

AI – Driven Audio Signal Analysis for Healthcare Monitoring

Mr. Prathamesh Patil, Ms. Swarada Athave, Ms. Sakshi Kumbhar, Mr. Nikhil Patra, Ms. Tahura Hakim

Department Of Computer Science And Engineering

Under The Guidance Of

Prof.S.R.Patil

Abstract -

Nowadays, healthcare systems are increasingly using Artificial Intelligence (AI) and Machine Learning (ML) technologies for smart disease detection and patient monitoring. This paper presents an AI-based cough detection system that analyze audio signals to identify cough and non-cough sounds. The proposed system uses both pre-recorded cough datasets and real-time microphone input for prediction. To improve the quality of audio signals, preprocessing techniques such as noise reduction, normalization, feature extraction, and spectrogram generation are applied. Different algorithms including Convolutional Neural Network (CNN), Random Forest, and Logistic Regression are implemented and compared to achieve better classification accuracy. Among these algorithms, CNN provides better performance because it can automatically learn complex audio patterns from spectrogram images. The developed system supports real-time cough detection, making it useful for healthcare monitoring and respiratory disease analysis. The proposed model provides a simple, fast, and cost-effective solution for intelligent healthcare applications and can be further enhanced for remote patient monitoring and telemedicine systems.

Key Words: Artificial Intelligence, Machine Learning, CNN, Cough Detection, Audio Signal Analysis, Healthcare Monitoring.

1.INTRODUCTION :

Artificial Intelligence (AI) and Machine Learning (ML) are rapidly transforming the healthcare industry by enabling smart, automated, and efficient systems for disease detection and patient monitoring. In recent years, there has been a growing demand for intelligent healthcare solutions that can assist in early diagnosis and

reduce the dependency on manual medical analysis. Respiratory diseases such as Asthma, Bronchitis, pneumonia, tuberculosis, and viral infections are increasing worldwide, and cough is one of the most common symptoms associated with these conditions. Early detection and continuous monitoring of cough sounds can significantly help in identifying respiratory health issues at an early stage and improving treatment outcomes.

Traditional healthcare systems rely heavily on clinical examinations, laboratory tests, and continuous observation by medical professionals. These methods are often time-consuming, expensive, and not always accessible in remote or rural areas. Moreover, manual monitoring of cough symptoms is not always practical due to the need for constant supervision. This creates a strong need for an automated, accurate, and cost-effective system that can analyze cough sounds using modern technologies.

This paper presents an AI-based cough detection system using audio signal processing and machine learning techniques. The system is designed to classify audio samples into two categories: cough and non-cough. It utilizes both dataset-based audio inputs and real-time microphone recordings to make the system more practical and user-friendly. To improve performance, audio preprocessing techniques such as noise reduction, normalization, feature extraction, and spectrogram generation are applied, which help in extracting meaningful features from audio signals.

In this project, different algorithms such as Convolutional Neural Network (CNN), Random Forest, and Logistic Regression are implemented and compared for performance analysis. CNN is used for deep learning-based feature extraction from spectrogram images, while Random Forest and Logistic Regression are used for traditional machine learning-based classification. The performance of these models is evaluated using metrics such as accuracy, precision, recall, and F1-score.

The system also supports real-time cough detection, making it suitable for continuous health monitoring applications. The proposed solution provides a fast, automated, and cost-effective approach for respiratory disease analysis. It can assist healthcare professionals in early diagnosis and reduce manual effort in patient monitoring. In future, this system can be extended to detect disease-specific cough patterns and integrated with telemedicine and smart healthcare devices for remote monitoring.

The system is designed to work efficiently even in real-world noisy environments by applying advanced audio preprocessing techniques. It aims to bridge the gap between healthcare and artificial intelligence by providing an accessible and intelligent diagnostic support tool. This approach helps in improving early detection systems and enhances remote healthcare monitoring capabilities.

2. Body of Paper :

2.1 Literature Review:

The use of Artificial Intelligence (AI) and Machine Learning (ML) in healthcare has increased significantly in recent years, especially for disease detection and patient monitoring. Researchers have focused on developing automated systems for analyzing respiratory sounds such as cough, breathing, and speech signals. These systems aim to support early diagnosis of respiratory diseases like asthma, pneumonia, tuberculosis, bronchitis, and COVID-19. Audio signal processing plays an important role in extracting useful features from sound data for classification and analysis.

Many traditional approaches use machine learning algorithms such as Logistic Regression, Support Vector Machine (SVM), Decision Tree, and Random Forest. These models generally depend on manually extracted features like MFCC (Mel Frequency Cepstral Coefficients), spectral features, and time-domain characteristics. Among these, Random Forest is widely used because it improves accuracy by combining multiple decision trees and reduces overfitting problems.

