International Research Journal of Education and Technology
Peer Reviewed Journal
ISSN 2581-7795

AN INTELLIGENT FRAMEWORK FOR MULTISOUND DETECTION AND ENHANCEMENT IN DIGITAL STETHOSCOPES

Shreyaa V, Zameer Ali S, Pragadeesh R, Dr.Deepa D

[1]*Studuent, Dept. of Biomedical Engineering, Anna University, IN*
[2]*Studuent, Dept. of Biomedical Engineering, Anna University, IN*
[3] *Studuent, Dept. of Biomedical Engineering, Anna University, IN*
[4] *Head of the Departmentt, Dept. of Biomedical Engineering, Anna University, IN*

---***---

Abstract - *Technological advancements in digital signal processing and AI have enhanced auscultation-based diagnostics. Conventional stethoscopes function but suffer from ambient noise and expert understanding. The desire for smart and automated diagnostic systems resulted in digital stethoscopes that provide better analysis of heart and lung sounds. This research seeks to overcome the deficiencies of traditional auscultation by combining machine learning with digital stethoscopes. The system applies Gaussian noise filtering and spectral subtraction for the elimination of noise and NMF for separation of sound. Subsequently, heart and lung disease classification is conducted based on derived features using a CNN-LSTM. An LLM based on GPT-2 offers interactive explanations of findings for enhanced understanding. Experimental outcomes indicate 91.46% accuracy which is significantly higher than with traditional auscultation techniques. The system is able to distinguish between heart and lung sounds and can accurately diagnose. The AI driven explanations further enhance the systems application in clinical practice. These outcomes demonstrate the capability of AI driven digital stethoscopes in non-invasive diagnosis and decision making for medical professionals.*

Key Words: *Digital stethoscope, CNN-LSTM, sound separation, signal processing, disease classification, AI.*

1.INTRODUCTION

Medical diagnostics has witnessed tremendous changes with the introduction of artificial intelligence (AI) and sophisticated signal processing methods. Auscultation, a basic diagnostic technique, has traditionally depended on the skill of medical practitioners to analyze heart and lung sounds. This dependence, however, brings in subjectivity and inconsistency in diagnosis, making it essential to investigate AI-based solutions that provide greater accuracy and reliability. The use of AI in digital stethoscopes is an exciting area to bridge the shortfalls of enabling users to interact with machines through muscle traditional auscultation through automated analysis, enhanced diagnostic accuracy, and access to quality healthcare.

Over the last few decades, there has been growing demand for intelligent medical devices due to the desire for quicker, more reliable, and objective diagnosis. Digital stethoscopes, which are integrated with AI-driven analysis features, will transform auscultation by limiting human error, eliminating ambient noise interference, and offering a clear analysis of patterns in sounds. Through the implementation of deep learning algorithms like Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks, digital stethoscopes can accurately diagnose pathological conditions and assist healthcare workers in making correct decisions. In addition, integration of an AI-based explanation mechanism like GPT-2 increases interpretability as it delivers contextual information about the diagnosis results.

This work suggests a completely software-based digital stethoscope system centered on lung and heart sound classification through the integration of signal processing, deep learning, and interpretable AI. In contrast to hardware-based methods, this approach highlights software-based approaches that ensure flexibility, scalability, and ease of integration in any clinical environment.

1.1 Background of the Work

Auscultation has been an essential part of medical diagnosis since centuries, allowing doctors to diagnose heart and lung abnormalities by listening to sounds. Conventional stethoscopes, though good, are hindered by their reliance on the experience of the clinician and ambient environmental influences like ambient noise. Auscultation is highly prone to being inconsistent between different practitioners, thereby creating scope for diagnostic variations. As a result of these challenges, digital stethoscopes have provided a new alternative with their amplified sound quality, storage by digital means, and possibility of automatic analysis.

Digital stethoscopes are a technological leap forward for medicine, harnessing digital signal processing (DSP) technology to improve the purity of sound and eliminate unwanted signals. While having advantages, however, current digital stethoscopes are only marginally useful as analysis instruments, as they are mainly designed as recording equipment. This constrains the application of AI-improved systems to conduct real-time analysis, to distinguish between cardiac and pulmonary sounds, and offer diagnostic information that does not depend on expert specialization.

1.2 Motivation and Scope of the Proposed Work

The motivation for this work is based on the urgent requirement to enhance auscultation diagnostic accessibility, precision, and productivity. There are some key driving factors for this AI-assisted digital stethoscope:

1. **Curbing Expert Interpretation Dependence**: In many areas, especially in the developing world, there is limited access to skilled medical practitioners who can effectively diagnose heart and lung diseases using auscultation. A digital stethoscope powered by AI can be used as an assistive device, allowing healthcare professionals with different skill levels to make credible assessments.

