MYSTETH - Home-based Heart Monitoring

Kopal Jain  
Department of Electrical Engineering  
Indian Institute of Technology, Kharagpur, India

[kopaljain@kgpian.iitkgp.ac.in](mailto:kopaljain@kgpian.iitkgp.ac.in)

Rohit Jain

Netaji Subhas University of Technology, New Delhi, India

rohitjn256@gmail.com

Salik Khwaja Mohammad

Netaji Subhas University of Technology, New Delhi, India

salik2109@gmail.com

Swati Aggarwal\*

Molde University College, Molde Norway

[swati.aggarwal@himolde.no](mailto:swati.aggarwal@himolde.no)

Orcid: 0000-0001-5986-6915

\*Corresponding author: Swati Aggarwal, Associate Professor, Postdoc Marie Curie Fellow, PhD

**Abstract**

The MySteth is an intelligent medical tool designed for cardiac disease screening, utilizing either a stethoscope or smartphone to record heart sounds. Normal heart sounds in healthy individuals consist of "lub" and "dub" noises, while murmurs—additional sounds during heartbeats—can indicate cardiac anomalies such as valve dysfunctions and rapid blood flow, categorized as systolic or diastolic.

**Method**

MySteth was developed and tested using heart sounds recorded via smartphone and digital stethoscope. To achieve high classification accuracy, MySteth employs a hybrid CNN-LSTM model combined with Linear Predictive Coding (LPC) for preprocessing. The study involves classifying recorded heart sounds into normal heartbeats and murmurs, with murmurs further divided into systolic and diastolic categories.

**Results**

The tool demonstrated an accuracy of 92% in distinguishing normal heartbeats from murmurs, 91% in classifying murmurs into systolic and diastolic types, and 90% in further categorizing systolic murmurs into Ejection Systolic Murmurs (ESM) and Pansystolic Murmurs (PSM). MySteth is accessible and affordable, requiring minimal equipment, as most individuals already own a smartphone, and digital stethoscopes are commonly available. This ease of use facilitates both professional and home-based heart monitoring, especially beneficial in remote areas with limited healthcare access.

**Conclusion**

MySteth is an at-home heart diagnostic tool that leverages deep learning to classify heart sounds into normal, ESM, PSM, and diastolic murmurs. Its user-friendly design and minimal hardware requirements ensure broad adoption across various healthcare settings, facilitating timely and accurate preliminary heart investigations. This capability is crucial in combating the global burden of cardiovascular diseases. MySteth's scalability and ease of deployment underscore its potential in early cardiovascular disease diagnosis, particularly in underserved regions, thereby promoting preventive healthcare.

**Keywords :** Cardiac Disease Screening, Heart Sounds, Murmurs, Systolic Murmurs, Diastolic Murmurs

**Highlights**

* MySteth is an innovative at-home heart diagnostic tool that uses deep learning to classify heart sounds into normal, systolic, and diastolic murmurs, with systolic murmurs further sub-classified into Ejection Systolic Murmur (ESM) and PanSystolic Murmur (PSM).
* By employing a hybrid CNN-LSTM model and Linear Predictive Coding (LPC) for preprocessing, MySteth achieves high classification accuracy, making it an effective tool for early heart condition detection, especially in underserved areas.
* Its user-friendly design and minimal hardware requirements ensure broad adoption across various healthcare settings, aiding timely and accurate preliminary heart investigations to combat the global burden of cardiovascular diseases.

**Introduction**

The two typical heart sounds in healthy people are a lub and a dub, which happen one after the other with each beating. It's common to refer to the lub as the first heart sound (S1) and a dub as the second heart sound (S2). Additional noises are heard in regular heart sounds (HS), which can be used in pathology diagnosis in circumstances when the heart is aberrant, such as valve dysfunctions and fast blood flow[1]. These extra noises, sometimes referred to as murmurs, exhibit distinct traits in relation to heart valve problems, which are circulatory heart illnesses[2]. The most common way to categorize cardiac murmurs is by timing; they can be classified as either systolic[3] or diastolic[4], depending on which portion of the heartbeat they occur during.

Murmurs of the heart that are audible during systole are known as systolic murmurs. The most common systolic murmur[5]:

* Ejection-systolic murmurs (ESM): Diamond-shaped or spindle-shaped. The intensity first increases and then decreases during S1.
* Pansystolic murmurs (PSM): Rectangular shaped. The intensity remains constant during S1.

The murmur heard in the heart during diastole is called diastolic heart murmur. Diastolic murmurs end at or before S1 and begin at or after S2[6].

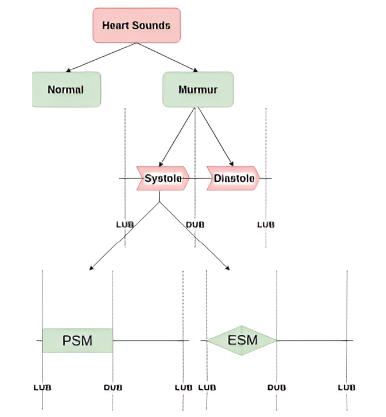
Heart murmurs are a problem that affects a large percentage of people worldwide. These murmurs might be an indicator of underlying cardiovascular disorders including valve dysfunctions. About 2.5% of Americans have heart valve disease, with the prevalence rising with age, according to the American Heart Association[7]. One of the primary reasons heart murmurs are not timely diagnosed is the lack of access to regular and comprehensive cardiac evaluations, particularly in underserved and rural areas[8]. Additionally, the subtle nature of some murmurs can make them difficult to detect without specialized equipment and expertise. The introduction of a home-based preliminary diagnostic tool for heart murmurs could be highly beneficial. Such a tool would enable individuals to monitor their heart sounds regularly, facilitating early detection of abnormalities and prompting timely medical consultations. This proactive approach could significantly reduce the burden of undiagnosed heart conditions, improve patient outcomes, and decrease healthcare costs associated with advanced cardiovascular diseases[9].

