Mohammed Nakib Ansari 174

Phase I Report Final Batch 174(3).pdf

Project Report - Phase 1

Capstone 25-26

KPR Institute of Engineering and Technology

Document Details

Submission ID

trn:oid:::1:3433093836

Submission Date

Dec 4, 2025, 1:47 PM GMT+5:30

Download Date

Dec 4, 2025, 1:49 PM GMT+5:30

File Name

Phase_I_Report_Final_Batch_174_3_.pdf

File Size

950.9 KB

50 Pages

9,172 Words

56,986 Characters

9% Overall Similarity

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

Filtered from the Report

- Bibliography
- Quoted Text

Match Groups

83 Not Cited or Quoted 9%
Matches with neither in-text citation nor quotation marks

1 Missing Quotations 0%
Matches that are still very similar to source material

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

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

Top Sources

6% Internet sources
6% Publications
0% Submitted works (Student Papers)

Match Groups

- **83** Not Cited or Quoted 9%
  Matches with neither in-text citation nor quotation marks
- **1** Missing Quotations 0%
  Matches that are still very similar to source material
- **0** Missing Citation 0%
  Matches that have quotation marks, but no in-text citation
- **0** Cited and Quoted 0%
  Matches with in-text citation present, but no quotation marks

Top Sources

6% Internet sources
6% Publications
0% Submitted works (Student Papers)

Top Sources

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

1	Internet	
arxiv.org		<1%
2	Internet	
pmc.ncbi.nlm.nih.gov		<1%
3	Publication	
Sukhpreet Kaur, Amanpreet Kaur, Manish Kumar. "Recent Advances in Computati...		<1%
4	Publication	
Mahmood, Arafat. "A Package of Smartphone and Sensor-Based Objective Measu...		<1%
5	Internet	
www.mdpi.com		<1%
6	Internet	
multiplatform.ai		<1%
7	Publication	
Deepak Kumar, Brijesh Bakariya, Chaman Verma, Zoltan Illes. "LivXAI-Net: An exp...		<1%
8	Publication	
Omar AlZoubi, Saja Khaled Tawalbeh, Mohammad AL-Smadi. "Affect detection fro...		<1%
9	Internet	
bth.diva-portal.org		<1%
10	Internet	
www.transparencymarketresearch.com		<1%

11	Internet	
hal.science		<1%
12	Publication	
"Artificial Intelligence in Medicine", Springer Science and Business Media LLC, 2023		<1%
13	Internet	
d-nb.info		<1%
14	Internet	
www.coursehero.com		<1%
15	Internet	
www.ijfmr.com		<1%
16	Publication	
Sardar Ansari, Negar Farzaneh, Marlena Duda, Kelsey Horan et al. "A Review of A...		<1%
17	Internet	
kth.diva-portal.org		<1%
18	Internet	
oxleas.koha-ptfs.co.uk		<1%
19	Internet	
thelittleaussiebakery.com		<1%
20	Internet	
ebin.pub		<1%
21	Internet	
ir.uitm.edu.my		<1%
22	Internet	
iscap.us		<1%
23	Publication	
Asali, Ehsan. "Spatiotemporal Multimodal Representation Learning for Activity U...		<1%
24	Publication	
Vila Muñoz, Maria del Mar. "Carotid Artery Ultrasound Image-Based Cardiovascul...		<1%

