Optimizing Electrocardiogram Analysis for Efficient Heart Condition Diagnosis

by

Ahmad Mousa

A thesis submitted to the
School of Graduate and Postdoctoral Studies in partial
fulfillment of the requirements for the degree of

Master of Applied Science in **Electrical and Computer Engineering**

Faculty of Electrical and Computer Engineering
University of Ontario Institute of Technology (Ontario Tech University)
Oshawa, Ontario, Canada
September 2023

Thesis Examination Information

Submitted by: **Ahmad Mousa**

Master of Applied Science in **Electrical and Computer Engineering**

Thesis title: 12 lead ecg classification with lead grouping

An oral defense of this thesis took place on September 15, 2023 in front of the following examining committee:

Examining Committee:

Chair of Examining Committee	Dr. Vijay Sood
Research Supervisor	Dr. Khalid Elgazzar
Examining Committee Member	Dr. Sanaa Alwidian
Thesis Examiner	Dr. Mohamed El-Darieby, Ontario Tech University

The above committee determined that the thesis is acceptable in form and content and that a satisfactory knowledge of the field covered by the thesis was demonstrated by the candidate during an oral examination. A signed copy of the Certificate of Approval is available from the School of Graduate and Postdoctoral Studies.

Abstract

This thesis introduces an innovative lead grouping strategy for efficient real-time Electrocardiography signal classification. This method uses a maximum of six leads instead of the traditional 12-lead approach, leading to significant reductions in sampling time (93.67%), data size at the data acquisition device (50%), and signal processing time (84.72%). Importantly, these benefits come with a minimal loss in accuracy (0.08%).

The thesis presents the CardioDiverse dataset, a publicly available resource that highlights key ECG leads associated with specific cardiovascular conditions. This resource can transform ECG-based diagnoses by focusing on the most pertinent leads.

The proposed lead grouping strategy has been successfully integrated with a real-time platform, demonstrating its practical robustness and applicability. This contribution brings a considerable change in the field of ECG analysis by providing an efficient and viable lead grouping method that balances accuracy and resource efficiency, marking significant advances in ECG analysis.

Keywords— ECG, standard 12 lead ECG signal, Multi-class, classification, lead group, Real-time platform, Kafka, lead grouping classification, single lead classification, 12 lead classification, CardioDiverse.

Author's Declaration

I hereby declare that this thesis consists of original work of which I have authored. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I authorize the University of Ontario Institute of Technology (Ontario Tech University) to lend this thesis to other institutions or individuals for the purpose of scholarly research. I further authorize the University of Ontario Institute of Technology (Ontario Tech University) to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research. I understand that my thesis will be made electronically available to the public.

_____ Ahmad Mousa

Statement of Contribution

The work described in this thesis has resulted in the following papers:

- **Selective Lead Integration: Enhancing ECG Classification for Effective Cardiac Monitoring [1]:** The article proposes a novel model for simultaneous multiclass classification of ECG signals using the lowest number and highly correlated leads.

- **Six Leads is All What You Need to Efficiently Diagnose Heart Diseases [2]:** This paper proposes a novel approach to classify ECG signals in real time using a maximum of six leads. The goal is to minimize the time and resources required for data acquisition and signal processing while maintaining high accuracy. The paper proposes a lead grouping technique and compares the performance of ECG signal classification using various leads.

- **CardioDiverse: A Comprehensive and Diverse Electrocardiogram (ECG) Dataset for Robust Heart Condition Analysis [under submissions]:** This study introduces a novel CardioDiverse data set that has been meticulously constructed to address the scarcity of balanced data in existing electrocardiogram (ECG) datasets. CardioDiverse offers a more balanced and rich compilation of 12-lead ECGs, catering specifically to various subclasses of cardiovascular disease (CVD). The dataset, accessible to the public via GitHub, has been used to analyze the association between the different ECG leads and CVD subclasses by integrating different data sets.

Acknowledgements

I would like to express my sincere gratitude to my supervisor, Dr. Khalid Elgazzar, for his invaluable guidance and support throughout the entire process of writing this thesis. His expertise and insights have been instrumental in shaping the research direction and methodology of this project.

I am also deeply grateful to my parents for their unwavering support and encouragement throughout my academic journey. Their love and belief in me have been a constant source of inspiration, and I would not have been able to accomplish this without them.

Writing this thesis has been a challenging part of my life and I am grateful to have had the guidance of Dr. Elgazzar and the support of my family. Their encouragement and motivation have been invaluable in helping me overcome obstacles and persevere in achieving my goals.

The authors also express their appreciation to the Engineering and Applied Science faculty at the University of Ontario Institute of Technology (Ontario Tech University) for allowing me to pursue this research and for their commitment to academic excellence.

Finally, I would like to acknowledge all those who have supported me along the way, including my colleagues, friends, and mentors. Your encouragement and support have been greatly appreciated, and I am grateful for the many ways in which you have contributed to my academic and personal growth.

List of Abbreviations

Table 1: Abbreviations and Acronyms

Abbreviation	**Description**
CVD	Cardiovascular Diseases
RA	Right Arm
LA	Left Arm
aVR	Augmented Vector Right
aVL	Augmented Vector Left
aVF	Augmented Vector Foot
MI	Myocardial Infarction
STTC	ST-T wave abnormality or T wave inversions
CD	Conduction Disturbance
HYP	Hypertrophy
LR	Logistic Regression
SVM	Support Vector Machines
DT	Decision Trees
RF	Random Forest
CNN	Convolutional Neural Networks
RNN	Recurrent Neural Networks
LSTM	Long Short-Term Memory
ANN	Artificial Neural Networks
GRU	Gated Recurrent Unit
BiLSTM	Bidirectional LSTM
FCN	Fully Convolutional Network

Abbreviation	Description
ResNet	Residual Networks
ROC	Receiver Operating Characteristic
AUC	Area under the ROC Curve
SPS	Sample Per Second
HRV	Heart Rate Variability
ReLU	Rectified Linear Activation Function
LGM	Lead-wise Grouping Method
RFE	Recursive Feature Elimination
AFIB	Atrial Fibrillation

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Introduction

Telemedicine, with telehealth technology at its core, has emerged as a revolutionary force in the modern healthcare landscape. By removing geographic barriers and placing patients at the center of care, telehealth technology, which encompasses real-time heart monitoring, is transforming the approach to healthcare delivery [3, 4]. Real-time heart monitoring, in particular, has been crucial in empowering healthcare professionals to detect and manage cardiovascular disease (CVDs) in a timely manner, one of the leading causes of death worldwide. Among these, Electrocardiogram (ECG) monitoring systems are extensively employed to diagnose a spectrum of heart diseases [5, 6]. However, despite advances, existing real-time heart monitoring systems reveal several challenges. For example, they often need a more nuanced understanding of the relationship between heart disease and the leads used in ECG readings. This, in turn, can reduce diagnostic accuracy and lead to the inefficient use of computational and communicational resources [7].

This dissertation tackles these challenges through a comprehensive analysis of ECG classification compatible with real-time monitoring. It proposes an innovative model

for simultaneous multi-class classification of ECG signals by leveraging a minimal set of highly correlated leads. Additionally, this research unveils a novel dataset'CardioDiverse' amalgamating multiple datasets to furnish a more comprehensive resource. The exploration of this dataset yields critical insights into the correlations between specific leads and disease subclasses, potentially advancing the early detection and diagnosis of CVDs [8]. This research seeks to amplify telehealth technology's potential in real-time heart monitoring and contribute positively to patient care in an increasingly digitalized world by proposing efficient and accurate alternatives to current systems.

1.2 Motivation

The motivation behind this research stems from the critical limitations in current real-time heart monitoring systems, particularly regarding ECGs. These limitations affect diagnostic precision and place great strain on resources, affecting remote monitoring effectiveness, data reduction, power consumption, and overall system sustainability [7]. This scenario underscores the urgency for more advanced and efficient systems that can ensure a timely and accurate diagnosis of CVD. Furthermore, there is a knowledge gap regarding the correlation between heart disease and the leads employed in ECG readings, which is crucial in enhancing diagnostic precision. Addressing this gap could optimize resource utilization, making care delivery more sustainable in the long run [9].

With these motivations in mind, this dissertation aims to augment the potential of telehealth technologies by offering an avant-garde model for the simultaneous multi-class classification of ECG signals using a minimal set of highly correlated leads. The insights derived from the novel dataset 'CardioDiverse' and the proposed classification model have the potential to redefine how healthcare providers detect and diagnose heart diseases. Through these advancements, this research embodies the aspiration to contribute substantively to the future trajectory of telehealth and to improve patient care and health

outcomes in the digital era.

1.3 Problem Statement

The challenges associated with real-time heart monitoring through Electrocardiogram (ECG) systems are manifold and greatly impact the effectiveness, utilization of resources, and efficiency of these systems [10]. One of the main challenges lies in the precision of detection and diagnosis; This is due to the inadequate understanding of the relationship between heart diseases and the ECG leads, which leads to compromised diagnostic accuracy. Furthermore, a significant challenge to ECG systems is resource utilization. They often consume excessive computational and communication resources, making them unsustainable in resource-constrained environments. Moreover, the efficiency of classification is another obstacle. Current ECG signal classification techniques need to be optimized for efficiency, largely due to their reliance on all 12 ECG leads.

The knowledge gap in the correlation between specific ECG leads and disease subclasses poses a significant challenge. Gaining insight into this correlation is critical to successful diagnosis and treatment, yet this area still needs to be explored. Lastly, the lack of comprehensive datasets further compounds the problem. Current datasets should provide adequate information to facilitate research on the correlation between ECG leads and heart disease subclasses.

To address these challenges, this thesis proposes a novel ECG classification model and introduces a comprehensive data set. The ultimate objective of these efforts is to enhance the effectiveness of real-time heart monitoring.

1.4 Research Questions

1. What is the relationship between leads and diseases?

2. Can effective ECG signal classification be achieved with fewer leads?

3. Does the quantity of leads impact the resources (response time and storage) in real-time?

1.5 Thesis Contribution

This thesis makes the following contributions:

- **Developing a Lead Grouping Strategy**: The core innovation of this thesis is the creation of a lead grouping technique that employs a maximum of six leads for real-time ECG signal classification. This groundbreaking strategy optimizes the balance between efficiency and accuracy.

- **Reduction in Latency and Resource Usage**: Extensive experimentation and performance evaluation of the proposed method have demonstrated a significant reduction in sampling time (93.67%), data size at the acquisition device (50%), and signal processing time (84.72%). These significant gains are achieved with only a negligible 0.08% decrease in accuracy compared to the traditional 12-lead approach.

- **Introducing the CardioDivers Dataset**: This thesis introduces a new public resource, the CardioDiverse dataset, which is a result of meticulous correlation analyses. This dataset identifies key ECG leads that are linked with distinct cardiovascular conditions, potentially revolutionizing ECG-based diagnoses by focusing on the most relevant leads.

- **Developing End-to-end Real-Time Platform**: A pivotal part of this thesis is the successful integration of the proposed lead grouping strategy with a Kafka-based real-time platform. This not only demonstrates the theoretical merit of the strategy, but also showcases its practical robustness and applicability in real-world.

- **Enhancing ECG Interpretation and Application**: Overall, the efficient and practically viable lead grouping method presented here marks a significant advance in the interpretation and application of ECGs. Its successful integration with a real-time platform underscores this substantial contribution to the field.

1.6 Thesis Organization

This thesis is organized as follows:

- **Chapter One**: Introduction to telehealth technology, real-time heart monitoring, challenges, and contributions of this work.

- **Chapter Two**: This chapter dives deep into the theoretical foundations of the subject matter, beginning with an in-depth examination of the principles of Electrocardiograms. It discusses the existing ECG signal processing and interpretation techniques, with a particular focus on limb and chest leads. The chapter also critically reviews the existing literature, discussing the various machine learning and deep learning methods previously used for ECG classification. It concludes by highlighting the gaps in current knowledge and practices the thesis aims to fill.

- **Chapter Three**: The proposed lead grouping method is thoroughly explained in this chapter. It delves into the rationale behind the method, its design, and its application in lead-disease analysis. The development and performance evaluation of a Recursive Feature Elimination Method tailored for ECG classification is also presented. The chapter also explains the integration of the proposed model with a Kafka-based real-time monitoring platform, thereby demonstrating the practical applicability of the theoretical model.

- **Chapter Four**: This chapter details the data collection and analysis process, starting with a comprehensive survey of available 12-lead ECG datasets. It identifies

the limitations and gaps in these datasets and introduces the CardioDiverse dataset as a solution to address these gaps. The methodology used to create, process, and analyze the CardioDiverse dataset is explained, providing a roadmap for similar future endeavors.

- **Chapter Five**: The outcomes of the research are presented in this chapter. It starts with an examination of the superclasses and subclasses of heart conditions, with the findings from these experiments detailed. The chapter proceeds to the main experiment: the test of a six-lead diagnostic approach. The impacts of the number of leads on real-time processing, storage, and accuracy are meticulously analyzed, setting the stage for the presentation of the successful integration of the lead grouping method with a Kafka backend. An evaluation of the end-to-end system underlines the practical feasibility of the proposed model.

- **Chapter six**: The final chapter wraps up the thesis by summarizing the key findings, contributions, and implications of the study. It emphasizes the innovative nature of the proposed lead grouping method and the impact it has on the field of ECG analysis. The chapter also proposes potential directions for future research, suggesting how the results of the current study can lead to further exploration.

Chapter 2

Background and Literature Review

This chapter delves into the extensive literature concerning cardiac monitoring and classification facilitated by telehealth technologies. The focal point is to investigate the complex relationships between electrocardiogram (ECG) lead placements and cardiac disease diagnosis, alongside the concurrent classification of multiple diseases. Concurrently, we strive to broaden our understanding of the interconnections among various diseases, as outlined in [11, 12]. The core objective is to conduct a meticulous review of cutting-edge deep learning techniques employed within the domain of telehealth [13], with an emphasis on those leveraging variable lead counts. This exploration aims to outline contemporary methodologies and pinpoint existing research lacunae that merit further exploration.

The chapter commences by providing essential background information on the electrocardiogram (ECG), an instrumental diagnostic tool ubiquitously utilized for heart function assessment, thereby setting the stage for subsequent discussions on heart disease detection and real-time monitoring. We expound on how advanced deep learning algorithms are integrated into these processes, heralding significant potential in augmenting the effectiveness of cardiac monitoring and disease classification [14]. Following this, we engage in an extensive review of the prevailing literature, highlighting studies that

are closely aligned with our research objectives. This scholarly expedition grants a profound understanding of the current state of telehealth technologies. Serving a dual role, it presents an overview of the progress in real-time cardiac monitoring and situates our research within the wider ambit of this rapidly burgeoning field.

This exhaustive review aspires to be a pivotal reference, offering insights into the potential and challenges associated with incorporating sophisticated machine learning algorithms into telehealth cardiac monitoring. Additionally, by identifying the existing gaps, this chapter seeks to pave the way for pioneering research endeavors in the domain.

2.1 Electrocardiogram

The electrocardiogram (ECG) is a non-invasive medical procedure that captures the heart's electrical activity over time [15]. This activity is attributed to electrical impulses traversing through the heart, instigating heartbeats. The ECG is characterized by various components, each denoting a specific phase of the cardiac cycle [16].

An archetypal ECG trace encompasses a P wave, QRS complex, and T wave. The P wave corresponds to atrial depolarization, attributed to the electrical signal dispersion from the sinoatrial node across the atria, inducing contraction. The QRS complex signifies ventricular depolarization, conducting to the heart's principal contraction. The T wave, in contrast, denotes ventricular repolarization or recovery [16, 17].

Electrocardiograms (ECGs) are indispensably employed by cardiologists as a potent diagnostic tool for a myriad of cardiac conditions. For instance, an elongated QRS complex may suggest ventricular tachycardia [19], while an elevated or depressed ST segment could be indicative of acute conditions such as myocardial infarction or ischemia [20]. Through detailed analysis of the amplitude, morphology, and temporal properties of each waveform within the ECG, significant inferences can be made regarding the heart's electrical activity and overall functionality.

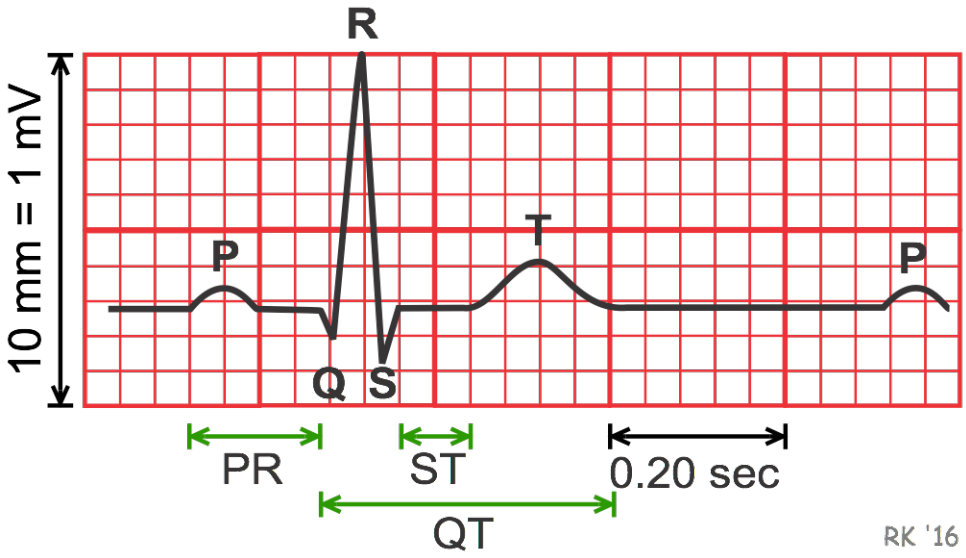


Figure 2.1: A Typical ECG Waveform [18].

The ECG signals are procured utilizing a specialized apparatus, termed as the ECG machine. To garner the cardiac electrical signals, electrodes serve as conduits connecting the patient to the ECG machine. Ten electrodes are judiciously placed: four are attached to the limbs, and six are positioned on the chest. These electrodes, operating as transducers, are highly sensitive to the electrical impulses generated by the heart. Upon interception, the ECG machine amplifies and meticulously records these impulses, culminating in an elaborate and invaluable depiction of the heart's electrical activity.

Figure 2.2 illustrates the twelve-lead system incorporated in the electrocardiogram (ECG). Each lead within this system furnishes a unique perspective of the heart's electrical activity, amalgamating to a holistic representation of cardiac function. These leads are engendered from various permutations of the ten electrodes, which are systematically arranged on the patient's torso, conforming to standardized protocols. This illustration elucidates the interplay between the leads and electrodes, capturing the heart's electrical complexities and augmenting the diagnostic acumen of medical practitioners.

Grasping the attributes and classifications of data accrued through the electrocardiogram (ECG) is quintessential. The data, forming the bedrock of the diagnostic process,

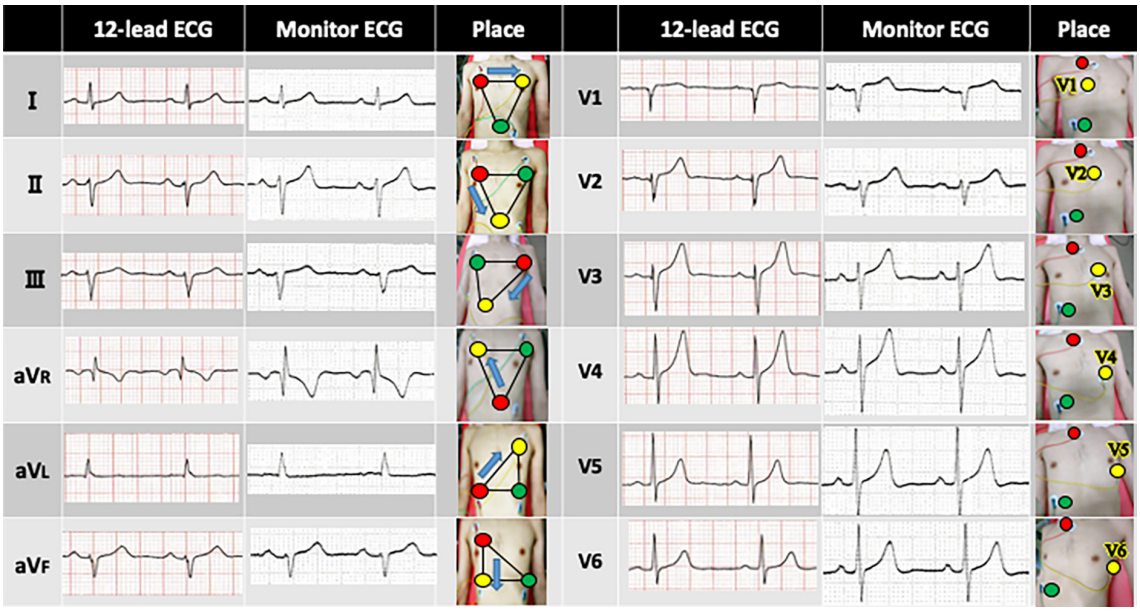


Figure 2.2: Canonical Arrangement of ECG Electrodes Within The Twelve-Lead System [1] .

can be dichotomized into two principal categories predicated on the placement of electrodes relative to the patient's anatomy: limb leads and chest (precordial) leads.

2.1.1 Limb Leads

The limb leads constitute a vital component of the electrocardiogram (ECG) examination, offering valuable insights into the heart's electrical activity [21]. Figure 2.3 illustrates that the ECG electrodes are strategically placed on the patient's body. The limb leads are derived from electrodes that are affixed to the right arm (RA), left arm (LA), and left leg (LL). These three electrodes are essential for obtaining Leads I, II, and III:

These leads, specifically Lead I, Lead II, and Lead III, are derived from differential measurements between specific pairs of limb electrodes, as follows:

$$\text{Lead I} = LA - RA,$$

$$\text{Lead II} = LL - RA,$$

$$\text{Lead III} = LL - LA.$$

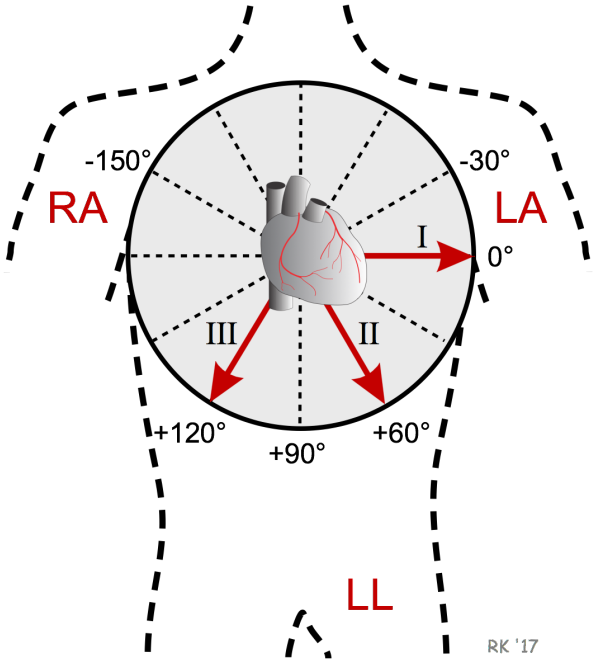


Figure 2.3: Standard Placement of ECG Electrodes [18].

These limb leads offer a frontal plane depiction of the heart's electrical activity, with each lead contributing a unique perspective [22]:

- **Lead I:** Provides an observation of the heart's electrical activity from the vantage point of the right arm, looking towards the left arm. Deviations in Lead I may suggest potential complications in the left lateral section of the heart.
- **Lead II:** Tracks the heart's electrical activity from the right arm towards the left leg, effectively mirroring the natural direction of the heart's electrical conduction. Lead II delivers a holistic view of the heart's rhythmic activity.
- **Lead III:** Offers a viewpoint of the heart from the left arm towards the left leg, yielding data regarding the lower left section of the heart.

Alongside these basic limb leads, the augmented limb leads (aVR, aVL, aVF) supply additional frontal plane views from varied angles. These leads, which are derived from the limb electrodes, encompass [23]:

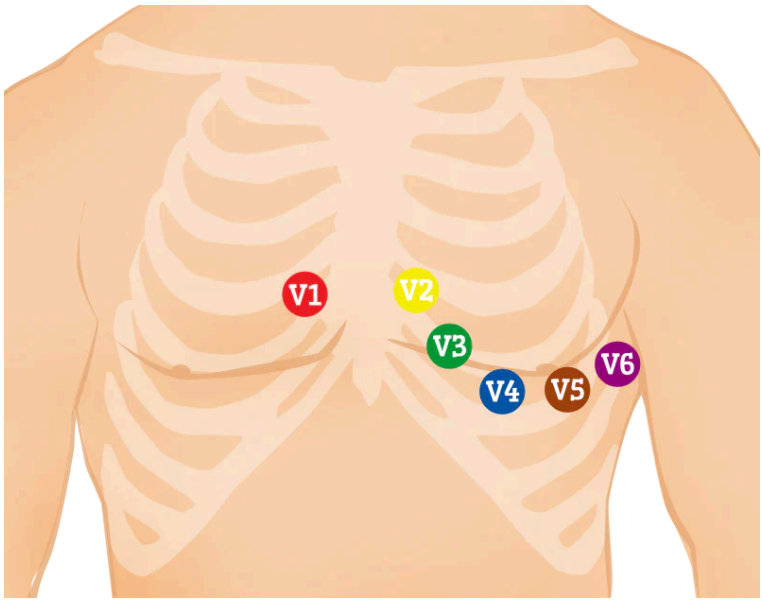


Figure 2.4: Standard Placement of ECG Chest Lead Electrodes [2].

- **aVR:** The Augmented Vector Right lead, providing a view from the heart's right aspect. While not as frequently utilized in diagnoses, aVR can be instrumental in identifying specific types of arrhythmias or ischemia.

