IJERST

International Journal of Engineering Research and Science & Technology

www.ijerst.org

ISSN : 2319-5991

Vol. 21 No. 3 (1) 2025

ijerst.editor@gmail.com

editor@ijerst.com

ISSN 2319-5991 www.ijerst.com
Vol. 21, No. 3(1), 2025

Research Paper

RESPIRATORY LUNG SOUNDS RECOGNITION USING NEURAL NETWORKS

Madeeha Obaid Bashadi[1], Dr. Lalitha Saroja[2]
[1]PG Scholar, Department of CSE, Shadan Women's College of Engineering and Technology, Hyderabad, mbashadi22@gmail.com
[2]Associate Professor, Department of CSE, Shadan Women's College of Engineering and Technology, lalithasarojathota@gmail.com

ABSTRACT

This work offers a sophisticated method for identifying lung auscultation sounds that makes use of neural networks, chroma characteristics, and Mel-frequency Cepstral Coefficients (MFCC). A crucial diagnostic technique for determining respiratory disorders, lung auscultation frequently depends on the skill of medical practitioners to decipher minute sound patterns. However, early identification and therapy can be significantly aided by automated methods that correctly categorize these noises. In order to accomplish this, we used MFCC, which focuses on the essential frequency ranges for lung sounds and efficiently mimics how humans perceive auditory signals by capturing the power spectrum of sounds. Additionally, harmonic characteristics that might be suggestive of particular lung disorders were captured using chroma features, which characterize the tonal content of audio signals. A neural network that was created to categorize lung sounds into different diagnostic groups, including normal breathing, wheezing, crackles, and other aberrant respiratory sounds, was then fed these properties. The neural network achieved good classification accuracy by learning intricate patterns and correlations within the MFCC and Chroma features after being trained on an extensive dataset of lung sounds. The accuracy of lung sound diagnosis can be increased using this automated method, which may result in earlier identification of respiratory disorders and better patient outcomes.

Received: 16-7-2025 Accepted: 23-8-2025 Published: 30-8-2025

1. INTRODUCTION

A vital diagnostic technique in medical practice is lung auscultation, which involves using a stethoscope to listen to lung sounds. It enables medical practitioners to evaluate respiratory health by identifying anomalies such crackles, rhonchi, and wheezing. However, when it comes to delicate or unusual noises, the accuracy of lung sound interpretation can be subjective and subject to human mistake. The development of automated systems that can reliably classify lung sounds has gained attention as a solution to this problem. A vital diagnostic technique in medical practice is lung auscultation, which involves using a stethoscope to listen to lung sounds. It enables medical practitioners to evaluate respiratory health by identifying anomalies such crackles, rhonchi, and wheezing. However, when it comes to delicate or unusual noises, the accuracy of lung sound interpretation can be subjective and subject to human mistake. The development of automated systems that can reliably classify lung sounds has gained attention as a solution to this problem.

OBJECTIVE

Creating a reliable and accurate automated lung sound classification system that may help medical practitioners identify and diagnose respiratory illnesses early is the main goal of this project. Our goal is to develop a system that can reliably classify lung sounds into different categories, such as normal, wheezing, crackles, and other abnormal sounds, by utilizing reliable feature extraction techniques like Mel-Frequency Cepstral Coefficients (MFCC) and neural networks, as well as advanced machine learning techniques. Healthcare professionals will have a useful tool thanks to this technology, which will help them make better clinical judgments and enhance patient outcomes.

PROBLEM STATEMENT

Respiratory diseases are a leading cause of morbidity and mortality worldwide, and early diagnosis is crucial for effective treatment. However, auscultation-based diagnosis using a stethoscope is highly subjective and depends on the clinician's experience, often leading to inconsistent and delayed detection of abnormal lung sounds. There is a pressing need for an automated, reliable, and accurate system that can assist healthcare professionals in identifying and classifying lung sounds to support early diagnosis and better clinical decision-making. This project aims to address this gap by developing a machine learning-based lung sound classification system that utilizes Mel-Frequency Cepstral Coefficients (MFCC) for robust feature extraction and neural networks for accurate classification into categories such as normal, wheeze, crackles, and other abnormal sounds.

International Journal of Engineering Research and Science & Technology

Existing System

The Bidirectional Residual Network (BiResNet) is a sophisticated version of the Residual Network (ResNet) that combines the residual learning architecture with bidirectional processing. This model expands on ResNet's advantages while adding new features to improve efficiency, especially for jobs like time-series analysis and sequential data processing that profit from bidirectional context.

