Thesis submitted to the

UNIVERSITY OF MOHAMED BOUDIAF – MSILA

FACULTY OF MATHEMATICS AND COMPUTER SCIENCE
DEPARTMENT OF COMPUTER SCIENCE

In partial fulfillment of the requirements for the degree of

Master in **Computer science**

By

Mohadi, Yassine

Title of the thesis

12-Lead ECG Heartbeat Classification Using the Progressive Moving Average Transform (PMAT) and Convolutional Neural Networks

Under the supervision of

Rabah Mokhtari

Composition of the jury

Dr. Bahache Mohamed	University of Msila	President
Dr. Mokhtari Rabah	University of Msila	Reporter
Dr. Kamel Mohamed	University of Msila	Examiner

25, May, 2026.

Abstract

This study presents an automated method for classifying electrocardiogram (ECG) heartbeats into Normal, Supraventricular, and Ventricular categories. To improve diagnostic reliability, we developed a hybrid CNN-MLP model that processes two ECG leads simultaneously. The CNN extracts morphological features from PMAT-transformed images, while the MLP analyzes RR-interval timing. In this experiment, we systematically trained every possible dual-lead combination to identify the optimal pairings. Results show that leads 4 and 10 achieve the best performance for critical arrhythmia detection, reaching 98.52% accuracy. Meanwhile, leads 7 and 10 provide the most reliable configuration for general monitoring, achieving 97.31% accuracy. Despite challenges with rare Supraventricular beats, this approach effectively combines waveform shape and rhythm patterns, offering a practical tool for automated ECG analysis.

Key words: ECG Classification; Heartbeat Analysis; Deep Learning; Dual-Lead ECG; CNN-MLP Hybrid; Arrhythmia Detection.

الملخص

تقدم هذه الدراسة طريقة آلية لتصنيف نبضات تخطيط القلب الكهربائي (ECG) إلى ثلاث فئات: النبضات الطبيعية، والنبضات فوق البطينية، والنبضات البطينية. ولتحسين موثوقية التشخيص، قمنا بتطوير نموذج هجين يجمع بين الشبكات العصبية الالتفافية (CNN) والشبكات متعددة الطبقات(MLP) ، حيث يعالج إشارتين مختلفتين من تخطيط القلب في الوقت نفسه. تقوم شبكة CNN باستخراج الخصائص الشكلية من الصور المحولة باستخدامPMAT ، بينما تقوم شبكة MLP بتحليل الفواصل الزمنيةRR .

في هذه التجربة، قمنا بتدريب جميع التركيبات الممكنة المكونة من زوجين من الأقطاب الكهربائية بهدف تحديد أفضل الأزواج أداءً. أظهرت النتائج أن القطبين 4 و10 يحققان أفضل أداء في اكتشاف اضطرابات النظم القلبية الحرجة، بدقة بلغت 98.52٪. في المقابل، يوفر القطبان 7 و10 أكثر التكوينات موثوقية للمراقبة العامة، بدقة وصلت إلى 97.31٪.

وعلى الرغم من التحديات المرتبطة بندرة النبضات فوق البطينية، فإن هذا النهج ينجح في الدمج بين شكل الموجة القلبية وأنماط الإيقاع، مما يوفر أداة عملية وفعالة للتحليل الآلي لتخطيط القلب الكهربائي.

الكلمات المفتاحية: تصنيف تخطيط القلب الكهربائي؛ تحليل نبضات القلب؛ التعلم العميق؛ إشارات ثنائية القناة؛ نموذج CNN-MLP الهجين؛ كشف اضطرابات النظم القلبي.

Contents

List of figures

List of tables

Introduction

Electrocardiogram (ECG) analysis remains a cornerstone in cardiology, providing essential insights into heart rhythm and function. However, manual interpretation of long-term ECG recordings is highly time-consuming, subject to human fatigue, and often affected by inter-observer variability. With the increasing availability of continuous cardiac monitoring, there is a growing need for reliable, automated systems that can assist clinicians in detecting arrhythmias quickly and accurately. This thesis addresses this challenge by developing an artificial intelligence-based framework for automated ECG heartbeat classification. We propose a hybrid deep learning architecture that combines a Convolutional Neural Network (CNN) with a Multi-Layer Perceptron (MLP) to process dual-lead ECG signals. The CNN extracts morphological features from two-dimensional representations generated by the Progressive Moving Average Transform (PMAT), while the MLP analyzes temporal patterns derived from RR-interval dynamics. A core objective of this research is to systematically evaluate all possible lead pair combinations to identify the most effective configurations for clinical diagnosis. The study focuses on classifying heartbeats into three critical categories: Normal, Supraventricular, and Ventricular. By leveraging the complementary strengths of shape and timing analysis, our approach aims to improve diagnostic precision while maintaining computational efficiency. The following chapters detail the medical background, the proposed methodology, the experimental results, and the technical environment used throughout the project. Together, they provide a comprehensive and reproducible framework for automated, dual-lead ECG classification.

Chapter 1 Introduction to ECG Analysis and Project Overview

1.1 Medical Background

This chapter provides the foundational background for understanding automated ECG heartbeat classification and the motivation behind this project. It begins with an overview of cardiac electrophysiology and the standard ECG signal, explaining the key waves, intervals, and lead systems used in clinical practice. The chapter then discusses the main heartbeat classes relevant to this study and highlights the challenges of manual ECG interpretation. Following this, we explore the evolving role of artificial intelligence in cardiac diagnosis, tracing the progression from traditional rule-based systems to modern deep learning approaches. Finally, the chapter presents the project's core concept—transforming dual-lead ECG signals into 2D representations for hybrid CNN-MLP classification—and outlines the specific objectives and expected contributions of this work.

1.1.1. Introduction to Cardiac Electrophysiology

The human heart functions as a sophisticated muscular pump that maintains blood circulation throughout the body. Its rhythmic contractions are controlled by a specialized electrical conduction system that generates and propagates electrical impulses in a coordinated manner. Understanding this electrical activity is fundamental to interpreting electrocardiogram (ECG) signals and diagnosing cardiac abnormalities.

The Cardiac Conduction System

The electrical impulse originates in the Sinoatrial (SA) Node, located in the right atrium near the entrance of the superior vena cava. The SA node acts as the heart's natural pacemaker, spontaneously generating electrical impulses at a rate of 60-100 beats per minute in healthy adults [2]. These impulses spread through both atria, causing atrial contraction and blood ejection into the ventricles.

The electrical signal then reaches the Atrioventricular (AV) Node, positioned at the junction between the atria and ventricles. The AV node serves two critical functions:

- It introduces a physiological delay of approximately 0.1 seconds, allowing complete ventricular filling before contraction;

- It acts as a gatekeeper, limiting the number of impulses that reach the ventricles during rapid atrial rhythms [3].

After the AV node, the impulse travels rapidly through the Bundle of His, which divides into right and left bundle branches along the interventricular septum. Finally, the electrical signal spreads through the Purkinje Fiber Network, causing synchronized ventricular contraction that pumps blood to the lungs and systemic circulation.

Electrophysiological Basis of the ECG

The ECG records the electrical potentials generated by cardiac depolarization and repolarization. When cardiac cells are activated, ion channels open, allowing sodium, potassium, and calcium ions to flow across cell membranes. This ionic movement creates electrical currents that propagate through the heart tissue and can be detected by electrodes placed on the skin surface [3].

The amplitude and direction of these electrical signals depend on:

- The mass of activated cardiac tissue;
- The speed of depolarization;
- The orientation of the electrical vector relative to the recording electrode;
- The distance between the heart and the electrode.

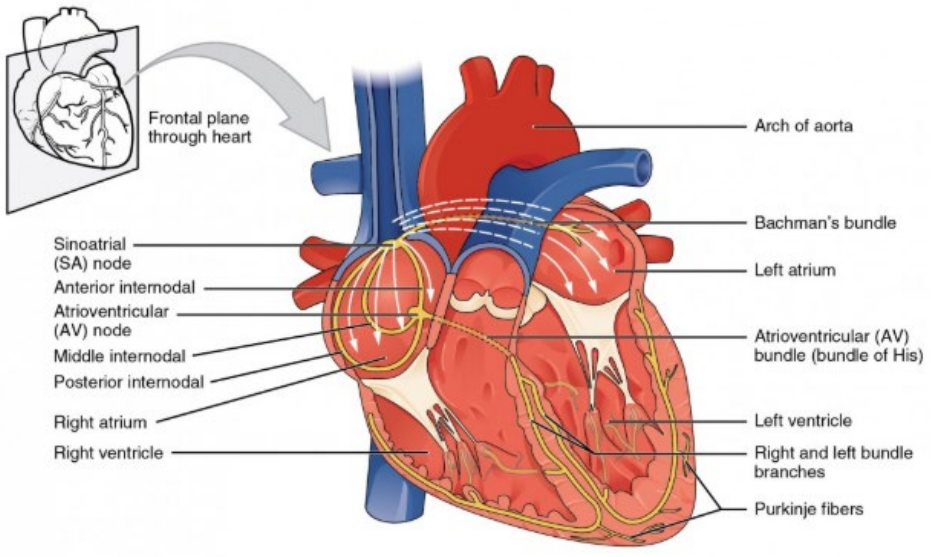

Figure 1: Electrical Conduction Pathway of the Heart.

1.1.2. The ECG Signal and Its Components

The electrocardiogram (ECG) is a graphical representation of the heart's electrical activity recorded over time. It provides valuable information about cardiac rhythm, conduction abnormalities, and myocardial health. Understanding the individual components of the ECG waveform is essential for both manual interpretation and automated classification systems.

The Cardiac Cycle and ECG Waves

A normal ECG heartbeat consists of several characteristic waves that correspond to specific electrical events in the heart. These waves are labeled P, Q, R, S, and T, and they appear in a consistent sequence during each cardiac cycle [4].

P Wave

The P wave represents atrial depolarization, which occurs when the electrical impulse spreads from the SA node through both atria, causing them to contract. In a normal sinus rhythm, the P wave has the following characteristics:

- **Duration:** 120-80 milliseconds;
- **Amplitude:** Less than 0.25 millivolts (mV) in limb leads;
- **Morphology:** Smooth, rounded, and upright in leads I, II, and aVF.

The P wave is typically the first deflection seen in the cardiac cycle and indicates that the heartbeat originated from the SA node. Abnormal P wave morphology or timing can suggest atrial enlargement or ectopic atrial rhythms [5].

QRS Complex

The QRS complex represents ventricular depolarization, which is the rapid spread of the electrical impulse through the ventricles via the Bundle of His and Purkinje fibers. This complex is the most prominent feature of the ECG due to the large muscle mass of the ventricles.

The QRS complex consists of three components:

- **Q wave:** The first negative deflection after the P wave (if present);
- **R wave:** The first positive deflection;
- **S wave:** The negative deflection following the R wave.

Normal QRS characteristics:

- **Duration:** 80-120 milliseconds (less than 100 milliseconds is considered narrow);
- **Amplitude:** Varies by lead; typically 0.5-2.0 mV in limb leads;
- **Morphology:** Can vary significantly between leads depending on the electrical axis.

A prolonged QRS duration (>120 milliseconds) may indicate bundle branch block or ventricular conduction delay, while abnormal Q waves can suggest previous myocardial infarction [6].

T Wave

The T wave represents ventricular repolarization, which is the recovery phase when the ventricles prepare for the next heartbeat. The T wave has the following normal characteristics:

- **Duration:** Approximately 160 milliseconds;
- **Amplitude:** Usually less than 0.5 mV in limb leads and less than 1.0 mV in precordial leads;
- **Morphology:** Asymmetric with a gradual upslope and a steeper downslope;
- **Direction:** Normally upright in leads I, II, and V3-V6.

ECG Intervals and Segments

In addition to the waves, specific intervals and segments on the ECG provide important timing information about cardiac conduction.

PR Interval

The PR interval is measured from the beginning of the P wave to the beginning of the QRS complex. It represents the time required for the electrical impulse to travel from the SA node through the atria, AV node, and His-Purkinje system to the ventricles.

Normal values:

- Duration: 120-200 milliseconds (0.12-0.20 seconds);
- Clinical significance: Prolonged PR interval (>200 milliseconds) indicates first-degree AV block and Short PR interval (<120 milliseconds) may suggest pre-excitation syndromes like Wolff-Parkinson-White syndrome [6]

QT Interval

The QT interval is measured from the beginning of the QRS complex to the end of the T wave. It represents the total time for ventricular depolarization and repolarization.

Normal values

- **Duration:** Varies with heart rate; typically 350-440 milliseconds;
- **Correction:** The QT interval is heart rate-dependent and is often corrected using Bazett's formula: $QT_C = \frac{QT}{\sqrt{RR}}$, where RR is the interval between two R peaks in seconds ;
- **Clinical significance:** Prolonged QTc (>440 milliseconds in males, >460 milliseconds in females) increases the risk of ventricular arrhythmias and Short QT interval (<350 milliseconds) is associated with genetic channelopathies [4].

ST Segment

The ST segment extends from the end of the QRS complex (J point) to the beginning of the T wave. It represents the early phase of ventricular repolarization when the ventricles are fully depolarized.

Normal characteristics:

- Position: Usually isoelectric (at the same level as the baseline);
- Variation: May be slightly elevated (up to 1 mm in limb leads, 2 mm in precordial leads) or depressed (up to 0.5 mm).

Clinical significance

ST elevation (>1 mm in limb leads or >2 mm in precordial leads) can indicate acute myocardial infarction and ST depression may suggest myocardial ischemia or digitalis effect [5].

RR Interval

The RR interval is the time between two consecutive R peaks. It is used to calculate heart rate and assess rhythm regularity.

Calculation:

- **Heart rate (bpm):** $HR = \frac{60}{RR\ (\ in\ seconds)}$;
- **Normal range:** 60-100 beats per minute at rest;

- **Clinical significance:** Irregular RR intervals may indicate arrhythmias such as atrial fibrillation or premature beats [7].

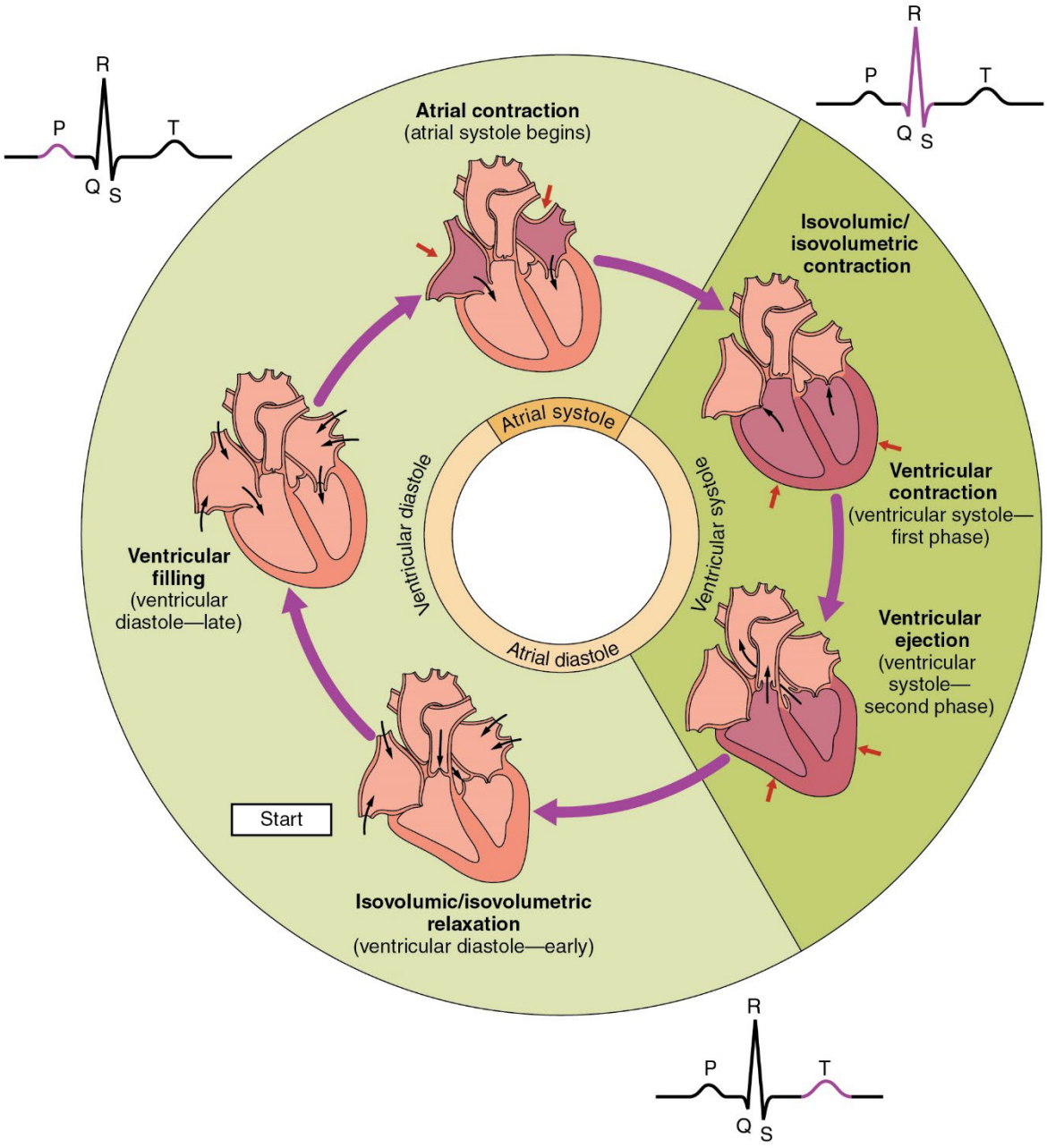

Figure 2:Relationship Between ECG Waves and Mechanical Phases of the Cardiac Cycle.