25 Publication
C. MUELLER. "Right bundle branch block and long-term mortality in patients with ... <1%

26 Internet
psasir.upm.edu.my <1%

27 Internet
www.keyurparikh.com <1%

28 Publication
Hui Yu, Jinqiu Li, Lixin Zhang, Yuzhen Cao, Xuyao Yu, Jinglai Sun. "Design of lung n... <1%

29 Publication
Wei Wang, Jing Ma, Kuanquan Wang, Yuxin Lin, Suyu Dong, Xiaokun Li, Gongning ... <1%

30 Internet
rodin.uca.es <1%

31 Internet
www.sciencexcel.com <1%

32 Internet
123dok.org <1%

33 Publication
Junsang Park, Junho An, Jinkook Kim, Sunghoon Jung, Yeongjoon Gil, Yoojin Jang, ... <1%

34 Publication
Li, Yixin. "Cyber-Physical System in Healthcare: Real-Time Cardiovascular Disease ... <1%

35 Internet
studentsrepo.um.edu.my <1%

36 Internet
www.diva-portal.org <1%

37 Internet
www.ijese.org <1%

38 Publication
Mohammad Faisal Khan, Mohammad AbaOud. "Blockchain-Integrated Security fo... <1%

39	Publication	
Muhammad Shakeel Akram, Bogaraju Sharatchandra Varma, Aqib Javed, Jim Har...		<1%
40	Publication	
Paul Innerd. "Physical Activity Assessment - A Lifecourse Approach", Routledge, 2...		<1%
41	Internet	
cinc.org		<1%
42	Internet	
docplayer.net		<1%
43	Internet	
ediss.uni-goettingen.de		<1%
44	Internet	
gyan.iitg.ac.in		<1%
45	Internet	
upcommons.upc.edu		<1%
46	Internet	
www.cureus.com		<1%
47	Internet	
www.frontiersin.org		<1%
48	Internet	
www.ijisrt.com		<1%
49	Internet	
www.medrxiv.org		<1%
50	Internet	
www.propulsiontechjournal.com		<1%
51	Publication	
"Euroanaesthesia 2004: Joint Meeting of the European Society of Anaesthesiologi...		<1%
52	Publication	
Gang Zhao, Dian Wang. "A Multiple Criteria Decision-Making Method Generated b...		<1%

53 Publication

Salwa Belaqziz, Salma El Hajjami, Hicham Amellal, Redouan Lahmyed, Lahcen Kou... <1%

54 Publication

Sukhpreet Kaur, Amanpreet Kaur, Manish Kumar. "Recent Advances in Computati... <1%

55 Publication

Kirti Aggarwal, Anuja Arora, Zahid Akhtar, Alessandro Bruno. "Computational Int... <1%

56 Publication

Sabuj, Hasibul Hasan. "Fed-Ensemble: Enhancing Federated Learning With Ensem... <1%

57 Publication

Ying Ge, Feilong Shen, Petr Sklenička, Jan Vymazal, Marek Baxa, Zhongbing Chen.... <1%

CARDIOVASCULAR DISEASE PREDICTION USING DEEP LEARNING TECHNIQUES

PHASE I REPORT

Submitted by

MOHAMMED NAKIB ANSARI (711322105089)
RAKESH YADAV (711322105130)
ADITYA PATEL (711322102004)

in partial fulfillment for the award of the degree of

BACHELOR OF ENGINEERING

in

BIOMEDICAL ENGINEERING

COMPUTER SCIENCE AND ENGINEERING

KPR INSTITUTE OF ENGINEERING AND TECHNOLOGY
(AUTONOMOUS)
ARASUR, COIMBATORE 641 407

ANNA UNIVERSITY :: CHENNAI 600 025

D E C E M B E R 2 0 2 5

ANNA UNIVERSITY : CHENNAI 600025

BONAFIDE CERTIFICATE

Certified that this project report titled **"CARDIOVASCULAR DISEASE PREDICTION USING DEEP LEARNING TECHNIQUES"** is the bonafide work of **"MOHAMMED NAKIB ANSARI, RAKESH YADAV and ADITYA PATEL"** who carried out the project work under my supervision.

SIGNATURE
Dr.DEVI PRIYA R

HEAD OF THE DEPARTMENT,
Department of Computer Science and Engineering,
KPR Institute of Engineering and Technology, Arasur,
Coimbatore – 641 407.

SIGNATURE
Ms.NANDHINI N
ASSISTANT PROFESSOR I / CS,
SUPERVISOR,
Department of Computer Science and Engineering,
KPR Institute of Engineering and Technology, Arasur,
Coimbatore – 641 407.

