SONOCARDIA (AI-Powered Mobile application for Heart Disease Detection & Predection Using Heart Sounds)

Ahmed AbdelRahman Gaber[1], Bassel AbdelRahim [2], Mohab Mohammed Saber[3], Omar AbdelAzim[4], Ragy Sameh Wasfy[5]
Supervised By Dr. Alaa Hamdy[6], Eng. Salma Osama[7]
Faculty of Computer Science
Misr International University, Cairo, Egypt
ahmed2103540[1], bassel2107433[2], mohab2100084[3], omar2100083[4], ragy2106213 [5]
alaa.hamdy[6], salma.osama[7],{@miuegypt.edu.eg}

***Abstract*—Cardiovascular disease remains a leading cause of death worldwide, highlighting the need for an early and accurate diagnosis. Traditional auscultation and ECG interpretation require expertise and are often subject to human error. This project introduces SONOCARDIA, a novel heart health monitoring system that integrates phonocardiogram (PCG) and electrocardiogram (ECG) signals with machine and deep learning models to detect murmurs and predict potential heart diseases. The system employs an embedded device based on ESP32, incorporating the MAX9814 microphone and the AD8232 ECG module to collect real-time data. The acquired signals are pre-processed and analyzed using multiple models, including XGBoost, Convolutional Neural Networks (CNN), and Bidirectional Long-Short-Term Memory (BiLSTM). The final system is deployed through a Flask web server to facilitate user interaction and visualization of the results. Performance is evaluated using key metrics such as accuracy, precision, recall, and F1 score, confirming the system's effectiveness in aiding early diagnosis.**

Keywords: Heart Murmur; ECG; CNN; XGBoost; BiLSTM; Deep Learning; Cardiac Health.

I. INTRODUCTION

Heart diseases are a major public health concern, with conditions such as murmurs and arrhythmias often going undetected in their early stages. Conventional diagnostic methods like auscultation with stethoscopes and manual ECG interpretation require trained clinicians and are sometimes insufficient in identifying subtle anomalies. This has led researchers to explore automated approaches that combine biomedical signal acquisition with machine learning to provide timely and reliable cardiac assessments.

The goal of this project is to design an end-to-end heart monitoring solution, named SONOCARDIA, which leverages both PCG and ECG data to perform murmur classification and heart disease prediction. The system uses affordable and portable hardware, including an ESP32 microcontroller, a MAX9814 microphone for heart sound acquisition, and an AD8232 ECG sensor. These signals are preprocessed using filtering and segmentation techniques to remove noise and highlight diagnostically relevant features.

Various machine learning and deep learning algorithms are employed to classify the conditions. XGBoost is used to distinguish between normal, murmur, and artifact PCG recordings. CNN models are implemented for finer murmur classification based on waveform characteristics. Additionally, a BiLSTM model is trained to analyze ECG sequences and predict heart diseases. The system's modularity, interpretability, and real-time performance make it suitable for integration into mobile health platforms, enhancing accessibility in remote or underserved regions.

The integration and comparative analysis of various heart sound and ECG-based diagnostic technologies is the main contribution of this paper. We evaluate machine learning and deep learning models on standardized cardiac datasets to identify the most effective approaches for murmur classification and heart disease prediction. This comprehensive study highlights the strengths and limitations of each model and offers insights into how combining signal modalities and model architectures can enhance diagnostic accuracy and system reliability in real-world applications.

The sections that follow take a logical order. introducing a wide range of papers that are only focused on the particular problem. Next, a thorough breakdown of the steps taken to address the problem, together with a description of the data sets and each algorithm that was tested. Next, each method's result is presented, along with a compelling comparison that makes determining which is the most effective easier. Lastly, summarizing the topics covered and citing relevant literature.

II. Related Work

The domain of heart sound classification and murmur detection & prediction has been widely studied, and numerous researchers proposed innovative approaches and achieved promising results. In our project, we build upon the knowledge and methodologies established in existing literature to enhance the accuracy and interpretability of murmur classification. Our work integrates insights from several key studies that have informed our design choices and model architecture. All relevant references are provided in the references section.

- This device[1],**"Eko Duo"** is a handheld smart stethoscope and electrocardiogram (ECG) monitor that captures heart sounds and electrical activity simultaneously. The system integrates machine learning algorithms to analyze both acoustic and ECG data for detecting heart murmurs, arrhythmias, and other abnormalities. Eko Duo's mobile application provides clinicians with real-time access to the data, and AI-driven analyses help in detecting conditions like atrial fibrillation, valvular disease, and heart failure. The system has been widely adopted in telemedicine applications, allowing for remote patient monitoring and diagnosis, particularly in areas with limited access to cardiovascular specialists.

- **Stethio AI**[2] is another AI-powered system that uses digital stethoscopes to record and analyze heart sounds. The system employs convolutional neural networks (CNNs) to classify heart sounds into normal and abnormal categories. Its primary focus is on the detection of valvular heart diseases, such as mitral valve prolapse and aortic stenosis. Stethio AI's mobile platform also includes a cloud-based solution, enabling physicians to store and share heart sound data with specialists for further review. Additionally, its real-time feedback mechanism allows healthcare professionals to assess heart conditions during routine checkups.

- This model[3] is a cloud-based AI solution that analyzes heart sounds to detect cardiac conditions such as heart murmurs, arrhythmias, and valve defects. The system relies on a hybrid AI model, combining both machine learning and signal processing techniques. By integrating with portable heart sound recording devices, CardioSounds allows patients and healthcare providers to collect data in home settings, facilitating early diagnosis and continuous monitoring. The mobile application presents results in an intuitive format, making it suitable for both clinical and non-clinical environments.

- **Medtronic**[4] has developed an AI-powered digital stethoscope integrated with advanced algorithms to detect heart sounds and identify conditions like heart murmurs and irregular heartbeats. This system uses a large database of heart sounds to improve its diagnostic accuracy over time. The data collected can be sent to a mobile application where machine learning algorithms provide a diagnostic assessment. Medtronic's stethoscope is used in both clinical and home settings, allowing for the early detection of heart conditions and reducing the burden on healthcare systems.

