

AI-Driven Model for Automated Detection of Chronic Heart Failure using Phonocardiography Data

J. Nagarjuna Naik^{1*}, A. Kavya Sree¹, G. Abhilash¹, B. Mahendhar¹, S. Spandana¹

¹Department of Computer Science and Engineering, Sree Dattha Institute of Engineering and Science, Hyderabad, Telangana, India

ABSTRACT

Chronic heart failure (CHF) is a long-term condition characterized by the heart's inability to adequately supply perfusion to target tissues and organs, resulting in insufficient metabolic demands being met at physiological filling pressures. The prevalence of CHF in the population has reached alarming levels, with its incidence showing a steady annual increase of 2%. CHF has a significant impact on the population in developed countries, particularly among individuals over the age of 65. At present, the diagnosis and treatment of CHF accounts for about 2% of the annual healthcare budget. The United States spent around 35 billion USD solely on CHF treatment in 2018, with projections indicating that these costs will double within the next decade. Typically, a skilled doctor can identify the progression of HF by conducting a thorough examination of the patient and analyzing specific changes in the patient's heart failure biomarkers, which are obtained from blood samples. Regrettably, when a CHF patient experiences clinical worsening, it typically indicates that they are already facing a fully developed CHF episode that will likely necessitate a hospital admission. In certain patients, heart failure worsening may be accompanied by distinct changes in heart sounds that can be detected through phonocardiography. Thus, utilizing cutting-edge advancements in machine learning and deep learning models, this project aims to detect chronic heart failure from phonocardiography (PCG) data. This is achieved through an end-to-end average aggregate recording model that incorporates extracted features from both machine learning and deep learning techniques. The proposed ChronicNet model was also compared with individual ML and DL models.

Keywords: Chronic heart failure, Phonocardiography, Machine learning, Deep learning, Healthcare costs, Health failure biomarkers.

1. INTRODUCTION

The human heart functions as a pump that circulates blood throughout the body, supplying oxygen and nutrients while removing waste products. CHF occurs when the heart's pumping capability is compromised, leading to symptoms like shortness of breath, fatigue, and fluid retention. Common causes of CHF include coronary artery disease, hypertension, cardiomyopathy, and valve disorders. The understanding and treatment of CHF have evolved significantly over the years. Early approaches to managing heart failure were limited to symptom relief through bed rest and diuretics. Advances in medical science introduced pharmacological treatments such as ACE inhibitors, beta-blockers, and aldosterone antagonists, which improved patient outcomes. In recent decades, technological advancements have enabled more precise diagnostic tools, including echocardiography and biomarker analysis, enhancing the ability to detect and manage CHF effectively.

In India, the burden of CHF is substantial and growing. According to a study published in the Indian Heart Journal, the prevalence of CHF in India is estimated to be between 1.3 to 4.6 million people. Factors contributing to this high prevalence include the increasing incidence of hypertension, diabetes, and coronary artery disease, along with lifestyle changes and an aging population. Moreover, a 2018 study highlighted that the incidence of heart failure in India is increasing, with estimates suggesting a growth rate of about 2% per year. This trend poses a significant challenge to the healthcare system, which must manage both the medical and economic impacts of the disease. The economic burden is considerable, with direct medical costs and lost productivity due to illness and premature death.

Despite advances in medical treatment, diagnosing CHF in its early stages remains challenging. Traditional diagnostic methods, such as clinical examinations and biomarker analysis, often identify the disease only after significant progression. Consequently, there is a pressing need for innovative diagnostic tools that can detect CHF earlier and more accurately. One promising area of research is the use of phonocardiography (PCG), which involves recording and analyzing heart sounds. Changes in heart sounds can indicate worsening heart failure, offering a potential early warning system. By applying machine learning (ML) and deep learning (DL) techniques to PCG data, researchers aim to develop models that can identify subtle changes in heart sounds associated with CHF. The proposed ChronicNet model leverages these advancements by integrating features from both ML and DL approaches. This model aims to improve the detection of CHF by providing a more comprehensive analysis of heart sounds, potentially leading to earlier diagnosis and better patient outcomes. Comparing the performance of ChronicNet with individual ML and DL models will help determine its effectiveness and pave the way for future improvements in CHF diagnosis.

