DETECTION OF CONGESTIVE HEART FAILURE USING MULTI-LEAD ECG SIGNAL ANALYSIS

A PROJECT REPORT

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF THE DEGREE

OF

MASTER OF TECHNOLOGY

IN

COMPUTER SCIENCE & ENGINEERING

Submitted by

ANKUR NIGAM

24/CSE/15

Under the Supervision of

Dr. Rajesh Kumar Yadav

Associate Professor, Department of Computer Science & Engineering
Delhi Technological University

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)
Bawana Road, Delhi-110042, India

MAY, 2026

DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Main Bawana Road, Delhi-110042

ACKNOWLEDGEMENT

I sincerely thank my supervisor, Dr. R. K. Yadav,Associate Professor, Department of Computer Science & Engineering, Delhi Technological University, for guiding this project from its initial idea to its final form. His suggestions helped me refine the problem statement, understand ECG-based diagnosis more clearly, and improve the experimental presentation of this work.

I also thank the faculty members, technical staff, and academic environment of the Department of Computer Science & Engineering for providing the support required to complete this dissertation. The coursework and discussions during the M.Tech. Computer Science & Engineering programme helped me connect signal processing concepts with machine learning methods used in this study.

This work uses publicly available ECG resources, especially the BIDMC Congestive Heart Failure Database and the MIT-BIH Normal Sinus Rhythm Database. I acknowledge the researchers and institutions who made these datasets available, as such resources make reproducible biomedical research possible.

I am also thankful to my classmates, friends, and family members for their encouragement during the project. Their patience and support helped me remain consistent throughout the preparation, implementation, and writing stages of this dissertation.

Place: Delhi ANKUR NIGAM

Date: 24/CSE/15

DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Main Bawana Road, Delhi-110042

CANDIDATE'S DECLARATION

I, **Ankur Nigam**, Roll No. **24/CSE/15**, a student of M.Tech. Computer Science & Engineering, declare that the dissertation titled **"Detection of congestive heart failure using multi-lead ECG signal analysis"** has been prepared by me as part of the requirements of Delhi Technological University. The work presented in this report is based on my own study, implementation, analysis, and writing, and all external material used for reference has been duly acknowledged through citations.

I further declare that this dissertation has not been submitted earlier, either fully or partially, for the award of any degree, diploma, fellowship, associateship, or equivalent academic recognition at this or any other institution.

Place: Delhi **Ankur Nigam**

Date: **24/CSE/15**

This is also certified that the revisions and suggestions communicated during evaluation have been addressed by the candidate, and the declaration made above is accepted to the best of our knowledge.

Signature of Supervisor **Signature of External Examiner**

DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Main Bawana Road, Delhi-110042

CERTIFICATE

This certificate confirms that ANKUR NIGAM, Roll No. 24/CSE/15, Department of Computer Science & Engineering, Delhi Technological University, has completed the project dissertation titled **"Detection of congestive heart failure using multi-lead ECG signal analysis"** under my supervision. The dissertation is submitted toward the requirements for the Master of Technology degree in Computer Science & Engineering.

Based on the work reviewed by me, the dissertation presents the student's project study, implementation, and analysis. To the best of my knowledge, the same work has not been submitted elsewhere, in full or in part, for any degree or diploma.

Place: Delhi

Date:

Dr. R. K. Yadav
Associate Professor

ABSTRACT

Congestive heart failure (CHF) is a condition in which the heart is unable to pump blood as well as it should to meet the needs of normal circulation. Over time this condition can decrease physical capacity, increase hospital visits and increase the risk of serious cardiac events. The early diagnosis of CHF is important for patient care and follow-up monitoring because of these consequences. ECG recording is a practical tool for this purpose, as it is inexpensive, non-invasive and suitable for repeated measurement. Multi-lead ECG data are especially useful, because different leads observe the electrical behaviour of the heart from different directions. Most conventional methods are hand-crafted measures and may not well represent morphological changes of ECG with time. Furthermore, evaluation on random segment splitting may provide an optimistic estimate, if the signals from one patient are present in the training and testing set. This dissertation investigates a multi-lead diagnostic attention-based recurrent neural network (MLDA-RNN) for classification of CHF. We use the BIDMC Congestive Heart Failure Database and the MIT-BIH Normal Sinus Rhythm Database. The data is separated patient wise, so that testing is done on unseen subjects. Integer Haar Wavelet Transform is included to get compact signal information into low computational cost. The MLDA-RNN model achieved an accuracy of 96.06 % on the blind test data. The result shows that the selected combination of patient-wise evaluation, multi-lead ECG representation, recurrent sequence learning, attention-based weighting and efficient wavelet processing can provide a solid framework for automated CHF screening. The approach can be further extended for decision-support tools, remote monitoring workflows and wearable ECG systems after broader clinical validation.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Background

1.1.1 Congestive Heart Failure

Congestive heart failure (CHF) is a long-term cardiac disorder in which the heart cannot supply blood at a level required by the body. Global estimates for 2017 report approximately 64.34 million CHF cases [3], showing that the disease is a major public-health burden. Diagnosis can be complicated because CHF may appear with reduced, preserved, or mid-range ejection fraction. In a healthy condition, the left ventricle ejects enough oxygenated blood during each beat; in CHF, this pumping ability declines and circulation becomes insufficient. Early detection is therefore important. Clinical assessment may involve chest X-ray, MRI, angiography, echocardiography, nuclear imaging, and other tests. ECG-based analysis is attractive because it is inexpensive, non-invasive, widely available, and suitable for repeated monitoring. For this reason, the present work focuses on ECG signals for CHF detection.

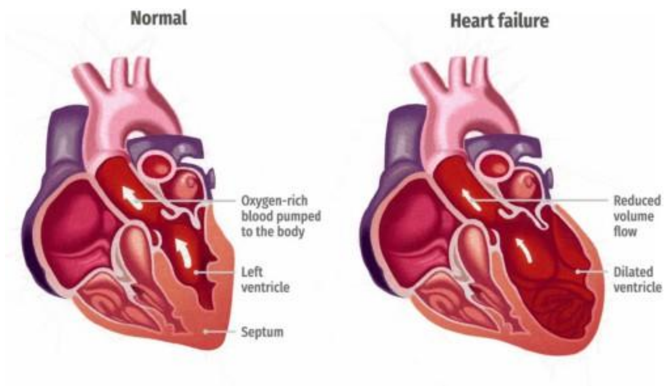

Figure 1.1. Normal heart vs Heart failure

1.1.2 Electrocardiogram Signals (ECG Signals)

An electrocardiogram records the electrical activity generated during the cardiac cycle. It is painless, quick to acquire, and widely used for observing heart rhythm and identifying abnormal cardiac behaviour. The human heart contains two atria and two ventricles, and their depolarization and repolarization create the visible ECG waveform. As illustrated in Fig. 1.2, the important waveform components include the P wave, QRS complex, and T wave. The P wave is related to atrial depolarization, the QRS complex corresponds to ventricular depolarization, and the T wave represents ventricular repolarization. Since ventricular muscle mass is larger than atrial muscle mass, the QRS complex usually has a larger amplitude. Variations in waveform shape, timing, and intervals can provide clinically useful information.

ECG machines are standard equipment. Some personal gadgets, such as smartwatches, also allow for ECG monitoring. ECG can be used in case of Chest pain, Dizziness, Rapid pulse, shortness of breath, Weakness, fatigue, or a decline in the ability to exercise. ECG offers vital information for the prognostic assessment, diagnosis, and therapy of individuals with congestive heart failure.

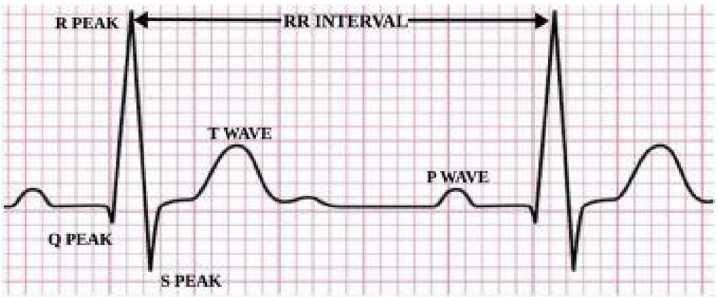

Figure 1.2. ECG Signal

1.2 Clinical and Technical Background

The diagnosis of congestive heart failure is not based on a single observation; rather, it is generally supported by a combination of symptoms, physical examination, imaging findings, laboratory investigations, and electrical activity recorded through ECG. Common clinical symptoms include shortness of breath, fatigue, swelling of the lower limbs, and reduced exercise tolerance. However, these symptoms can overlap with other respiratory or metabolic conditions. For this reason, automated ECG-based analysis can act as a useful complementary screening tool, especially when continuous or repeated monitoring is required.

CHF can have several effects on the ECG from a signal processing standpoint. The disease may change the timing, shape and variability of the cardiac cycles. Features such as QRS duration, RR interval variability, ST-T segment behaviour and beat-to-beat rhythm regularity can reflect changes in ventricular activation, autonomic regulation and myocardial conduction. A cardiologist can visually inspect the ECG traces, but subtle temporal changes over long recordings are difficult to assess manually. Therefore automated systems are useful as they can consistently process many ECG segments and extract hidden patterns that may not be obvious in a short visual inspection. Multi-lead ECG recordings are of special importance, since each lead looks at the heart from a different electrical point of view. One lead may be able to record rhythm information but not be able to pick up spatial abnormalities that can be seen in another lead. Combining multiple leads enables the model to learn the complementary representations of cardiac activity. In this project, the study of multi-lead ECG allows the model to learn both the temporal dependence of each lead and the diagnostic dependence between leads. This is one of the reasons why recurrent neural networks with attention mechanisms are suitable for the task. Calculation efficiency is of practical importance for the CHF detection system. ECG monitoring is gradually shifting to portable devices, bedside monitors and wearable healthcare systems. These platforms can have limited memory, battery capacity, and processing power. Thus, feature extraction techniques, such as the Integer Haar Wavelet Transform, are useful, as they can provide useful time-frequency information without the heavy computational cost of more complex transforms. This project exploits this idea to support a model design that is not just accurate but also relevant for low resource deployment.