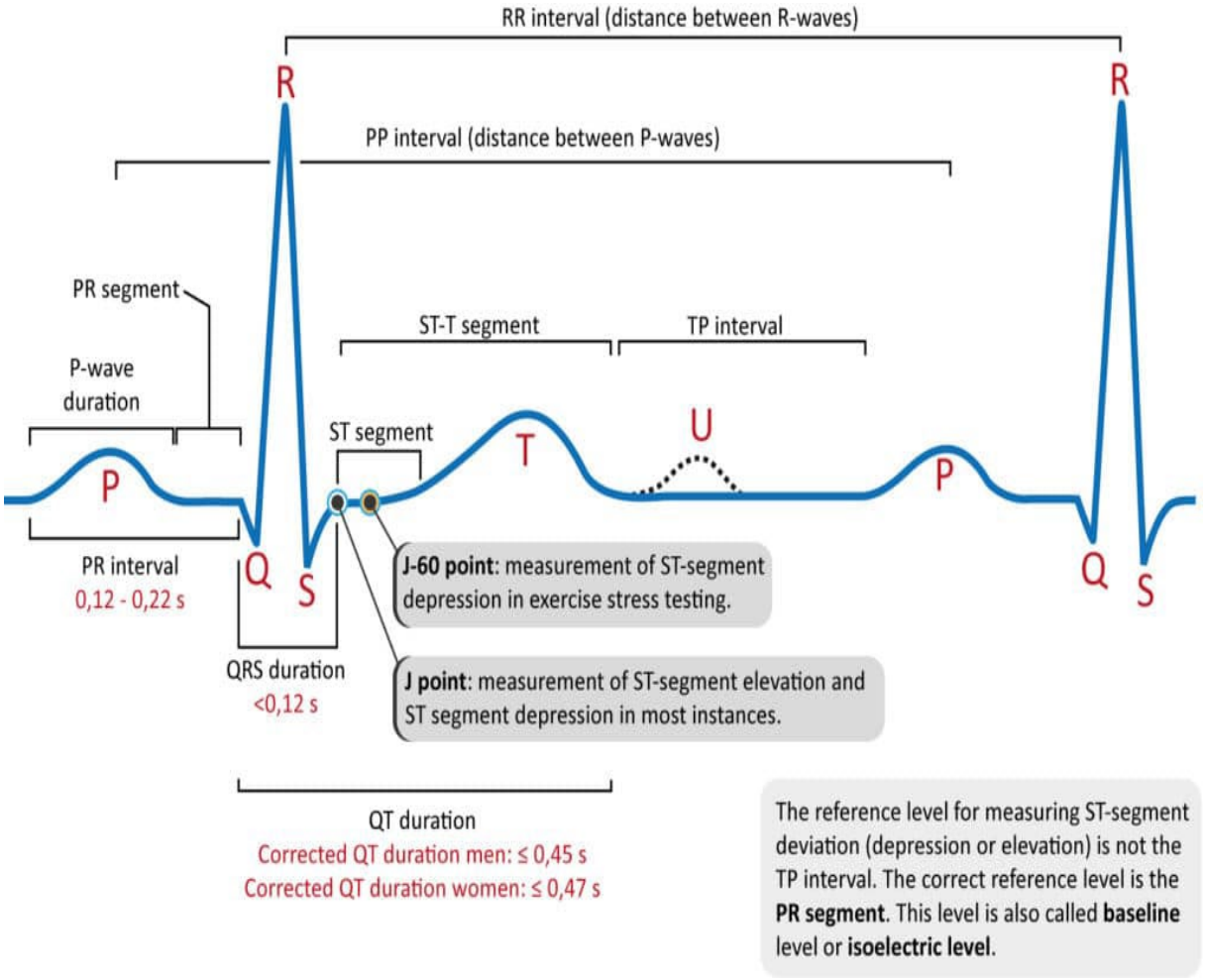

Figure 3:Normal ECG Waveform with Waves, Intervals, and Segments

Normal ECG Values Summary

Parameter	***Normal Range***	***Clinical Significance***
P wave duration	80-120 ms	Atrial depolarization time
PR interval	120- 200 ms	AV conduction time
QRS duration	80- 120 ms	Ventricular depolarization time
QT interval	350- 440 ms	Total ventricular activity
QTc (corrected)	<440 ms (M), <460 ms (F)	Risk of arrhythmias
ST segment	±0.5-1.0 mm	Myocardial ischemia/infarction
Heart rate	60-100 bpm	Sinus rhythm range

Table 1:Values of Normal ECG.

1.1.3. ECG Leads and Recording Systems

To capture the heart's electrical activity from different angles, the standard electrocardiogram uses multiple recording electrodes placed at specific locations on the body. Each pair of electrodes creates a "lead," which represents a unique electrical viewpoint of the heart. The standard 12-lead ECG system provides a comprehensive three-dimensional assessment of cardiac electrical activity by combining information from different spatial planes.

The Twelve Standard Leads

The 12-lead ECG system consists of two main categories: six limb leads and six precordial (chest) leads. The limb leads are further divided into bipolar leads (I, II, and III) and augmented unipolar leads (aVR, aVL, and aVF). The bipolar limb leads measure the potential difference between two limbs and form what is known as Einthoven's triangle, an equilateral triangle with the heart at its center. Lead I measures the potential between the right and left arms, Lead II between the right arm and left leg, and Lead III between the left arm and left leg [7].

The augmented unipolar leads (aVR, aVL, and aVF) were developed later to provide additional frontal plane views. These leads measure the electrical activity from the right arm (aVR), left arm (aVL), and left foot (aVF) relative to a combined reference point from the other limbs. The "a" stands for augmented, indicating that the signal amplitude is increased by 50% compared to the original unipolar recordings [7].

The six precordial leads (V1 through V6) are unipolar chest leads placed at specific positions across the anterior and lateral chest wall. These leads provide a horizontal or transverse view of the heart's electrical activity. V1 and V2 are positioned over the right ventricle, V3 and V4 over the interventricular septum and left ventricle, and V5 and V6 over the lateral wall of the left ventricle. This strategic placement allows for detailed assessment of different cardiac regions and is particularly valuable for detecting myocardial infarction and ventricular hypertrophy [4].

Einthoven's Triangle and the Frontal Plane

Einthoven's triangle is a fundamental concept in ECG interpretation that represents the relationship between the three standard limb leads. Willem Einthoven, who won the Nobel Prize for his work on the ECG, proposed that the heart's electrical activity could be represented as a vector within an equilateral triangle formed by the right arm, left arm, and left leg [10]. This geometric model allows clinicians to determine the heart's electrical axis in the frontal plane.

The relationship between the three limb leads is described by Einthoven's law, which states that the voltage in Lead II equals the sum of the voltages in Leads I and III at any given instant $II = I + III$.

This relationship holds true because the three leads form a closed electrical circuit around the heart. Understanding this principle is essential for verifying proper electrode placement and detecting recording errors [8].

Why Multiple Leads Are Necessary

The use of multiple leads is critical for accurate ECG interpretation because the heart is a three-dimensional structure, and its electrical activity propagates in multiple directions. A single lead provides only a one-dimensional view of this complex activity, which may miss important abnormalities or provide misleading information. For example, an ST-segment elevation indicating acute myocardial infarction may be clearly visible in leads V2-V4 (anterior leads) but completely normal in limb leads [4].

Different leads are sensitive to different cardiac regions:

- Inferior wall: Leads II, III, and aVF;
- Anterior wall: Leads V1-V4;
- Lateral wall: Leads I, aVL, V5 and V6;
- Septal region: Leads V1 and V2;
- Right ventricle: Lead V1 and right-sided leads (if recorded).

This regional specificity is crucial for localizing myocardial infarctions, identifying conduction abnormalities, and diagnosing chamber enlargement. In clinical practice, the 12-lead ECG remains the gold standard for initial cardiac assessment because it provides this comprehensive multi-angle view [7].

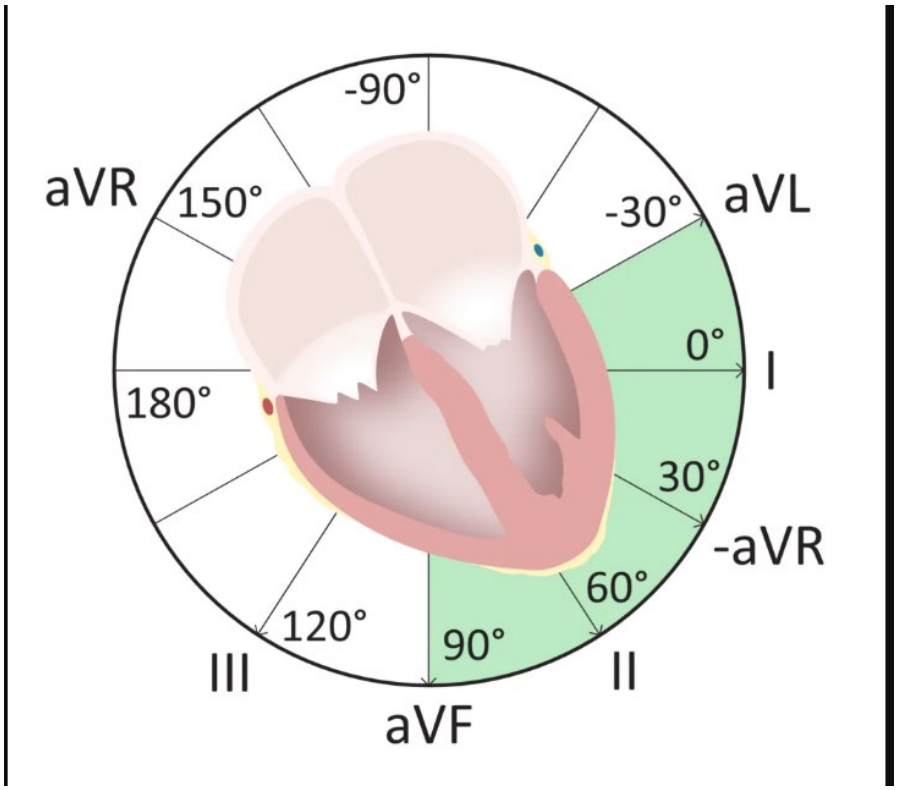

Figure 4: Hexaxial Reference System for Cardiac Electrical Axis Determination

1.1.4. Target Heartbeat Classes for Classification

For this project, heartbeats from the INCART database are divided into three important groups for automatic classification: Normal (N), Supraventricular (S), and Ventricular (V). Normal beats include sinus rhythm and some other beats that still keep a mostly normal ventricular conduction pattern . Supraventricular beats start above the ventricles and usually have a QRS shape similar to normal beats, but they often appear earlier than expected or are followed by pauses . Ventricular beats start inside the ventricles and usually have wider QRS complexes with unusual shapes because the electrical signal follows abnormal pathways.

Fusion beats are not included in this classification task. They are rare in the dataset, which can cause imbalance during training. Their mixed shape also makes them difficult for the model to classify correctly. In addition, detecting Normal, Supraventricular, and Ventricular beats is more important for early arrhythmia screening. Focusing on these three classes helps improve model performance and keeps the system simpler and more efficient.

The subtle differences between these classes, particularly between Normal and Supraventricular beats, highlight the importance of combining both morphological features and temporal dynamics in the classification approach. This challenge motivates the use of artificial intelligence techniques, which can learn complex patterns from data and assist clinicians in making more consistent and accurate diagnoses. The following section discusses the evolving role of AI in ECG analysis and sets the foundation for the hybrid methodology adopted in this project.

1.2 The Role of Artificial Intelligence in ECG Analysis

1.2.1. Challenges in Manual ECG Interpretation

Electrocardiogram interpretation remains a fundamental skill in clinical cardiology, yet manual analysis presents several significant challenges that can affect diagnostic accuracy and patient care. Despite decades of medical training and technological advancement, human interpretation of ECG signals continues to face limitations that artificial intelligence aims to address.

Inter-Observer Variability

One of the most critical challenges in manual ECG interpretation is the substantial variability between different clinicians. Studies have shown that even experienced cardiologists may disagree on the classification of certain arrhythmias, particularly when dealing with subtle morphological differences or borderline cases. This inter-observer variability can lead to inconsistent diagnoses, where the same ECG recording may be interpreted differently depending on the clinician's experience, training background, and subjective judgment.

Fatigue and Workload

In clinical settings, physicians often review dozens or even hundreds of ECG recordings during a single shift, particularly in emergency departments and cardiac care units. This high workload, combined with long working hours, can lead to fatigue and decreased attention to detail. Research indicates that diagnostic accuracy tends to decline as the number of consecutive interpretations increases, raising concerns about patient safety and the reliability of manual analysis in high-pressure environments.

Complexity of Long-Term Monitoring

Modern cardiac monitoring technologies, such as Holter monitors and implantable loop recorders, can generate continuous ECG data over periods ranging from 24 hours to several years. A single 24-hour Holter recording may contain over 100,000 heartbeats, making comprehensive manual review extremely time-consuming and practically impossible. Clinicians typically rely on automated algorithms to flag potential abnormalities, but these systems often generate false positives that still require human verification, creating a significant bottleneck in clinical workflow [7].

Rare Arrhythmia Detection

Certain cardiac arrhythmias occur infrequently, making them difficult to detect through manual review. A clinician may encounter only a handful of cases of specific arrhythmia types throughout their career, limiting their ability to recognize subtle patterns. This challenge is particularly relevant for conditions like supraventricular ectopic beats, which may represent less than 1% of total heartbeats in a patient's recording. The rarity of these events increases the likelihood of missed diagnoses during manual interpretation [8].

Time-Critical Decision Making

In emergency situations, such as acute myocardial infarction or life-threatening arrhythmias, rapid and accurate ECG interpretation is essential for timely intervention. However, the pressure to make quick decisions can compromise diagnostic accuracy. Delays in recognizing critical patterns, such as ST-segment elevation or ventricular tachycardia, can have severe consequences for patient outcomes. Artificial intelligence systems offer the potential for real-time analysis and immediate alerts, supporting clinicians in time-sensitive scenarios [4].

Need for Specialized Expertise

Accurate ECG interpretation requires extensive training and continuous practice. Not all healthcare settings have access to cardiologists or electrophysiologists with specialized expertise in complex arrhythmia diagnosis. In rural or resource-limited areas, general practitioners or nurses may be responsible for initial ECG assessment, increasing the risk of misinterpretation. This disparity in expertise highlights the need for decision support tools that can augment the capabilities of non-specialist healthcare providers [7].

These challenges underscore the growing need for automated, reliable, and efficient ECG analysis systems that can complement human expertise, reduce diagnostic variability, and improve patient care outcomes.

1.2.2. Traditional Computer-Assisted ECG Analysis

Before the emergence of machine learning and deep learning techniques, computer-assisted ECG analysis relied primarily on rule-based systems and traditional signal processing methods. These early approaches represented the first attempt to automate the interpretation of electrocardiogram signals and reduce the burden on clinicians.

Rule-Based Expert Systems

The earliest computerized ECG interpretation systems were based on explicit decision rules derived from clinical knowledge and expert cardiologist input. These systems used a series of if-then statements to evaluate specific ECG features such as heart rate, PR interval duration, QRS width, and ST-segment deviation. For example, a rule might state: "If the PR interval is greater than 200 ms, then classify as first-degree AV block" [6]. These expert systems were designed to mimic the decision-making process of experienced clinicians by encoding established diagnostic criteria into computer algorithms.

The main advantage of rule-based systems was their interpretability. Each diagnostic decision could be traced back to specific rules, making it easy for clinicians to understand how the system reached its conclusion. However, these systems had significant limitations. They required extensive manual programming of thousands of rules to cover all possible clinical scenarios, and they struggled with borderline cases or overlapping conditions where multiple rules might conflict [4].

Traditional Signal Processing Methods

Traditional ECG analysis also employed classical signal processing techniques to extract meaningful features from the raw signal. These methods included digital filtering to remove noise, derivative-based algorithms for QRS complex detection, and template matching for waveform classification. The most widely used QRS detection algorithm, developed by Pan and Tompkins in 1985, combined bandpass filtering, differentiation, squaring, and moving-window integration to identify R-peaks with high accuracy [9].