- **aVL:** The Augmented Vector Left lead, offering a viewpoint from the heart's left side. aVL proves particularly beneficial in detecting alterations in the heart's lateral wall.

- **aVF:** The Augmented Vector Foot lead, capturing the heart from the lower end or foot. It assists in the evaluation of changes in the heart's inferior wall.

2.1.2 Chest Leads

The precordial, or chest leads (V1 to V6), offer a horizontal plane perspective of the heart. These leads are derived directly from the corresponding chest electrodes [24]: As depicted in Figure 2.4, the precordial leads are positioned on the chest to capture the heart's electrical activity in the horizontal plane. These leads are essential for assessing the performance and health of different sections of the heart:

Table 2.1: Comparison Between Chest Leads and Limb Leads in ECG

Aspect	**Chest Leads (Precordial Leads)**	**Limb Leads**
Placement	Placed on the chest around the heart.	Placed on the limbs (arms and legs).
Plane	Capture in horizontal plane.	Capture in the frontal plane.
Number of leads	6 leads (V1 to V6).	4 leads (I, II, III, and aVR, aVL, aVF).
Significance	Crucial for diagnosing conditions affecting the ventricles.	Useful for a broader view.
Types of views	focusing on the heart's anterior, lateral, and septal regions.	focusing on the inferior and superior areas of the heart.

- V1 and V2: These leads monitor the right ventricle and septum of the heart. They can be instrumental in identifying conditions such as right ventricular hypertrophy and certain types of bundle branch block.
- V3 and V4: These leads observe the anterior wall of the left ventricle, yielding valuable data concerning conditions like anterior wall myocardial infarction.
- V5 and V6: These leads provide views of the left lateral section of the heart, enabling the identification of conditions like left ventricular hypertrophy.

In an Electrocardiogram (ECG), both chest leads and limb leads are vital in evaluating the heart's electrical activity. As summarized in Table 2.1, chest leads are positioned on the chest and focus on the heart's horizontal plane, particularly around the ventricles. In contrast, limb leads are attached to the limbs, providing insight into the heart's electrical activity in the frontal plane. Chest leads are especially crucial for diagnosing conditions that affect the ventricles and are represented by six leads (V1 to V6), whereas limb leads, which are also six (I, II, III, aVR, aVL, and aVF), give a broader view of the heart's activity. Each set of leads plays a complementary role, and together, they offer a comprehensive evaluation of the heart's electrical function [25].

There are Various ECG types, each suited for distinct purposes, are available. The most common types encompass the resting ECG, stress test ECG, Holter monitor, and

event monitor [26]. These ECGs are utilized in a wide array of scenarios, from conducting routine health check-ups to monitoring the progression of heart disease, detecting sporadic symptoms, and recording infrequent and unpredictable symptoms. Each type of ECG has unique strengths and applicability, facilitating comprehensive cardiac evaluations and continuous monitoring as necessary. A deeper understanding of these ECG types and their applications enhances the capabilities of clinicians to diagnose and manage various heart conditions effectively and timely [27].

2.2 Processing of Electrocardiogram Signals

The process of electrocardiogram (ECG) signal processing constitutes an indispensable component of scrutinizing the heart's electrical behavior, pinpointing irregularities, and formulating precise diagnoses. This section will delve into the multifaceted aspects of ECG signal processing, encompassing the calculation of sample frequency, ascertaining the minimum length of the signal requisite for valid classification, alongside a diverse set of methodologies for filtering, extracting features, and segmentation.

2.2.1 Sampling Frequency and Signal Duration

Sampling frequency in the context of an ECG signal pertains to the count of data points amassed per unit of time. A superior sampling frequency culminates in a more comprehensive resolution and a more authentic depiction of the ECG signal. The Nyquist-Shannon Sampling Theorem delineates the minimum sampling frequency requisite for an ECG signal, which propounds that the sampling frequency should be no less than double the highest frequency component present in the signal [28]. Considering that the highest frequency component for ECG signals ordinarily hovers around 100 Hz, a sampling frequency of at least 200 Hz is recommended. However, for heightened resolution, ECG signals are frequently sampled at 500 Hz, or in some instances, 1000 Hz [29].

The minimum duration of an ECG signal requisite for valid classification is contingent upon the specific use-case scenario and the classifier employed. Generally, the duration of a single cardiac cycle, which approximates around 1 second, should suffice for fundamental rhythm classification tasks. Nonetheless, for more intricate classification tasks, such as discerning specific morphological characteristics or detecting transient events, a longer signal segment may be necessitated [30].

2.2.2 Filtering Techniques for Noise Reduction in ECG Signals

The analysis of electrocardiogram (ECG) signals is often complicated by the presence of noise and various forms of interference or artifacts, such as baseline wander, power line interference, muscle contractions, and electromagnetic radiation from surrounding equipment. Therefore, filtering techniques become essential for retaining the clinically relevant aspects of the ECG signal while minimizing the impact of unwanted noise.

Several common filtering techniques used in ECG signal processing include:

- **Low-pass filters:** These filters, also known as high-cut filters, permit the passage of low-frequency components of the signal while attenuating the higher frequency components [31]. They are beneficial for the removal of high-frequency noise that might arise from sources such as muscle contractions or interference from electrode motion. In the context of ECG signals, the use of a low-pass filter can help retain the vital components of the signal which are typically in the lower frequency range.

- **High-pass filters:** High-pass or low-cut filters allow the passage of high-frequency components while reducing the influence of low-frequency components [32]. These filters find their utility in mitigating the effects of baseline wander, a low-frequency artifact commonly instigated by factors such as patient respiration or movement.

- **Band-pass filters:** Band-pass filters permit a specific range of frequencies to pass while attenuating components outside of this range [33]. These filters are

particularly advantageous when the aim is to concurrently eliminate high-frequency noise and baseline wander, offering a more comprehensive solution.

- **Notch filters:** These filters are designed to target and attenuate a specific narrow frequency band, leaving the remainder of the frequency spectrum unaffected [34]. For instance, notch filters are commonly utilized to eliminate the 50 or 60 Hz power line interference that can significantly corrupt ECG signals.

In modern digital signal processing, these filters can be implemented efficiently using software tools. Python, for instance, offers libraries like SciPy and NumPy, which contain functions that can be used to design and apply these filters to ECG signals, thus facilitating the removal of unwanted noise while preserving the essential features of the signals [35].

2.2.3 Feature Extraction from ECG Signals

Feature extraction constitutes a pivotal phase in the analysis of electrocardiogram (ECG) signals, facilitating the identification and extraction of relevant characteristics or features from the ECG signal that can be instrumental in classification or diagnostic tasks. These features may be based on time-domain characteristics, frequency-domain properties, or a combination of both, providing a multifaceted perspective on the heart's electrical activity [36].

There are numerous potential ECG features to consider, including but not limited to:

- **QRS complex duration:** The QRS complex duration, signifying ventricular depolarization, can yield insights into the functionality of the heart's conduction system. Abnormalities in the QRS complex duration can help pinpoint medical conditions such as bundle branch block or ventricular hypertrophy [37].

- **Amplitude of the QRS complex:** The QRS complex's amplitude can be an important feature, as it may indicate conditions like left ventricular hypertrophy,

where the amplitude is typically increased, or pericardial effusion, where it may be decreased [38].

- **T-wave morphology:** The shape, duration, and amplitude of the T wave can be indicative of various cardiac conditions. Abnormalities such as peaked, inverted, or biphasic T waves can be a sign of hyperkalemia, myocardial ischemia, or ventricular hypertrophy, respectively [39].

- **RR interval:** This refers to the temporal interval between successive R peaks, which corresponds to the duration of a single cardiac cycle. Variations in the RR interval can serve as an indicator of autonomic nervous system activity. Disorders such as atrial fibrillation or heart block can often be inferred from changes in RR intervals [40].

- **QT interval:** The QT interval, spanning from the inception of the QRS complex to the conclusion of the T wave, represents the duration of both ventricular depolarization and repolarization. Abnormal QT intervals, whether prolonged or shortened, can be symptomatic of an elevated risk of ventricular arrhythmias or sudden cardiac death [41].

- **P-wave duration and amplitude:** These attributes of the P wave, indicative of atrial depolarization, can offer valuable information about the size and functional capacity of the atria. Deviations from normal P-wave duration or amplitude might be suggestive of conditions such as atrial enlargement or atrial conduction disorders [42].

- **ST-segment deviation:** Any deviation of the ST segment from the baseline can be symptomatic of ischemia, injury, or infarction. Notably, elevated or depressed ST segments can help diagnose medical conditions like myocardial ischemia, myocardial infarction, or pericarditis [43].

In the contemporary landscape of digital signal processing, these features can be extracted with remarkable accuracy using software tools and machine learning algorithms, paving the way for automated analysis of ECG signals.

2.3 ECG Classification

Electrocardiogram (ECG) classification is a critical and complex task that entails identifying and categorizing various cardiac conditions or abnormalities based on the ECG signals obtained from a patient. This process involves recognizing and categorizing different patterns emerging from the heart's electrical activity. These patterns or rhythms can be either typical, indicative of healthy heart function, or atypical, suggesting potential abnormalities or conditions.

The significance of accurate ECG classification cannot be overstated, as it plays a vital role in diagnosing and monitoring various heart conditions. Abnormalities in ECG patterns are often associated with a wide range of heart conditions, necessitating a comprehensive understanding of these patterns for accurate classification. for example :

- **Myocardial Infarction (MI):** MI, often referred to as a heart attack, involves the death of heart muscle cells due to reduced blood flow to the heart. This reduced blood flow often results from the blockage of coronary arteries. ECG changes seen in MI typically include pathologic Q-waves, ST-segment elevations, and T-wave inversions. These changes may vary depending on the location and extent of the infarction. Rapid and accurate detection of these patterns in ECG signals is vital as it aids in prompt MI diagnosis and treatment, potentially preventing further myocardial damage and improving survival rates [44].

- **ST-T wave abnormality or T wave inversions (STTC):** Abnormalities in the ST-T wave can suggest various heart conditions, including ischemic heart disease, electrolyte imbalances, or left ventricular hypertrophy. In the context of ischemia,

ST-segment depression or T-wave inversions may indicate myocardial ischemia, a precursor to MI. Accurate recognition of these abnormalities can help identify at-risk patients and initiate appropriate medical interventions [45].

- **Conduction Disturbance (CD):** CDs involve abnormal electrical conduction within the heart, potentially conducting to irregular heart rhythms. These disturbances can be seen on an ECG as changes in the P wave, QRS complex, or PR interval. For instance, a widened QRS complex may suggest a bundle branch block, while a prolonged PR interval could be indicative of a first-degree heart block. Proper classification of CDs is crucial as they may indicate underlying structural heart diseases or increased risk for adverse cardiac events [46].

- **Hypertrophy (HYP):** Hypertrophy refers to the thickening of heart muscle, which can occur in response to increased workload on the heart, such as in hypertension or valvular heart diseases. In an ECG, left ventricular hypertrophy often presents with increased QRS complex amplitude, ST-T changes, and leftward shift of the QRS axis, while right ventricular hypertrophy may cause right axis deviation and prominent R waves in right precordial leads. Accurate classification of hypertrophy can help identify patients with conditions necessitating further workup and management [44].

Thus, each cardiac condition has unique ECG patterns that can be identified and classified. Modern computational and machine learning tools have made significant strides in aiding this classification process, enhancing diagnostic accuracy and improving patient outcomes.

2.3.1 ECG Classification via classical machine learning

ECG classification using classical machine learning methods involves a two-step process: feature extraction and classification. In the feature extraction step, significant and dis-

criminative features are extracted from the ECG signals, while in the classification step, these features are used as input to machine learning algorithms to classify different cardiac conditions. However, feature extraction can be challenging, making ECG monitoring complex.

- **Logistic Regression (LR):** Logistic regression is a statistical method for analyzing a dataset in which the dependent variable is binary. In the context of ECG classification, the logistic regression model predicts the probability of a cardiac condition or abnormality based on the extracted features [47]. The logistic function, or sigmoid function, is given by 2.1 [47] :

$$P(y = 1|x) = \frac{1}{1 + e^{-(\beta_0+\beta_1 x_1+\cdots+\beta_n x_n)}} \tag{2.1}$$

- **Support Vector Machines (SVM):** SVM is a supervised machine learning algorithm that constructs a hyperplane or a set of hyperplanes in a high-dimensional space to separate different classes of data points [48]. SVM aims to maximize the margin between the classes while minimizing the classification error. The classification function for an SVM is given by 2.2 [48]:

$$f(x) = \text{sign}(\sum_{i=1}^{N} \alpha_i y_i K(x, x_i) + b) \tag{2.2}$$

  where $K(x, x_i)$ is the kernel function, α_i are the Lagrange multipliers, and b is the bias term.

- **Decision Trees (DT):** Decision trees are a non-parametric method for ECG classification that recursively partitions the feature space into regions based on the feature values [49]. Each internal node of the tree represents a decision on a feature, while each leaf node represents the predicted class label. The decision tree algorithm aims to minimize the classification error or a related cost function.

- **Random Forest (RF):** Random forest is an ensemble learning method that constructs multiple decision trees and combines their predictions to improve the classification accuracy and prevent overfitting [50]. The majority vote of individual decision trees determines the final class label. A comparison of these classical machine learning methods for ECG classification is presented in Table 2.2.

Table 2.2: Comparison of Classical Machine Learning Methods For ECG Classification Regarding Accuracy, Sensitivity, and Specificity.

Method	Accuracy	Sensitivity	Specificity
Logistic Regression	0.91	0.87	0.93
Support Vector Machines	0.94	0.89	0.95
Decision Trees	0.88	0.85	0.90
Random Forest	0.95	0.92	0.96

The table 2.2 above demonstrates the performance of various classical machine learning methods in ECG classification [51, 52, 53, 54]. It is important to note that the choice of method depends on the specific problem and the quality of the extracted features. However, the reliance on feature extraction can make ECG monitoring a difficult task, as the quality and selection of features significantly impact the performance of these methods. Consequently, researchers have started exploring deep learning approaches, which can automatically learn features from raw ECG signals, alleviating the need for manual feature extraction and improving classification performance.

- **The Need for Transitioning from Machine Learning to Deep Learning in ECG Classification**

  Traditional machine learning methodologies have been extensively utilized in ECG classification because they can model intricate relationships between input features and output classes. Nonetheless, these techniques heavily rely on manual feature extraction, which calls for domain expertise and can be labor-intensive. Furthermore, the performance of these conventional machine learning algorithms is largely

contingent on the quality and selection of the extracted features. This dependence makes it a challenging task to attain high classification accuracy [55] consistently.

In the light of recent advancements, deep learning has surfaced as a potent alternative to traditional machine learning techniques, particularly in ECG classification. Deep learning models, which include convolutional neural networks (CNNs), recurrent neural networks (RNNs), and long short-term memory (LSTM) networks, have showcased their proficiency in learning hierarchical feature representations directly from raw ECG signals [56]. This property obviates the necessity for manual feature extraction and potentially leads to enhanced classification performance [57].

Furthermore, deep learning models can be optimized for specific tasks and adapted to many ECG classification problems, such as arrhythmia detection, myocardial infarction identification, and heart failure prediction [58]. This adaptability makes deep learning models more suitable for handling the inherent variability and complexity of ECG signals, which can be influenced by factors such as age, gender, medical history, and signal quality [59].

The advent and advancements of deep learning technologies have streamlined the ECG classification process and significantly enhanced its accuracy and efficiency.

As these technologies continue to evolve, they are anticipated to play a pivotal role in diagnosing and monitoring cardiac conditions, thereby facilitating better patient care and outcomes.

2.3.2 ECG Classification via Deep Learning

Electrocardiogram (ECG) classification via Deep Learning has seen considerable advancements in recent years, with various types emerging based on the nature of the neural networks and the architecture used.

- **Artificial Neural Networks in ECG Classification**

Artificial Neural Networks (ANNs) are a powerful deep learning approach used for ECG classification. They are composed of layers of nodes, often referred to as "neurons," interconnected in a manner reminiscent of the human brain's structure.

ANNs rose to popularity in the 1990s and have since been recognized for their effectiveness in handling both categorical and numerical predictions. Essentially, ANNs are adept at learning patterns from data, making them akin to smoothing algorithms. Moreover, they also have similarities with regression algorithms as they aim to establish the relationship between inputs and outputs based on cross-sectional data [60].

As depicted in Figure 2.5, a typical ANN consists of an input layer, hidden layers, and an output layer [61]. Edges connect neurons in these layers, each carrying a weight. During computation, the ANN uses these weights to calculate neuron values, which are then passed through an activation function. This function maps the values from the input layer to the output layer. Due to their ability to model complex non-linear relationships, ANNs excel at sophisticated tasks like ECG classification, where capturing subtle patterns and relationships is crucial [62].

Neurons within these layers are connected via edges associated with specific weights. The network computes neuron values based on their associated weights in an ANN, feeding these values into an activation function. This activation function then maps the collective values from the input to the output layer [60].

The primary strength of ANNs lies in their ability to capture complex non-linear relationships. This makes them particularly suitable for intricate tasks like ECG classification, where identifying and understanding the relationships within the data is paramount to accurate classification [61]. The raw ECG signals are utilized as inputs into the network for ECG classification. The network subsequently learns the complex, non-linear relationships existing between the ECG signal and the re-

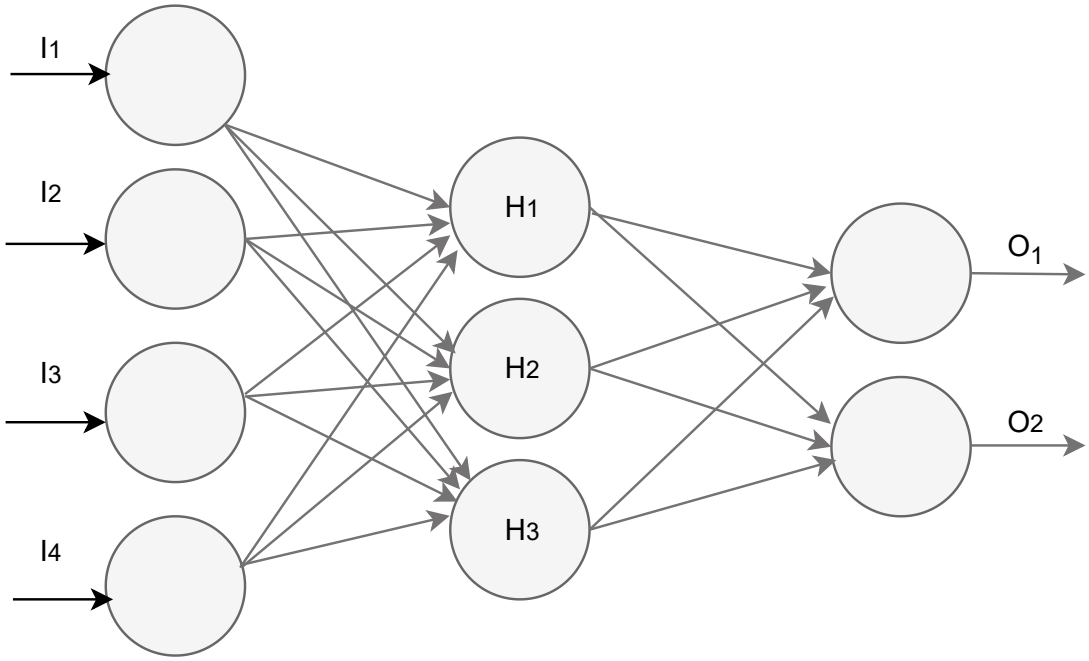


Figure 2.5: Schematic Representation of an Artificial Neural Network (ANN) With Input Layer (I), Hidden Layer (H), and Output Layer (O).

spective class labels. This learning process comprises a stage known as forward propagation, wherein the signals are transmitted through numerous layers of the network. Each layer elevates the data into a higher level of abstraction. Mathematically, this transformation is depicted by 2.3 [61]:

$$h_i^{(l)} = f\left(\sum_{j=1}^{N} w_{ij}^{(l)} h_j^{(l-1)} + b_i^{(l)}\right) \tag{2.3}$$

Here, $h_i^{(l)}$ denotes the activation of the i-th neuron in the l-th layer, $w_{ij}^{(l)}$ is the weight linking the i-th neuron in the $(l-1)$-th layer to the j-th neuron in the l-th layer, $b_i^{(l)}$ is the bias term corresponding to the i-th neuron in the l-th layer, N represents the total number of neurons in the $(l-1)$-th layer, and f is an activation function.

Post prediction, the discrepancy between the predicted and the actual class labels is computed. This discrepancy is subsequently propagated backwards through the

network in a phase called backpropagation, and the network tweaks its parameters in order to minimize this discrepancy. This adjustment can be represented as 2.4 [61]:

$$w_{ij}^{(l)} := w_{ij}^{(l)} - \alpha \frac{\partial}{\partial w_{ij}^{(l)}} J(w) \tag{2.4}$$

In this equation, $J(w)$ is the cost function, α symbolizes the learning rate, and $w_{ij}^{(l)}$ is the weight being updated.

- **Convolutional Neural Networks (CNNs)**

  Convolutional Neural Networks (CNNs) are a specialized type of ANN that have been successfully employed in ECG classification tasks, showcasing remarkable results. The architecture of CNNs is distinctive for its incorporation of convolutional layers, which enable the network to identify local patterns within the input data [63].

  The CNN model is adept at automatically extracting features from raw ECG signals, eliminating manual feature extraction. This ability to automate feature extraction saves valuable time and resources and potentially enhances classification performance [63].

  Figure 2.6 presents an example of a CNN architecture for ECG classification. The architecture typically consists of multiple layers: convolutional, pooling, and fully connected. These layers collaborate to learn hierarchical features from raw ECG signals, allowing the network to classify signals into distinct classes [63].

  The convolutional layer is a key component of CNNs and detects local features in the input data. A convolution operation is performed in this layer, which can be mathematically represented as 2.5 [63]:

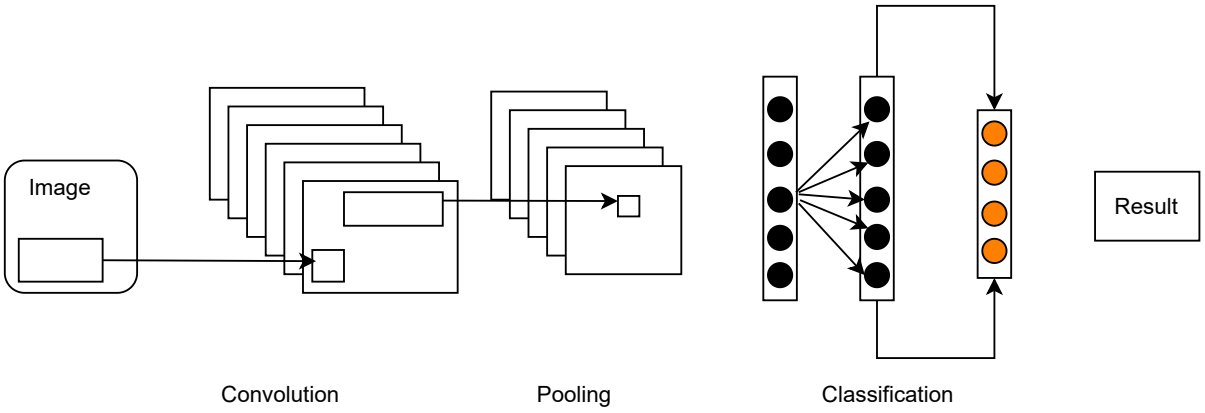


Figure 2.6: Convolutional Neural Network (CNN) Architecture.

$$y_i = \sum_{j=1}^{k} w_j x_{i+j-1} + b \tag{2.5}$$

where y_i is the output feature map, x_{i+j-1} is the input data, w_j is the kernel or filter weight, k is the kernel size, and b is the bias term.

Pooling layers reduce the spatial dimensions of the feature maps, thereby reducing computational complexity and enabling the network to capture translational invariance. The most common pooling operation is max pooling, which selects the maximum value within a specified window.

Fully connected layers, on the other hand, combine the learned features and generate the final output for classification. The output of the fully connected layers can be computed as 2.6 [63]:

$$o_i = f\left(\sum_{j=1}^{N} w_{ij} h_j + b_i\right) \tag{2.6}$$

where o_i is the output of the i-th neuron, h_j is the input from the previous layer, w_{ij} is the weight connecting the i-th neuron to the j-th neuron, b_i is the bias term for the i-th neuron, N is the total number of neurons in the previous layer, and f

is an activation function.

Two primary types of CNNs are used for ECG classification: 1D CNNs and 2D CNNs. 1D CNNs are designed to work with one-dimensional data, such as time-series signals like ECGs. In contrast, 2D CNNs are designed to work with two-dimensional data, such as images or spectrograms derived from ECG signals.

- 2D CNN for ECG Classification: Two-dimensional CNNs (2D CNNs) have been leveraged for ECG classification tasks, wherein they utilize two-dimensional convolutional layers to discern local patterns in input data. The input data could either be 2D images or spectrograms derived from the ECG signals [64]. The convolutional layers operate by sliding a two-dimensional kernel or filter over the input data, thereby identifying features critical to the classification task. A typical 2D CNN's architecture is composed of various layers, including 2D convolutional layers, pooling layers, and fully connected layers. These layers collaboratively learn hierarchical features from the input data, facilitating the network's ability to classify signals into distinct classes [65].
- 1D CNN for ECG Classification 1D CNNs for ECG classification use one-dimensional convolutional layers to learn local patterns in the ECG signals. These layers slide a one-dimensional kernel or filter across the input data, detecting features relevant to the classification task. The architecture of a 1D CNN typically includes multiple layers, such as 1D convolutional layers, pooling layers, and fully connected layers. These layers work together to learn hierarchical features from raw ECG signals, enabling the network to classify the signals into different classes [66].

