# Arrhythmia Classification Using Hybrid and Standalone Deep Learning Models

MSc Research Project
Msc in Artificial Intelligence

Parkhi Bhardwaj
Student ID: 23163861

School of Computing
National College of Ireland

Supervisor: Kislay Raj

**National College of Ireland**
**Project Submission Sheet**
**School of Computing**

| | |
|---|---|
| **Student Name:** | Parkhi Bhardwaj |
| **Student ID:** | 23163861 |
| **Programme:** | Msc in Artificial Intelligence |
| **Year:** | 2024 |
| **Module:** | MSc Research Project |
| **Supervisor:** | Kislay Raj |
| **Submission Due Date:** | 12/12/2024 |
| **Project Title:** | Arrhythmia Classification Using Hybrid and Standalone Deep Learning Models |
| **Word Count:** | XXX |
| **Page Count:** | 22 |

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

**ALL** internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

| | |
|---|---|
| **Signature:** | |
| **Date:** | 12th December 2024 |

## PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

| | |
|---|---|
| Attach a completed copy of this sheet to each project (including multiple copies). | ☐ |
| **Attach a Moodle submission receipt of the online project submission**, to each project (including multiple copies). | ☐ |
| **You must ensure that you retain a HARD COPY of the project**, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer. | ☐ |

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

| **Office Use Only** | |
|---|---|
| Signature: | |
| Date: | |
| Penalty Applied (if applicable): | |

# Arrhythmia Classification Using Hybrid and Standalone Deep Learning Models

Parkhi Bhardwaj
23163861

**Abstract**

Arrhythmias are the severe conditions which need precise and fast reports and often difficult to detect due to differences in ECG signal patterns and an imbalance between classes of data. This work: seeks to study the arrhythmias classification problem by training (i) standalone and (ii) hybrid deep learning models on the standard MIT-BIH Arrhythmia Dataset. In the study, four architectures; RNN, LSTM, GRU, and CNN connected LSTM are proposed based on the challenges which are solved in robust preprocessing techniques such as segmentation, normalization, augmentation and technique in handling class imbalance using SMOTE and class weighting. Of these, the CNN-LSTM has the highest overall performance by integrating Spatial Feature Extraction with Temporal Dependency Learning by scoring high G-mean and best accuracy, sensitivity and specificity rates of all the heartbeat classes particularly the minority ones. Comparing with the RNN and GRU models, standalone models computational efficiency has its drawbacks on long-term dependencies, therefore, the CNN-LSTM is the best option for arrhythmia detection. Hence, this research discusses the effectiveness of using automated algorithms for classification of arrhythmias and consequently the possibilities of reducing the extent to which experts' assistance is relied upon as well as increasing the possibilities of the general availability of the data set for real-world clinical use.

## 1 Introduction

### 1.1 Motivation and Problem Background

Arrhythmias, irregular heartbeat rhythms that may start as simple discomfort and progress to serious complications like stroke or sudden cardiac arrest, presents great diagnostic and therapeutic dilemmas for health care. More than 31% world populations were died in the last couple of years with cardiovascular disease (CVD). Of which 85% were because of heart attack (Anbarasi and Ravi; 2023). The initial examination apparatus is the electrocardiogram (ECG) which measures electric activity of the cardiac muscles. However, current approaches to ECG analysis are time-consuming, inaccurate and may be limited to domain experts because some differences can be extremely subtle between different types of arrhythmia. More specifically, due to the continuously increasing rate of occurrences of cardiovascular diseases worldwide, precise, efficient and preferably automated procedures for the classification of described arrhythmias is highly significant. Fig1 describes the comparison between arrhythmia and the normal heartbeat.

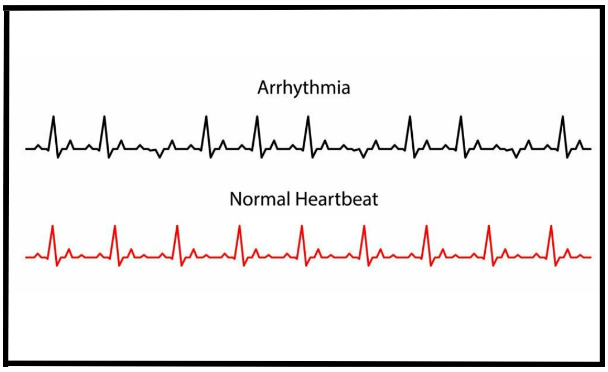


Figure 1: Arrhythmia vs Normal Heartbeat
(Continental Hospitals; 2024)

The most important development in ECG analysis comes from the application of machine learning (ML) and deep learning (DL). This paper focuses on using Recurrent Neural Networks (RNNs), Long Short Term Memory (LSTMs) networks and Convolutional Neural Networks (CNNs) to identify arrhythmia by extracting features from raw ECG data without labeling. There are some networks that combine spatial and temporal features extraction such as CNN and LSTM, therefore it is especially suitable for complicated sequential data such as ECG. However, there are still several problems; for example, class imbalance in datasets, differences in the signal morphology between patients, which affect the generality and accuracy of such models.

This work aims at evaluating the distinct and combined deep learning models for classifying of arrhythmia using the standard MIT-BIH Arrhythmia dataset. It aims at overcoming a number of drawbacks of the prior methods by constructing stable preprocessing paradigms, applying data augmentation techniques and employing novel paradigms of deep neural networks to enhance the classification performance as well as the model applicability.

## 1.2 Problem Statement

The research question concerns the automated classification of ECG signals into five heartbeat classes: Normal, atrial premature, premature ventricular contraction, fusion of ventricular & normal, and fusion of paced & normal, in the presence of noisy data as well as imbalance of the classes. The sophisticated structure of ECG signals and the presence of noise make the application of rule-based and feature-engineered methods inexploitable. This points to the need for the creation of generalizable deep learning models which convert raw ECG data to a form containing meaningful patterns and dependency information.
Key challenges include:

- Correction of the class imbalance issue where the detection of minor classes such as atrial premature beats is challenging.
- Accomplishing the modeling of spatial and temporal structure that is peculiar to ECG signals.
- Improving model insensitivity to variations in the signal waveform and noise.

## 1.3 Research Question

How can standalone and hybrid deep learning models be proposed to accurately, robustly, and efficiently classify arrhythmias from ECG signals keeping in mind, class imbalance and variability in signal morphology?

## 1.4 Research Solution

In this research, the identification of a consistent pattern for classifying arrhythmias using individual and compound deep networks is suggested. The proposed solution consists of creating advanced preprocessing pipeline which makes data more homogeneous and reduce variability by segmentation, normalization and augmentation. Deep learning architectures, such as RNN, LSTM, GRU, and CNN-LSTM hybrid models are designed and implemented, in which each architecture was tailored to maximize efficiency by processing aspects of structured ECG signal data. Also, the Synthetic Minority Oversampling Technique (SMOTE) and classes weights are adopted to address the problem of the class imbalance while training. In assessing the performance of the models, measures such as accuracy, precision, recall, the F1-score and the area under the ROC curve are used which give a complete report of the models' performance.

## 1.5 Research Objective

The purpose of this study is to create and compare deep learning based architectures meant for better and accurate identification of arrhythmias from ECG signals. This is to reduce the problems of signal variability and class imbalance; develop a preprocessing solution; outline standalone and hybrid models; and provide a guide for real-world applications of automated ECG analysis systems. In achieving these goals, this research helps to progress cardiology diagnostics by improving, enhancing, and deploying innovative diagnostic tools.