With 17.9 million deaths from Cardiovascular Diseases (CVDs) per year, or 31% of all fatalities globally, CVD is a major public health concern[10]. Early detection is key since cardiac disorders can worsen over time and necessitate more involved forms of care. For instance, coronary heart disease, one of the most common cardiac conditions in the United States, can worsen over time and eventually necessitate coronary artery bypass grafting (CABG)[11]. Preventative detection of heart diseases is essential, and medical professionals often start by checking the patient’s heartbeat and abnormalities. Further tests, such as blood pressure and fasting protein profile tests, are then performed for further analysis[12]. Currently, there is no easy method for heart screening at home without specialized medical personnel. Heart health monitoring and the availability of at-home testing options is crucial for promoting heart health awareness. At-home diagnostics can significantly contribute to heart health promotion and better outcomes for those at risk of cardiovascular problems by enabling individuals to adopt proactive measures towards lowering their risk of heart disease[13]. Numerous studies have employed machine learning and deep learning techniques to categorize heartbeat sounds; most of these studies have focused on data from phonocardiography (PCG), a specialized device used for medical diagnostics[14,15,16,17,18,19]. However, this technology is not accessible to the average consumer and cannot be performed at home.

Advancements in technology have led to smartphone applications like SensiCardiac[20], Mobile Stethoscope[21], and iStethoscope Pro[22], which allow heart sounds to be conveniently recorded. Some studies have also used Electrocardiogram (ECG) signals, which are obtained from specialized instruments in medical facilities. Some studies have also used audio from electronic stethoscopes and mobile phones[23]. Only a few categories have been used to categorize murmurs: artifact, extra-heart sound, extrasystole, murmur, normal heartbeat, moderate, severe, or normal, aberrant. There has never been an attempt to further categorize murmurs into systolic, diastolic, and systolic murmurs as well as ESM and PSM. Because there aren't enough datasets available, the majority of these studies have limitations[24]. The models are trained and validated on specific datasets which may not encompass the full variability seen in global populations. Without prior patient information, other classifications of murmurs—such as mitral valve prolapse, mitral regurgitation, and aortic stenosis—cannot be made. These classifications require further tests such as ECG, ultrasound, and cardiac CT[25]. The key to reducing healthcare costs from CVD and increasing patient outcomes lies in early detection, prevention, and access to quality health services[26]. Unfortunately, emergency rooms and hospitals are overcrowded, while affordable healthcare clinics are scarce. This created the need for the development of in-home health monitoring and CVD management programs[27]. In this work, authors have proposed MySteth as an innovative at-home heart diagnostic tool that aims to address the gap in care by providing a convenient and accessible solution for preliminary heart investigations. While there are other heart testing options available, Mysteth offers several benefits that set it apart from its competitors. It is a first-of-its-kind screening method that uses deep learning techniques and recorded heartbeat sounds to detect a wide class of heart diseases using a phone or digital stethoscope. By employing deep learning techniques, Mysteth can classify heart murmurs with greater granularity, potentially distinguishing between various types of murmurs such as systolic, diastolic, ESM, and PSM, which have not been extensively categorized in previously published studies. The most prevalent valvular heart illnesses in the world, including arrhythmia, mitral regurgitation, and coronary heart disease, can all be effectively detected at home with this technique.

**Methods**

The purpose of this work is to introduce Mysteth, an at-home cardiac diagnostic instrument created to enable the at-home screening of heart diseases. It includes conditions of the heart and blood vessels, such as coronary artery disease, hypertension, heart failure, and stroke. Globally, CVD is a major cause of morbidity and mortality that has a significant effect on public health and healthcare systems. Early detection and prevention are crucial because CVD accounts for 17.9 million deaths yearly[28]. Medical technology has advanced, but there are still no easily available, user-friendly techniques for doing at-home cardiac screenings[29]. This study looks at how Mysteth uses digital stethoscope technology and deep learning methods to offer a quick and easy way to do initial cardiac investigations. The goal is to raise awareness of heart health issues and improve the lives of people who are at risk of CVDs.



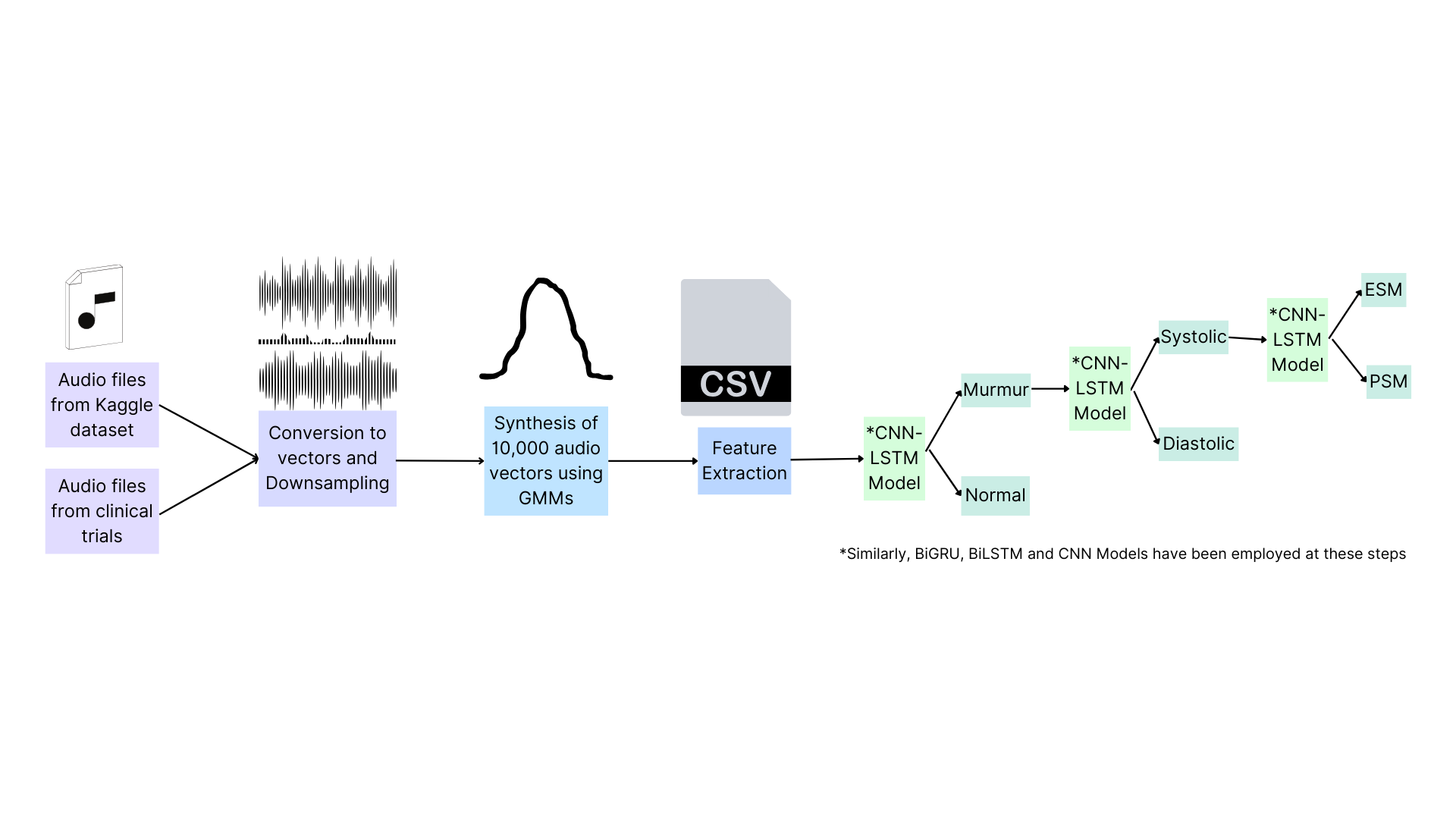
**Fig. 1** Classification of Heart Sounds into Normal and Murmurs, and further classifications to ESM and PSM

