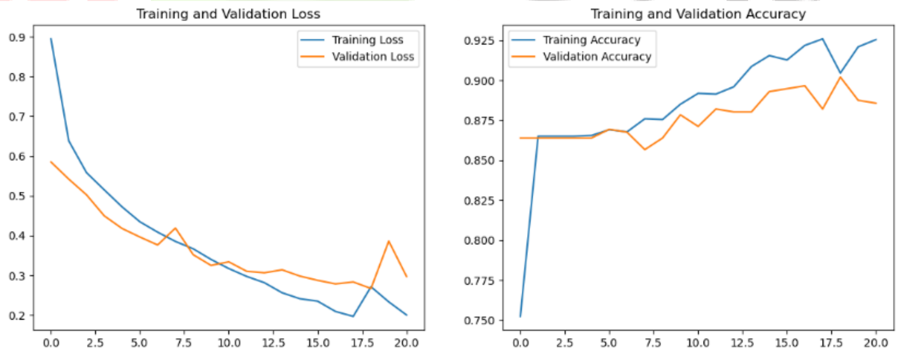

Fig. 2 Loss VS Accuracy

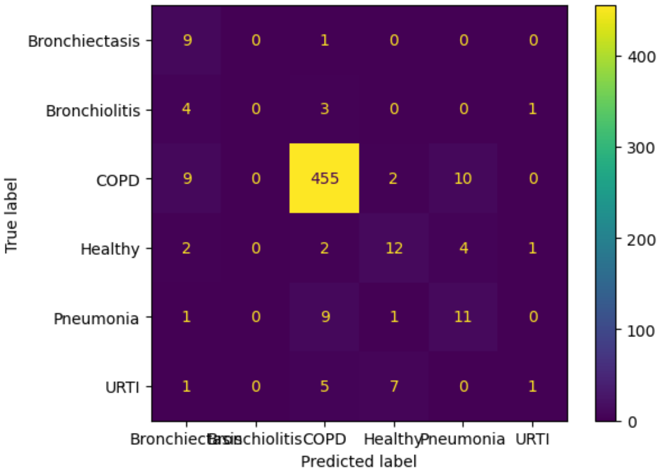

Fig.3 True Label VS Predicted Label

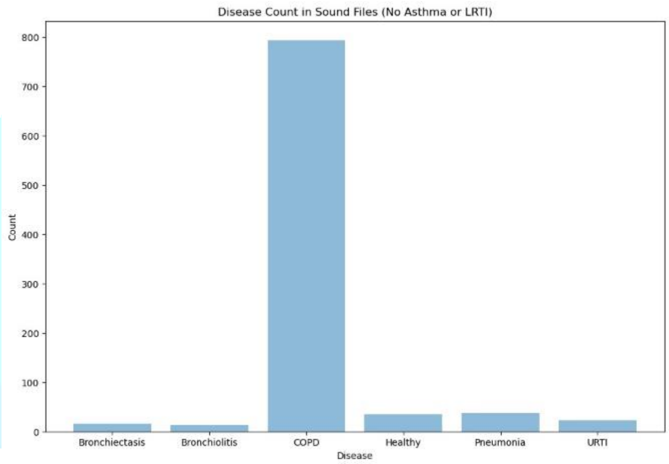

Fig.4 Disease Count in Sound Files

VI. Analysis Of Existing Challenges And RESPITECH's Contribution

The use of respiratory sound analysis for disease detection has been a growing area of research, yet several limitations have persisted across earlier systems. Traditional diagnostic approaches such as auscultation and spirometry, while widely practiced, rely heavily on subjective clinical judgment, require specialized equipment, and are often inaccessible in low-resource or remote settings. Furthermore, the similarity in acoustic patterns among different respiratory conditions—especially in early stages—makes accurate classification a challenging task, frequently resulting in misdiagnosis or delayed treatment.

Prior studies utilizing machine learning techniques like Support Vector Machines (SVM) and handcrafted features such as MFCCs have achieved promising results but still fall short in practical application. Many models struggle with noisy audio inputs, suffer from low generalization across diverse patient profiles, and lack the ability to capture the temporal dynamics of respiratory sounds. Additionally, most of these systems are not designed for real-time use or integration into scalable healthcare frameworks such as telemedicine platforms.

RESPITECH was designed to address these gaps directly. By combining CNN-based feature extraction with a BiLSTM network, it captures both spatial and temporal patterns in respiratory sounds, improving classification performance even in acoustically similar conditions like asthma, COPD, LRTI, and URTI. The system is further enhanced with real-time noise filtering and the evaluation of multiple cough segments per patient to reduce diagnostic errors.