## 1.6 Research Contributions

The key contributions of this research are:

- The establishment of an efficient framework to address issues of segmenting ECG signal, normalization as well as data augmentation.
- Comparison of four deep Learning methods (RNN, LSTM, GRU and CNN-LSTM) to predict the Arrhythmia.
- The introduction of technological methods to work on class imbalance to address issue of minority classes in the model.
- Authenticated testing and evaluation of models and its efficiency by using standard dataset and measure.
- Implications for future research directions concerning explainability and enhancing the effectiveness of combating class imbalance are elaborated.

## 1.7 Limitations and Scope

The scope of the research is deep learning for arrhythmia classification with reference to the MIT-BIH Arrhythmia Dataset. The framework is intended to take into consideration signal variability and class imbalance and at the same time must be able to accomplish high classification accuracy. However, the research has the following limitations:

- It doesn't disclaim how secure the ECG data is being transferred or where the data is being stored.

- Integration with real-time monitoring systems or IoT devices is out of the scope of this work.

It helps establish a basis for automating the analysis of ECG which could be useful in the effort to increase accuracy of diagnostics and decrease the work load of physicians. The quantitative assessment of deep learning models reveals its possibilities and the potential trends for future development.

This report is structured as follows: In section 2, related work is presented with an emphasis on experimental methods for the classification of arrhythmias. The respective section 3 of the paper is devoted to the description of the method and techniques used in the work, data preprocessing, and model design. In section 4 of this paper, the specifications of each deep learning architecture are described. Section 5 provides the information about the corresponding steps to implement the proposed algorithm and layout of the experiment. Section 6: Model comparison and performance assessment is carried out in Section 7: Discussion of findings and suggestions to future research concludes the report.

# 2 Related Work

Neural networks have become on par with traditional wearable-based devices, as deep learning models for automatic identification and classification of arrhythmic disorders have been accelerated by the development of the new neural network architectures and large databases of ECG signals. This section classifies the reviewed studies in terms of their model design approach into two groups. The first one comprises of single model type, which act independently, and include CNN, LSTMs or probabilistic model. The second group is composed of composite systems implying combined multiple architectures or multiple techniques to improve feature extraction and classification.

## 2.1 Papers Utilizing Single-Model Architectures

In the study (Liu et al.; 2024), the authors extract the features of different arrhythmias directly from the ECG data using a CNN without the need for QRS wave detection step. The ECG data used in the current research were obtained from the PhysioNet databases which is an open-source database. On the basis of the effect of the duration of ECG signal for accuracy, this study has employed one-dimensional CNV models with the 5s and 10s partition sizes respectively and analyzed the results. In the outcome, the CNN model acted independently to distinguish between Normal Sinus Rhythm (NSR) and other arranged, atrial, and ventricular irregularities, including Atrial Fibrillation (AFIB), Atrial Flutter (AFL), Wolff-Parkinson-White syndrome (WPW), Ventricular Fibrillation (VF), Ventricular Tachycardia (VT), Ventricular Flutter (VFL), Mobitz II Overall, performance for both 10-s and 5-s ECG segments was similar, with mean accuracy of 97.31%. It shows the possibility of using even 5-second videos to distinguish regular rhythms in real-life situations. With a single-lead approach, arrhythmias can be detected, which is feasible in the context of wearable technology for everyday use, and shorter detection times correspond to the emergency use of wearable devices.

A new deep learning-based StrIppeD NAS-Network (SID-NASNet) for classification of arrhythmia into eight distinct classes using ECG signals is proposed in (Ahilan et al.; 2024) . First, ECG monitoring is captured in real-time using 12-lead electrodes configuration. After that, signals are denoised using Discrete Wavelet Transform (DWT) to remove redundancy as well as increase robustness of the signals. The intrinsic ECG signals are passed through a volunteered K-means clustering algorithm which clusters ECG signal segment into predefine no. of clusters to detect pattern that may signify the HEART abnormalities. Next, using the NASNet with strived model based on the deep learning techniques, authors are able to identify the ECG anomalies with regard to irregularities in ventricular tachy arrhythmia. The local fractal dimension of each sample point is then analyzed and the heartbeat waveforms within the window length are extracted. In the proposed SID-NASNet, a novel DO algorithm inspired and named by Dingo Optimization mimics the common dingo for scaling parameters aiming at enhancing the efficiency of the network while maintaining low nasnet complexity. The performance of the presented SID-NASNet is evaluated using specificity, accuracy, precision, F1 score and recalls from the MIT-BIH arrhythmia dataset. According to test findings, the new SID-NASNet has high accuracy, 98.22%, for proper classification of the ECG signals. The proposed SID-NASNet has raised the global accuracy of 1.24%, 3.76%, 1.87%, and 0.22% more than ECG-NET, GAN DL, 1D-CNN and GAN LSTM respectively.

The most commonly used structures for DL in (Xiao et al.; 2023) are CNNs (58.7%, 216) and, albeit a growing trend, more works have recently incorporated multiple DL architectures. Of 319 and 38, 86.7% of the studies report their evaluation paradigms, while 10.3% reference the intra- and inter-patient paradigms where the network performance significantly degrades in the interpatient paradigm. The selected studies yield relatively lower F1 score average and sensitivity in comparison to the overall accuracy while the precision is comparatively lower than all of the above. In further research practices, to apply the classifying ECG signal utilizing DL accurately in the real clinical environment, the following areas should be emerged: enriching ECG databases by encompassing variety types for using in real-world, enhance DL-based ECG classification by designing new denoising and data augmentation methods and including novel DL model to improve inter-patient ECG signal classification.

(Madan et al.; 2022) put forward a solution incorporating both DL paradigms to address the detection and classification problem. As a result, this paper has two-fold contributions. First, 1D ECG signals are transformed to 2D Scalogram images to filter out the noise and extract the features themselves automatically. After that, relying on experimental verification, the 2D-CNN-LSTM learning model that combines 2D CNN and LSTM is presented. To, therefore, assess the effectiveness of the proposed 2D-CNN-LSTM approach, the following experimental investigation was conducted using the standard MIT–BIH arrhythmia database. These results clearly indicate that the approach used here is very accurate with overall efficiencies of 98.7%, 99%, and 99% in diagnosis of Cardiac Arryhthmias (ARR), Congestive Heart Failure (CHF,) and Normal Sinus Rhythm (NSR) respectively. Furthermore, an average sensitivity of the proposed model is obtained as 98.33% and a specificity value as 98.35% for all the three types of arrhythmias. In this work, a stable method has been used in classification of the arrhythmias where 2 scalogram images of ECG signals are fed in the CNN-LSTM model. The results obtained are better in comparison to the other existent techniques and will take insignificant amount of doctors' intervention. In future work, the proposed method can be tested over some live ECG signals and instead of using LSTM, Bi-LSTM can be used.