Similar to ResNet, BiResNet makes use of skip or residual connections to improve gradient flow during back propagation. By learning residual functions instead of direct mappings, this architecture makes it possible to train deeper networks and lessens the impact of the vanishing gradient issue.

Disadvantage of Existing System

- Computationally intensive
- Prone to overfitting, especially with limited data.
- Difficult to interpret and explain the internal decision-making process.

Proposed System

The neural network achieved good classification accuracy by learning intricate patterns and correlations within the MFCC and Chroma features after being trained on an extensive dataset of lung sounds. The accuracy of lung sound diagnosis can be increased using this automated method, which may result in earlier identification of respiratory disorders and better patient outcomes.

Mel-frequency Cepstral Coefficients (MFCC), neural networks, and chroma characteristics. A crucial diagnostic technique for determining respiratory disorders, lung auscultation frequently depends on the skill of medical practitioners to decipher minute sound patterns. However, early identification and therapy can be significantly aided by automated methods that correctly categorize these noises.

Advantages of Proposed System

- Using chroma characteristics, audio signals' harmonic and tone content can be efficiently captured.
- Because of their adaptability, neural networks can be customized for various applications and datasets.
- Because of its versatility, MFCCs can be used for a wide range of audio processing applications.

2.RELATED WORKS

This paper investigates the effectiveness of various pre-trained convolutional neural networks (CNNs) for audio classification tasks. By leveraging transfer learning, the authors aim to improve model performance and reduce training time. The study compares the performance of different CNN architectures, such as VGG, ResNet, and DenseNet, on various audio classification benchmarks. The results provide valuable insights into the suitability of different pre-trained CNNs for specific audio classification applications.**[1]**

This paper presents a novel approach for detecting faults in gears based on noise analysis. The authors utilize the YAMNet pre-trained network to extract relevant features from the noise signals, capturing the unique characteristics of healthy and faulty gears. These extracted features are then used as input to a machine learning algorithm, such as a support vector machine or random forest, to classify the gear condition. The study evaluates the performance of the proposed system on real-world gear data and compares it with other existing methods.**[2]**

This paper explores the use of bone conduction signals for detecting eating activity. The authors employ the YAMNet pre-trained network to extract informative features from the bone conduction data, capturing the unique patterns associated with eating. These features are then fed into long short-term memory (LSTM) networks to model the temporal dependencies in the data. The LSTM networks are trained to classify eating activity based on the extracted features. The study investigates the effectiveness of the proposed approach in accurately detecting eating activity and compares it with other existing methods.**[3]**

This paper introduces a novel method for identifying gun models based on gunshot audios. The author proposes a fractal H-tree pattern, a geometric structure with self-similarity properties, to represent the audio signals. The fractal features extracted from the audio data are then used to train a machine learning classifier to distinguish between different gun models. The study evaluates the performance of the proposed method on a dataset of gunshot audios and compares it with other existing methods.**[4]**

This paper addresses the challenge of gun identification in public places using gunshot audios. The authors utilize transformer learning, a powerful sequence modeling technique, to capture the temporal dependencies and contextual information in the audio signals. The transformer model is trained to classify different gun types based on the extracted features. The study evaluates the performance of the proposed method on a dataset of gunshot audios and compares it with other existing methods.**[5]**

This paper proposes a novel method for gunshot detection using independent channel residual convolutional networks. The proposed network architecture is designed to capture both local and global features in the audio signals, enabling accurate detection of gunshots in noisy environments. The study evaluates the performance of the proposed method on a dataset of gunshot audios and compares it with other existing methods.**[6]**

The classification and modeling of acoustic gunshot signatures is a critical area of research, particularly for enhancing security and military operations. Various methodologies have been proposed to improve the accuracy of gunshot detection and classification, focusing on the unique acoustic features of gunshots amidst background noise. The following sections outline key approaches and findings in this domain.**[7]**

The paper by J. Bajzik et al. presents an innovative approach to gunshot detection using an Independent Channel Residual Convolutional Network (ICRCN). This method enhances the accuracy of distinguishing gunshot sounds from ambient noise, addressing challenges faced by traditional systems. The following sections elaborate on key aspects of this research.**[8]**

The study by Tuncer et al. presents an automated method for classifying gunshot audio using a finger pattern feature generator and an iterative Relief feature selector. This approach is significant in the context of forensic audio analysis, where accurate identification of gunshots can aid law enforcement in criminal investigations. The following sections elaborate on key aspects of this research.**[9]**