1.2.1 Need for Automated CHF Detection

The increasing burden of cardiovascular disease and the need for early warning systems require automated CHF detection. Patients usually do not go to hospital until the symptoms are serious. ECG based screening system can help in detecting abnormal patterns earlier and can help in further medical examination. Such a system does not replace the physician, but it allows to prioritize cases, to reduce the manual workload and to offer objective measurements at the time of repeated monitoring

The main advantages of ECG-based automated CHF detection are, as follows: • Non-invasive measurement: ECG recording does not require surgical intervention and is accepted in clinical practice. • Cost-Effective: ECG devices are generally cheaper than advanced imaging techniques like MRI or nuclear imaging. • Continuous monitoring: ECG signals can be recorded for extended periods of time, making them suitable for remote and wearable monitoring. • Objective analysis: Machine learning models can process large amounts of ECG data reliably. • Early detection: Automated systems can detect suspicious patterns and assist clinicians in making timely decisions. With these advantages, building a robust automated system is a challenge. ECG signals are noisy, patient dependent and affected by electrode placement, movement, baseline drift and physiological variations. Further, a model trained on randomly partitioned ECG segments may perform well on testing but not on unseen patients. This project overcomes this limitation by using patient-wise separation, which is more stringent and closer to real-world deployment.

1.2.2 Challenges in ECG Based Classification

There are some challenges to be addressed but ECG signals can be effectively used for CHF classification. First, ECG signals are often contaminated by artifacts due to muscle activity, power-line interference, and electrode motion. If not controlled, the model may learn irrelevant patterns due to such noise. Second, the ECG recordings of different patients may have different amplitude, heart rate and morphological features. If the model is not trained carefully these variations can degrade generalization. Third, CHF-associated changes might be subtle and occur over time rather than a single, obvious abnormality. Yet another big problem is data leakage. If there are beats or segments from the same patient in both train and test sets, the model may learn patient-specific patterns instead of disease-specific patterns.This can lead to overly optimistic results. Therefore, patient-wise data splitting is important for evaluating how the model behaves on truly unseen subjects. In the present work, blind test data are created to reduce this risk and to provide a more meaningful estimate of performance.

Table 1.1. Major challenges in ECG-based CHF detection

Challenge	Cause	Effect on model
Noise and artifacts	Electrode movement, muscle activity, power-line interference	Distorts ECG morphology and may reduce classification accuracy
Patient variability	Differences in age, heart rate, disease severity, and physiology	Makes generalization to unseen patients difficult
Lead variation	Different leads capture different electrical views of the heart	Important diagnostic information may be missed by single-lead models
Temporal dependency	CHF patterns may appear across multiple beats	Static features may fail to capture long-term signal behaviour
Data leakage	Random segment-level splitting of ECG data	Inflates testing performance and reduces real-world reliability

1.3 Research Motivation

Congestive heart failure is a very serious problem and its diagnosis is very important as soon as possible because doing so may help to avoid long-term consequences and unexpected heart mortality. So here we are using ECG signals, and incorporating different models to get the best accuracy.

The novelty of our work is to come up with a multi-lead diagnostic attention-based recurrent neural network for the classification of CHF. The method analyses the two-lead ECGs to extract the multi-scale temporal dependencies from each lead for better classification.

1.4 Organization Of The Report

The remaining chapters are arranged to present the study in a logical sequence. Chapter 2 reviews earlier ECG-based CHF detection studies and introduces the datasets used in this dissertation. Chapter 3 explains the signal-processing method, the feature-generation procedure, and the classification models. Chapter 4 presents the experiments, compares the obtained results, and discusses the behaviour of the models. Chapter 5 summarizes the major findings and outlines possible extensions of the work.

CHAPTER 2

REVIEW OF PRIOR WORKS

2.1 Literature Review

Recent studies on CHF detection from ECG abandoned the traditional hand-crafted heart-rate-variability analysis in favor of modern deep learning models that directly operate on ECG sequences. CHF is a major cardiovascular disease because of its large population of patients and the risk of rehospitalization and sudden deterioration if not identified early. Many researchers have explored the potential of using ECG-derived patterns to discriminate CHF subjects from normal subjects and other cardiac conditions due to its low cost, non-invasive nature, and availability for ECG recording. Early studies have been mainly based on manually designed descriptors. Melillo et al. [4] reported high sensitivity but relatively lower specificity with long-term HRV measurements with a classification and regression tree . Orhan [5] classified CHF and normal sinus rhythm patterns based on feature extraction using amplitude and time partitioning. Masetic and Subasi [6] used the autoregressive Burg features with a decision tree classifier and obtained high classification accuracy. These approaches prove that ECG and HRV data can provide useful diagnostic information, but the performance is strongly depending on the selected features. Machine learning methods such as KNN, LDA, MLP, SVM and RBF-based neural networks have also been explored for CHF identification. Isler and Kuntalp [7] focused on short-term and long-term HRV information, and achieved good results using classical classifiers. Fuadah et al. [8] employed Hjorth descriptors and entropy-based features with different classifiers on the MIT-BIH and BIDMC records. Their results show that well-designed features can support accurate CHF detection, but feature engineering may require domain knowledge and may not generalize equally well across datasets.

With the growth of deep learning, researchers began to reduce dependence on manually selected features. Acharya et al. [4] applied a convolutional neural network directly to raw ECG data, showing that deep models can learn discriminative representations automatically. Darmawahyuni et al. [5] used an LSTM-based model for sequential ECG analysis, while Prabhakararao and Dandapat [6] used a multilayer recurrent neural network that accepts ECG signals as input. These studies demonstrate the suitability of recurrent and convolutional architectures for ECG classification tasks.

Wavelet-based methods remain relevant because ECG signals are non-stationary. Nahak et al. [7] compared HRV features and wavelet-based features for a three-class classification problem involving arrhythmia, CHF, and normal sinus rhythm. They observed that wavelet information contributed strongly to classification performance. Zou et al. [8] proposed a multiscale residual UNet++ approach, highlighting the importance of multiresolution analysis and hardware-oriented design considerations. Similarly, the MLDA-RNN method proposed by Prabhakararao and Dandapat [1] uses multi-lead ECG information with an attention-based recurrent structure, which is closely related to the direction followed in this project.

Recent studies have extended ECG classification using transformer-based and hybrid architectures. Liu et al. [9] proposed ECVT-Net, which combines convolutional processing with a vision transformer for extracting high-level ECG features. Prusty et al. [10] converted ECG information into a form suitable for a 2D CNN using scale-invariant feature transform ideas. Mahmoud et al. [11] used a CNN-LSTM framework with frequency-domain preprocessing, and Odinaka et al. [12] investigated transfer learning using pretrained image models. More recently, Lee et al. [13] demonstrated the potential of deep learning on standard 12-lead ECG recordings for heart-failure-related diagnosis in clinical settings.

Overall, the literature suggests three important observations. First, ECG signals contain measurable information related to CHF. Second, deep learning models can reduce manual feature-design effort and can capture nonlinear temporal relationships. Third, evaluation strategy is critical: random segment-level splitting may overestimate performance if patient identity information leaks into the testing set. Motivated by these points, the present work uses patient-wise data separation and evaluates an MLDA-RNN-based approach using multi-lead ECG signals.

2.2 Datasets Used

2.2.1 MIT-BIH Normal Sinus Rhythm Database

This database contains long-term ECG recordings of 18 patients which includes 5 males and 13 women from Boston's Beth Israel Hospital's Arrhythmia Laboratory. Their patients were found to have no symptoms of arrhythmias. Each recording contains 2 ECG leads and is sampled at 128 samples per second.

Table 2.1. Previous Works of Congestive Heart Failure Detection

Work	Year	Features	Extraction Method / Technique used	Database Used	Accuracy	Model Used
[6]	2022	ECG signal	—	MIT-BIH, BIDMC, PTBDB	98.57%	multilayered RNN
[5]	2020	—	—	MIT-BIH, BIDMC	99.86%	LSTM
[14]	2022	Hjorth Descriptor and Entropy	—	MIT-BIH, BIDMC	RBFN: 97%, remaining 100%	RBFN, SVM, kNN, RF, and ANN
[4]	2019	—	—	NSRDB, Fantasia, BIDMC	98.97%	CNN
[1]	2020	Direct ECG lead	—	STAFF III and PTB diagnostic	97.79%	MLDA-RNN
[8]	2022	HRV features	—	NSRRR and CHF-RR interval database	89.83%	Multiscale Residual UNet++
[7]	2020	HRV features, wavelet based features	—	MIT-BIH Arrhythmia, BIDMC and MIT-BIH NSRD	fusion of HRV and wavelet based features produced 93.33% accuracy	SVM
[9]	2022	ECG signal	CNN + Vision Transformer	MIT-BIH, BIDMC	98.88%	ECVT-Net
[12]	2023	ECG signal (time-frequency)	Transfer learning	MIT-BIH, BIDMC	99.2%	ResNet-50, AlexNet
[10]	2024	SIFT image features	Scale-Invariant Feature Transform	PhysioNet (ARR, CHF, NSR)	99.78%	2D-Deep CNN
[11]	2024	FFT + raw ECG	FFT preprocessing	MIT-BIH, BIDMC	—	CNN-LSTM
[13]	2025	12-lead ECG waveforms	Deep learning (end-to-end)	Samsung Medical Center cohort	—	Deep learning (AI-ECG)

2.2.2 BIDMC Congestive Heart Failure Database

This database contains long-term ECG recordings from 15 patients with severe congestive heart failure, including 11 men and 4 women, ages 22 to 71 (NYHA class 3–4). Each recording consists of 20 hours in duration and contains 2 ECG leads, sampled at 250 samples per second.

2.3 Dataset Preparation and Study Design

The reliability of any machine learning model depends strongly on how the dataset is prepared. In biomedical signal classification, dataset preparation is especially important because the model may otherwise learn artifacts, recording conditions, or patient identities instead of the disease pattern. For this reason, the proposed work follows a structured preparation strategy consisting of signal selection, patient-wise grouping, preprocessing, segmentation, feature extraction, and model evaluation.

The MIT-BIH Normal Sinus Rhythm Database contains normal subjects and the BIDMC Congestive Heart Failure Database contains CHF subjects. These two datasets provide ECG recordings from different physiological groups and are therefore suitable for binary classification of CHF versus normal sinus rhythm. The classification task is to determine whether a given ECG segment is of the CHF class or of the normal class. Since ECG signals are time series, the model must learn not only isolated values but also temporal patterns that change over multiple samples and beat

2.3.1 Patient-wise Data Separation

We made a conscious decision to separate patients in this design. Many ECG classification studies divide long recordings into several short segments and then separate them into training and testing sets randomly. This generates many samples, but also means that segments from the same patient could end up in both sets. In such a case, the test set is not truly independent, because the model has already seen the signal characteristics of the same patient during training. This can lead to the reported accuracy being inflated relative to the true real-world performance.