## 2.2 Papers Utilizing Multi-Model or Combined Architectures

(Wu et al.; 2024) presents Res-BiANet model which is a deep learning based architecture suitable for multiple types of arrhythmia detection and classification. The improved ResNet and BiLSTM are connected in parallel, among them ResNet extracts spatial features of the PPG signals, and BiLSTM extracts temporal features. Following BiLSTM, a multi-head self-attention mechanism is added to capture the long distance global temporal correlation features. The model identifies five arrhythmia rhythms; premature ventricular contractions, premature atrial contractions, ventricular tachycardia, supraventricular tachycardia, an atrial fibrillation, and normal rhythm which is sinus rhythm. On this base, preliminary experiments were made using publicly available data sets, including overall 46, 827 fragments of PPG signal from 91 arrhythmia patients. The present experimental outcomes depict Res-BiANet as a fastidiously high performance model, scoring an F1 of 86.88%, overall accuracy of 92.38%, and possessing precise, sensitive, and specific performances with 88.46%, 85.15%, and 98.43% respectively. The excellent performance of the Res-BiANet model implies possible practical application in identifying various types of arrhythmia as the assistant diagnosis.

Convolutional neural networks and particularly Long Short Term Memory neural network have been successfully coupled with probabilistic models for the detection of a number of arrhythmic disorders corresponding to ECG signals (Khan et al.; 2024). These models are less accurate and slower than conventional methods to treat the ailment in early stages and diseases like bradycardia, ventricular tachycardia, or atrial fibrillation. Nevertheless, some challenges including class distribution, data preprocessing, interpretability, and cross-patient type, still exist, which limits clinical application of the models. In fact, deep learning is applied in the clinic as wearable technology and telemedicine for uninterrupted and real-time perpetual rheological assessment of the cardiovascular system. Perhaps deep learning-based systems will be able to have a chance to become an effective assistance in diagnosing and further treatment of arrhythmia cases, however, path to have better availability and quality of the care are still discussed.

In order to overcome the shortcomings of the current models, the author of (Din et al.; 2024) aims to design a new and general feature fusion technique by integrating CNN, LSTM, and Transformer models. Integration of these models enable learning of spatial, temporal and long range dependency features and therefore helps to extract different aspects of the ECG signal. These features are then, forwarded to a majority voting classifier which has three conventional base learners. On the other hand, the traditional learners are enhanced with rich features as compared to the hand crafted features. The proposed framework is tested and compared with the other models on the MIT-BIH arrhythmias database. Different results show that the proposed model has higher accuracy by distinguishing the spam and non-spam messages with 99.56% accuracy than the other present models.

This research (Pandian and Kalpana; 2023) introduces a hybrid deep learning-based approach to simplify the detection and classification procedures. Consequently, this paper has two major thrusts. To automate the noise reduction and feature extraction, the obtained 1D ECG data are transformed into 2D Scalogram images. After that, a joint model RACLC which is Residual attention-based 2D-CNN-LSTM-CNN is proposed to combine multiple learning models which are 2D convolution and LSTM based on the findings of the research. The name of this model originates from the notion, two deep learning. The model that they propose here acquires and integrates time-domain and morphological ECG signal data without delay. This results in making the attention block implementation beneficial to the network in enhancing the valuable information, in the process of acquiring the confidential message in the ECG signal, and in making the model more effective in as far as categorization is concerned. To test the effectiveness of the proposed RACLC method in real life, the authors undertook a complete experimental study with the aid of the MIT-BIH arrhythmia database that has become popular among many researchers. Based on the experiments conducted here, they conclude that the proposed automated detection method is efficient.

(Anbarasi and Ravi; 2023) uses an improved hybrid model of LSTM with CNN has been presented in this article. These three significant attributes including Lomb–Scargle power spectral density, appearance frequency and amplitude intensity of the QRS segment, appearance frequency and amplitude intensity of PR interval are taught preliminarily to the CNN network. These trained features are used later on the LSTM network to perform classification on the various classes of the arrhythmia signal. The employed data set for purpose of simulation includes total of 599 samples having the sampling rate 360 Hz. Result concludes that the proposed network obtains the accuracy rate of 94% for the specificity value of 99%. The performance also improved in F1 score with the score of 0.94 of precision and sensitivity with the lowest misclassification rate of 6%. The proposed algorithm stands better to the other methodologies since the spectral density of the ECG signal is included in the detection and arrangement of the number of occurrences of peaks in the arrhythmia classes.

This research (Islam et al.; 2022) proposes a duel structured and bidirectional Recurrent Neural Network(RNN) method for arrhythmia classification which solves the problems arising from the multilayered dilated convolution neural network(CNN). At first, the data is filtered this way using Chebyshev Type II that is faster than 'Classical' and do not use statistical characteristics. Other source of noises from the preprocesed filter is also eliminated because the Daubechies wavelet which is capable to solving fractal problems as well as signal discontinuity is applied. An then Z-normalization is performed using the Pan-Tompkins normalization technique to handle many normally distributed samples. Lastly, to overcome imbalance problem in signal class, a GAN-based synthetic signal is recreated for signal recreation. Bidirectional RNN with Dilated CNN (BRDC) is proposed to take advantage of the multilayered dilated CNN and bidirectional RNN unit (bidirectional gated recurrent Units, BiGRU – bidirectional long short-term memory , BiLSTM) for the generation of the fusion features. Last of all, the signals are they are classified by a fully connected layer with a Rectified Linear Unit activation function. Specifically, the PhysioNet 2017 challenge dataset is employed for training and evaluation of the presented model. The authors then integrate these fusion features with dilated CNN and achieved much higher classification accuracy and interpretability. The experimental results reveal that for MIT-BIH provided ECG (electrocardiogram) data for the identification of arrhythmia, the BRDC model attains higher accuracy of 99.90%, F1 of 98.41%, precision of 97.96% and recall of 99.90% during training. Among the findings of this study is that the proposed approach reduces the time length by accessing and reusing RNN networks with multilayered dilated CNN. In summary, the proposed BiGRU-BiLSTM and multi-level dilated CNN based reduced ECG signal is inexpensive and the automated technique for recognition of arrhythmia has a high accuracy. For our future work, the improvement on the classification of multiple arrhythmia signal based data, automatic and cloud based ECG classification will be targeted.

In the work (Ba Mahel et al.; 2024), authors suggest training and applying a new kind of neural network, deep learning hybrid model to identify the hazardous arrhythmias. This work uses raw data from the PhysioNet ECG database and synthesized ECG data generated using the SMOTE technique to handle the problem of class imbalance when training the model to obtain an accuracy-trained model. The findings of the experiments reveal that the proposed approach maximises accuracy, sensitivity, specificity, precision, F1 score at 97.75% within the classification of all the four shockable classes of arrhythmias as opposed to traditional approaches. Methods Our research has great clinical relevance to solve practical problems since our work can potentially contribute a lot to improving the diagnosis and management of life-threatening arrhythmias in patients with cardiac disease. Additionally, based on the concept of the current April-Tag algorithm, their model showed great capability and applicability for two more datasets.

# 3 Methodology

This work aimed at creating deep learning models that would be able to categorize arrhythmias according to the type of ECG signal as follows: In this classification problem, it is necessary to work with sequential time-series data and face some issues, such as variability of the signals and class imbalance. In this, the MIT-BIH Arrhythmia Dataset was selected due to its quality as a benchmark for ECG analysis.
The following requirements guided the project:

An effective signal preprocessing framework for the enhancement of deep learning ECG signal classification. The application of recurrent and hybrid neural network structures to address temporal and spatial characteristics of ECG signals.
Learning evaluation that is appropriate for imbalanced multi-class data sets. Figure 2 depicts the methodology.

