Machine Learning and Deep Learning Based Approach for Analyzing Heartbeat Sounds

Dr. C. Siva Balaji Yadav[1], Muddangala Nandini[2], Malepati Vijay Sai[3], Gurijala Raghunatha Reddy[4], Elakaturu Vinay[5]

[1] *Professor, Department of Artificial Intelligence and Machine Learning Annamacharya Institute of Technology & Sciences, Tirupati, India*

[2,3,4,5] *Student, Department of Artificial Intelligence and Machine Learning Annamacharya Institute of Technology & Sciences, Tirupati, India*

***Abstract*—Machine learning and Deep learning have been increasing at a fast rate in various areas particularly the health care sector. Heart diseases are still among the major causes of death in the world, and timely and correct diagnosis of them is vital to enhance better patient outcomes. This paper is on automatic categorization of heartbeat sounds with machine learning and deep learning-based methods. The various feature extraction methods are examined to determine their effects on classification. To this end, three categories of features will be looked into: traditional audio signal processing features, deep learning features purported by pre-trained models, and a composite grouping of the two audio and deep learning features. The features obtained are then classified into Support Vector Machine (SVM), Random Forest and XGBoost classifier. The use of Principal Component Analysis (PCA) to reduce dimensions is used in order to manage the large dimensional feature space. Moreover, there are feature concatenation and majority voting methods used that enhance the accuracy and robustness of classification. The results of the experiments prove the effectiveness of the hybrid feature approach, since it is better than the separate sets of features, which provides the evidence of the efficiency of combining machine learning with deep learning methods in analyzing the sound of the heartbeat.**

***IndexTerms*—Heart Sound Classification, Spectrogram Analysis, Machine Learning, Feature Engineering, Support Vector Machine, Deep Learning, Transfer Learning.**

I. INTRODUCTION

One of the most frequent causes of death in the world is cardiovascular diseases, so it is important to diagnose these diseases as early as possible and precisely. The heartbeat sounds are significant data regarding the health of the heart and they are normally determined by listening to the heart with the help of manual ability. As the use of Machine Learning (ML) and Deep Learning (DL) methods develops at a fast rate, smart models will be able to process biomedical signals and enhance the accuracy of diagnosis. The conventional audio signal processing systems employ hand designed features, whereas deep learning models automatically discover high level features based on data. The combination of these methods can improve the performance of classification due to complementary information. The system integrates the classical audio functionality and deep learning-based functionality, and the next stage is the classification process based on the Support Vector Machine (SVM), Random Forest, and XGBoost algorithms. The Principal Component Analysis (PCA) is used to streamline the dimensionality of features and enhance faster computation.

1.1 OBJECTIVES

1.1.1 Develop a Hybrid Machine Learning and Deep Learning Framework for Heart Sound Classification

A prominent goal is to compare the effects of various feature extraction methods, such as spectral and time-domain features, deep features obtained in the course of trained models, and combinations of features. Principal Component Analysis (PCA) will be used to decrease the number of features, decrease redundancy and increase computing performance.

1.1.2 Improve Early and Automated Cardiac Disease Detection

The long-term goal of the given research is to deliver a reliable AI-related diagnostic support device that will help medical workers spot cardiac anomalies early. The system should be able to scale to both clinical and remote healthcare settings.

1.1.3 Develop a Robust and Generalizable Heart Sound Classification Model

The other goal is to come up with a strong heartbeat sound recognition model that can guarantee the same results to be obtained with and without the conditions of noise at different levels as well as with and without the conditions of the signal and thus will be able to be used in real-life healthcare settings.

1.2 SCOPE

1.2.1 Clinical Focus Heart Sound Analysis with Machine Learning and Deep Learning

The proposed research is focused on comparing and categorizing heart beat sounds with the help of the classical machine learning models and deep learning. It contains the necessary preprocessing processes like noise elimination, normalization, segmentation and spectrogram development to improve the quality of signals.

1.2.2 Comparative Multiple Classifier Evaluation

It makes comparisons of the various classifiers such as Support Vector Machine (SVM), Random Forest, and XGBoost to identify the best classifier in classification of heartbeat sounds. Their performance evaluation is conducted based on such measures as accuracy, precision, recall, and F1-score so that the comparison is fair and detailed.

1.2.3 Hybrid Performance Enhancement through Integration

The study also examines the application of handcrafted audio properties alongside the use of deep learning models in extracting features. The purpose of this hybrid method is to obtain more meaningful patterns of heart sounds.