To avoid this issue, the present work separates the data at the patient level. All ECG segments belonging to a particular patient are placed either in the training set or in the testing set, but not both. This produces blind test data and gives a better estimate of how the model will behave when it receives ECG data from a new patient. The method is more challenging because the model must learn general CHF-related characteristics instead of memorizing individual patient patterns.

Table 2.2. Difference between random split and patient-wise split

Criterion	Random segment split	Patient-wise split
Data division	ECG segments are randomly divided	Complete patient recordings are separated
Risk of leakage	High, because same patient may appear in train and test sets	Low, because test patients are unseen
Reported accuracy	May be overly optimistic	More realistic and difficult
Clinical relevance	Limited for unseen-patient deployment	Better suited for real screening systems
Model requirement	Can exploit patient-specific patterns	Must learn disease-specific patterns

2.3.2 Signal Preprocessing

Raw ECG signals are not directly useful for classification as they may contain noise and baseline fluctuations. Preprocessing enhances the signal quality and directs the model toward diagnostically relevant information. The main preprocessing steps considered in this work are baseline correction, normalization, segmentation and quality checking. Baseline correction can reduce slow drift caused by respiration or electrode movement. Normalization reduces the amplitude differences across recordings, so the model is less biased towards signal magnitude. Segmentation divides long ECG recordings into smaller windows to be processed by the neural network. The quality checking is also important as poor quality ECG segments may contain artifacts which are not representative of cardiac activity . If the model is trained on these segments, it can learn spurious patterns. In a practical system, noisy seg-ments should be either removed or separately dealt with by a signal quality assessment module. The quality-aware block in the proposed framework is designed to support this requirement by ensuring that usable ECG segments are propagated for classification

2.3.3 Segmentation Strategy

The neural networks usually take fixed-size input windows, while ECG recordings are continuous signals. Hence, segmentation is performed to convert each recording into smaller windows. The length of the segment must be chosen carefully. Too short segment may not contain enough cardiac cycles to represent temporal behaviour. If it is too long, the computational cost increases and the number of training samples decreases. A balanced segment length is one where the model can see multiple beats and keep computation manageable. Segment level classification is useful for detection of CHF, because it allows to analyse long patient recordings in smaller parts. Finally, a patient-level decision can be made by aggregating segment predictions by majority voting or averaging of probabilities. Although the focus of this project is on segment-level classification, the framework can be generalized to patient-level screening by aggregating predictions over time.

2.3.4 Feature Representation

Feature representation is a crucial part of the suggested methodology. ECG signals contain information in time and frequency domains. In the time domain, wave shape, beat interval, and amplitude variation are patterns; in the frequency domain, rhythm irregularities and signal energy distribution are patterns. Wavelet-based methods are well suited for ECG analysis due to the non-stationary nature of ECG and its time-varying frequency content. The Integer Haar Wavelet Transform gives a simple and efficient representation of the signal at different resolutions.

In addition to wavelet-based features, the recurrent neural network learns temporal representations directly from the ECG sequence. The attention mechanism further helps the model assign greater importance to diagnostically meaningful parts of the signal. Thus, the proposed method combines handcrafted signal processing principles with trainable deep learning representations.

CHAPTER 3

ALGORITHMS AND MODELS USED FOR DETECTION OF CHF

For efficient hardware implementation of algorithms, the feature extraction module is supposed to be computationally inexpensive which means utilizing minimum memory. So our work consists of a computationally efficient algorithm that is memory and power-efficient. The algorithm is appropriate for low-power wearable devices with limited resources.

3.1 Wavelet Transform

ECG is a non-stationary biomedical signal, which means that its frequency content changes with time. A method that only describes the signal in the time domain may miss frequency-related information, while a method that only describes the frequency domain may ignore when a particular event occurs. Wavelet transform is useful for such signals because it provides a time-frequency representation. It decomposes the signal into components that describe both slow variations and rapid changes.

In ECG analysis, wavelet decomposition is useful because different parts of the ECG waveform have different frequency characteristics. The QRS complex is sharp and usually contains higher-frequency information, whereas baseline trends and rhythm-related variations are represented more strongly in lower-frequency components. By separating approximation and detail coefficients, wavelet processing can provide a compact representation for classification.

A general wavelet-based feature extraction system applies low-pass and high-pass filtering followed by downsampling. The low-pass branch produces approximation coefficients, while the high-pass branch produces detail coefficients. Repeating this process over multiple levels gives a multiresolution view of the ECG signal. This project uses this idea through the Integer Haar Transform so that feature extraction remains computationally simple.

3.1.1 Mallat's Algorithm

Mallat's algorithm is a fast implementation of the discrete wavelet transform. The input signal at each level is filtered by a low pass filter and a high pass filter.

Then the outputs of these filters are downsampled by a factor of 2. The low pass output gives the approximation part of the signal. The high pass output gives the detail part. Again we can decompose the approximation output to get the next level of coefficients.

This repeated filtering and downsampling structure makes the algorithm efficient and applicable to multilevel analysis; In ECG processing, the lower level detail coefficients may retain sharp transitions, while the deeper approximation coefficients retain slower signal trends. The number of decomposition levels is chosen based on the sampling rate, length of segment and the type of features required by the classifier.

There exist several wavelet families for signal analysis such as Haar, Daubechies, Symlets and Coiflets. The choice of the wavelet impacts the compactness and interpretability of the extracted features. In this project, Haar-based processing is chosen because of its simple structure and its suitability for low-complexity implementation

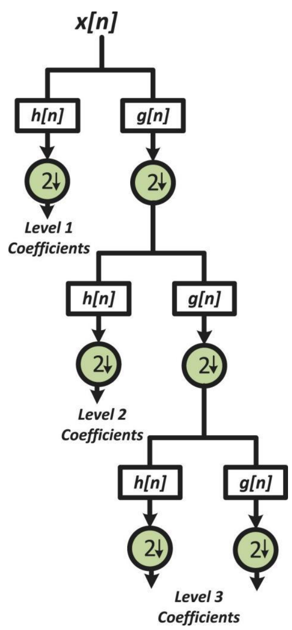

Figure 3.1. Mallat's Algorithm [2]

3.1.2 Integer Haar Transform (IHT)

The Haar wavelet is the simplest wavelet basis and is often used when computational efficiency is important. The low-pass and high-pass filter coefficients can be expressed as[2] follows [2]:

$$g = \frac{1}{\sqrt{2}}[1,\, 1] \tag{3.1}$$

$$h = \left[\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right] \tag{3.2}$$

The standard Haar transform uses floating point operations because of the scaling factor. Floating-point arithmetic may increase area, energy and implementation complexity for low-power hardware. This difficulty is overcome by Integer Haar Transform using integer operations and shift based computation . This makes it attractive for wearable ECG systems with limited memory and power resources.

The approximation coefficient (CA) and detail coefficient (CD) used in this work are represented as [2]:

$$C_A[n] = \left\lfloor \frac{1}{2}\left(x[2n] + \frac{1}{2}x[2n+1]\right)\right\rfloor \tag{3.3}$$

$$C_D[n] = x[2n] - x[2n+1] \tag{3.4}$$

Division by powers of two can be implemented using right-shift operations in digital hardware. Therefore, IHT is suitable for extracting ECG features in systems where the classifier may be deployed on a resource-constrained device.

3.2 Methodology Used

The objective of the methodology is to classify ECG records into CHF and normal categories using a structured signal-processing and machine-learning pipeline. Three types of models are considered: a decision tree, a deep neural network, and the multi-lead diagnostic attention-based recurrent neural network (MLDA-RNN). The decision tree and DNN are used as baseline models, while MLDA-RNN is treated as the main model because it is better suited for sequential multi-lead ECG data.

The data are divided at the patient level into training, validation, and testing subsets. This design is intentionally used to reduce data leakage. For the BIDMC CHF database, 10 patients are used for training, 2 patients are used for validation, and 3 patients are kept for testing. For the MIT-BIH Normal Sinus Rhythm database, 12 patients are used for training, 3 patients for validation, and 3 patients for testing. Since complete patients are assigned to only one subset, the final testing set behaves as blind data.

The BIDMC recordings represent the CHF class. For each patient, ECG signals are read, filtered, and stored in an array for further processing. The filtered signal is then divided into fixed-length windows of 750 samples. Each window is arranged as one row in the feature-generation pipeline. For the BIDMC training subset, this produces an array of size 238885×750. Integer Haar Transform is then applied repeatedly to obtain approximation coefficients (CA) and detail coefficients (CD). After six iterations of the transform, the CA and CD feature matrices for BIDMC training have dimensions 238885×12 each.

The same procedure is followed for validation and testing records of the BIDMC database. The validation feature matrices for CA and CD have dimensions 47782×12 and 47787×12, respectively. The testing feature matrices for both CA and CD have dimensions 71984×12. These values show that the pipeline converts long ECG recordings into compact wavelet-based representations while maintaining patient-wise separation.

The MIT-BIH Normal Sinus Rhythm database is processed in a similar manner. Since the BIDMC recordings are sampled at 250 Hz and the MIT-BIH recordings are sampled at 128 Hz, MIT-BIH signals are first resampled to 250 Hz to maintain consistency between the two classes. After resampling, ECG filtering is performed, and the resulting signal is segmented into windows of 750 samples. For the MIT-BIH training subset, the segmented array has size 181325×750.

Integer Haar Transform is then applied to the MIT-BIH segments. The CA and CD matrices for the training subset both have size 181325×12. For validation, both CA and CD matrices have size 44068×12. For testing, both feature matrices have size 43395×12. These transformed features are used to construct input sets for the machine learning and deep learning models.

Three feature configurations are evaluated so that the contribution of wavelet components can be compared. In the first configuration, CA and CD are concatenated and used together. In the second configuration, only CA features are used. In the third configuration, only CD features are used. This comparison is useful because CA represents lower-frequency approximation information, whereas CD represents detail information. The results show how each representation affects model behaviour.

1. Combined CA and CD features are used for training, validation, and testing.

2. Only approximation coefficient features (CA) are used.

3. Only detail coefficient features (CD) are used.

The same training and testing protocol is applied to all selected models. This ensures that the comparison among the decision tree, DNN, and MLDA-RNN is based on the same data division and feature configurations. The final analysis is then performed using accuracy, sensitivity, specificity, confusion matrices, and ROC curves.