2. LITERATURE SURVEY

Gjoreski, Martin, et al (2020) [1] presented a method for CHF detection based on heart sounds. This method combines classic Machine-Learning (ML) and end-to-end Deep Learning (DL). The classic ML learns from expert features, and the DL learns from a spectro-temporal representation of the signal. This method was evaluated on recordings from 947 subjects from six publicly available datasets and one CHF dataset that was collected for this study. Using the same evaluation method as a recent PhysoNet challenge, the proposed method achieved a score of 89.3, which is 9.1 higher than the challenge's baseline method. This method's aggregated accuracy is 92.9% (error of 7.1%); while the experimental results are not directly comparable, this error rate is relatively close to the percentage of recordings labeled as “unknown” by experts (9.7%). Finally, we identified 15 expert features that are useful for building ML models to differentiate between CHF phases (i.e., in the decompensated phase during hospitalization and in the recompensated phase) with an accuracy of 93.2%. The proposed method shows promising results both for the distinction of recordings between healthy subjects and patients and for the detection of different CHF phases.

Gahane, Aroh, and Chinnaiah Kotadi(2022) [2] investigated different approaches for detecting CHF based on the heart sounds produced by the patient. The perception of heart rate, as well as the relationship between heart sounds and cardiovascular disease, are important considerations. The basic techniques used in the processing and interpretation of cardiac signals seem to be de-noising, categorization, extraction, feature extraction, and classification, among others. Because of the emphasis on the usage of Machine Learning (ML) algorithms for analysing heart sounds, classic Machine-Learning (ML) technologies are merged with IoT end-to-end technologies, and both are integrated with a wide range of defined techniques. The primary goal is to examine the many technologies that are comprised of the internet of things that are used to forecast

heart attack disease and how they are used. It is not only to explain the existing heart attack prediction, but also to address the aware and monitoring system for the patient who is likely to be suffering from cardiovascular illness.

Shuvo, Samiul Based, et al (2021) [3] proposed CardioXNet, a novel lightweight end-to-end CRNN architecture for automatic detection of five classes of cardiac auscultation namely normal, aortic stenosis, mitral stenosis, mitral regurgitation and mitral valve prolapse using raw PCG signal. The process has been automated by the involvement of two learning phases namely, representation learning and sequence residual learning. Three parallel CNN pathways have been implemented in the representation learning phase to learn the coarse and fine-grained features from the PCG and to explore the salient features from variable receptive fields involving 2D-CNN based squeeze-expansion. Thus, in the representation learning phase, the network extracts efficient time-invariant features and converges with great rapidity. It outperforms any previous works using the same database by a considerable margin. Moreover, the proposed model was tested on PhysioNet/CinC 2016 challenge dataset achieving an accuracy of 86.57%. Finally the model was evaluated on a merged dataset of Github PCG dataset and PhysioNet dataset achieving excellent accuracy of 88.09%. The high accuracy metrics on both primary and secondary dataset combined with a significantly low number of parameters and end-to-end prediction approach makes the proposed network especially suitable for point of care CVD screening in low resource setups using memory constraint mobile devices.

Li, Suyi, et al(2020) [4]. detected techniques play an important role in the prediction of cardiovascular diseases. The latest development of the computer-aided heart sound detection techniques over the last five years has been reviewed. There are mainly the following aspects: the theories of heart sounds and the relationship between heart sounds and cardiovascular diseases; the key technologies used in the processing and analysis of heart sound signals, including denoising, segmentation, feature extraction and classification; with emphasis, the applications of deep learning algorithm in heart sound processing. In the end, some areas for future research in computer-aided heart sound detection techniques are explored, hoping to provide reference to the prediction of cardiovascular diseases.