In recent studies, deep learning techniques have shown better performance compared to traditional methods. Convolutional Neural Networks (CNNs) are commonly used for analyzing spectrogram images of audio signals. CNN models automatically learn important features from data without manual feature extraction, making them more efficient and accurate. Many research works have proven that CNN-based models provide higher accuracy

in cough detection tasks due to their ability to capture complex audio patterns.

Researchers have also developed hybrid approaches that combine machine learning and deep learning techniques to improve performance. In addition, real-time cough detection systems using smartphones. These systems became more important during the COVID-19 pandemic for quick and contactless screening.

However, existing systems still face challenges such as background noise, limited dataset size, and variation in cough sounds among individuals. These issues reduce accuracy and generalization performance. Therefore, there is a need for a more robust and efficient system. The proposed system addresses these limitations by using CNN, Random Forest, and Logistic Regression along with real-time audio processing to improve accuracy and reliability.

2.2 Proposed System:

The proposed system aims to develop an AI-based cough detection model using audio signal analysis and machine learning techniques. The system is designed to classify audio inputs into two categories: cough and non-cough. It works on both pre-recorded dataset audio and real-time microphone input, making it suitable for practical healthcare monitoring applications.

The overall methodology of the system consists of several stages. First, audio data is collected from a cough dataset as well as real-time recording using a microphone. The collected audio signals are then preprocessed to improve quality and remove unwanted noise. Preprocessing techniques such as noise reduction, normalization, silence removal, and resampling are applied. After preprocessing, feature extraction is performed using methods like Mel Frequency Cepstral Coefficients (MFCC) and spectrogram generation, which convert audio signals into meaningful numerical representations suitable for model training.

In the next stage, the extracted features are used to train machine learning and deep learning models. The system implements three main algorithms: Convolutional Neural Network (CNN), Random Forest, and Logistic Regression. CNN is used for deep learning-based classification using spectrogram images, while Random Forest and Logistic Regression are used as traditional machine learning classifiers for comparison. The dataset

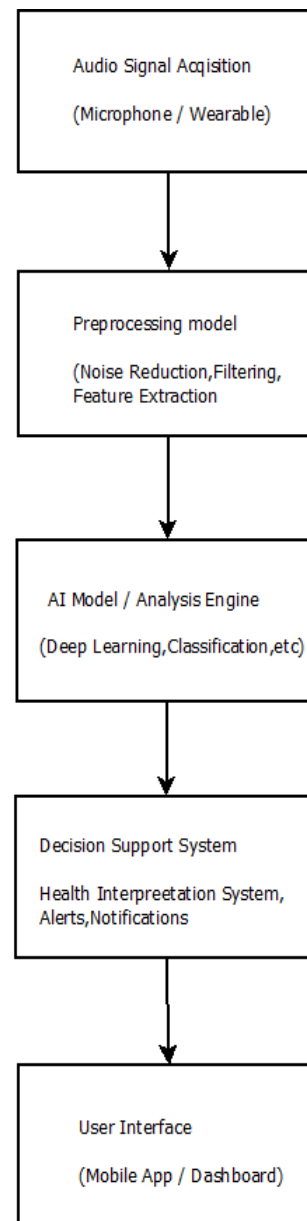
is divided into training and testing sets to evaluate model performance.

After training, the models are tested using unseen audio samples and real-time inputs. The system predicts whether the input audio contains a cough sound or not. The performance of each model is evaluated using metrics such as accuracy, precision, recall, and F1-score. This comparison helps identify the most effective algorithm for cough detection.

Finally, the real-time module captures audio through a microphone and processes it through the trained model to provide instant predictions. The proposed system is efficient, scalable, and suitable for healthcare applications such as remote patient monitoring and early respiratory disease detection. It demonstrates the practical use of Artificial Intelligence and audio signal processing in building intelligent healthcare solutions.

The system is designed to be scalable and can be easily extended for detecting other respiratory sound patterns in future enhancements. It also ensures faster processing and reliable predictions, making it suitable for real-time healthcare applications and smart monitoring systems. Furthermore, the integration of multiple machine learning models enhances the reliability and robustness of the system by enabling performance comparison and selection of the best-performing approach. This work highlights the potential of AI-driven audio analysis in improving early diagnosis and preventive healthcare solutions.

2.3 System Architecture:



2.4 Methodology:

The methodology of the proposed cough detection system is designed to process audio signals, extract meaningful features, and classify them using machine learning and deep learning models. The overall process consists of data collection, preprocessing, feature extraction, model training, and prediction.

2.4.1 Data Collection :

The audio dataset consists of cough and non-cough sound recordings collected from publicly available datasets as well as real-time audio samples recorded using

a microphone. Each audio sample is labeled as either “cough” or “non-cough” to enable supervised learning. The dataset includes variations in noise levels, speakers, and recording environments to improve model robustness.