- **AliveCor**[5] has developed a device which is a portable ECG device that can detect multiple types of arrhythmias, including atrial fibrillation, bradycardia, and tachycardia. While primarily an ECG-based tool, it can also be used in conjunction with digital stethoscopes for heart sound analysis. This hybrid approach provides a more comprehensive cardiovascular assessment. The device works with a smartphone app, enabling users to send their ECG and heart sound data to their doctors for analysis. The AI-driven app provides instant feedback to the user, notifying them of potential cardiac abnormalities.

- **eKuore Pro**[6] is a digital stethoscope that pairs with a mobile app to record and share heart and lung sounds. It is particularly useful for telemedicine and remote consultation but does not incorporate AI-driven diagnostics or personalized health monitoring features. Its primary audience is healthcare professionals rather than individual users.

- **EchoWear**[7]: EchoWear is a wearable technology that integrates with smartwatches to monitor heart sounds and rhythms. While it provides continuous monitoring, its diagnostic capabilities are limited to general health metrics and do not include detailed analysis for specific heart conditions. The system primarily targets fitness enthusiasts rather than patients or healthcare providers.

- **iStethoscope Pro**[8]: The iStethoscope Pro is a mobile application that uses a smartphone's built-in microphone or an external stethoscope attachment to record heart sounds. It provides basic visualizations of heart sound waveforms and allows users to save and share recordings. However, it does not include advanced

AI-based diagnostic features or comprehensive heart health insights.

III. Proposed Methodology

This system is designed to enhance early detection and prediction of heart-related conditions by leveraging heart sound recordings and ECG signals.

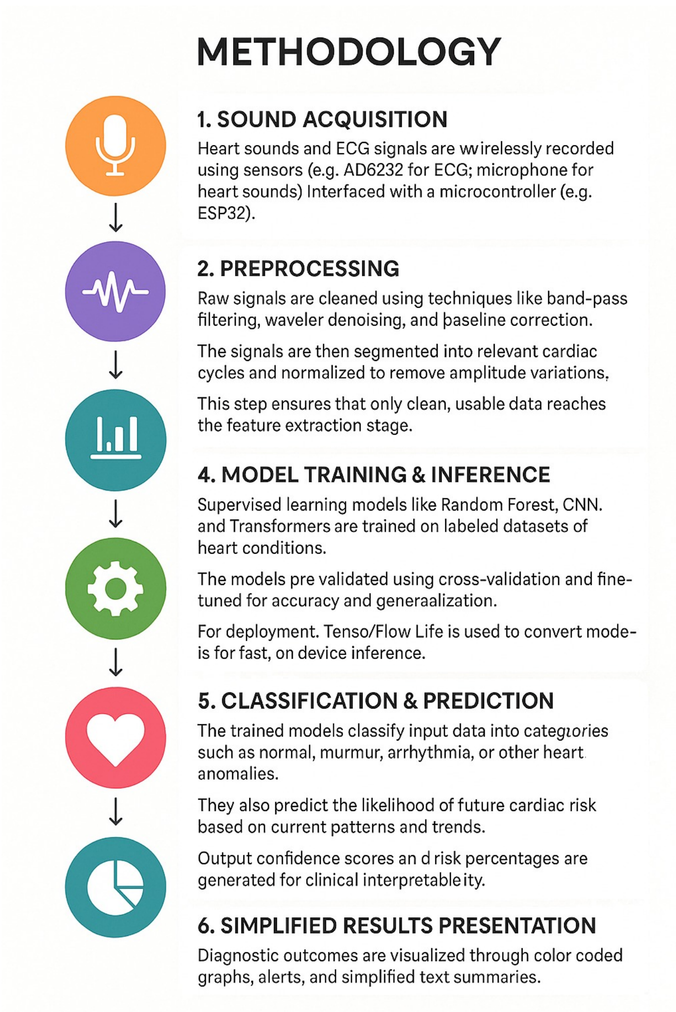


Fig. 1. Methodology Diagram

Starting with data acquisition through an embedded stethoscope and ECG device, the system preprocesses the signals using specialized libraries, then applies deep & machine learning models for abnormality detection. It outputs both prediction and classification results, which are then presented in a simplified format to support proactive patient monitoring and timely alerts.