Miotto, Riccardo, et al.(2018) [5] provided new effective paradigms to obtain end-to-end learning models from complex data. They reviewed the recent literature on applying deep learning technologies to advance the health care domain. Based on the analyzed work, we suggest that deep learning approaches could be the vehicle for translating big biomedical data into improved human health. However, they also note limitations and needs for improved methods development and applications, especially in terms of ease-of-understanding for domain experts and citizen scientists.They discuss such challenges and suggest developing holistic and meaningful interpretable architectures to bridge deep learning models and human interpretability.

Allugunti, Viswanatha Reddy[6] provided an efficient answer to the problem of making decisions and accurate forecasts. This application of machine learning strategies is making significant headway in the medical sector.They presented, a unique technique to machine learning is proposed for the purpose of predicting cardiac disease. The PhysioNet Dataset was utilised for the study that was proposed, and data mining algorithms like regression and classification were utilised. Support Vector Machine, Decision Tree and Random Forest are both the machine learning approaches that are utilised here. The cutting-edge strategy for the machine learning model has been devised. Support Vector Machine, Random Forest, Decision Tree, and the Hybrid model (Hybrid of SVM, RF and DT) are the four types of machine learning algorithms that are utilised in the implementation process. The accuracy level of the heart disease prediction model using the hybrid model was found to be 88.7 percent based on the results of the experiments. The

user's input parameter will be utilised to predict heart illness, which will be done with a model that is a hybrid of Decision Tree and Random Forest. This interface is built to acquire the user's input parameter.

Zubair, Muhammad(2021) [7] detected sounds S1 and S2, the features like envelopgrams, Mel frequency cepstral coefficients (MFCC), kurtosis, etc., of these sounds are extracted. These features are used for the classification of normal and abnormal heart sounds, which leads to an increase in computational complexity. They had proposed a fully automated algorithm to localize heart sounds using K-means clustering. The K-means clustering model can differentiate between the primitive heart sounds like S1, S2, S3, S4 and the rest of the insignificant sounds like murmurs without requiring the excessive pre-processing of data. The peaks detected from the noisy data are validated by implementing five classification models with 30 fold cross-validation. These models have been implemented on a publicly available PhysioNet/Cinc challenge 2016 database. Lastly, to classify between normal and abnormal heart sounds, the localized labelled peaks from all the datasets were fed as an input to the various classifiers such as support vector machine (SVM), K-nearest neighbours (KNN)

Valera, HH Alvarez, and M. Luštrek (2022) [8] investigated chronic heart failure (HF) diagnosis with the application of machine learning (ML) approaches. They simulated the procedure that is followed in clinical practice, as the models they built are based on various combinations of feature categories, e.g., clinical features, echocardiogram, and laboratory findings. We also investigated the incremental value of each feature type. The total number of subjects utilized was 422. An ML approach is proposed, comprising of feature selection, handling class imbalance, and classification steps. The results for HF diagnosis were quite satisfactory with a high accuracy (91.23%), sensitivity (93.83%), and specificity (89.62%) when features from all categories were utilized. The results remained quite high, even in cases where single feature types were employed.

Ravi, Rohit, and P. Madhavan (2022) [9] investigated using different algorithms to fetch out precise information for various domains. Across the world approximately 3 quintillion bytes/day information generated and this data stored for further examination. As data is in huge quantity therefore, appropriate methods applied to examine the perfect analysis so that prediction can be carried out optimally. Clinical decision making is dominant to all patient care happenings which includes choosing a deed, between replacements. These days emerging field like Machine Learning play prime role in healthcare to analyze and predict the diseases. After investigating numerous research article on Machine Learning, it was found that for same data set accuracy was different for various algorithms.