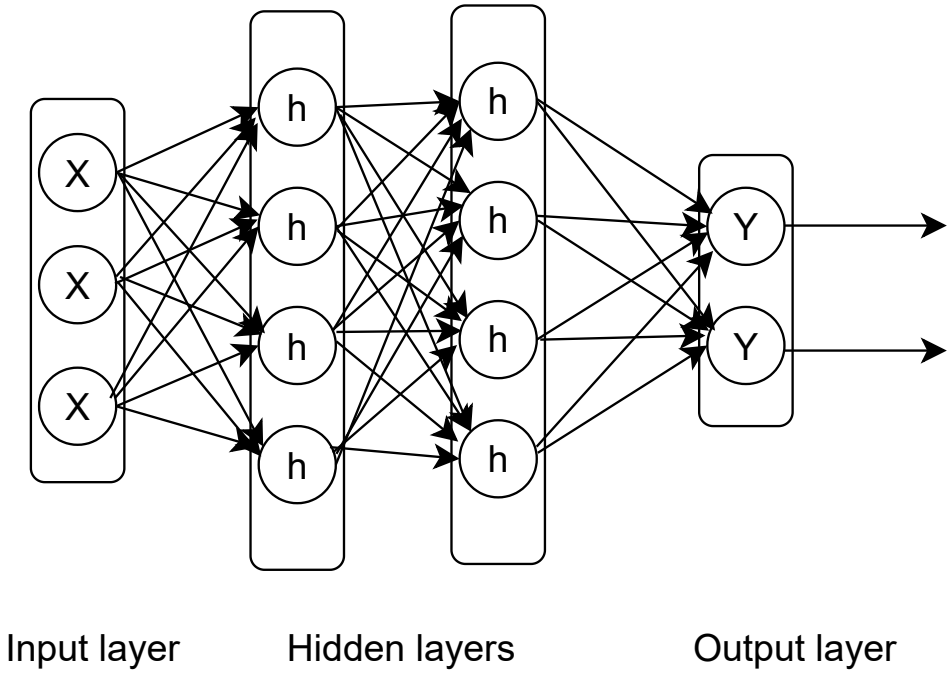
Deep learning models such as Artificial Neural Networks (ANNs), One-Dimensional Convolutional Neural Networks (1D CNNs), and Two-Dimensional Convolutional Neural Networks (2D CNNs) have shown significant promise in ECG classifica-

tion, outperforming traditional machine learning techniques in several aspects. They have the capability to learn features autonomously from raw data, manage large-scale datasets, and support end-to-end learning, conducting to a significant enhancement in the accuracy and robustness of ECG classification systems [67][65][66]. However, while these models have exhibited superior performance, they may still need help dealing with sequential data. Both ANNs and CNNs might need help with capturing the temporal dependencies in the ECG signals, which is a critical factor in accurate classification. Furthermore, the size and complexity of CNNs can also make them computationally expensive and difficult to interpret. Therefore, despite the advantages of ANNs, 1D CNNs, and 2D CNNs, a gap needs to be addressed - the effective analysis of time-dependent data inherent in ECG signals. This opens up the opportunity to explore other deep learning architectures, such as Recurrent Neural Networks (RNNs), which are specifically designed to handle sequential data and might offer a solution to this problem.

- **Recurrent Neural Networks (RNN) for ECG Classification** Recurrent Neural Networks (RNNs) [68] are a variant of artificial neural networks designed specifically to handle sequential data by capturing temporal dependencies. This characteristic renders RNNs particularly effective for ECG classification tasks, given their capacity to model the time-dependent nature of ECG signals.

  Figure 2.7 showcases an example of RNN architecture for ECG classification. This architecture generally comprises of an input layer, recurrent layers, and an output layer. Collaboratively, these layers process raw ECG signals, effectively capturing temporal dependencies and enabling the network to distinguish between different classes of signals [69].

  One of the crucial components of RNNs is the recurrent layer, which encapsulates temporal information from the input data. This layer utilizes a recurrent unit that



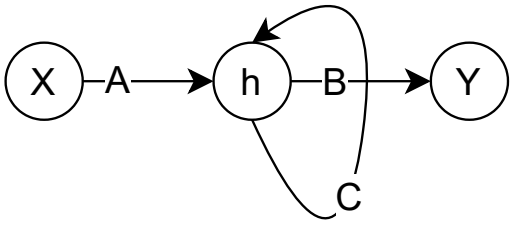


Figure 2.7: Recurrent Neural Network (RNN) Architecture.

consistently maintains and updates hidden states over time. The output from this recurrent layer at a given time step, denoted as t, can be represented mathematically as 2.7 [69]:

$$h_t = f(W_{xh}x_t + W_{hh}h_{t-1} + b) \tag{2.7}$$

where h_t is the hidden state at time step t, x_t is the input at time step t, W_{xh} and W_{hh} are the input-to-hidden and hidden-to-hidden weight matrices, respectively, b is the bias term, and f is an activation function.

The output layer in an RNN is responsible for generating the final output for classification. The output of the output layer can be computed as 2.8 [69]:

$$o_i = f\left(\sum_{j=1}^{N} w_{ij} h_j + b_i\right) \tag{2.8}$$

where o_i is the output of the i-th neuron, h_j is the input from the previous layer, w_{ij} is the weight connecting the i-th neuron to the j-th neuron, b_i is the bias term for the i-th neuron, N is the total number of neurons in the previous layer, and f is an activation function.

RNNs are effective for ECG classification, as they can model the time-varying nature of ECG signals and capture the complex temporal dependencies between different signal segments. However, traditional RNNs can suffer from vanishing or exploding gradient problems, making it challenging to learn long-range dependencies. To overcome these issues, advanced RNN architectures, such as Long Short-Term Memory (LSTM) networks and Gated Recurrent Unit (GRU) networks, have been proposed and demonstrated improved performance in ECG classification tasks [70].

- **Long Short-Term Memory (LSTM) Networks for ECG Classification**

  Long Short-Term Memory (LSTM) networks represent a specialized version of RNNs designed to overcome a central limitation of conventional RNNs: the difficulty in learning long-term dependencies due to the infamous vanishing gradient problem. This issue refers to the propensity for gradients to become exceedingly small during the backpropagation process, which can effectively stall the network's learning [71] [72].

  LSTMs address this problem by integrating a memory cell capable of preserving information for extended periods. This feature makes them exceptionally capable

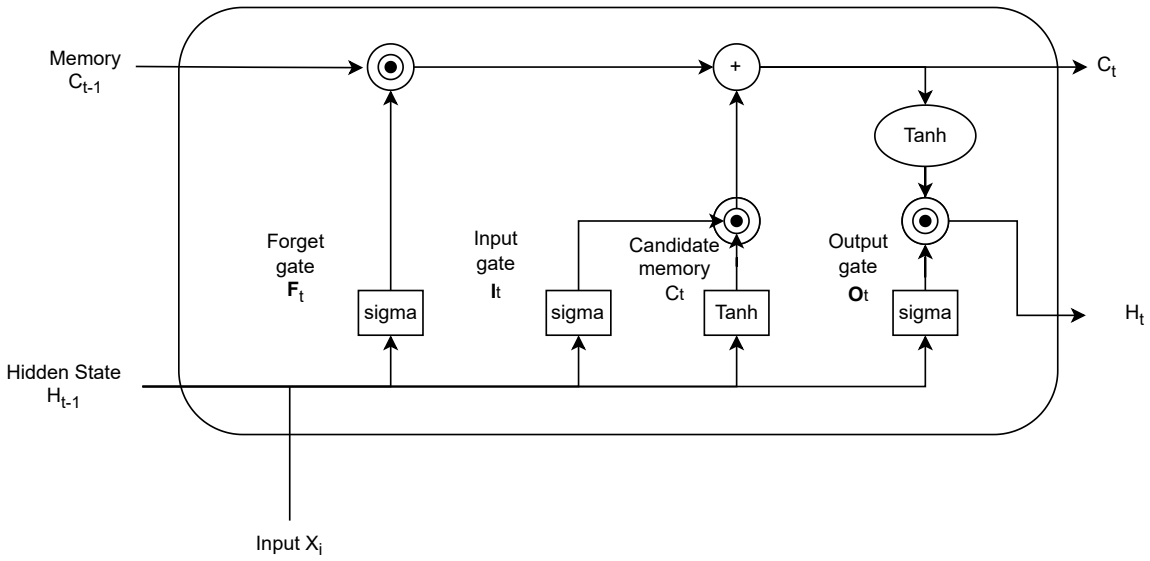


Figure 2.8: A Long Short-Term Memory (LSTM) Network Architecture.

of learning from data where substantial time gaps distance critical information [71]. Within the context of ECG classification, an LSTM network can input a series of ECG signals and make predictions grounded in both the current signal and preceding signals it has processed, akin to a traditional RNN. However, due to the integrated memory cell, it can also recall vital information from the distant past, which may be instrumental for accurate predictions [67].

The nucleus of an LSTM network is the LSTM unit, composed of a memory cell and three gates, namely the input gate, forget gate, and output gate. These gates regulate the memory cell's influx, internal, and efflux of information. The following equations dictate the behavior of an LSTM unit 2.9 - 2.13 [67]:

$$f_t = \sigma_g(W_f[h_{t-1}, x_t] + b_f) \tag{2.9}$$

$$i_t = \sigma_g(W_i[h_{t-1}, x_t] + b_i) \tag{2.10}$$

$$o_t = \sigma_g(W_o[h_{t-1}, x_t] + b_o) \tag{2.11}$$

$$c_t = f_t * c_{t-1} + i_t * \sigma_c(W_c[h_{t-1}, x_t] + b_c) \tag{2.12}$$

$$h_t = o_t * \sigma_h(c_t) \tag{2.13}$$

In these equations, x_t represents the input at time t, h_{t-1} signifies the output from the previous time step, c_t denotes the memory cell state at time t, and f_t, i_t, and o_t are the forget, input, and output gates respectively, which control the information flow within the memory cell. W and b represent weights and biases, and σ_g, σ_c, and σ_h are the activation functions (typically sigmoid for the gates and hyperbolic tangent for the cell state and output) [67].

The capacity of the LSTM's memory cell to recall long-term dependencies allows it to capture critical information from the distant past within ECG signals, which is often crucial for accurate ECG classification [67]. For instance, if a patient's heart rhythm exhibits a periodic abnormality, an LSTM network can detect this pattern and leverage it for classification, even if it isn't present in the most recent ECG signals. This capability, coupled with LSTM's resilience to noise and variability in ECG signals, renders them an effective tool for ECG classification [67].

While the LSTM network is proficient at capturing temporal dependencies in ECG signals, it is inherently designed to learn from past information to predict future outcomes. However, future information can also supply valuable context for interpreting the past in many clinical situations. This observation led to the development of the Bidirectional LSTM (BiLSTM), an extension of the traditional LSTM. The BiLSTM introduces backward information passed from future states to augment the learning process, providing the network with a broader contextual understanding for accurate classification [73].

- **Bidirectional LSTM for ECG Classification**

  Bidirectional LSTMs (BiLSTMs) are an extension of the standard LSTM architecture, designed to improve the model's ability to capture information from past and future time steps in a sequence. BiLSTMs consist of two separate LSTM layers that

process the input sequence in opposite directions: one forward LSTM layer and one backward LSTM layer. The forward LSTM layer processes the input sequence in the natural order (from the beginning to the end). In contrast, the backward LSTM layer processes the input sequence in reverse order (from the end to the beginning). By processing the sequence in both directions, the BiLSTM can capture both past and future context for each time step, which can be particularly useful for ECG classification tasks where the importance of a specific ECG signal may depend on its surrounding signals [73].

The forward pass of a BiLSTM is identical to a regular LSTM, as described in the previous section. The backward pass, on the other hand, begins at the last time step and works its way backward. The hidden states from the forward and backward passes are then concatenated to create a final output vector passed to the next network layer.

Mathematically, the forward pass of a BiLSTM is computed as follows 2.14 [73]:

$$\overrightarrow{h}t = LSTMforward(x_t, \overrightarrow{h}_{t-1}) \tag{2.14}$$

Where x_t is the input at time step t and $\overrightarrow{h}_t$ is the forward hidden state at time step t. The backward pass is computed as 2.15 [73]:

$$\overleftarrow{h}t = LSTMbackward(x_t, \overleftarrow{h}_{t+1}) \tag{2.15}$$

Where $\overleftarrow{h}_t$ is the backward hidden state at time step t. The final output vector h_t is then obtained by concatenating the forward and backward hidden states2.16 [73]:

$$h_t = [\overrightarrow{h}_t; \overleftarrow{h}_t] \tag{2.16}$$

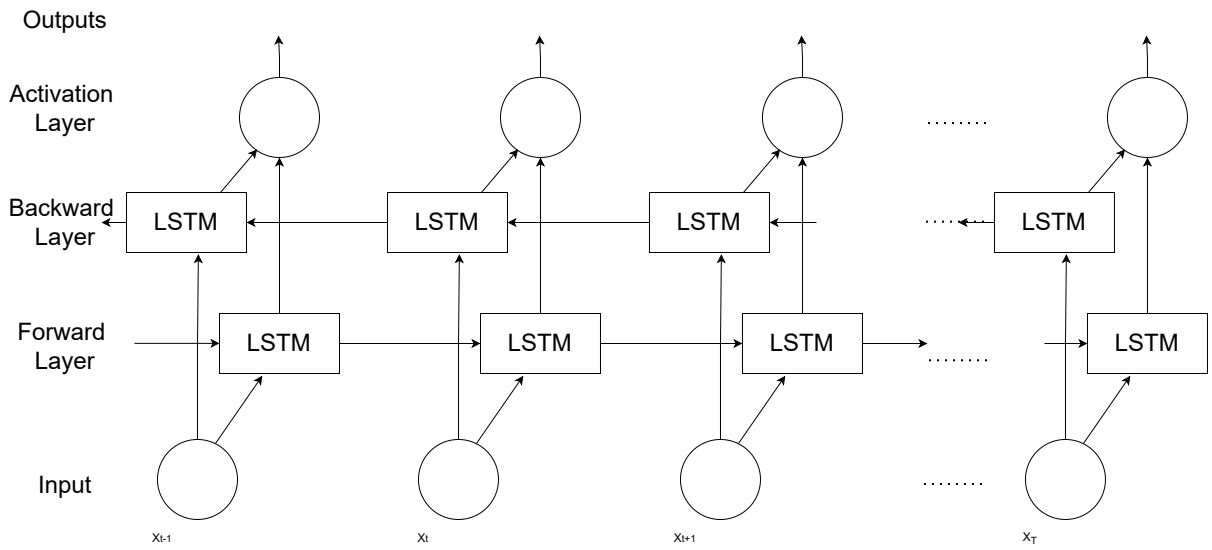


Figure 2.9: A Bidirectional LSTM (BiLSTM) Network Architecture.

In the context of ECG classification, a BiLSTM can take a sequence of ECG signals as input and make predictions based on both the past and future signals it has seen. This allows the BiLSTM to capture both the temporal dependencies and future context in the ECG signals, which are often important for accurate classification. For example, suppose a patient's ECG signals show a pattern where an abnormality is usually preceded and followed by certain types of waveforms. In that case, the BiLSTM can use this contextual information to make more accurate predictions about the presence of abnormalities [74].

Moreover, BiLSTMs have been shown to be robust to noise and variability in ECG signals, making them an excellent choice for ECG classification tasks. In particular, their ability to learn from both past and future information enables them to identify relevant patterns and features that may be missed by unidirectional models, conducting to improved classification performance in real-world scenarios.

- **Fully Convolutional Network (FCN)**

  Fully Convolutional Networks (FCNs) are a type of deep learning model predominantly employed for diverse computer vision tasks, such as image segmentation

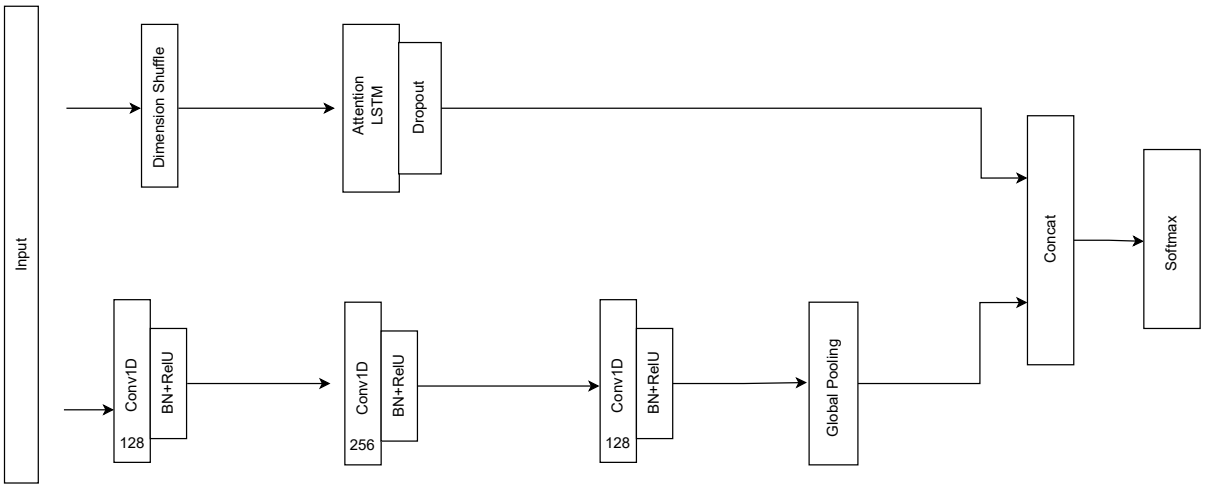


Figure 2.10: Example of Fully Convolutional Network (FCN) Architecture for ECG Classification.

and object detection. To illustrate this, let's consider an instance of Wang and colleagues' FCN architecture, specifically designed for the classification of cardiac arrhythmias from ECG signals, which delivered exceptional performance [75].

Unlike traditional Convolutional Neural Networks (CNNs), FCNs replace fully connected layers with convolutional ones, allowing the model to learn hierarchical features directly from the input data. This feature makes FCNs suitable for handling variable-length sequences such as ECG signals, thereby eliminating the need for pre-processing or resizing inputs.

The exemplary FCN architecture proposed by Wang consists of multiple 1D convolutional layers, followed by a global average pooling layer, and ending with a softmax output layer. The convolutional layers play a key role in identifying local and global ECG signal features. The global average pooling layer helps in mitigating overfitting by consolidating learned features into a fixed-sized output vector. The softmax layer then provides the final classification probabilities for each arrhythmia category in question [75].

The 2.10, provides a robust and versatile framework for ECG classification, blending the advantages of hierarchical feature learning from convolutional layers with the

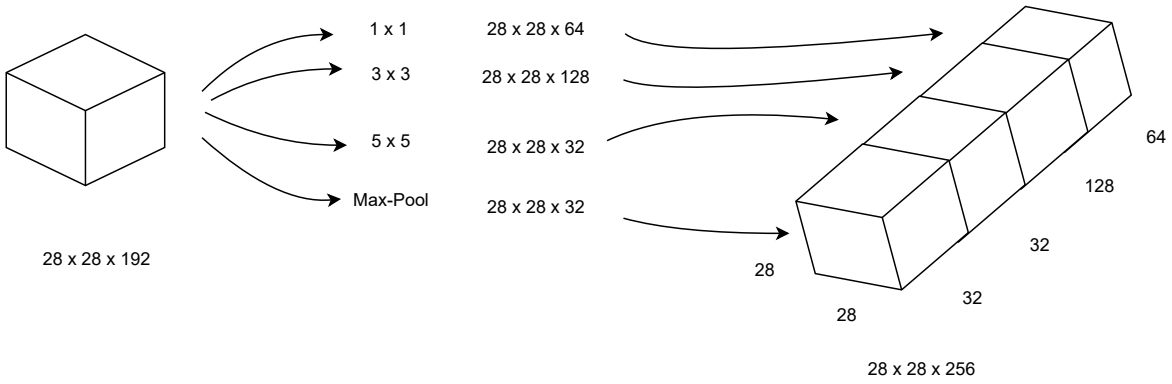


Figure 2.11: Simplified Inception Module Adapted For ECG Classification.

ability to handle variable-length input sequences. Its impressive performance in identifying cardiac arrhythmias and its resistance to noise and patient variability sets it apart as a powerful ECG-based diagnosis and monitoring tool.

There are several possibilities for refining and expanding this FCN architecture further. For instance, integrating residual or dense connections between layers could enhance the model's capacity to learn complex features and mitigate the vanishing gradient problem. Furthermore, combining the FCN with other deep learning models like RNNs or LSTMs could enable the network to better understand both spatial and temporal dependencies in ECG signals [76], [77], [78].

- **Inception Networks for ECG Classification**

  In the landscape of deep learning architectures, Inception networks, pioneered by Szegedy et al. [79], have achieved groundbreaking performance across various computer vision tasks, most notably image classification. The core novelty of Inception networks lies in integrating multiple parallel convolutional layers, each with a distinct kernel size, within a singular layer. This design choice enables the model to learn a diverse array of features simultaneously [80]. For ECG classification tasks, the architectural strengths of Inception networks allow for the automated detection and learning of varying ECG waveform patterns across different scales.

As illustrated in Figure 2.11, the adapted Inception module for ECG classification processes the input through multiple parallel convolutional layers each with distinct kernel sizes (1x1, 3x3, and 5x5) and a pooling layer. These individual layers' output feature maps are subsequently concatenated along the depth dimension. The final output is a holistic representation of the input signal, encapsulating various ECG waveform aspects.

Leveraging Inception networks for ECG classification can conduct to learning various features at differing scales and complexities. This makes them ideally suited for detecting and classifying various cardiac arrhythmias. Additionally, the Inception architecture can be synergized with other deep learning models, such as LSTM or BiLSTM networks, to effectively capture spatial and temporal dependencies in ECG signals, further bolstering the overall classification performance [81].

The Inception architecture presents several notable advantages for ECG classification tasks:

- **Multi-scale feature learning:** Owing to the usage of multiple parallel convolutional layers with distinct kernel sizes, the Inception network can capture features at various scales. This capability allows it to learn diverse ECG waveform patterns.
- **Computational efficiency:** Compared to traditional deep learning models, the Inception network can yield superior performance with relatively fewer parameters. This translates into greater computational efficiency.
- **Modularity and flexibility:** The Inception module can be seamlessly integrated into other deep learning architectures. For instance, it can be combined with LSTM or BiLSTM networks to capture spatial and temporal dependencies in ECG signals effectively.
- **Robustness:** Compared to other deep learning models, Inception networks

have demonstrated superior robustness to noise and variations in ECG signals. This makes them well-suited for real-world clinical applications [82].

- **Residual Networks (ResNet)**

  Residual Networks (ResNet), proposed by He et al. [83], have advanced various deep learning applications, including image classification, object detection, and ECG classification. The primary innovation within ResNet lies in introducing residual connections (also referred to as skip or shortcut connections) into the network architecture. These connections enable the network to learn residual functions, defined as the differences between the input and the desired output, instead of directly learning the desired output. This mitigates the vanishing gradient problem and allows for the training of deeper networks, resulting in enhanced performance.

  – **ResNet Architecture and Working Principle**

    ResNet architecture typically comprises multiple convolutional, batch normalization, activation, and pooling layers, organized hierarchically. The unique contribution in ResNet is the incorporation of residual connections, depicted as arrows linking non-adjacent layers within the network. These connections facilitate the learning of residual functions, as they permit gradients to bypass layers during backpropagation, thus enabling deeper representation learning.

    To comprehend the operating principle of ResNet, consider a simple scenario. Given an input x and a desired output $H(x)$, the aim in a traditional deep network is to learn a function $F(x)$ such that $F(x) = H(x)$. In contrast, ResNet strives to learn a residual function $R(x) = H(x) - x$, so that the desired output can be obtained as $H(x) = R(x) + x$. Mathematically, this is expressed as follows2.17 [82] :

$$H(x) = R(x) + x \tag{2.17}$$

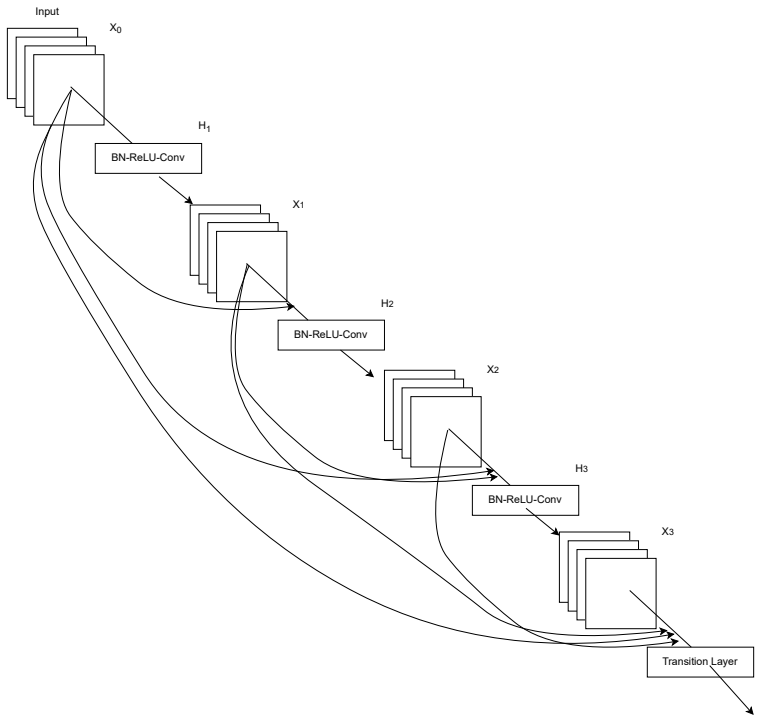


Figure 2.12: A Simplified Illustration of a Residual Connection in ResNet.