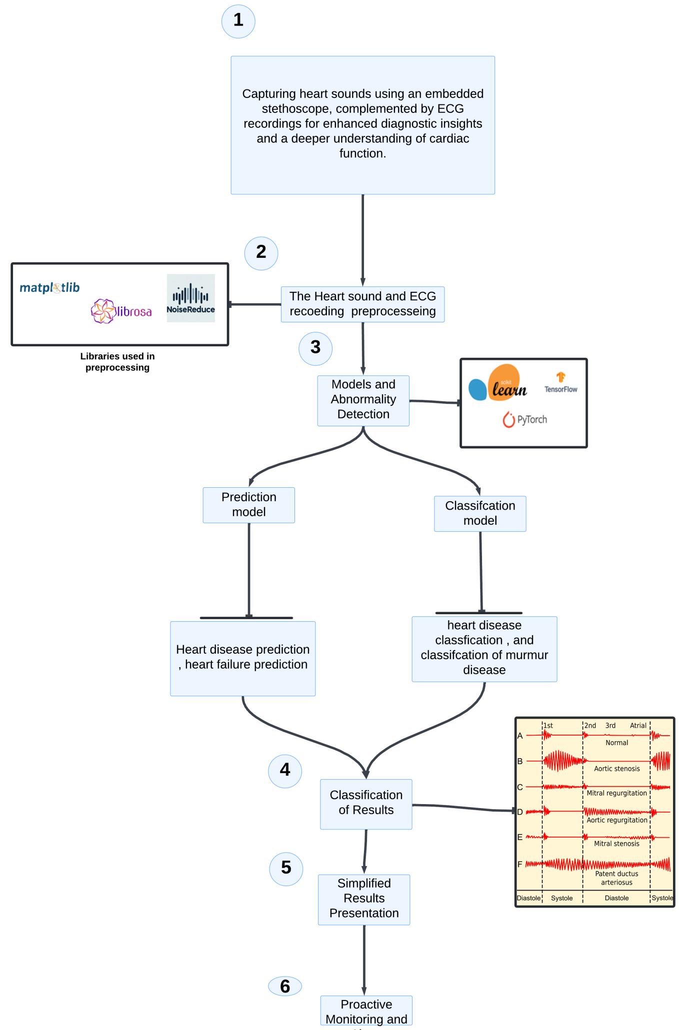


Fig. 2. System Overview

A. *Capturing data through an embedded stethoscope and ECG device*

Before preprocessing & model inference begins, the SONOCARDIA system collects both heart sound and ECG data directly from the custom hardware we built. The device integrates a medical-grade stethoscope connected to a MAX9814 microphone amplifier for capturing high-quality heart sounds, and an AD8232 ECG module for recording the heart's electrical activity. An ESP32 microcontroller collects and transmits this data wirelessly to the cloud, where the models perform their analysis.

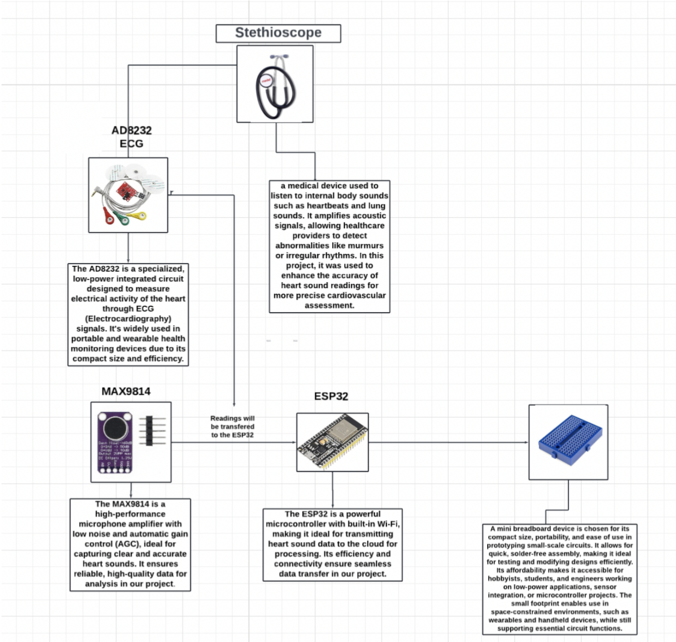


Fig. 3. Hardware

The SONOCARDIA device system consists of the following main components:

- **Stethoscope:** A medical device used to listen to internal body sounds such as heartbeats and lung sounds. In this project, it is used to capture clear acoustic signals for analysis.
- **AD8232 ECG Module:** A low-power integrated circuit designed for measuring the electrical activity of the heart through ECG signals. It provides a clear and amplified signal for accurate monitoring.
- **MAX9814 Microphone Amplifier:** A high-performance microphone amplifier with low noise and automatic gain control (AGC), ensuring reliable, high-quality audio data for analysis.
- **ESP32 Microcontroller:** A powerful microcontroller with built-in Wi-Fi, used to transmit heart sound data to the cloud for further processing.
- **Breadboard:** A compact prototyping board used for testing and assembling the circuit without the need for soldering.

1) The stethoscope captures heart sounds and transmits them to the MAX9814 microphone amplifier.
2) The MAX9814 amplifier enhances the sound quality by applying automatic gain control (AGC), ensuring clear audio signals.
3) Simultaneously, the AD8232 ECG module records the electrical activity of the heart, providing complementary data.
4) Both audio and ECG data are transmitted to the ESP32 microcontroller, which processes and sends the data to the cloud for analysis.
5) The breadboard is used to assemble and connect all components securely during the prototyping stage.

B. *Preprocessing and Feature Extraction*

In machine learning and signal processing, preprocessing plays a crucial role in preparing raw data for effective analysis and model training. This section outlines the detailed preprocessing steps applied to both audio and ECG datasets used in the prediction models. These preprocessing techniques ensure that the data is clean, standardized, and feature-rich, which is essential for achieving accurate predictions. The process is divided into two main parts: preprocessing of the audio dataset and preprocessing of the ECG dataset. Each part includes specific methods tailored to handle the unique characteristics and challenges of audio and ECG signals.

1) Preprocessing Audio dataset:

a) **Traditional Machine Learning Preprocessing**
For the traditional approach:
- Audio files are loaded using `librosa` at a 22050 Hz sampling rate.
- Files are standardized to a 10-second duration.
- Standardization is achieved through zero-padding for shorter files and truncation for longer ones, ensuring consistent input lengths.

b) **YAMNet Preprocessing**
For the YAMNet-specific preprocessing:
- The audio is resampled to 16kHz to meet YAMNet's requirements.
- The audio is standardized to 3 seconds (48000 samples).
- The audio is normalized by dividing by the maximum absolute value to ensure consistent amplitude ranges.
- Data is converted to `float32` format for TensorFlow compatibility.

c) **Basic Audio Loading and Standardization**
- Audio files are loaded using `librosa` at a 22050 Hz sampling rate.
- Audio is normalized using `librosa.util.normalize()`.
- Files are standardized to a maximum length of 200 frames.

d) **Noise Filtering and Cleaning**
- Audio normalization, Bandpass filtering (20-400Hz) to isolate heart sounds
- High-pass filtering to remove baseline wander
- Uses Butterworth filter for signal cleaning.
- Applies high-pass filtering to remove low-frequency noise.

e) **Feature Extraction**
- **MFCCs (Mel-frequency cepstral coefficients)**:

– Extracts 13 MFCC features.
– Captures spectral characteristics.
- **Spectral Features**:
  – Spectral centroids.
  – Spectral rolloff.
  – Spectral bandwidth.
  – Zero-crossing rate.

f) **Feature Normalization**
- Features are normalized using mean and standard deviation.
- Zero-padding or truncation is applied to ensure consistent length.

g) **Data Augmentation**
- **Noise Addition**: Adds controlled Gaussian noise with a factor ranging from 0.002 to 0.015.
- **Pitch Shifting**: Random pitch shifts between -3 and +3 semitones, preserving audio quality while creating variations.

h) **Data Preparation**
- Handles multiple audio formats (`.wav` and `.mp3`).
- Processes recordings from different valve locations (AV, PV, TV, MV).
- Combines features from all valve locations.
- Implements data augmentation with a configurable augmentation factor.

i) **Feature Standardization**
- Normalizes features using mean and standard deviation.
- Handles missing values by using zero-padding.
- Ensures consistent feature dimensions across all samples.