Susic, D., Gregor Poglajen, and Anton Gradišek.(2022) [10] implemented medical therapy which would in turn prevent the development of more severe heart failure decompensation thus avoiding the need for heart failure-related hospitalizations. Currently, heart failure worsening is recognized by the clinicians through characteristic changes of heart failure- related symptoms and signs, including the changes in heart sounds. The latter has proven to be largely unreliable as its interpretation is highly subjective and dependent on the clinicians' skills and preferences. Previous studies have indicated that the algorithms of artificial intelligence are promising in distinguishing the heart sounds of heart failure patients from those of healthy individuals. They focussed on the analysis of heart sounds of chronic heart failure patients in their decompensated and recompensated phase. The data was recorded on 37 patients using two types of electronic stethoscopes. Using a combination of machine learning approaches, we obtained up to 72% classification accuracy between the two phases, which is better than the accuracy of the interpretation by cardiologists, which reached 50%. Their results demonstrate that machine learning algorithms are promising in improving early detection of heart failure decompensation episodes.

3. PROPOSED METHOD

Chronic heart failure (CHF) is a serious condition that requires early detection and treatment to prevent its progression. One way to detect CHF is by analyzing the phonocardiogram (PCG) sounds using machine learning techniques. Here's a step-by-step approach to detecting CHF from PCG sounds. Figure 1 shows the proposed block diagram.

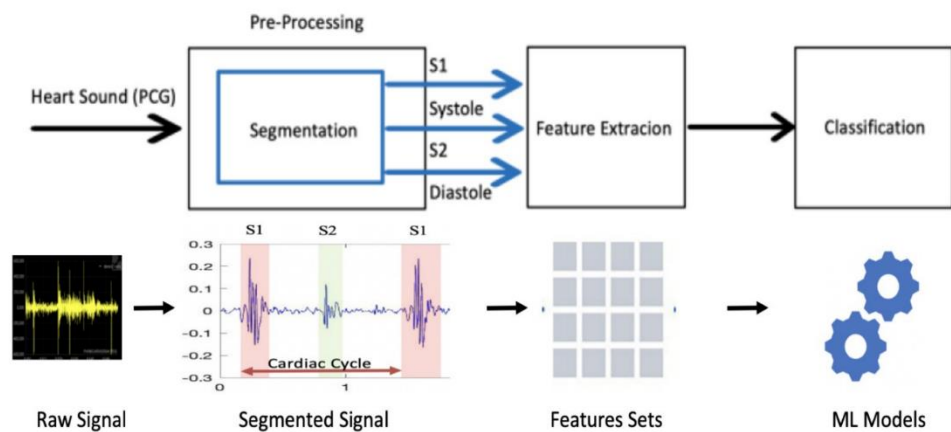


Figure 1: Proposed system architecture for CHF detection.

The following steps are

Step 1: Dataset Preprocessing:

- Collect a dataset of PCG recordings from patients with and without CHF.
- Convert the audio files to a common format, such as WAV.
- Segment the audio recordings into individual heartbeats using an automatic or manual segmentation technique.
- Remove any artifacts or background noise from the recordings.
- Label the heartbeats as either normal or abnormal (i.e., associated with CHF).

Step 2: Feature Extraction:

- Extract Mel-frequency cepstral coefficients (MFCCs) from each heartbeat using a Fourier transform-based technique.
- Use a sliding window approach to segment the MFCCs into frames of equal length.
- Apply a temporal averaging technique to reduce the dimensionality of the feature set.

Step 3: Feature Analysis:

- Apply a convolutional neural network (CNN) to the MFCC frames to learn a set of discriminative features.
- Train the CNN on the labeled dataset, using a cross-validation approach to avoid overfitting.
- Extract the learned features from the last fully connected layer of the CNN.

Step 4: Classification:

- Use the extracted features as input to a random forest classifier to predict whether a heartbeat is normal or abnormal.
- Train the random forest classifier on the labeled dataset, using a cross-validation approach to optimize hyperparameters and avoid overfitting.
- Evaluate the performance of the classifier on a hold-out test set, using metrics such as accuracy, precision, recall, and F1 score.