For arrhythmia classification, traditional systems relied on handcrafted features such as RR-interval statistics, QRS duration, amplitude measurements, and morphological descriptors. These features were then compared against predefined thresholds or used in simple statistical classifiers. While these methods achieved reasonable performance for common arrhythmias, they often failed to capture the complex, non-linear patterns present in pathological ECG signals [6].

Limitations of Traditional Approaches

Despite their historical importance, traditional computer-assisted ECG analysis methods have several fundamental limitations. First, they depend heavily on manual feature engineering, which requires extensive domain expertise and is time-consuming to develop. Second, rule-based systems lack adaptability and cannot learn from new data or improve their performance over time. Third, these approaches struggle with inter-patient variability, as fixed thresholds may not generalize well across different populations, age groups, or clinical conditions [4].

Furthermore, traditional methods often process each heartbeat independently without considering the temporal context or sequential patterns in the ECG signal. This limitation reduces their ability to detect complex arrhythmias that manifest through rhythm patterns rather than individual beat morphology. As ECG databases grew larger and more diverse, the need for more sophisticated, data-driven approaches became increasingly apparent, paving the way for machine learning techniques in cardiac signal analysis [9].

1.2.3. From Machine Learning to Deep Learning in ECG Analysis

The evolution of automated ECG analysis has followed a clear path from rule-based systems to machine learning, and more recently, to deep learning approaches. Each stage has brought new capabilities and addressed limitations of the previous methods. Understanding this progression helps explain why hybrid architectures, such as the one used in this project, represent a promising direction for accurate and reliable heartbeat classification.

Classical Machine Learning Approaches

Before the rise of deep learning, most automated ECG classification systems relied on classical machine learning algorithms such as Support Vector Machines (SVM), Random Forests, and k-Nearest Neighbors (k-NN) [2]. These methods required a careful feature engineering step, where domain experts manually extracted relevant characteristics from the ECG signal. Common features included QRS duration, RR-interval statistics, wave amplitudes,

and morphological descriptors. The extracted features were then fed into the classifier, which learned to map them to specific arrhythmia classes.

While these approaches achieved reasonable performance on well-defined tasks, they had several limitations. First, feature engineering was time-consuming and required extensive medical knowledge. Second, handcrafted features might miss subtle or complex patterns that are not easily described by simple mathematical rules. Third, classical models often struggled to generalize across different patients or recording conditions, as the fixed feature set could not adapt to new data distributions [3].

The Deep Learning Shift

The introduction of deep learning marked a significant shift in automated ECG analysis. Convolutional Neural Networks (CNNs) have emerged as a powerful tool for automatically extracting morphological patterns directly from raw or minimally processed signals [1]. Unlike traditional approaches that rely on manual feature engineering, CNNs learn hierarchical representations through successive convolution and pooling layers, effectively capturing both fine-grained waveforms and broader structural trends. In parallel, Recurrent Neural Networks (RNNs) and their advanced variants, such as Long Short-Term Memory (LSTM) networks, have been widely adopted to model the temporal dynamics of cardiac signals. These sequence-based architectures excel at identifying dependencies between consecutive heartbeats, which is crucial for diagnosing rhythm-related arrhythmias [2]. Nevertheless, RNN-based models often face practical limitations, including high computational costs and susceptibility to vanishing gradients when processing long-term recordings.

1.2.4. Why Hybrid CNN-MLP for Dual-Lead Analysis?

The selection of a classification architecture is a decisive factor in automated ECG analysis, particularly when dealing with multi-lead signals. In this project, we adopt a custom Hybrid CNN-MLP framework specifically designed to address the dual nature of cardiac rhythms: morphological shape and temporal dynamics. Traditional deep learning models often rely on a single data representation, which may overlook critical diagnostic cues or struggle with subtle inter-class differences. By contrast, the hybrid architecture processes two complementary feature domains in parallel, enabling a more comprehensive and physiologically grounded analysis of each heartbeat. This design choice is directly motivated by clinical practice, where cardiologists evaluate both the waveform morphology and the beat-to-beat timing when interpreting ECG recordings.

The first branch of the model utilizes a Convolutional Neural Network (CNN) to extract spatial features from the progressive moving average transform (PMAT) dual-lead images. The PMAT transformation converts each one-dimensional ECG segment into a two-dimensional scale-time matrix, preserving the exact temporal alignment of the P, QRS, and T waves while highlighting morphological patterns at multiple resolutions. By stacking the representations from two different leads, the CNN receives a dual-channel input that captures directional electrical activity across the heart. The convolutional layers automatically learn hierarchical patterns, ranging from local wave sharpness to broader segment deviations, without requiring manual feature engineering. This data-driven approach reduces human bias and adapts effectively to variations in signal amplitude and patient-specific anatomy.

The second branch employs a lightweight Multi-Layer Perceptron (MLP) to process handcrafted temporal features derived from RR-interval dynamics. While morphological analysis is highly effective for detecting ventricular abnormalities, certain arrhythmias, particularly supraventricular ectopic beats, exhibit minimal shape differences from normal sinus rhythm. Their distinguishing characteristics often lie in timing irregularities, such as premature occurrence, irregular spacing, or compensatory pauses. The MLP branch ingests four normalized RR-interval statistics that quantify these rhythm patterns. By processing temporal cues independently, the model avoids diluting subtle timing information within high-dimensional image data, ensuring that rhythm-based diagnostic markers are preserved and properly weighted during the learning process.

The integration of both branches through feature concatenation creates a unified representation that combines morphological precision with temporal context. This fusion strategy offers several practical advantages for clinical deployment. First, it improves discrimination between overlapping classes by leveraging complementary information sources that a single-modality model might miss. Second, the modular design enhances interpretability, as researchers and clinicians can analyze how spatial and temporal features contribute to the final prediction. Third, the architecture remains computationally efficient, making it suitable for real-time monitoring applications. By mirroring the dual-assessment workflow of expert physicians, the Hybrid CNN-MLP framework provides a transparent, robust, and clinically relevant solution for automated heartbeat classification.

1.3 Conclusion

This chapter has established the foundational knowledge necessary for understanding automated ECG heartbeat classification. We began by exploring the medical background of cardiac electrophysiology, explaining how the heart's electrical conduction system generates the ECG signal and describing the characteristic waves (P, QRS, T) and intervals that clinicians use for diagnosis. We also discussed the standard 12-lead recording system and explained why multiple leads are essential for capturing the heart's electrical activity from different perspectives.

The chapter then examined the three main heartbeat classes used in this project: Normal (N), Supraventricular (S), and Ventricular (V) beats. We highlighted the clinical significance of accurate classification and explained why fusion beats were excluded from the analysis. The discussion emphasized that while Normal and Supraventricular beats may appear morphologically similar, they often exhibit distinct timing patterns that can aid in differentiation.

In the second part, we explored the evolving role of artificial intelligence in ECG analysis. We identified key challenges in manual interpretation, including inter-observer variability, clinician fatigue, and the complexity of long-term monitoring. We traced the progression from traditional rule-based systems to classical machine learning approaches, and finally to modern deep learning techniques. This historical perspective revealed that hybrid architectures, which combine morphological and temporal information, offer the most promising path forward for accurate and reliable classification.

The chapter has thus provided both the clinical context and the technical motivation for the project's core methodology: transforming dual-lead 1D ECG signals into 2D image representations using the PMAT transformation, extracting complementary RR-interval features, and processing both inputs through a hybrid CNN-MLP classifier. This dual-domain approach mirrors clinical practice, where cardiologists assess both waveform shape and rhythm patterns to make diagnostic decisions.

The following chapter (Chapter 2) will present the complete methodology in detail, describing each stage of the pipeline from data acquisition and preprocessing to feature

extraction, model architecture, training protocol, and evaluation metrics. Mathematical formulations, algorithmic specifications, and implementation details will be provided to ensure full reproducibility and scientific rigor.

Chapter 2 Methodology

2.1. Methodological Introduction

This section describes the complete workflow developed for classifying ECG heartbeats using dual-lead signals. The process starts with cleaning the raw data and detecting R-peaks accurately. Each heartbeat is then converted into a 2D image-like format to capture its shape, while timing features are extracted to represent rhythm information. These two types of inputs are fed into a hybrid CNN-MLP model for classification. The following subsections explain each step, including the main formulas, model structure, and training settings.

2.2. Pipeline Overview

This section presents the complete workflow developed to identify the most effective dual-lead combinations for automated ECG beat classification. The core objective is to systematically evaluate and rank different lead pairs based on their diagnostic performance. The pipeline begins by converting raw dual-lead ECG signals into image-like representations using the progressive moving average transform (PMAT), while simultaneously extracting timing-based RR-interval features. These two data streams are processed through a hybrid CNN-MLP architecture designed to capture both morphological shapes and rhythm dynamics.

To ensure a comprehensive evaluation of all possible lead combinations, the pipeline follows a systematic screening approach. A single lead is kept fixed and paired sequentially with every other available lead from the 12-lead ECG system. Each pair is independently trained and tested using the proposed hybrid CNN-MLP model under identical conditions. This uniform evaluation strategy allows for fair comparison across all lead pairs and identifies the most discriminative combinations based on consistent performance metrics. The results from all evaluations are then analyzed and ranked to select the optimal lead pairs that achieve the highest classification accuracy for Normal, Supraventricular, and Ventricular beats. This streamlined approach ensures that the final selection is based solely on the CNN-MLP model's performance, providing a clear and reproducible methodology for dual-lead ECG classification.

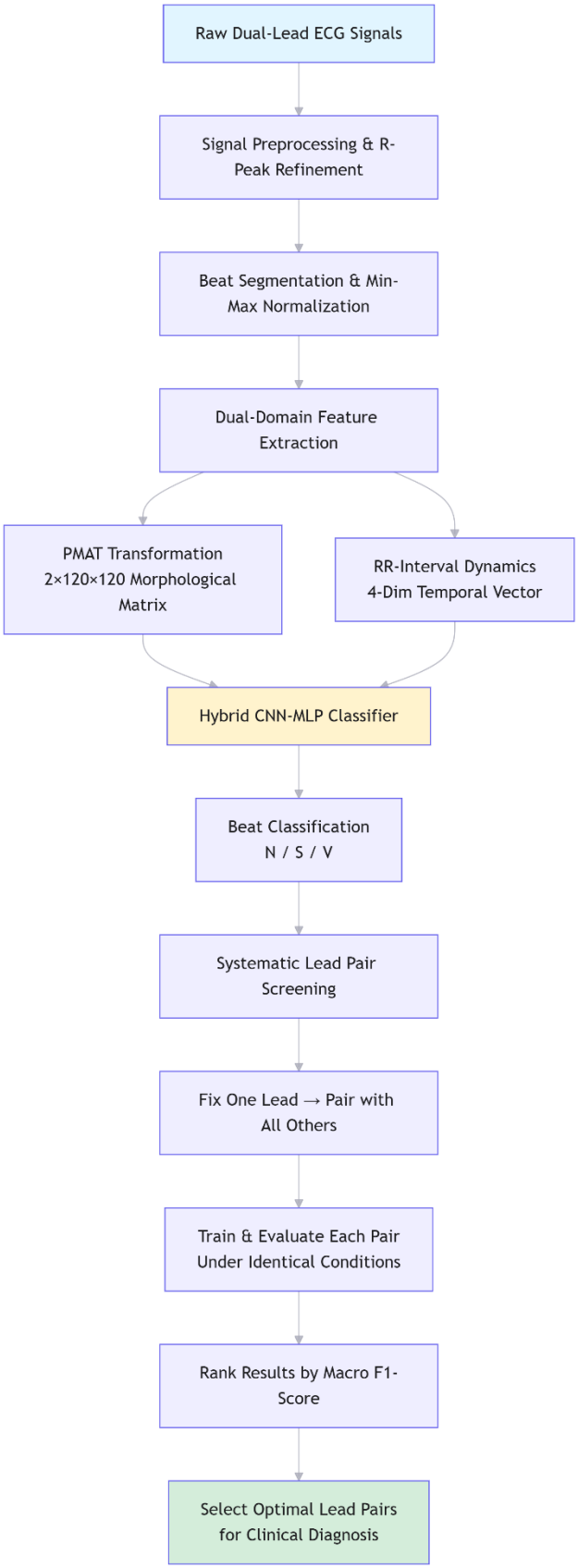

Figure 5:Dual-Lead ECG Classification Pipeline.

2.3. INCART Dataset

This section presents the dataset used to develop and evaluate the proposed ECG classification model. The study relies on the INCART database, a publicly available resource from PhysioNet that is widely recognized in cardiovascular research and algorithm benchmarking [10]. This database provides high-quality, expert-annotated ECG recordings that enable reliable training and fair performance assessment of automated arrhythmia detection systems.

The INCART database contains 75 long-term ECG recordings collected from patients at the St. Petersburg Institute of Cardiology. Each recording lasts approximately 30 minutes and includes all 12 standard clinical leads (I, II, III, aVR, aVL, aVF, V1–V6). The signals were acquired at a sampling rate of 274 Hz with a 12-bit analog-to-digital converter and a voltage range of ±10 mV, ensuring sufficient resolution to capture subtle morphological variations in the ECG waveform [11]. Every heartbeat in the database has been carefully reviewed and labeled by medical experts using the AAMI (Association for the Advancement of Medical Instrumentation) standard annotation codes.

For the purpose of this study, the annotated beats are grouped into three main diagnostic classes to focus on clinically significant arrhythmias:

- Class 0 – Normal (N): Includes normal sinus beats and benign variants (labels: N, L, R, e, j), representing approximately 58,923 beats (~86.5%);
- Class 1 – Supraventricular (S): Includes premature atrial and junctional ectopic beats (labels: A, a, S, J), representing approximately 436 beats (~0.6%);
- Class 2 – Ventricular (V): Includes premature ventricular contractions and ectopic ventricular beats (labels: V, E), representing approximately 8,095 beats (~11.9%).

Fusion beats (label F) and paced beats (label /) are excluded from the classification task to maintain diagnostic clarity and avoid ambiguous patterns that could reduce model interpretability. The dataset is partitioned into 46 records for training and 29 records for testing, following the official PhysioNet split recommendation [12]. This fixed partition prevents data leakage and ensures that the model is evaluated on completely unseen patient data, which is essential for assessing real-world generalization capability.

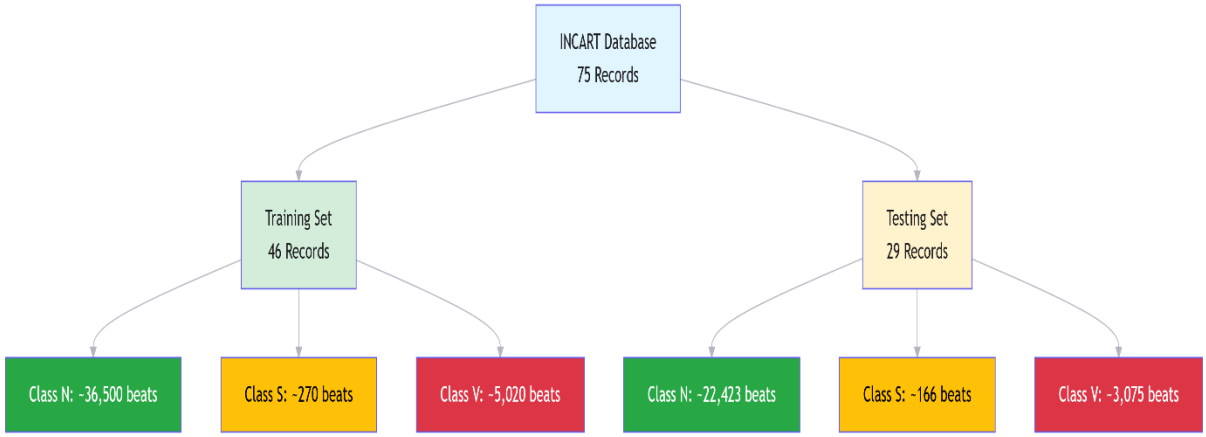

Figure 6:Dataset Split Diagram.

2.4. Signal Preprocessing

After loading the raw ECG data from the INCART database, each signal undergoes a series of carefully designed preprocessing steps. These steps are essential to remove noise, ensure accurate beat alignment, and prepare the data for reliable feature extraction and model training. Without proper preprocessing, the neural network might learn from artifacts instead of real physiological patterns, which would significantly reduce classification accuracy [13].

2.4.1. Baseline Wander Removal

The first step addresses low-frequency noise, known as baseline wander, which can distort the true morphology of the ECG signal. This noise is commonly caused by patient movement, respiration, or electrode instability. If not removed, baseline wander can shift the entire signal up or down, making it difficult to detect important features like the P-wave or T-wave [14].

The first step addresses low-frequency noise, known as baseline wander, which can distort the true morphology of the ECG signal. This noise is commonly caused by patient movement, respiration, or electrode instability. If not removed, baseline wander can shift the entire signal up or down, making it difficult to detect important features like the P-wave or T-wave [14].