2) **Preprocessing ECG dataset:**

a) **ECG Signal Cleaning:**
- Bandpass filtering (0.5-50 Hz) to remove noise and artifacts.
- Butterworth filter implementation for signal smoothing.
- Removal of baseline wander and high-frequency noise.
- Signal normalization using mean and standard deviation.

b) **Data Standardization:**
- Window-based processing with fixed window size (500 samples).
- Normalization of signals using z-score normalization.
- Handling of missing or invalid data points.
- Standardization of RR intervals.

c) **Beat Type Processing:**
- Filtering of common beat types (N, V, /, L, R, A, F).
- One-hot encoding of beat types.
- Mapping of beat types to numerical indices.
- Handling of unknown or rare beat types.

d) **Feature Extraction:**

i) **Temporal Features:**
- RR intervals (time between consecutive beats).
- Beat-to-beat intervals.
- Signal duration features.
- Window-based signal segments.

ii) **Signal Features:**
- Dual-lead ECG processing (MLII/ML2 and V leads).
- Signal amplitude features.
- Signal morphology features.
- Cross-lead correlation features.

iii) **Sequence Features:**
- Beat type sequences.
- RR interval variations.
- Consecutive beat patterns.
- Beat type transitions.

iv) **Statistical Features:**
- Mean and standard deviation of RR intervals.
- Beat type distribution statistics.
- Signal amplitude statistics.
- Variability measures.

v) **Window-based Features:**
- Fixed-size windows (500 samples).
- Overlapping window processing.
- Window normalization.
- Feature concatenation across leads.

C. *Used Algorithms*

In this paper, we employ a combination of machine learning and deep learning algorithms tailored to different tasks. An XGBoost model is used to classify heart sounds into three categories: normal, murmur, and artifact, based on audio data. Additionally, two convolutional neural network (CNN) models are utilized—one for classifying murmur severity levels using heart sound recordings, and another for predicting potential heart conditions using ECG signals. This multi-model approach allows for comprehensive analysis of both audio and ECG data to support accurate diagnosis and monitoring.

1) **XGBoost:**
XGBoost (Extreme Gradient Boosting) is a robust and efficient machine learning algorithm based on ensemble decision trees. In our project, we use XGBoost to

classify heart sounds after extracting structured audio features such as MFCCs, spectral roll-off, chroma, and zero-crossing rates. XGBoost builds an ensemble of decision trees sequentially, with each new tree trained to correct the errors of the previous ones using gradient descent. Its built-in regularization helps prevent overfitting, and its parallel processing makes it computationally efficient.

2) **YAMNet:**
YAMNet is a deep neural network model for audio event classification, built on the MobileNetV1 architecture and pre-trained on the AudioSet dataset. In our project, we use YAMNet as a deep feature extractor by feeding raw heart sound recordings into the model. The audio is converted into log mel-spectrograms, and YAMNet processes these through its convolutional layers to extract meaningful representations. These features are then passed to a custom classification layer trained specifically for murmur detection. YAMNet achieved a close percentage in our evaluation, making it a strong deep learning baseline for comparison against traditional models like XGBoost.

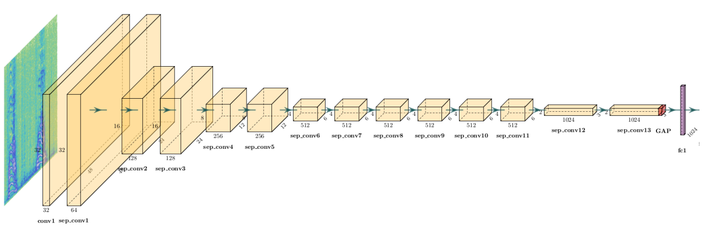


Fig. 4. YAMNet Structure

3) **Convolutional Neural Networks:**
Convolutional Neural Networks (CNN) are a class of deep learning models designed primarily for image and spatial data processing, but they are also highly effective for analyzing audio data represented as spectrograms. CNNs take mel-spectrograms of heart sound recordings as input, treating them like images. The convolutional layers automatically learn spatial patterns such as frequency and time-based features of murmurs, which are then passed through pooling and fully connected layers for classification. CNNs excel at capturing local dependencies and have proven effective in recognizing complex acoustic patterns in heart sounds.

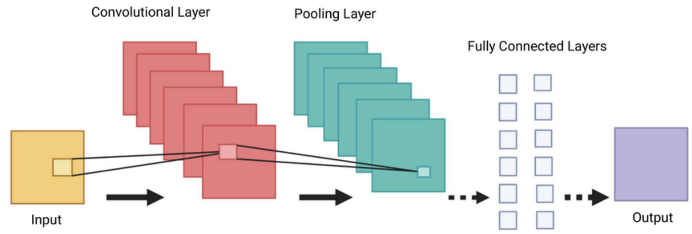


Fig. 5. CNN Structure

4) **Convolutional Neural Networks(A Multi-Layered Approach):**

Convolutional Neural Networks (CNNs) are deep learning architectures designed for image and signal processing tasks. They consist of multiple layers, including convolutional layers that detect features, pooling layers that reduce spatial dimensions, and fully connected layers for final predictions. CNNs automatically learn hierarchical features from raw input, enabling them to effectively classify or predict based on complex patterns, making them highly suited for tasks such as image recognition and signal analysis.