3.3 Models Used

3.3.1 Deep Neural Network (DNN)

A deep neural network is a multilayer computational model that learns a mapping between input features and output classes. In this project, the DNN is used as a baseline deep learning classifier for wavelet-derived ECG features. Unlike a decision tree, which splits data using explicit threshold rules, a DNN learns nonlinear combinations of features through trainable weights and activation functions.

The main components of a DNN are described below:

1. **Input layer:** The input layer receives the feature vector generated from ECG segments. Depending on the experiment, this vector may contain CA features, CD features, or a combination of both.

2. **Hidden layers:** Hidden layers transform the input into progressively more useful internal representations. Each neuron computes a weighted sum of its inputs and passes it through an activation function. In this work, hidden dense layers are used to learn relationships among ECG features.

3. **Activation function:** Activation functions introduce nonlinearity into the network. Without them, the entire network would behave like a linear model. ReLU is used in hidden layers because it is simple and effective for deep learning models.

4. **Output layer:** The output layer produces the final class probability. Since the task is binary classification, a sigmoid activation is suitable for predicting whether a segment belongs to the CHF or normal class.

5. **Loss function:** Binary cross-entropy is used to compare predicted probabilities with true labels. The training process minimizes this loss.

6. **Optimization:** The Adam optimizer updates the network parameters during training. It adapts the learning rate for different parameters and is widely used for neural network training.

7. **Regularization:** Regularization methods such as dropout or weight penalties can reduce overfitting. This is important because biomedical datasets may contain limited patient-level diversity.

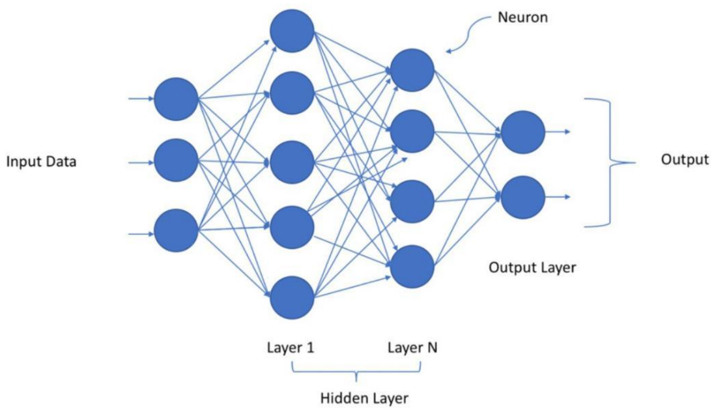

Figure 3.2. Deep Neural Network

Deep neural architecture with three hidden layers is shown below:

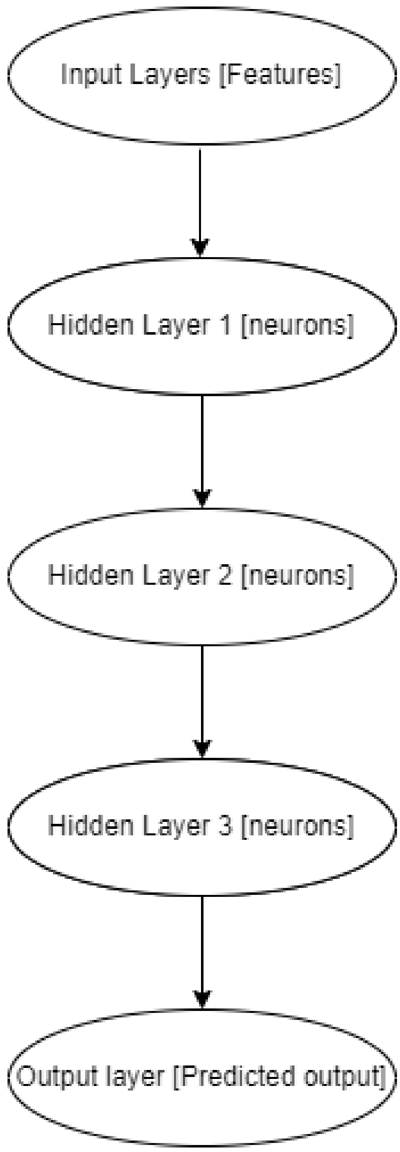

Figure 3.3. Deep Neural Network Architecture

In this project, the DNN input size depends on the number of selected features. The model uses dense hidden layers with 64 and 32 neurons, respectively. ReLU is used in the hidden layers, and sigmoid activation is used in the output layer. The model is trained using Adam optimization, binary cross-entropy loss, 10 epochs, and a batch size of 32. The DNN helps evaluate how well non-sequential deep learning performs on the extracted ECG features.

3.3.2 Decision Tree

A decision tree is a supervised learning model that predicts the output class by applying a sequence of feature-based decisions. Each internal node of the tree checks a condition on a feature, each branch corresponds to the outcome of that condition, and each leaf node represents a predicted class. Decision trees are easy to interpret because the path from root to leaf can be examined as a set of rules.

For ECG classification, a decision tree can be used to test whether wavelet-derived features contain enough discriminative information for separating CHF and normal signals. The model is computationally light and does not require complex training. However, its rule-based structure may not capture subtle temporal relationships in ECG signals. It may also overfit if the tree grows too deep.

In this work, the decision tree classifier is implemented using the `DecisionTreeClassifier` class from the scikit-learn library. The input to the classifier is generated from the same Integer Haar Transform pipeline used for the other models. This allows a fair comparison between a classical machine learning baseline and neural network-based models.

3.3.3 Multi-lead Diagnostic Attention-Based Recurrent Neural Network (MLDA-RNN) [1]

The MLDA-RNN model is selected because ECG data are sequential and multi-lead in nature. A recurrent network is more appropriate than a purely dense model when the order of samples carries diagnostic meaning. The model can process a sequence of ECG values and maintain hidden states that summarize previously observed information. This helps in learning rhythm-related and morphology-related changes across time.

A multi-lead design is useful because each ECG lead records cardiac electrical activity from a different viewpoint. Some leads may contain stronger evidence of CHF-related abnormalities than others. Therefore, the model should not treat all lead information as equally informative in every case. The attention mechanism addresses this by assigning greater importance to more useful parts of the learned representation.

The proposed architecture contains the following conceptual blocks:

Input Block. The input block receives preprocessed ECG segments. Depending on the experiment, the input may contain approximation coefficients, detail coefficients, or their combination. The purpose of this block is to arrange the ECG-derived features in a form that can be processed by the recurrent network.

Quality Aware Block. QualityAware Block. ECG recordings can have noise, baseline drift, or artifacts. A quality awareness stage helps to provide useful ECG segments to the model. The probability that the classifier learns non-cardiac noise patterns is reduced.

Recurrent Neural Network Encoding Block. Recurrent Neural Network Encoder Unit The recurrent encoding block learns the temporal dependencies in the sequence of ECG. However, basic RNN units can learn sequential data, but they often struggle with vanishing or exploding gradients when long dependencies are involved. Gated recurrent units (GRUs) solve this problem by gating mechanisms, but with fewer parameters than LSTM units. Thus, GRU is a good choice for ECG datasets with importance on computational simplicity and generalization.

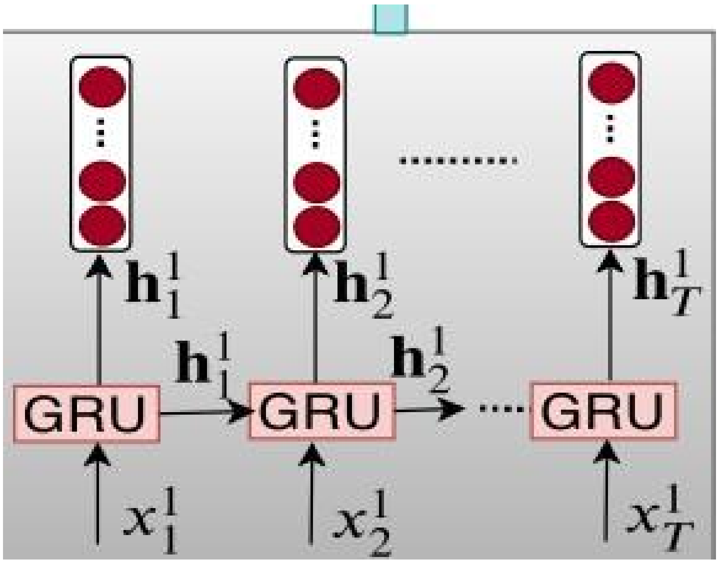

Figure 3.4. RNN Encoding block [1]

Let the ECG sequence for lead m be denoted as $\mathbf{x}^m = [x_1^m, x_2^m, \ldots, x_T^m]$. Here, x_t^m represents the value or feature at time step t for lead m. The GRU uses an update gate and a reset gate to decide how much past information should be retained and how much new information should be added.

The hidden state is updated as follows [1]:

$$\mathbf{h}_t^m = (1 - \mathbf{z}_t^m) \odot \mathbf{h}_{t-1}^m + \mathbf{z}_t^m \odot \widetilde{\mathbf{h}}_t^m \tag{3.5}$$

$$\mathbf{z}_t^m = \sigma(\mathbf{W}_{zi}^m x_t^m + \mathbf{W}_{zh}^m \mathbf{h}_{t-1}^m + \mathbf{b}_z^m) \tag{3.6}$$

In these expressions, $\sigma(\cdot)$ is the sigmoid function, $\mathbf{z}_t^m$ is the update gate, $\mathbf{h}_{t-1}^m$ is the previous hidden state, and $\widetilde{\mathbf{h}}_t^m$ is the candidate hidden state. The candidate hidden state is computed as:

$$\widetilde{\mathbf{h}}_t^m = \tanh\left(\mathbf{W}_{ci}^m x_t^m + \mathbf{W}_{cr}^m(\mathbf{r}_t^m \odot \mathbf{h}_{t-1}^m) + \mathbf{b}_c^m\right) \tag{3.7}$$

Here, $\mathbf{r}_t^m$ is the reset gate, $\odot$ indicates element-wise multiplication, and the $\mathbf{W}$ and $\mathbf{b}$ terms represent trainable weights and biases. These equations allow the GRU to control the flow of historical and current information through the sequence.