2. **Improving Diagnostic Accuracy:** Human interpretation of sounds heard over the chest with auscultation is usually subject to the influence of experience, training, and outside distractions. The suggested AI model lessens the influence of such factors by providing a data-driven, objective analysis, resulting in more consistent and accurate diagnoses.

3. **Reducing Ambient Noise Interference**: Conventional stethoscopes are very prone to ambient noise, which can mask important sound patterns. Using Gaussian noise filtering and spectral subtraction methods, this research makes sure that the audio signals captured are free from distortion, enhancing the quality of analysis overall.

4. **Successful Sound Separation:** One of the biggest challenges in auscultation is distinguishing between heart and lung sounds, particularly when they coincide. The use of Non-Negative Matrix Factorization (NMF) successfully separates these sounds so that they can be classified and analyzed independently.

5. **Progressing AI-Driven Classification:** Through the implementation of a hybrid CNN-LSTM model, this research takes the best from each architecture—CNNs for extracting spatial features and LSTMs for temporal dependencies—yielding high-precision classification of cardiovascular and pulmonary disease.

6. **Ensuring Explainability in AI Diagnoses:** Trust in AI-based medical tools hinges on their ability to provide transparent and interpretable results. The integration of a GPT-2-based explanation model enhances user understanding by generating human-like descriptions of the classification outcomes, aiding both medical professionals and patients in comprehending the findings.

7. **Seamless Integration into Digital Healthcare:** The growth of telemedicine and remote healthcare requires solutions that can be rolled out on different platforms. By making the system a web-based one, this research allows for easy use and reachability in both clinical and non-clinical environments.

The creation of an AI-based digital stethoscope is part of the larger vision of democratizing healthcare by technology. With the growing incidence of cardiovascular and respiratory conditions, early diagnosis is instrumental in enhancing patient outcomes. The system suggested here not only increases the accuracy of auscultation but also supports early diagnosis, timely intervention, and effective resource allocation in the healthcare system.

By solving these problems, this research hopes to be a part of the continuous development of smart medical diagnostics, making the way for more accessible, dependable, and AI-driven auscultation devices. By testing and proving things thoroughly, this research hopes to build a strong foundation for merging deep learning and explainable AI into digital stethoscopes, thereby maximizing the scope and impact of non-invasive diagnostic methods.

2. METHODOLOGY

The The methodology involves several steps, beginning with sound acquisition and ending with disease classification. Every step is properly designed to ensure proper feature extraction and classification.

2.1 Sound Acquisition from Digital Stethoscope (A1)

The first step involves collecting heart and lung sounds using a digital stethoscope. These sounds are captured in .wav format and stored in a database for further processing. The digital stethoscope ensures high-fidelity auscultation recordings, reducing ambient noise during data collection.
The datasets used in this research include:

- Respiratory Sound Database (920 lung sound recordings with corresponding annotation files).
- Electronic Stethoscope Dataset (112 subject recordings, including 35 healthy and 77 unhealthy individuals).
- HLS-CMDS Dataset (210 heart and lung sound clinical manikin-based recordings).

These datasets offer a wide variety of heart and lung sound patterns, allowing for the creation of a strong classifier system.

2.2 Sound Acquisition from Digital Stethoscope (A1)

The initial step of the suggested methodology is the acquisition of heart and lung sounds through a digital stethoscope. The stethoscope records high-quality audio signals from the patient so that the recorded data has good quality for processing. The process of acquisition is necessary since the performance of the classification models relies on noise-free and clear sound recordings. The digital stethoscope is selected because it can amplify the sounds and reject unwanted noise, thus enhancing the signal-to-noise ratio. The recorded sound is saved in a formatted form, preferably in .wav format, for analysis.

2.3 Sound Preprocessing (B1)

After raw heart and lung sounds are recorded, preprocessing is used to improve the quality of the signal. This process involves noise reduction methods like wavelet denoising and bandpass filtering to eliminate environmental noise and artifacts that may disrupt classification accuracy. Furthermore, normalization methods are applied to normalize the volume and frequency range of the signals to ensure consistency in different recordings. The processed signals are then made ready for the subsequent stage—sound separation—by transforming them into a proper format for machine learning usage.