2.4.2 Audio Preprocessing:

Raw audio signals cannot be directly used for machine learning models; therefore, preprocessing is applied. This includes:

- Conversion of audio into a uniform sampling rate
- Noise reduction to remove background interference
- Normalization of audio signals
- Trimming of silent portions from recordings

These steps ensure that the audio data is clean and consistent for feature extraction.

2.4.3 Feature Extraction:

To convert audio signals into a machine-readable format, feature extraction is performed. The primary feature used in this study is Mel-Frequency Cepstral Coefficients (MFCC), which captures the frequency characteristics of audio signals.

In addition, spectrogram representations are generated for training the CNN model. These features help in converting audio signals into numerical and visual formats suitable for classification.

2.4.4. Model Training:

Three different models are trained and evaluated in this study:

1.Logistic Regression: Used as a baseline linear classifier for MFCC features.

2,Random Forest Classifier: An ensemble model that improves classification accuracy by combining multiple decision trees.

3.Convolutional Neural Network (CNN): A deep learning model trained on spectrogram images to automatically learn spatial features from audio data.

2.4.5 Real-Time Prediction System:

After training, the best-performing model is deployed for real-time prediction. The system captures live audio input through a microphone, processes it using the same preprocessing and feature extraction pipeline, and then feeds it into the trained model. The model then classifies the input audio as either “cough” or “non-cough” and displays the result instantly.

2.4.6 Evaluation Process:

The performance of all models is evaluated using standard metrics such as accuracy, precision, recall, F1-score, and confusion matrix. These metrics help in comparing the effectiveness of different models and selecting the most suitable one for cough detection.

3. Future Work:

- Increase the dataset size with more cough and non-cough audio samples from different environments, age groups, and health conditions to improve model accuracy and generalization.
- Extend the system from binary classification (cough / no cough) to multi-class classification such as dry cough, wet cough, and disease-specific cough (e.g., COVID-19, tuberculosis).
- Implement advanced deep learning models like LSTM, GRU, and Transformer-based architectures to improve sequential audio feature learning.
- Improve real-time performance by optimizing model size and reducing computation time for faster predictions.
- Deploy the system on mobile devices or wearable health monitoring systems for continuous real-time tracking.
- Integrate cloud-based storage and processing for large-scale health monitoring and remote patient analysis.
- Enhance noise robustness by training the model with more noisy and real-world audio samples.

4.Conclusion:

The proposed cough detection system presents an effective approach for automated respiratory sound analysis using machine learning and deep learning techniques. The system is designed to classify audio inputs into cough and non-cough categories by leveraging

multiple classification models, including Logistic Regression, Random Forest, and Convolutional Neural Networks (CNN). The workflow includes essential stages such as audio preprocessing, noise reduction, normalization, and feature extraction using Mel-Frequency Cepstral Coefficients (MFCC) and spectrogram representations.

The comparative analysis of models demonstrates that traditional machine learning algorithms such as Logistic Regression and Random Forest perform reasonably well on engineered features; however, they are limited in capturing complex temporal and spatial patterns present in audio data. In contrast, the CNN model achieves superior performance by automatically learning hierarchical feature representations from spectrogram images, making it more suitable for audio-based classification tasks.

Furthermore, the system is capable of handling both offline dataset evaluation and real-time audio input, which enhances its practical usability in real-world health monitoring scenarios. The results indicate that AI-based audio classification systems can significantly contribute to early detection and monitoring of respiratory conditions, thereby supporting healthcare applications.

Despite the promising results, the system still has limitations such as dependency on dataset quality, sensitivity to background noise, and limited generalization across diverse real-world conditions. These challenges highlight the need for more robust training datasets and improved model architectures.

Overall, this work establishes a strong foundation for intelligent cough detection systems and demonstrates the potential of deep learning in healthcare audio analytics. Future enhancements such as large-scale dataset integration, lightweight model deployment, and mobile-based real-time monitoring systems can further improve the applicability and efficiency of the proposed approach. Overall, the proposed system demonstrates that machine learning and deep learning techniques can effectively be applied to audio-based healthcare applications with high reliability. This work can serve as a foundation for developing more advanced and real-time intelligent health monitoring systems in the future.

5. References:

1. T. Giannakopoulos and A. Pikrakis, *Introduction to Audio Analysis: A MATLAB Approach*, Academic Press, 2014.
2. B. McFee et al., “librosa: Audio and Music Signal Analysis in Python,” *Proceedings of the 14th Python in Science Conference*, 2015.
3. I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, MIT Press, 2016.
4. S. B. Davis and P. Mermelstein, “Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences,” *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 1980.
5. L. Breiman, “Random Forests,” *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
6. K. Simonyan and A. Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition,” *ICLR*, 2015.
7. T. Heittola et al., “Audio Set and Audio Event Detection,” *IEEE ICASSP*, 2018.
8. Scikit-learn Developers, “Scikit-learn: Machine Learning in Python,” <https://scikit-learn.org>