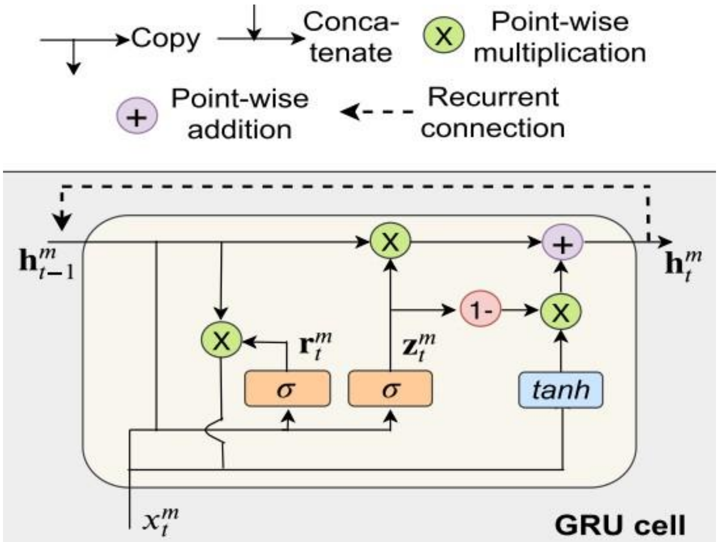

Figure 3.5. GRU Cell [1]

After recurrent encoding, the attention mechanism assigns weights to learned features so that the model can emphasize diagnostically important information. The final classification layer then predicts whether the ECG segment belongs to the CHF class or the normal sinus rhythm class.

3.4 Detailed Proposed Workflow

The workflow proposed here aims to transform raw multi-lead ECG recordings into reliable CHF predictions. The whole process consists of five main stages, namely data acquisition, pre-processing, feature extraction, temporal modelling and classification. Every step adds to the ultimate performance of the system. Any weakness at any stage can affect the final output, therefore the workflow is designed as a structured pipeline rather than a single isolated classifier.

3.4.1 Stage 1: Data Acquisition

In the first stage, the ECG records are taken from selected public databases. Normal class is represented by recordings from subjects with normal sinus rhythm, and

CHF class is represented by recordings from patients diagnosed with congestive heart failure. The signals need to be consistently processed before training as the different datasets may have different sampling frequencies and recording durations. In this step the recordings are grouped according to class labels and patients' identities.

3.4.2 Stage 2: Preprocessing and Normalization

The second step is to improve the quality of ECG signals. Biomedical signals are often contaminated by noise and baseline drift which will have negative impact on both feature extraction and neural network learning. Preprocessing removes these unwanted components. Once the noise reduction normalization is performed to put the signals on a common scale. This is important because the amplitude of ECG signals can vary due to electrode placement, patient body characteristics, and recording equipment. Normalization helps the model focus on relative changes in the waveform rather than the magnitude of the signal

3.4.3 Stage 3: Integer Haar Wavelet Transform

The third stage employs Integer Haar Wavelet Transform to obtain compact and informative features. The Haar transform decomposes a signal into approximation and detail components. The approximation coefficients capture the low frequency trend, while the detail coefficients capture the fast variations. Both of these components can be meaningful for ECG analysis. The approximation coefficients may relate to the overall rhythm. The detail coefficients may relate to the sharp QRS complexes and other abrupt transitions in the waveform. Integer implementation is good because it reduces computational complexity and avoids floating point heavy operations.

3.4.4 Stage 4: Recurrent Neural Encoding

The fourth stage employs a recurrent neural encoding block for learning temporal dependencies. ECG is a sequential signal. Interpretation of one sample or beat often depends on the previous and future patterns. We use GRU based recurrent units as they model temporal dependencies with less parameters than LSTM units. This is one reason why GRU is suitable for biomedical data sets where the amount of labelled data may be limited. The recurrent block produces hidden representations that summarize temporal information for each ECG lead.

3.4.5 Stage 5: Attention and Classification

The fifth stage uses attention and classification. The attention mechanism enables the model to attend more on more informative time steps or lead-specific features. In multi-lead ECG classification, not all leads are equally useful for all patients and segments. Some leads may show more marked abnormalities related to CHF than others. The attention mechanism helps the network to assign a larger weight to these useful features. The last layer is the classification layer which gives the probability of the segment of ECG to belong to the CHF class.

Table 3.1. Summary of the proposed CHF detection workflow

Stage	Purpose	Expected output
Data acquisition	Collect normal and CHF ECG recordings	Labelled patient-wise ECG dataset
Preprocessing	Remove noise and normalize amplitude	Clean ECG segments
Wavelet transform	Extract time-frequency information	Approximation and detail features
RNN encoding	Learn temporal dependencies	Hidden sequence representation
Attention and classification	Focus on informative patterns and classify	CHF or normal prediction

3.5 Evaluation Metrics

Accuracy is not sufficient to evaluate the performance of the model, especially in the medical classification. The classifier may be accurate if one class dominates the data set, but will not detect positive cases that are of clinical importance. Therefore, in this work we consider multiple evaluation metrics including accuracy, sensitivity, specificity, confusion matrix, receiver operating characteristic curve. The accuracy is the total number of correctly classified samples. Sensitivity (or true positive rate) is a measure of the ability of the model to correctly detect CHF cases. This is clinically significant as a missed CHF case may lead to a delay in treatment. Specificity is the ability of the model to correctly recognize normal cases. High specificity is also important, as too many false alarms may lead to loss of trust in the system and increase unnecessary clinical workload. So the confusion matrix gives a more detailed overview of right and wrong predictions. It is made up of true positives, true negatives, false positives and false negatives. The ROC curve shows the trade-off between the sensitivity and the false positive rate at different threshold values. A curve closer to the top-left corner would be a better model. Together, the metrics offer a more complete picture of the model than any single number.

Table 3.2. Evaluation metrics used for CHF classification

Metric	**Formula/Meaning**	**Importance**
Accuracy	Ratio of correct predictions to total predictions	Shows overall classification performance
Sensitivity	$TP/(TP+FN)$	Measures ability to detect CHF cases
Specificity	$TN/(TN+FP)$	Measures ability to identify normal subjects
False positive rate	$FP/(FP+TN)$	Indicates how often normal samples are misclassified
Confusion matrix	Counts TP, TN, FP, and FN	Gives detailed error distribution
ROC curve	Sensitivity versus false positive rate	Evaluates threshold-dependent behaviour

3.6 Implementation Considerations

The proposed model implementation requires careful selection of hyperparameters. The key parameters are segment length , batch size , number of recurrent units , number of dense neurons , learning rate , number of epochs and drop out rate . If the model is too small, it can underfit and fail to learn useful patterns in the ECGs. If it is too large it may overfit and perform poorly on blind test patients. Thus, the architecture has to trade off learning capacity against generalization. Another important factor to consider is class imbalance. If one class has more segments than the other, the model can be biased towards the majority class. This problem might be dealt with by class balancing, careful sampling or class-weighting strategies. Also, training curves should be monitored to detect overfitting. If training accuracy goes up and validation accuracy goes down, the model is memorizing training data instead of learning general features. Inference speed is also important for real-time applications. Fast prediction and low power consumption are the requirements of a wearable or portable monitoring system. Efficient wavelet processing and a GRU-based recurrent model are used to cut down the computational load compared with heavier architectures. This makes the proposed approach more convenient for practical health monitoring systems.

CHAPTER 4

EXPERIMENTS AND RESULTS

4.1 Machine Learning Models

In the previous chapters, we have described the techniques and models in a detailed way. This chapter will contain the results of those experiments. At first, we will display the results of deep neural networks (DNN) and decision trees in three different ways.

1. By using only the detailed coefficient (CD) of Haar integer wavelet and generating the training and testing dataset.
2. By using only the approximate coefficient (CA) of the Haar integer wavelet and generating the training and testing dataset.
3. By combining detailed coefficient and approximate coefficient of Haar integer wavelet and generating the training and testing dataset.

Table 4.1. Metrics result of different models by using CD of HIT

Models	**Accuracy**	**Sensitivity**	**Specificity**	**F1 score**	**AUC**
Decision Tree	0.70025	0.6975	0.70225	0.6626405	0.69988
DNN	0.7884	0.739204	0.824289	0.746738	0.7817468

Table 4.2. Metrics result of different models by using CA of HIT

Models	**Accuracy**	**Sensitivity**	**Specificity**	**F1 score**	**AUC**
Decision Tree	0.8696787	0.76854	0.9435339	0.832717	0.856038
DNN	0.891	0.817214	0.94485309	0.863528	0.8810336

Table 4.3. Metrics result of different models by combining CA and CD of HIT together

Models	Accuracy	Sensitivity	Specificity	F1 score	AUC
Decision Tree	0.92538	0.8911311	0.950397	0.90975517	0.92076
DNN	0.9586	0.94729	0.9669	0.9508	0.95709
ARNN	0.960668	0.9622998	0.9596849	0.9484646	0.9609923

After observing the results of different models by three different methods – (1) by considering only CA, (2) by considering only CD, and (3) by combining both CA and CD – we can observe that all models DNN, Decision tree, and RNN give better results when CA and CD are combined. However, without combining CA and CD, models are giving better results with CA as compared to CD. So it is better to take both CA and CD into consideration; however, there is a problem with taking both CA and CD in that the algorithm will take more space and more time to compute the output. Compared to all three models we can see that ARNN outstands every other model by giving an accuracy of 96.066%.

The most important part of our results as compared to previous work carried out till now is that our results are produced on blind data. While dividing our datasets into training and testing parts we have not used any train-test split technique which divides our dataset randomly; instead, we have divided our dataset based on patients. So in this way, our model is completely unaware and is not friendly with our dataset. As per the literature survey we carried out, we can clearly state that the way we are dividing our dataset into training and testing has not been used till now, and the accuracy we have obtained with our ARNN model is remarkable on our blind dataset.

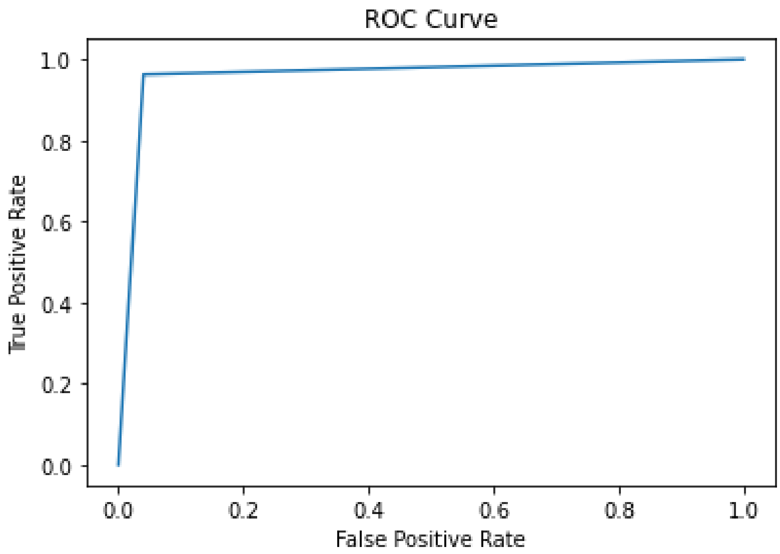

Figure 4.1. ROC Curve of ARNN Model

The above AUC-ROC figure is also called an Area Under the Receiver Operating Characteristics. ROC is a probability curve and AUC represents the degree or measure of separability for the model; it generally indicates how much our model is capable of separating different classes. It is plotted between the sensitivity and (1 − specificity). The higher the AUC, the better the model is capable of separating between different classes. The ROC curve of the ARNN model is approximately near to 1, which means that our model has a very good measure of separability and is very much capable of separating patients with the disease and without the disease.