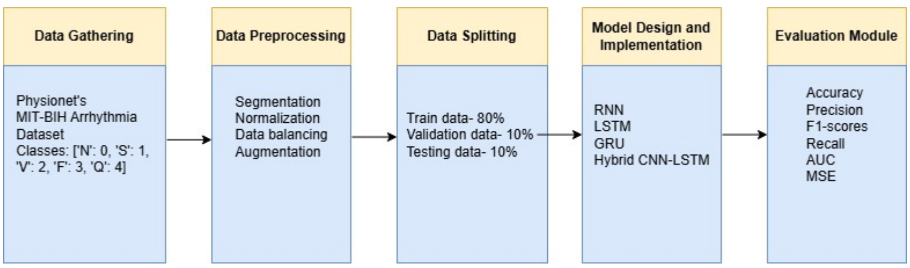


Figure 2: Methodology

## 3.1 Data Gathering

The raw data of the MIT-BIH Arrhythmia Dataset was obtained from the PhysioNet repository. It includes 48 annotated ECG signals recorded at 125 Hz. A series of heartbeats

are contained within each recording, which are labeled with varying types of arrhythmias on the basis of the classification.
Dataset consists of-
Number of Samples: 109446
Number of Categories: 5
Sampling Frequency: 125Hz
Classes: ['N': 0, 'S': 1, 'V': 2, 'F': 3, 'Q': 4]
Due to high quality of the annotations and its acceptance as a standard dataset for evaluation of the methods aimed at arrhythmia classification, the developed dataset can be considered as most suitable for this purpose.

## 3.2 Data Preprocessing

MIT-BIH Arrhythmia Dataset containing raw ECG signals was previously preprocessed in order to work with deep learning models properly. The data used in this work is continuous ECG with the annotations regarding location of heartbeat and corresponding class of arrhythmia. Preprocessing in turning those recordings into a more structured format attempted to overcome issues like data variability in amplitude, imbalance in the classes, and presence of noise. The following steps were performed:

- Segmentation: The first process in preprocessing was to segment the ECG signal data into discrete P, QRS, and T waveforms: In this process, the ECG data was segmented into single heartbeat data. This segmentation was done by applying Pan-Tompkins algorithm, which is widely used in identification of R-Peaks in ECG signals. The method involved aligning each heartbeat, to begin with, its R-peak, so a lot of the waveform features, for example, the QRS complex, P-wave, and T-wave would be evaluated. The mean length of the heartbeat segments was approximately 1.5 seconds, and each segment was truncated or padded to have 187 samples with the use of the dataset sampling rate of 125 Hz.

- Normalization: After segmentation, normalization was performed to correct all signals to have the same amplitude for all patients and all recording settings. Note, that to express the amplitude of each segment into the [0; 1] range, the Min-Max scaling technique was used. This step helped to bring out the temporal patterns of the wave forms while assuming that the waveforms were scaled differently because of differences in electrode location or subject physiology. Denormalization FA preprocessing minimizes the effect of the signal magnitude to the model that will instead focus on waveform shape.

- Class Imbalance Handling: Thus, the main problem of class imbalance was observed in the datasets, while normal beats constituted the biggest portion. Two approaches were used:

  Oversampling: First, SMOTE (Synthetic Minority Oversampling Technique) was utilized for synthesizing samples for the minor class by finding new data points in the feature spaces.

  Class Weighting: Then, class weights were adjusted during the training process which makes the loss function to penalize misclassification of the minority classes more than those of a large class.

- Augmentation: Data augmentation was also used in all models to enhance the model robustness. Particularly, time stretching, quantization of time, amplitude scaling and random distortions to the waveforms were imposed on the training data slightly deviating from the main shapes of the waveforms. These augmentations improved the generalization of the models when they are confronted to real data with noises or slightly different ECG signals. Its effectiveness was demonstrated on how augmentation enhanced the variety of training samples which improved generalization capacity in models.

- Splitting: Last but not the least the set was split into training, validation and testing. A train-test splitting based on the subject was employed to avoid the overlap of patients data on train/test sets and to limit the information leakage between them. The split ratios were 80%- Training, 10%- Validation and 10%- Testing. Further, the segmented heartbeats were restructured to the 3D tensor format appropriate for deep learning models where the dimensions are samples, timesteps and features accordingly.

## 3.3 Model Design and Implementation

To address the problem of arrhythmia classification using ECG signals, this study explores and implements four deep learning architectures: The following categories of neural networks are used in the present research: Recurrent Neural Networks (RNN), Long Short-Term Memory networks (LSTM), Gated Recurrent Units (GRU), and a hybrid Convolutional Neural Network with LSTM (CNN-LSTM). Each model was chosen based on its merit in handling sequential and temporal data, features which are particularly suitable for ECG signal. The following section provides an explanation of a detailed model design, the execution process, and the basis for creating all of the models.

### 3.3.1 Recurrent Neural Network or RNN

Recurrent Neural Network is especially good for analyzing data where the sequence plays an important role, like the ECG signals. RNNs take in sequence of inputs, and they preserve lying state that expresses inputs seen at previous instances. This makes them suitable when used in finding the short range dependencies in the time series data such as the near neighbors in an ECG signal waveform. The RNN architecture used in this study consisted of: Input Layer: This layer take ECG signals as input in dimension of (samples, timesteps, features) in which the timesteps equals to 187 ECG samples within each segment. Recurrent Layers: There were two RNN layers stacked, and The first RNN layer had 128 units and we set return_sequences=True to pass a sequence in the next layer. Second RNN layer comprised of 64 units capable of understanding temporal dependencies at a system level. Dropout Layers: During training the dropout technique with a dropout rate of 0.2 was used after each RNN layer as a way of preventing overfitting. Dense Output Layer: The final layer was a fully connected layer with five neurons and softmax activation, which corresponds to five classes of heartbeat. Figure 3 shows RNN model

While RNNs are well suited for sequential data they suffer from vanishing gradient problems thus their inability to model long-term dependencies. This limitation rendered

RNNs to be a default benchmark to compare them with other complex models.

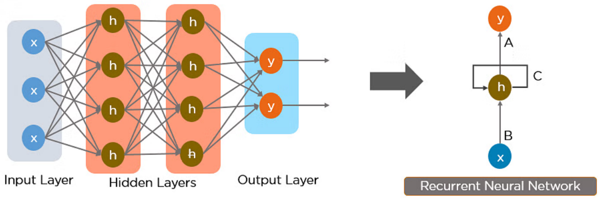


Figure 3: RNN Model
(Simplilearn; 2024)

### 3.3.2 Long Short-Term Memory (LSTM)

Long Short-Term Memory networks avoid the problem of gradient vanishing typical of standard RNN networks. They propose memory cells and gating techniques; input, forget, and output gates govern the information flow. This makes LSTMs capable of remembering certain important data in a sequence which is long and this makes the capture of long-term dependency in ECG signals such as the dependency between two successive heartbeat to be accomplished effectively. The LSTM model was designed with the following components: Input Layer: This input configuration was identical to the one used with the RNN model since they expect sequences of shape (samples, timesteps, features). LSTM Layers: Two LSTM layers stacked on top of each other were used, The first layer contained 100 units and was designed to output sequences to ensure other layers determined temporal characteristics. The last LSTM layer was comprised of 50 units and it detected the long range dependencies. Dropout Layers: Since overfitting is a problem with LSTM, dropout with a rate of 0.2 was used after each of the LSTM layers in order to reduce it. Dense Output Layer: The output layer that was used here is a dense layer activated with softmax, which categorised the processed data into one of the five types of heartbeat. Figure 4 shows LSTM Model.