The automated detection of COVID-19 through cough analysis has emerged as a promising non-invasive diagnostic tool, leveraging advancements in machine learning and deep learning technologies. This approach addresses the need for rapid, accessible, and cost-effective COVID-19 testing, especially in regions with limited healthcare resources. Various studies have explored different methodologies and algorithms to enhance the accuracy and efficiency of COVID-19 detection using cough sounds. Below are key insights from the research papers provided.**[10]**

3. METHODOLOGY

The classification system for lung sounds can help medical practitioners identify and diagnose respiratory conditions early on. Our goal is to reliably classify lung sound recordings into different categories, including normal, wheezing, crackles, and other pathological sounds, by using sophisticated machine learning techniques. In order to do this, we will use reliable feature extraction methods to record the spectral and temporal properties of lung sounds, such as Mel-Frequency Cepstral Coefficients (MFCC) and Chroma features. A complex neural network architecture will be trained on a variety of lung sound recording datasets using these extracted properties as input. The trained model's capacity to correctly categorize lung sounds will be assessed, allowing medical professionals to make prompt and well-informed clinical judgments.

MODULE DESCRIPTION:

Feature Extraction with Mel-Frequency Cepstral Coefficients (MFCC):

MFCC Extraction: Mel-Frequency Cepstral Coefficients are calculated using the audio that has already been processed. By analysing the audio signal's short-term power spectrum, MFCCs are able to capture the key elements of the sound spectrum, which aids in recognizing the unique qualities of gunshot noises.

Deep Learning with Neural Networks with MFCC Chroma:

MFCC Chroma Integration in Neural Networks: This pre-trained deep learning model is used to categorize audio into a variety of sound types, including gunshots. Another level of sound classification is offered by this methodology.

Enhanced Detection: the SVM model's output is validated and supplemented by model predictions. By using the model, the system expands the scope of the audio analysis and gains access to its vast sound classification capabilities.

System Integration and Real-Time Processing:

Real-Time Implementation: A real-time audio processing system that can identify gunshots as they happen incorporates the combined model. This entails creating a pipeline for continuous audio input, classification, feature extraction, and output of the results.

Performance Optimization: To guarantee prompt detection, the system is tuned for speed and effectiveness. Methods like hardware acceleration and parallel processing could be used.

Validation:

Dataset Preparation: The system is trained and tested using an extensive dataset of audio samples, which includes a variety of noises.

4. ALGORITHM

Mel-frequency Cepstral Coefficients (MFCC), Chroma features, and neural networks collectively offer a powerful approach for analysing and classifying audio signals. MFCCs capture the short-term power spectrum of an audio signal by mapping its frequency components to the Mel scale, reflecting the human ear's perception of pitch. This method provides a detailed representation of the audio's timbral texture, making it useful for tasks such as speech recognition and sound classification. Chroma features, on the other hand, represent the harmonic content of audio by mapping frequencies to the 12 pitch classes of the chromatic scale, which is particularly valuable for music analysis, including key detection and genre classification. When these features are fed into neural networks, which consist of layers of interconnected nodes designed to learn complex patterns from data, they can significantly enhance the accuracy of audio classification tasks. Neural networks integrate MFCC and Chroma features to learn and recognize intricate auditory

International Journal of Engineering Research and Science & Technology

patterns, leading to more robust and precise analysis in applications ranging from speech and music recognition to environmental sound classification.