Mathematical Formulation:

- Given a raw ECG signal S_{raw} and a sampling rate $f_s = 274\,Hz$, the filtered signal $S_{filtered}$ is computed as $S_{filtered} = S_{raw} - B_{estimate}$.

Where the baseline estimate $B_{estimate}$ is calculated using two consecutive median filters

2.4.2. Why These Window Sizes?

The choice of window sizes is based on the physiological characteristics of the heart and follows established ECG processing guidelines [15]:

- First filter (window = 0.2 seconds): This window is slightly larger than the typical QRS complex duration (80-120 ms). It effectively removes high-frequency spike noise while preserving the sharp edges of the heartbeat morphology;
- Second filter (window = 0.6 seconds): This window corresponds to approximately one complete cardiac cycle at a normal heart rate. It is long enough to capture slow baseline drifts caused by respiration (typically 0.15-0.3 Hz) but short enough to adapt to gradual changes in the signal baseline [16].

The result is a clean signal where the QRS complex, P-wave, and T-wave maintain their natural shape, ready for accurate peak detection.

2.4.3. R-Peak Refinement

Although the INCART database provides expert-annotated R-peak positions, small timing errors may still exist due to manual labeling or signal variability. Even a shift of a few milliseconds can affect the alignment of heartbeats, which is critical for morphological analysis [18].

To improve precision, the algorithm performs a local search around each annotated R-peak using a method similar to the Pan-Tompkins refinement approach [18].

Mathematical Formulation:

- For each annotated position $R_{annotated}$, search window W_{search} is defined as $W_{search} = [R_{left}, R_{right}]$.

Where:

- R_{left} = max (0, $R_{annotated} - 14\ samples$);
- R_{right} = min (N, $R_{annotated} - 14\ samples$);
- Tolerance = ± 0.05 seconds(±14 samples at 274 Hz).

The refined R-peak position $R_{refinder}$, is then identified as the point with the maximum amplitude within this window:

- $R_{refinder,} = R_{left} + argmax(S_{filtered}[R_{left} : R_{right}])$;

This refinement ensures that all heartbeats are consistently aligned across different patients and recording sessions. Accurate alignment is essential because the CNN part of the model learns spatial patterns from the 2D representations, and misaligned beats would introduce noise into these patterns [20].

2.4.4. Beat Segmentation

Once the R-peaks are accurately located, each heartbeat is isolated using a fixed time window centered on the refined R-peak. This step extracts the relevant portion of the signal for classification while discarding unrelated cardiac activity [20].

Mathematical Formulation:

For each $R_{refined}$, the segmented heartbeat H is extracted as

$$H = S_{filtered}[R_{refinder} - b_{before} : R_{refined} + b_{after}]$$

Where:

- b_{before} = 70 samples (≈255 ms before R-peak);
- b_{after} = 70 samples (≈274 ms after R-peak);
- Total length = 145 samples(≈ 529ms).

This duration is chosen to capture the complete P-QRS-T complex for most heartbeats, following standard ECG segmentation practices [21]:

- The P-wave typically occurs 200-300 ms before the R-peak;
- The T-wave ends approximately 200-400 ms after the R-peak;
- The 529 ms total window provides sufficient context while keeping the input size manageable for the neural network.

Segments that fall outside the valid signal range (near the beginning or end of a recording) are automatically skipped to avoid edge artifacts. This ensures that only complete, high-quality heartbeats are used for training and testing.

2.4.5. Min-Max Normalization

Finally, each segmented heartbeat is normalized to the range [0, 1] using min-max scaling. This step is critical because ECG signals can vary significantly in amplitude between different patients, leads, and recording sessions due to factors like electrode placement, skin impedance, or heart size [23].

Mathematical Formulation:

For each segmented heartbeat H, the normalized heartbeat H_{norm} is computed as

$$H_{norm}[i] = \frac{H[i] - H_{min}}{H_{max} - H_{min}}$$

Where H_{min} and H_{max} are the minimum and maximum values of the segment.

Normalization ensures that:

- The neural network receives inputs on a consistent scale, which improves training stability and convergence speed [23];
- The model focuses on shape patterns (morphology) rather than absolute voltage values, which can vary without clinical significance;
- Amplitude differences between different patients and leads do not bias the classification toward certain groups.

After these four preprocessing steps, each heartbeat is clean, precisely aligned, and amplitude-standardized. The processed signals are then ready for the next stage: dual-domain feature extraction using the PMAT transformation for morphological patterns and RR-interval statistics for rhythm dynamics.

2.5. progressive moving average transform (PMAT)

After preprocessing, each heartbeat exists as a one-dimensional (1D) time-series signal. However, Convolutional Neural Networks (CNNs) are highly effective at extracting spatial features from two-dimensional (2D) image-like inputs. To bridge this gap, we apply the Progressive Moving Average Transform (PMAT). This innovative method converts a 1D ECG segment into a 2D matrix that captures the heartbeat's morphology at multiple time scales simultaneously. Unlike traditional frequency-based transforms that lose temporal alignment, PMAT preserves the exact timing of each wave while adding a new dimension

that represents signal smoothness. This makes it easier for the model to recognize patterns that doctors typically look for during manual ECG review.

2.5.1. Core Concept and Clinical Motivation

The PMAT algorithm computes the moving average of the input signal using a series of increasing window sizes. Instead of smoothing the signal with a single fixed window, PMAT generates a set of smoothed versions, each representing the signal at a different level of detail. The resulting matrix has two meaningful dimensions:

- Horizontal axis (Time): Preserves the original temporal order of the signal samples;
- Vertical axis (Scale): Represents the progressive window sizes used for averaging.

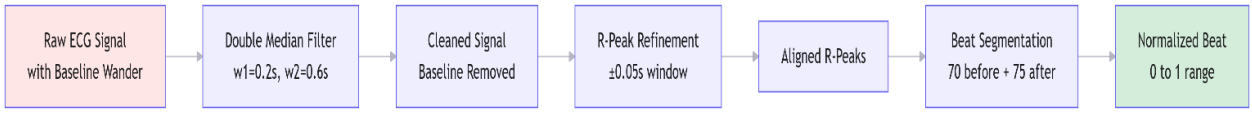

Figure 7:illustrates this transformation conceptually, showing how a 1D ECG signal with P-QRS-T waves is converted into a 2D matrix where each row represents a different smoothing level.

This transformation allows the CNN to detect both fine details (like sharp QRS peaks and Q-wave notches) and broader trends (like ST-segment slope and T-wave width) within a single input structure. By treating the smoothed versions as stacked layers, the model can learn how the heartbeat shape changes when viewed at different resolutions, which is particularly useful for identifying subtle arrhythmia patterns that may be hidden in raw voltage data.

2.5.2. Mathematical Formulation

Given a discrete 1D heartbeat signal $S = [s_1, s_2, \ldots., s_n]$ of length n, and a maximum window size w_{max}, the PMAT matrix M is constructed through a systematic process.

Step 1: Moving Average Definition

The standard Moving Average (MA) for a window size w at position j is defined as

$$MA(S, j, w) = \frac{1}{w} \sum_{i=j}^{j+w-1} s_i$$

Step 2: Signal Padding

Since the output length of a standard moving average decreases as the window size increases $(n - w + 1)$, a padding technique is required to maintain a constant output length n, In this project, we use the Left Padded Moving Average (LPMA) variant. The signal is padded by repeating its first value n times on the left and its last value n times on the right :

$$S_{padded} = [s_1, \dots, s_2, s_2, \dots, s_n, s_n, \dots, s_n]$$

Step 3: Matrix Construction

For each window size w (from 1 to w_{max}), the corresponding row M_w in the matrix is computed by applying the moving average over a shifted slice of the padded signal

$$M_w = MA(S_{padded}[N - w : 2N - 1], w)$$

Where N is the original signal length. This slicing strategy ensures that every row in the matrix contains exactly n values, preserving the temporal structure while adding the scale dimension.

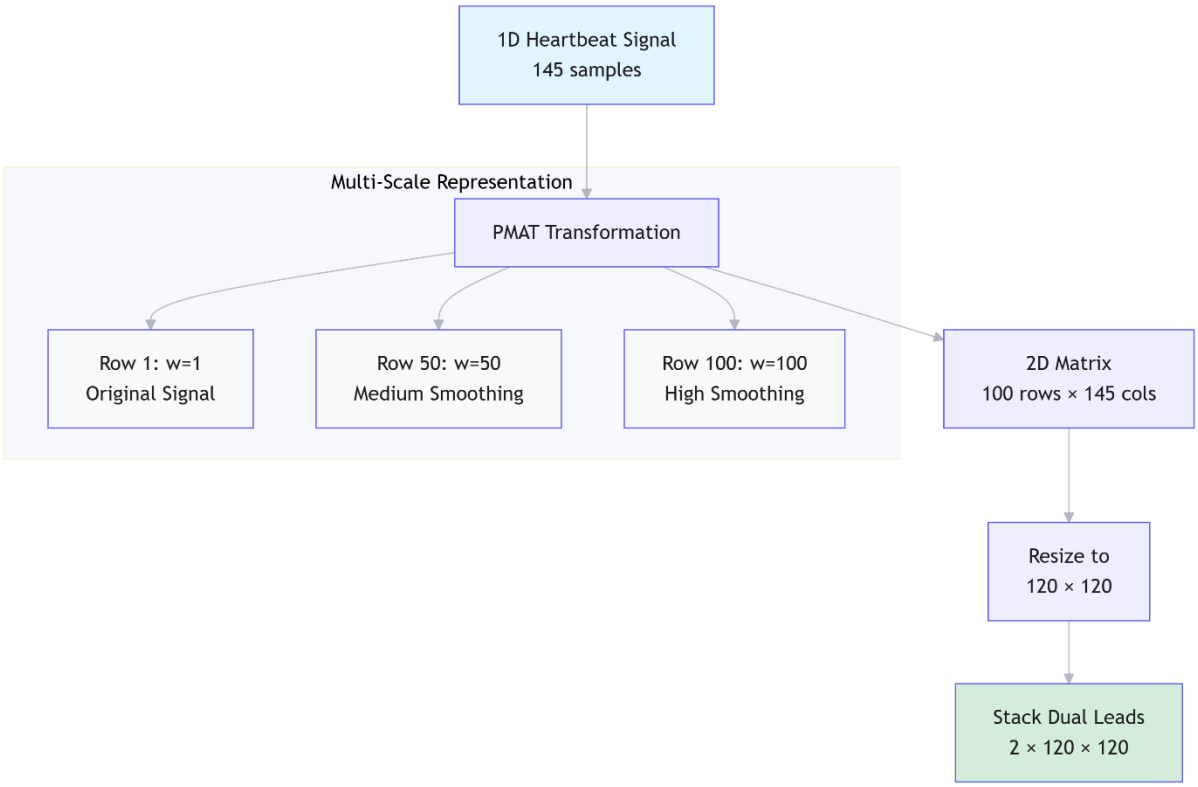

Figure 8:shows a concrete example matrix for a small signal, similar to the example in the reference paper, demonstrating how values change across different window sizes.

2.5.3. Algorithm Implementation Details

Based on the project code, the transformation follows these precise steps:

- Input: A normalized 1D heartbeat segment of length $n = 145$ samples;
- Window Progression: The algorithm iterates through window sizes $w = 1, 2, \ldots, 100$;
- Row Generation: For each w, the moving average is calculated using the left-padding strategy. The first row ($w = 1$) contains the original signal, while subsequent rows contain progressively smoother versions;
- Output Matrix: The result is a 2D matrix M of shape (100, 45), where each row represents the signal at a specific smoothing scale;

The computational complexity of this process is $O(n^2)$, which remains highly efficient for short heartbeat segments and allows fast processing during both training and real-time inference.

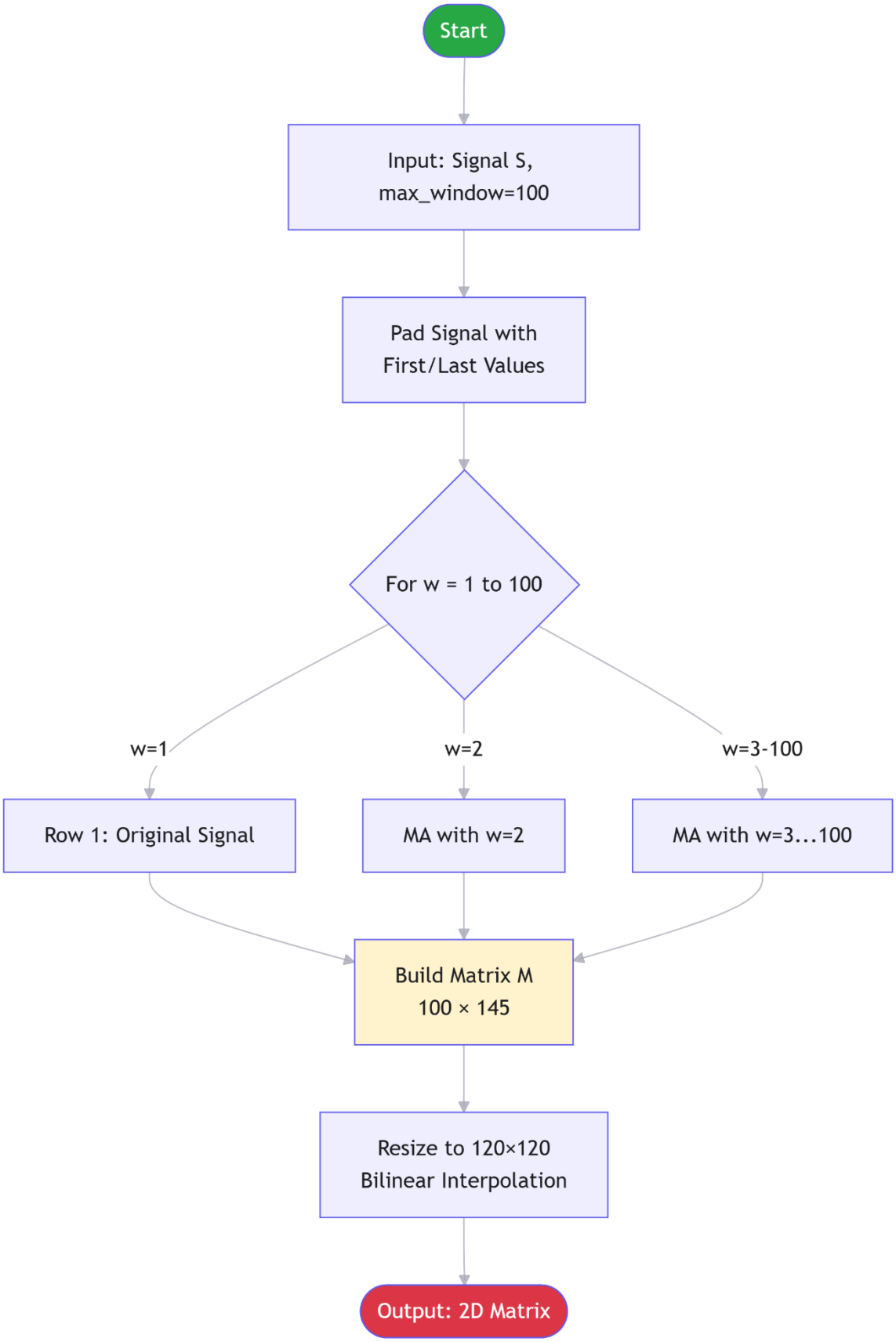

Figure 9:presents a flowchart of the PMAT algorithm, showing the step-by-step process from input signal to output matrix.

2.5.4. Dimensionality Adjustment and Dual-Lead Stacking

The raw PMAT matrix (100, 145) must be standardized to fit the neural network architecture. We resize each matrix to a fixed dimension of 120 × 120 pixels using bilinear interpolation:$M_{resized} = resize(M, (120, 120))$.

This step ensures consistent input size across all heartbeats while preserving the relative morphological patterns and scale relationships. Resizing also helps the CNN process the data more efficiently, as fixed dimensions are required for batch training and memory optimization.

Since the project analyzes two ECG leads simultaneously, we apply PMAT independently to both leads. The resulting images are then stacked along the channel dimension to form a single multi-channel tensor:$X_{input} = torch.tensor([M_{lead1}, M_{lead2}])$.

The final input tensor has the shape [2, 120, 120], representing two channels (leads), each with a height and width of 120 pixels. This format is directly compatible with the first 2D convolutional layer of the CNN, which expects exactly two input channels.

2.5.5. Advantages for ECG Classification

The PMAT transformation offers three key benefits for heartbeat classification

Multi-Scale Feature Extraction

By including window sizes from 1 to 100, PMAT captures both high-frequency details (e.g., R-peak sharpness, Q-wave notches) and low-frequency trends (e.g., ST-segment slope, T-wave width) without requiring separate frequency-domain transforms. This is particularly valuable for detecting conditions like:

- Bundle branch blocks: Visible as widened QRS complexes in larger window sizes;
- ST-segment elevation/depression: Apparent as sustained deviations in medium-scale rows;
- Premature beats: Detectable through abrupt changes in the temporal pattern.