In our work heartbeat is divided into two categories: murmurs, and normal heart sounds. We next divide the murmurs into systolic and diastolic murmurs. We further classify systolic murmurs into PSM and ESM. We don’t need further categorization of diastolic murmurs, as most of the murmurs in this category are pathologic in nature and hence severe[30]. The categorization shown in Fig. 1 is the one identified by most of the doctors when they first examine a patient using a stethoscope. It is good enough to manifest evidence for a variety of heart disease. This procedure involves three classification steps to progressively refine the detection and categorization of heart sounds.

The complete procedure used to perform the classification in Mysteth is explained into two main parts: Data Preprocessing and the MySteth Architecture.

**Part A: Data Collection, Labeling, Preprocessing, Refining, and Data Synthesis**

This part includes the steps shown in Fig. 2, as explained below, which are applied on the original dataset to build a suitable Neural Network Model.

**Fig. 2** Preliminary steps for Data Preprocessing on the original dataset available

1. **Data Collection:** The authors used a publicly available Kaggle dataset (https://www.kaggle.com/kinguistics/heartbeat-sounds) to identify murmurs in heartbeat sound audios. The dataset contains 832 distinct heartbeats, of which 150 were selected for the use case. This dataset was gathered from the general public via the stethoscope Pro iPhone app and a clinic trial in hospitals using the digital stethoscope, DigiScope.
2. **Data Labeling:** The publicly available dataset (https://www.kaggle.com/kinguistics/heartbeat-sounds) was annotated by Dr. Nishant Thakur, a super-specialised cardiologist from Max Hospital, I.P. Extension, Delhi, India, and re-annotated and cross-checked by Dr. Rajat Jain, a super-specialised cardiologist from Safdarjung Hospital, Delhi, India. Since the dataset is publically available so no ethical approvals were required.
3. **Audio Processing and Refining:** Raw audios, sampled at 22050 Hz, were downsampled to 4 kHz. This downsampling reduces computational load and storage requirements while retaining essential information for heartbeat analysis[31]. Only the first 3 seconds of each audio were preserved to capture a complete cardiac cycle (S1 to S2 to S1), ensuring that the analysis encompasses all critical heart sounds . Audios shorter than 3 seconds were repeated to reach or exceed the 3-second duration, maintaining consistency in input length for the model. The study transformed audio signals into numerical data through the extraction of distinct features representative of signal characteristics, including amplitude, frequency, and duration, using the librosa library. Librosa is a widely-used Python library for audio analysis, known for its robust feature extraction capabilities, which facilitate effective signal characterization for subsequent classification[32].
4. **Data Synthesis:** Given the small initial dataset, Gaussian Mixture Models (GMM) were used to increase the dataset size to 10,000 audio vectors. This approach is beneficial as GMMs can generate new, realistic data points by modeling the probability distribution of the existing data, thus enhancing the dataset without additional data collection efforts[33]. Out of the 10,000 vectors, 3,600 were murmurs, out of which 730 were systolic murmurs. GMMs were used again to increase the number of audio vectors represented by systolic murmurs to a size of 5,000. This targeted augmentation ensures that the dataset is well-balanced, particularly for the systolic murmur class, which is crucial for training a robust and unbiased classification model[34].
5. **Model Training:** Various models were trained on the refined datasets obtained from each of the following classification tasks. The train test ratios for all tasks were kept constant at a 70–30 percent split:
6. Classification Task 1: Applied on the original dataset to separate the heartbeat sounds into normal heartbeats and murmurs
7. Classification Task 2: Applied on the Murmurs obtained from classification task 1 to obtain systolic and diastolic murmurs
8. Classification Task 3: Applied on the Systolic Murmurs obtained from classification task 2 to divide them into Pansystolic Murmurs (PSM) and Ejection Systolic Murmurs (ESM)

Various neural network models were applied on the dataset to obtain the best possible results:

1. CNN-LSTM: CNNs can reduce noise by focusing on important features through convolutional filters, which makes the subsequent LSTM layers more effective in learning the temporal dependencies of the cleaned signal[35]. Details of the models are as follows:
2. Input Layer: Processed numerical data representing the heartbeat audio signals.
3. Intermediate CNN and LSTM Layers as shown in Table 1.
4. Output Layer: Softmax activation function to classify the audio signals into categories (e.g., normal heartbeat, murmur).
5. BiLSTM: BiLSTMs have been successfully applied to various medical signal classification tasks, including ECG and phonocardiography (PCG) signals. Their effectiveness in capturing the temporal dynamics and dependencies in such data makes them a reliable choice for heartbeat classification[36]. The details of the model are as follows:
6. Input Layer: Processed numerical data representing the heartbeat audio signals.
7. Intermediate BiLSTM and Dense Layers as shown in Table 1.
8. Output Layer of Size 2 Units: Softmax activation to classify the audio signals into categories (e.g., systolic murmur, diastolic murmur).
9. CNN: Heartbeat signals can exhibit significant variability in both time and frequency domains. CNNs, with their ability to apply convolutional filters across the input signal, can robustly handle such variations and capture essential characteristics of the heartbeat patterns[37]. The details of the model are as follows:
10. Input Layer: Processed numerical data representing the heartbeat audio signals.
11. Intermediate CNN Layers as shown in Table 1.
12. Output Layer: Softmax activation function to provide classification probabilities. (e.g., Normal Heartbeat, Murmur)
13. BiGRU: Heartbeat signals are sequential in nature, and it is crucial to capture the temporal dependencies within the data. BiGRUs can process the input in both forward and backward directions, capturing dependencies from both past and future contexts, which is particularly beneficial for heartbeat classification[38]. The details of the models are as follows:
14. Input Layer: Processed numerical data representing the heartbeat audio signals.
15. Intermediate BiLSTM and Dense Layers as shown in Table 1.
16. Output Layer of Size 2 Units: Softmax activation to classify the audio signals into categories (e.g., systolic murmur, diastolic murmur).