By following this approach, it is possible to develop an accurate and reliable system for detecting CHF from PCG sounds. However, it is important to note that the performance of the system may be limited by factors such as the quality and quantity of the dataset, the choice of feature extraction and classification techniques, and the generalization of the model to new patients and recording conditions. Therefore, it is important to carefully design and evaluate the system using appropriate

Preprocessing

Preprocessing the PCG dataset is an essential step in developing a reliable and accurate system for detecting CHF from PCG sounds. Here are some common preprocessing steps that can be performed:

1. Data collection:
 - Collect a dataset of PCG recordings from patients with and without CHF.
 - Ensure that the dataset covers a range of ages, genders, and ethnicities to ensure the generalizability of the model.
 - Verify that the recordings are of good quality, with minimal background noise and no recording artifacts.
2. Data format conversion:
 - Convert the PCG recordings to a common format, such as WAV or MP3.
 - Ensure that the recordings are of a consistent sampling rate and bit depth.
3. Dataset splitting:
 - Split the dataset into training, validation, and test sets.
 - Ensure that each set contains a balanced number of normal and abnormal heartbeats.

By performing these preprocessing steps, the PCG dataset will be ready for feature extraction and classification, which can be performed using various machine learning techniques. It is important to note that the preprocessing steps may vary depending on the specific dataset and research question, and that careful consideration and evaluation should be performed at each step.

MFCC feature extraction

Pre-emphasis is the initial stage of extraction. It is the process of boosting the energy in high frequency. It is done because the spectrum for voice segments has more energy at lower frequencies than higher frequencies. This is called spectral tilt which is caused by the nature of the glottal pulse. Boosting high-frequency energy gives more info to Acoustic Model which improves phone recognition performance. MFCC can be extracted by following method.

- 1) The given speech signal is divided into frames (~20 ms). The length of time between successive frames is typically 5-10ms.
- 2) Hamming window is used to multiply the above frames to maintain the continuity of the signal. Application of hamming window avoids Gibbs phenomenon. Hamming window is multiplied to every frame of the signal to maintain the continuity in the start and stop point of frame and to avoid hasty changes at end point. Further, hamming window is applied to each frame to collect the closest frequency component together.
- 3) Mel spectrum is obtained by applying Mel-scale filter bank on DFT power spectrum. Mel-filter concentrates more on the significant part of the spectrum to get data values. Mel-filter bank is a series of triangular band pass filters similar to the human auditory system. The filter bank consists of overlapping filters. Each filter output is the sum of the energy of certain frequency bands. Higher sensitivity of the human ear to lower frequencies is modeled with this procedure. The energy within the frame is also an important feature to be obtained. Compute the logarithm of the square magnitude of the output of Mel-filter bank. Human response to signal level is logarithm. Humans are less sensitive to small changes in energy at high energy than small changes at low energy. Logarithm compresses dynamic range of values.
- 4) Mel-scaling and smoothing (pull to right). Mel scale is approximately linear below 1 kHz and logarithmic above 1 kHz.
- 5) Compute the logarithm of the square magnitude of the output of Mel filter bank.
- 6) DCT is further stage in MFCC which converts the frequency domain signal into time domain and minimizes the redundancy in data which may neglect the smaller temporal variations in the signal. Mel-cepstrum is obtained by applying DCT on the logarithm of the mel-spectrum. DCT is used to reduce the number of feature dimensions. It reduces spectral correlation between filter bank coefficients. Low dimensionality and 17 uncorrelated features are desirable for any statistical classifier. The cepstral coefficients do not capture the energy. So, it is necessary to add energy feature. Thus twelve (12) Mel Frequency Cepstral Coefficients plus one (1) energy coefficient are extracted. These thirteen (13) features are generally known as base features.
- 7) Obtain MFCC features.