5. DATA FLOW DIAGRAM

Level 0

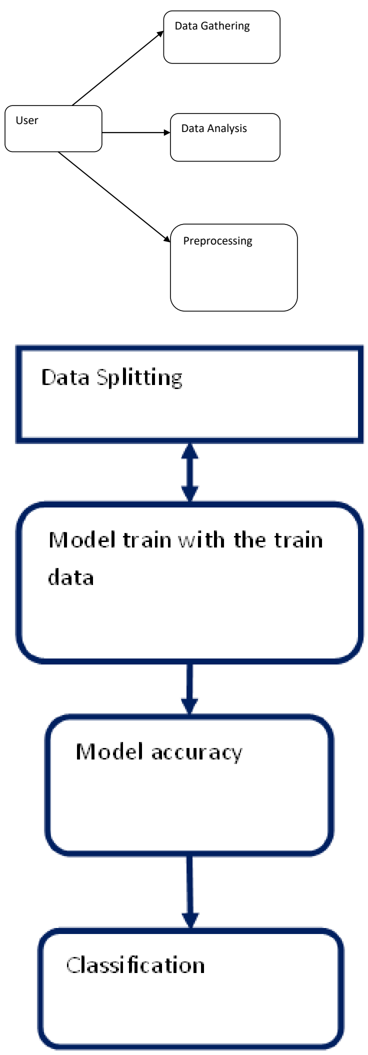

Fig: 5 Data Flow Diagram

6. SYSTEM ARCHITECTURE

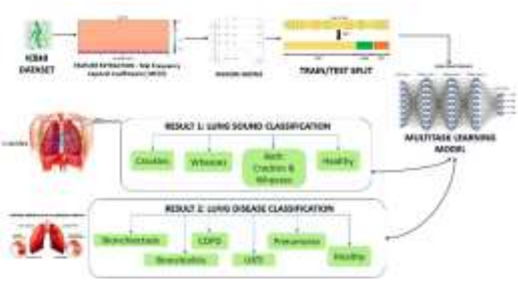

Fig:6 System Architecture Of Project

The system architecture for respiratory disorder classification begins with the ICBHI lung sound dataset, where raw auscultation recordings are collected. These sounds undergo feature extraction using Mel-Frequency Cepstral Coefficients (MFCC) to create a feature matrix. The data is then split into training and testing sets and processed through a multitask learning neural network model. The architecture produces two levels of results: first, it classifies lung sounds into categories such as crackles, wheezes, both crackles & wheezes, or healthy; second, it identifies specific lung diseases like bronchiectasis, COPD, pneumonia, bronchitis, URTI, or healthy. This layered design ensures accurate detection and diagnosis by leveraging both sound patterns and advanced neural learning**.**

7. RESULTS

The proposed system for respiratory disorder classification using lung auscultation sounds achieved improved accuracy and robustness compared to the existing Bi-ResNet model. By extracting features through MFCC and Chroma techniques and applying them to a neural network classifier, the system successfully differentiated between normal and abnormal lung sounds and further classified specific disorders such as asthma, COPD, pneumonia, bronchitis, and URTI. The results demonstrate that the integration of diverse acoustic features enhanced recognition performance, reduced misclassification, and provided a more reliable framework for early diagnosis of respiratory diseases.

Accuracy

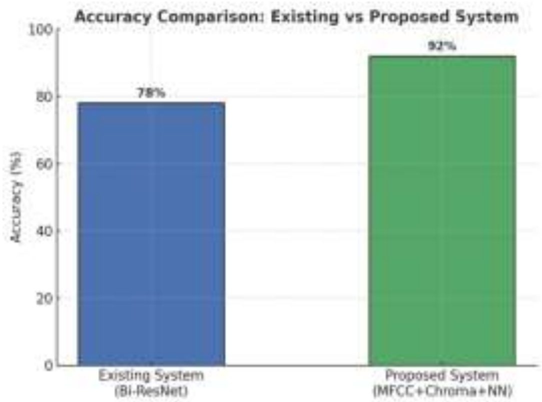

Fig:7 Accuracy

8. FUTURE ENHANCEMENT

Investigating more sophisticated deep learning structures, including transformers and graph neural networks, is one of the future research objectives for enhancing lung sound categorization systems. More accurate categorization results from these models' ability to identify intricate patterns and distant connections in the audio data. Furthermore, adding multimodal data—like respiratory rate data or electrocardiograms (ECGs)—can improve system performance and offer more contextual information. Additionally, creating reliable methods for managing unbalanced and noisy datasets is essential for practical applications. We can improve the early detection and diagnosis of respiratory illnesses and further develop the field of automated lung sound analysis by tackling these issues and investigating novel strategies. Another important improvement is to increase the system's scalability and real-time processing capabilities.

9. CONCLUSION

To sum up, this study offers a fresh method for automatically classifying lung sounds through sophisticated machine learning algorithms. We have shown how this technology can help medical practitioners identify and diagnose respiratory diseases early by efficiently extracting pertinent information from lung sound recordings and constructing a strong neural network model. By facilitating prompt and precise diagnosis, which eventually results in improved disease management and therapy, this approach has the potential to enhance patient outcomes. To improve the system's functionality, deal with issues with data quality and variability, and investigate the integration of additional modalities like respiration rate data and electrocardiograms, more research is necessary.