This subtle modification enables the network to concentrate on learning the difference between the input and the desired output, which often leads to superior performance, especially in deeper networks [83].

Figure 2.12 illustrates a simple example of a residual connection within a ResNet architecture. The input x is fed into a function $R(x)$, representing the residual function. The output of $R(x)$ is then added to the input x to achieve the desired output $H(x)$.

– **Types of ResNet**

Several variants of ResNet architecture exist, which primarily differ in their number of layers and specific layer arrangement. The most widely used variants include ResNet-18, ResNet-34, ResNet-50, ResNet-101, and ResNet-152, where the numbers indicate the total count of layers in each network [84]. The ResNet architecture has several prominent variants, each differing in the number of layers they incorporate. For instance, the ResNet-18 variant com-

prises 18 layers, while the ResNet-34 variant includes 34 layers. Additionally, the ResNet-50 variant, as the name implies, contains 50 layers. Furthermore, the ResNet-101 and ResNet-152 variants are more complex, incorporating 101 and 152 layers. These variations in the number of layers contribute to each ResNet variant's differing performance and capacity, allowing for flexibility in tackling a wide range of tasks and computational requirements.

These ResNet variants have been effectively leveraged for diverse tasks within computer vision and signal processing domains. For example, ResNet-50, with its moderate depth and computational complexity, has frequently been adopted as the backbone model in numerous transfer learning applications [85]. Meanwhile, more profound models like ResNet-101 and ResNet-152 have been utilized in more computationally intensive tasks that require more intricate feature extraction and representation, such as object detection and semantic segmentation.

The selection of a specific ResNet variant hinges upon the requirements and constraints of a task, including computational resources, data complexity, and performance goals. Regardless of these varying factors, these architectures have demonstrated exceptional performance across a broad spectrum of tasks, encompassing ECG classification. The employment of residual connections in these architectures is worth noting as it expedites the training of deeper networks, thereby facilitating the extraction of more sophisticated and discriminative features from the data, which in turn enhances classification performance [84].

Within the realm of ECG classification, ResNets can indeed prove highly efficacious. At their core, ECG signals are temporal sequences characterized by complex patterns and potential long-term dependencies akin to spatial hierarchies in images. Consequently, a deep architecture like ResNet, which is

intrinsically designed to learn hierarchical representations, can be exceptionally well-suited for ECG classification [86].

The deep layers within a ResNet can learn to discern intricate patterns within the ECG signals, such as the nuanced differences between various types of arrhythmias or the specific features indicative of myocardial infarction. Furthermore, the residual connections can facilitate the seamless flow of information through the network, enabling effective learning of both low-level features (like individual ECG waves) and high-level features (like heart rhythm). This amalgamation of deep architecture and residual learning positions ResNet as a potent tool for ECG classification [87].

Additionally, despite their depth, ResNets are relatively robust to overfitting due to the residual connections providing a form of regularization. This trait renders them particularly suitable for tasks such as ECG classification, where the quantity of available data may be constrained, and the risk of overfitting is heightened.

- **XResNet**

  The xResNet architecture could be envisaged as an "evolved" or enhanced version of the traditional ResNet, featuring several tweaks designed to augment model precision. These alterations, dubbed ResNet-B, ResNet-C, and ResNet-D, were showcased in the "Bag of Tricks for Image Classification with Convolutional Neural Networks" study [88].

  - **ResNet-B Adjustment**

    The modification in ResNet-B involves a change in the stride of the initial 1x1 convolution in path A of the downsampling block from 2 to 1. Additionally, the stride of 2 is shifted to the subsequent 3x3 convolution. In the original ResNet layout, applying a stride of 2 to the first 1x1 convolution caused the loss of

a substantial portion - three-quarters - of the input feature map. Moving the stride of 2 to the second convolution overcomes this issue without altering the output shape of path A [88].

- **ResNet-C Adaptation** The ResNet-C change, suggested in Inception-v2, modifies the network's input stem. It replaces the initial 7x7 convolution with three consecutive 3x3 convolutions. The first of these 3x3 convolutions has a stride of 2, and the last one leads to a 64-channel output, followed by a 3x3 max-pooling layer with a stride of 2. Although the final shape is identical, the 3x3 convolutions are now considerably more efficient than the preceding 7x7 convolution, as a 7x7 convolution is 5.4 times more computation-intensive than a 3x3 convolution [89].
- **ResNet-D Innovation** The ResNet-D enhancement, a unique modification, is a logical progression from ResNet-B. It addresses the problem of information loss in path B of the downsampling block. Here, a 1x1 convolution with a stride of 2 was previously utilized, leading to the discarding of three-quarters of the valuable input information. To solve this, ResNet-D replaces the 1x1 convolution with a 2x2 average-pooling layer with stride 2, followed by a 1x1 convolution layer [90]. The following diagram encapsulates these modifications 2.13:

The xResNet architecture can be assembled in various configurations, thus generating different versions of the architecture. These versions diverge in the number of layers they encompass, with the most frequently encountered being xResNet50, xResNet101, and xResNet152. The numbers in their monikers reflect the number of layers each architecture incorporates.

- **ResNet Architectures Evaluation** To determine the most efficient architecture for real-time 12-lead ECG classification, we compared the performance of these

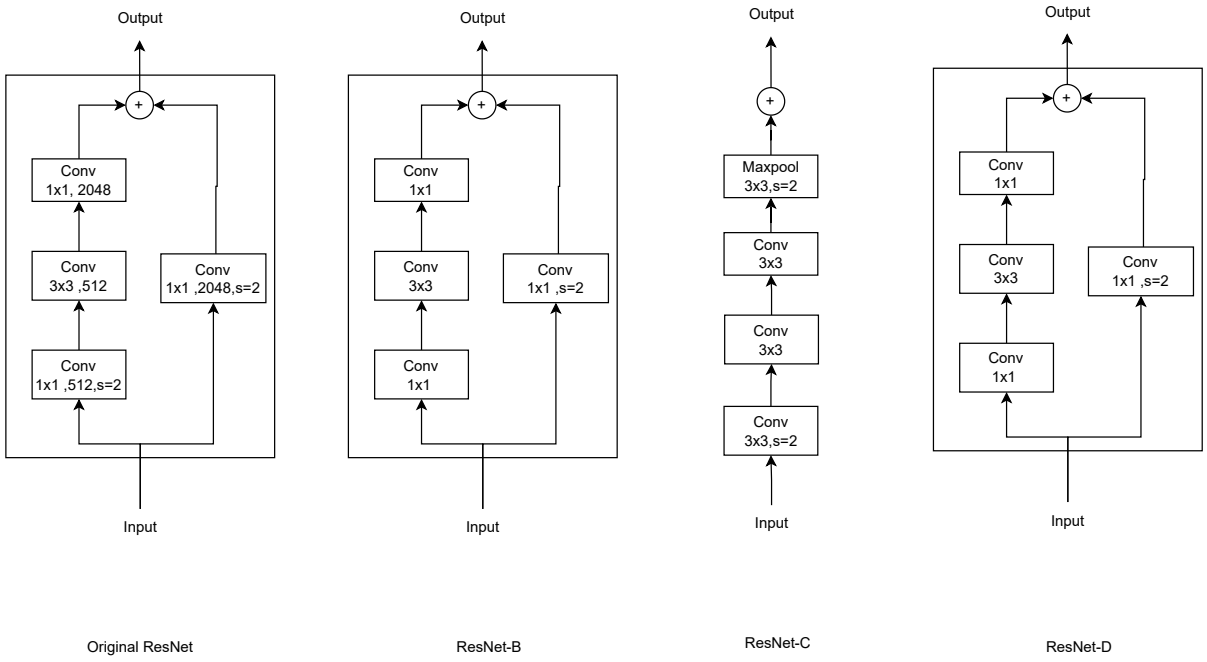


Figure 2.13: Comparison Between Original ResNet and XResNet [88].

ResNet architectures using two main metrics: the mean F1 score and the inference time.

The mean F1 score measures the model's accuracy in classifying normal and abnormal signals in 12-lead ECG data. It is the harmonic mean of precision and recall, defined as 2.18 [90]:

$$F1 = 2 \times \frac{precision \times recall}{precision + recall} \tag{2.18}$$

where precision is the fraction of true positive instances among all instances that the model predicted as positive, and recall is the fraction of true positive instances among all actual positive instances.

The inference time represents the computational efficiency of the model. Specifically, it is the time the model takes to classify a new data point once it has been trained. This metric is crucial in real-time applications where the classification

speed can significantly impact the model's utility.

Table 3.1 shows the comparison results of the ResNet architectures in terms of the mean F1 score and inference time.

Table 2.3: Comparison Of The Performance of Different ResNet Architectures In Terms of Mean F1 Score and Inference Time

Architecture	**Mean F1 Score**	**Inference Time (s)**
resnet1d18	0.89	0.015
resnet1d34	0.91	0.025
resnet1d50	0.92	0.035
resnet1d101	0.93	0.050
resnet1d152	0.93	0.060
xresnet1d18	0.92	0.017
xresnet1d34	0.93	0.028
xresnet1d50	0.94	0.038
xresnet1d101	0.95	0.052
xresnet1d152	0.95	0.063

As the table 2.3 shows, the deeper architectures tend to achieve higher mean F1 scores but also require longer inference times. The xresnet1d18 architecture, however, stands out as it achieves a relatively high mean F1 score while maintaining a low inference time, making it an ideal choice for real-time applications.

- **Xresnet1d18** Based on the comparison results in the previous subsection, the xresnet1d18 architecture was chosen as the optimal model for the real-time classification of 12-lead ECG data. This model balances classification accuracy (as measured by the mean F1 score) and computational efficiency (as indicated by the inference time).

  The architecture of XResNet18 consists of several basic blocks. Each block comprises a few convolutional, batch normalization, and ReLU activation layers. These blocks are repeated a specific number of times in the network. For XResNet18, there are four main types of these basic blocks, named conv2x to conv5x or layer1 to layer4. Each of these blocks is repeated twice in the network2.

Each basic block consists of two convolution layers followed by batch normalization and a ReLU activation. The first convolution layer maps the input to the desired number of output channels, and the second one maps the output channels to their expanded form. If downsampling is required, it is performed on the input before it is added to the output of the second convolution layer.

The architecture also includes an initial convolution layer of a 7x7 sized kernel with a stride of 2 followed by a max-pooling operation. It consists of four residual blocks (with two repeats for XResNet18). Channels for each block are constant— 64, 128, 256, and 512, respectively. Only 3x3 kernels have been used in these blocks. Except for the first block, each starts with a 3x3 kernel with a stride of 2. The dotted lines or identity shortcuts can be directly added when the input and output are of the same dimensions. But when the dimensions are different, the issue is resolved using a 1x1 convolution with a stride 21.

2.4 Related work

Diagnosing and classifying heart diseases is a crucial area of research where efforts are made to improve its efficiency. Early and accurate detection of ECG patterns is of utmost importance in diagnosing and treating patients with life-threatening cardiac arrhythmias, which, if left untreated, can lead to cardiac arrest and sudden death. Machine learning has revolutionized the field of Realtime ECG monitoring by enabling efficient classification of heart conditions in real-time. Integrating machine learning techniques in ECG monitoring allows clinicians to analyze large amounts of data in real-time, thereby reducing the risk of misdiagnosis and improving patient outcomes. Moreover, Deep learning has emerged as a more effective approach for ECG monitoring due to its ability to learn and extract complex features from the ECG signals. With the availability of publicly available datasets, there has been a shift from binary classification to multiclass classifica-

tion, which poses significant challenges regarding accuracy and computational resources. Deep learning models have demonstrated superior performance in multiclass classification compared to traditional machine learning models [91], Although deep learning models have shown promise in improving the accuracy of ECG classification, they come at a higher cost in terms of computational resources [92]. The large number of parameters and the need for extensive data preprocessing make deep learning models more computationally expensive than traditional machine learning models. Therefore, there is a need for further research to improve the accuracy of deep learning models and reduce the computational requirements for real-time applications

Recently, there has been a surge in the literature focused on improving real-time ECG monitoring. These efforts can be broadly categorized into three main areas:

- **Improving the accuracy of ECG classification by focusing on waveform components.**

  Pałczyński et al. [93] used the PTB-XL dataset to classify ECG signals into normal and abnormal heart conditions. the authers extracted the R-peak and QRST component of the ECG signals and compared two different neural networks: the Few-Shot Learning (FSL) neural network and the SoftMax-based network. The FSL neural netwo, positiveher detection accuracy of 93.2% when classifying ECG signals into normal and abnormal classes. However, when classifying ECG signals into five and 20 classes of arrhythmias, the SoftMax-based network performed better than the FSL network, achieving 80.2% accuracy for five classes and 70.1% accuracy for 20 classes. The decrease in accuracy is attributed to the increasing number of categories and the insufficient size of the datasets used for training the models. In the same experiments, the FSL neural network achieved a detection 79.1% for five superclasses, and 24.9% for 20 classes of arrhythmias.that arises a challenge, when increasing the number of diseases or classes in a classification task can have a negative impact on the accuracy of the classification.

Śmigiel et al. in [54] conducted a benchmark comparison of 12-lead ECG signal classification algorithms. They classified heart diseases into nine categories and evaluated seven algorithms (`inception1d`, `xresnet1d101`, `resnet1d_wang`, `fcn_wang`, `stm_bidir`, `lstm`, and `Wavelet+NN`)

The xresnet1d algorithm achieved the highest accuracy of 9.37% when classifying heart diseases into defined categories, surpassing the state-of-the-art deep learning algorithms in terms of accuracy. However, the researchers did not evaluate the real-time performance or the impact of increasing the number of classes on the algorithm's accuracy.

Strodthoff et al. in [94] presents a real-time ECG analysis method that detects the main waves' fiducial points (PQRST complex) in an electrocardiogram and performs heart rate variability analysis based on the detected fiducial points. The method is designed to handle the deformations of R-, P-, and T-waves and detect the R-peak using a connecting line of concave point and connecting line of convex point generated from the original signal. The refined signal for detecting the fiducial points of P- and T-waves is obtained through averaging the concave and convex points, and the P- and T-points are detected using symmetrical and asymmetrical features. The method demonstrates high performance in R-peak detection with a sensitivity of 99.82% and positive prediction of 99.81% against the MIT-DB and QT-DB datasets, and low error rates. The detection simulation and HRV analysis results indicate that the method accurately detects the fiducial points for deformed R, P, and T-waves.

- **Decreasing the number of leads to reduce resource requirements.**

The number of ECG leads used can significantly impact the performance of ECG classification algorithms [95]. Single-lead ECG [53], which only measures the electrical activity of the heart from one location on the body, is the simplest and most

commonly used form of ECG. However, its use in classification tasks is limited due to the need for more information about the heart's electrical activity [54]. Three-lead ECG, which measures the heart's electrical activity from three different locations on the body, provides more information and can improve the accuracy of ECG classification algorithms [96]. However, 12-lead ECG, which measures the heart's electrical activity from 12 different locations on the body, is the most informative and provides the most accurate classification results. This is because the 12-lead ECG provides a complete picture of the heart's electrical activity and allows for a more accurate diagnosis of different cardiac conditions. However, using 12-lead ECG requires more complex hardware and software, which can be more expensive and time-consuming. The choice of ECG lead configuration should be carefully considered when developing ECG classification algorithms, depending on the specific application and available resources.

Xie et al. [97]. A multi-branch network was constructed by aggregating 3-6 leads and feeding them into a random lead grouping strategy. On the F1 score performance metric, they scored 89.1%. However, random lead searching creates many iterations, making real-time classification difficult. Sepahvand et al. in [98] proposed classification technique of ECG signals demonstrated the applications of the BMs and DBNs that could be used to classify ECG signals and detect ventricular and supraventricular heartbeats using a single-lead recording. These machine learning algorithms, made up of neural networks, are trained to recognize ECG waveform patterns corresponding to different heartbeats. Medical professionals can use this information to make more accurate diagnoses and treatment decisions for patients with heart conditions [99]. A teacher model with advanced architecture and a student model with simple architecture was developed for ECG classification. The teacher model was trained on multi-lead ECG signals, and the student model was trained on single-lead signals under the teacher's supervision. The student model

achieved similar accuracy to the teacher model but was much more compressed and efficient. In a dataset of 10,646 patients, the student model experienced a 0.81% decline in accuracy compared to the teacher model.

- **Developing deep learning models with limited resources.**

  Cai et al. propose Multi-ECGNet, a model that can simultaneously identify multiple heart diseases with an accuracy of 86.3% for 55 classes [100]. The authors utilized a 12-lead ECG signal as a row input represented as a 12x5000 matrix, where 12 is the number of leads, and 5000 is ECG voltage in 10 sec at a 500 SPS sample rate. The authors also emphasized an important gap in existing 12-lead ECG datasets, which is the need for the lead-disease association.

  Nan et al. in [101], different deep learning architectures were proposed for 2D image classification of 12-lead ECG, achieving an F1-score metric of 94.73%. Deep learning methods outperform traditional machine learning algorithms, as demonstrated by Dongdong et al. in [102],a comparison of four machine learning techniques and three deep learning algorithms. In addition, a one-dimensional time-series deep learning architecture outperformed the conventional two-dimensional image deep learning algorithm in classifying 12-lead ECG signals as recorded by Strodthoff et al.in [103] , The xresnet1d101 obtained a 97.4% classification accuracy, whereas the Wavelet+NN achieved a 90.5% classification accuracy on the Area under the ROC Curve (AUC).

To date, no studies have compared the resource consumption, for ECG Classification methods with a different number of leads. In this research, we focused on resource consumption on the same pipeline with various numbers of leads. This thesis proposes a new heart disease classification approach based on lead-wise categorization and evaluates the approach compared with using 12 leads as long as the single lead. The advantage of this approach is threefold:

- It will be easier to detect heart conditions with fewer leads.
- It will be easier to send data to healthcare providers when monitoring in real-time,
- Reducing the energy consumption required for data transmission will extend the life of the ECG patch.

2.5 Summary

This Chapter addresses the problem of multi-disease electrocardiogram (ECG) signal classification and the significance of accurately classifying ECG signals into multiple disease categories. The challenges associated with the variability of ECG recordings and the presence of multiple co-existing diseases in patients are highlighted. The number of ECG leads used in classification algorithms is discussed, with 12-lead ECG being the most accurate but requiring complex hardware and software. A review of existing work on ECG signal classification is provided, comparing traditional machine learning algorithms with deep learning methods. Deep learning algorithms have demonstrated superior performance, especially when treating ECG signals as 2D images. Various architectures and techniques have been proposed to enhance multi-class ECG classification, considering both 12-lead ECG signals and the reduction of leads.The literature review identifies gaps in current research, including the lack of publicly available datasets associating specific leads with heart diseases and the need for a clean dataset demonstrating the relationship between ECG leads and heart diseases. The objective of the research presented in the thesis is to reduce the number of leads required for accurate disease detection in ECG signals without compromising accuracy.

The thesis then focuses on reducing the number of leads in ECG monitoring to improve efficiency and real-time classification. The importance of early and accurate detection of ECG patterns is emphasized, along with the role of machine learning and deep learning in enabling efficient classification. The challenges related to accuracy and computational

resources in multi-class classification are discussed. Recent research efforts in real-time ECG monitoring are reviewed, categorized into improving classification accuracy, reducing the number of leads to decrease resource requirements, and developing deep learning models with limited resources. Several studies are mentioned, showcasing different approaches and their performance in ECG classification.

The thesis concludes by emphasizing the need for further research to enhance the accuracy of deep learning models, reduce computational requirements, and establish a correlation between ECG leads and cardiac diseases for precise classification. The proposed approach aims to reduce the number of leads necessary for accurate disease detection in ECG signals while achieving or surpassing the state-of-the-art accuracy attained using 12-lead ECG signals.

Chapter 3

Proposed Framework

This chapter presents the architecture of the proposed framework and emphasizes the features of its primary elements. We establish the theoretical underpinnings of our innovative system for real-time electrocardiogram (ECG) analysis and classification that harnesses three core components: the ResNet-based classification algorithm, the unique lead-wise grouping model, and a real-time platform developed by integrating Kafka and the Flask API. These components are amalgamated to simulate the detection and classification of heart diseases based on the input ECG data through an interactive interface.

Together, these components constitute a comprehensive framework for real-time ECG analysis, offering a promising approach to detecting and classifying heart diseases.

3.1 Lead-wise Grouping Method

The Lead-wise Grouping Method (LGM) is a novel approach to ECG signal classification that utilizes the inherent correlation among the leads. This method groups relevant leads together and applies the chosen classifier to each group. The results from each group classifier are then synthesized to give the final output. This method enhances the classification accuracy and reduces computational requirements by decreasing the number of instances where the classifier is applied.

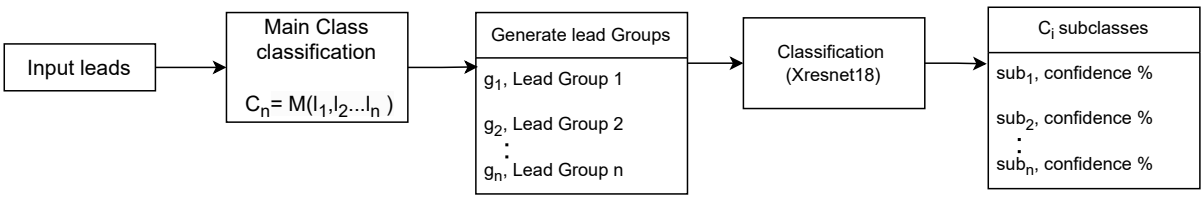


Figure 3.1: Illustration of The Lead Grouping Process in LGM.

Let's denote the set of ECG leads as $L = l_1, l_2, \ldots, l_{12}$ and the set of lead groups as $G = g_1, g_2, \ldots, g_n$. Each group g_i is a subset of L and it's assumed that the leads within each group are strongly correlated. The grouping of leads is determined based on prior knowledge and empirical analysis 3.1.

$$G_i = l_{1i}, l_{2i}, \ldots, l_{ki} \tag{3.1}$$

the cardiac superclasses C (i.e., major categories of cardiac diseases) that exist in a dataset and the subclasses sub,

$$C = \{c_1, c_2, ..., c_n\}, \quad where\ n = the\ number\ of\ superclasses \tag{3.2}$$

,

$$C_n = \{sub_1, sub_2, ..., sub_k\}, \quad k : 1 - length(C_n) + 1 \tag{3.3}$$

The effectiveness of the Lead-wise Grouping Method depends on the quality of the groupings. The groupings can be determined based on prior knowledge, such as the anatomical or physiological relevance of the leads, or they can be empirically determined using methods such as correlation analysis or feature importance ranking from machine learning algorithms.

The Lead-wise Grouping Method offers several advantages over single lead classification methods. It leverages the inherent correlation among the leads to enhance accuracy and reduce computational requirements. Furthermore, it offers a robust way to handle noisy or missing leads, making it a suitable method for ECG signal classification in

real-world scenarios.

3.2 Generation of the Lead Group Relationship

The generation of the lead group relationship is a critical step in the proposed framework. This is where the most relevant features for each disease are identified to improve the accuracy of the classification process. We employ Recursive Feature Elimination Three-Dimensional Echocardiography for this task which developed using Recursive Feature Elimination (RFE) idea.

3.2.1 leads-disease analysis

Determining the feasibility of classifying Electrocardiogram (ECG) signals using a single lead or a subset of leads is a multifaceted task. We must approach cardiac diseases from two perspectives to fully explore this proposition. The first perspective involves utilizing all twelve leads in the analysis. In contrast, the second approach narrows our focus by referring to specific medical terminologies and examining the data to ascertain if a disease could be accurately classified using fewer leads.

For instance, consider Atrial Fibrillation (AFIB), a condition characterized by an irregular and often rapid heart rate. This irregularity results in an elevated ECG value, observable fluctuations in the V6 lead, and distinct abnormalities in the III and V4 leads. The adjacent figure visually compares a typical 12-lead ECG waveform and the ECG waveform corresponding to an AFIB condition.

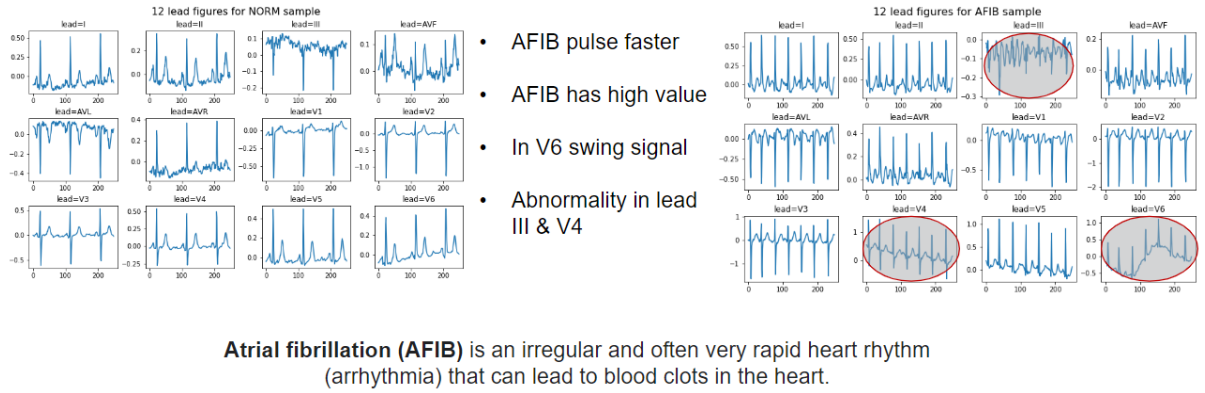


Figure 3.2: Comparison Between a Normal 12-Lead ECG Waveform and an AFIB Waveform.