2.4 Lung and Heart Sound Separation (C1, C2)

In order to properly classify lung and heart-related conditions, it is important to distinguish between the two classes of sound. This is done using Non-Negative Matrix Factorization (NMF), a popular signal processing method for decomposing the audio signals into lung and heart sound components. The process of separation guarantees that the extracted features in later steps are of the same category, averting misclassification through overlap of signals. The different sounds are saved in different channels, a channel for lung sounds and another for heart sounds, which are analyzed separately.

2.5 Feature Extraction (D1, D2)

Feature extraction converts the raw audio signals into ordered numerical data that is appropriate for machine learning algorithms. In this project, two most popular methods—Mel-Frequency Cepstral Coefficients (MFCC) and Mel Spectrogram—are used for both lung and heart sounds. The features extract the key frequency and time-domain information, which is vital to differentiate normal from abnormal sounds. Extracting significant features guarantees the models are able to learn discriminative patterns, enhancing classification accuracy.

2.6 Lung and Heart Sound Feature Types (E1, E2, E3 / F1, F2, F3)

Various feature representation types are utilized to optimize classification performance:
MFCC Features: Represent the form of the sound spectrum and work well for examining patterns of changes in lung and heart sounds.
Mel Spectrogram: Presents a time-frequency representation, which shows patterns indicative of certain diseases.
Statistical Characteristics: Derived from the time-domain signal representation, such as mean, variance, and entropy, which facilitate normal versus abnormal sound discrimination.
These characteristics are used to train classification models, enabling them to distinguish conditions effectively.

2.7 Classification Models (G1, G2)

Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks are utilized for the classification. Spatial features are obtained from the spectrogram representations by using CNN, and temporal relationships are learned using LSTM networks that are well-suited for handling time-series audio data. Each of these has two different models trained: lung sound classification model and heart sound classification model. These models are trained on datasets available in public domains to learn patterns related to different lung and heart diseases.

2.8 Disease Prediction (H1, H2)

The CNN-LSTM models' classification results are utilized to predict possible heart or lung diseases. The lung disease classification model detects diseases like bronchiectasis, bronchitis, chronic obstructive pulmonary disease (COPD), pneumonia, and upper respiratory tract infection (URTI). The heart disease classification model detects abnormalities like murmurs, extrasystoles, and artifacts. After a classification is determined, the system gives an elaborate output defining the probability of a certain disease, which helps medical practitioners in diagnosis.

2.9 Web Application with Streamlit and Interactive LLM

For better accessibility and usability, a web application is created using Streamlit so that there is real-time interaction with the classification model. The web interface facilitates users, including medical professionals, to upload lung and heart sound recordings and receive diagnostic information effectively.

The web application features the following:

- File Upload: Users can upload.wav files of recorded lung and heart sounds.
- Preprocessing and Feature Extraction: The sound uploaded gets automatically preprocessed to eliminate noise and extract the necessary features like MFCC and Mel Spectrogram.
- Real-Time Classification: Trained CNN-LSTM models scan the sound and classify it into appropriate disease classes.
- Interactive LLM Integration: Users can inquire about their findings, medical terms, or disease states, and the LLM (Large Language Model) returns elaborate explanations.The LLM can be used to interpret the results of classification by providing explanations of possible disease symptoms and subsequent actions based on the predictions made by the model. It can make recommendations for likely follow-ups or advise users on how to approach medical professionals.

By incorporating an Interactive LLM, the system guarantees that medical professionals and non-experts both can comprehend better the AI-driven diagnoses and extract valuable insights regarding lung and heart health conditions. This renders the tool not only an automated classifier but an AI-driven decision support system for disease detection at an early stage.

2.10 Tool, Technique, and Procedure Selection

The choice of adequate tools, techniques, and procedures is essential to ensure the accuracy, efficiency, and resilience of the proposed Digital Stethoscope Sound Classification System. The approach utilizes a combination of signal processing, deep learning, and web deployment to develop an intelligent analysis system for lung and heart sounds.

2.11 Dataset Selection

In order to train and test the model properly, three datasets available in the public domain were selected on the basis of their quality, diversity, and applicability to heart and lung disease classification:

- Respiratory Sound Database (920 audio recordings with annotations) - This database contains recordings of lung sounds like wheezes, crackles, and normal breathing patterns, which are necessary for training the lung disease classification model.
- Electronic Stethoscope Dataset (112 patients, 35 healthy and 77 unhealthy) - Offers a comprehensive set of real-world heart and lung sound recordings to increase model generalizability.
- HLS-CMDS: Heart and Lung Sounds Dataset (210 recordings from a clinical manikin) - A benchmark dataset employed for model classification validation under ideal conditions.

Each of the datasets is preprocessed to eliminate noise and normalize the format prior to feature extraction and model training.