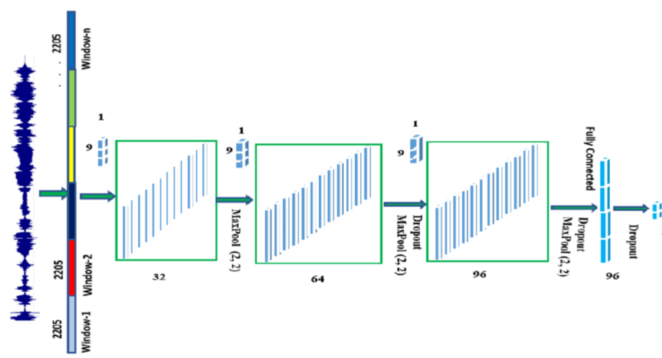


Fig. 6. CNN multi layers

5) **Bidirectional LSTM :**
Bidirectional Long Short-Term Memory (BiLSTM) networks are a type of Recurrent Neural Network (RNN) that can learn long-term dependencies in sequential data—such as audio signals. Unlike standard LSTMs, BiLSTMs process input sequences in both forward and backward directions, allowing the model to capture context from the past and the future. BiLSTM is applied to sequences of audio features (e.g., MFCCs or spectrogram slices) to understand temporal patterns and variations across time. This makes it particularly effective in detecting murmurs that may occur at different times within the cardiac cycle, enhancing the model's understanding of time-based dependencies in heart sound recordings.

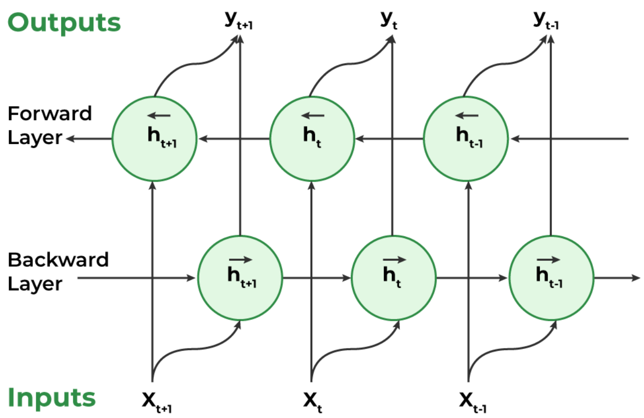


Fig. 7. Bidirectional LSTM Structure

D. *Models Created & Integrated*

The system is built upon the integration of three deep and machine learning models, each serving a distinct yet complementary purpose:

1) **Heart Sound Detection Model (XGBoost-based):** This model performs an initial classification of heart sound recordings into three categories: **Normal**, **Murmur**, or **Artifact**. It uses handcrafted features extracted from the audio signals and employs the XGBoost classifier for its robust performance on structured data.

2) **Murmur Severity Classification Model (CNN-based):**

Activated only if the output of the first model is Murmur, the model uses a Convolutional Neural Network (CNN) designed to classify several murmur characteristics from heart sound recordings. It processes input from four valve auscultation positions and outputs clinically meaningful predictions.

- Model Architecture:

- **Input Layer:** Accepts feature matrices (e.g., MFCC, spectrograms) from all four valve positions (AV, MV, PV, TV).
- **Convolutional Layers:** Multiple convolutional blocks with batch normalization and max pooling for hierarchical feature extraction.
- **Shared Layers:** Common convolutional layers are used to extract shared features.
- **Branching Heads:** Separate classification branches for each murmur characteristic.
- **Output Layers:** Each branch uses a softmax layer for multi-class classification.

- Training Methodology:

- **5-Fold Cross-Validation:** Ensures robustness and generalizability of the model.
- **Early Stopping & Learning Rate Reduction:** Applied to avoid overfitting and improve convergence.
- **Class Weighting:** Used to address class imbalance in underrepresented murmur characteristics.
- **Patient-Based Splitting:** Ensures that recordings from the same patient are not split across training and validation sets, preventing data leakage.

- Labels and Their Clinical Significance:

The model predicts six key murmur characteristics:

a) 1. Murmur Location (Simplified Categories)::

- Single valve: AV, MV, PV, TV
- Left heart: AV + MV
- Right heart: PV + TV
- AV + right heart valves
- MV + right heart valves
- Multiple valves (3+ valves)
- Other combinations

b) 2. Systolic Murmur Timing::

- Early-systolic
- Mid-systolic
- Late-systolic
- Holosystolic / Pansystolic
- Unknown

c) 3. Systolic Murmur Shape::

- Crescendo (increasing intensity)
- Decrescendo (decreasing intensity)
- Crescendo-decrescendo (diamond-shaped)
- Plateau (consistent intensity)
- Unknown

d) 4. Systolic Murmur Grading::

- Grade I – Barely audible
- Grade II – Quiet but clearly audible
- Grade III – Moderately loud
- Grade IV – Loud with palpable thrill
- Grade V – Very loud, audible with stethoscope partly off chest
- Grade VI – Audible without stethoscope
- Unknown

e) 5. Systolic Murmur Pitch::

- Low
- Medium
- High
- Unknown

f) 6. Systolic Murmur Quality::

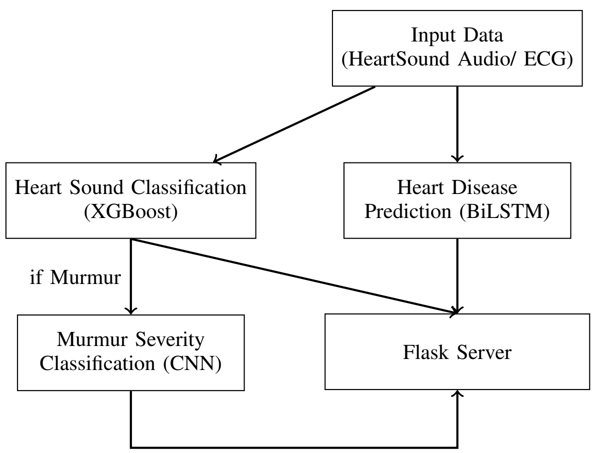
- Blowing
- Harsh
- Musical
- Unknown

A key methodological innovation is the **simplification of murmur location categories**, which improves generalization and reduces overfitting by grouping functionally and anatomically related valves.