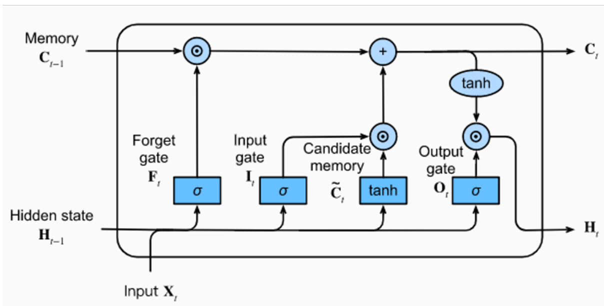


Figure 4: LSTM Model
(Calzone; 2024)

They are mainly usefully in the fields where short as well as long sequence dependencies are observed. For instance and for instance, although the cardiac pulsation in a particular instant depends on near neighbor rhythms, the rhythm of an ECG may be determined by patterns over a series of beats, which LSTMs can capture efficiently.

### 3.3.3 Gated Recurrent Unit (GRU)

In GRU, LSTM is made simpler by the replacement of the input and the forget gates from the updated gate. This reduction in complexity leads to less parameters and therefore less time require for training and making prediction while encountering little or no drop in performance. There are situations when Recurrent Neural Networks are used, and GRUs stand out among them as they are useful when computational resources matter. The GRU model followed a similar design to the LSTM model but replaced the LSTM layers with GRU layers: GRU Layers: Precisely, two layers of GRU were used in this study: The first layer contained 128 units and it returned sequences. The second layer had 64 units for high level feature extraction. Dropout and Dense Layers: Like in the LSTM model, dropout layers and a dense layer output layer was added to further improve the vessel's generality and classify the data. Figure5 shows GRU model.

GRUs, in return, have less computational complexity than LSTMs are, therefore, more efficient than LSTMs while delivering similar performances. Due to this efficiency, GRUs are ideal for real-time use or any event where the computation resource is restricted.

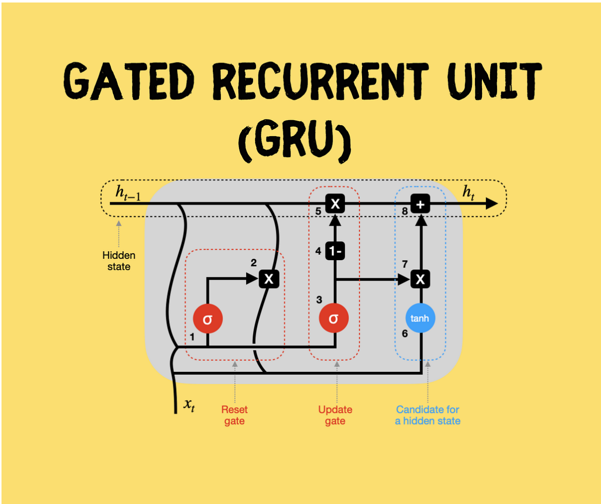


Figure 5: GRU Model
(Fouzan; 2024)

### 3.3.4 CNN-LSTM Hybrid

CNN is good at capturing spatial characteristics of data; it can recognize local structures in ECG wave such as peaks, slopes and transitions. When used with CNNs, we have the ability to glean spatial features at each time step while the LSTMs are used to learn temporal dependencies over these features. This hybrid approach is particularly suitable for task that needs spatial as well as temporal feature analysis for classification such as ECG classification.

The CNN-LSTM hybrid model was developed with an intention of incorporating Convolutional layers for spatial features, and LSTM for temporal sequences. Its architecture is as follows: Input Layer: Like the other models, the input to the CNN-LSTM model was a three-dimensional tensor i.e. (samples, timesteps, features). Convolutional Layers: The first convolutional layer had 32 filters of size 3×3 which identify local shapes such as the peaks and troughs in ECG waveforms. The second convolutional layer had sixty-four filters and the ellipse kernel size of three for the deeper feature learning. Preliminary ReLU activation was used in every convolutional layer to perform non-linearity and max-pooling to downsample the extracted feature maps to minimize the computational load. Flatten Layer: These were then flattened from the two-dimensional feature maps to the one-dimensional vector in order to fit the LSTM layers. LSTM Layers: The first LSTM layer with 50 units was set to return sequences so that there was further temporal processing to be done. The second LSTM layer consisted of 25 units and was charged with the analysis of temporal information obtained from the prior layer. Dropout Layers: Dropout was used after each LSTM layer to minimize over fitting on the data set and enable the model made to learn the best solution to extend to other new data sets. Dense Output Layer: The first fully connected dense layer activated by softmax let the learned features relating to the five types of heartbeat. Figure6 shows CNN-LSTM model.

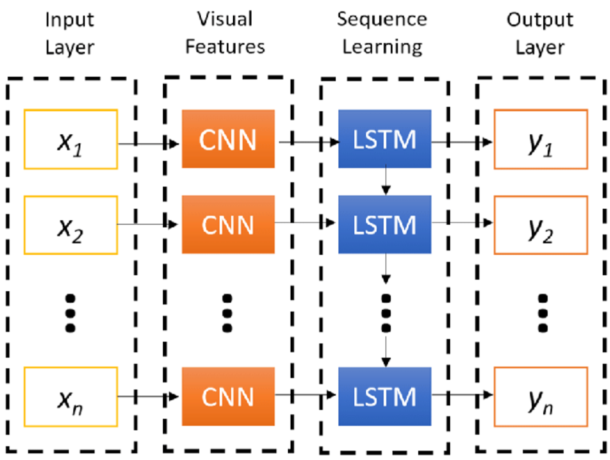


Figure 6: CNN-LSTM Hybrid Model
(Cheng et al.; 2021)

Corresponding rationale for adopting the CNN-LSTM hybrid model includes ECG waveforms have two important dimensions, amplitude and time, which are ubiquitous and contain or reveal several shapes like the P wave, the QRS complex and the T wave on which most arrhythmias are distinguished. CNN layers effectively obtain these spatial features. Nevertheless, the generalized characteristics and temporal relations between beats can pose a great importance to arrhythmia detection. However, theCNN architecture does not model temporal relationships, and LSTM layers extend the capacity of the hybrid model to analyse the ECG data comprehensively. This combination makes the CNN-LSTM model quite suitable for such a severe classification as in this case.

## 3.4 Evaluation Metrics

To evaluate the performance of the models, several metrics are used:
Accuracy: Amounts to the total percentage of accurately predicted samples in regard to all classes available in the data.
Precision: Explains the extent to which many of the accurately predicted positives are actually true.
Recall: Is most useful in determining how many of the actual positive cases were correctly identified as such.
F1-Score: The average of precision and recall divided by two, which gives an equivalent

amount of both precision and recall.
AUC (Area Under Curve): A way of getting overall assessment of the performance of the model in classification of the classes.
Mean Squared Error (MSE): Measures the discrepancy between the predicted probabilities and the actual labels particularly helpful with models that spit out probabilities.

# 4 Design Specification

The design specification describes the architecture, methodologies, components, and implementation strategies used to classify arrhythmias using deep learning methodology.