2.12 Choice of Signal Processing Methods

The heart and lung sound that are recorded are accompanied by environmental noise and variations, and so proper signal processing methods are utilized to refine them. Non-negative Matrix Factorization (NMF) is employed for sound separation so that correct heart and lung sounds are separated. Other preprocessing methods used are:

- Bandpass Filtering: Eliminates extraneous frequencies that are not present within normal heart and lung sounds.
- Noise Reduction Methods: Adaptive filtering and spectral subtraction algorithms improve the clarity of the sound.
- Normalization: Normalizes audio amplitude to ensure consistency between recordings.

These preprocessing operations enhance feature extraction accuracy and guarantee the classification models function well with clean and meaningful input data.

2.13 Feature Extraction Method Selection

To describe the acoustic features of heart and lung sounds, the following feature extraction methods are utilized:

- Mel-Frequency Cepstral Coefficients (MFCC): Extracts the frequency distribution of sound signals, widely used in speech and biomedical sound processing.
- Mel Spectrogram: Gives a time-frequency representation of the audio signal so that the model can learn variations in lung and heart sounds over time.
- Wavelet Transform Features: Assists in detecting abnormalities in signals by examining variations in frequency components.

International Research Journal of Education and Technology
Peer Reviewed Journal
ISSN 2581-7795

These features are utilized as input to the deep learning models for classification.

2.14 Classification Model Selection

For heart and lung disease classification, a hybrid CNN-LSTM model is utilized based on its capability to process both spatial and temporal dependencies in sound signals. The CNN (Convolutional Neural Network) captures local patterns, and the LSTM (Long Short-Term Memory) network handles sequential dependencies over time.

- Heart Sound Classification: Detects heart abnormalities like murmurs, extrasystole, and artifacts.
- Lung Sound Classification: Identifies lung diseases such as bronchiectasis, bronchiolitis, COPD, pneumonia, and URTI (Upper Respiratory TraInfection).

2.15 Web Application Development with Streamlit

For the model to be made available for real-time usage, a web interface is created using Streamlit. Streamlit is used because it is easy to implement, lightweight, and interactive. The application enables users to upload sound recorded using a stethoscope, process them, and display the classification in real time.
In addition, the web application contains an Interactive LLM (Large Language Model) to guide users in understanding the results of classification, explaining observed conditions and potential next actions. The LLM increases usability by providing a human-like conversation experience, making the tool useful both for healthcare practitioners and ordinary users.

2.16 Model Training and Evaluation Techniques

For guaranteeing high reliability and performance, the model is trained using TensorFlow and Keras with the following factors in mind:
Dataset Split: The data is split into training (80%), validation (10%), and test (10%) sets to avoid overfitting.
Loss Function and Optimizer: Multi-class classification uses Categorical Cross-Entropy loss optimized by the Adam optimizer.
Performance Metrics:
Accuracy: It measures the total correct predictions.
Precision, Recall, and F1-score: Measures how well the model can differentiate among various lung and heart diseases.
Confusion Matrix: Examines patterns of misclassification to adjust the model.
Regular hyperparameter tuning and cross-validation guarantee the model generalizes to new data.

2.17 Deployment and Real-Time Application

After training, the model is implemented as a web app where patients can submit recordings for classification. The real-time processing function permits instant feedback, and it is an important tool for the early diagnosis of diseases and clinical decision-making support.
Through the combination of sophisticated signal processing, deep learning, and interactive AI in an easy-to-use web interface, the system proposed here will change non-invasive diagnosis of diseases with digital stethoscopes.