This model has strong potential as a **clinical decision support tool**, offering objective, consistent murmur assessment.

3) **Heart Disease Prediction Model (BiLSTM-based):** This model uses ECG signal data to predict the **likelihood of heart disease** in the future. It is built with a Bidirectional Long Short-Term Memory (BiLSTM) architecture to capture temporal patterns in ECG sequences, and it outputs a prediction along with a **confidence score**.

All three models are seamlessly integrated into a unified system using a **Flask server**, which acts as the backend API. The Flask server handles incoming data (audio and ECG), routes it to the appropriate models based on the pipeline logic, and returns the final diagnostic results in a structured and interpretable format. This integration allows for real-time interaction and easy deployment in web or mobile applications.



E. *Performance Metrics*

The models are measured with many metrics. These metrics consist of accuracy, precision, f1 score, and recall. Accuracy is the percentage of correctly expected data from all the data. The precision can be defined as the total number of accurately expected positives minus the anticipated positives. Recall is the number of correctly anticipated positives out of all the true positives. The number of accurately predicted negatives among all expected negatives is how specificity is determined.

$$\text{Accuracy} = \frac{\text{TN} + \text{TP}}{\text{TN} + \text{TP} + \text{FN} + \text{FP}} \tag{1}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{2}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{3}$$

$$\text{F1 Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$

IV. DATASETS DESCRIPTION

Our dataset, collected via digital stethoscopes and from another resources, includes normal and abnormal heartbeats with patient metadata. Designed for heart disease classification, it supports murmur detection, heart sound segmentation (S1, S2, murmurs), and disease prediction. High-quality recordings make it valuable for machine learning in cardiac diagnostics.

- **The CirCor DigiScope Phonocardiogram Dataset**[9]: The data were collected from a pediatric population in Northeast Brazil in July-August 2014 and June-July 2015. The target population was individuals who were 21 years old or younger who presented voluntarily for screening with a signed parental or legal guardian consent form. All participants completed a sociodemographic questionnaire and subsequently underwent a clinical examination, a nursing assessment, and cardiac investigations.

- **Heartbeat Sounds**[10]: This dataset was originally for a machine learning challenge to classify heart beat sounds. The data was gathered from two sources: (A) from the general public via the iStethoscope Pro iPhone app, and (B) from a clinic trial in hospitals using the digital stethoscope DigiScope.

- **Classification of Heart Sound Recordings: The PhysioNet/Computing in Cardiology Challenge 2016**[11]: The heart sound recordings were collected from different locations on the body. The typical four locations are aortic area, pulmonic area, tricuspid area and mitral area, but could be one of nine different locations. In both training and test sets, heart sound recordings were divided into two types: normal and abnormal heart sound recordings. The normal recordings were from healthy subjects and the abnormal ones were from patients with a confirmed cardiac diagnosis. Heart valve defects include mitral valve prolapse, mitral regurgitation, aortic stenosis and valvular surgery.

- **MIT-BIH Arrhythmia Database**[12]: This database contains 48 half-hour two-channel ECG recordings from 47 subjects, collected by the BIH Arrhythmia Laboratory (1975–1979). It includes 23 randomly selected recordings from 4,000 ECGs at Beth Israel Hospital (60% inpatients, 40% outpatients), while 25 were chosen for rare arrhythmias. Signals were digitized at 360 samples/second with 11-bit resolution over 10 mV, and 110,000 heartbeats were annotated by cardiologists for arrhythmia research. (15 beat types, **grouped into 5 classes**: **Normal (N)** (N, L, R, e, j), **Supraventricular**

(S) (A, a, J, S), **Ventricular (V)** (V, E), **Fusion (F)** (F), and **Unknown (Q)** (/, f, Q).)

- **The PTB Diagnostic ECG Database**[13]: contains 549 ECG recordings from 290 subjects (ages 17–87, mean 57.2), with up to five recordings per subject. Signals were collected using a PTB prototype recorder with 16 channels (14 for ECG, 1 for respiration, 1 for line voltage), digitized at 1000 samples/second with 16-bit resolution. Each recording includes 15 leads: the standard 12-lead ECG plus 3 Frank leads (Vx, Vy, Vz). Clinical summaries, including diagnosis, medical history, and interventions, are available for most subjects, except 22 cases.

V. Results and Analysis

This section presents the outcomes of the implemented methods, providing a comprehensive analysis of the results obtained. It begins with a detailed examination of the performance metrics, followed by a discussion of the results, highlighting key observations, patterns, and potential insights. The analysis aims to interpret the findings in the context of the research objectives, offering a clear understanding of their significance.

1) **The first model Heart Sound Detection Model:**
We used two advanced models, XGBoost and YAMNet, were evaluated for heart sound classification. YAMNet, a deep learning model known for its proficiency in audio classification, demonstrated notable performance by leveraging pre-trained audio embeddings. However, despite its sophisticated architecture, YAMNet slightly lagged behind in accuracy compared to the XGBoost model. The XGBoost model, leveraging handcrafted audio features and optimized for structured data, consistently outperformed YAMNet, achieving an impressive 89% accuracy. This result underscores the effectiveness of the XGBoost model in capturing the intricate characteristics of heart sounds, making it the superior choice for this classification task.
There was also a combination between the two models, but it did not achieve the same level of accuracy as the XGBoost model alone.