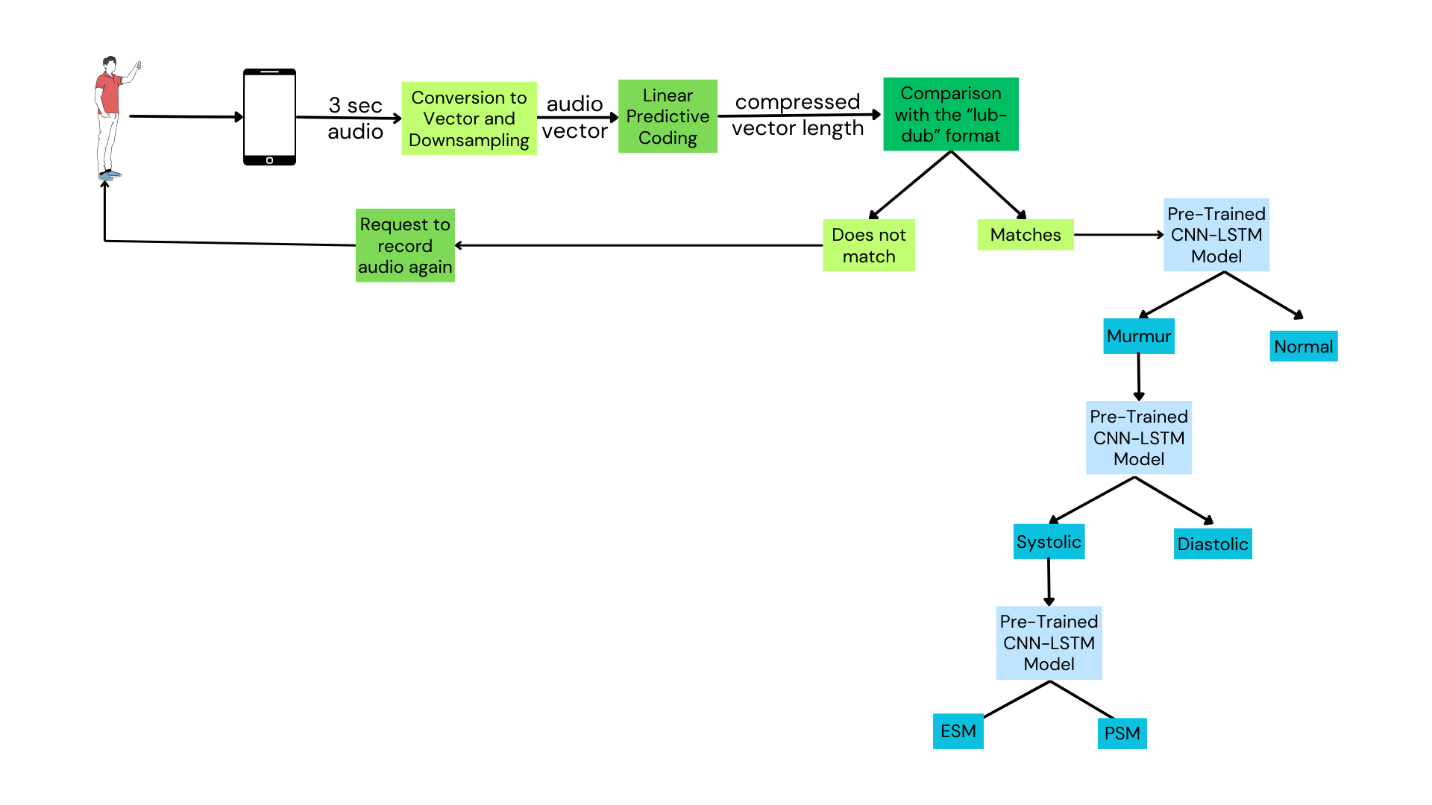
**Table 1** Model Architectures employed for training on the audio vector and feature dataset.

|  |  |  |
| --- | --- | --- |
| S. No. | Model | Architecture |
| 1 | CNN-LSTM | Layer 1: CNN Layer with 9 filters, ReLU activation  Layer 2: CNN Layer with 64 filters, ReLU activation  Layer 3: CNN Layer with 32 filters, ReLU activation  Layer 4: LSTM Layer with 8 neurons, tanh activation (default)  Layer 5: LSTM Layer with 4 neurons, tanh activation (default) |
| 2 | BiLSTM | Layer 1: BiLSTM Layer with 128 neurons, tanh activation (default)  Layer 2: BiLSTM Layer with 64 neurons, tanh activation (default)  Layer 3: Dense Layer with 64 neurons, ReLU activation  Layer 4: Dense Layer with 32 neurons, ReLU activation |
| 3 | CNN | Layer 1: CNN Layer with 9 filters, ReLU activation  Layer 2: CNN Layer with 64 filters, ReLU activation  Layer 3: CNN Layer with 32 filters, ReLU activation |
| 4 | BiGRU | Layer 1: BiGRU Layer with 128 neurons, tanh activation (default)  Layer 2: BiGRU Layer with 64 neurons, tanh activation (default)  Layer 3: Dense Layer with 64 neurons, ReLU activation  Layer 4: Dense Layer with 32 neurons, ReLU activation |

The results obtained from the models for each of the classification are articulated in the Results section. As the CNN-LSTM model gave the best accuracy scores, it was employed in the MySteth Architecture.