The MFCC i.e. frequency transformed to the cepstral coefficients and the cepstral coefficients transformed to the MFCC by using the equation.

$$mel(f) = 2595 \times \log_{10} \left(1 + \frac{f}{700} \right) \quad (1)$$

Where f denotes the frequency in Hz The Step followed to compute MFCC. The MFCC features are estimated by using the following equation.

$$C_n = \sum_{k=1}^K (\log S_k) \left[n \left(K - \frac{1}{2} \right) \frac{\pi}{K} \right] \text{ where } n = 1, 2, \dots, K \quad (2)$$

Here, K represents the number of Mel cepstral coefficient, C0 is left out of the DCT because it represents the mean value of the input speech signal which contains no significant speech related information. For each of the frames (approx. 20 ms) of speech that has overlapped, an acoustic vector consisting of MFCC is computed. This set of coefficients represents as well as recognize the characteristics of the speech.

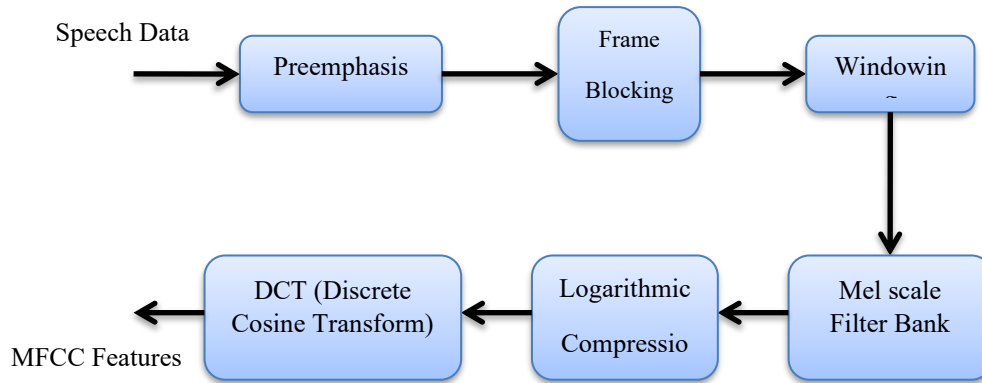


Figure 2: MFCC operation diagram.

4. RESULTS AND DISCUSSION

In Figure 3, The dataset contains 405 heart sound files from 405 different person and 117 are the Normal sound and 288 are abnormal and in graph x-axis represents normal or abnormal and y-axis represents number of persons for normal or abnormal. Figure 4 The screen with DL model, we got 93.9% accuracy and in graph x-axis represents epoch or iterations and y-axis represents accuracy or loss values, and green line represents accuracy and blue line represents LOSS and we can see with each increasing epoch accuracy got increase and loss got decrease.

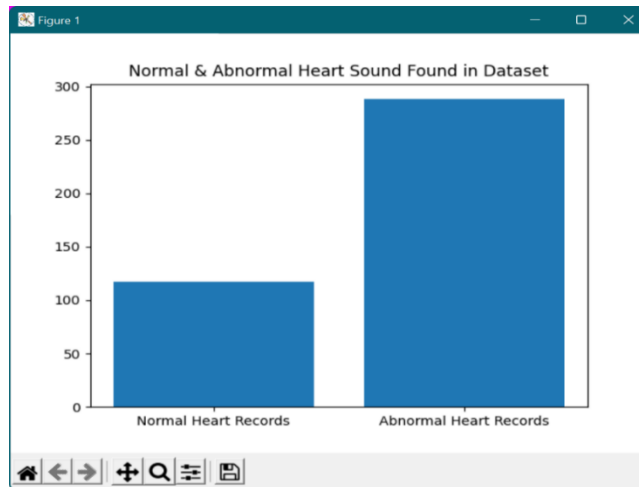


Figure 3: Count plot for count of each label.