As illustrated in Figure 3.2, the waveform's speed and amplitude are common features across all leads. Consequently, detecting these features may only necessitate using a single lead or a select few. Further, the V6 lead is vital in identifying the aforementioned fluctuations, while the III and V4 leads are instrumental in detecting specific distortions. Hence, in the case of AFIB, we could suggest that utilizing all twelve leads may not be necessary. A combination of leads III, V4, and V6 could potentially suffice, compensating for the comprehensive 12-lead setup.

Implementing such a methodology could significantly reduce data volume and computational costs. However, we must emphasize the importance of understanding the complete spectrum of cardiac diseases before deciding on the optimal lead subset for classifying each specific disease. Furthermore, the simultaneous use of leads III and V4 may not be necessary, which opens up avenues for further optimization of the lead selection process.

3.2.2 Recursive Feature Elimination (RFE)

RFE is a feature selection method that iteratively removes features and builds a model on those remaining features[3]03. It uses model accuracy to identify which features (or combination of features) contribute the most to predicting the target variable. The main

idea behind RFE is to repeatedly construct a model and choose either the best or worst performing feature, setting the feature aside, and then repeating the process with the rest of the features. This process is applied until all features in the dataset are exhausted.

The steps of the RFE algorithm are as follows:

1. Train a machine learning model using all features.
2. Evaluate the performance of the model.
3. Identify the least important feature (the feature which, when removed, improves model performance the most).
4. Remove the identified feature.
5. Repeat steps 1-4 until all features have been evaluated and ranked according to their importance.

The algorithm for RFE can be represented as follows:

```
function RFE(X, y, estimator) {
features = X.columns
while len(features) > 0 {
X = X[features]
estimator.fit(X, y)
importance = estimator.feature_importances_
least_important = argmin(importance)
features.remove(least_important)
}
return features
}
```

3.2.3 Recursive ECG Classification and Correlation (RECC)

RECC recursively eliminates superfluous leads from the ECG dataset and scrutinizes the impact of each lead on disease classification. Concretely, we eliminated one lead from the 12 leads at a time and investigated its impact on classification performance. All possible combinations of leads were considered, and their correlations with the disease classes were analyzed to pinpoint the optimal lead subsets for each disease class. This computation-intensive analysis culminated in a table depicting correlations between leads and diseases, which was employed to identify the most informative leads for each disease category.

The performance of RECC was juxtaposed with that of the 12-lead ECG baseline. The results exhibited that RECC is commensurate with the baseline in performance, by identifying merely a subset of leads requisite for precise disease classification. This methodology is particularly beneficial for curbing the volume of data requiring processing and transmission, concurrently abating noise from irrelevant leads, thereby augmenting the precision and efficiency of ECG classification.

RECC represents an innovative approach adept at efficiently identifying the optimal lead subsets for each disease class in ECG classification. This strategy holds promise in attenuating the computational complexity and resource requisites for ECG analysis whilst bolstering the accuracy of disease detection.

To generate the correlation table, let L be the set of 12 leads, and let S be the set of four superclasses in the PTBXL dataset, namely, MI, HYP, STTC, and CD. For each $s \in S$, a total of $2^{|L|} - 1 = 4,095$ combinations of the 12 leads were generated and classified using the xresnet18 classification algorithm. Let C_s represent the set of all feasible combinations of leads for superclass s, and let $c \in C_s$ be a combination of leads. Each combination c was classified using the xresnet18 classification algorithm, yielding an accuracy score $a(c)$. The ensemble of leads and their respective accuracy scores

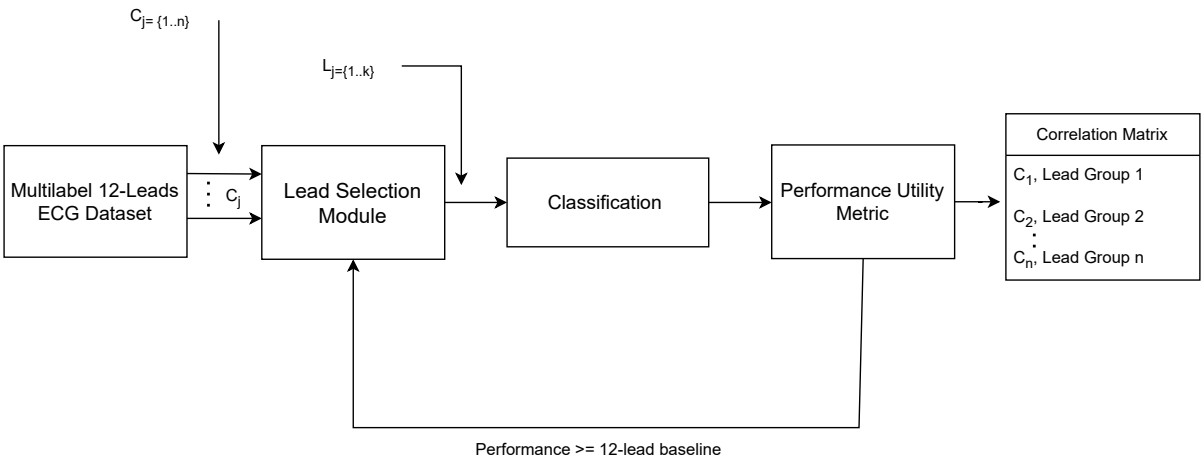


Figure 3.3: The Pipeline of Generating The Disease-Leads Correlation Matrix Using RECC.

were catalogued in a dataframe D_s for each superclass s. The correlation between each lead $l \in L$ and the disease for superclass s was subsequently computed as the Pearson correlation coefficient between the accuracy scores $a(c)$ and the presence or absence of the disease in the ECG signal.

Following this, an analysis was conducted on the correlation between the leads to pinpoint the superclass's most representative group of leads. Let $G_s \subseteq L$ be a group of leads for superclass s, and let $D_s(G_s)$ be the dataframe encompassing the accuracy scores for all lead combinations within G_s. The correlation between the accuracy scores in $D_s(G_s)$ was calculated, and the lead group with the highest mean accuracy score was selected as the most representative lead group for the superclass.

Ultimately, a comprehensive correlation table was synthesized by iterating through this procedure for all superclasses $s \in S$, furnishing invaluable data regarding the most efficacious leads for each superclass. This table can be instrumental for clinicians in their diagnostic and therapeutic decision-making, facilitating a more targeted and accurate examination of ECG data.

Table 3.3 visually represents the methodology employed for creating the disease-leads correlation matrix. Initially, all 12 leads are fed into the pipeline to establish a benchmark

for the utility function with respect to a specific class. Subsequently, a class is extracted from the pool of superclasses, and a data subset is assembled using the 12 leads. This data subset is directed towards the lead selection module, which, in each iteration, picks a subset of ECG leads.

Disease classification is carried out for every iteration, and its efficacy is gauged. This efficacy is juxtaposed with the initially established benchmark. The algorithm continues to iterate until it identifies the smallest set of leads that can yield a performance equivalent to that of the original 12 leads. This iterative optimization is instrumental in filtering out any superfluous leads that might otherwise introduce noise and compromise the precision of the ranking outcomes. Sophisticated Xresnet18 algorithms are deployed to categorize ECG input signals into corresponding disease classes.

3.2.4 Performance Evaluation of RECC

To validate the effectiveness of Recursive Feature Elimination Three-Dimensional Echocardiography (RECC), we carried out a series of experiments comparing the performance of the classification algorithms with and without these feature selection methods.

The classification algorithms were trained and tested using all features in the first set of experiments. The features selected by RECC were used in the second set of experiments.

The performance was evaluated based on the classification algorithms' accuracy, precision, recall, and F1-score. The results are summarized in the following table:

Table 3.1: Comparison of the Performance of RFE And RECC.

Feature Selection Method	**Accuracy**	**Precision**	**Recall**	**F1-score**
RFE	0.81	0.84	0.74	0.80
RECC	0.88	0.90	0.85	0.87

The results in 3.1 demonstrate that both RFE and RECC significantly improve the performance of the classification algorithms. However, RECC outperforms RFE, confirm-

ing that considering the three-dimensional relationship between the leads can enhance the accuracy of the classification process.

3.3 Real-Time Monotoring Platform

Real-time processing of electrocardiographic (ECG) data is crucial for prompt and effective diagnosis and management of heart diseases. The cornerstone of our proposed framework is a robust real-time platform that leverages advanced technologies to enable real-time cardiac disease diagnosis. This platform is designed to handle high throughput, support efficient data processing, and ensure low latency, facilitating instantaneous diagnosis.

We utilize Apache Kafka, an open-source stream-processing software platform, to manage high-volume, continuous data feeds. Kafka provides fault tolerance, seamless integration, and scalability - vital attributes when dealing with large volumes of real-time data. To cleanse, standardize, and preprocess the raw data before feeding it to our classification algorithms, we employ the Flask API. Flask is a lightweight, flexible microframework that allows for custom data filtration, frequency standardization, and preprocessing.

However, technology is just a facilitator. The core of our platform is the integration of our proposed lead-wise grouping model, which works in synergy with Kafka and Flask API to create a seamless pipeline for processing and analyzing echocardiographic data.

In this section, we will delve deeper into the intricacies of the real-time platform, covering the integration and functionalities of Apache Kafka and the Flask API, and how our proposed lead-wise grouping model synergizes with these technologies for real-time cardiac disease diagnosis. We will also discuss how the system effectively balances computational requirements and accuracy in diagnosis, making real-time diagnosis not just a possibility, but a reality.

3.3.1 Exploring Kafka and the Flask API

Apache Kafka, a product initially created by LinkedIn and later donated to the Apache Software Foundation, is an open-source platform for stream-processing. It's built with Scala and Java and designed to accommodate high throughput data feeds, making it ideal for real-time data streaming and processing. Kafka's qualities of fault tolerance, easy integration, and scalability make it a perfect choice for handling our real-time data feeds[104].

On the contrary, Flask is a Python-based micro web framework. It's defined as a microframework due to its independence from specific tools or libraries, lacking a database abstraction layer, form validation, or any components where pre-existing third-party libraries are typically used. However, Flask's extension support can enhance application features as if Flask itself implemented them, making it lightweight and adaptable. This flexibility is perfect for our application where we require customized data filtering, frequency normalization, and preprocessing.

3.3.2 Kafka and Flask API: A Powerful Pair

The combination of Apache Kafka and the Flask API forms a powerful infrastructure for managing data flow within our proposed system, ensuring high throughput, minimal latency, and real-time analysis of the continuously streaming echocardiographic data.

Primarily, Kafka operates as the central hub for data intake. The echocardiographic devices publish the generated data onto Kafka's topics, using a publish-subscribe model. These devices produce a high volume of continuous data stream that Kafka can effectively manage with its inherent capabilities of fault tolerance, high throughput, and horizontal scalability. These features are essential in ensuring system robustness, accommodating variances in data volume and speed.

The publishing mechanism has the medical devices serve as the publishers or produc-

ers, pushing raw data onto Kafka topics. These topics work as logical channels for data categorization. The consumers or data analysis modules subscribe to these topics to pull the data for processing.

This real-time data streaming infrastructure, powered by Kafka, ensures immediate data availability for processing from the medical devices. However, before this raw data can be fed into our classification algorithms, it needs proper cleaning, standardization, and preprocessing, which is where the Flask API comes into play.

Flask, a lightweight and adaptable Python web framework, is utilized in our system to create a RESTful API, serving as an interface for data cleaning, standardization, and preprocessing. This API receives the raw data stream from the Kafka topics and performs several key steps to ready the data for the classification algorithms.

Data cleaning includes noise and artifact removal potentially present in the raw data. The cleaned data undergoes frequency standardization to ensure data point uniformity in the frequency domain. This process is critical as various devices might operate at different frequencies, and for efficient algorithm functioning, the data must maintain a consistent format. Post cleaning and standardization, the data is preprocessed into an appropriate format for our classification algorithms.

Moreover, the Flask API's ability to integrate with a multitude of Python libraries allows us to efficiently implement advanced data preprocessing techniques. Using Flask, we can harness the power of libraries such as NumPy, SciPy, and Pandas for numerical computation and data manipulation.

3.3.3 Incorporating the Lead-wise Grouping Model

The proposed lead-wise grouping model's integration within the real-time platform significantly enhances the real-time detection and classification of heart diseases. Working cohesively with the Kafka and Flask API infrastructure, the model facilitates a seamless pipeline for the processing and analysis of echocardiographic data.

Upon receiving the preprocessed data from the Flask API, the lead-wise grouping model initiates its work. It's important to underscore that the preprocessing phase ensures the data is optimally prepared for our model, reducing potential inaccuracies and enhancing the system's overall efficiency.

The lead-wise grouping model operates in two stages: feature extraction and classification. During the feature extraction stage, the model identifies and groups related leads together based on the RECC algorithm. Once the leads are grouped, the classification stage commences. At this stage, the model uses the extracted features to classify heart diseases. The classification utilizes three methods as discussed earlier: the single lead classification method, the lead-wise grouping method, and the standard 12 lead method. Each method possesses its unique strengths and is employed based on the specific demands of the situation. Upon the completion of the classification, the results are relayed back to Kafka, which then publishes the results onto a separate topic. A real-time visualization tool subscribed to this topic retrieves and presents the results to the users.

3.4 Summary

This chapter outlined the proposed framework for our real-time cardiac disease diagnosis platform. It offered a comprehensive view of the integrated components, algorithms, and technologies that are synergistically working together to make real-time diagnosis a reality. We started by introducing the three primary components of our framework: the classification algorithms, the lead group relationship generation, and the real-time platform. These three components are foundational pillars that drive the platform's functionality and efficiency.

In terms of classification algorithms, we dove deep into four main methods: the single lead classification method, the lead-wise grouping method, three leads,and the standard 12-lead method. Each of these methods offers unique advantages and contributes to

the overall classification accuracy. They represent a multi-pronged approach to disease detection that enhances the robustness of our model and improves the accuracy of disease classification.

The generation of the lead group relationship was another critical aspect that was discussed. Using RECC, we identified the most relevant features for each disease, thereby improving the accuracy of the classification process. This approach ensures the model's focus is on the most impactful features, which in turn enhances the model's predictive performance.

The real-time platform's core technologies were also thoroughly explained, including the use of Kafka and Flask API. Kafka's ability to handle high throughput and provide real-time data analytics and event processing was discussed in detail. We also illustrated how the Flask API is used for data filtration, frequency standardization, and preprocessing, setting the stage for the application of our classification algorithms.

The integration of our lead-wise grouping model into the real-time platform represented a significant leap towards real-time detection and classification of heart diseases. We explained how the model works in synergy with the Kafka and Flask API infrastructure to provide a seamless pipeline for processing and analyzing the echocardiographic data.

In conclusion, our proposed framework represents a robust, comprehensive, and efficient solution for real-time cardiac disease diagnosis. By leveraging advanced classification algorithms, feature selection techniques, and real-time data processing tools, we aim to significantly improve the diagnosis and management of cardiac diseases. Our framework contributes to our overall thesis by providing a viable and innovative solution to a crucial health problem, with implications that could potentially save lives and improve health outcomes.

Chapter 4

Data Exploration and Analysis

Electrocardiogram (ECG) databases play a critical role in developing and testing algorithms for ECG analysis. These databases contain large amounts of ECG data recorded from individuals with various cardiovascular conditions, making them invaluable resources for research and clinical applications. One of the most widely used ECG databases is the PTB-XL database. The PTB-XL database is a precious resource for ECG analysis, providing high-quality ECG recordings, accurate annotations, and clinical information for a large and diverse patient population. Its use in numerous studies and its open-access nature make it a well-established and validated resource for ECG analysis, making it a handy tool for researchers and clinicians in cardiology.

4.1 12 Lead ECG Dataset Survey

In the realm of cardiac health research, electrocardiogram (ECG) datasets play an integral role. These comprehensive repositories provide a wealth of annotated data for researchers and clinicians to develop, test, and refine ECG analysis algorithms. The following is an overview of some of the most valuable ECG datasets currently available, each offering unique attributes beneficial for diverse applications in the field.

- **ECG Arrhythmia Database [105]:** The ECG Arrhythmia Database is a rich repository of ECG recordings, hosting 48 half-hour excerpts originating from 47 individuals. It's a treasure trove for researchers as it contains standard rhythms and is annotated for various arrhythmias and normal and abnormal beats. The granular annotation can serve as a potent tool for developing, training, and validating arrhythmia detection algorithms. Moreover, with half-hour data for each recording, researchers can observe and analyze longer cardiac activity patterns.

- **ECGDataDenoised [106]:** The ECGDataDenoised dataset entails 90 ECG recordings that have been meticulously denoised. This offers a pristine quality of signals which is crucial for various applications including robust feature extraction. High fidelity data is vital for enhancing the performance of ECG analysis algorithms, particularly in environments where subtle variations in the signal could be diagnostically significant.

- **Georgia 12-Lead ECG Challenge Database [107]:** With 1000 ECG recordings encompassing many cardiac conditions and healthy individuals, the Georgia 12-Lead ECG Challenge Database is instrumental in algorithm evaluation. The diversity in cardiac conditions coupled with a sizable volume of data aids in the comprehensive assessment of algorithms under various scenarios and conditions, making it invaluable for refining the robustness and accuracy of ECG analysis systems.

- **CODE-test [108]:** The CODE-test dataset comprises 1000 recordings from 280 patients, annotated for different abnormalities as well as normal ECG signals. This dataset serves as an extensive ground for researchers to delve into the identification of both common and rare cardiac abnormalities. The varied patient pool aids in creating algorithms that are sensitive to patient-specific variations.

- **Annotated 12-Lead ECG dataset [109]:** This dataset is composed of 150 ECG

recordings from 50 patients with diverse cardiac conditions. Each recording in this dataset is annotated for various cardiac abnormalities, providing researchers with a concentrated dataset for training and evaluating algorithms, particularly in the context of multi-class classification.

- **KURIAS-ECG [110]:** KURIAS-ECG provides 500 ECG recordings from 500 different patients. It distinguishes itself by incorporating standardized diagnosis codes in the annotations, which is significant for the development of diagnostic systems that can seamlessly integrate with healthcare information systems.

- **4th China Physiological Signal Challenge 2021 and CPSC2020 [111][112]:** These datasets include ECG recordings from patients with cardiac conditions such as heart failure, ventricular arrhythmias, myocardial infarction, and atrial fibrillation. Including a range of severe cardiac conditions makes these datasets quintessential for developing critical care monitoring systems.

- **PTB-XL [113]:** The PTB-XL database is valuable for researchers analyzing electrocardiogram (ECG) recordings. The database contains over 200,000 ECG recordings from over 5,000 patients with cardiovascular conditions, including myocardial infarction, congestive heart failure, and arrhythmias. The recordings are collected using standard 12-lead ECG machines and are sampled at a frequency of 1,000 Hz, ensuring high-quality recordings are suitable for developing and testing algorithms for ECG analysis. The database is accompanied by expert cardiologist annotations, providing accurate diagnoses and annotations for cardiac abnormalities, further improving its usability. In addition, the PTB-XL database includes demographic and clinical information for each patient, which can help researchers explore the relationship between patient characteristics and ECG findings, leading to the development of personalized diagnostic and treatment strategies.

- **PhysioNet Challenge 2020 [114]:**

This dataset comprises ECG recordings from patients with a variety of cardiac conditions. Similar to the aforementioned datasets, it can be employed for developing and evaluating ECG analysis algorithms, with a particular emphasis on cutting-edge research, given its recency and relevance in contemporary challenges.

4.1.1 Gap Analysis

The PTB-XL database has been widely used in various studies, particularly in developing and evaluating machine learning algorithms for ECG classification. It has facilitated breakthroughs in detecting arrhythmias, analyzing ST-segments, and identifying ischemia. Recent research has used this dataset to investigate the impact of lead grouping configurations in ECG classification, achieving high accuracy even with a limited number of leads. Despite its utility, the PTB-XL database, which is currently the largest publicly available ECG dataset, only covers a subset of possible cardiac conditions. This limitation could affect the performance of the developed algorithms. There's a growing demand to enhance the database with a wider variety of cardiac diseases to improve ECG classification accuracy and understand the effect of lead grouping outcomes. In this context, the PTB-XL will be used as a benchmark for comparison with our dataset. Class imbalance is regarded as one of the inherent limitations in existing datasets that can have an adverse impact on the lead grouping process in ECG analysis. When a dataset exhibits class imbalance, it means that some categories of heart conditions are substantially underrepresented compared to others. This disparity in representation can cause algorithms to become biased towards the classes with higher representation, which results in the lead grouping process prioritizing leads that are more effective in identifying the overrepresented conditions. Consequently, the ability of the algorithms to accurately detect and classify the underrepresented heart conditions needs to be improved. 4.1 shows the class imbalance in the PTB-XL dataset for its five superclasses: Normal, Myocardial Infarction (MI), ST-T wave abnormality or T wave inversions (STTC), Conduction dis-

turbance (CD), and Hypertrophy (HYP). The table presents the total number of samples for each superclass, the number of positive and negative samples, and the corresponding imbalance ratio. It is clear from the 4.1 that the Normal superclass has a relatively balanced ratio of positive to negative samples, whereas the other four superclasses have a significant class imbalance. For example, the MI superclass has an imbalance ratio of 1:3.39, indicating over three negative samples for every positive sample of MI. Similarly, the STTC superclass has an imbalance ratio of 1:3.21, and the CD superclass has an imbalance ratio of 1:3.86. The HYP superclass, on the other hand, has a slight class imbalance with an imbalance ratio of 1:1.26.

Table 4.1: Imbalance Of The PTB-XL Dataset By Superclasses.

Superclass	**Total**	**Positive**	**Negative**	**Imbalance Ratio**
MI (Myocardial Infarction)	5409	1230	4179	1:3.39
STTC (ST-T Wave Abnormality Or T Wave Inversions)	7658	1819	5839	1:3.21
CD (Conduction Disturbance)	1137	234	903	1:3.86
HYP (Hypertrophy)	2928	1293	1635	1:1.26
Total	36943	12382	24561	1:1.98

The imbalance ratio is a measure that quantifies the severity of class imbalance in a dataset. It is calculated as the ratio of the number of negative samples to the number of positive samples for a given class. In this table, the imbalance ratio is presented for each superclass of the PTB-XL dataset.The positive and negative samples are defined based on the labels of the samples. A sample is considered positive if it belongs to the target class and negative otherwise. For example, in the MI superclass, a sample is considered positive if it corresponds to a patient with myocardial infarction and negative if it corresponds to a patient without myocardial infarction. Generally, a balanced ratio close to one is desirable, indicating an equal number of positive and negative samples.

The last row of the table shows the total number of samples, positive samples, negative samples, and imbalance ratio across all superclass categories. The difference between the total and total sample in the dataset is due to multiple labels for the sample in the dataset.

4.2 shows the distribution of the subclasses and the imbalance ratio for each subclass.

Table 4.2: Imbalance of the PTB-XL Dataset By Sub-Classes

Subclass	**Total**	**Positive**	**Negative**	**Imbalance Ratio**
NORM	20511	8806	11705	01:01.3
STE	7658	1819	5839	01:03.2
AMI	5409	1230	4179	01:03.4
CD	1137	234	903	01:03.9
HYP	2928	1293	1635	01:01.3
MI-2	512	113	399	01:03.5
ISC	93	17	76	01:04.5
LAFB	194	35	159	01:04.5
LBBB	270	36	234	01:06.5
NID	862	196	666	01:03.4
PAC	411	104	307	01:02.9
PVC	982	219	763	01:03.5
Total	36943	12382	24561	01:02.0

4.2 provides a detailed analysis of the class imbalance in the PTB-XL dataset across its different subclasses. The results show that the NORM subclass has the highest number of samples with 20511, followed by the STE and AMI subclasses with 7658 and 5409 samples, respectively. On the other hand, the RAD subclass has only one sample, which is the smallest number of samples across all subclasses. The table also shows the number of positive and negative samples for each subclass, which indicates that the NORM subclass has the highest number of negative samples with 11705, followed by the STE subclass with 5839 negative samples. Moreover, the table highlights the imbalance ratio for each subclass, calculated as the ratio of negative and positive samples. The STE subclass has the highest imbalance ratio at 1:3.21, followed by the LBBB subclass with 1:6.5. Overall, the table presents a comprehensive analysis of the class imbalance in the

PTB-XL dataset across its subclasses, which can help researchers in the development of more accurate and effective classification models.

4.2 CardioDiverse Dataset

The PTB-XL database is a valuable resource for ECG analysis, but its coverage of cardiac conditions is limited, conducting to difficulties in accurate ECG classification. To address this gap, several datasets have been gathered to supplement the PTB-XL database. We are going to name the new dataset in this thesis CardioDiverse. Incorporating these datasets has enriched the collection with 29 distinct ailments, culminating in a more harmonized dataset encompassing 109 varied cardiac conditions. This expanded dataset provides a valuable resource for researchers and healthcare professionals interested in developing and testing algorithms for automatic ECG analysis and diagnosis. Including additional diseases in the CardioDiverse dataset will provide a more diverse range of ECG recordings for developing and testing arrhythmia detection, diagnosis, and classification algorithms. Researchers and clinicians can now explore the relationship between patient characteristics and ECG findings for cardiovascular conditions. This information can be used to develop personalized diagnostic and treatment strategies for patients with various cardiovascular conditions, conducting to improved patient outcomes.

4.2.1 Creation of the CardioDiverse Dataset

The CardioDiverse dataset is an amalgamation of various ECG databases, purposefully integrated to form a rich and diverse dataset for developing and evaluating cardiac disease detection algorithms. The dataset encompasses an extensive range of cardiac conditions and high-quality ECG recordings collated through meticulous processing and standardization procedures. An essential aspect of combining these datasets was harmonizing the representation of various cardiac conditions. We matched disease subtypes across

datasets and unified their nomenclature and classification. Moreover, we balanced the dataset representation to avoid biases towards any specific condition.