Submitted for project viva-voce examination conducted on …….

INTERNAL EXAMINER **EXTERNAL EXAMINER**

ACKNOWLEDGMENT

We express our sincere thanks and profound gratitude to **Dr. K. P. Ramasamy**, Chairman, and **Ms. Gayathri Anandakrishnan**, Secretary, KPR Group of Institutions, for providing an excellent academic environment and for their visionary leadership and unwavering support for research and innovation.

We are deeply grateful to **Dr. A. M. Natarajan**, Chief Executive; **Dr. R. Devi Priya**, Principal; and **Dr. G. Anusha**, Vice Principal, KPR Institute of Engineering and Technology, for their esteemed guidance and for fostering a culture of academic excellence.

Our sincere gratitude to the project leadership team, including **Mr. G. K. Jakir Hussain**, Institute Project Coordinator; **Dr. R. Sampathkumar** and **Mr. D. Premkumar**, Institute Project Associate Coordinators; **Dr. Gayathri**, Project Chief Mentor and **Dr. V. Priya**, Project Coordinator, for their structured oversight and timely advice.

We are especially thankful to **Dr. Devi Priya R**, Head, Department of **Computer Science and Engineering**, for her support. We owe our deepest gratitude for our project supervisor, **Ms. Nandhini N, Assistant Professor,** Department of **Computer Science and Engineering**, whose expert guidance, immense patience, and constant encouragement have been the cornerstones of this endeavour.

Finally, we acknowledge **Mr. Vasanth N**, **CTO of Pinesphere**, for his valuable industrial insights. We also extend our heartfelt thanks to our beloved family, faculty members, supporting staffs and friends for their unwavering support and motivation.

MOHAMMED NAKIB ANSARI (711322105089)
RAKESH YADAV (711322105130)
ADITYA PATEL (711322102004)

ABSTRACT

Cardiovascular diseases (CVDs) remain the leading cause of global mortality, claiming an estimated 17.9 million lives annually according to the World Health Organization (WHO). The Electrocardiogram (ECG) is the primary diagnostic tool for detecting cardiac anomalies; however, manual interpretation of 12-lead ECG signals is time-consuming, prone to inter-observer variability, and susceptible to errors caused by clinician fatigue. This project proposes and implements a Hybrid Deep Learning System for the multi-label classification of cardiovascular diseases. Unlike traditional models that rely solely on raw signal processing, this system employs a Multi-Modal Data Fusion architecture. The model integrates temporal features extracted from 12-lead ECG signals using a 1D-Convolutional Neural Network (1D-CNN) with clinical demographic data (Age and Biological Sex) processed via a Dense Neural Network. This fusion approach mimics clinical decision-making, where patient history is vital for accurate diagnosis.The model was trained and validated on the PTB-XL dataset, a large-scale, annotated database of clinical ECGs. The system is designed to detect five specific, critical cardiac conditions: Normal Sinus Rhythm, Myocardial Infarction, Atrial Fibrillation, ST-Depression, and general Arrhythmias. The final system is deployed as a user-friendly web application using Streamlit, featuring dual modes of operation: a Single-Patient Analysis mode for detailed clinician review.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

ABBREVIATION	FULL FORM
AFIB	Atrial Fibrillation
AUC	Area Under the Curve
BCE	Binary Cross-Entropy
CAD	Coronary Artery Disease
DWT	Discrete Wavelet Transform
ECG	Electrocardiogram
CVD	Cardiovascular Disease
GRU	Gated Recurrent Unit
PTB-XL	Physikalisch-Technische Bundesanstalt ECG Dataset
ROC	Receiver Operating Characteristic
SNR	Signal-to-Noise Ratio
SCP-ECG	Standard Communications Protocol for ECG
STTC	ST/T Wave Changes
SNOMED-CT	Systematized Nomenclature of Medicine - Clinical Terms
SVR/SVM	Support Vector Regression / Support Vector Machine
WFDB	Waveform Database (PhysioNet Library)
XAI	Explainable Artificial Intelligence