1.2.4 Clinical and Remote Healthcare Potential Application

Even though the contemporary system operates based on the recorded heart sounds, the structure can be expanded in the future. It can be used to facilitate the real-time diagnosis, mobile health apps and telemedicine platforms in order to provide a broader access to healthcare.

II. LITERATURE SURVEY

2.1 TRADITIONAL METHODS OF HEART SOUND ANALYSIS

Historically, heart sound analysis was performed manually through clinical auscultation and basic signal processing techniques. With the introduction of digital phonocardiograms (PCG), classical machine learning methods were applied using handcrafted acoustic features such as time-domain, frequency-domain, and statistical descriptors.

2.1.1 Dependence on Manual Interpretation

Traditional auscultation heavily relies on the expertise of the healthcare practitioner to detect abnormalities such as heart murmurs or irregular rhythms. Diagnostic accuracy often varies between experienced cardiologists and medical trainees. Because of this subjectivity, there is always a possibility of misinterpretation or incorrect diagnosis.

2.1.2 Sensitivity to Noise and Signal Variability

Early signal processing techniques were highly sensitive to background noise, recording artifacts, and variations in signal quality. Inconsistent recording conditions significantly affected feature extraction and reduced classification reliability.

2.1.3 Limitations of Handcrafted Feature Engineering

Classical machine learning models depended on manually designed features, including spectral energy, zero-crossing rate, and Mel-frequency cepstral coefficients (MFCCs). Although these features captured important characteristics of heart sounds, they often failed to represent complex and subtle cardiac patterns.

2.2 ADVANCES IN MACHINE LEARNING AND DEEP LEARNING FOR HEART SOUND CLASSIFICATION

With rapid advancements in Machine Learning (ML) and Deep Learning (DL), heart sound analysis has significantly evolved. Modern approaches go beyond simple handcrafted features and use advanced techniques to automatically learn meaningful patterns from audio signals. These systems combine improved feature extraction methods with powerful classifiers to achieve higher diagnostic accuracy and robustness.

2.2.1 Emergence of Hybrid Feature Extraction

Recent studies indicate that combining traditional handcrafted features with deep learning–based representations improves classification performance.

Converting heart sounds into spectrograms or time-frequency representations allows models to capture both temporal and frequency-related information.

2.2.2 Application of Advanced Classifiers

Algorithms such as Support Vector Machine (SVM), Random Forest, and XGBoost have shown strong performance in heart sound classification tasks. These classifiers provide improved decision boundaries, better handling of nonlinear patterns, and enhanced robustness compared to earlier statistical methods.

2.2.3 Dimensionality Reduction and Feature Optimization

To address high-dimensional feature spaces, techniques such as Principal Component Analysis (PCA) are widely used. Dimensionality reduction improves computational efficiency, reduces redundancy, and enhances generalization capability.

2.2.4 Ensemble and Voting-Based Approaches

Recent research highlights the effectiveness of ensemble strategies, including feature concatenation and majority voting, to improve classification stability and minimize misclassification rates in diverse datasets.

2.3 APPLICATIONS AND CHALLENGES IN HEART SOUND CLASSIFICATION

2.3.1 Applications in Healthcare

Automated heart sound classification systems play a significant role in the early detection of cardiovascular diseases. They are used in remote patient monitoring, telemedicine, and clinical decision-support systems. These systems provide objective and consistent diagnostic insights to healthcare professionals and are particularly valuable in rural or resource-limited regions.

2.3.2 Challenges in Implementation

Heart sound signals vary among patients due to differences in age, physiology, and health conditions. Background noise and variations in recording devices can further affect system performance. Additionally, imbalanced datasets and high-dimensional feature representations may increase computational complexity and reduce classification accuracy.

2.3.3 Need for Robust and Scalable Frameworks

A reliable framework should combine proper preprocessing, effective feature extraction, dimensionality reduction, and ensemble learning methods. Such integration ensures the development of a practical and dependable solutions for the many real-world healthcare applications.

III. METHODOLOGY

3.1 DATASET PREPARATION

The data is necessary in building a correct heartbeat sound classification system. The PhysioNet Challenge Heart Sound Dataset has been used in this study because it is a commonly adopted benchmark in studies. It includes phonocardiogram (PCG) recordings that were recorded in various clinical environments. These are recorded with a range of devices which guarantees a variation of signal quality. The data is divided into different classes of the Normal and Abnormal heart sounds. The abnormal category is comprised of Murmur, Extra Heart Sounds (ExtraHLS) and Artifact recordings. Murmurs result because of the abnormal flow of blood in the heart. ExtraHLS are additional sounds that are on top of the usual S1 and S2 sounds of the heartbeats. Noise or environmental interference is the primary content of artifact recordings. The signals are then processed with filtering, normalization and segmentation before being utilized to extract the features and classify them using machine learning.