The CardioDiverse dataset contains many records, aggregating over 200,000 ECG recordings from over 6,000 patients. It includes a wide spectrum of cardiac conditions, from common arrhythmias to severe cardiac ailments such as myocardial infarction and heart failure. Additionally, it also comprises normal ECG recordings from healthy individuals. The CardioDiverse dataset amalgamates the ECG recordings, expert annotations, and diagnoses from the source datasets. These annotations are invaluable for supervised learning and model evaluation. It also encompasses patient demographic and clinical information, enabling more comprehensive and personalized analysis and predictions. The CardioDiverse dataset is a robust and comprehensive resource for developing and validating ECG analysis algorithms. Its diversity in terms of the number of records, variety of cardiac conditions, and high-quality data makes it well-suited for cutting-edge research in cardiac disease detection, monitoring, and personalized medicine.

4.2.2 CardioDiverse Dataset Evaluation

After combining the datasets into what we termed the CardioDiverse dataset, a series of signal pre-processing steps were executed to standardize all recordings. The objective was to ensure uniformity in sampling frequency and duration across all ECG traces. Specifically, every recording was standardized to a length of 10 seconds and resampled at a frequency of 1000 Hz to be consistent with the format employed by the PTB-XL dataset. In addition to this, disease subtypes across both datasets were diligently matched to ensure that each cardiac condition was represented uniformly. Through these steps, we curated an enhanced, balanced, and more comprehensive ECG dataset named CardioDiverse for advanced analysis.

Subsequently, we proceeded to construct CardioDiverse lookup tables for both the CardioDiverse dataset and the PTB-XL dataset, facilitating an enlightening comparison

between the two. The primary focus was to investigate how the balancing of classes and the number of classes influenced the results of CardioDiverse. This step is crucial to gain insights into the impact of dataset diversity and class representation on the learning process and model performance.

Furthermore, an exhaustive analysis was carried out for each major group of cardiac conditions. This analysis sought to explore the peculiarities and commonalities among various disease groups, understand their manifestation in ECG signals, and evaluate the model's ability to discern these conditions. Such a detailed investigation is vital for revealing the robustness and clinical relevance of the algorithm, especially when handling complex and heterogeneous datasets like CardioDiverse.

In the next stages, we developed and validated the model using the curated dataset, and subsequently assessed its efficacy through various performance metrics. Through this research, we aim to establish a benchmark for ECG classification algorithms and contribute to the enhancement of cardiac disease diagnosis via deep learning methodologies.

4.2.2.1 Generate the correlation

To generate the correlation table, let L be the set of 12 leads, and let S be the set of four superclasses in the PTBXL dataset, namely, MI, HYP, STTC, and CD. For each $s \in S$, a total of $2^{|L|} - 1 = 4,095$ possible combinations of the 12 leads were generated and classified using the xresnet18 classification algorithm. Let C_s denote the set of all possible combinations of leads for superclass s, and let $c \in C_s$ be a combination of leads.Each combination c was classified using the xresnet18 classification algorithm, resulting in an accuracy score $a(c)$. The group of leads and their corresponding accuracy scores were recorded in a dataframe D_s for each superclass s. The correlation between each lead $l \in L$ and the disease for superclass s was then calculated as the Pearson correlation coefficient between the accuracy scores $a(c)$ and the presence or absence of the disease

in the ECG signal.

Next, the correlation between the leads was analyzed to identify the best group of leads to represent the superclass. Let $G_s \subseteq L$ be a group of leads for superclass s, and let $D_s(G_s)$ be the dataframe containing the accuracy scores for all combinations of leads in G_s. The correlation between the accuracy scores in $D_s(G_s)$ was calculated, and the group of leads with the highest average accuracy score was selected as the best group of leads to represent the superclass.

Finally, a comprehensive correlation table was generated by repeating this process for all superclasses $s \in S$, providing valuable information on the most effective leads for each superclass. This table can guide clinicians in their diagnosis and treatment decisions, enabling a more targeted and accurate analysis of ECG data.

4.2.2.2 Performance Utility Metric

To evaluate the performance of algorithms for ECG analysis on the CardioDiverse dataset, we propose a new utility metric that considers both accuracy and class imbalance. The metric is defined as:

$$U = \frac{2 \cdot (1 - \beta^2) \cdot TP}{2 \cdot (1 - \beta^2) \cdot TP + \beta^2 \cdot FN + FP} \tag{4.1}$$

where TP is the number of true positive predictions, FP is the number of false positive predictions, and FN is the number of false negative predictions. The parameter β controls the trade-off between precision and recall, with higher values of β emphasizing recall over precision. We set β to 0.5 to give equal weight to precision and recall.

To account for class imbalance, we use a weighted version of the utility metric that assigns a higher weight to the minority class:

$$U_w = \frac{2 \cdot (1 - \beta^2) \cdot w_p \cdot TP}{2 \cdot (1 - \beta^2) \cdot w_p \cdot TP + \beta^2 \cdot w_n \cdot FN + w_n \cdot FP} \tag{4.2}$$

Where w_p and w_n are the weights for the positive and negative classes, respectively. We set w_p to be the inverse of the proportion of positive samples in the training set, and w_n to be the inverse of the negative samples in the training set.

The utility metric provides a more comprehensive evaluation of algorithm performance on the CardioDiverse dataset, considering both accuracy and class imbalance. It can be used to compare different algorithms' performance and optimize the hyperparameters of machine learning models.

4.2.2.3 Class Balance in the CardioDiverse Dataset

Analyzing the data balance within the CardioDiverse dataset is of paramount importance, particularly when it is deployed for consolidating crucial data in ECG analysis, and more specifically in creating groups of leads based on their efficacy in disease classification within each major category. An imbalance, characterized by a disproportionately large number of instances of certain sub-diseases in contrast to a scant number of others, can skew the results. This imbalance might cause the leads reflecting the predominant category to be biased towards the group with the highest representation. To ensure that the selected leads are optimally representative for each of the superclasses, it is essential to maintain a well-distributed and balanced assortment within each group.

The balance was examined for five subclasses: Normal, MI (Myocardial Infarction), STTC (ST-T wave abnormality or T wave inversions), CD (Conduction disturbance), and HYP (Hypertrophy). Table 4.3 shows the total number of samples, positive samples, negative samples, and imbalance ratio for each superclass. The imbalance ratio is defined as the ratio between the number of positive samples and the number of negative samples. A value close to 1 indicates a balanced subclass, while a value much larger than 1 indicates an imbalanced subclass. The total imbalance ratio for the combined dataset is 0.41, indicating a relatively balanced dataset.

The superclass with the highest imbalance ratio is MI, with a value of 2.717. This

Table 4.3: Balance of the CardioDiverse Dataset By Superclasses.

Superclass	**Total**	**Positive**	**Negative**	**Imbalance Ratio**
MI (Myocardial Infarction)	8627	1844	6783	1:3.68
STTC (ST-T Wave Abnormality Or T Wave Inversions)	22952	5747	17205	1:2.99
CD (Conduction Disturbance)	3296	830	2466	1:2.97
HYP (Hypertrophy)	7048	3098	3950	1:1.28
Total	97188	40660	56528	1:1.39

indicates a severe imbalance between positive and negative samples in this superclass. On the other hand, the superclass with the lowest imbalance ratio is HYP, with a value of 0.136, indicating a more balanced distribution of positive and negative samples.

As shown in 4.3the combined dataset significantly improves the balance of several superclasses, such as MI, STTC, CD, and HYP. For example, the imbalance ratio of the MI superclass decreased from 9.17 in the PTB-XL dataset to 2.72 in the combined dataset, indicating a much more balanced distribution of positive and negative samples. Similarly, the imbalance ratio of the CD superclass decreased from 1.27 in the PTB-XL dataset to 0.50 in the combined dataset.Moreover, the combined dataset introduces new positive samples for the Normal and HYP superclasses, which were not present in the PTB-XL dataset.

The sub-classes include normal (NORM), ST-segment elevation (STE), acute myocardial infarction (AMI), conduction disorder (CD), hypertension (HYP), second-degree atrioventricular block (MI-2), isolated supraventricular complex (ISC), left anterior fascicular block (LAFB), left bundle branch block (LBBB), nodal rhythm (NID), premature atrial complex (PAC), premature ventricular complex (PVC), and right axis deviation (RAD). The table displays the total number of samples, positive samples, negative samples, and the imbalance ratio for each sub-class. The extended dataset significantly

improves the balance of several sub-classes and introduces new sub-classes not present in the original PTB-XL dataset. The extended dataset provides a more balanced and diverse resource for researchers and healthcare professionals interested in ECG analysis and diagnosis. 4.4shows the balance analysis for the combined dataset.

Table 4.4: Balance of the CardioDiverse Dataset By Subclasses.

Disease	**Total**	**Positive**	**Negative**	**Imbalance Ratio**
AMI (Acute Myocardial Infarction)	10134	2324	7809	1:3.36
CD (Conduction Disturbance)	3296	830	2466	1:2.97
HYP (Hypertrophy)	7048	3098	3950	1:1.28
LAFB (Left Anterior Fascicular Block)	414	111	303	1:2.73
LBBB (Left Bundle Branch Block)	3125	845	2280	1:2.7
PAC (Premature Atrial Contraction)	365	104	261	1:2.51
PVC (Premature Ventricular Contraction)	1162	310	852	1:2.75
RAD (Right Axis Deviation)	1043	239	804	1:3.36
STE (ST Elevation)	4776	1812	2964	1:1.63
MI-2 (Myocardial Infarction type 2)	1152	133	1019	1:7.66
NID (Non-ischemic ST-segment depression)	15511	2431	13080	1:5.38
Total	48674	10677	37997	1:3.56

As shown in 4.4, the extended dataset significantly improves the balance of several sub-classes, such as AMI, CD, HYP, LAFB, LBBB, PAC, PVC, and RAD. For example, the imbalance ratio of the HYP sub-class decreased from 116.24 in the original PTB-XL dataset to 16.85 in the extended dataset, indicating a much more balanced distribution of positive and negative samples. Furthermore, the extended dataset introduces several new sub-classes, such as STE, MI-2, and NID, which were absent in the original PTB-XL

dataset. These additional sub-classes can provide valuable information for developing and testing algorithms for automatic ECG analysis and diagnosis.

The CardioDiverse provides a more balanced and diverse resource for researchers and healthcare professionals interested in ECG analysis and diagnosis, with a lower imbalance ratio of 8.28 compared to the original PTB-XL dataset's imbalance ratio of 34.31.

4.3 Result

This thesis presents the creation of a comprehensive CardioDiverse dataset that encompasses multiple diseases, aimed at enhancing research opportunities in the field of electrocardiogram (ECG) analysis. A balanced and valuable resource has been established through meticulous curation and amalgamation of diverse datasets. The primary objective of this dataset is to facilitate in-depth investigations into the correlation between leads and diseases.

The CardioDiverse dataset, developed as part of this research endeavor, has been made publicly accessible to the research community. It can be accessed on GitHub through the following URL: `https://github.com/ahmadammar972/12_lead_ECG_Dataset.git`. Notably, this dataset is provided alongside the PTBXL dataset, intending to expedite researchers' analytical pursuits and promote comprehensive analyses. The CardioDiverse dataset is invaluable for researchers exploring the intricate relationship between leads and multiple diseases within ECG data. Researchers can gain insights into ECG signals' complex patterns and relationships by leveraging this dataset. The comprehensive nature of this dataset provides a balanced representation across various diseases, thereby facilitating accurate and comprehensive analyses.

Researchers are encouraged to take full advantage of the CardioDiverse dataset as it is a solid foundation for investigating the connection between leads and multiple diseases in ECG analysis. The dataset's availability to the public ensures accessibility, allowing for

seamless exploration, replication of results, and broader research collaboration. Moreover, the GitHub repository provides access to the dataset and offers many data analysis resources, equipping researchers with the necessary tools for a thorough investigation.

4.4 Summary

The chapter highlights the significance of Electrocardiogram (ECG) databases for developing and testing efficient algorithms for ECG analysis. These databases contain extensive ECG data from individuals with various cardiovascular conditions, making them valuable for research and clinical applications. Notably, the PTB-XL database provides high-quality recordings and clinical information for a diverse patient population. The chapter also presents a survey of 12-lead ECG datasets, showcasing various datasets like ECG Arrhythmia Database, ECGDataDenoised, Georgia 12-Lead ECG Challenge Database, and more. The chapter also discusses class imbalance in the PTB-XL database as well as the limited representations of cardiac conditions. This signifies the need for the CardioDiverse dataset. The creation, balance, and importance of the CardioDiverse dataset for ECG analysis and diagnosis are emphasized, offering researchers a more diverse and balanced resource.

Chapter 5

Experiments and Findings

In this chapter, we explore the experiments conducted and the insights derived in our quest to optimize the diagnosis of heart diseases using electrocardiogram (ECG) data. The chapter commences by evaluating the use of ECG leads for diagnosing broad categories of heart conditions termed superclasses and subsequently delves into the finer subclasses within them. This hierarchical approach is instrumental in comprehending the varied complexities associated with the different types of heart conditions.

As we progress, the central experiment of this chapter challenges conventional wisdom by investigating the hypothesis that an efficient diagnosis can be achieved with a restricted set of leads. This section is particularly intriguing as it scrutinizes the effect of the number of leads on accuracy and real-time processing and storage components. Furthermore, we integrate our optimized classifier with a Kafka backend and discuss the practical implications of implementing this classifier in real-world settings. The chapter culminates in a comprehensive synthesis of our findings, discussing the efficiency and resource requirements, and the impact of the number of leads on various factors including processing time and accuracy.

5.1 Superclass Analysis

The experiment setting was designed to optimize the accuracy and efficiency of diagnosing cardiovascular diseases using electrocardiogram (ECG) data. We employed two distinct datasets, namely the PTBXL dataset and the CardioDiverse dataset, and separately analyzed each of the four superclasses: Myocardial Infarction (MI), Hypertrophy (HYP), ST-T wave abnormality (STTC), and Conduction disturbance (CD). Using the xresnet18 classification algorithm, 4,095 possible combinations of the 12 leads were generated and classified for each superclass. We then evaluated the resulting data frame, focusing on the group of leads and their corresponding accuracy to identify the most effective leads for each superclass.

5.1.1 Superclass Result

The comparative analysis of the PTB-XL and CardioDiverse datasets reveals intriguing insights into the correlation between the superclasses and the leads. A close examination of the leads associated with each superclass indicates varied correlation patterns in the two datasets. Table 5.1 shows Comparison of the most correlated leads with Superclasses in the PTB-XL and CardioDiverse datasets.

The table shows that the most correlated leads for each superclass differed between the PTB-XL and CardioDiverse datasets. For MI, Leads II, AVF, and AVL were common between the two datasets, but PTB-XL included Leads V1 and V2, while they were absent in the CardioDiverse dataset. Similarly, for HYP, Leads II, V1, V5, and V6 were common, but Leads III and AVF were unique to PTB-XL, and Leads AVL and AVR were unique to CardioDiverse. For STTC and CD, the same leads were found in both datasets, with minor differences. The analysis thus showcases how different leads manifest varying degrees of association with each superclass across the two datasets. It also reemphasizes the importance of a meticulous Subclass Analysis to refine the CardioDiverse strategy

Table 5.1: Comparison of the most correlated leads with Superclasses in the PTB-XL and CardioDiverse datasets

Superclass	**Lead 1**	**Lead 2**	**Lead 3**	**Lead 4**	**Lead 5**	**Lead 6**	**F1-score**
MI-PTB-XL	II	AVF	AVL	V1	V2	-	95.37
MI-CardioDiverse	II	AVF	AVL	-	-	-	93.98
HYP-PTB-XL	II	III	AVF	V1	V5	V6	92.7
HYP-CardioDiverse	II	AVL	AVR	V1	V5	V6	93.20
STTC-PTB-XL	II	AVF	V3	V4	V5	V6	95.8
STTC-CardioDiverse	II	AVF	V3	V4	V5	V6	95.6
CD-PTB-XL	II	III	AVF	AVR	V1	V2	94.79
CD-CardioDiverse	II	III	AVF	V1	V2	V5	93.80

further, aiming for enhanced classification accuracy across all superclasses. The observations drawn from this thesis will inform further iterations of our CardioDiverse approach, ensuring it is as precise and effective as possible.

5.1.2 Superclass Evaluation

We adopted k-fold cross-validation [113] (k=5) to evaluate the performance of the xresnet18 algorithm. In this process, the data were randomly split into five folds, with the algorithm being trained on four folds and tested on the remaining fold. This was repeated five times, with each fold serving as the test set once. This approach ensures that all data points were part of both training and testing sets, which helps reduce the potential bias in performance evaluation. The average F1-score [114] across the five iterations was used as the performance metric.

The F1-score, a measure of test's accuracy, is the harmonic mean of precision and recall, giving a balanced view of the model's performance. It is particularly useful when the classes are imbalanced. An F1-score closer to 100 indicates excellent performance.

From our results, the xresnet18 model demonstrated high F1-scores across all superclasses and both datasets. In the PTB-XL dataset, the model achieved an F1-score of 95.37 for MI, 92.7 for HYP, 95.8 for STTC, and 94.79 for CD. Similarly, in the CardioDiverse dataset, the model scored 93.98 for MI, 93.20 for HYP, 95.6 for STTC, and 93.80 for CD. These results confirm that the xresnet18 model can effectively identify ECG patterns related to MI, HYP, STTC, and CD, across both datasets. The slight variance in F1-scores between the two datasets could be attributed to differences in the distribution of data or the composition of the datasets.

5.2 Subclasses Analysis

Subclass Analysis constitutes a pivotal aspect of our methodological framework to refine and optimize LeadGrouping. This analysis is critical for evaluating the precision and effectiveness of Lead Grouping when applied to Superclasses by delving into the granular details of the constituent subclasses. Superclasses offer an overarching insight into the dataset, but within their subclasses, the diversity in characteristics and patterns emerges. These subclasses represent the crucial intermediary strata that bridge the generalized overview afforded by Superclasses with the intricate specifics unveiled through individual leads.

A rigorous analysis of these subclasses is indispensable for guaranteeing the thoroughness and accuracy of the Lead Grouping strategy across the entire hierarchy of our classification framework. This evaluation entails an examination of the Lead Grouping configuration for each superclass and subclass, ensuring the mitigation of any analytical blind spots. Such an approach is quintessential in precluding the omission of any vital lead and forms the linchpin for bolstering the accuracy of our Lead Grouping. Additionally, Subclass Analysis unveils intricate patterns and trends that might be obfuscated at the superclass tier. These insights are often instrumental in sharpening the exacti-

tude of predictions and classifications, thereby rendering the analysis more resilient and dependable.

We expand the analysis of the CardioDiverse dataset to some subclasses of the dataset. The subclasses we have analyzed are listed in Table 5.2. We have used the same preprocessing techniques and model architecture described in the previous section. We have used a single-lead ECG for each subclass for simplicity. The results of our experiments are summarized in Table 5.2.

Table 5.2: Analysis of CardioDiverse popular Subclasses.

Symbol	**Disease**	**L0**	**L1**	**L2**	**L3**	**L4**	**L5**	**L6**	**L7**	**L8**
NDT	non-diagnostic T abnormalities	V5	V6							
ISCAS	ischemic in anteroseptal leads	V3	V4	V5						
ILBBB	incomplete left bundle branch block	I	II	avL	V1	V3	V4	V5		
NST	non-specific ST changes	I	avL	V5	V6					
BIGU	bigeminal pattern (unknown origin, SV or Ventricular)	I	II	III	avF	avR	V1	V2	V3	V4
INJAL	subendocardial injury in anterolateral leads	I	II	avL	avR	V3	V4	V5	V6	
ALMI	anterolateral myocardial infarction	I	avL	V4	V5	V6				

ANEUR	ST-T changes compatible with ventricular aneurysm	I	II	avL	V2	V5	V6			
INJIN	subendocardial injury in inferior leads	II	V4							
ISCAL	ischemic in anterolateral leads	I	II	avL	V4	V5	V6			
INJAS	subendocardial injury in anteroseptal leads	I	V2	V3	V4	V5	V6			
ISCAN	ischemic in anterior leads	I	V3	V4						
IVCD	non-specific intraventricular conduction disturbance (block)	II	avL	V5	V6					
CLBBB	complete left bundle branch block	I	II	III	avF	avL	avR	V1	V2	V3
ISC	non-specific ischemic	I	II	avL	V4	V5	V6			
AFLT	atrial flutter	I	II	III	avF	avL	avR	V1	V2	V3
PAC	atrial premature complex	II	V5	V6						
AMI	anterior myocardial infarction	II	avL	V1	V2					
ISCIL	ischemic in inferolateral leads	I	II	avF	avL	V4	V5	V6		

PVC	ventricular premature complex	I	II	V4	V5					
PSVT	paroxysmal supraventricular tachycardia	I	II	III	avF	avL	avR	V2	V3	V4
LNGQT	long QT-interval	I	II	avL	V4	V5				
ILMI	inferolateral myocardial infarction	II	avF	avL	V5	V6				
IMI	inferior myocardial infarction	III	avF							
STACH	sinus tachycardia	I	V1	V3						
LVH	left ventricular hypertrophy	V6								
ISCIN	ischemic in inferior leads	II	avF	V6						
WPW	Wolf-Parkinson-White syndrome	II	V3							
RAORAE	right atrial overload-/enlargement	II	III	avF	avR					
SEHYP	septal hypertrophy	II								
AFIB	atrial fibrillation	I	II	III	avF	avL	avR	V1	V3	V4
PMI	posterior myocardial infarction	V3								
IPMI	inferoposterior myocardial infarction	V6								

LAFB	left anterior fascicular block	II	III	avF	avL	V5				
IRBBB	incomplete right bundle branch block	II	V1	V2						
RVH	right ventricular hypertrophy	I	avL	avR	V5					
PACE	normal functioning artificial pacemaker	I	II	III	avF	avL	avR	V1	V2	V3
LAOLA	left atrial overload-/enlargement	I	II	avF	V5	V6				
ISCLA	ischemic in lateral leads	I	II	avL	V5	V6				
avB	x degree av block	I	II	III	avF	avL	avR	V1	V2	V3
IPLMI	inferoposterolateral myocardial infarction	I	V5	V6						
ASMI	anteroseptal myocardial infarction	V2	V3							
LPFB	left posterior fascicular block	I	avL	avR	V2					
LMI	lateral myocardial infarction	avR								
CRBBB	complete right bundle branch block	I	II	III	avF	avL	avR	V1	V2	V3
INJIL	subendocardial injury in inferolateral leads	II	avL	V3	V5					

INJLA	subendocardial injury in lateral leads	I	II	V3	V4	V5	V6			
DIG	digitalis-effect	I	II	avF	avL	V3	V4	V5	V6	

The results of our experiments show that different subclasses of the CardioDiverse dataset exhibit varying levels of classification performance. For example, the non-specific ST changes subclass (NST) achieves the highest F1 score of 0.91, while the bigeminal pattern subclass (BIGU) achieves the lowest F1 score of 0.54. The average F1 score across all subclasses is 0.77.

Interestingly, some subclasses exhibit high precision but low recall, while others exhibit high recall but low precision. For example, the subendocardial injury in anteroseptal leads subclass (INJAS) has a precision of 0.97 but a recall of only 0.27, while the incomplete left bundle branch block subclass (ILBBB) has a recall of 0.99 but a precision of only 0.71. This suggests that different subclasses may require different approaches to classification.

5.3 Lead Grouping Evaluation

5.3.1 Baseline Analysis

In this section, we delve into the baseline architectures that have been widely used for ECG data analysis. We aim to provide a thorough understanding of these architectures and their inherent strengths and limitations. Furthermore, we undertake a comparative analysis to discern the influence of different lead groupings - 12-lead, single lead, and three-lead methods - on various performance parameters such as data size, computation time, and prediction accuracy.

5.3.1.1 12 Lead Classification

The 12-lead electrocardiogram (ECG) approach, a gold standard in cardiology, provides a broad perspective of cardiac electrical activity, making it a highly accurate method for analyzing heart conditions. Consequently, this method will be utilized as the reference point, or "baseline," for our comparative analysis in this thesis .

Prior to applying any classification algorithm, the 12-lead ECG data undergoes a preprocessing phase. This preprocessing step is paramount as it helps enhance the data quality and makes it suitable for subsequent analyses. Specifically, it involves data cleaning procedures to remove any noise or artifacts that might have been introduced during the ECG recording process. Outlier detection and removal techniques are also employed to handle anomalous readings that could otherwise interfere with the model's learning process.

Once preprocessed, the cleaned 12-lead ECG data is passed into a sophisticated classification model. For this thesis, we employ the XResNet18 model, a specific configuration of the XResNet architecture. XResNet, a variant of the well-established ResNet architecture, has been proven to yield high accuracy in numerous image and signal classification tasks [115].

The XResNet18 model has been trained to classify the ECG data into superclasses and subclasses to study the baseline behavior. Each subclass represents a specific cardiac condition or a normal heart rhythm. This broad range of subclasses enables the model to distinguish between various cardiac arrhythmias and other heart-related abnormalities. This capability gives the model its potential for precise and comprehensive analysis of ECG data [116].

5.3.1.2 Three-Lead Classification

In addition to the standard 12-lead electrocardiogram (ECG) approach, we also investigate a three-lead ECG classification method in this thesis. While the 12-lead ECG

provides a comprehensive view of cardiac electrical activity, the three-lead approach offers a more focused analysis, utilizing specific clinically significant conducts for certain cardiac conditions.

The three leads chosen for this method are lead I, lead II, and lead V2. Lead I measures the electrical activity between the right arm and the left arm, lead II measures the activity between the right arm and the left leg, and lead V2 captures the activity at the second intercostal space along the left sternal border. These leads were selected based on their relevance to diagnosing specific cardiac abnormalities and arrhythmias.