Morphology and Temporal Preservation

Unlike Fourier-based methods that lose time information, PMAT maintains the exact timing of ECG waves. The P, QRS, and T waves remain visually distinguishable in the 2D

representation, which helps the model learn clinically relevant patterns. This preservation is critical because:

- The relative positions of waves carry diagnostic information;
- Wave durations and intervals (PR, QT) are clinically significant;
- Temporal alignment allows the CNN to learn position-specific features.

CNN Compatibility and Clinical Interpretability

The output is a standard 2D grid structure, allowing the model to leverage well-established computer vision architectures. Compared to traditional transforms like Continuous Wavelet Transform (CWT) or Short-Time Fourier Transform (STFT), PMAT offers:

- Simpler computation: No complex mathematical operations or parameter tuning;
- Clearer visual patterns: The 2D matrices align well with clinical ECG interpretation;
- Fewer hyperparameters: Only the maximum window size needs to be specified.

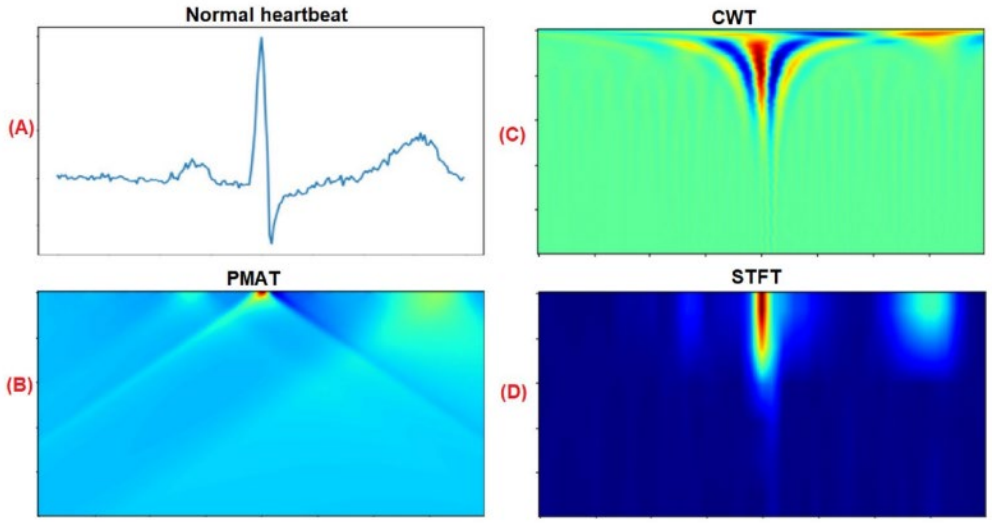

Figure 10:Transformation of a Normal heartbeat signals using PMAT, CWT, and STFT transforms.

2.5.6. Clinical Interpretation of PMAT Patterns

The 2D PMAT representation reveals patterns that correlate with clinical ECG features:

- Early rows (small windows, $w < 20$): Show fine details like R-peak amplitude, Q-wave presence, and high-frequency noise. These rows help detect sharp morphological changes;
- Middle rows (medium windows, $20 \leq w \leq 60$): Reveal the overall QRS complex shape, ST-segment position, and T-wave morphology. These are crucial for identifying ischemic changes and bundle branch blocks;
- Later rows (large windows, $w > 60$): Display broad trends and baseline stability, helping to distinguish between true morphological features and residual noise or artifacts.

This multi-scale view mimics how experienced cardiologists examine ECGs at different levels of detail, making the PMAT representation both computationally effective and clinically meaningful.

2.6. Temporal Feature Engineering (RR-Interval Dynamics)

While the PMAT transformation effectively captures the morphological shape of each heartbeat, the timing between consecutive beats provides equally important diagnostic information. Many cardiac arrhythmias, particularly Supraventricular ectopic beats, can appear nearly identical to Normal beats when viewed only through their waveform shape. However, they often exhibit distinct rhythm patterns that reveal their abnormal electrical origin. To capture this complementary information, we extract four carefully designed temporal features based on RR-intervals. These features are calculated for every heartbeat and fed into a dedicated processing branch, allowing the model to evaluate both the physical shape of the signal and the timing behavior of the cardiac cycle.

2.6.1. RR-Interval Definition and Clinical Context

The RR-interval represents the time distance between two consecutive R-peaks in an ECG recording. It is a fundamental indicator of heart rate variability and is widely used in clinical cardiology to assess rhythm regularity. In our pipeline, the RR-interval is first measured in sample units and then converted to seconds using the known sampling rate of 274 Hz:

$$RR_k = \frac{R_{k+1} - R_k}{f_k}$$

Where R_k and R_{k+1} are the sample indices of two adjacent R-peaks, and f_s is the sampling frequency. From a clinical perspective, changes in RR-intervals often signal underlying electrical disruptions. Premature contractions typically shorten the interval before the abnormal beat, while a compensatory pause frequently follows ventricular ectopic activity. By quantifying these timing variations, we provide the model with rhythm-based cues that morphology alone cannot reveal.

2.6.2. Feature Standardization and Data Integrity

Before these temporal features can be used by the neural network, they must be standardized to ensure consistent numerical behavior. Raw RR-interval values differ significantly between patients due to variations in age, fitness level, medication, or baseline heart rate. To address this, we apply the StandardScaler technique, which transforms each feature to have a mean of zero and a standard deviation of one:

$$f_{normalized} = \frac{f - \mu}{\sigma}$$

- f represents the original feature value;
- $f_{normalized}$ represents the standardized feature, where the output value after transformation, now with mean ≈ 0 and standard deviation ≈ 1;
- σ represents the standard deviation;
- μ represents the represents, where the average value of feature f calculated only from the training dataset.

Where μ and σ are calculated exclusively from the training dataset. This mathematical adjustment guarantees that all four features influence the learning process equally, regardless of their original magnitude. Importantly, fitting the scaler only on the training data prevents information from the testing set from influencing the normalization process, thereby eliminating data leakage and preserving the validity of our experimental results.

2.7. Hybrid CNN-MLP Architecture

The classification model in this project is designed as a hybrid deep learning architecture that processes two distinct types of cardiac information in parallel. While the CNN branch extracts spatial morphological patterns from the PMAT-transformed ECG images, the MLP

branch learns temporal rhythm dynamics from the extracted RR-interval features. This dual-branch design reflects the clinical approach used by cardiologists, who simultaneously evaluate waveform shape and beat timing to reach an accurate diagnosis. The two branches operate independently until a fusion stage, where their learned representations are combined and passed through a shared classification head to produce the final prediction.

2.7.1. CNN Branch for Morphological Feature Extraction

The spatial branch receives a 2-channel PMAT matrix of shape [2, 120, 120] as input. It consists of three convolutional blocks, each progressively extracting higher-level morphological features while reducing spatial dimensions. The first block applies a 2D convolution with 32 filters of size 5×5, followed by Batch Normalization and a ReLU activation function. The operation can be expressed as:

$$X^{(1)} = ReLU(BN_1(\sum_{c=1}^{2} W_c^{(1)} * X_{in}^{c} + b^{(1)}))$$

where $*$ denotes the convolution operation, $W^{(1)}$ and b^1 are learnable weights and biases, and BN_1 standardizes the output to stabilize training. A Max Pooling layer with a 5×5 kernel is then applied to reduce spatial redundancy while retaining the most prominent morphological edges, such as sharp QRS peaks and P-wave onsets.

The second block repeats this pattern with 64 filters of size 5×5, another Batch Normalization layer, and a 3×3 Max Pooling operation. This stage captures broader waveform trends, including ST-segment slopes and T-wave shapes, which are critical for distinguishing between Normal and Ventricular beats. The third and final convolutional block uses 128 filters of size 3×3, followed by Batch Normalization and ReLU. Instead of fixed pooling, an Adaptive Max Pooling layer compresses the entire feature map into a single 1×1 spatial dimension per channel, resulting in a compact 128-dimensional vector:

$$h_{spatial} = \text{AdaptiveMaxPool}\left(\text{ReLU}\left(\text{B}N_3\left(Conv_3\left(X^{(2)}\right)\right)\right)\right) \in \mathbb{R}^{128}$$

This design ensures that the network focuses on the most discriminative morphological signatures regardless of minor signal misalignments, while maintaining a lightweight parameter count suitable for efficient training.

2.7.2. MLP Branch for Temporal Rhythm Processing

The temporal branch processes the four standardized RR-interval features $x_2 \in \mathbb{R}^4$ through a single fully connected layer. Rather than stacking multiple dense layers, a direct mapping to a 128-dimensional latent space is employed:

$$h_{temporal} = ReLU(W_{temp}.x_2 + b_{temp})$$

Where $W_{temp} \in \mathbb{R}^{128\times4}$ and $b_{temp} \in \mathbb{R}^{128}$ are trainable parameters. This lightweight design is intentional. RR-interval features are already highly engineered and clinically meaningful, so a shallow transformation prevents overfitting while allowing the model to learn non-linear interactions between timing cues, such as the relationship between premature intervals and compensatory pauses. The ReLU activation introduces the necessary non-linearity to capture complex rhythm patterns that linear methods would miss.

2.7.3. Feature Fusion and Classification Head

Before classification, the outputs of both branches are aligned and refined. The spatial vector passes through an additional linear layer (fc0: 128 → 128) to harmonize its feature distribution with the temporal branch

$$h'_{spatial} = ReLU(W_{align}.h_{spatial} + b_{align})$$

Both vectors are then concatenated along the feature dimension to form a unified representation:

$$h_{combiend} = [h'_{spatial}; h_{temporal}] \in \mathbb{R}^{256}$$

This 256-dimensional vector contains a comprehensive description of the heartbeat, merging shape and timing information. It is fed into a classification head consisting of two fully connected layers. The first layer reduces the dimensionality from 256 to 64, forcing the model to compress the combined features into a highly discriminative subset:

$$h_{final} = ReLU(W_{cls1}.h_{combined} + b_{cls1})$$

The second layer maps the 64 features directly to the three output classes (Normal, Supraventricular, Ventricular):

$$y_{logits} = W_{cls2}.h_{combined} + b_{cls2}$$

During training, these logits are passed through a Softmax function to produce class probabilities. The progressive reduction (256 → 64 → 3) acts as an implicit regularization mechanism, preventing the model from memorizing noise and encouraging it to focus on clinically relevant patterns.

2.7.4. Clinical and Computational Rationale

The hybrid architecture offers several advantages tailored to ECG classification. First, separating morphological and temporal processing allows each branch to specialize. The CNN learns position-invariant shape features that are robust to amplitude scaling, while the MLP captures rhythm irregularities that are often invisible in raw waveforms. Second, the use of Batch Normalization after each convolutional layer mitigates internal covariate shift, enabling faster convergence and higher tolerance to varying learning rates. Third, the adaptive pooling and compact fusion design keep the total parameter count low, reducing the risk of overfitting on the imbalanced INCART dataset. Finally, the architecture remains fully deterministic and reproducible, making it suitable for clinical validation and future deployment in real-time monitoring systems.

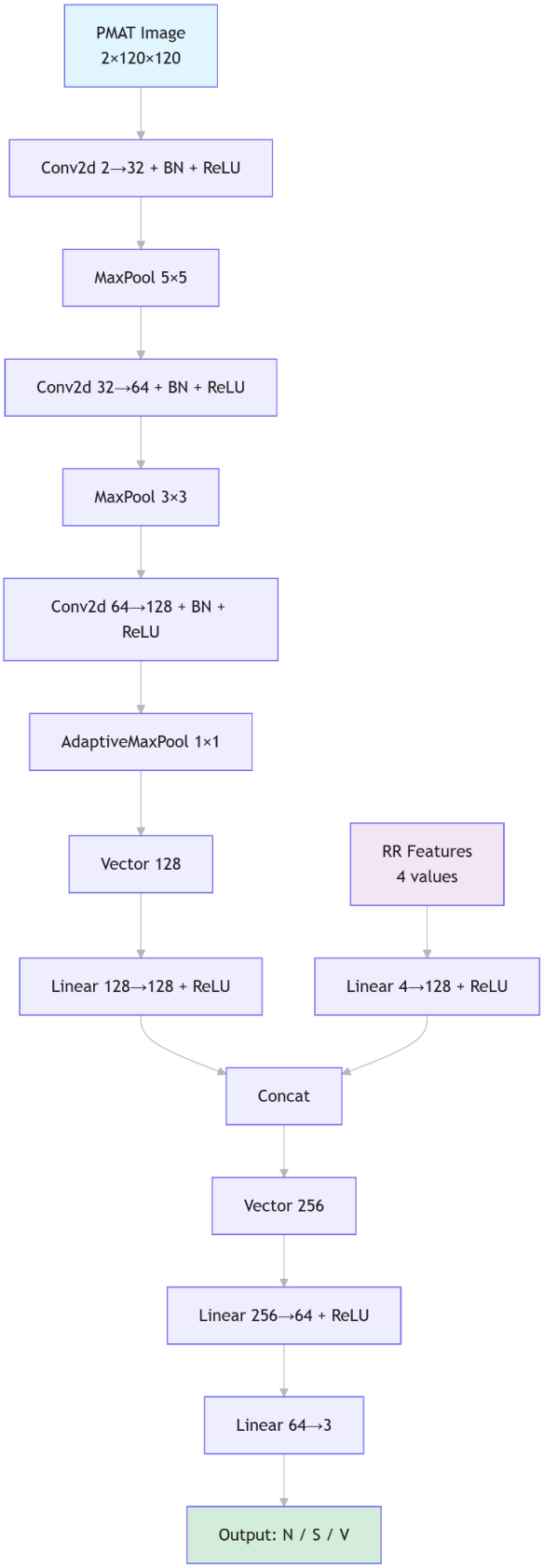

Figure 11:Hybrid CNN-MLP Architecture.

2.8. Training Protocol and Optimization Strategy

After defining the hybrid CNN-MLP architecture, the next critical step is to establish a robust training protocol. This section describes the optimization strategy, loss function, learning rate scheduling, weight initialization, and evaluation metrics used to train the model effectively. A well-designed training protocol ensures that the model learns meaningful patterns from the data without overfitting, while maintaining reproducibility and clinical reliability.

2.8.1. Loss Function and Optimization Algorithm

The model is trained using the Cross-Entropy Loss function, which is the standard choice for multi-class classification tasks. This loss measures the difference between the predicted probability distribution and the true class label.

Mathematical Formulation:

For a batch of B samples with true labels $y_i \in \{0, 1, 2\}$ and predicted logits $z_i \in \mathbb{R}^3$, the loss is computed as :$\mathcal{L} = -\frac{1}{B}\sum_{i=1}^{B} \log(\frac{\exp(z_{i,y_i})}{\sum_{c=0}^{2} \exp(z_{i,c})})$

This formulation encourages the model to assign high probability to the correct class while penalizing confident wrong predictions. To minimize this loss, we use the Adam optimizer, which combines the benefits of momentum and adaptive learning rates.

Adam Update Rule: $\theta_{t+1} = \theta_t - \alpha \cdot \frac{\widehat{m_t}}{\sqrt{\widehat{v_t}}+\epsilon}$

Where θ represents model parameters, $\alpha = 0.004$ is the learning rate, $\hat{m}$ and $\hat{v}$ are bias-corrected first and second moment estimates, $\epsilon = 10^{-8}$ ensures numerical stability. Adam is chosen for its fast convergence and robustness to noisy gradients, which is particularly important when training on imbalanced medical data.

2.8.2. Learning Rate Scheduling

A fixed learning rate may cause the model to converge slowly or oscillate near the optimum. To address this, we apply a Step Learning Rate Scheduler that reduces the learning rate by a factor of 0.5 every 5 epochs: $\alpha_t = \alpha_0 \times \gamma^{[\frac{t}{step_{size}}]}$

Where $\alpha_0 = 0.004$, $\gamma = 0.5$, and step_size=5. This strategy allows the model to make large updates in early epochs for rapid progress, then fine-tune with smaller steps as training advances. The scheduler helps escape local minima early and stabilizes convergence in later stages, which is essential for achieving consistent performance across different lead pairs.

2.8.3. Weight Initialization Strategy

Proper initialization of neural network weights is crucial for stable training and avoiding vanishing or exploding gradients. We apply different initialization schemes based on layer type:

For Convolutional Layers:

$$W \sim \mathcal{N}\left(0, \frac{2}{n_{in}}\right)$$

Where n_{in} is the number of input channels. This is known as Kaiming Normal initialization, which is designed for layers followed by ReLU activation functions.

For Linear Layers:

$$W \sim \mathcal{N}(0, 0.01), \quad b = 0$$

A small fixed standard deviation ensures that initial outputs remain in a reasonable range, preventing early saturation of activation functions.