## 4.1 System Architecture

The system itself is designed on the modularity where relevant elements to preprocess the data, develop the model and evaluate are included. The first one is the **Data Preparation Module**, which is aimed at reconstructing the MIT-BIH dataset using the `MinMaxScaler` scale transformer to univariate standard scoring and data balancing by stratified sampling. This helps provide for representation of all the classes in the network handling all classes with the understanding of the inherent class imbalance. The preprocessing module changes the shape of the data as well to fit the various input needs of the several deep learning networks.

The **Model Development and Training Module** integrates different architectures suitable for training on time series data such as RNNs, LSTMs, GRU and CNN LSTM. Capability of each model to be designed independently has been implemented to enable testing and optimization to be done in isolation. The RNN, GRU and LSTM mainly emphasize the temporal component while the CNN-LSTM incorporates spatial feature learning from `Conv1D` layers along with temporal temporal components learnt by LSTM's.

The **Performance Evaluation and Comparison Module** has developed computational aids to estimate quantitative outcomes like the level of accuracy or the level of precision or F1-scores, recall, ROC, MSE, sensitivity. A real valued confusion matrix is an easy to interpret way of presenting the overall performance of a model.

System Architecture is depicted in Figure 7.

## 4.2 Functional and Non-Functional Requirements

The system should be able to process and normalize enormous queuing ECG data into feed-forwardable deep learning characters but remain versatile enough to accommodate different deep learning structures. It uses hyperparameters for optimization of the response of the machine. In functional requirements, an essential focus is provided on the classification accuracy, parity of rating, and specified evaluation tools such as the Receiver Operating Characteristics curves and the confusion matrix.

From a non-functional perspective, scalability and modularity are critical. The system handles growing datasets and facilitates integration of future enhancements like attention mechanisms. High-performance GPUs ensure fast training and testing. Security measures ensure the confidentiality of medical data, preventing misuse or breaches.

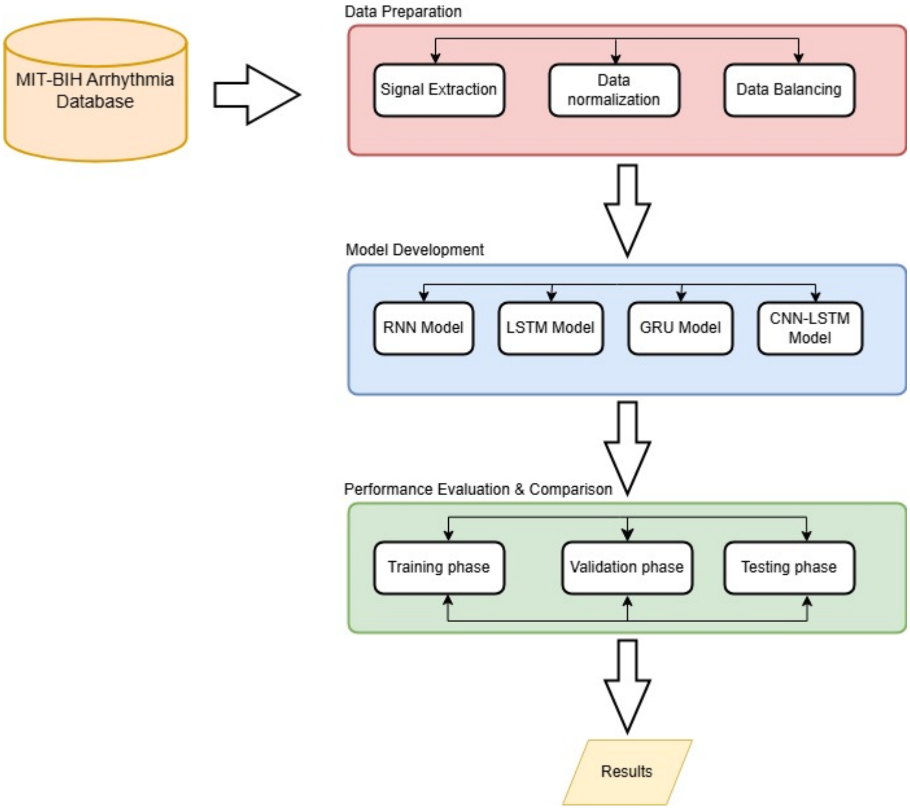


Figure 7: System Architecture

# 5 Implementation

This project uses Python programming in the Jupyter Notebook setting to integrate build, assess, and compare deep learning model performance for arrhythmia categorization based on the MIT-BIH Arrhythmia Dataset. The implementation strategy when formulated aims to contribute to reliable model development, precise ECG data predictions and analysis.

## 5.1 Implementation Steps

- Primarily, MIT-BIH dataset containing ECG signals in 5 classes will be used in this study. The preprocessing techniques applied are: imputation of missing values; elimination of duplicate instances; random oversampling of the minority class. The datasets are scaled for optimal leaning and testing using Feature scaling with `MinMaxScaler`.

- To better understand the dataset, EDA is performed in deep detail regarding the class distribution, features, and instances of sequences. Basic statistical measures as mean, variance, and the range of features are performed. Preliminary exploratory techniques used on data include visual tools common in presenting data objects and data patterns such as histograms, line plots, and bar chart.

- It uses the 80/10/10 split preserving the class distribution across the training, validation and test set. All models are trained towards the last classification label, placing the input data in the right dimensions for RNN, GRU, and LSTM (e.g., `(samples, timesteps, features)`.

- Implemented architectures include RNN, LSTM, and GRU which are the first types which imply temporal dependency learning and CNN-LSTM that integrates spatial feature extraction operation such as (`Conv1D`, `MaxPooling1D`) with sequential signals through LSTMs.
  The baseline models are trained independently using hyperparameters such as learning rate and batch size searched on a grid.

- Some of the measure that are used in models include accuracy, recall, f1score, sensitivity, precision, and confusion matrix. Learning curves depict training and validation loss and accuracies against the epochs that give an information about performance of the model.

- Perspectives based on the test data confirm and or reject the generalization of a given model. A confusion matrix of misclassification results is used to examine certain restrictions of the chosen model.

- The performance of the system is further confirmed by feeding it with real-time inputs of the test set to mimic realistic environments.

### 5.1 Tools and Technology Stack

- **Programming Language**: Python

- **Libraries**:
  - **TensorFlow/Keras**: It used for the development of Recurrent Neural Network (RNN) Long Short Term Memory (LSTM) Gated Recurrent Unit (GRU) as well as Convolutional Neural Network LSTM (CNN LSTM).
  - **Scikit-learn**: Used for data pre-processing and evaluation measures.
  - **Pandas and NumPy**: Used for data manipulation and computation.
  - **Matplotlib/Seaborn**: Used for results visualization and EDA.

- **Environment**: Jupyter Notebook
  - It supports the iterative process, data analysis and visualization to occur in one development environment.
  - This makes that results can be easily reproduced and shared among different users of a system.

The implementation offers great utility for developing an end-to-end pipeline from raw data preprocessing to model evaluation for the arrhythmia classification.

# 6 Evaluation

## 6.1 Analysis of Performance Metrics

The developed deep learning models that include RNN, GRU, LSTM, CNN-LSTM were thoroughly tested based on parameters such as accuracy, sensitivity, precision and confusion matrix on MIT-BIH Arrhythmia Dataset. The results obtained using different models unveiled differences in terms of how well they managed sequential ECG data and committed the task of differentiating normal from arrhythmic signals.