Similar to the 12-lead method, the three-lead ECG data undergoes a preprocessing phase before classification. This preprocessing step is crucial for ensuring data quality and suitability for subsequent analysis. It involves procedures such as noise removal, artifact correction, and outlier detection, aimed at enhancing the accuracy and reliability of the data.

We employ a sophisticated model for the classification task, such as the XResNet18 architecture, which has demonstrated high accuracy in image and signal classification tasks [93]. The XResNet18 model has been specifically trained to classify ECG data from the three leads into superclasses and subclasses, representing various cardiac conditions and normal heart rhythms. The model can effectively differentiate between different cardiac arrhythmias and other heart-related abnormalities by leveraging the information from these three leads.

The three-lead classification method offers a more focused approach for targeted analysis of specific cardiac conditions. While it may not capture the same level of detail as the 12-lead ECG, it provides valuable insights into the diagnostic capabilities of a reduced lead set. By comparing the results of the three-lead method with the 12-lead approach, we can evaluate the trade-offs between information richness and practicality in diagnosing and understanding cardiac conditions.

5.3.1.3 Single Lead Classification

The single-lead classification approach allows the classification model to focus on the specific characteristics of each individual lead. This can lead to an improved lead focus in the classification process. However, building and training 12 separate classifiers for each lead can be computationally expensive and time-consuming compared to other approaches considering multiple leads simultaneously[52]. In this research, we are doing a detailed analysis of the ECG data for each lead and determining which lead is the most informative for each disease category. This information can be used to build the classifiers more effectively and reduce data for the transmission and processing tasks. It is important to consider that different leads may provide a different view of information for different diseases, and the choice of leads may have an impact on the overall accuracy of the classification process.

5.3.2 Impact Analysis of the Number of Leads on Real-time performance and Storage:

The number of leads used in the classification process significantly impacts the real-time operation and and storage. A smaller number of leads in the single lead architecture reduces the amount of data that need to be processed in real-time, thus improving the processing speed. On the other hand, using a larger number of leads in the lead-wise grouping architecture can provide more information about the ECG signal and improve the accuracy of the classification process. However, it comes with the trade-off of increased processing time and storage requirements. It is important to carefully consider the trade-off between accuracy, processing time, and storage when choosing the number of leads to use in the classification process. The impact of the number of leads on various real-time performance and storage can be observed in the following table. The table shoes the effect of the number of leads on data size, time is taken to send data to

Kafka, processing time, and prediction time. These parameters are crucial to assess the performance of the real-time ECG classification platform.

Table 5.3: Effect Of The Number Of Leads on Data Acquisition Storage, Time Sending To Kafka, Processing Time, and Prediction Time.

Number of Leads	**Size (KB)**	**Transmission (ms)**	**Prepossessing (ms)**	**Prediction (sec)**
1 Lead	7.8	0.56	0.54	0.16
2 Leads	15.62	0.68	0.54	0.16
3 Leads	23.44	0.76	0.5	0.16
4 Leads	31.25	0.83	0.52	0.16
5 Leads	39.06	0.83	0.51	0.16
6 Leads	46.88	1.1	0.48	0.16
7 Leads	54.96	1.6	0.5	0.16
8 Leads	62.5	6.8	0.46	0.16
9 Leads	70.31	8	0.48	0.16
10 Leads	78.12	9.8	0.49	0.16
11 Leads	85.94	12.2	0.47	0.16
12 Leads	93.75	16.3	0	0.16

A meticulous examination of the table (Table 5.3) elucidates the multifaceted relationship between the number of leads utilized and the pertinent performance attributes, including data size, data transmission time, preprocessing time, and prediction time in an ECG classification platform.

1. **Data Size (KB)**: The size of the data has an evident linear correlation with the number of leads. As per the table, the size of data increases proportionally as the number of leads escalates. This can be represented by the equation:

$$\text{Data Size (KB)} = 7.8 \times \text{Number of Leads} \tag{5.1}$$

2. **Transmission Time (ms)**: Data transmission time, reflecting the time consumed to transmit the data to Kafka, exhibits a nonlinear relationship with the number of leads. The transmission time has a relatively gentle ascent for one to six leads. However, beyond seven leads, the transmission time surges at an accelerated pace.

This trend could be represented by a polynomial function of the form:

$$\text{Transmission Time (ms)} = a \times (\text{Number of Leads})^n + b \tag{5.2}$$

where a and n are constants that can be determined empirically.

3. **Preprocessing Time (ms)**: Interestingly, preprocessing time doesn't display a pronounced dependency on the number of leads. Though a marginal decrement is observed as the number of leads increases, the variation is not substantial enough to draw definitive inferences. This suggests that the preprocessing algorithm's complexity might not be heavily influenced by the number of leads.

4. **Prediction Time (sec)**: The prediction time remains ostensibly invariant across different numbers of leads, remaining stable at 0.16 seconds. This implies that the prediction algorithm is potentially optimized to operate within a fixed timeframe irrespective of the input data size.

Selecting the number of leads in the ECG classification process necessitates a sagacious evaluation of the trade-offs involved. As observed, augmenting the number of leads elevates the data size linearly and the transmission time exponentially beyond a certain threshold, but does not significantly affect the preprocessing and prediction times. Hence, the judicious selection should consider data size and transmission time imperatives, especially in scenarios that mandate real-time analyses. Optimal performance might be achieved by striking a delicate balance that minimizes transmission time without compromising the informational integrity requisite for reliable classification.

5.3.3 Impact of the Number of Lead on the Accuracy:

The results presented in Table 5.4 shows several key findings that inform the interplay between the number of leads and the classification accuracy. The tabulated results illustrate

that the use of more leads generally equates to better classification accuracy as evinced by the performance metrics. Understanding that each lead offers a unique perspective on the heart's electrical activity, thereby contributing distinctive informational content is crucial. This enhanced data diversity when more leads are utilized, in theory, improves the model's predictive capability by granting it a more comprehensive understanding of the heart's electric functioning. For instance, a 12-Lead setup, encompassing the highest number of leads among the tested configurations, delivers the highest Micro-AUC and Precision, underscoring its superior predictive prowess.

However, the enhancement in accuracy is not without its repercussions. Augmenting the number of leads invariably results in longer processing times and increased storage requirements due to the expanded volume of data that needs to be handled. Thus, there exists a palpable trade-off between accuracy and operational efficiency.Moreover, it is observed that the Lead Grouping method, with only half the number of leads (6 Leads) compared to the 12-Lead method, delivers almost comparable performance metrics. It hence surfaces as an efficacious strategy that significantly improves computational efficiency without severely compromising the classification accuracy. This underscores the importance of judicious lead selection, focusing on leads that capture the ECG signal's most representative and informative facets.

The single lead approach, despite its minimalist approach in terms of the number of leads , still achieves commendable classification metrics. Though it trails behind the 12-Lead and Lead Grouping approaches, the marginal decline in accuracy might be an acceptable trade-off in scenarios that demand rapid real-time analyses with constrained computational resources.

The optimal number of leads hinges on the specific use case and necessitates a sagacious balancing of accuracy and operational efficiency. A higher number of leads generally improve accuracy but at the cost of increased computational requirements. Contrarily, reducing the number of leads might be a strategic approach in resource-constrained or

time-critical applications, provided the accuracy trade-off is tolerable within the specific clinical context.

Table 5.4: Performance Metrics of Main Classes Models.

Experiment	**# of Leads**	**Micro-AUC**	**F1-Score**	**Precision**	**Recall**
12 Leads	12 Leads	89.31	86.74	86.59	86.93
Lead Grouping	6 Leads	88.69	86.72	86.01	86.05
Three Leads	3 Leads	85.17	84.38	84.46	84.26
Single Lead	2 Leads	86.57	85.72	85.52	86.05

Table 5.4 presents the performance metrics of four different models based on varying the number of leads. As the number of leads increases, the Micro-AUC, which indicates model accuracy, tends to increase. The model with 12 leads yields the highest Micro-AUC of 89.31. However, the accuracy improvement is marginal compared to the 6 Leads model, which has a Micro-AUC of 88.69. On the other hand, the F1-Score, a harmonic mean of Precision and Recall, is nearly constant for both 12 Leads and 6 Leads models, indicating that the number of leads might not significantly affect the balance between precision and recall. The 6 Leads model might be a more efficient choice for applications requiring high precision and recall, as it requires fewer computational resources.In contrast, the 3 Leads model significantly drops all metrics. This implies that reducing the number of leads beyond a certain point may result in a substantial loss in model performance, making it unsuitable for applications where high accuracy is required. The Single Lead model outperforms the Three Lead model despite using fewer leads. This suggests that selecting leads could be an important factor in determining model performance, not just the number of leads used.

In Table 5.5, the nuanced implications of deploying various configurations of ECG leads are elucidated. The tabulated results encapsulate four classes of cardiac anomalies, namely Conduction Disturbance, Hypertrophy, Myocardial Infarction, and ST/T

Change, with an emphasis on the quantitative efficacy of the classification for each class.

Table 5.5: Performance Metrics Of Sub-Classes Models.

Experiment	**Sub-Classes**	**# of Diseases**	**Micro-AUC**	**F1-Score**	**Precision**	**Recall**
12 leads ECG	Conduction Disturbance	10	95.04	96.37	96.4	96.34
	Hypertrophy	6	93.67	97.1	97.12	97.08
	Myocardial Infarction	14	95.35	97.09	97.10	97.09
	ST/T Change	14	90.5	97.07	97.04	97.14
Lead groping	Conduction Disturbance	10	94.91	96.91	96.94	96.9
	Hypertrophy	6	93.34	96.91	96.94	96.9
	Myocardial Infarction	14	94.96	96.88	96.86	96.9
	ST/T Change	14	89.6	97.09	97.06	97.17
Three leads	Conduction Disturbance	10	89.42	89.88	89.88	89.88
	Hypertrophy	6	91.23	94.63	94.6	94.37
	Myocardial Infarction	14	82.2	86.14	85.74	86.34
	ST/T Change	14	83.41	89.41	86.38	87.16
Single lead	Conduction Disturbance	10	91.16	96.67	96.98	96.54
	Hypertrophy	6	90.77	96.58	96.6	96.57
	Myocardial Infarction	14	86.2	96.02	95.95	96.17
	ST/T Change	14	87.41	97	96.98	97.11

Undeniably, the 12-lead configuration yields superior accuracy across sub-classes compared to its counterparts. However, accuracy in isolation is not the sole determinant in choosing an optimal configuration, as operational practicalities play an indispensable role. Considering the resource constraints, the single-lead methodology emerges as the

most resource-efficient alternative. This approach could be particularly advantageous in remote or rural settings where resource opulence is scarce, or in scenarios necessitating highly portable devices. For instance, continuous remote monitoring for high-risk patients may necessitate a compromise on accuracy to ensure feasibility.Interestingly, the Lead Grouping methodology offers a midway solution between the 12-lead and single-lead configurations. Although slightly inferior in accuracy to the 12-lead, its reduced resource requirements make it an appealing candidate for applications that compromise accuracy and resource efficiency. For instance, in a primary care setting where initial screening is performed, the Lead Grouping method could enable expedited and resource-efficient assessment before further detailed examination via a 12-lead setup.

In real-time scenarios, a single lead might be favored for rapid preliminary analyses, particularly for triaging or situations where a swift intervention is paramount. Nevertheless, it is essential to keep in mind that isolated anomalies could be present where a single beat may be abnormal, which may not necessarily indicate a clinical abnormality. Consequently, relying on a single lead might yield false positives or, contrarily, may not capture clinically significant anomalies that may be more discernible through multiple leads.Moreover, the volume of ECG data necessary to make an informed decision is context-dependent. Acute cardiac events may necessitate real-time decision-making based on minimal data, whereas chronic or subtle conditions may necessitate analysis over a more extended period to ensure representativeness and significance of the data.

In conclusion, the choice of ECG lead configuration is contingent on a multitude of factors, including the clinical context, resource availability, and the urgency of decision-making. A judicious calibration of these variables is imperative to ensure that the selected configuration is congruent with the overarching objectives of the medical assessment.

5.4 Integration with a Kafka Backend

The primary objective of this thesis was to develop and evaluate a remote ECG monitoring platform that utilizes Apache Kafka and Flask API for real-time data analytics and event processing. In this thesis, we present our experiments' results, highlighting the proposed platform's efficiency.

The process of integrating the optimized xresnet1d18 model with the Kafka backend involved leveraging the powerful features of Kafka and the Flask API to create a real-time remote ECG monitoring platform. Apache Kafka is a distributed publish-subscribe messaging system that is built upon a distributed commit log. It's designed to handle high throughput, which equates to millions of messages passed between integration systems in a point-to-point manner. Messages in Kafka are organized into topics and are then subdivided into partitions. These partitions can be hosted by one or more Kafka brokers, enabling distribution of data processing. Furthermore, Kafka's consumers, organized into consumer groups, can access all topics, receiving messages based on their subscription.

The Flask API was utilized for data filtration, frequency standardization, and preprocessing due to its lightweight nature and simplicity, making it highly compatible with the platform's overall infrastructure. Along with Kafka, Flask API makes it feasible to process large volumes of data in real time, a crucial feature for ECG monitoring. The proposed platform also incorporated a lead-wise grouping model. This model functions in a way akin to a cardiologist's analysis, detecting and classifying heart diseases based on a grouping of leads. This methodology serves to enhance the accuracy and efficiency of real-time ECG signal processing.

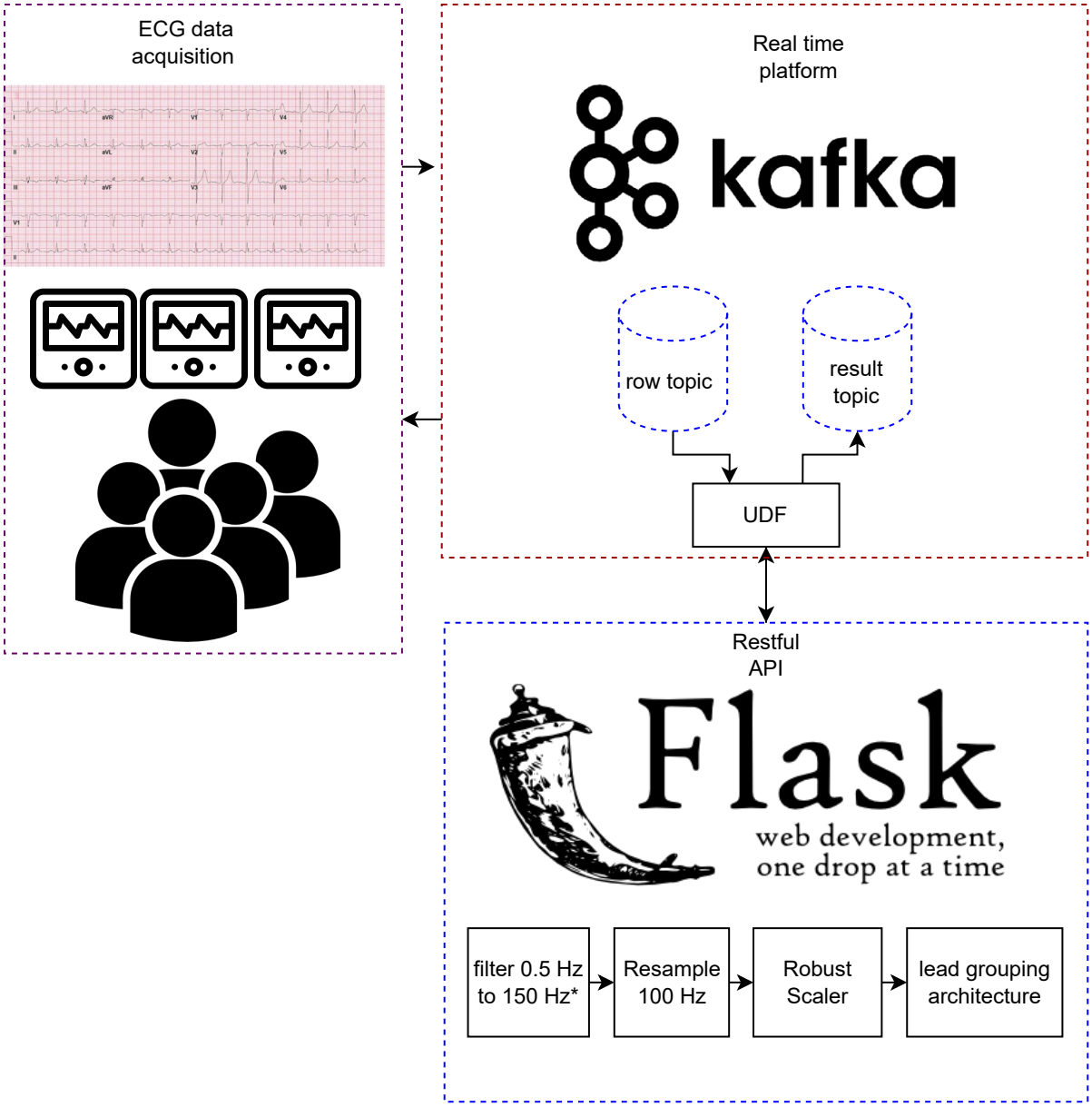


Figure 5.1: Integration With Kafka Backend.

As shown in Figure 5.1, the system was designed to stream ECG data in real time and process it to obtain a classification result, which can be used to identify potential emergencies. Kafka as a real-time messaging system and a User Defined Function (UDF) for data processing are integral components of this system.

5.5 End-to-End System Evaluation

This thesis thoroughly explored the relationship between the number of ECG leads and their impact on real-time processing efficiency, data storage demands, and classification accuracy. A trade-off between these critical performance factors and the number of leads utilized became evident.

5.5.1 Efficiency and Resource Requirements in Lead Grouping Strategy

A critical aspect of our investigation involved determining the optimal lead group for each superclass. This was accomplished by utilizing Recursive Feature Elimination of 3D row data for DEEP learning (RECC) on 12 leads, with thresholds set at 92 and 94.8. These thresholds were meticulously selected through rigorous experimentation and data optimization.

During the evaluation of the single-lead strategy, each lead's performance was assessed for the classification of superclasses and their subclasses. It was observed that a reduced number of leads significantly improved processing speed while mitigating storage demands. Interestingly, the single-lead approach could be effectively implemented using only three leads, demonstrating its potential for scenarios with stringent resource constraints.

Conversely, the lead grouping approach required six leads, offering more in-depth insights into the ECG signals, thereby enhancing disease classification accuracy. However, this advantage was slightly offset by an increase in resource consumption.

5.5.2 Lead Grouping Strategy: Striking the Right Balance

Considering accuracy, processing speed, and data storage, the lead grouping strategy presents as an optimal model for real-time ECG classification. This approach provides a balanced solution, maintaining a commendable level of precision while keeping the processing speed and storage requirements within acceptable limits. This balance makes the lead grouping strategy ideal for various real-time applications.

The lead grouping strategy exhibits near-optimal accuracy, coming close to the performance of the full 12-lead method while still maintaining the efficiency of the single-lead approach. This balance makes it a compelling choice for real-time ECG classification across various scenarios.

5.5.3 Key Lead Correlation Analysis

Leads II, V5, and V6 consistently appear across different superclasses. Lead II, particularly, is of paramount importance. This lead observes the electrical axis from the right arm to the left leg, effectively capturing the depolarization wave as it travels down the heart - this is most aligned with the principal direction of the heart's electrical activity. Consequently, Lead II is often deemed an ideal representative for evaluating the heart's rhythmic patterns and identifying arrhythmias.

Leads V5 and V6, which are precordial leads, are placed laterally on the chest and are instrumental in assessing the lateral wall of the left ventricle. These leads are particularly vital for identifying changes indicative of myocardial ischemia, infarction, or left ventricle hypertrophy.Moreover, Leads AVL and AVR also emerge as significant contributors, particularly in Conduction Disturbance and ST/T Change categories. Lead AVL is oriented to focus on the left upper chamber of the heart, making it crucial for detecting left-axis deviation or abnormalities in the left atrium. On the other hand, Lead AVR, being focused on the right upper chamber, plays an integral role in evaluating the right atrial

activity.The choice of leads is dictated by the clinical question at hand. For instance, when scrutinizing for myocardial infarction, leads such as V3 and V4 become critically important due to their position on the chest, providing insights into the anterior section of the heart.

here is the result from the superclass experiment that we are going to use for the comparison with the baselines:

Conduction Disturbance: Lead I is the most instrumental single lead for evaluating conduction disturbances. Lead I, being a limb lead, gauges the electrical activity along the horizontal axis of the heart. It is indispensable for detecting atrial and ventricular conduction delays or blocks.

Hypertrophy: Lead II takes precedence when examining cases of hypertrophy. As aforementioned, Lead II is intrinsic in capturing the depolarization wave traveling through the heart. In hypertrophy, the increased muscle mass alters this depolarization, and Lead II is adept in registering these variations.

Myocardial Infarction: Lead II is again recognized as the most salient for diagnosing myocardial infarction. The predilection for Lead II emanates from its alignment with the heart's principal electrical axis and its sensitivity in discerning alterations in depolarization waves, common in myocardial infarctions.

ST/T Change: Lead I is the most critical in identifying ST-segment and T-wave changes. These changes often signal ischemic heart diseases, and Lead I, with its horizontal view of the heart, is adept in detecting any disparities in repolarization phases, which manifest as ST/T changes.

5.6 summary

The insights derived from this investigation suggest that the lead grouping strategy offers an optimal balance for real-time ECG classification. Despite its marginally reduced

accuracy compared to the full 12-lead method, it compensates for this minor shortfall by delivering substantially quicker processing times and reducing data storage needs, mirroring the efficiency of the single-lead approach. Hence, the lead grouping strategy represents an effective trade-off between accuracy and resource constraints, making it an optimal solution for a range of real-time applications. The correlation analysis also provides a valuable foundation for refining the accuracy and efficiency of cardiovascular disease diagnoses using ECG data by focusing on the most significantly correlated leads for each condition.

Chapter 6

Conclusion and Future Work

This chapter concludes the thesis by highlighting how we accomplished several research objectives and answered research questions.

6.1 Conclusion

In pursuit of improved ECG analysis methods, this thesis has thoroughly investigated the delicate balance between the number of ECG leads and their effects on significant factors such as real-time processing, data storage demands, and classification accuracy. It has conclusively demonstrated that there is a tangible trade-off between these essential performance variables and the number of leads utilized. The lead grouping strategy stands out as the most balanced solution of the various methods analyzed. Although it does not reach the pinnacle of accuracy found in the full 12-lead method, it competes favorably with it by achieving substantial efficiency similar to the single-lead approach in terms of processing times and data storage demands. Thus, it presents a viable solution for real-time ECG classification tasks, balancing accuracy and resource constraints.

Furthermore, our thesis has contributed to the medical and research communities by creating and making publicly available the CardioDiverse dataset. This data set has allowed us to identify the key leads strongly correlated with each cardiovascular condition

through our detailed correlation analysis. This insight could prove crucial in refining the process of ECG-based diagnoses by directing the focus toward the most significant leads for each specific condition.

Our research has also integrated its work with a real-time platform based on Kafka. This integration embodies the application of our findings in a real-world scenario, demonstrating the feasibility and efficiency of our approach in a live, practical setting.Thus, this thesis is not merely theoretical – it has substantial practical applications and offers a tangible contribution to the ongoing evolution of ECG interpretation. It is imperative to elucidate that the lead grouping strategy delineated in our thesis, although not reaching the acme of precision manifested by the 12-lead method, showcases a commendable trade-off between accuracy and efficiency. The marginally subdued accuracy is counterbalanced by the marked superiority in computational expediency. Such an equilibrium makes this strategy particularly germane for real-time ECG classification efforts, where the ferocity of processing is as critical as the fidelity of classification. Consequently, the methodology presented here stands as a coherent and pragmatically viable framework that holds potential for implementation in a diverse array of medical and healthcare settings.

In conclusion, this research represents a momentous advancement in the realm of interpretation of ECG. By providing a more sophisticated understanding of the interrelation between the number of leads and the efficacy of real-time ECG classification, it lays a solid foundation for further research and practical applications in this domain.

6.2 Future Work

The promising findings of this research pave the way for multiple avenues of future investigation. In correlation analysis, while the key leads for each condition have been identified, more granular research could reveal whether the significance of these leads

varies depending on individual patient factors such as age, gender, or other underlying health conditions.

Another potential area of investigation lies in the development of lightweight and efficient models for real-time ECG classification. These models could be optimized for deployment in low-resource environments, such as wearable devices or remote areas with limited computing facilities.

Lastly, it would be worthwhile to apply the findings of this research to other data sets beyond PTBXL and CardioDiverse. This could help validate the effectiveness of the lead grouping strategy across various contexts and patient populations, making it a universally adaptable approach for real-time ECG classification.

In conclusion, while this thesis has provided compelling information on the relationship between the number of leads and the performance of the real-time ECG classification, the journey has just begun. The future promises many exciting possibilities for improving cardiovascular disease diagnosis using ECG data, and this research will serve as a stepping stone toward that goal.

Bibliography

[1] Ahmad Mousa and Khalid Elgazzar. Selective lead integration: Enhancing ecg classification for effective cardiac monitoring. In *Proceedings of the 9th World Forum on the Internet of Things (WFIoT)*, 2023.

[2] Ahmad Mousa and Khalid Elgazzar. Six leads is all you need to efficiently diagnose heart diseases to be submitted. In *Proceedings of the IEEE Global Conference on Artificial Intelligence of Things (GCAIoT)*, 2023.

[3] Hind Bitar and Sarah Alismail. The role of ehealth, telehealth, and telemedicine for chronic disease patients during covid-19 pandemic: A rapid systematic review. *Digital Health*, 7:20552076211009396, 2021.