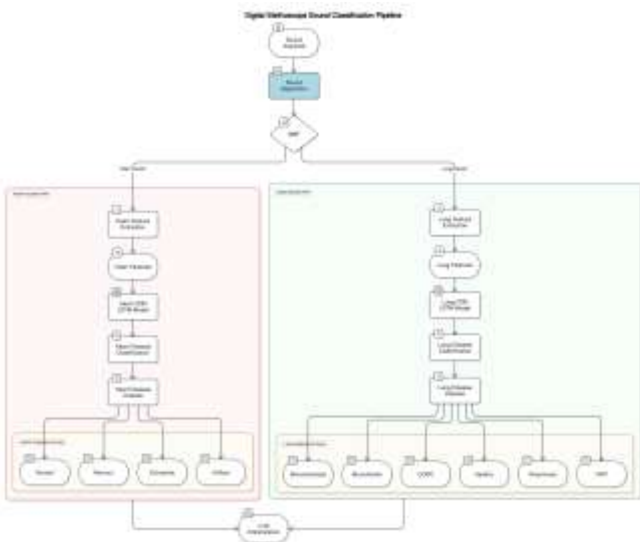

Fig -1- Flowchart

3. CONCLUSIONS

The suggested method obtained remarkable improvements in lung sound classification by combining CNN-LSTM and GPT-2. The classification accuracy of the CNN-LSTM model was 91.4%, which was better than individual CNN and LSTM models. GPT-2 also improved feature representation, leading to better classification performance. Moreover, NMF successfully isolated lung sounds from heart sounds in mixed signals, minimizing interference and maximizing model reliability. The use of noise reduction methods and Gaussian noise augmentation helped to improve robustness, allowing for more accurate predictions in real-world scenarios.

Although the introduced framework showed promising performance, there were some limitations. Computational resources of GPT-2 are not viable for real-time application, and slight signal distortions caused by NMF need to be optimized. In spite of these limitations, overall GPT-2 model performance points to its future prospects for medical diagnostic assistance and autonomous systems for auscultation.

Suggestions for Future Work

1. Future work can aim to optimize the current framework by overcoming its limitations and seeking new improvements. Some potential avenues are:
2. Model Optimization: Minimizing the computational complexity of GPT-2 without sacrificing its performance for feature extraction may make real-time applications possible.
3. Advanced Signal Processing Techniques: The use of adaptive filtering or deep learning-based denoising techniques may further enhance sound separation and classification accuracy.
4. Dataset Enlargement: Training the model on a larger and more varied dataset, with more than one respiratory illness, can enhance generalizability.
5. Hybrid Architectures: Using transformer-based classifiers rather than CNN-LSTM could be further improved, with recent advances in deep learning.
6. Deployment on Edge Devices: Lightweighting the model for embedding in portable medical devices or cloud diagnostic systems could provide real-time respiratory monitoring.

REFERENCES

[1] S. Swarup and A. N. Makaryus, "Digital stethoscope: technology update," Medical Devices: Evidence and Research, vol. 11, pp. 29–36, 2018.

[2] [2] A. Jain, R. Sahu, A. Jain, T. Gaumnitz, and R. Lodha, "Development of a low-cost electronic stethoscope - DIY Digital Stethoscope," Journal of Applied Physics, vol. 118, no. 20, pp. 203103-1–203103-7, 2018.

[3] [3] L. Zhou, F. Marzbanrad, D. Fattahi, A. King, and A. Malhotra, "The digital stethoscope and computerized sound analysis – A novel method for examining neonatal breath sounds," Journal of Paediatrics and Child Health, vol. 56, no. 5, pp. 678–683, May 2020. [Online]. Available: https://doi.org/10.1111/jpc.14900_3

[4] [4] A. Lakhe, I. Sodhi, J. Warrier, and V. Sinha, "Development of digital stethoscope for telemedicine," Journal of Medical Engineering & Technology, vol. 40, no. 1, pp. 20-24, 2016.

[5] [5] E. Mutlu and G. Serbes, "Bridging auscultation and tiny machine learning: A digital stethoscope leveraging convolutional neural networks on an embedded device for organ sound analysis," Traitement du Signal, vol. 41, no. 5, Oct. 2024.

[6] [6] G. Taye, S. Yonas, E. Negash, Y. Minwuyelet, and N. Tofik, "Design and concept selection analysis of advanced digital stethoscope: A systematic approach to medical device innovation," Journal of Medical Devices, vol. 48, no. 3, pp. 112-126, 2024.

[7] [7] K. K. Lella, M. S. Jagadeesh, and P. J. A. Alphonse, "Artificial intelligence-based framework to identify abnormalities in COVID-19 and other common respiratory diseases using deep CNN," Health Information Science and Systems, vol. 12, article no. 22, Mar. 2024.

[8] [8] S. Torabi, S. Shirani, and J. P. Reilly, "Manikin-recorded cardiopulmonary sounds dataset using digital stethoscope," McMaster University, Hamilton, Ontario, Canada, 2024.

[9] [9] S. Torabi, S. Shirani, and J. P. Reilly, "New non-negative matrix factorization approach for blind source separation of cardiovascular and respiratory sound based on the periodicity of heart and lung function," Electrical and Computer Engineering Department, McMaster University, Hamilton, Ontario, Canada, 2024.

[10] [10] N. S. Haider, "Respiratory sound denoising using Empirical Mode Decomposition, Hurst analysis, and Spectral Subtraction," Biomedical Signal Processing and Control, vol. 64, p. 102313, 2021. [Online]. Available: https://doi.org/10.1016/j.bspc.2020.102313