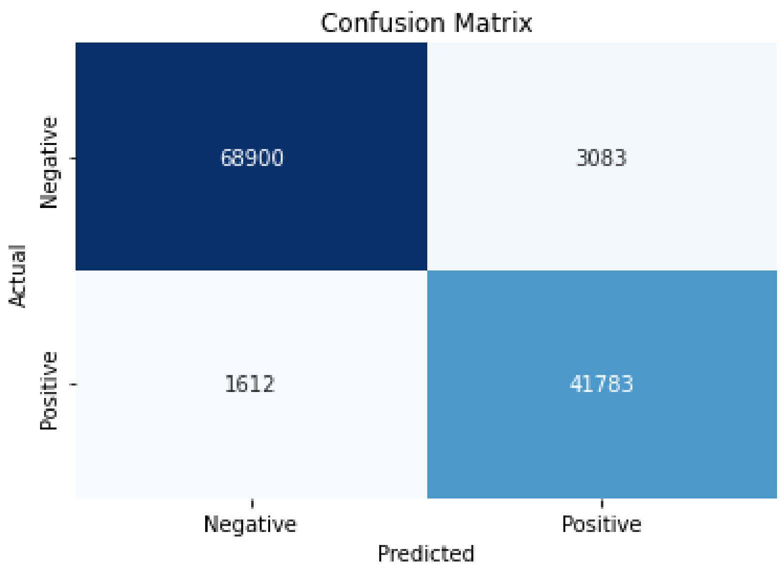

Figure 4.2. Confusion matrix of ARNN Model

A confusion matrix displays the counts of true positives, true negatives, false positives, and false negatives to summarise the performance of a classification model. In the above confusion matrix of the ARNN model, we can see that values of TN (true negative) and TP (true positive) are so high as compared to FN (False negative) and FP (False positive), hence our model is doing a very good classification.

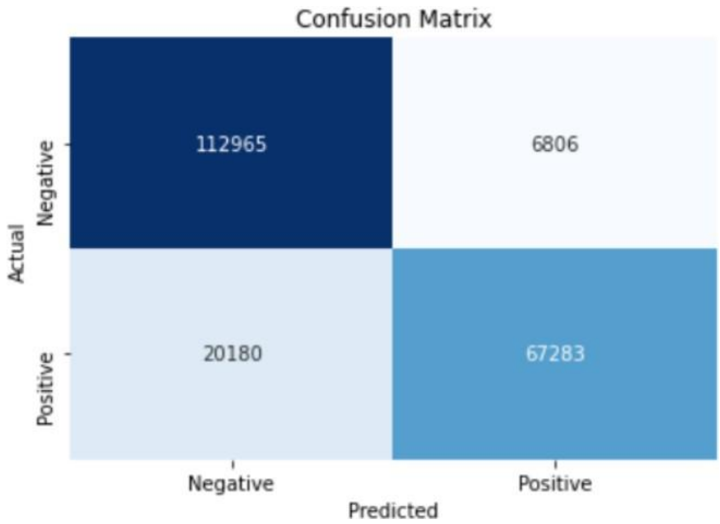

Figure 4.3. Confusion matrix of DT by taking CA

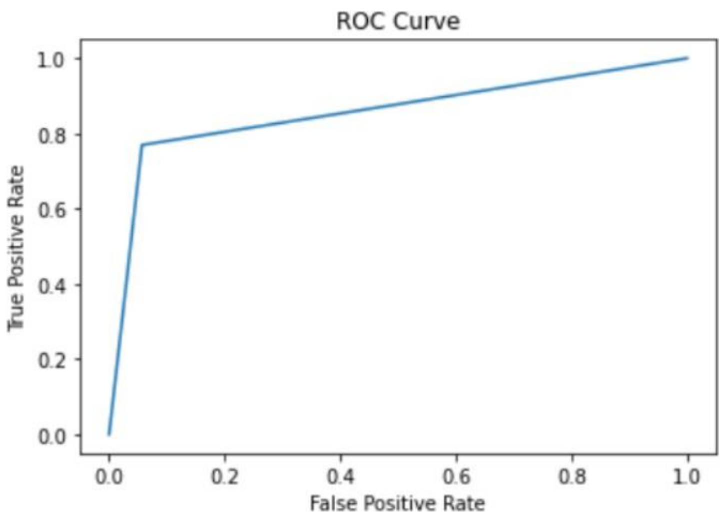

Figure 4.4. ROC curve of decision tree by taking CA

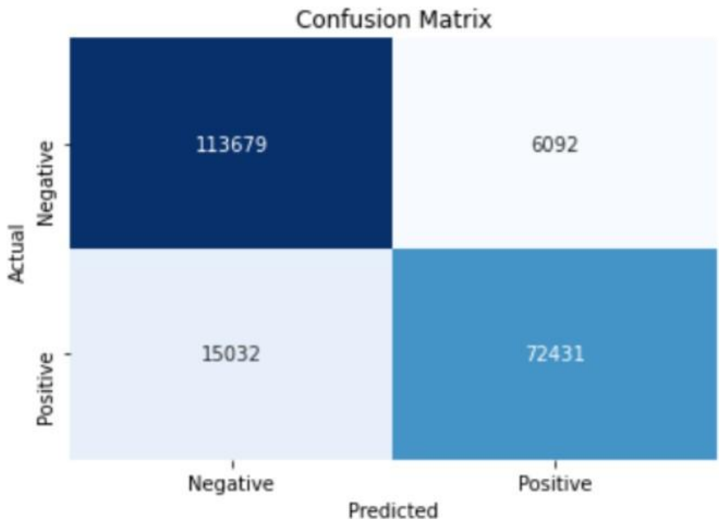

Figure 4.5. Confusion matrix of DNN by taking CA

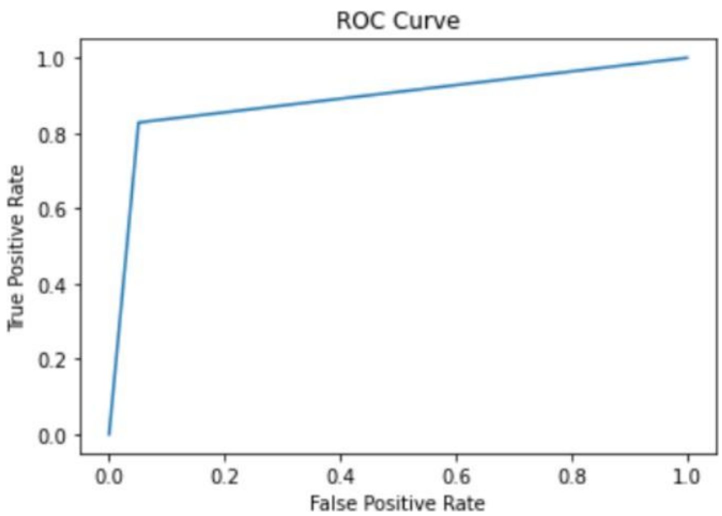

Figure 4.6. ROC curve of DNN by taking CA

Confusion matrix and ROC curve of DNN by taking CA are showing much better results as compared to confusion matrix and ROC curve of DT by taking CA. ARNN ROC curve and confusion matrix are showing better results as compared to ROC curve and confusion matrix of both DNN and DT.

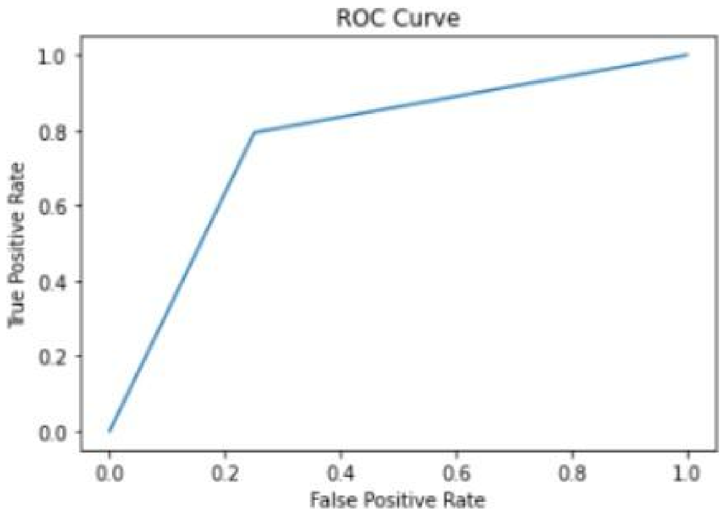

Figure 4.7. ROC curve of DNN by taking CD

From this figure, we can see that there is approximately 80% chance that our DNN model will be able to distinguish between positive class and negative class. Also, we can see that the ROC curve of DNN by considering CD is behaving poorly as compared to the ROC of DNN by taking CA.

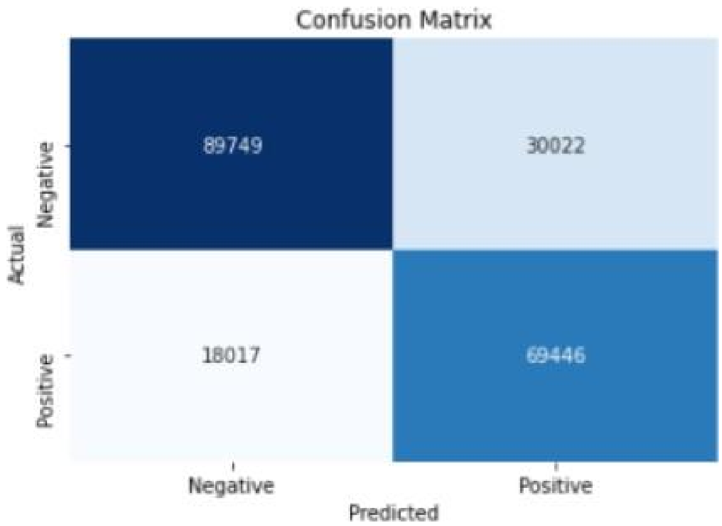

Figure 4.8. Confusion matrix of DNN by taking CD

We can see here that the DNN model by considering only CA was able to identify more TP and TN as compared to the DNN model by considering only CD.