In addition to its technical capabilities, RESPITECH is built with a focus on usability and accessibility. It leverages a secure, cloud-based infrastructure and supports remote data input through mobile or IoT devices, enabling deployment in real-world healthcare environments. By offering a fast, reliable, and automated analysis pipeline, RESPITECH contributes meaningfully to overcoming the diagnostic challenges faced in previous research—bringing us closer to scalable, AI-driven respiratory care solutions.

VII. CONCLUSION

In this study, we have presented **RESPITECH**, a lightweight and computationally efficient deep learning-based assistive framework designed to support healthcare professionals in the early detection of all the respiratory disease through analysis of respiratory sounds. The proposed system leverages a Convolutional Neural Network (CNN) architecture in combination with advanced audio feature extraction techniques, utilizing Librosa's powerful signal processing capabilities. Among the features evaluated—MFCC, Mel-Spectrogram, Chroma, Constant-Q Chroma, and Chroma CENS—MFCC demonstrated superior accuracy and robustness in classifying pathological respiratory patterns associated with them.

The promising results from our experiments affirm the potential of RESPITECH as a scalable and accessible diagnostic aid, especially in resource-constrained settings. The framework is used to detect a broader range of cardiopulmonary conditions, such as asthma, pneumonia, and even early signs of cardiac distress through heart sound analysis. Future iterations may also integrate disease severity estimation, advanced data augmentation strategies, and real-time breath monitoring to enhance diagnostic precision and user experience.

To ensure clinical adoption and trust, it will be essential to incorporate robust security protocols and privacy-preserving mechanisms, protecting patient data while resisting adversarial threats. As RESPITECH evolves, it holds significant promise for transforming remote healthcare delivery and advancing AI-assisted diagnostic tools in the field of respiratory medicine.

VIII. REFERENCES

[1] R. Palaniappan and K. Sundaraj, "Respiratory sound classification using cepstral features and support vector machine," Computers in Biology and Medicine, vol. 75, pp. 118–129, Mar. 2016.

[2] Ida Ayu Putu Ari Crisdayanti and Seong-Eun Kim.Multilevel Feature Fusion-based Convolutional Neural Network for Anomaly Classification of Respiratory Sound,IEEE International Conference 2022.

[3] J. Amoh and K. Odame, "Deep neural networks for identifying cough sounds," in Proc. IEEE Biomedical Circuits and Systems Conf. (BioCAS), 2016, pp. 1–4.

[4] J. Laguarta, F. Hueto, and B. Subirana, "COVID-19 artificial intelligence diagnosis using only cough recordings," IEEE Open Journal of Engineering in Medicine and Biology, vol. 1, pp. 275–281, 2020.

[5] Murat Aykanat ,Özkan Kılıç,Bahar Kurt and Sevgi Saryal.Classification of lung sounds using convolutional neural networks, EURASIP Journal 2017.

[6] Rizwana Zulfiqar,Fiaz Majeed,Rizwana Irfan,Hafiz Tayyab Rauf,Elhadj Benkhelifa and Abdelkader Nasreddine Belkacem.Abnormal Respiratory Sounds Classification Using Deep CNN Through Artificial Noise Addition,Pulmonary Medicine 2021.

[7] Arpan Srivastava,Sonakshi Jain,Ryan Miranda,Shruti Patil,Sharnil Pandya and Ketan Kotecha.Deep learning based respiratory sound analysis for detection of chronic obstructive pulmonary disease, Srivastava 2021.

[8] Ahmad H. Sabry,Omar I. Dallal Bashi,N.H. Nik Ali,Yasir Mahmood Al Kubaisi.Lung disease recognition methods using audio-based analysis with machine learning ,Elsevier 2024.

[9] Adithya Sunil Edakkadan and Abhishek Srivastava.Deep Learning Based Portable Respiratory Sound Classification System,IEEE 2023.

[10] Carlos A. Galindo Meza,Juan A. del Hoyo Ontiveros and Paulo Lopez-Meyer.Classification of Respiration Sounds Using Deep Pre-trained Audio Embeddings,IEEE 2021.

[11] Ankita Bose and B.K. Tripathy.Deep Learning for Audio Signal Classification,Research Gate 2020.

[12] Balamurali, B.T.; Hee, H.I.; Kapoor, S.; Teoh, O.H.; Teng, S.S.; Lee, K.P.; Herremans, D.; Chen, J.M. Deep Neural Network-Based Respiratory Pathology Classification Using Cough Sounds, Sensors 2021.