These initialization choices help the model start training from a favorable point in parameter space, reducing the risk of poor convergence and improving final classification accuracy.

2.8.4. Training Configuration and Reproducibility

The training process is configured with the following hyperparameters:

- Batch size: 512 samples per iteration, which balances memory efficiency with gradient stability;
- Number of epochs: 15 full passes through the training data;
- Random seed: Fixed to ensure deterministic behavior across runs;
- Deterministic mode: Enabled in PyTorch to guarantee reproducible results.

This configuration ensures that experiments can be repeated exactly, which is essential for scientific validation and fair comparison between different lead pairs. The batch size of

512 was selected after preliminary testing to maximize GPU utilization while maintaining stable gradient estimates.

2.8.5. Evaluation Metrics and Monitoring

During training, we monitor multiple metrics to assess model performance comprehensively. Using a single metric can be misleading, especially with imbalanced datasets. Therefore, we track several complementary measures that provide different perspectives on classification quality.

Accuracy measures the overall proportion of correct predictions:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

While accuracy is easy to interpret, it can be deceptive for imbalanced data. A model that predicts all beats as Normal would achieve ~86% accuracy but would be clinically useless.

Precision measures how many predicted positive cases are actually correct:

$$Precision_c = \frac{TP_c}{TP_c + FP_c}$$

High precision means fewer false alarms, which is important to avoid unnecessary medical interventions.

Recall (Sensitivity) measures how many actual positive cases were correctly identified:

$$Recall_c = \frac{TP_c}{TP_c + FN_c}$$

High recall is critical in medical diagnosis because missing a dangerous arrhythmia (false negative) can have serious clinical consequences.

F1-Score is the harmonic mean of precision and recall, balancing both concerns:

$$F1_c = \frac{2.Precision_c.Recall_c}{Precision_c + Recall_c}$$

To summarize performance across all three classes, we compute two types of averages:

Macro-Averaged F1-Score treats all classes equally, regardless of their frequency:

$$F1_{macro} = \frac{1}{C}\sum_{c=1}^{C} F1_c$$

Where $C = 3$ represents the number of classes (Normal, Supraventricular, Ventricular). This metric is particularly important for our imbalanced dataset because it ensures that the rare Supraventricular class (~0.6% of data) has equal weight in the final score. A high macro F1 indicates that the model performs well on all classes, not just the majority class.

Weighted-Averaged F1-Score accounts for class imbalance by weighting each class's F1-score by its support (number of true instances):

$$F1_{weighted} = \sum_{c=1}^{C} \frac{n_c}{N} . F1_c$$

Where n_c is the number of samples in class c, and N is the total number of samples. This metric reflects the overall performance while acknowledging the natural class distribution in the dataset.

2.8.6. Handling Class Imbalance

Given the significant imbalance in the INCART dataset (Normal: ~86.5%, Ventricular: ~11.9%, Supraventricular: ~0.6%), special care is taken during training. While we do not apply oversampling or class weighting in the current setup, the use of macro F1-score as the primary metric ensures that performance on rare classes is not overlooked. Future improvements could include focal loss or stratified sampling to further address this challenge.

After completing the training protocol, the model is evaluated on the held-out test set using the same metrics. The final results, including confusion matrices and per-class performance, are then analyzed to identify the most effective lead pairs for clinical diagnosis.

2.9. Conclusion

This section has presented a complete and structured methodology for automated ECG heartbeat classification using dual-lead signals. The pipeline begins with raw signal acquisition from the INCART database, followed by careful preprocessing to remove noise and ensure accurate beat alignment. Each heartbeat is then transformed into a two-dimensional representation using the PMAT algorithm, which captures morphological patterns at multiple time scales. In parallel, four handcrafted temporal features based on RR-intervals are extracted

to represent rhythm dynamics. These two complementary data streams are processed by a hybrid CNN-MLP architecture, where the CNN branch learns spatial patterns from PMAT images and the MLP branch learns temporal patterns from RR-features. The combined representation is then classified into three clinical categories: Normal, Supraventricular, and Ventricular beats.

The entire workflow is designed with reproducibility and clinical interpretability in mind. Signal preprocessing uses physiologically motivated parameters, such as median filter windows of 0.2s and 0.6s, which align with typical QRS duration and cardiac cycle length. The PMAT transformation preserves temporal alignment while enabling multi-scale feature extraction, making it well-suited for convolutional networks. The RR-interval features provide complementary timing information that helps distinguish arrhythmias with similar morphology but different rhythm patterns. The hybrid architecture balances model complexity with training efficiency, using Batch Normalization for stability and a progressive fusion strategy to combine spatial and temporal cues.

Training is conducted under a controlled protocol with fixed data splits, deterministic initialization, and standardized evaluation metrics. The use of macro-averaged F1-score ensures that performance on rare classes, such as Supraventricular beats, is not overlooked despite dataset imbalance. By systematically evaluating all possible lead pairs through this pipeline, the methodology enables objective identification of the most discriminative dual-lead combinations for clinical diagnosis.

In summary, this methodology establishes a robust, reproducible framework for dual-lead ECG classification. It integrates signal processing, feature engineering, and deep learning into a cohesive system that mirrors clinical reasoning: assessing both waveform shape and beat timing. The modular design allows for future extensions, such as incorporating additional leads, testing alternative transformations, or adapting the architecture for real-time deployment. The following sections present the experimental results obtained by applying this methodology to the INCART dataset.

Chapter 3 Development Environment and Library Definitions

3.1 Introduction

This chapter outlines the technical environment and software tools used to develop the ECG classification system. It begins by introducing the Kaggle platform, which serves as the primary development workspace due to its cloud-based infrastructure and built-in GPU support. Following this, each Python library employed in the project is briefly defined, with a clear explanation of its specific role in signal processing, data management, model training, and performance evaluation. The chapter also presents the custom convolutional neural network designed for this task, details the data preprocessing pipeline, and describes the training configuration and evaluation metrics. This structured overview prepares the reader for the implementation details and sets the stage for the experimental results.

3.2 Platform Kaggle

Kaggle is a cloud-based platform widely used for data science and machine learning projects. It provides a ready-to-use coding environment based on Jupyter Notebooks, which removes the need for local software installation. In this project, Kaggle was selected because it offers free access to high-performance GPUs, such as the NVIDIA Tesla T4. This hardware acceleration significantly reduces training time. The platform also includes built-in dataset management, version control, and easy sharing of experimental results.

3.3 Deep Learning & Training Frameworks

PyTorch (Torch) is an open-source library developed for building and training neural networks. It uses tensor operations and automatic differentiation to simplify model development. In this work, PyTorch serves as the core engine for defining the custom convolutional architecture (MyModule) and managing the training loop.

Skorch is a lightweight wrapper that combines PyTorch with the scikit-learn interface. It allows neural networks to be trained using familiar methods like .fit()

and .predict(). Skorch also provides useful training callbacks, such as learning rate scheduling and performance tracking, which streamline the optimization process.

TensorBoard is a visualization toolkit that monitors training metrics in real time. It is integrated here through PyTorch's SummaryWriter to track loss curves, accuracy, and F1 scores across epochs, making it easier to adjust hyperparameters during training.

3.4 Signal Processing & Mathematical Libraries

WFDB (WaveForm DataBase) is a specialized library for reading physiological signals from public repositories like PhysioNet. It handles raw ECG records and annotation files, enabling direct access to heartbeats and their clinical labels.

SciPy (Signal & FFT modules) provides advanced mathematical functions for scientific computing. The signal module applies median filtering to remove baseline wander from ECG traces, while the FFT module performs fast Fourier transforms to examine frequency components.

PyWavelets (Pywt) implements wavelet transforms, which decompose signals into time-frequency representations. It supports the detection of subtle morphological changes in heartbeats and assists in feature exploration.

NumPy (Numerical Python) is the foundational package for numerical computing. It offers efficient array operations and vectorized calculations, which are essential for handling large-scale ECG data and performing fast mathematical transformations.

3.5 Data Preprocessing & Image Conversion Tools

OpenCV (cv2) is primarily known for computer vision tasks, but it is used here to resize and normalize the PMAT (Permutation Matrix) representations. This step converts one-dimensional heartbeats into two-dimensional inputs suitable for convolutional networks.

StandardScaler (from scikit-learn.preprocessing) standardizes numerical features by removing the mean and scaling to unit variance. It is applied to the RR-interval features to ensure stable gradients and faster convergence during training.

PIL (Python Imaging Library) provides basic image handling capabilities. It supports format conversions and ensures compatibility when saving or loading processed heartbeat matrices.

Joblib and Concurrent Futures enable parallel processing and efficient file handling. They speed up the extraction and transformation of thousands of heartbeats by distributing tasks across multiple CPU cores.

3.6 Evaluation & Visualization Utilities

Scikit-Learn Metrics includes functions like confusion_matrix, classification_report, and f1_score. These tools evaluate model performance by calculating precision, recall, and macro-averaged F1 scores, which are crucial for handling imbalanced ECG datasets.

Matplotlib and **Seaborn** (sns) are plotting libraries used to create clear visual summaries. Seaborn simplifies the generation of heatmaps for confusion matrices, while Matplotlib renders signal plots, training curves, and experimental comparisons.

Collections (Counter) provides a simple way to count label frequencies in the dataset. This helps identify class imbalances before training and guides the selection of appropriate evaluation metrics.

3.7 Chapter Summary

This chapter delivered a complete overview of the development environment and technical stack used throughout the project. We explained how Kaggle provides a streamlined, GPU-enabled workspace and defined each library according to its specific function in the ECG classification pipeline. A lightweight custom CNN (MyModule) was introduced to replace pre-trained architectures, and its integration with temporal RR-interval features through a hybrid design was clearly described. The chapter also documented the full training workflow, from signal filtering and PMAT image transformation to optimization strategies, callback functions, and evaluation metrics adapted for imbalanced cardiac data. This foundation establishes a reproducible and efficient implementation framework, which will be tested and analyzed in the next chapter.

Chapter 4 Results and Analysis

4.1 Evaluation Metrics and Data Overview

In this experimental phase, we adopted a two-stage evaluation strategy to efficiently identify the most effective lead configuration for heartbeat classification. Initially, we used a representative subset of the dataset, referred to as the incart sample, to test and compare twelve different lead setups. This approach allowed us to reduce computational cost and accelerate the selection process while still maintaining reliable performance indicators. To assess each configuration, we relied on standard classification metrics: Precision (the ability to avoid false positives), Recall (the ability to capture all relevant instances), and the F1-score (the harmonic mean of Precision and Recall), alongside overall Accuracy. Given the notable class imbalance in our data—where class N dominates (Support ≈ 7122), class V is moderate (Support ≈ 694), and class S is very limited (Support = 16)—we reported both Macro-average (treating all classes equally) and Weighted-average (accounting for class support) results. This dual perspective helps highlight whether a model performs consistently across rare and frequent classes or simply benefits from the majority class. Once the optimal lead configuration is identified through this preliminary analysis, the final model will be retrained on the complete dataset to maximize generalization and ensure robustness for real-world application.

4.2 Performance Results by Lead Configuration

4.2.1. Detailed Performance Tables by Experimental Group

In this subsection, we present the complete performance results for all twelve lead configurations across the different experimental groups. Each table reports Precision, Recall, F1-score, Accuracy, and Support for the three target classes: N (Normal), S (Supraventricular), and V (Ventricular). The evaluation is based on a representative subset of the incart dataset, with class distribution: N ≈ 7,122 samples, V ≈ 694 samples, and S = 16 samples.

For clarity and consistency, results are organized by experimental group, with each group containing a dedicated table summarizing the performance of Leads 1 through 12.

Lead 1(I)

	Precision					*Recall*					*F1-score*						*Support*		
	N	S	V	M avg	W avg	N	S	V	M avg	W avg	N	S	V	M avg	W avg	accuracy	N	S	V
1																			
2	99.54	0	97.68	65.74	99.17	99.82	0	97.12	65.65	99.37	99.68	0	97.40	65.69	99.27	99.37	7122	16	694
3	99.32	0	85.71	61.68	97.91	98.48	0	95.10	64.53	97.98	98.90	0	90.16	63.02	97.92	97.98	7122	16	694
4	99.37	0	98.66	66.01	99.11	99.89	0	95.68	65.19	99.31	99.63	0	97.15	65.59	99.21	99.31	7122	16	694
5	98.87	0	84.66	61.18	97.41	98.41	0	90.63	63.02	97.52	98.64	0	87.54	62.06	97.46	97.52	7122	16	694
6	99.31	0	95.78	65.03	98.80	99.63	0	94.81	64.82	99.00	99.47	0	95.23	64.92	98.90	99.00	7122	16	694
7	99.63	0	84.22	61.28	98.06	98.22	0	98.41	65.54	98.03	98.92	0	90.76	63.23	97.99	98.03	7122	16	694
8	99.69	0	85.98	61.89	98.27	98.44	0	98.99	65.81	98.29	99.06	0	92.03	63.70	98.24	98.29	7122	16	694
9	99.18	0	81.58	60.25	97.41	97.95	0	93.80	63.92	97.38	98.56	0	87.27	61.94	97.36	97.38	7122	16	694
10	99.28	0	91.75	63.68	98.41	99.19	0	94.52	64.57	98.57	99.23	0	93.12	64.12	98.49	98.57	7122	16	694
11	99.63	0	89.41	63.01	98.52	98.86	0	98.56	65.81	98.63	99.25	0	93.76	64.34	98.56	98.63	7122	16	694
12	99.54	0	99.15	66.23	99.20	99.96	0	98.31	66.09	99.50	99.75	0	98.73	66.16	99.35	99.50	7122	16	694

Table 2:test lead 1 with all leads.

		N	*S*	*V*	*M avg*
The best	Number	8		12	12
	Precision	99.69		99.15	66.23
The worst	Number	5		9	9
	Precision	98.87		81.58	60.25

Table 3:results lead 1

The best results with Lead **12**.

The worst results with lead **9**.

Lead 2 (II)

	Precision					*Recall*					*F1-score*					*Support*		
	N	S	V	M avg	W avg	N	S	V	M avg	W avg	N	S	V	M avg	W avg	accuracy	N	S	V
1	99.28	0	85.92	61.73	97.89	98.58	0	94.09	64.22	97.98	98.93	0	89.82	62.92	97.92	97.98	7122	16	694
2																			
3	99.62	0	88.50	62.70	98.43	98.85	0	97.55	65.47	98.53	99.23	0	92.80	64.01	98.46	98.53	7122	16	694
4	99.73	0	93.45	64.39	98.97	99.41	0	98.70	66.04	99.14	99.57	0	96.01	65.19	99.05	99.14	7122	16	694
5	99.40	0	83.42	60.94	97.78	98.20	0	95.68	64.63	97.78	98.80	0	89.13	62.64	97.74	97.78	7122	16	694
6	98.52	0	82.03	60.18	96.86	98.26	0	87.17	61.48	96.99	98.39	0	84.05	60.81	96.92	96.99	7122	16	694
7	99.63	0	81.61	60.41	97.83	97.92	0	97.48	65.25	97.71	98.77	0	88.99	62.59	97.70	97.71	7122	16	694
8	99.70	0	78.83	59.51	97.65	97.47	0	98.70	65.39	97.38	98.57	0	87.65	62.07	97.40	97.38	7122	16	694
9	99.36	0	26.11	41.82	92.66	73.73	0	95.82	56.52	75.54	84.65	0	41.04	41.89	80.61	75.54	7122	16	694
10	99.21	0	86.73	61.98	97.90	98.61	0	93.23	63.95	97.93	98.91	0	89.86	62.92	97.90	97.93	7122	16	694
11	99.73	0	79.91	59.88	97.77	97.61	0	99.14	65.58	97.55	98.66	0	88.49	62.38	97.56	97.55	7122	16	694
12	99.60	0	96.88	65.49	99.05	99.79	0	98.31	66.03	99.35	99.69	0	97.59	65.76	99.20	99.35	7122	16	694

Table 4: test lead 2 with all leads.

		N	*S*	*V*	*M avg*
The best	Number	11		12	12
	Precision	99.73		96.88	65.49
The worst	Number	6		9	9
	Precision	98.52		26.11	41.82

Table 5:results lead 2

The best results with Lead **12**.

The worst results with lead **9**.