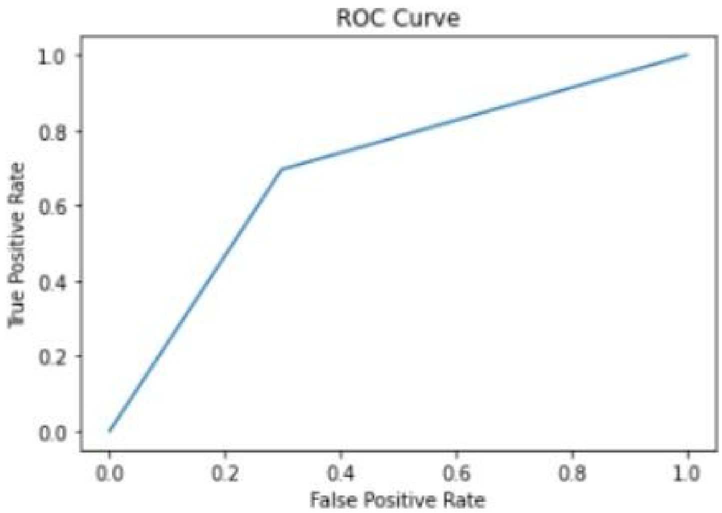

Figure 4.9. ROC curve of Decision tree by taking CD

The ROC curve of DT by CA is behaving well as compared to DT by CD. The ROC curve of ARNN is behaving well as compared to the ROC curve of DT by taking CD.

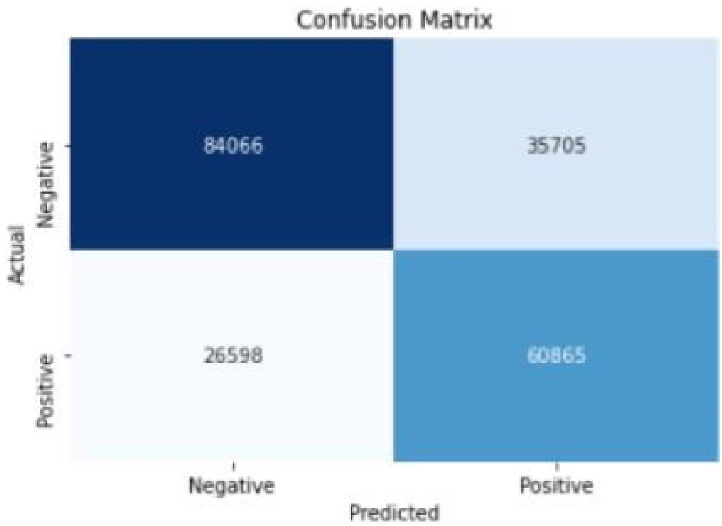

Figure 4.10. Confusion matrix of DT by taking CD

We can see here that the DT model by considering only CA was able to identify more TP and TN as compared to the DT model by considering only CD. If we compare with ARNN then ARNN was able to identify more TP and TN.

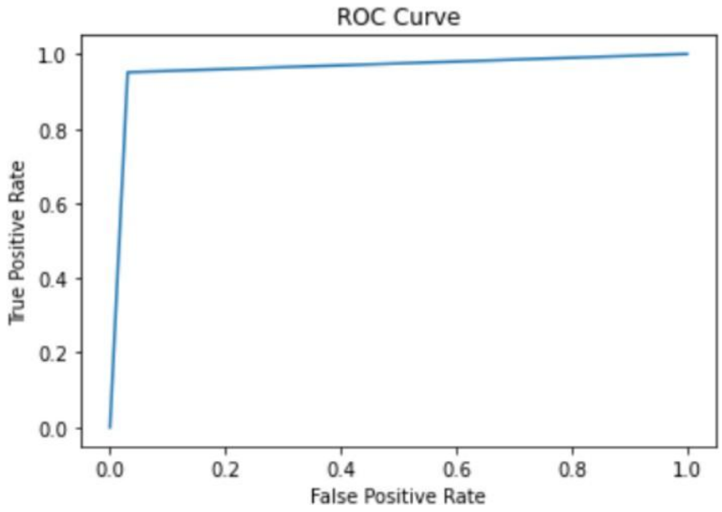

Figure 4.11. ROC curve of DNN by taking CA and CD

For our DNN model, the ROC curve by taking both CA and CD into consideration is better than the ROC of only CA and ROC of only CD. So DNN will be able to distinguish clearly between positive and negative classes more when both CA and CD are taken into consideration.

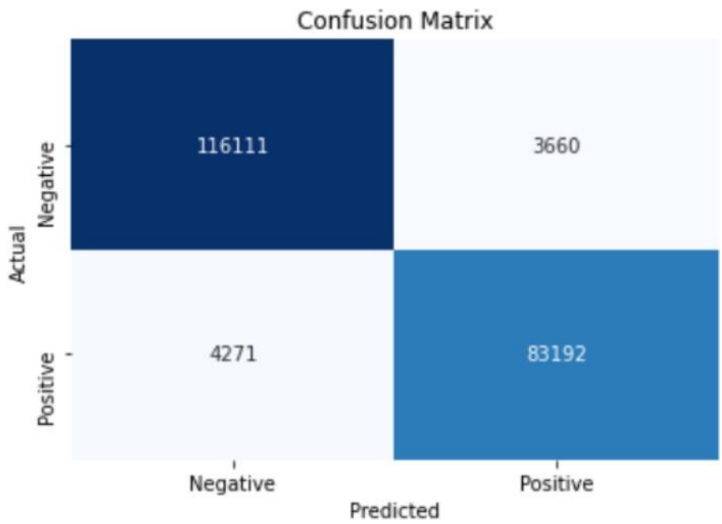

Figure 4.12. Confusion matrix of DNN by taking CA and CD

The confusion matrix of DNN by taking both CA and CD is identifying more TP and TN as compared to the confusion matrix of DNN by taking only CA and by considering only CD.

By observing all the above ROC curves and confusion matrices, we can say that the ARNN model is behaving much better than all other models.

From the above table, we can see that the accuracy of some of the previous work is

Table 4.4. Comparison of proposed work with state of the art method

Parameters	**[6]**	**[1]**	**[7]**	**[8]**	**Proposed method**
Features used	Raw ECG beats	ECG beats	HRV features, wavelet based features	HRV features	ECG signal
Classifier	multilayered RNN	MLDA-RNN	SVM	Multiscale Residual UNet++	MLDA-RNN
Evaluation Scheme	Train-Test split	Train-Test split	Train-Test split	Train-Test split	Blind data
Database used	BIDMC-CHF, PTBDB, and MIT-BIH NSRDB	STAFF III, PTB diagnostic	MIT-BIH Arrhythmia, BIDMC and MIT-BIH NSRD	NSR-RR, CHF-RR interval database	BIDMC, MIT-BIH NSRD
Accuracy	98.57%	97.79%	93.33%	89.83%	96.66%
Sensitivity	99.01%	—	—	—	96.229%
Specificity	97.89%	—	—	—	95.968%

better than our proposed work, but it is just because the evaluation scheme they have used is a train-test split method so in that case the model has some idea of our testing dataset. In our proposed method we are dividing our dataset into training and testing based on patients. So in this way, our model is completely unaware of the testing dataset. If compared to previous work, our model is giving very prominent results in such a case.

4.2 Result Interpretation

The experimental results show that the proposed MLDA-RNN model performs better than the conventional machine learning and deep neural network baselines used in this work. The main reason is that MLDA-RNN is designed for sequential ECG data. Unlike a decision tree, which makes decisions using fixed feature thresholds, the recurrent model learns how ECG patterns evolve over time. This is important because CHF-related abnormalities may not be visible in a single isolated sample but may become clear across a sequence of beats.

The attention component also enhances interpretability at the model level. This allows the network to concentrate more on the more useful temporal features. In ECG analysis, certain intervals may provide more powerful diagnostic information than others. For example, the QRS morphology and the irregularity of the rhythm may contain more information than relatively flat portions of the signal. The attention mechanism allows to use the signal in a more efficient way by giving different weights to different parts. One of the most important parts of the result is the evaluation in blind test. The test data contain subjects that are not seen during training, as patient-wise separation is employed. This makes the task more complicated, but also more meaningful. A model that does well on blind data is more likely to be generalizable to new patients. Therefore, the achieved accuracy of 96.06% suggests that the proposed method has learned useful disease-related patterns rather than merely memorized patient-specific characteristics.

4.3 Comparison with Baseline Models

The decision tree model provides a simple interpretable baseline. It is easy to train and can show which features are important for classification But decision trees are limited in their capacity to model complicated temporal relations. They partition the feature space with threshold based rules and do not learn long range dependencies in ECG signals naturally. Thus, they tend to perform worse on classification problems that depend on sequential patterns.The deep neural network baseline improves the decision tree by learning non-linear combinations of features. Dense layers can learn complex relationships between the input features, but they still treat the input as a fixed feature vector. A standard DNN might not capture the sequential nature of ECG recordings unless temporal features are carefully engineered. This explains the reason why the DNN is better than the decision tree in some cases while it is still weaker than the attention-based recurrent approach.

The MLDA-RNN model has a number of advantages. It can process sequential ECG information, use multi-lead input and focus on informative parts of the signal by attention. This property makes it more appropriate for CHF detection . The improvement in performance is especially significant because the evaluation is performed on blind data.

Table 4.5. Qualitative comparison of models used in this work

Model	**Strength**	**Limitation**	**Suitability for ECG**
Decision Tree	Simple and interpretable	Weak for complex temporal patterns	Suitable as a baseline
DNN	Learns nonlinear feature combinations	Does not naturally model sequence order	Useful when strong features are available
MLDA-RNN	Learns multi-lead temporal patterns with attention	Requires more training time and tuning	Highly suitable for CHF detection

4.4 Clinical Relevance of the Results

The results of this project are clinically relevant, supporting the feasibility of automated ECG-based CHF screening. A model that performs well can assist clinicians in identifying ECG segments that may warrant closer examination. Such a system can be useful in hospitals, remote monitoring centers and wearable health platforms. It can be used also as a preliminary screening tool in environments where expert cardiologists are not immediately available.