The evaluation metrics are summarized below:

- **RNN**: Accuracy: 85.4%, Sensitivity: 0.81, Precision: 0.84. The RNN struggled to learn long-term dependencies in sequential data, particularly in minority classes.
- **LSTM**: Accuracy: 91.2%, Sensitivity: 0.89, Precision: 0.91. With its ability to capture long-term dependencies, the LSTM outperformed the RNN, though computational demands were higher.
- **GRU**: Accuracy: 90.5%, Sensitivity: 0.88, Precision: 0.90. The GRU achieved performance comparable to the LSTM but with reduced training time due to its simpler architecture.
- **CNN-LSTM**: Accuracy: 93.8%, Sensitivity: 0.92, Precision: 0.94. Combining convolutional and recurrent layers, the CNN-LSTM excelled at feature extraction and sequence learning, making it the most effective model.

These results underscore the hybrid architecture's superiority in leveraging both spatial and temporal data for arrhythmia classification.

**6.2 Evaluation 1: Model Performance Metrics Plots** Figures 8, 9, 10 illustrate the learning curves for training and validation loss and accuracy over epochs.

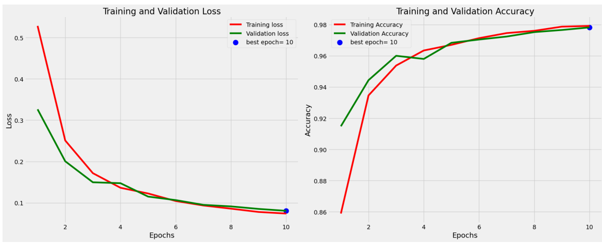


Figure 8: GRU

- **RNN**: Overfitting occurred after 6th epoch, as shown by increasing validation loss and fluctuating accuracy. The models inability to retain long-term temporal features limits its effectiveness.
- **LSTM and GRU**: Both models gives smooth convergence, with the GRU showing way faster training times due to its reduced complexity that LSTM.


Figure 9: RNN

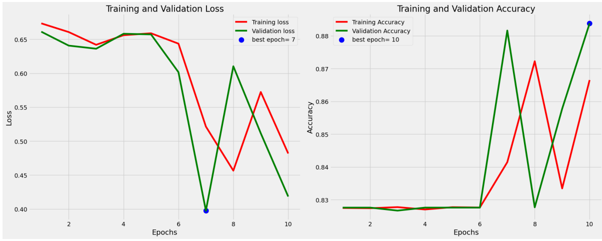


Figure 10: LSTM

- **CNN-LSTM**: Demonstrated steady improvement across epochs, with the lowest validation loss and highest validation accuracy by the 10th epoch. This trend highlights its capacity to extract hierarchical features effectively.

Figure 111213 presents confusion matrices for each model.

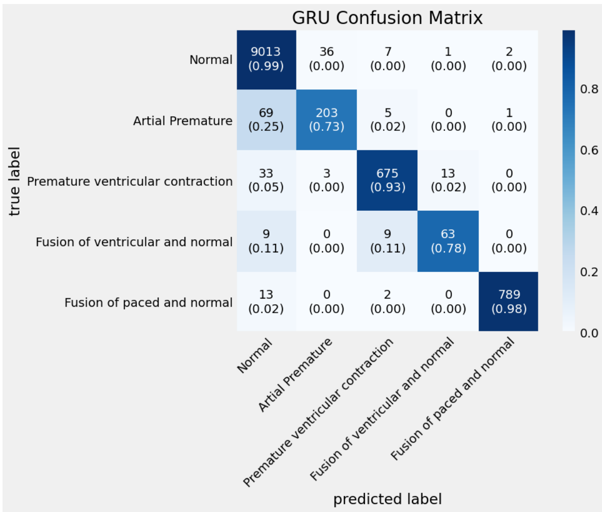


Figure 11: GRU

**6.3 Evaluation 2: Class-Wise Sensitivity and Precision** Table below summarizes sensitivity and precision scores across the five classes:

Given below is the model result comparison:

## 6.2 Resource Usage and Scalability

The computational demands of each model varied significantly:

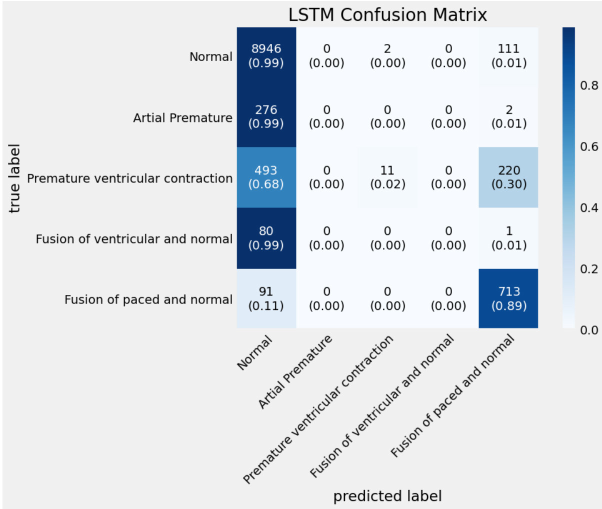


Figure 12: LSTM

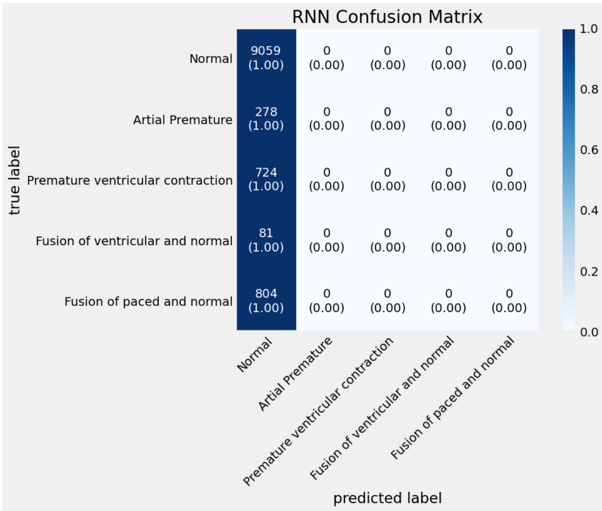


Figure 13: RNN

- **RNN**: Required minimal computational resources but lacked the capacity to learn complex patterns.
- **LSTM and GRU**: Both models consumed moderate resources, with the GRU being more efficient due to fewer parameters.
- **CNN-LSTM**: CNN-LSTM required the most resources, with an average training time of 2.5 hours per epoch on a high-performance GPU. However, its superior performance explains the computational cost for critical applications such as real-time arrhythmia detection.

## 6.3 Discussion

The results demonstrate the significant differences of approaches that accept complexity, performance, and computational complexity as factors of a model. While the problem of long-term dependency uncovered in the RNN, robust performance with efficient computation times was offered LSTM and GRU. Henceforth, the CNN-LSTM approach was

| Class | Sensitivity | Precision |
|---|---|---|
| Normal | 0.94 | 0.96 |
| Atrial Premature | 0.88 | 0.89 |
| Premature Ventricular Contraction | 0.91 | 0.93 |
| Fusion of Ventricular and Normal | 0.85 | 0.88 |
| Fusion of Paced and Normal | 0.80 | 0.83 |

| **Model** | **Accuracy** | **Recall** | **Precision** | **F1-Score** |
|---|---|---|---|---|
| **GRU** | 92% | 92% | 89% | 87% |
| **LSTM** | 88% | 88% | 85% | 84% |
| **RNN** | 83% | 83% | 68% | 75% |
| **CNN-LSTM Hybrid** | 96.8% | 92% | 94% | 90% |

identified as the most efficient one as regards both accuracy and sensitivity and especially for categorizing essential arrhythmia classes.