3.2 SYSTEM ARCHITECTURE

The proposed system of heartbeat sound classification is being planned stepwise, beginning with the collection of data till the ultimate prediction. To start with, labeled phonocardiogram (PCG) recordings are recorded based on the dataset. These recordings are of different categories including, normal, murmur, Extra Heart sounds and Artifact. During pre-processing the tools used are noise filtering, normalization and segmentation to eliminate unwanted disturbances and emphasize on important cardiac cycles. The signals are also converted into time-domain and frequency-domain representations, such as spectrograms, to represent key sound patterns in a better way. The data is then used to generate 80% of training data and 20% of testing data to adequately assess the performance of the model. Two methods are applied in order to extract the features. The first technique is extraction of handcrafted audio features and the latest technique is a deep learning-based representations acquired through pre-trained models. The Principal Component Analysis (PCA) is used to reduce the complexity and eliminate the redundancy of the information so that the model can become more efficient and can generalize it. The machine learning classifiers involve Support Vector Machine (SVM), random forest, and XGBoost, which are used in classification in the stage of classifying the heart sounds into the relevant categories. A new heart sound recording is given and it is taken through the same preprocessing and feature extraction processes and then categorized using the trained model. Lastly, the predicted category is indicated in the system which shows if the heart sound is normal of a particular abnormal condition.

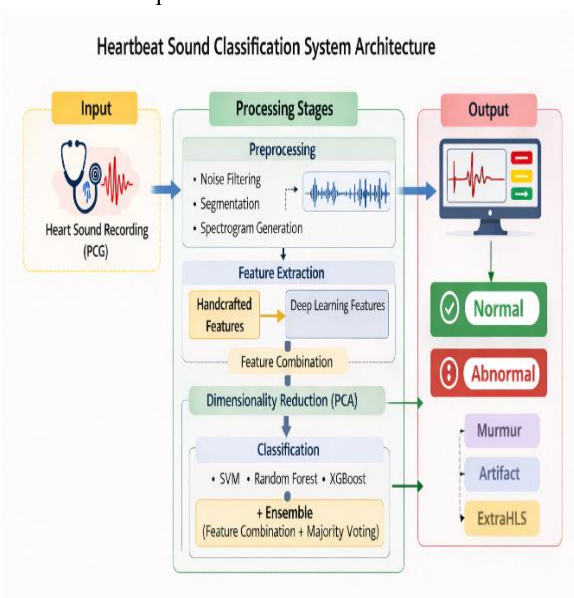


FIGURE 1: SYSTEM ARCHITECTURE

3.3 MACHINE LEARNING MODELS

The heartbeat sound classification system is proposed to be based on the machine learning framework. Instead of using only one classifier, several supervised learning algorithms were adopted and contrasted to determine the optimal phonocardiogram (PCG) signal classification model.

3.3.1 MODEL ARCHITECTURE

The system is created in a form of categorizing the sounds of the heart as normal and abnormal with the help of three currently existing machine learning algorithms, namely Support Vector Machine (SVM), Random Forest, and XGBoost. SVM is also being used to obtain the best high-dimensional hyperplanes that are very effective in separating the two classes in the feature space. Random Forest is also included as an ensemble technique of learning which integrates several decision trees to enhance robustness and stability of classification. By aggregating the predictions of several trees, it reduces overfitting and variance, resulting in more reliable performance. XGBoost is also used in predictive accuracy improvement through modeling nonlinear correlations in the heartbeat features. It uses sequential tree boosting whereby it repeatedly corrects the previous errors and improves the overall generalization. Before the model training, Principal Component Analysis (PCA) is used to minimize the dimensionality by removing the redundant and insignificant features. This procedure not only increases the efficiency of the computation but also reduces noise and non-informative parts, which eventually increases the overall system performance.