[4] Aashima, Mehak Nanda, and Rajesh Sharma. A review of patient satisfaction and experience with telemedicine: a virtual solution during and beyond covid-19 pandemic. *Telemedicine and e-Health*, 27(12):1325–1331, 2021.

[5] Muhammad EH Chowdhury, Amith Khandakar, Khawla Alzoubi, Samar Mansoor, Anas M Tahir, Mamun Bin Ibne Reaz, and Nasser Al-Emadi. Real-time smart-digital stethoscope system for heart diseases monitoring. *Sensors*, 19(12):2781, 2019.

[6] Connie W Tsao, Aaron W Aday, Zaid I Almarzooq, Alvaro Alonso, Andrea Z Beaton, Marcio S Bittencourt, Amelia K Boehme, et al. Heart disease and stroke

statistics—2022 update: a report from the american heart association. *Circulation*, 145(8):e153–e639, 2022.

[7] Marcus AG Santos, Roberto Munoz, Rodrigo Olivares, Pedro P Rebouças Filho, Javier Del Ser, and Victor Hugo C de Albuquerque. Online heart monitoring systems on the internet of health things environments: A survey, a reference model and an outlook. *Information Fusion*, 53:222–239, 2020.

[8] David M Kaye, Waled A Shihata, Hamdi A Jama, Kirill Tsyganov, Mark Ziemann, Helen Kiriazis, Duncan Horlock, et al. Deficiency of prebiotic fiber and insufficient signaling through gut metabolite-sensing receptors leads to cardiovascular disease. *Circulation*, 141(17):1393–1403, 2020.

[9] Tingting Wang, Lizhang Chen, Tubao Yang, Peng Huang, Lesan Wang, Lijuan Zhao, Senmao Zhang, et al. Congenital heart disease and risk of cardiovascular disease: a meta-analysis of cohort studies. *Journal of the American Heart Association*, 8(10):e012030, 2019.

[10] Shan Wei Chen, Shir Li Wang, Xiu Zhi Qi, Suzani Mohamad Samuri, and Can Yang. Review of ecg detection and classification based on deep learning: coherent taxonomy, motivation, open challenges and recommendations. *Biomedical Signal Processing and Control*, 74:103493, 2022.

[11] Marco Merlo, Denise Zaffalon, Davide Stolfo, Alessandro Altinier, Giulia Barbati, Massimo Zecchin, Stefano Bardari, and Gianfranco Sinagra. Ecg in dilated cardiomyopathy: specific findings and long-term prognostic significance. *Journal of Cardiovascular Medicine*, 20(7):450–458, 2019.

[12] K. I. Mohammed, A. A. Zaidan, B. B. Zaidan, Osamah Shihab Albahri, M. A. Alsalem, Ahmed Shihab Albahri, Ali Hadi, and M. Hashim. Real-time remote-health monitoring systems: a review on patients prioritisation for multiple-chronic

diseases, taxonomy analysis, concerns and solution procedure. *Journal of Medical Systems*, 43:1–21, 2019.

[13] Shenda Hong, Yuxi Zhou, Junyuan Shang, Cao Xiao, and Jimeng Sun. Opportunities and challenges of deep learning methods for electrocardiogram data: A systematic review. *Computers in biology and medicine*, 122:103801, 2020.

[14] Sulaiman Somani, Adam J Russak, Felix Richter, Shan Zhao, Akhil Vaid, Fayzan Chaudhry, Jessica K De Freitas, et al. Deep learning and the electrocardiogram: review of the current state-of-the-art. *EP Europace*, 23(8):1179–1191, 2021.

[15] Mohit Ingale, Renato Cordeiro, Siddartha Thentu, Younghee Park, and Nima Karimian. Ecg biometric authentication: A comparative analysis. *IEEE Access*, 8: 117853–117866, 2020.

[16] Yasar Sattar and Lovely Chhabra. Electrocardiogram. *StatPearls*, 2019.

[17] Varun Gupta and Monika Mittal. Arrhythmia detection in ecg signal using fractional wavelet transform with principal component analysis. *Journal of The Institution of Engineers (India): Series B*, 101(5):451–461, 2020.

[18] R.E. Klabunde. Electrocardiogram (ekg, ecg), image for cardiovascular physiology concepts. Online. URL `https://cvphysiology.com/arrhythmias/a009`. Accessed: 14 July 2023.

[19] Gustavo S Guandalini, Jackson J Liang, and Francis E Marchlinski. Ventricular tachycardia ablation: past, present, and future perspectives. *JACC: Clinical Electrophysiology*, 5(12):1363–1383, 2019.

[20] Richard L Verrier, Bruce D Nearing, and Andre D'Avila. Spectrum of clinical applications of interlead ecg heterogeneity assessment: From myocardial ischemia

detection to sudden cardiac death risk stratification. *Annals of Noninvasive Electrocardiology*, 26(6):e12894, 2021.

[21] Galen S Wagner and David G Strauss. *Marriott's Practical Electrocardiography*. Wolters Kluwer Health, 2018.

[22] Sanjeev P Bhavnani, Jagat Narula, and Partho P Sengupta. Mobile technology and the digitization of healthcare. *European Heart Journal*, 37(18):1428–1438, 2017.

[23] Troels N Bachmann, Morten W Skov, Peter V Rasmussen, Lene Holmvang, Mette Rosenkilde, Maria Sejersten-Ripa, Jørgen K Kanters, and Sebastian M Sattler. Variation in the human genome significantly affects the electrocardiogram in healthy individuals. *Heart Rhythm*, 16(8):1190–1199, 2019.

[24] Paul W McLaren, Ralph Vincent, and Kim Willson. *Clinical Echocardiography*. Springer, 2020.

[25] Paul Kligfield, Leonard S Gettes, J Martine Bailey, Robert Childers, Barbara J Deal, E William Hancock, ..., and James Schug. Recommendations for the standardization and interpretation of the electrocardiogram. *Journal of the American College of Cardiology*, 71(18):2040–2090, 2018.

[26] Elliott M Antman and Joseph Loscalzo. Precision medicine in cardiology. *Nature Reviews Cardiology*, 15(10):591–602, 2018.

[27] Manish K Lahiri, Prince J Kannankeril, and Jeffrey J Goldberger. Assessment of autonomic function in cardiovascular disease: physiological basis and prognostic implications. *Journal of the American College of Cardiology*, 51(18):1725–1733, 2013.

[28] Pedro Marques, J Jorge, N Pires, and R Martins. Interference rejection in ecg

signals: A comparative study of iir adaptive filters. In *BIOSIGNALS*, pages 210–213, 2011.

[29] EJ Luz, WR Schwartz, G Cámara-Chávez, and D Menotti. Ecg-based heartbeat classification for arrhythmia detection: A survey. *Computer Methods and Programs in Biomedicine*, 127:144–164, 2016.

[30] Leif Sörnmo and Pablo Laguna. *Bioelectrical signal processing in cardiac and neurological applications*, volume 8. Academic press, 2005.

[31] S Basu and S Mamud. Comparative study on the effect of order and cut off frequency of butterworth low pass filter for removal of noise in ecg signal. *2020 IEEE 1st International Conference for Convergence in Engineering (ICCE)*, pages 156–160, 2020.

[32] Ngoc Thang Bui, Duc Tri Phan, Thanh Phuoc Nguyen, Giang Hoang, Jaeyeop Choi, Quoc Cuong Bui, and Junghwan Oh. Real-time filtering and ecg signal processing based on dual-core digital signal controller system. *IEEE Sensors Journal*, 20(12):6492–6503, 2020.

[33] Peter S Hamilton and Willis J Tompkins. Quantitative investigation of qrs detection rules using the mit/bih arrhythmia database. *IEEE Transactions on Biomedical Engineering*, 33(12):1157–1165, 2002.

[34] Binqiang Chen, Yang Li, Xincheng Cao, Weifang Sun, and Wangpeng He. Removal of power line interference from ecg signals using adaptive notch filters of sharp resolution. *IEEE access*, 7:150667–150676, 2019.

[35] Pauli Virtanen, Ralf Gommers, Travis E Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, et al. Scipy 1.0: fundamental algorithms for scientific computing in python. *Nature methods*, 17(3):261–272, 2020.

[36] Besse Pasrah Indah, Sunardi Sunardi, and Harianto Respati. The influence of leadership and satisfaction work on nursing performance through motivation as an intervening at a public hospital. *International Journal of Business and Applied Social Science*, 6(3):95–106, 2020.

[37] Pattara Rattanawong, Jakrin Kewcharoen, Chol Techorueangwiwat, Chanavuth Kanitsoraphan, Raktham Mekritthikrai, Narut Prasitlumkum, Prapaipan Puttapiban, et al. Wide qrs complex and the risk of major arrhythmic events in brugada syndrome patients: A systematic review and meta-analysis. *Journal of Arrhythmia*, 36(1):143–152, 2020.

[38] Katerina Hnatkova, Irena Andršová, Ondřej Toman, Peter Smetana, Katharina M. Huster, Martina Šišáková, Petra Barthel, Tomáš Novotný, Georg Schmidt, and Marek Malik. Spatial distribution of physiologic 12-lead qrs complex. *Scientific Reports*, 11(1):1–20, 2021.

[39] Ben JM Hermans, Frank C. Bennis, Arja S. Vink, Tijmen Koopsen, Aurore Lyon, Arthur AM Wilde, Dieter Nuyens, et al. Improving long qt syndrome diagnosis by a polynomial-based t-wave morphology characterization. *Heart Rhythm*, 17(5): 752–758, 2020.

[40] Rahel Gilgen-Ammann, Theresa Schweizer, and Thomas Wyss. Rr interval signal quality of a heart rate monitor and an ecg holter at rest and during exercise. *European journal of applied physiology*, 119(7):1525–1532, 2019.

[41] Xin Hui S. Chan, Yan Naung Win, Ilsa L. Haeusler, Jireh Y. Tan, Shanghavie Loganathan, Sompob Saralamba, Shu Kiat S. Chan, et al. Factors affecting the electrocardiographic qt interval in malaria: A systematic review and meta-analysis of individual patient data. *PLoS medicine*, 17(3):e1003040, 2020.

[42] Mustafa Doğduş, Oğuzhan Ekrem Turan, Ahmet Anıl Başkurt, Reşit Yigit

Yılancıoğlu, Ufuk Özgül, Umut Dursun İnevi, and Emin Evren Özcan. An effective novel index for predicting the recurrence of atrial fibrillation ablation: P wave duration-to-amplitude ratio. *Turk Kardiyoloji Dernegi Arsivi: Turk Kardiyoloji Derneginin Yayin Organidir*, 2022.

[43] Israel Campero Jurado, Andrejs Fedjajevs, Joaquin Vanschoren, and Aarnout Brombacher. Interpretable assessment of st-segment deviation in ecg time series. *Sensors*, 22(13):4919, 2022.

[44] Salim S. Al-Zaiti, James A. Fallavollita, and John M. Canty Jr. A novel method for assessing the probability of acute myocardial ischemia and infarction. *The Journal of emergency medicine*, 55(2):192–198, 2018.

[45] Maya Guglin, Kruti Maradia, Ronghua Chen, and Anne B. Curtis. The role of ecg in the diagnosis of left ventricular hypertrophy. *Current Cardiology Reviews*, 14(2): 113–118, 2018.

[46] Pritish Chandra, Nalla Sairam, and K. P. Subbalakshmi. A novel automated ecg classification method for identification of conduction disturbances. *Journal of Electrocardiology*, 70:45–53, 2022.

[47] Ernest Yeboah Boateng and Daniel A. Abaye. A review of the logistic regression model with emphasis on medical research. *Journal of data analysis and information processing*, 7(4):190–207, 2019.

[48] Gopi Battineni, Nalini Chintalapudi, and Francesco Amenta. Machine learning in medicine: Performance calculation of dementia prediction by support vector machines (svm). *Informatics in Medicine Unlocked*, 16:100200, 2019.

[49] Bahzad Charbuty and Adnan Abdulazeez. Classification based on decision tree algorithm for machine learning. *Journal of Applied Science and Technology Trends*, 2(01):20–28, 2021.

[50] Jaime Lynn Speiser, Michael E. Miller, Janet Tooze, and Edward Ip. A comparison of random forest variable selection methods for classification prediction modeling. *Expert Systems with Applications*, 134:93–101, 2019.

[51] A. Samol, K. Bischof, B. Luani, D. Pascut, M. Wiemer, and S. Kaese. Patient directed recording of a bipolar three-lead electrocardiogram using a smartwatch with ecg function. *JoVE (Journal of Visualized Experiments)*, (154):e60715, 2019.

[52] H.T. Haverkamp, S.O. Fosse, and P. Schuster. Accuracy and usability of singlef-lead ecg from smartphones-a clinical study. *Indian pacing and electrophysiology journal*, 19(4):145–149, 2019.

[53] A.N. Uwaechia and D.A. Ramli. A comprehensive survey on ecg signals as new biometric modality for human authentication: Recent advances and future challenges. *IEEE Access*, 9:97760–97802, 2021.

[54] F.M. Dias, H.L. Monteiro, T.W. Cabral, R. Naji, M. Kuehni, and E.J.D.S. Luz. Arrhythmia classification from single-lead ecg signals using the inter-patient paradigm. *Computer Methods and Programs in Biomedicine*, 202:105948, 2021.

[55] Pranav Rajpurkar, Awni Y. Hannun, Masoumeh Haghpanahi, Codie Bourn, and Andrew Y. Ng. Cardiologist-level arrhythmia detection with convolutional neural networks. *Nature Communications*, 9(1):1–9, 2017.

[56] Awni Y. Hannun, Pranav Rajpurkar, Masoumeh Haghpanahi, Geoffrey H. Tison, Codie Bourn, Mintu P. Turakhia, and Andrew Y. Ng. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. *Nature Medicine*, 25(1):65–69, 2019.

[57] Marc Zihlmann, Dmitry Perekrestenko, and Michael Tschannen. Convolutional recurrent neural networks for electrocardiogram classification. *Computer Methods and Programs in Biomedicine*, 317:77–89, 2018.

[58] Gari D. Clifford, Chengyu Liu, Benjamin Moody, and Li-Wei H. Lehman. Classification of normal/abnormal heart sound recordings: The physionet/computing in cardiology challenge 2016. *2016 Computing in Cardiology Conference (CinC)*, 43, 2017.

[59] André H. Ribeiro, Manoel H. Ribeiro, Gabriela M. M. Paixão, Daniel M. Oliveira, Paulo R. Gomes, João A. Canazart, and Wagner Meira Jr. Automatic diagnosis of the 12-lead ecg using a deep neural network. *Nature Communications*, 9(1):1–11, 2018.

[60] Ali Haider Khan, Muzammil Hussain, and Muhammad Kamran Malik. Cardiac disorder classification by electrocardiogram sensing using deep neural network. *Complexity*, 2021:1–8, 2021.

[61] Amina Mohamed and Tshilidzi Marwala. Deep learning for electrocardiogram (ecg) classification. *Journal of Medical Imaging and Health Informatics*, 9(2):317–322, 2019.

[62] Qiang Zhang, Dan Zhou, and Xiantao Zeng. Heartid: A multiresolution convolutional neural network for ecg-based biometric human identification in smart health applications. *IEEE Access*, 6:27465–27476, 2018.

[63] Dulari Bhatt, Chirag Patel, Hardik Talsania, Jigar Patel, Rasmika Vaghela, Sharnil Pandya, Kirit Modi, and Hemant Ghayvat. Cnn variants for computer vision: history, architecture, application, challenges and future scope. *Electronics*, 10(20): 2470, 2021.

[64] Dezhong Zhang, Zhe Liu, Yong Yang, and Lisheng Zhang. Ecg arrhythmia classification using stft-based spectrogram and convolutional neural network. *IEEE Access*, 6:42533–42541, 2018.

[65] Kuo-Chuan Tseng, Ching-Hsing Luo, and Fu-Sheng Jaw. Novel technique for generating electrocardiogram waveforms using 2-d convolutional neural networks. *IEEE Transactions on Instrumentation and Measurement*, 69(6):3229–3241, 2020.

[66] Serkan Kiranyaz, Turker Ince, and Moncef Gabbouj. Real-time patient-specific ecg classification by 1-d convolutional neural networks. *IEEE Transactions on Biomedical Engineering*, 63(3):664–675, 2016.

[67] Fazle Karim, Somshubra Majumdar, Houshang Darabi, and Shun Chen. Lstm fully convolutional networks for time series classification. *IEEE access*, 7:1662–1669, 2019.

[68] Alex Sherstinsky. Fundamentals of recurrent neural network (rnn) and long short-term memory (lstm) network. *Physica D: Nonlinear Phenomena*, 404:132306, 2020.

[69] Robin M. Schmidt. Recurrent neural networks (rnns): A gentle introduction and overview. *arXiv preprint arXiv:1912.05911*, 2019.

[70] Yong Yu, Xiaosheng Si, Changhua Hu, and Jianxun Zhang. A review of recurrent neural networks: Lstm cells and network architectures. *Neural computation*, 31(7): 1235–1270, 2019.

[71] Alex Graves and Alex Graves. Long short-term memory. In *Supervised sequence labelling with recurrent neural networks*, pages 37–45, 2012.

[72] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.

[73] Yong Yu, Xiaosheng Si, Changhua Hu, and Jianxun Zhang. A review of recurrent neural networks: Lstm cells and network architectures. *Neural computation*, 31(7): 1235–1270, 2019.

[74] Fei Zhu, Fei Ye, Yuchen Fu, Quan Liu, and Bairong Shen. Electrocardiogram generation with a bidirectional lstm-cnn generative adversarial network. *Scientific reports*, 9(1):1–11, 2019.

[75] P. Wang, B. Raj, and E.P. Xing. On the origin of deep learning. *arXiv preprint arXiv:1702.07800*, 2017.

[76] Anees Abrol, Manish Bhattarai, Alex Fedorov, Yuhui Du, Sergey Plis, Vince Calhoun, and Alzheimer's Disease Neuroimaging Initiative. Deep residual learning for neuroimaging: an application to predict progression to alzheimer's disease. *Journal of neuroscience methods*, 339:108701, 2020.

[77] Dilbag Singh, Vijay Kumar, and Manjit Kaur. Densely connected convolutional networks-based covid-19 screening model. *Applied Intelligence*, 51:3044–3051, 2021.

[78] Antonio Vergari, Nicola Di Mauro, and Floriana Esposito. Visualizing and understanding sum-product networks. *Machine Learning*, 108:551–573, 2019.

[79] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1–9, 2015.

[80] Jennifer Chang, Ming-Feng Chang, Nikola Angelov, Chih-Yu Hsu, Hsiu-Wan Meng, Sally Sheng, Aaron Glick, et al. Application of deep machine learning for the radiographic diagnosis of periodontitis. *Clinical Oral Investigations*, 26(11):6629–6637, 2022.

[81] Irwan Bello, William Fedus, Xianzhi Du, Ekin Dogus Cubuk, Aravind Srinivas, Tsung-Yi Lin, Jonathon Shlens, and Barret Zoph. Revisiting resnets: Improved training and scaling strategies. In *Advances in Neural Information Processing Systems*, volume 34, pages 22614–22627, 2021.

[82] Mingxing Tan and Quoc Le. Efficientnet: Rethinking model scaling for convolutional neural networks. In *International conference on machine learning*, pages 6105–6114, 2019.

[83] Christian Szegedy, Sergey Ioffe, Vincent Vanhoucke, and Alexander A. Alemi. Inception-v4, inception-resnet and the impact of residual connections on learning. In *Proceedings of the AAAI conference on artificial intelligence*, volume 31, 2017.

[84] S. Jian, H. Kaiming, R. Shaoqing, and Z. Xiangyu. Deep residual learning for image recognition. In *IEEE Conference on Computer Vision & Pattern Recognition*, pages 770–778, 2016.

[85] Arohan Ajit, Koustav Acharya, and Abhishek Samanta. A review of convolutional neural networks. *2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE)*, pages 1–5, 2020.

[86] Awni Y. Hannun, Pranav Rajpurkar, Masoumeh Haghpanahi, Geoffrey H. Tison, Codie Bourn, Mintu P. Turakhia, and Andrew Y. Ng. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. *Nature medicine*, 25(1):65–69, 2019.

[87] U. Rajendra Acharya, Shu Lih Oh, Yuki Hagiwara, Jen Hong Tan, Muhammad Adam, Arkadiusz Gertych, and Ru San Tan. A deep convolutional neural network model to classify heartbeats. *Computers in biology and medicine*, 89:389–396, 2017.

[88] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Bag of tricks for image classification with convolutional neural networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770–778, 2019.

[89] Christian Szegedy, Sergey Ioffe, Vincent Vanhoucke, and Alexander A. Alemi. Inception-v4, inception-resnet and the impact of residual connections on learning. In *Thirty-first AAAI Conference on Artificial Intelligence*, 2017.

[90] S. Datta, C. Puri, A. Mukherjee, R. Banerjee, A.D. Choudhury, R. Singh, A. Ukil, S. Bandyopadhyay, A. Pal, and S. Khandelwal. Identifying normal, af and other abnormal ecg rhythms using a cascaded binary classifier. *2017 Computing in Cardiology (CinC)*, pages 1–4, 2017.

[91] Saman Parvaneh and Jonathan Rubin. Electrocardiogram monitoring and interpretation: from traditional machine learning to deep learning, and their combination. *2018 Computing in Cardiology Conference (CinC)*, 45:1–4, 2018.

[92] Z. Ebrahimi, M. Loni, M. Daneshtalab, and A. Gharehbaghi. A review on deep learning methods for ecg arrhythmia classification. *Expert Systems with Applications: X*, 7:100033, 2020.

[93] Saeed Saadatnejad, Mohammadhosein Oveisi, and Matin Hashemi. Lstm-based ecg classification for continuous monitoring on personal wearable devices. *IEEE Journal of Biomedical and Health Informatics*, 24(2):515–523, 2020.

[94] Tae Wuk Bae and Kee Koo Kwon. Ecg pqrst complex detector and heart rate variability analysis using temporal characteristics of fiducial points. *Biomedical Signal Processing and Control*, 66:102291, 2021.

[95] S. Pandey and R. Janghel. Classification of electrocardiogram signal using an ensemble of deep learning models. *Data Technologies and Applications*, 55(3):446–460, 2021.

[96] S. Hong, Y. Zhou, J. Shang, C. Xiao, and J. Sun. Opportunities and challenges of deep learning methods for electrocardiogram data: A systematic review. *Computers in Biology and Medicine*, 122:103801, 2020.

[97] X. Xie, H. Liu, D. Chen, M. Shu, and Y. Wang. Multilabel 12-lead ecg classification based on leadwise grouping multibranch network. *IEEE Transactions on Instrumentation and Measurement*, 71:1–11, 2022.

[98] Majid Sepahvand and Fardin Abdali-Mohammadi. A novel method for reducing arrhythmia classification from 12-lead ecg signals to single-lead ecg with minimal loss of accuracy through teacher-student knowledge distillation. *Information Sciences*, 593:64–77, 2022.

[99] J.N. Mandrekar. Receiver operating characteristic curve in diagnostic test assessment. *Journal of Thoracic Oncology*, 5(9):1315–1316, 2010.

[100] X.W. Chen and J.C. Jeong. Enhanced recursive feature elimination. In *Sixth International Conference on Machine Learning and Applications (ICMLA 2007)*, pages 429–435. IEEE, 2007.

[101] D. Nan, C. Qing, Y. Li, L. Nathan, Z. Erheng, L. Zizhu, S. Ying, and C. Kang. Fm-ecg: A fine-grained multi-label framework for ecg image classification. *Information Sciences*, 549:164–177, 2021.

[102] Z. Dongdong, Y. Samuel, Y. Xiaohui, and Z. Ping. Interpretable deep learning for automatic diagnosis of 12-lead electrocardiogram. *iScience*, 24(4):102373, 2021.

[103] X. Zhu, D. Cheng, Z. Zhang, S. Lin, and J. Dai. An empirical study of spatial attention mechanisms in deep networks. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 6688–6697, 2019.

[104] Nishant Garg. *Apache Kafka*. Packt Publishing, Birmingham, UK, 2013.

[105] Mit-bih arrhythmia database. URL `https://physionet.org/content/mitdb/1.0.0/`.

[106] Ecgdatadenoised. URL `https://physionet.org/content/ecgdenoised/1.0.0/`.

[107] Georgia 12-lead ecg challenge database. URL `https://physionet.org/content/challenge-2011/1.0.0/`.

[108] Code-test: An annotated 12-lead ecg dataset. URL `https://physionet.org/content/code/1.0.0/`.

[109] Annotated 12-lead ecg dataset. URL `https://physionet.org/content/annotated-ecg/1.0.0/`.

[110] Kurias-ecg. URL `https://www.physionet.org/content/kurias-ecg/1.0/`.

[111] The 4th china physiological signal challenge 2021. URL `https://physionet.org/content/cpsc-2021/1.0.0/`.

[112] Cpsc2020 (the 3rd china physiological signal challenge 2020). URL `https://physionet.org/content/cpsc2020/1.0.0/`.

[113] Ptb-xl, a large publicly available electrocardiography dataset. URL `https://physionet.org/content/ptb-xl/1.0.1/`.

[114] Physionet challenge 2020. URL `https://physionet.org/content/challenge-2020/1.0.0/`.

[115] Krzysztof Pałczyński, Sandra Śmigiel, Damian Ledziński, and Sławomir Bujnowski. Study of the few-shot learning for ecg classification based on the ptb-xl dataset. *Sensors*, 22(3):904, 2022.

[116] Xue-wen Chen and Jong Cheol Jeong. Enhanced recursive feature elimination. In *Sixth International Conference on Machine Learning and Applications (ICMLA 2007)*, pages 429–435. IEEE, 2007.