**Part B: MySteth Architecture**

The MySteth Architecture was designed to handle the refined datasets and perform the classifications at each step. The following steps are part of the process, which is depicted in Fig. 3:



**Fig. 3** MySteth Architecture combining CNNs and LSTMs to classify heartbeats

1. **Recording:** A person records their heartbeat using a smartphone in a silent environment, capturing a 3-second audio clip.
2. **Downsampling and Compression:** Linear Predictive Coding (LPC) facilitates a two-step procedure of downsampling and compression of the recorded audio before training the models. By lowering the audio's sampling rate, downsampling effectively minimizes the amount of data and computing load while preserving crucial information. The spectral envelope of the digital voice signals is then compressed using LPC compression. By preserving important components of the heart sounds, this approach improves the efficacy of feature extraction[39]. LPC minimizes the amount of data while preserving important information, which makes processing and analyzing the cardiac sounds simpler and quicker. In order to ensure that deep learning models can effectively capture and learn from the key elements of the heart sounds throughout the training phase, this preliminary step optimizes the data for the models[40]. The integration of LPC for downsampling and compression in the preprocessing pipeline optimizes the data for deep learning models[41].
3. **Heartbeat Signal Verification:** The extracted features are compared to a reference model of a normal "lub-dub" heartbeat. The patient is asked to re-record the heartbeat if the signal does not match the classic "lub-dub" pattern.
4. **Model Training:** The audio vector and the extracted featurespasses through a pre-trained CNN-LSTM model that has shown the best results (as tabulated in Tables 1, 2 and 3). This model had been trained previously on huge amounts of data (as explained in Part A). CNN and LSTM were applied serially due to the complementary nature of their roles in feature extraction and sequence modeling. The CNNs can preprocess and distill the essential features, which the LSTMs can then analyze in a temporal context[42].
5. Static feature extraction is done by Convolutional Neural Networks (CNN).
6. Temporal characteristics are extracted using Long Short-Term Memory (LSTM).
7. Three convolution layers with kernel sizes of 9, 64, and 32 are applied, followed by batch normalization after each convolution layer. The details of the layers are given in Part A.
8. For LSTM-based models, two layers of sizes 8 and 4 are added, followed by a dense layer. The details of the layers are given in Part A.
9. **Classification Task 1:** Two categories are created from the processed audio: murmur and normal heartbeat.
10. **Classification Task 2:** Murmurs are further divided into diastolic and systolic forms.
11. **Classification Task 3:** Systolic Murmurs are classified into Pansystolic Murmur (PSM) and Ejection Systolic Murmur (ESM).

**Results**

This section presents the results obtained from classifying heartbeat sounds using different models (as explained in Part A of the Methods section) into the following categories: Normal heartbeat, Murmurs, Systolic Murmur, Diastolic Murmur, PSM (PanSystolic Murmur), and ESM (Ejection Systolic Murmur).

1. **Classification results for Normal Heartbeat, and Murmur:** The classification accuracy for Normal Heartbeat and Murmurs was evaluated using different models on the compressed audio representations obtained through Linear Predictive Coding (LPC). Table 2 provides an overview of the outcomes.

The classification of heartbeat sounds is an essential task in the medical field, as it helps healthcare professionals diagnose various cardiovascular conditions. Although numerous research activities have been carried out to enhance the precision of heartbeat sound categorization, the majority of them have concentrated on refining data pre-processing methods or employing a single primary method such as neural networks, support vector machines, or hidden Markov models[43].

**Table 2** Accuracy scores for Mysteth for classifying heart sounds into Normal Heartbeat, and Murmur.

|  |  |
| --- | --- |
| **Model** | **Accuracy** |
| BiLSTM | 68% |
| CNN | 72% |
| BiGRU | 88% |
| CNN and LSTM | 92% |

1. **Classification results for systolic murmur, and diastolic murmur:** The classification accuracy for distinguishing between Systolic and Diastolic Murmurs was assessed using various models on the compressed representations of the murmur audio. Table 3 displays the results.

**Table 3** Accuracy scores for Mysteth for classifying murmurs into Systolic Murmur and Diastolic Murmur.

|  |  |
| --- | --- |
| **Model** | **Accuracy** |
| CNN | 68% |
| BiLSTM | 72% |
| BiGRU | 84% |
| CNN and LSTM | 91% |

1. **Classification results for PSM, and ESM:** The accuracy for further classifying systolic murmurs into PSM and ESM was evaluated using various models on the compressed feature representations obtained from LPC feature extraction method. Table 4 displays the outcomes.

**Table 4** Accuracy scores for Mysteth for classifying systolic murmurs further into PSM, ESM using various models.

| **Model** | **Accuracy** |
| --- | --- |
| BiLSTM | 51% |
| BiGRU | 62% |
| CNN and LSTM | 71% |
| CNN | 90% |

A hybrid classifier can significantly enhance classification accuracy.When combined in a hybrid CNN-LSTM model, these models can effectively extract deep features and contextual time data from Phonocardiogram (PCG) signals. The CNN component handles feature extraction, while the LSTM module extracts time-dependent features[44].

This hybrid approach has been shown to outperform single CNN or LSTM-based methods, producing richer and more concentrated models with higher performance and fewer parameters. These findings demonstrate how well different recurrent neural networks function in conjunction with convolutional neural networks to tackle challenging audio categorization problems. The utilization of LPC for feature extraction significantly contributes to the models' performance, especially in distinguishing subtle differences in heart sounds[45].

**Discussion**

The introduction of MySteth as an innovative at-home heart diagnostic tool represents an advancement in the field of cardiac care, addressing critical gaps in the accessibility and convenience of preliminary heart investigations. This discussion focuses on the unique aspects and justifications for our approach, emphasizing the integration of CNN-LSTM architectures with Linear Predictive Coding (LPC) preprocessing, and the impact of these choices on the efficacy and practicality of MySteth.

The primary motivation for employing a hybrid CNN-LSTM model stems from the complementary strengths of Convolutional Neural Networks (CNNs) and Long Short-Term Memory networks (LSTMs) in handling the complexities of heartbeat sound classification[46]. CNNs are adept at extracting spatial features from the input data, capturing local patterns and significant characteristics of the heart sounds. This capability is crucial for identifying the nuanced features present in heartbeat signals, such as murmurs and other anomalies. LSTMs, on the other hand, excel at modeling temporal dependencies and sequential patterns within the data. By integrating LSTMs with CNNs, we ensure that the model not only recognizes spatial features but also understands how these features evolve over time[47]. This combination is particularly effective for analyzing heartbeat sounds, which inherently possess both spatial and temporal dimensions.

Linear Predictive Coding (LPC) plays a pivotal role in our approach by facilitating data compression and enhancing feature extraction. LPC reduces the complexity of the raw audio data while preserving essential information, making the subsequent processing by CNN and LSTM layers more efficient. This preprocessing step is crucial for improving the model's ability to detect subtle patterns and anomalies in the cardiac sounds, thereby enhancing classification accuracy[48]. By incorporating LPC, we address the challenge of high data volume and computational load, enabling the use of advanced deep learning models even in resource-constrained environments. This efficiency is particularly beneficial for at-home diagnostic tools like MySteth, where minimal hardware requirements and quick processing are critical for user adoption and practicality[49].