CHAPTER 1

INTRODUCTION

1.1 Overview of Cardiovascular Diseases

The cardiovascular system plays a central role in sustaining human life. It is made up of the heart an intricate muscular pump and an extensive network of blood vessels that extend throughout the body. Together, they maintain the continuous flow of blood necessary for transporting oxygen, nutrients, hormones, and immune cells to tissues, while at the same time removing waste products such as carbon dioxide and other metabolic by-products. Any disruption in this delicate system can have immediate consequences because the body's organs rely on a steady and well-regulated blood supply to function normally.

Cardiovascular Diseases (CVDs) encompass a broad range of disorders that affect the heart or the blood vessels. These conditions range from structural problems, such as valve defects, to electrical abnormalities in the heart's rhythm, to blockages in the coronary arteries. As emphasized by the World Health Organization (WHO), CVDs are the leading cause of global mortality, responsible for approximately 17.9 million deaths every year, a staggering 32% of all global deaths. This statistic highlights the magnitude of the problem and demonstrates how critical it is to develop better diagnostic and preventive strategies.

The burden of cardiovascular illness extends far beyond mortality figures.

CVDs contribute significantly to disability, reduced quality of life, and financial strain for patients and their families. In many low- and middle-income nations, the situation is further complicated by limited access to specialized cardiac care, a shortage of trained cardiologists, and inadequate healthcare infrastructure. These constraints often delay diagnosis and treatment, leading to worse outcomes.

Among the wide spectrum of cardiovascular disorders, certain conditions are particularly common and life-threatening. Coronary Artery Disease (CAD), which results from the narrowing or blockage of coronary arteries, remains the most prevalent. Arrhythmias, which refer to abnormalities in the heart's electrical rhythm, and Heart Failure, where the heart is unable to pump efficiently, are also major contributors to morbidity. Within these, two conditions are especially critical: Myocardial Infarction (MI) and Atrial Fibrillation (AFIB). MI or heart attack results from a sudden blockage in a coronary artery, depriving heart tissue of oxygen. AFIB, a rapid and irregular beating of the atria, significantly increases the risk of stroke if left untreated. Both conditions require timely and accurate diagnosis to avoid permanent damage or death.

The causes of CVDs are rarely singular. Instead, they arise from an interplay of multiple factors. Genetics may predispose certain individuals to heart disease, while lifestyle choices such as smoking, poor diet, excessive alcohol intake, and physical inactivity further increase risk. In addition, biological factors including age and sex play a significant role. The likelihood of cardiovascular events increases with age, and symptoms often manifest differently in males and females, making diagnosis more challenging. For example, younger males may show classical chest pain, whereas females may present with atypical symptoms such as fatigue or nausea.

Given these complexities, a diagnostic system must consider not only the electrical or mechanical signals of the heart but also demographic and clinical context. A heart rate that appears low for one patient may be normal for another depending on age, fitness level, or underlying medical conditions. Therefore, integrating signal data with demographic information is critical for producing accurate and clinically meaningful assessments.