Model	Accuracy (%)
XGBoost	89%
YAMNet	84%
Combined Models	87%

TABLE I
PERFORMANCE COMPARISON OF HEART SOUND CLASSIFICATION MODELS

2) **Second Our Severity Classification Model:**
We uses a Convolutional Neural Network (CNN) designed to classify several murmur characteristics from heart sound recordings. It is multi-output model, because it gives 6 output as every label has it's own accuracy, and overall the model achieved an accuracy of 92% upon evaluation, demonstrating strong predictive performance. The Labels are :

a) Locations
b) Quality
c) Timing
d) Pitch
e) Shape
f) Grading

This high accuracy underscores the model's effectiveness in reliably assessing severity levels from the input data.

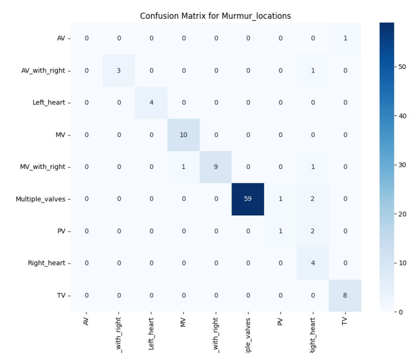


Fig. 8. Confusion matrix fo murmur **locations** in severity model

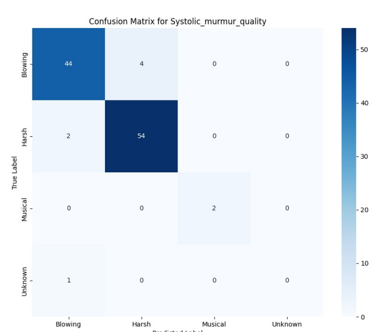


Fig. 9. Confusion matrix fo murmur **quality** in severity model

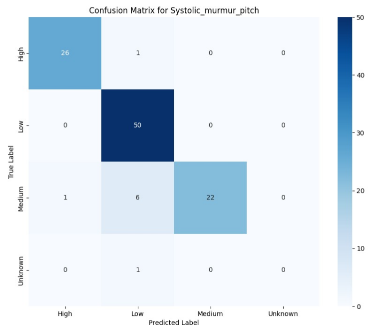


Fig. 11. Confusion matrix fo murmur **pitch** in severity model

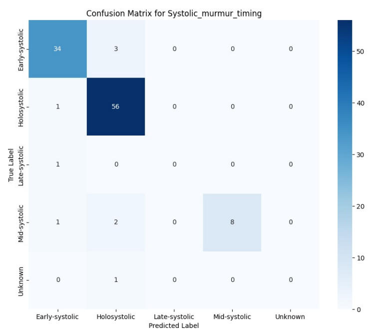


Fig. 10. Confusion matrix fo murmur **timing** in severity model

3) **Finally, The arrhythmia prediction model:** exhibited strong predictive capability, reflecting its effectiveness in identifying abnormal heart rhythms with high accuracy. The consistent performance across multiple test cases further validated its reliability for the detection of arrhythmias. This performance demonstrates its potential for the early detection of arrhythmic conditions using ECG data.

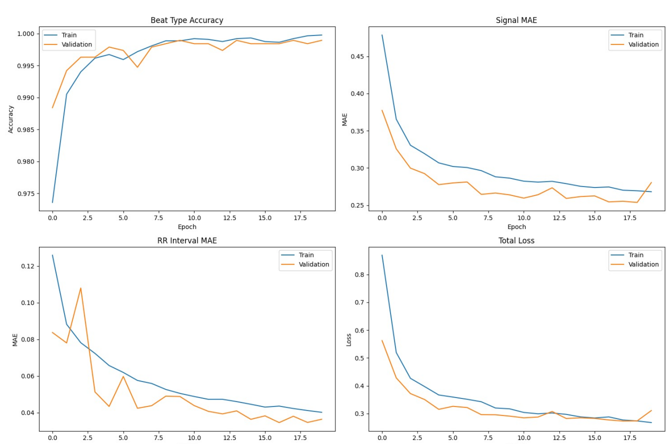


Fig. 12. Training and Validation Performance Metrics for Multi-Output Arrythmia Prediction Model