Compared to conventional methods of heartbeat classification, which often rely on manual feature extraction and traditional machine learning algorithms, our CNN-LSTM approach offers several distinct advantages. Traditional methods can be limited by their dependency on handcrafted features and their inability to fully capture the complexity of the heartbeat signals. In contrast, deep learning models, particularly the CNN-LSTM combination, automatically learn relevant features from the data, leading to more accurate and robust classifications[50]. Moreover, the ability to handle large and complex datasets without significant manual intervention makes our approach more scalable and adaptable to different healthcare settings. High accuracy rates have been achieved by our models, including the exceptional performance of the CNN-LSTM model with a 92% accuracy for classifying the heartbeats into normal and murmurs, and 91% for classifying the murmurs into systolic and diastolic murmurs. This underscores the effectiveness of our method in differentiating between normal and pathological heart sounds as well as finer distinctions such as various types of murmurs.

The minimal hardware requirements and straightforward implementation of MySteth mean it can be readily adopted in various healthcare environments, including remote or under-resourced areas. This accessibility addresses a critical need in global healthcare, providing reliable and early detection tools for heart disease, which remains a leading cause of mortality worldwide[51].

While our work presents significant advancements, there are limitations and areas for improvement. As already stated, not much study has been done in the area of categorizing mobile phone heartbeat sounds[52]. To improve research in this area, a larger and more realistic dataset must be created[53]. The models are trained and validated on specific datasets which may not encompass the full variability seen in global populations. Future work should focus on incorporating more diverse datasets to enhance generalizability. In this work, the audio was encoded using linear predictive coding. To compress audio, more encoding methods can be employed, such as auto-encoders. While our method is efficient, optimizing it further for real-time processing and deployment on portable devices could enhance its practical application. Future research should explore seamless integration with existing clinical workflows, ensuring that the technology is user-friendly for healthcare professionals. To confirm the models' long-term dependability and efficacy in practical situations, longitudinal research and comprehensive clinical trials are required[54].

**Conclusions**

MySteth is a tool that our study introduces. Using deep learning algorithms and just the sound of a heartbeat recorded using a phone or digital stethoscope, the authors investigated the screening of a broad class of heart disorders. Heartbeats can be categorized by MySteth into three categories: normal, systolic, and diastolic murmurs. Systolic murmurs can also be further classified into two categories: Ejection Systolic Murmur (ESM) and PanSystolic Murmur (PSM).

In order to keep the condition from getting worse to the point where it becomes fatal or irreversible, this effort can be very helpful in identifying the onset of a wide class of cardiovascular heart diseases[55]. MySteth, a tool in the field of heart sound classification, can significantly contribute to preventive healthcare and lower the total burden of cardiovascular illnesses by enabling at-home early screening and precise diagnosis of heart murmurs and other irregularities. Because of its architecture, the instrument can be deployed in a variety of contexts, such as remote and rural locations, allowing disadvantaged groups to benefit from modern diagnostics. The scalable and adaptable nature of MySteth ensures it can be integrated into different healthcare environments, from large urban hospitals to small rural clinics and even home-based care. By addressing both the technological and practical challenges in this domain, MySteth stands out as a viable solution with significant potential for improving global health outcomes.

**Declarations**

**Abbreviations**

HS Heart Sounds

ESM Ejection Systolic Murmurs

PSM PanSystolic Murmurs

CVD Cardiovascular Disease

CABG Coronary Artery Bypass Grafting

MFCC Mel-Frequency Cepstral Coefficient

RMSE Root Mean Squared Error

ECG Electrocardiogram

CT Computed Tomography

GMM Gaussian Mixture Model

STFT Short-Term Fourier Transformation

LPC Linear Predictive Coding

PCG Phonocardiogram

CNN Convolutional Neural Networks

LSTM Long Short Term Memory

**Ethics approval and consent to participate**

This study uses publicly available Kaggle dataset (https://www.kaggle.com/kinguistics/heartbeat-sounds) to identify murmurs in heartbeat sound audios. As the dataset is in the public domain, so no ethical approval was required.

**Consent for publication**

Not applicable

**Availability of data and materials**

The authors used a publicly available Kaggle dataset (https://www.kaggle.com/kinguistics/heartbeat-sounds) to identify murmurs in heartbeat sound audios. The dataset contains 832 distinct heartbeats, of which 150 were selected for the use case. This dataset was gathered from the general public via the stethoscope Pro iPhone app and a clinic trial in hospitals using the digital stethoscope, DigiScope. The dataset was annotated by Dr. Nishant Thakur, a super-specialised cardiologist from Max Hospital, I.P. Extension, Delhi, India, and re-annotated and cross-checked by Dr. Rajat Jain, a super-specialised cardiologist from Safdarjung Hospital, Delhi, India. Annotated data can be made available on reasonable request.

**Competing interests**

The authors declare that they have no competing interests.

**Funding**

There was no financial support provided for this study. APC will be supported by Molde University College, Norway.

**Author’s Contribution**

K.J. (Kopal Jain): Conducted the primary research, developed the model and improved the first draft of the manucript.

R.J. (Rohit Jain): Conducted the primary research, wrote the initial draft of the manuscript and also coordinated the data annotation process.

S.K.M. (Salik Khwaja Mohammad): Conducted the primary research, wrote the initial draft of the manuscript and also coordinated the data annotation process.

S.A. (Swati Aggarwal): Designed the study framework, provided overall guidance and supervision for the project. S.A. reviewed and edited the manuscript, ensuring the scientific rigor and coherence of the study. S.A. also handled the correspondence and submission process.

All authors reviewed and approved the final manuscript.

**Acknowledgements**

The authors would like to thank Vishal Kumar Lal, Department of Biosciences and Biotechnology, Indian Institute of Technology, Kharagpur, for his assistance in creating a synthetic dataset using the Librosa library, extracting advanced audio features, data preprocessing, and applying various classical ML models.

**References**

[1] Roy JK, Roy TS, Mukhopadhyay SC. Heart sound: Detection and analytical approach towards diseases. Modern Sensing Technologies. 2019:103-45.

[2] Kumar D, Carvalho P, Couceiro R, Antunes M, Paiva RP, Henriques J. Heart murmur classification using complexity signatures. In2010 20th International Conference on Pattern Recognition 2010 Aug 23 (pp. 2564-2567). IEEE.