But the model should be applied as a decision-support tool, not a substitute for clinical diagnosis. Diagnosis of CHF requires medical history, symptoms, imaging and laboratory findings, in addition to ECG. The proposed system can be useful to generate alerts and assist interpretation but the final decision should be made by medical professionals. Future validation on larger and more diverse datasets will further strengthen the clinical applicability of the model.

4.5 Error Analysis

Misclassification can occur for a number of reasons. False positive results may occur in some normal subjects with ECG patterns resembling abnormal rhythms. Likewise, some CHF patients may have segments that appear reasonably normal giving rise to false negatives. Predictions may also be affected by noise and artifacts, and differences in recording conditions. It is important to know about these mistakes because medical systems must be reliable and transparent.

False negatives are missed CHF cases and are particularly important in CHF detection. Longer ECG windows, additional clinical features, or aggregation at the patient level may reduce false negatives. On the other hand, false positives may lead to unnecessary follow-up but are generally less dangerous than missed disease cases. Therefore a practical system may select a threshold which yields higher sensitivity at an acceptable specificity level.

4.6 Limitations of the Present Work

The proposed method yields promising results but has some limitations. First, the study is based on publicly available datasets and the number of patients may be limited compared to large hospital-scale datasets. Second, the model is tested for binary classification between CHF and normal sinus rhythm. The real clinical environments could be composed of other conditions like arrhythmia, myocardial infarction and conduction abnormalities. Third, ECG acquisition settings may differ from device to device which may influence model generalization.

Another con is that deep learning models require fine tuning and can be difficult to understand completely. While attention improves interpretability, it does not provide the same level of explanation as a rule-based clinical decision. Future work can include explanations techniques such as saliency maps, lead-wise attention visualization and beat-level contribution analysis.

4.7 Practical Deployment Considerations

For deployment in a real healthcare system, the model needs to be integrated with an ECG acquisition device and a signal quality assessment module. First, the system should check if the incoming ECG is usable. If the signal is noisy, the device can request re-recording or ignore the segment. These clean segments can then be passed into the feature extraction and classification pipeline. The final output may be a risk score, rather than a simple yes/no label.

Data privacy is another important deployment issue. ECG data are sensitive medical records and must be stored and transmitted in a secure way. But if processing is done in the cloud, encryption and access control are needed. On-device processing can mitigate the privacy risk and communication cost for wearable devices. The computational efficiency of Integer Haar Wavelet Transform and GRU-based modelling supports this type of deployment.

Further enhancement of proposed system can be done by adding patient level aggregation.

Instead of a decision based on one segment, predictions of several segments can be combined. This can reduce the effect of noisy or uncertain segments and lead to a more stable final decision This is useful for long term monitoring where ECG is recorded continuously over minutes or hours.

CHAPTER 5

CONCLUSION AND FUTURE WORK

5.1 Conclusion

In this dissertation, a signal-processing features and machine learning model-based approach for congestive heart failure detection using ECG was proposed. The study starts with the clinical motivation for automated CHF screening. Then, the prior art in ECG classification methods is reviewed and the BIDMC and MIT-BIH datasets are described.The methodology then described preprocessing, patient-wise data separation, Integer Haar Transform feature extraction, and the models evaluated in this work. Among the tested models, MLDA-RNN achieved the strongest performance with an accuracy of 96.06%.

A major contribution of this work is the use of patient-wise testing. Instead of randomly mixing ECG segments, complete patient recordings were assigned to training, validation, or testing groups. This makes the evaluation more difficult because the model is tested on unseen subjects rather than on segments from already observed patients. The obtained result is therefore more meaningful for practical screening than a simple random split. The performance of the attention-based recurrent model suggests that temporal multi-lead ECG information is valuable for CHF recognition under a stricter evaluation setting.

5.2 Future Work

As a direction for future work, the following aspects could be explored:

- Extending the model to handle multi-class classification of different severity levels of congestive heart failure.
- Testing the proposed approach on larger and more diverse ECG datasets to improve the generalizability of the model.
- Investigating the deployment of the Integer Haar Wavelet Transform based fea-

ture extraction on actual low-power wearable hardware platforms.

- Exploring more advanced attention mechanisms and newer RNN architectures such as transformer-based models for further accuracy improvement.
- Combining ECG-based features with other clinical parameters to achieve a more holistic diagnosis of CHF.

BIBLIOGRAPHY

[1] E. Prabhakararao and S. Dandapat, “Myocardial infarction severity stages classification from ECG signals using attentional recurrent neural network,” *IEEE Sensors Journal*, vol. 20, no. 15, pp. 8711–8720, 2020.

[2] M. Janveja and G. Trivedi, “An area and power efficient VLSI architecture for ECG feature extraction for wearable IoT healthcare applications,” *Integration*, vol. 82, pp. 96–103, 2022.

[3] G. Lippi and F. Sanchis-Gomar, “Global epidemiology and future trends of heart failure,” *AME Medical Journal*, vol. 5, 2020.

[4] U. R. Acharya, H. Fujita, S. L. Oh, Y. Hagiwara, J. H. Tan, M. Adam, and R. S. Tan, “Deep convolutional neural network for the automated diagnosis of congestive heart failure using ECG signals,” *Applied Intelligence*, vol. 49, 2019.

[5] A. Darmawahyuni, S. Nurmaini, M. Yuwandini, M. N. Rachmatullah, F. Masdung, and B. Tutuko, “Congestive heart failure waveform classification based on short time-step analysis with recurrent network,” *Informatics in Medicine Unlocked*, vol. 21, 2020.

[6] E. Prabhakararao and S. Dandapat, “Congestive heart failure detection from ECG signals using deep residual neural network,” *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 53, no. 5, pp. 3008–3018, 2023.

[7] S. Nahak and G. Saha, “A fusion based classification of normal, arrhythmia and congestive heart failure in ECG,” in *2020 National Conference on Communications (NCC)*, 2020, pp. 1–6.

[8] L. Zou, Z. Huang, X. Yu, J. Zheng, A. Liu, and M. Lei, “Automatic detection of congestive heart failure based on multiscale residual UNet++: From centralized learning to federated learning,” *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–13, 2023.

[9] T. Liu, Y. Si, W. Yang, J. Huang, Y. Yu, G. Zhang, and R. Zhou, “Inter-patient congestive heart failure detection using ECG-convolution-vision transformer network,” *Sensors*, vol. 22, no. 9, p. 3283, 2022.

[10] M. R. Prusty, T. N. Pandey, P. S. Lekha, G. Lellapalli, and A. Gupta, "Scalar invariant transform based deep learning framework for detecting heart failures using ECG signals," *Scientific Reports*, vol. 14, p. 2765, 2024.

[11] A. Mahmoud *et al.*, "Electrocardiogram signals classification using deep-learning-based incorporated convolutional neural network and long short-term memory framework," *Computers*, vol. 13, no. 2, p. 55, 2024.

[12] I. Odinaka *et al.*, "Cardiac arrhythmia detection using deep learning and time-frequency representation of ECG signals," *BMC Medical Informatics and Decision Making*, vol. 23, p. 230, 2023.

[13] S. M. Lee, T. Kim, M. Shin, J.-O. Choi, M. J. Chung, and D. Kim, "Deep learning algorithm for detection of acute heart failure using standard ECG waveforms," *European Heart Journal – Digital Health*, 2025.

[14] Y. N. Fuadah and K. M. Lim, "Optimal classification of atrial fibrillation and congestive heart failure using machine learning," *Frontiers in Physiology*, vol. 12, 2022.

Aman jolly

AnkurFinalThesis

PHD

Document Details

Submission ID

trn:oid:::27535:140988706

Submission Date

May 30, 2026, 12:01 AM GMT+5:30

Download Date

May 30, 2026, 12:05 AM GMT+5:30

File Name

AnkurFinalThesis.pdf

File Size

1.7 MB

52 Pages

11,431 Words

63,664 Characters

14% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

Filtered from the Report

- Bibliography
- Quoted Text
- Small Matches (less than 8 words)

Match Groups

128 Not Cited or Quoted 13%
Matches with neither in-text citation nor quotation marks

7 Missing Quotations 1%
Matches that are still very similar to source material

0 Missing Citation 0%
Matches that have quotation marks, but no in-text citation

0 Cited and Quoted 0%
Matches with in-text citation present, but no quotation marks

Top Sources

9% Internet sources

8% Publications

10% Submitted works (Student Papers)

Integrity Flags

0 Integrity Flags for Review

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.

Match Groups

- **128** Not Cited or Quoted 13%
  Matches with neither in-text citation nor quotation marks
- **7** Missing Quotations 1%
  Matches that are still very similar to source material
- **0** Missing Citation 0%
  Matches that have quotation marks, but no in-text citation
- **0** Cited and Quoted 0%
  Matches with in-text citation present, but no quotation marks

Top Sources

9% Internet sources
8% Publications
10% Submitted works (Student Papers)

Top Sources

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	Internet	dspace.dtu.ac.in:8080	1%
2	Publication	Terzi, Merve Begüm. "Artificial Intelligence-Based Hybrid Anomaly Detection and ...	<1%
3	Publication	Rushik Parmar, Meenali Janveja, Jan Pidanic, Gaurav Trivedi. "Design of DNN-Bas...	<1%
4	Student papers	dtusimilarity on 2024-05-29	<1%
5	Publication	L. B. Fertig, J. H. McClellan. "Frequency Modulation", Wiley, 1999	<1%
6	Internet	www.hindawi.com	<1%
7	Internet	ebin.pub	<1%
8	Internet	lmu.pressbooks.pub	<1%
9	Student papers	Delhi Technological University on 2026-05-26	<1%
10	Internet	lisa.unical.it	<1%

Aman jolly

AnkurFinalThesis

PHD

Document Details

Submission ID

trn:oid:::27535:140988706

Submission Date

May 30, 2026, 12:01 AM GMT+5:30

Download Date

May 30, 2026, 12:07 AM GMT+5:30

File Name

AnkurFinalThesis.pdf

File Size

1.7 MB

52 Pages

11,431 Words

63,664 Characters

*% detected as AI

AI detection includes the possibility of false positives. Although some text in this submission is likely AI generated, scores below the 20% threshold are not surfaced because they have a higher likelihood of false positives.

Caution: Review required.

It is essential to understand the limitations of AI detection before making decisions about a student's work. We encourage you to learn more about Turnitin's AI detection capabilities before using the tool.