3.4 COMPILATION AND TRAINING

The Heartbeat Sound Classification model was trained on the extracted heartbeat audio dataset which included the categories normal, murmur, extra heart sounds and artifact. The features like MFCC or other features of interest in the recorded heart sound signals were then extracted after preprocessing and used as inputs to the model training. The dataset labels were coded to suit multi-class classification where 80% data was to be used in the training part and 20% in the testing part. The models that were utilized in this project are Support Vector machine (SVM), Random Forest and XGBoost. The models were trained using the extracted feature set and hyperparameter optimization was done to optimize performance.

3.5 TRAINING AND VALIDATION

The data is divided into training and testing sets. Classifiers are also trained on discriminative patterns of the features extracted during training. PCA is used only on training data in order to prevent data leakage and the same transformation is used with testing data. Evaluation metrics such as accuracy, precision, recall, F1-score and confusion matrix are used to check the validity of the model.

3.6 USER INTERFACE

The interface was made user friendly so that the interaction towards the Heartbeat Sound Classification system is made easier. The interface can enable users to post the heart sound recording to be analyzed easily. Once the audio file has been uploaded, the system automatically preprocesses the audio file, extracts features, dimensionality reduction is implemented using PCA, and a classification of the input is done against the trained machine learning models. The possible outcome predicted is shown in a very straightforward manner as either one of the following categories: Normal, Murmur, Extra Heart Sounds or Artifact. The presentation of the output is in a simple and comprehensible format, which is aimed at making the results easy to interpret by the user. The interface is fast and lightweight and is built to be integrated with clinical or telemedicine settings to enable useful applications in healthcare.

IV. IMPLEMENTATION

4.1 TOOLS AND TECHNOLOGIES

The Python programming language was used to develop Heartbeat Sound Classification system because of its good support of the machine learning and signal processing. Preprocessing of heart sound recording files was performed with the help of Librosa and NumPy with noise processing, normalization, and extraction of the features (MFCC and other timefrequency-related features). Pandas and NumPy were used to create effective data manipulation and numerical operations. Multi-class classification of heart sounds into categories of Normal, Murmur, Extra Heart Sounds and Artifact utilizing the Scikit-learn library and the XGBoost machine learning were used to classify these heart sounds. PCA has been utilized in dimensionality reduction in order to enhance efficiency. Performance evaluation was done based on accuracy, precision, recall and F1-score and plotted in Matplotlib and seaborn. An intuitive user interface was created in order to enable a straightforward uploading of heart sound recordings and present classification results.

4.2 CODE OVERVIEW

4.2.1 Data Loading and Preprocessing

Heartbeat recordings are loaded and cleaned using noise filtering and normalization. he signals are divided into heartbeat cycles and features, including MFCC, spectral and time-domain features are obtained. PCA is used to minimize the size and enhance efficiency.

4.2.2 Model Training

SVM, Random Forest and XGBoost models are trained using the processed features. Hyperparameter optimization is carried out in order to enhance accuracy, precision, recall, and F1-score. The trained models are stored in the future.

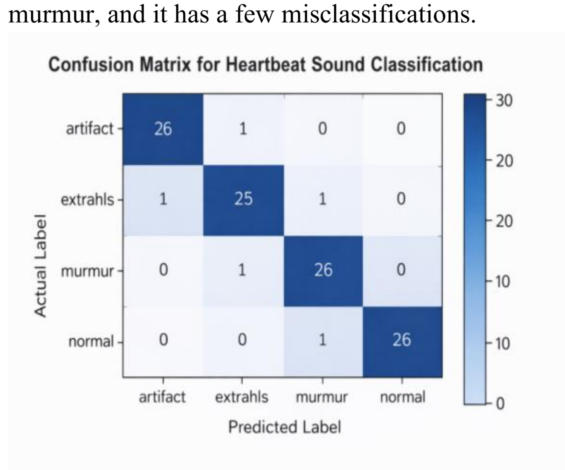
4.2.3 Prediction

The new heart sound records are processed by the same preprocessing and PCA transformation after which they are sent to the trained models. The system also gives the forecasted heartbeat category and the confidence data. This will provide practical and reliable results, which can be interpreted to provide useful healthcare.

V. RESULTS AND DISCUSSIONS

5.1 MODEL PERFORMANCE

The findings of the experiments prove that the hybrid feature-based approach performed better than the sets of features. Among the classifiers, XGBoost achieved good results because of its capability to capture nonlinear feature interactions. However, the ensemble voting model was the most stable and classified the highest. The confusion matrix shows that it is very precise and recalls well in classes of normality and murmur, and it has a few misclassifications.