[3] Reimer-Kent J. Heart Sounds: Are You Listening? Part 2. Canadian Journal of Cardiovascular Nursing. 2013 Aug 1;23(3).

[4] Walker HK, Hall WD, Hurst JW. Clinical methods: the history, physical, and laboratory examinations.

[5] Deng L, O'Shaughnessy D. Speech processing: a dynamic and optimization-oriented approach. CRC Press; 2003 Jun 18.

[6] Patnaik AN. The diastolic murmurs. Indian Journal of Cardiovascular Disease in Women-WINCARS. 2019 Dec;4(04):228-32.

[7] Nishimura RA, Otto CM, Bonow RO, Carabello BA, Erwin III JP, Guyton RA, O’Gara PT, Ruiz CE, Skubas NJ, Sorajja P, Sundt III TM. 2014 AHA/ACC guideline for the management of patients with valvular heart disease: executive summary: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. Circulation. 2014 Jun 10;129(23):2440-92.

[8] Benjamin EJ, Muntner P, Alonso A, Bittencourt MS, Callaway CW, Carson AP, Chamberlain AM, Chang AR, Cheng S, Das SR, Delling FN. Heart disease and stroke statistics—2019 update: a report from the American Heart Association. Circulation. 2019 Mar 5;139(10):e56-28.

[9] Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nature medicine. 2019 Jan;25(1):44-56.

[10] Hussain MM, Rafi U, Imran A, Rehman MU, Abbas SK. Risk Factors Associated with Cardiovascular Disorders: Risk Factors Associated with Cardiovascular Disorders. Pakistan BioMedical Journal. 2024 Feb 29:03-10.

[11] Satler LF, Green CE, Wallace RB, Rackley CE. Coronary artery disease in the elderly. The American journal of cardiology. 1989 Jan 15;63(3):245-8.

[12] Vasan RS. Biomarkers of cardiovascular disease: molecular basis and practical considerations. Circulation. 2006 May 16;113(19):2335-62.

[13] Guasti L, Dilaveris P, Mamas MA, Richter D, Christodorescu R, Lumens J, Schuuring MJ, Carugo S, Afilalo J, Ferrini M, Asteggiano R. Digital health in older adults for the prevention and management of cardiovascular diseases and frailty. A clinical consensus statement from the ESC Council for Cardiology Practice/Taskforce on Geriatric Cardiology, the ESC Digital Health Committee and the ESC Working Group on e‐Cardiology. ESC heart failure. 2022 Oct;9(5):2808-22.

[14] Latif S, Usman M, Rana R, Qadir J. Phonocardiographic sensing using deep learning for abnormal heartbeat detection. IEEE Sensors Journal. 2018 Sep 16;18(22):9393-400.

[15] Altaf A, Mahdin H, Alive AM, Ninggal MI, Altaf A, Javid I. Systematic Review for Phonocardiography Classification Based on Machine Learning. Int. J. Adv. Comput. Sci. Appl. 2023;14:806-17.

[16] Brites IS, da Silva LM, Barbosa JL, Rigo SJ, Correia SD, Leithardt VR. Machine learning and iot applied to cardiovascular diseases identification through heart sounds: A literature review. InInformatics 2021 Oct 30 (Vol. 8, No. 4, p. 73). MDPI.

[17] Abdelmageed S. Wrist-based Phonocardiogram Diagnosis Leveraging Machine Learning.

[18] Alkahtani HK, Haq IU, Ghadi YY, Innab N, Alajmi M, Nurbapa M. Precision Diagnosis: An Automated Method for Detecting Congenital Heart Diseases in Children from Phonocardiogram Signals Employing Deep Neural Network. IEEE Access. 2024 Apr 30.

[19] Li H, Wang X, Liu C, Zeng Q, Zheng Y, Chu X, Yao L, Wang J, Jiao Y, Karmakar C. A fusion framework based on multi-domain features and deep learning features of phonocardiogram for coronary artery disease detection. Computers in biology and medicine. 2020 May 1;120:103733.

[20] Azmeen A, Vakilzadian H, Haider H, Mathers DH, Zimmerman R, Bedi S, O’Leary EL. Heart sounds: Past, present, and future from a technological and clinical perspective–a systematic review. Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine. 2023 Jun;237(6):669-82.

[21] Chamberlain D, Mofor J, Fletcher R, Kodgule R. Mobile stethoscope and signal processing algorithms for pulmonary screening and diagnostics. In2015 IEEE Global Humanitarian Technology Conference (GHTC) 2015 Oct 8 (pp. 385-392). IEEE.

[22] Luo H, Lamata P, Bazin S, Bautista T, Barclay N, Shahmohammadi M, Lubrecht JM, Delhaas T, Prinzen FW. Smartphone as an electronic stethoscope: factors influencing heart sound quality. European Heart Journal-Digital Health. 2022 Sep 28;3(3):473-80.

[23] Wu YC, Han CC, Chang CS, Chang FL, Chen SF, Shieh TY, Chen HM, Lin JY. Development of an electronic stethoscope and a classification algorithm for cardiopulmonary sounds. Sensors. 2022 Jun 3;22(11):4263.

[24] Dwivedi S, Mohammad A. Heartbeat Pattern and Arrhythmia Classification: A Review. Frontiers in Biomedical Technologies. 2024;11(1):130-48.

[25] Stillman AE, Oudkerk M, Ackerman M, Becker CR, Buszman PE, De Feyter PJ, Hoffmann U, Keadey MT, Marano R, Lipton MJ, Raff GL. Use of multidetector computed tomography for the assessment of acute chest pain: a consensus statement of the North American Society of Cardiac Imaging and the European Society of Cardiac Radiology. The International Journal of Cardiovascular Imaging. 2007 Aug;23:415-27.

[26] Eyre H, Kahn R, Robertson RM, ACS/ADA/AHA Collaborative Writing Committee, ACS/ADA/AHA Collaborative Writing Committee Members, Clark NG, Doyle C, Hong Y, Gansler T, Glynn T, Smith RA. Preventing cancer, cardiovascular disease, and diabetes: a common agenda for the American Cancer Society, the American Diabetes Association, and the American Heart Association. Circulation. 2004 Jun 29;109(25):3244-55.

[27] Moses JC, Adibi S, Angelova M, Islam SM. Smart home technology solutions for cardiovascular diseases: a systematic review. Applied System Innovation. 2022 May 27;5(3):51.