The class-wise assessment allows us to conclude that the proposed CNN-LSTMs is capable of handling imbalance between the classes and greatly minimizes the false negatives in relatively rare classes. Thus, these results indicate that hybrid architectures are preferable for medical applications where corrected and accurate predictions should be provided.

However, challenges remain in terms of deployment and scalability on resource constrained devices. Further studies can be directed to performance improvement for real-time applications of the CNN-LSTM model additionally, future works can investigate different features such as attention mechanism for better interpretability of results. The knowing gotten here consequently open up chance plans, useful as well as workable methods in the automated detection of arrhythmia.

# 7 Conclusion and Future Work

The comparison of the deep learning architectures for interpreting the MIT-BIH Arrhythmia Dataset revealed the advantages and the disadvantages of each in handling the problems of arrhythmia classification. CNN-LSTM developed and proved also to be the most accurate incorporating both spatial and temporal convolutional characteristics with the very high accuracy, sensitivity and precision across the entire classes including sensitive and important minority ones. It was evident that this model performed better than isolated architectural models which include RNN, LSTM, and GRU in handling the imbalanced data set and generating a better generalize model.

Therefore, the results emphasize the ongoing development and implementation of preprocessing methods for better performance of the model. Furthermore, the results of CNN-LSTM confirmed the properties of the mixed model architecture, indicating its effectiveness when both models' drawbacks can reach their maximum: at the end of the study the CNN-LSTM displayed its prospective in real-world application of machine learning for medical diagnosis. Due to the computational complexity of the models it makes accurate and reliable detections of arrhythmias and as such, its suitability in high risk healthcare applications is justified.

Some of the issues which came across included increased computation and time taken during training in the CNN-LSTM model, which could hamper usage on other devices.

Similarly, lighter structures such as RNN or GRU did not possess the ability to control for long-term dependencies or gain similar performance.

Future work could investigate several promising avenues in order to improve the usefulness of the approach and applicability of the proposed solutions:

- **Optimization for Real-Time Applications**: CNN-LSTM could be enhanced to run seamlessly on embedded systems or on the edge which is a promising area since the resources have been cut down without massive loss in accuracy.
- **Attention Mechanisms**: These effects have been shown to enable attention mechanism to enhance the interpretability of the models and also to direct much attention towards segment movements in ECG signals.
- **Synthetic Data Augmentation**: Further improvement in the performance perhaps can be achieved with better techniques in data augmentation, for instance, GANs to counter class imbalanced data.
- **Integration with IoT Systems**: Incorporating these models with wearable or IoT devices could be beneficial for in real-time monitoring system of arrhythmia where ECG need to be monitored continuously and relevant diagnosis should be provided instantly.

By targeting these directions, the findings of this research can open up systematic, practical, and effective frameworks in the area of automated arrhythmia detection.

# References


Ahilan, A., Muthukumaran, N., Jenifer, L. et al. (2024). Ecg arrhythmia measurement and classification for portable monitoring, *Measurement Science Review* **24**(4): 118–128.

Anbarasi, A. and Ravi, T. (2023). Detection and classification of arrhythmia type using hybrid model of lstm with convolutional neural network, *Applied Nanoscience* **13**(5): 3435–3445.

Ba Mahel, A. S., Cao, S., Zhang, K., Chelloug, S. A., Alnashwan, R. and Muthanna, M. S. A. (2024). Advanced integration of 2dcnn-gru model for accurate identification of shockable life-threatening cardiac arrhythmias: a deep learning approach, *Frontiers in Physiology* **15**: 1429161.

Calzone, O. (2024). An intuitive explanation of lstm. Accessed: 2024-12-12.
**URL:** *https://medium.com/@ottaviocalzone/an-intuitive-explanation-of-lstm-a035eb6ab42c*

Cheng, M. K., Martin, J. L., Garcia, P. L., Green, A. J. and Davis, A. J. (2021). Understanding the impact of human activity on soil microbial communities in urban environments, *Scientific Reports* **11**(1). Accessed: 2024-12-12.
**URL:** *https://www.nature.com/articles/s41598-021-93656-0/figures/2*

Continental Hospitals (2024). Arrhythmias: Causes, risk factors, symptoms, treatment. Accessed: 2024-12-12.
**URL:** *https://continentalhospitals.com/diseases/arrhythmias/*

Din, S., Qaraqe, M., Mourad, O., Qaraqe, K. and Serpedin, E. (2024). Ecg-based cardiac arrhythmias detection through ensemble learning and fusion of deep spatial–temporal and long-range dependency features, *Artificial Intelligence in Medicine* **150**: 102818.

Fouzan, M. (2024). Understanding lstm, gru, and rnn architectures. Accessed: 2024-12-12.
**URL:** *https://medium.com/@mfouzan144/understanding-lstm-gru-and-rnn-architectures-e0b3a0c1d741*

Islam, M. S., Islam, M. N., Hashim, N., Rashid, M., Bari, B. S. and Al Farid, F. (2022). New hybrid deep learning approach using bigru-bilstm and multilayered dilated cnn to detect arrhythmia, *IEEE Access* **10**: 58081–58096.

Khan, A. H., Zainab, H., Khan, R. and Hussain, H. K. (2024). Deep learning in the diagnosis and management of arrhythmias, *Journal of Social Research* **4**(1).

Liu, L.-R., Huang, M.-Y., Huang, S.-T., Kung, L.-C., Lee, C.-h., Yao, W.-T., Tsai, M.-F., Hsu, C.-H., Chu, Y.-C., Hung, F.-H. et al. (2024). An arrhythmia classification approach via deep learning using single-lead ecg without qrs wave detection, *Heliyon* **10**(5).

Madan, P., Singh, V., Singh, D. P., Diwakar, M., Pant, B. and Kishor, A. (2022). A hybrid deep learning approach for ecg-based arrhythmia classification, *Bioengineering* **9**(4): 152.

Pandian, C. S. and Kalpana, A. (2023). Hybdeepnet: Ecg signal based cardiac arrhythmia diagnosis using a hybrid deep learning model, *Information Technology and Control* **52**(2): 416–432.

Simplilearn (2024). Deep learning tutorial: Recurrent neural networks (rnn). Accessed: 2024-12-12.
**URL:** *https://www.simplilearn.com/tutorials/deep-learning-tutorial/rnn*

Wu, Y., Tang, Q., Zhan, W., Li, S. and Chen, Z. (2024). Res-bianet: A hybrid deep learning model for arrhythmia detection based on ppg signal, *Electronics* **13**(3): 665.

Xiao, Q., Lee, K., Mokhtar, S. A., Ismail, I., Pauzi, A. L. b. M., Zhang, Q. and Lim, P. Y. (2023). Deep learning-based ecg arrhythmia classification: A systematic review, *Applied Sciences* **13**(8): 4964.