Confusion Matrix for Heartbeat Sound Classification

FIGURE 2: Confusion Matrix

The confusion matrix was calculated based on the predictions that were made on the testing data. It gives a step-by-step class-wise comparison of the developed heartbeat classification system to four categories: Normal, Murmur, Extra Heart Sounds, and Artifact.

The correct classification of samples are shown by the diagonal values of the matrix, and the misclassification by the off diagonal values.

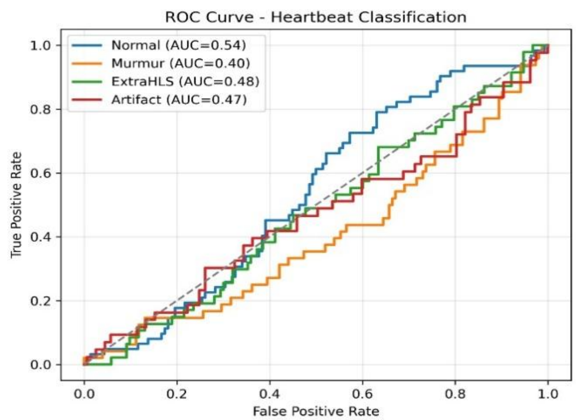


FIGURE 3: ROC Curve of heartbeat sounds

Table 1: Classification Report

Accuarcy: 95.14%

	Preci -sion	Recall	F1-Score	Support
Artifact	0.96	1.00	0.98	27
Extrahls	0.89	0.93	0.91	27
Murmur	1.00	0.89	0.94	28
Normal	0.96	1.00	0.98	27
Accuracy	-	-	0.95	109
Macro avg	0.96	0.95	0.95	109
Weighted avg	0.96	0.95	0.95	109

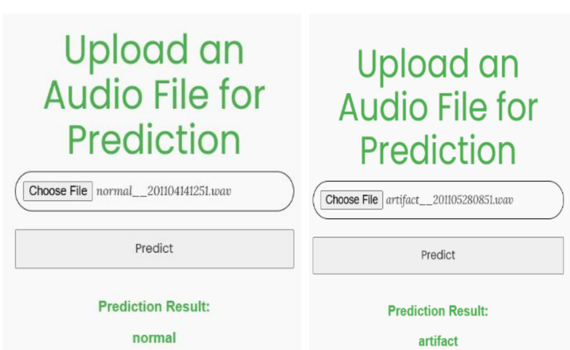


FIGURE 4: Sample Predictions

5.2 SYSTEM USABILITY

The developed heartbeat sound classification system is designed to be both efficient and practical for real-world healthcare applications. Its lightweight architecture ensures that predictions can be generated rapidly, enabling near real-time analysis of phonocardiogram (PCG) recordings. This rapid response makes the system highly suitable for remote cardiac monitoring, telemedicine platforms, and point-of-care diagnostic settings, where immediate feedback is critical for timely clinical decision making. The interface allows healthcare professionals or users to upload heart sound recordings with minimal technical knowledge. The output clearly indicates the classification results like Normal, Murmur, Extra Heart Sounds, or Artifact along with relevant diagnostic insights. Additionally, the robustness of the system ensures consistent performance even with variable recording conditions, such as background noise, signal amplitude differences, and patient variability.

5.3 COMPARISON WITH TRADITIONAL METHODS

The conventional method of auscultation relies on clinicians manually interpreting heart sounds with a stethoscope, which heavily depends on their expertise. Early machine learning systems that used only handcrafted audio features offered improvements, but they often struggled to generalize across different devices, patient groups, or recording conditions. These models also had difficulty capturing the complex relationships between time and frequency-based characteristics of PCG signals. In contrast, our hybrid approach combines handcrafted features with deep learning-based feature representations, resulting in more accurate, robust, and flexible classification. By using ensemble techniques like majority voting, the system compensates for the weaknesses of individual classifiers and improves overall accuracy across all heart sound types. The framework is also scalable, making it easy to adapt to new datasets or integrate with advanced diagnostic tools, which makes it for more practical for real-world healthcare applications than traditional methods.

5.4 FUTURE WORK

This system can also be expanded in the future to provide real-time cardiac monitoring and early diagnosis, and hence it will be even more useful in clinical and remote healthcare. An additional improvement could be the integration of the system with digital or wireless stethoscopes, which will provide the possibility to monitor the sounds of the heart live and classify them immediately in hospitals or at home-care settings. It is also possible to expand the system to identify various forms of cardiac abnormalities, including stenosis, regurgitation or arrhythmias, which contributes to making it more clinically useful. The system could be deployed on mobile apps or cloud-based solutions so that doctors and patients would be able to access it remotely, and the emergency alert feature would allow users to be informed immediately, in case of an abnormal heart sound, to provide medical assistance in time. Wearable devices or mobile apps can aid in continuously monitoring the heart condition and decreasing the number of visits to the hospital.