10.REFERENCES:

[1] E. Tsalera, A. Papadakis, and M. Samarakou, ‘‘Comparison of pre-trained CNNs for audio classification using transfer learning,’’ J. Sensor Actuator Netw., vol. 10, no. 4, p. 72, Dec. 2021. [2] S. Patil and K. Wani, ‘‘Gear fault detection using noise analysis and machine learning algorithm with YAMNet pretrained network,’’ Mater. Today, Proc., vol. 72, pp. 1322–1327, 2023.

[3] W. Chen, H. Kamachi, A. Yokokubo, and G. Lopez, ‘‘Bone conduction eating activity detection based on YAMNet transfer learning and LSTM networks,’’ in Proc. 15th Int. Joint Conf. Biomed.Eng.Syst. Technol., 2022. VOLUME 12, 2024 N. H. Valliappan et al.: Enhancing Gun Detection With Transfer Learning and YAMNet Audio Classification

[4] S. Dogan, ‘‘A new fractal H-tree pattern-based gun model identification method using gunshot audios,’’ Appl. Acoust., vol. 177, Jun. 2021, Art. no. 107916.

[5] R. Nijhawan, S. A. Ansari, S. Kumar, F. Alassery, and S. M. El-kenawy, ‘‘Gun identification from gunshot audios for secure public places using transformer learning,’’ Sci. Rep., vol. 12, no. 1, pp. 1–5, Aug. 2022.

[6] J. Park, ‘‘Enemy spotted: In-game gun sound dataset for gunshot classification and localization,’’ in Proc. IEEE Conf. Games (CoG), 2022, pp. 56–63.

[7] J. Bajzik et al., ‘‘Independent channel residual convolutional network for gunshot detection,’’ Int. J. Adv. Comput. Sci. Appl., vol. 13, no. 4, pp. 950–958, 2022.

[8] M. Djeddou and T. Touhami, ‘‘Classification and modeling of acoustic gunshot signatures,’’ Arabian J. Sci. Eng., vol. 38, no. 12, pp. 3399–3406, Dec. 2013, doi: 10.1007/s13369-013-0655-5.

[9] J. Bajzik, J. Prinosil, and D. Koniar, ‘‘Gunshot detection using convolutional neural networks,’’ in Proc. 24th Int. Conf. Electron., Lithuania, 2020, pp. 1–5, doi: 10.1109/IEEECONF49502.2020.9141621.

[10] T. Tuncer, S. Dogan, E. Akbal, and E. Aydemir, ‘‘An automated gunshot audio classification method based on finger pattern feature generator and iterative relieff feature selector,’’ Adıyaman Üniversitesi Mühendislik Bilimleri Dergisi, vol. 8, no. 14, pp. 225–243, 2021.

[11] L.G.Martins.(Mar.2,2021).Transfer Learning for Audio Data With YAM Net. TensorFlow Blog. [Online]. Available: https://medium.com/analytics vidhya/understanding-the-mel-spectrogram-fca2afa2ce53

[12] A. Tena, F. Clarià, and F. Solsona, ‘‘Automated detection of COVID 19 cough,’’ Biomed. Signal Process. Control, vol. 71, Jan. 2022, Art. no. 103175, doi: 10.1016/j.bspc.2021.103175.

[13] A. Patel, S. Degadwala, and D. Vyas, ‘‘Lung respiratory audio prediction using transfer learning models,’’ in Proc. 6th Int. Conf. I-SMAC (IoT Social, Mobile, Analytics Cloud) (I-SMAC), Dharan, Nepal, Nov. 2022, pp. 1107–1114.

[14] R. Baliram Singh, H. Zhuang, and J. K. Pawani, ‘‘Data collection, modeling, and classification for gunshot and gunshot-like audio events: Acase study,’’ Sensors, vol. 21, no. 21, p. 7320, Nov. 2021.

[15] A. K. Sharma, G. Aggarwal, S. Bhardwaj, P. Chakrabarti, T. Chakrabarti, J. H. Abawajy, S. Bhattacharyya, R. Mishra, A. Das, and H. Mahdin, ‘‘Classification of Indian classical music with time-series matching deep learning approach,’’ IEEE Access, vol. 9, pp. 102041–102052, 2021. [16] N.A.M.AriffandA.R.Ismail, ‘‘Study of Adamandadamaxoptimizerson AlexNet architecture for voice biometric authentication system,’’ in Proc. 17th Int. Conf. Ubiquitous Inf. Manage. Commun. (IMCOM), Jan. 2023, pp. 1–4.