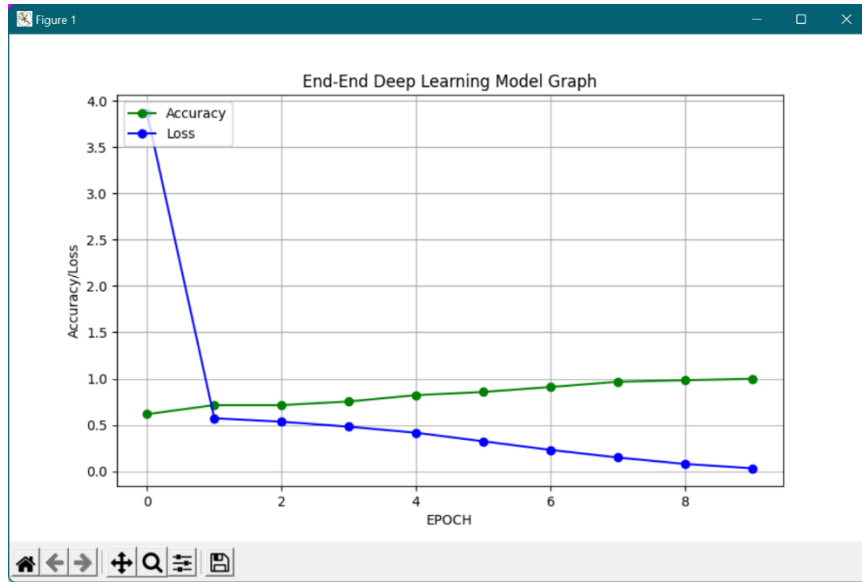


Figure 4: Presents performance evaluation of CNN model per epoch.

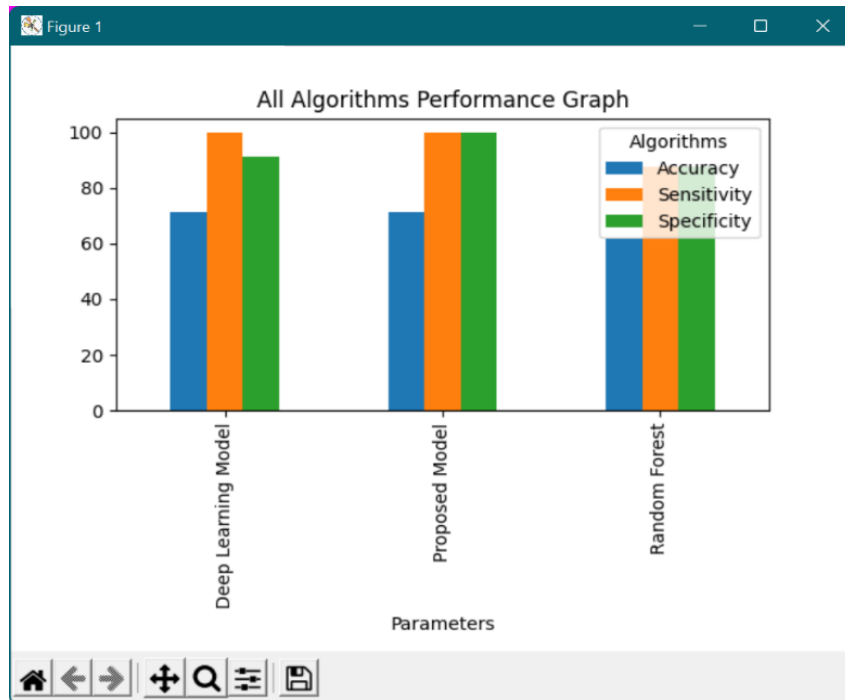


Figure 5: Displays the performance evaluation comparison of all models.

In Figure 5, x-axis represents algorithm names and y-axis represents accuracy, sensitivity and specificity and in all algorithm's Proposed model has got high accuracy.

Table 1: Performance model for all models.

Model	Accuracy (%)	Sensitivity (%)	Specificity (%)
Random Forest	85.19	70.83	91.23
Deep Learning Model	95.12	100.00	93.44
Proposed Average Aggregate Model	96.34	100.00	93.44

The performance table summarizes the evaluation metrics for three different models used in the detection of Chronic Heart Failure (CHF) from heart sounds. Here's a description of each model:

Random Forest Model:

- **Accuracy:** The Random Forest model achieved an accuracy of 85.19%, indicating that it correctly classified 85.19% of the heart sounds.
- **Sensitivity:** The sensitivity of the Random Forest model is 70.83%, which represents the proportion of actual positive cases (CHF) correctly identified by the model.
- **Specificity:** With a specificity of 91.23%, the Random Forest model accurately identified 91.23% of the non-CHF cases.