[28] Reddy KS, Mathur MR. Global Burden of CVD: Prevalence, Pattern, and Trends. InHandbook of Global Health 2021 May 12 (pp. 423-437). Cham: Springer International Publishing.

[29] Miller JC, Skoll D, Saxon LA. Home monitoring of cardiac devices in the era of COVID-19. Current Cardiology Reports. 2021 Jan;23:1-9.

[30] Ahlstrom C, Hult P, Rask P, Karlsson JE, Nylander E, Dahlström U, Ask P. Feature extraction for systolic heart murmur classification. Annals of biomedical engineering. 2006 Nov;34:1666-77.

[31] Mesin L. Heartbeat monitoring from adaptively down-sampled electrocardiogram. Computers in Biology and Medicine. 2017 May 1;84:217-25.

[32] Verma P, Yadav A, Sharma SD, Hooda D. Signal Processing and Feature Engineering of Respiratory Disease.

[33] Chen T, Morris J, Martin E. Probability density estimation via an infinite Gaussian mixture model: application to statistical process monitoring. Journal of the Royal Statistical Society Series C: Applied Statistics. 2006 Nov;55(5):699-715.

[34] Moreno-Barea FJ, Franco L, Elizondo D, Grootveld M. Application of data augmentation techniques towards metabolomics. Computers in Biology and Medicine. 2022 Sep 1;148:105916.

[35] Borré A, Seman LO, Camponogara E, Stefenon SF, Mariani VC, Coelho LD. Machine fault detection using a hybrid CNN-LSTM attention-based model. Sensors. 2023 May 5;23(9):4512.

[36] Aldughayfiq B, Ashfaq F, Jhanjhi NZ, Humayun M. A deep learning approach for atrial fibrillation classification using multi-feature time series data from ecg and ppg. Diagnostics. 2023 Jul 21;13(14):2442.

[37] Lee J, Park J, Kim KL, Nam J. SampleCNN: End-to-end deep convolutional neural networks using very small filters for music classification. Applied Sciences. 2018 Jan 22;8(1):150.

[38] Chen J, Shen M, Ma W, Zheng W. A spatio-temporal learning-based model for sleep apnea detection using single-lead ECG signals. Frontiers in Neuroscience. 2022 Aug 5;16:972581.

[39] Ibarra-Hernández RF, Alonso-Arévalo MA, Cruz-Gutiérrez A, Licona-Chávez AL, Villarreal-Reyes S. Design and evaluation of a parametric model for cardiac sounds. Computers in biology and medicine. 2017 Oct 1;89:170-80.

[40] Shipp S. Neural elements for predictive coding. Frontiers in psychology. 2016 Nov 18;7:1792.

[41] Singh AK, Krishnan S. ECG signal feature extraction trends in methods and applications. BioMedical Engineering OnLine. 2023 Mar 8;22(1):22.

[42] Ahmad T, Wu J, Alwageed HS, Khan F, Khan J, Lee Y. Human activity recognition based on deep-temporal learning using convolution neural networks features and bidirectional gated recurrent unit with features selection. IEEE access. 2023 Mar 30;11:33148-59.

[43] Chen J, Guo Z, Xu X, Jeon G, Camacho D. Artificial intelligence for heart sound classification: A review. Expert Systems. 2024 Apr;41(4):e13535.

[44] Koşar E, Barshan B. A new CNN-LSTM architecture for activity recognition employing wearable motion sensor data: Enabling diverse feature extraction. Engineering Applications of Artificial Intelligence. 2023 Sep 1;124:106529.

[45] Khan FA, Abid A, Khan MS. Automatic heart sound classification from segmented/unsegmented phonocardiogram signals using time and frequency features. Physiological measurement. 2020 Jun 2;41(5):055006.

[46] Lyu X, Rani S, Manimurugan S, Maple C, Feng Y. A Deep Neuro-Fuzzy Method for ECG Big Data Analysis via Exploring Multimodal Feature Fusion. IEEE Transactions on Fuzzy Systems. 2024 Jun 18.

[47] Barros B, Lacerda P, Albuquerque C, Conci A. Pulmonary COVID-19: learning spatiotemporal features combining CNN and LSTM networks for lung ultrasound video classification. Sensors. 2021 Aug 14;21(16):5486.

[48] Paschalis SM, Hutapea DK, Bachri KO. Heart Sound Processing for Early Diagnostic of Heart Abnormalities using Support Vector Machine. Jurnal ELTIKOM: Jurnal Teknik Elektro, Teknologi Informasi dan Komputer. 2024 Jun 30;8(1):57-65.

[49] Benjamin EJ, Virani SS, Callaway CW, Chamberlain AM, Chang AR, Cheng S, Chiuve SE, Cushman M, Delling FN, Deo R, De Ferranti SD. Heart disease and stroke statistics—2018 update: a report from the American Heart Association. circulation. 2018 Mar 20;137(12):e67-492.

[50] Chen C, Hua Z, Zhang R, Liu G, Wen W. Automated arrhythmia classification based on a combination network of CNN and LSTM. Biomedical Signal Processing and Control. 2020 Mar 1;57:101819.

[51] Nowbar AN, Gitto M, Howard JP, Francis DP, Al-Lamee R. Mortality from ischemic heart disease: Analysis of data from the World Health Organization and coronary artery disease risk factors From NCD Risk Factor Collaboration. Circulation: cardiovascular quality and outcomes. 2019 Jun;12(6):e005375.

[52] Li KH, White FA, Tipoe T, Liu T, Wong MC, Jesuthasan A, Baranchuk A, Tse G, Yan BP. The current state of mobile phone apps for monitoring heart rate, heart rate variability, and atrial fibrillation: narrative review. JMIR mHealth and uHealth. 2019 Feb 15;7(2):e11606.

[53] De Chazal P, O'Dwyer M, Reilly RB. Automatic classification of heartbeats using ECG morphology and heartbeat interval features. IEEE transactions on biomedical engineering. 2004 Jun 21;51(7):1196-206.

[54] Mallinckrod CH, Lane PW, Schnell D, Peng Y, Mancuso JP. Recommendations for the primary analysis of continuous endpoints in longitudinal clinical trials. Drug Information Journal. 2008 Jul;42(4):303-19.

[55] Nashif S, Raihan MR, Islam MR, Imam MH. Heart disease detection by using machine learning algorithms and a real-time cardiovascular health monitoring system. World Journal of Engineering and Technology. 2018 Sep 12;6(4):854-73.