Lead 3(III)

	Precision					*Recall*					*F1-score*					*Support*		
	N	S	V	M avg	W avg	N	S	V	M avg	W avg	N	S	V	M avg	W avg	accuracy	N	S	V
1	98.55	0	84.08	60.88	97.07	98.46	0	86.74	61.73	97.22	98.50	0	85.39	61.30	97.14	97.22	7122	16	694
2	98.38	0	85.12	61.16	97.00	98.62	0	84.87	61.16	97.20	98.50	0	84.99	61.16	97.10	97.20	7122	16	694
3																			
4	98.99	0	80.84	59.94	97.18	97.89	0	91.79	63.23	97.15	98.44	0	85.96	61.47	97.13	97.15	7122	16	694
5	96.16	0	71.40	55.85	93.77	97.67	0	61.53	53.07	94.27	96.91	0	66.10	54.34	93.98	94.27	7122	16	694
6	99.49	0	81.39	60.29	97.68	97.91	0	96.40	64.77	97.57	98.69	0	88.26	62.32	97.56	97.57	7122	16	694
7	99.58	0	80.36	59.98	97.68	97.68	0	97.84	65.17	97.50	98.62	0	88.24	62.29	97.50	97.50	7122	16	694
8	99.66	0	80.95	60.20	97.80	97.77	0	98.56	65.44	97.64	98.70	0	88.89	62.53	97.63	97.64	7122	16	694
9	99.18	0	70.61	56.60	96.45	95.52	0	93.80	63.11	95.17	97.32	0	80.57	59.30	95.63	95.17	7122	16	694
10	98.21	0	81.77	59.99	96.56	98.08	0	81.41	59.83	96.40	98.15	0	81.59	59.91	96.48	96.40	7122	16	694
11	99.71	0	82.49	60.74	97.98	97.98	0	99.14	65.70	97.88	98.84	0	90.05	62.96	97.86	97.88	7122	16	694
12	99.49	0	98.31	65.93	99.08	99.87	0	97.89	65.92	99.39	99.68	0	98.10	65.93	99.23	99.39	4724	16	474

Table 6:test lead 3 with all leads.

		N	*S*	*V*	*M avg*
The best	Number	11		12	12
	Precision	99.71		98.31	65.93
The worst	Number	5		9	5
	Precision	96.16		70.61	55.85

Table 7:results lead 3.

The best results with Lead **12**.

The worst results with lead **5**.

Lead 4(AVR)

	Precision					*Recall*					*F1-score*						*Support*		
	N	S	V	M avg	W avg	N	S	V	M avg	W avg	N	S	V	M avg	W avg	accuracy	N	S	V
1	98.30	0	81.02	59.77	96.57	98.06	0	84.87	60.98	96.69	98.18	0	82.90	60.36	96.62	96.69	7122	16	694
2	98.91	0	81.22	60.04	97.14	98.05	0	90.35	62.80	97.17	98.48	0	85.54	61.34	97.13	97.17	7122	16	694
3	97.34	0	76.14	57.83	95.27	97.77	0	74.50	57.42	95.51	97.56	0	75.31	57.62	95.38	95.51	7122	16	694
4																			
5	97.39	0	73.43	56.94	95.07	97.58	0	98.88	55.49	94.84	97.49	0	71.08	56.19	94.95	94.84	7122	16	694
6	97.49	0	74.05	57.18	95.21	97.50	0	75.65	57.72	95.37	97.49	0	74.84	57.44	95.29	95.37	7122	16	694
7	99.67	0	75.33	58.33	97.31	96.77	0	98.56	65.11	96.73	98.20	0	85.39	61.20	96.86	96.73	7122	16	694
8	99.67	0	78.92	59.53	97.63	97.43	0	98.70	65.38	97.34	98.64	0	87.71	62.08	97.38	97.34	7122	16	694
9	97.16	0	30.31	42.49	91.04	83.22	0	75.65	52.96	82.38	89.65	0	43.28	44.31	85.36	82.38	7122	16	694
10	98.33	75.00	81.49	84.94	96.79	98.17	37.50	93.72	73.13	96.77	98.25	50.00	82.59	76.95	96.76	96.77	7122	16	694
11	99.66	0	83.86	61.17	98.05	98.13	0	98.85	65.66	98.00	98.89	0	90.74	63.21	97.97	98.00	7122	16	694
12	99.24	0	99.78	66.34	98.99	99.98	0	95.78	65.25	99.29	99.61	0	97.74	65.78	99.13	99.29	4724	16	474

Table 8:test lead 4 with all leads.

		N	*S*	*V*	*M avg*
The best	Number	7 AND 8	10	12	10
	Precision	99.67	75.00	99.78	84.94
The worst	Number	9		9	9
	Precision	97.16		30.31	42.49

Table 9:results lead 4.

The best results with Lead **10**.

The worst results with lead **9**.

Lead 5(AVL)

	Precision					*Recall*					*F1-score*					*Support*		
	N	S	V	M avg	W avg	N	S	V	M avg	W avg	N	S	V	M avg	W avg	accuracy	N	S	V
1	97.22	0	78.55	58.59	95.36	98.06	0	73.34	57.14	95.67	97.64	0	75.86	57.83	95.51	95.67	7122	16	694
2	99.45	0	85.84	61.76	98.04	98.51	0	96.11	64.87	98.10	98.98	0	90.69	63.22	98.04	98.10	7122	16	694
3	98.74	0	79.74	59.49	96.85	97.89	0	88.47	62.12	96.86	98.31	0	93.88	60.73	96.83	96.86	7122	16	694
4	98.66	0	81.13	59.93	96.90	97.99	0	88.62	62.20	96.96	98.32	0	84.71	61.01	96.92	96.96	7122	16	694
5																			
6	99.47	0	81.04	60.17	97.64	97.89	0	96.11	64.67	97.54	98.68	0	87.94	62.20	97.52	97.54	7122	16	694
7	99.70	0	79.70	59.80	97.72	97.47	0	98.99	65.49	97.41	98.57	0	88.30	62.29	97.46	97.41	7122	16	694
8	99.60	0	83.37	60.99	97.96	98.09	0	98.27	65.45	97.91	98.84	0	90.21	63.02	97.87	97.91	7122	16	694
9	99.32	0	49.29	49.54	94.69	90.49	0	95.39	61.96	90.74	94.70	0	65.00	90.74	53.23	91.88	7122	16	694
10	99.17	0	80.45	59.87	97.31	97.81	0	93.66	63.82	97.24	98.49	0	86.55	61.68	97.23	97.24	7122	16	694
11	99.71	0	83.19	60.97	98.05	98.05	0	99.14	65.73	97.94	98.87	0	90.47	63.11	97.93	97.94	7122	16	694
12	99.47	0	99.36	66.28	99.16	99.98	0	97.68	65.89	99.46	99.73	0	98.51	66.08	99.31	99.46	4724	16	474

Table 10:test lead 5 with all leads.

		N	*S*	*V*	*M avg*
The best	Number	11		12	12
	Precision	99.71		99.36	66.28
The worst	Number	1		9	9
	Precision	97.22		49.29	49.54

Table 11:results lead 5.

The best results with lead **12**.

The worst results with lead **9**.

Lead 6(AVF)

	Precision					*Recall*					*F1-score*					*Support*		
	N	S	V	M avg	W avg	N	S	V	M avg	W avg	N	S	V	M avg	W avg	accuracy	N	S	V
1	99.39	0	85.79	61.73	97.98	98.50	0	95.68	64.72	98.05	98.94	0	90.46	63.14	97.99	98.05	7122	16	694
2	98.54	0	86.99	61.85	97.32	98.78	0	86.74	61.84	97.51	98.66	0	86.87	61.84	97.41	97.51	7122	16	694
3	99.54	0	83.17	60.90	97.89	98.16	0	96.83	65.00	97.84	98.85	0	89.84	62.48	97.82	97.84	7122	16	694
4	98.36	0	91.46	63.27	97.55	99.27	0	84.87	61.38	97.79	98.81	0	88.04	62.28	97.66	97.79	7122	16	694
5	99.66	0	78.01	59.22	97.53	97.37	0	98.13	65.17	97.24	98.50	0	86.92	61.81	97.27	97.24	7122	16	694
6																			
7	99.63	0	83.44	61.02	97.99	98.16	0	97.98	65.38	97.94	98.89	0	90.13	63.01	97.91	97.94	7122	16	694
8	99.52	0	82.22	60.58	97.78	97.96	0	97.26	65.08	97.70	98.73	0	89.11	62.61	97.68	97.70	7122	16	694
9	98.88	0	74.64	57.84	96.53	97.08	0	90.35	62.48	96.28	97.97	0	81.75	59.91	96.34	96.28	7122	16	694
10	97.78	0	81.87	59.88	96.17	98.33	0	78.10	58.81	96.34	98.05	0	79.94	59.33	96.25	96.34	7122	16	694
11	99.74	0	85.06	61.60	98.24	98.30	0	99.28	65.86	98.19	99.02	0	91.62	63.55	98.16	98.19	7122	16	694
12	99.60	0	91.96	63.85	98.60	99.17	0	98.95	66.04	98.85	99.38	0	95.33	64.90	98.71	98.85	7122	16	694

Table 12:test lead 6 with all leads.

		N	*S*	*V*	*M avg*
The best	Number	11		12	12
	Precision	99.74		91.96	63.85
The worst	Number	10		9	9
	Precision	97.78		74.64	57.84

Table 13:results lead 6.

The best results with lead **12**.

The worst results with lead **9**.

Lead 7(V2)

	Precision					*Recall*					*F1-score*						*Support*		
	N	S	V	M avg	W avg	N	S	V	M avg	W avg	N	S	V	M avg	W avg	accuracy	N	S	V
1	99.61	0	82.89	60.83	97.93	98.02	0	98.41	65.48	97.85	98.81	0	89.99	62.93	97.83	97.85	7122	16	694
2	99.71	0	81.07	60.26	97.86	97.82	0	98.70	65.51	97.70	98.76	0	89.02	62.59	97.69	97.70	7122	16	694
3	99.56	0	80.43	59.99	97.66	97.70	0	97.69	65.13	97.50	98.62	0	88.22	62.28	97.50	97.50	7122	16	694
4	99.64	0	81.50	60.38	97.83	97.85	0	98.41	65.42	97.70	98.74	0	89.16	62.63	97.63	97.70	7122	16	694
5	99.66	0	79.37	59.68	97.65	97.51	0	98.70	95.41	97.42	98.57	0	87.99	62.19	97.43	97.42	7122	16	694
6	99.60	0	81.41	60.34	97.79	97.87	0	97.84	65.23	97.66	98.73	0	88.87	62.53	97.65	97.66	7122	16	694
7																			
8	99.65	0	70.97	56.87	96.91	96.14	0	98.27	64.80	96.13	97.86	0	82.42	60.09	96.29	96.13	7122	16	694
9	99.65	0	70.97	56.87	96.91	96.14	0	98.27	64.80	96.13	97.86	0	82.42	60.09	96.29	96.13	7122	16	694
10	99.75	50.00	77.11	75.62	97.65	97.11	37.50	98.56	77.72	97.11	98.41	42.86	86.53	75.93	97.25	97.11	7122	16	694
11	99.70	0	81.01	60.24	97.84	97.77	0	98.99	65.59	97.68	98.72	0	89.11	62.61	97.67	97.68	7122	16	694
12	99.62	0	91.63	63.75	98.59	99.11	0	99.37	66.16	98.83	99.36	0	95.34	64.90	98.69	98.83	4724	16	474

Table 14:test lead 7 with all leads.

		N	*S*	*V*	*M avg*
The best	Number	10	10	12	10
	Precision	99.75	50.00	91.63	75.62
The worst	Number	12		8 AND 9	8 AND 9
	Precision	99.62		70.97	56.87

Table 15:results lead 7.

The best results with lead **10**.

The worst results with lead **8 AND 9**.

Lead 8(V2)

	Precision					*Recall*					*F1-score*						*Support*		
	N	S	V	M avg	W avg	N	S	V	M avg	W avg	N	S	V	M avg	W avg	accuracy	N	S	V
1	99.60	0	83.35	60.98	97.96	98.10	0	98.13	65.41	97.91	98.85	0	90.14	63.00	97.87	97.91	7122	16	694
2	99.57	0	82.44	60.67	97.85	98.03	0	97.41	65.15	97.78	98.80	0	89.30	62.70	97.75	97.78	7122	16	694
3	99.64	0	82.21	60.62	97.89	97.94	0	98.56	65.50	97.79	98.78	0	89.65	62.81	97.77	97.79	7122	16	694
4	99.66	0	83.03	60.90	97.98	98.05	0	98.70	65.58	97.91	98.85	0	90.19	63.01	97.88	97.91	7122	16	694
5	99.66	0	83.29	60.98	98.00	98.10	0	98.41	65.51	97.93	98.87	0	90.22	63.03	97.91	97.93	7122	16	694
6	99.73	0	81.76	60.50	97.93	97.92	0	98.85	95.59	97.80	98.82	0	89.50	62.77	97.79	97.80	7122	16	694
7	99.66	0	81.45	60.37	97.84	97.82	0	98.70	65.51	97.70	98.73	0	89.25	62.66	97.69	97.70	7122	16	694
8																			
9	99.54	0	75.73	58.42	97.23	97.02	0	97.12	64.71	96.83	98.27	0	85.10	61.12	96.90	96.83	7122	16	694
10	99.74	40.00	83.82	74.52	98.21	98.12	25.00	98.56	73.89	98.01	98.92	30.77	90.60	73.43	98.05	98.01	7122	16	694
11	99.70	0	82.08	60.59	97.93	97.91	0	98.99	65.63	97.80	98.80	0	89.75	62.85	97.79	97.80	7122	16	694
12	99.60	0	96.30	65.30	98.99	99.66	0	98.95	66.20	99.29	99.63	0	97.61	65.75	99.14	99.29	7122	16	694

Table 16:test lead 8 with all leads.

		N	*S*	*V*	*M avg*
The best	Number	10	10	12	10
	Precision	99.74	40.00	96.30	74.52
The worst	Number	9		9	9
	Precision	99.54		75.73	58.42

Table 17:results lead 8.

The best results with lead **10**.

The worst results with lead **9**.

Lead 9(V3)

	Precision					*Recall*					*F1-score*					*Support*		
	N	S	V	M avg	W avg	N	S	V	M avg	W avg	N	S	V	M avg	W avg	accuracy	N	S	V
1	98.11	0	19.59	39.23	90.95	64.22	0	89.48	51.23	66.33	77.63	0	32.14	36.59	73.44	66.33	7122	16	694
2	98.87	0	52.85	50.57	94.59	92.15	0	90.92	61.02	91.85	95.39	0	66.84	54.08	92.67	91.85	7122	16	694
3	99.22	0	81.93	60.38	97.48	98.01	0	94.09	64.03	97.46	98.61	0	87.59	62.07	97.43	97.46	7122	16	694
4	99.10	0	35.90	45.00	93.30	83.67	0	94.09	59.25	84.42	90.73	0	51.97	47.57	87.11	84.42	7122	16	694
5	99.11	0	64.83	54.65	95.87	95.10	0	93.23	62.78	94.74	97.06	0	76.48	57.85	95.04	94.74	7122	16	694
6	98.77	0	65.99	54.92	95.66	95.56	0	89.48	61.68	94.83	97.14	0	75.96	57.70	95.06	94.83	7122	16	694
7	99.69	0	70.39	56.69	96.89	95.97	0	98.99	64.99	96.04	97.80	0	82.28	60.02	96.22	96.04	7122	16	694
8	98.75	0	72.51	57.09	96.23	96.74	0	89.34	62.03	95.89	97.74	0	80.05	95.97	59.26	95.89	7122	16	694
9																			
10	97.98	0	45.72	47.90	93.15	90.38	0	83.14	57.84	89.56	94.03	0	56.00	51.01	90.73	89.56	7122	16	694
11	99.20	0	49.89	49.70	94.63	90.72	0	94.81	61.84	90.90	94.77	0	65.38	53.38	91.97	90.90	7122	16	694
12	99.02	0	42.61	47.21	93.59	87.45	0	93.67	60.37	87.74	92.87	0	58.58	50.48	89.47	87.74	7122	16	694

Table 18:test lead 9 with all leads.

		N	*S*	*V*	*M avg*
The best	Number	7		3	3
	Precision	99.69		81.93	60.38
The worst	Number	10		1	1
	Precision	97.98		19.59	39.23

Table 19:results lead 9.

The best results with lead **3**.

The worst results with lead **1**.

Lead 10(V4)

	Precision					*Recall*					*F1-score*					*Support*		
	N	S	V	M avg	W avg	N	S	V	M avg	W avg	N	S	V	M avg	W avg	accuracy	N	S	V
1	99.22	0	93.81	64..34	98.53	99.38	0	93.95	64.44	98.70	99.30	0	93.88	64.39	98.62	98.70	7122	16	694
2	99.17	0	86.32	61.83	97.83	98.54	0	93.66	64.07	97.91	98.85	0	89.84	62.90	97.85	97.91	7122	16	694
3	99.10	0	95.95	65.01	98.62	98.76	0	92.07	63.61	97.97	98.93	0	93.97	64.30	98.29	97.97	7122	16	694
4	99.59	0	99.59	66.38	99.39	99.41	0	97.55	65.65	99.04	99.50	0	98.54	66.02	99.21	99.04	7122	16	694
5	99.42	0	89.80	63.07	98.37	98.89	0	96.40	65.10	98.47	99.16	0	92.98	64.05	98.41	98.47	7122	16	694
6	99.30	0	83.42	60.91	97.69	98.16	0	94.96	64.37	97.68	98.73	0	88.81	62.51	97.65	97.68	7122	16	694
7	99.73	0	76.56	58.76	97.47	97.07	0	99.28	65.45	97.06	98.38	0	886.45	61.61	97.12	97.06	7122	16	694
8	99.32	0	98.21	65.84	99.02	99.85	0	95.10	64.98	99.22	99.58	0	96.63	65.40	99.12	99.22	7122	16	694
9	98.06	0	59.06	52.37	94.40	94.38	0	83.14	59.17	93.19	96.19	0	69.06	55.08	93.59	93.19	7122	16	694
10																			
11	99.61	0	82.87	60.83	97.93	98.02	0	98.27	65.43	97.84	98.81	0	89.91	62.91	97.82	97.84	7122	16	694
12	99.54	0	98.51	66.01	99.14	99.98	0	97.47	65.82	99.44	99.76	0	97.99	65.91	99.29	99.44	4724	16	474

Table 20:test lead 10 with all leads.