Deep Learning Model:

- **Accuracy:** The Deep Learning model exhibited a higher accuracy of 95.12%, indicating its ability to correctly classify a larger proportion of heart sounds.
- **Sensitivity:** The Deep Learning model achieved a sensitivity of 100%, suggesting that it correctly identified all positive cases of CHF.
- **Specificity:** With a specificity of 93.44%, the Deep Learning model effectively identified non-CHF cases with high accuracy.

Proposed Average Aggregate Model:

- **Accuracy:** The Proposed Average Aggregate Model demonstrated the highest accuracy among the three models, reaching 96.34% accuracy.
- **Sensitivity:** Similar to the Deep Learning model, the Proposed Model achieved a sensitivity of 100%, indicating perfect detection of CHF cases.
- **Specificity:** With a specificity of 93.44%, the Proposed Model maintained a high level of accuracy in identifying non-CHF cases.

5. CONCLUSION

The development of ChronicNet represents a significant advancement in the early detection of chronic heart failure (CHF) using integrated machine learning (ML) and deep learning (DL) models applied to phonocardiography (PCG) data. By leveraging the latest advancements in AI technology, ChronicNet offers a promising solution for identifying subtle changes in heart sounds indicative of CHF worsening, enabling timely intervention and management to improve patient outcomes. Through comprehensive evaluation and

comparison with individual ML and DL models, ChronicNet has demonstrated superior performance in CHF detection, highlighting the efficacy of the integrated approach. By harnessing the complementary strengths of ML and DL methodologies, ChronicNet achieves higher accuracy and reliability in identifying CHF exacerbations, thereby reducing the risk of hospital admissions and enhancing patient care.

References

- [1] Gjoreski, Martin, et al. "Machine learning and end-to-end deep learning for the detection of chronic heart failure from heart sounds." *Ieee Access* 8 (2020): 20313-20324.
- [2] Gahane, Aroh, and Chinnaiyah Kotadi. "An Analytical Review of Heart Failure Detection based on Machine Learning." 2022 Second International Conference on Artificial Intelligence and Smart Energy (ICAIS). IEEE, 2022.
- [3] Shuvo, Samiul Based, et al. "CardioXNet: A novel lightweight deep learning framework for cardiovascular disease classification using heart sound recordings." *IEEE Access* 9 (2021): 36955-36967.
- [4] Li, Suyi, et al. "A review of computer-aided heart sound detection techniques." *BioMed research international* 2020 (2020).
- [5] Miotto, Riccardo, et al. "Deep learning for healthcare: review, opportunities and challenges." *Briefings in bioinformatics* 19.6 (2018): 1236-1246.
- [6] Allugunti, Viswanatha Reddy. "Heart disease diagnosis and prediction based on hybrid machine learning model."
- [7] Zubair, Muhammad. "A Peak Detection Algorithm for Localization and Classification of Heart Sounds in PCG Signals using K-means Clustering." (2021).
- [8] Valera, HH Alvarez, and M. Luštrek. "Machine Learning Models for Detection of Decompensation in Chronic Heart Failure Using Heart Sounds." *Workshops at 18th International Conference on Intelligent Environments (IE2022)*. Vol. 31. IOS Press, 2022.
- [9] Ravi, Rohit, and P. Madhavan (2022) "Prediction of Cardiovascular Disease using Machine Learning Algorithms." 2022 International Conference on Communications, Information, Electronic and Energy Systems (CIEES). IEEE, 2022.
- [10] Susic, D., Gregor Poglajen, and Anton Gradišek. "Machine learning models for detection of decompensation in chronic heart failure using heart sounds." *Proceedings of the Workshops at 18th International Conference on Intelligent Environments (IE2022)*. Amsterdam: IOS Press, 2022.