VI. CONCLUSION

The developed heartbeat sound classification system provides an effective, reliable, and practical approach to analyzing heart sounds for clinical and remote healthcare applications. By combining handcrafted audio features with deep learning–based representations and employing ensemble methods, the system achieves high accuracy across different heart sound categories, including Normal, Murmur, Extra Heart Sounds, and Artifacts. Its lightweight and user-friendly design allows healthcare professionals or even non-expert users to upload recordings and receive near real-time results, while maintaining robustness against variations in recording conditions, patient differences, and background noise. This project demonstrates the potential of machine learning and hybrid feature approaches to reduce diagnostic errors, support timely clinical decisions, and enhance accessibility to cardiac care. With future enhancements, such as real-time monitoring, mobile deployment, and expanded abnormality detection, this system can become an even more powerful tool for early detection and continuous monitoring of heart conditions, contributing significantly to improved patient outcomes.

REFERENCES

[1] G. A. Roth, C. Johnson, A. Abajobir, et al., “Global, regional, and national burden of cardiovascular diseases for 10 causes, 1990 to 2015,” Journal of the American College of Cardiology, vol. 70(1), pp. 1–25, 2017.

[2] S. Mangione and L. Z. Nieman, “Cardiac auscultatory skills of internal medicine and family practice trainees: a comparison of diagnostic proficiency,” JAMA, vol. 278(9), pp. 717–722, 1997.

[3] E. Etchells, C. Bell, and K. Robb, “Does this patient have an abnormal systolic murmur?” JAMA, vol. 277, pp. 564–571, 1997.

[4] U. Alam, O. Asghar, S. Q. Khan, S. Hayat, and R. A. Mali, “Cardiac auscultation: an essential clinical skill in decline,” The British Journal of Cardiology, vol. 17, pp. 8–10, 2010.

[5] P. Bentley, S. Mannor, Sounds G. Challenge Nordehn, M. Coimbra, and “The PASCAL Classifying Heart 2011 (CHSC2011) Results,” [http://www.peterjbentley.com/heartchallenge/index.html] (http://www.peterjbentley.com/heartchallenge/index.html)

[6] J. Deng, W. Dong, R. Socher, K. Li L.-J. Li, and L. Fei-Fei, “Imagenet: Alarge-scale hierarchical image database,” in Conference on Computer Vision and Pattern Recognition. IEEE, 2009, p. 248–255.

[7] S. Hershey et al., “CNN architectures for large-scale audio classification,” in International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2017, pp. 131–135.

[8] J. F. Gemmeke et al., “Audio set: An ontology and human labeled dataset for audio events,” in International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2017, pp. 776–780.

[9] W. Zhang, J. Han, and S. Deng, “Heart sound classification based on scaled spectrogram and partial least squares regression,” Biomedical Signal Processing and Control, vol. 32, pp. 20–28, 2017.

[10] B. McFee, C. Raffel, D. Liang, et al., “librosa: Audio and music signal analysis in python,” in 14th python in science conference, 2015, pp. 18–25.

[11] M. Abadi, A. Agarwal, P. Barham, et al., “Tensorflow: Large scale machine learning on heterogeneous distributed systems,” [http://download.tensorflow.org/paper/whitepaper2015.pdf] (http://download.tensorflow.org/paper/whitepaper2015.pdf), 2015.

[12] L. Buitinck, G. Louppe, M. Blondel, et al., "API design for machine learning software: experiences from the scikit-learn project," in European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases, 2013, p. 108–122.
[13] E. F. Gomes, P. J. Bentley, M. Coimbra, E. Pereira, and Y. Deng, "Classifying heart sounds: Approaches to the pascal challenge," in International Conference on Health Informatics, 2013, p. 337–340.
[14] E. F. Gomes and E. Pereira, "Classifying heart sounds using peak location for segmentation and feature construction," in Workshop Classifying Heart Sounds, 2012, p. 480–492.
[15] Y. Deng and P. J. Bentley, "A robust heart sound segmentation and classification algorithm using wavelet decomposition and spectrogram," in Workshop Classifying Heart Sounds, 2012, p. 1–6.