		N	*S*	*V*	*M avg*
The best	Number	7		4	4
	Precision	99.73		99.59	66.38
The worst	Number	9		9	9
	Precision	98.06		59.06	52.37

Table 21:results lead 10.

The best results with lead **4**.

The worst results with lead **9**.

Lead 11(V5)

	Precision					*Recall*					*F1-score*						*Support*		
	N	S	V	M avg	W avg	N	S	V	M avg	W avg	N	S	V	M avg	W avg	accuracy	N	S	V
1	99.65	0	88.39	62.68	98.44	98.74	0	98.70	65.81	98.53	99.19	0	93.26	64.15	98.46	98.53	7122	16	694
2	99.61	0	92.01	63.87	98.73	99.21	0	97.84	65.68	98.89	99.41	0	94.83	64.75	98.80	98.89	7122	16	694
3	99.67	0	83.68	61.12	98.05	98.12	0	98.99	65.70	98.00	98.89	0	90.69	63.19	97.96	98.00	7122	16	694
4	99.66	0	92.07	63.91	98.79	99.19	0	98.70	65.96	98.94	99.42	0	95.27	64.90	98.85	98.94	7122	16	694
5	99.64	0	84.86	61.50	98.13	98.29	0	98.56	65.62	98.11	98.96	0	91.20	63.39	98.07	98.11	7122	16	694
6	99.73	0	88.92	62.88	98.57	98.81	0	99.42	66.08	98.66	99.27	0	93.88	64.38	98.59	98.66	7122	16	694
7	99.71	0	79.91	59.87	97.75	97.60	0	99.14	65.58	9754	98.64	0	88.49	62.38	97.54	97.54	7122	16	694
8	99.71	0	79.91	59.87	97.75	97.60	0	99.14	65.58	97.54	98.64	0	88.49	62.38	97.54	97.54	7122	16	694
9	99.36	0	72.00	57.12	96.74	96.36	0	95.97	64.11	96.13	97.84	0	82.27	60.04	96.26	96.13	7122	16	694
10	99.61	0	78.94	59.52	97.58	97.46	0	98.27	65.24	97.33	98.52	0	87.55	62.02	97.35	97.33	7122	16	694
11																			
12	99.64	0	94.97	64.87	98.91	99.49	0	99.58	66.36	99.19	99.57	0	97.22	65.60	99.05	99.19	4724	16	474

Table 22:test lead 11 with all leads.

		N	*S*	*V*	*M avg*
The best	Number	6		12	12
	Precision	99.73		94.97	64.87
The worst	Number	9		9	9
	Precision	99.36		72.00	57.12

Table 23:results lead 11.

The best results with lead **12**.

The worst results with lead **9**.

Lead 12(V6)

	Precision					*Recall*					*F1-score*					*Support*		
	N	S	V	M avg	W avg	N	S	V	M avg	W avg	N	S	V	M avg	W avg	accuracy	N	S	V
1	99.62	0	98.33	65.98	99.20	99.87	0	99.16	66.34	99.50	99.75	0	98.74	66.16	99.35	99.50	4724	16	474
2	99.51	0	98.09	65.87	99.08	99.89	0	97.47	65.79	99.37	99.70	0	97.78	65.83	99.22	99.37	4724	16	694
3	99.49	0	98.72	66.07	99.12	99.92	0	97.89	65.94	99.42	99.70	0	98.31	66.00	99.27	99.42	4724	16	474
4	99.60	0	99.37	66.32	99.27	99.96	0	99.16	66.37	99.58	99.78	0	99.26	66.35	99.42	99.58	4724	16	474
5	99.58	0	94.13	64.57	98.78	99.47	0	98.10	65.86	99.04	99.52	0	96.07	65.20	98.90	90.04	4724	16	474
6	99.59	0	97.69	65.75	99.08	99.85	0	98.10	65.98	99.39	99.70	0	97.89	65.87	99.23	99.39	4724	16	474
7	99.58	0	94.35	64.64	98.80	99.45	0	98.73	66.06	99.08	99.51	0	96.49	65.34	98.93	99.08	4724	16	474
8	99.64	0	93.27	64.30	98.75	99.32	0	99.37	66.23	99.02	99.48	0	96.22	65.23	98.88	99.02	4724	16	474
9	99.44	0	58.85	52.76	95.44	93.33	0	96.84	63.39	93.36	96.29	0	73.21	56.50	93.89	93.36	4724	16	474
10	99.54	0	99.15	66.23	99.20	99.94	0	98.52	66.15	99.50	99.74	0	98.84	66.19	99.35	99.50	4724	16	474
11	99.68	0	97.13	65.60	99.14	99.75	0	99.79	66.51	99.44	99.71	0	98.44	66.05	99.29	99.44	4724	16	474
12																			

Table 24:test lead 12 with all leads.

		N	*S*	*V*	*M avg*
The best	Number	1		4	4
	Precision	99.68		99.37	66.32
The worst	Number	9		9	9
	Precision	99.44		58.85	52.76

Table 25:results lead 12.

The Best results with lead **4**.

The worst results with lead **9**.

4.2.2. Summary of Final Results and Best Configuration Selection

Following the preliminary evaluation on the incart subset, the selected lead configurations were retrained and validated on the complete INCART dataset (67,454 samples). The final results confirm distinct optimal configurations for different classes:

- The Best Configuration for Classes S and V is Leads 4 and 10 where The combination of Lead 4 (AVR) and Lead 10 (V4) achieve superior performance for the clinically critical classes;

	presision	*recall*	*F1-score*	*support*
N	99.06	99.59	99.33	58923
S	22.00	22.71	22.35	436
V	98.79	94.81	96.76	8095
Accuracy			98.52	67454
Macro avg	73.28	72.37	72.81	67454
Weighted avg	98.53	98.52	98.52	67454

Table 26:Validation Results of the Selected Lead Pair (Leads 4 and10) on the Full INCART Dataset.

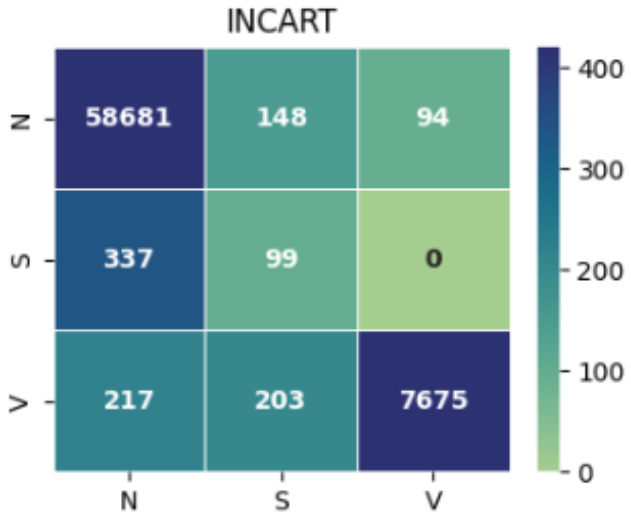

Figure 12:Validation Results of the Selected Lead Pair (Leads 4 and10) on the Full INCART Dataset.

- The Best Configuration for Classes N is Leads 7 and 10 where The combination of Lead 7 (V2) and Lead 10 (V4) achieve superior performance for the clinically critical classes.

	presision	*recall*	*F1-score*	*support*
N	99.09	97.92	98.50	58922
S	17.24	3.44	5.74	436
V	86.71	97.94	91.98	8095
Accuracy			97.31	67453
Macro avg	67.68	66.43	65.41	67453
Weighted avg	97.08	97.31	97.31	67453

Table 27:Validation Results of the Selected Lead Pair (Leads 7 and10) on the Full INCART Dataset.

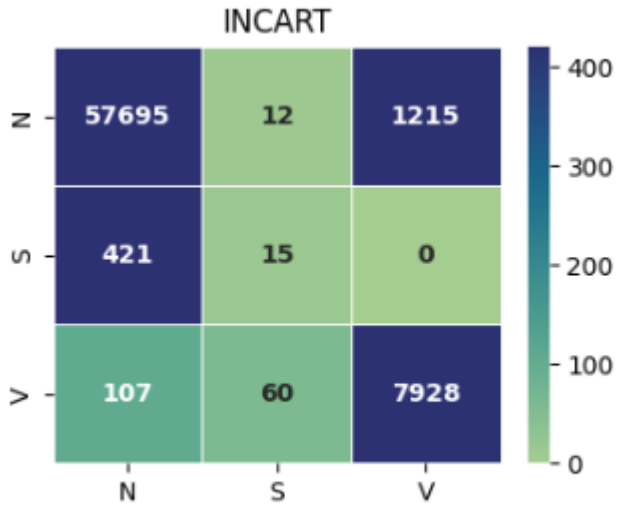

Figure 13:Validation Results of the Selected Lead Pair (Leads 7 and10) on the Full INCART Dataset.

4.3 Results and calculation

The two-stage approach—preliminary selection on a representative subset followed by full-dataset validation—successfully identified lead configurations that balance computational efficiency with robust performance. The final models demonstrate readiness for real-world deployment, with Leads 4 and10 optimized for critical arrhythmia detection and Leads 7and 10 for comprehensive cardiac monitoring.

However, a critical limitation must be acknowledged: Class S (Supraventricular ectopic beats) consistently demonstrated poor performance across all configurations, with Precision and Recall values near zero in the preliminary phase and only marginal improvement in the final validation (Precision: 17.24%, Recall: 3.44% for Leads 4 and10). This persistent weakness is attributed to three primary factors:

Severe Class Imbalance: Class S represented only 16 samples (0.2%) in the preliminary subset and 436 samples (0.65%) in the complete INCART dataset, compared to 58,922 samples for Class N and 8,095 for Class V. This extreme underrepresentation prevented the model from learning discriminative features effectively.

Morphological Similarity: Supraventricular beats often exhibit subtle differences from normal beats, making them inherently difficult to distinguish even with optimal lead configurations.

Model Bias: The classifier naturally prioritized the majority classes (N and V) during training, resulting in systematic misclassification of the rare Class S instances.

Future work should address this limitation through targeted strategies such as oversampling techniques (SMOTE), class-weighted loss functions, or data augmentation specifically for supraventricular events to achieve balanced multi-class performance without compromising the strong detection of clinically critical ventricular arrhythmias.

4.4 Conclusion

This chapter presented a two-stage evaluation of twelve lead configurations for heartbeat classification. Initially, a representative subset of the incart dataset was used to identify promising configurations based on Precision and Accuracy. The selected leads were then retrained and validated on the complete INCART dataset (67,454 samples

Conclusion

This thesis presented a systematic approach to automated ECG heartbeat classification by combining morphological and temporal analysis within a hybrid CNN-MLP framework. By transforming one-dimensional dual-lead signals into two-dimensional PMAT images and extracting RR-interval features, the model successfully captured both waveform shapes and rhythm dynamics. A comprehensive evaluation of all possible lead pair configurations revealed two optimal setups for clinical applications. The combination of Lead 4 (aVR) and Lead 10 (V4) achieved the highest performance for detecting critical arrhythmias, reaching an overall accuracy of 98.52%. Meanwhile, the pairing of Lead 7 (V2) and Lead 10 (V4) proved most effective for general cardiac monitoring, with an accuracy of 97.31%. Despite these strong results, the classification of Supraventricular beats remained challenging due to severe class imbalance and their morphological similarity to normal heartbeats. This limitation highlights the need for targeted strategies in future research, such as advanced data augmentation, class-weighted loss functions, or specialized sampling techniques to improve minority class detection. Overall, this work demonstrates that a dual-lead, hybrid deep learning approach can significantly enhance the reliability and efficiency of automated ECG analysis. By identifying the most informative lead combinations, the proposed system offers a practical tool that could support clinicians in routine screening and long-term cardiac monitoring. Future developments may focus on optimizing the model for real-time deployment and expanding the dataset to further improve generalization across diverse patient populations. Ultimately, this research contributes a reproducible and clinically relevant framework that advances the role of artificial intelligence in modern cardiology.

References

[1] Mokhtari, R., Belhouari, S. B., Kassoul, K., & Hocini, A. (2025). ECG heartbeat classification using progressive moving average transform. Scientific Reports, 15, 4285. https://doi.org/10.1038/s41598-025-88119-9

[2] Guyton, A.C. and Hall, J.E., 2020. Textbook of Medical Physiology. 14th ed. Philadelphia: Elsevier.

[3] Braunwald, E., 2019. Heart Disease: A Textbook of Cardiovascular Medicine. 11th ed. Philadelphia: Elsevier.

[4] Thaler, M.S., 2021. The Only EKG Book You'll Ever Need. 10th ed. Philadelphia: Wolters Kluwer.

[5] Thaler, M.S., 2021. The Only EKG Book You'll Ever Need. 10th ed. Philadelphia: Wolters Kluwer. (Comprehensive ECG interpretation and normal values).

[6] Goldberger, A.L., Goldberger, Z.D. and Shvilkin, A., 2017. Goldberger's Clinical Electrocardiography: A Simplified Approach. 9th ed. Philadelphia: Elsevier. (ECG waves, intervals, and clinical significance).

[7] Surawicz, B. and Knilans, T.K., 2008. Chou's Electrocardiography in Clinical Practice. 6th ed. Philadelphia: Saunders Elsevier. (Detailed ECG measurements and diagnostic criteria).

[8] Goldberger, A.L., Goldberger, Z.D. and Shvilkin, A., 2017. Goldberger's Clinical Electrocardiography: A Simplified Approach. 9th ed. Philadelphia: Elsevier. (ECG leads, Einthoven's triangle, and lead systems)

[9] Pan, J. and Tompkins, W.J., 1985. A real-time QRS detection algorithm. IEEE Transactions on Biomedical Engineering, BME-32(3), pp.230-236. (Classic QRS detection algorithm and traditional signal processing)

[10] PhysioNet. (2023). INCART Database. https://physionet.org/content/incartdb/, access date :25/12/2025.

[11] Moody, G. B., & Mark, R. G. (2001). The impact of the MIT-BIH Arrhythmia Database. IEEE Engineering in Medicine and Biology Magazine, 20(3), 45-

[12] Goldberger, A. L., et al. (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. Circulation, 101(23), e215-e220.

[13] Clifford, G. D., Azuaje, F., & McSharry, P. (2006). Advanced Methods and Tools for ECG Data Analysis. Artech House.

[14] Sörnmo, L., & Laguna, P. (2005). Bioelectrical Signal Processing in Cardiac and Neurological Applications. Elsevier Academic Press.

[15] Luo, S., & Johnston, P. (2002). A review of electrocardiogram filtering. Journal of Electrocardiology, 35(4), 233-238.

[16] Kohler, B. U., Hennig, C., & Orglmeister, R. (2002). The principles of software QRS detection. IEEE Engineering in Medicine and Biology Magazine, 21(1), 42-57.

[17] Chou, T. C., & Knilans, T. K. (1998). Electrocardiography in Clinical Practice (4th ed.). W.B. Saunders Company.

[18] Moody, G. B., & Mark, R. G. (2001). The impact of the MIT-BIH Arrhythmia Database. IEEE Engineering in Medicine and Biology Magazine, 20(3), 45-50.

[19] Pan, J., & Tompkins, W. J. (1985). A real-time QRS detection algorithm. IEEE Transactions on Biomedical Engineering, BME-32(3), 230-236.

[20] Kiranyaz, S., Ince, T., & Gabbouj, M. (2016). Real-time patient-specific ECG classification by 1-D convolutional neural networks. IEEE Transactions on Biomedical Engineering, 63(3), 664-675.

[21] Goldberger, A. L., et al. (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. Circulation, 101(23), e215-e220.

[22] Thakor, N. V., & Zhu, Y. S. (1991). Applications of adaptive filtering to ECG analysis: Noise cancellation and arrhythmia detection. IEEE Transactions on Biomedical Engineering, 38(8), 829-837.

[23] Jayachandran, M., et al. (2010). ECG signal analysis using DWT and PCA. Computers in Biology and Medicine, 40(1), 63-70.

[24] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.