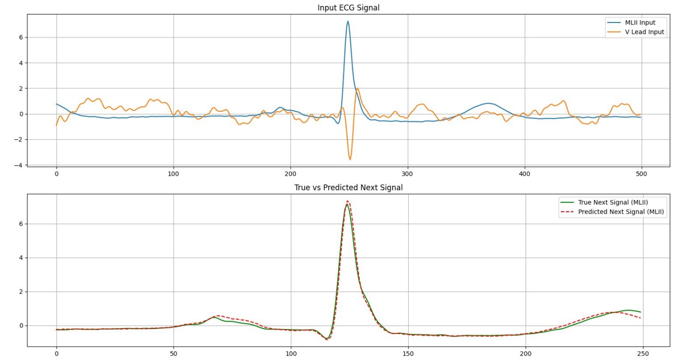


Fig. 13. Test Case 1

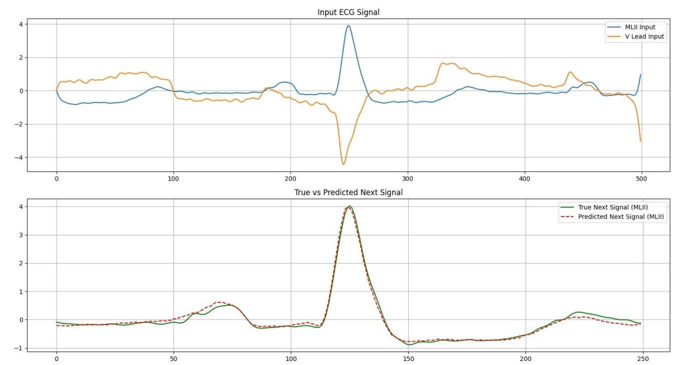


Fig. 14. Test Case 2

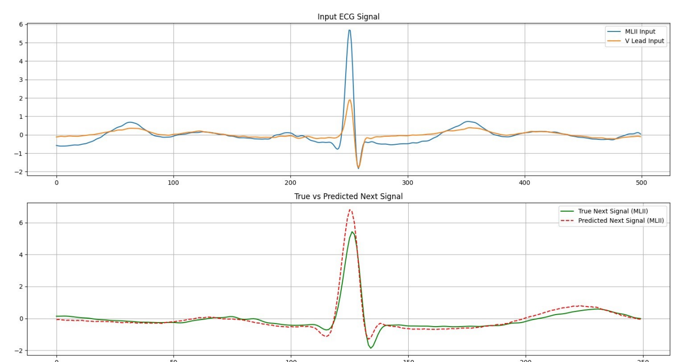


Fig. 15. Test Case 3

Overall, the results highlight the superiority of the XGBoost model for heart sound classification, the strong potential of deep learning techniques such as CNNs for complex multi-output tasks like murmur classification, and the robustness of the arrhythmia prediction model for accurately detecting abnormal heart rhythms. These findings underscore the importance of selecting models that align with the nature of the data and the specific classification objectives.

VI. Conclusion

In conclusion, the fusion of machine learning and deep learning techniques offers a groundbreaking approach to diagnosing heart conditions, especially through the analysis of heart sounds and ECG signals. The comparative results from our study reveal that deep learning architectures such as CNN and BiLSTM demonstrate superior performance in murmur classification, while XGBoost stands out among machine learning models for ECG-based predictions. These outcomes not only validate the reliability of such models in real-world diagnostics but also emphasize their potential to assist medical professionals in making quicker and more accurate decisions. As these technologies continue to evolve, they hold the promise of significantly enhancing early detection, treatment planning, and overall patient outcomes in cardiovascular healthcare.

VII. Acknowledgment

At the end, we can't forget to express our heartfelt gratitude to the Computer Science instructors and staff at Misr International University (MIU) for their exceptional dedication and continuous support throughout our academic journey. We are especially thankful to Professor Ayman Nabil, Dean of the Faculty of Computer Science, and Professor Abdelnasser Zaied, Vice Dean of Student Affairs, for their invaluable efforts in fostering a productive and supportive educational environment. Finally, we extend our deepest appreciation to Dr. Alaa Hamdy, Professor of Artificial Intelligence, and Eng. Salma Osama, Teaching Assistant, for their guidance, encouragement, and unwavering support throughout the development of this project.

References

[1] J. R. Selvaraj, A. Sathyan, N. Plakkal, and K. Sivavignesh, "Feasibility and utility of single-lead electrocardiogram recorded with a handheld device for screening of neonates: A pilot study," *International Journal of Advanced Medical and Health Research*, pp. 10–4103, 2024.

[2] S. Benjamens, P. Dhunnoo, and B. Meskó, "The state of artificial intelligence-based fda-approved medical devices and algorithms: an online database," *NPJ digital medicine*, vol. 3, no. 1, p. 118, 2020.

[3] Y. Kim, M. Moon, S. Moon, and W. Moon, "Effects of precise cardio sounds on the success rate of phonocardiography," *Plos one*, vol. 19, no. 7, p. e0305404, 2024.

[4] D. Adedinsewo, A. C. Morales-Lara, H. Hardway, P. Johnson, K. A. Young, W. T. Garzon-Siatoya, Y. S. B. Tobah, C. H. Rose, D. Burnette, K. Seccombe *et al.*, "Artificial intelligence–based screening for cardiomyopathy in an obstetric population: A pilot study," *Cardiovascular Digital Health Journal*, 2024.

[5] B. Krzowski, K. Skoczylas, G. Osak, N. Żurawska, M. Peller, Ł. Kołtowski, A. Zych, R. Główczyńska, P. Lodziński, M. Grabowski *et al.*, "Kardia mobile and istel hr applicability in clinical practice: a comparison of kardia mobile, istel hr, and standard 12-lead electrocardiogram records in 98 consecutive patients of a tertiary cardiovascular care centre," *European Heart Journal-Digital Health*, vol. 2, no. 3, pp. 467–476, 2021.

[6] G. Dimauro, D. Caivano, M. M. Ciccone, G. Dalena, and F. Girardi, "Classification of cardiac tones of mechanical and native mitral valves," in *Ambient Assisted Living: Italian Forum 2019 10*. Springer, 2021, pp. 211–222.

[7] H. Dubey, J. C. Goldberg, M. Abtahi, L. Mahler, and K. Mankodiya, "Echowear: smartwatch technology for voice and speech treatments of patients with parkinson's disease," in *Proceedings of the conference on Wireless Health*, 2015, pp. 1–8.

[8] P. J. Bentley, "istethoscope: a demonstration of the use of mobile devices for auscultation," *Mobile Health Technologies: Methods and Protocols*, pp. 293–303, 2015.

[9] P. D. C. M. N. C. O. C. F. A. J. S. M. T. H. T. T. A. E. A. B. R. R. S. G. D. C. M. T. C. Jorge Oliveira, Francesco Renna, "The circor digiscope phonocardiogram dataset v2," 2023. [Online]. Available: https://www.kaggle.com/datasets/bjoernjostein/the-circor-digiscope-phonocardiogram-dataset-v2

[10] N. G. C. M. Bentley, P. and S. Mannor, ""the PASCAL Classifying Heart Sounds Challenge 2011 (CHSC2011) Results"," 2011. [Online]. Available: https://www.kaggle.com/datasets/kinguistics/heartbeat-sounds

[11] G. D. C. David Liu, Afshin Samani *et al.*, "Classification of heart sound recordings: The physionet/computing in cardiology challenge 2016," *Computing in Cardiology*, vol. 43, pp. 609–612, 2016. [Online]. Available: https://physionet.org/content/challenge-2016/1.0.0/

[12] T. Yoon, "Mit-bih arrhythmia database," 2023. [Online]. Available: https://www.kaggle.com/datasets/taejoongyoon/mitbit-arrhythmia-database

[13] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "Ptb diagnostic ecg database," 2001. [Online]. Available: https://physionet.org/content/ptbdb/1.0.0/