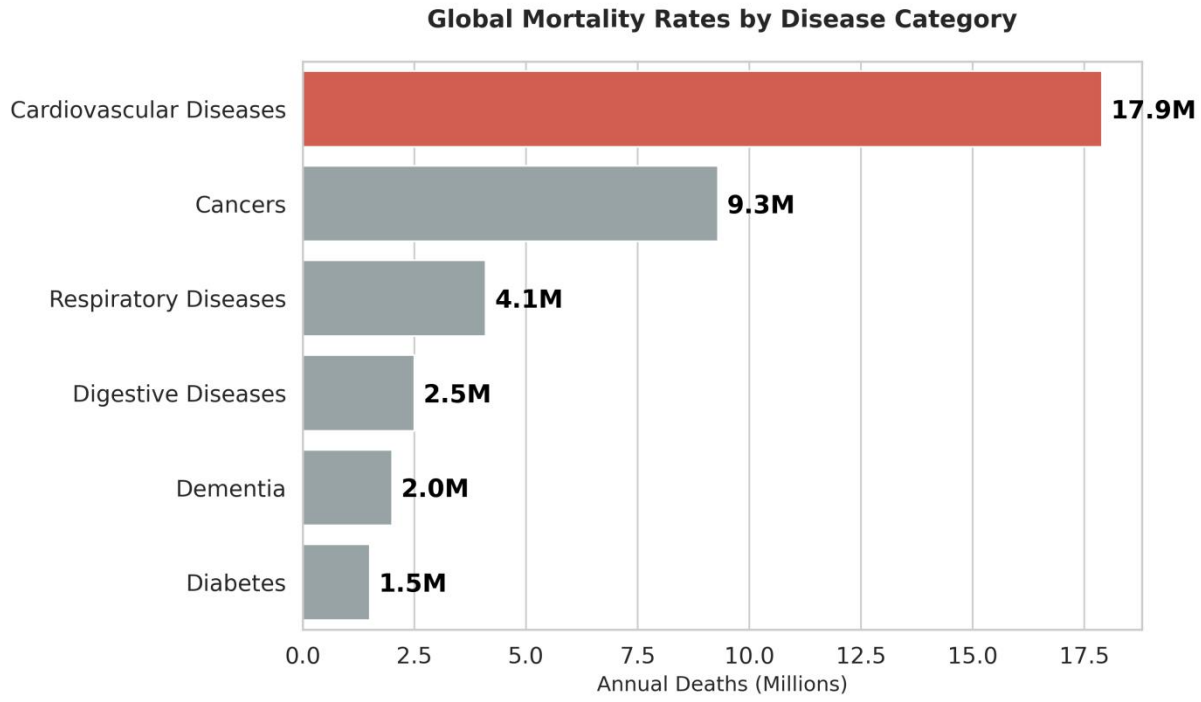

FIGURE 1.1: Chart showing Global Mortality Rates by Disease

1.2 The Electrocardiogram (ECG) as a Diagnostic Standard

25 The Electrocardiogram (ECG) has been a cornerstone of clinical cardiology for more than a century. Introduced by Willem Einthoven in the early 1900s, its ability to monitor the heart's electrical activity in a non-invasive manner revolutionized cardiac diagnosis. Since then, the ECG has remained one of the most accessible, affordable, and reliable tools for detecting a wide range of cardiac conditions.

At its core, the ECG measures the tiny electrical currents generated by the heart as it undergoes depolarization and repolarization. Electrodes placed at specific locations on the skin capture these signals, which are then displayed as waveforms. The shape, timing, and magnitude of these waves reveal vital information about the heart's rhythm, the conduction of electrical impulses, and even the presence of structural abnormalities. Because many cardiac diseases produce changes in these electrical patterns, the ECG remains an indispensable tool in emergency care, routine screening, and long-term monitoring.

1.2.1 Physiology of the ECG Waveform

A standard ECG waveform consists of several distinct components, each corresponding to a particular phase of the cardiac cycle. The P-wave is a small, rounded wave that represents the depolarization of the atria, the upper chambers of the heart. In a normal heart, it appears smooth and uniform, but when the P wave becomes absent, chaotic, or irregular, it can indicate conditions such as atrial fibrillation or atrial enlargement. The QRS complex is the tallest and sharpest portion of the ECG, reflecting the rapid depolarization of the ventricles, which are responsible for pumping blood to the lungs and the rest of the body. Variations in the duration, shape, or amplitude of the QRS complex may suggest conduction disorders such as bundle branch blocks or ventricular hypertrophy. Following the QRS complex, the T-wave represents ventricular repolarization, a phase when the ventricles reset in preparation for the next heartbeat. A T-wave that appears inverted, elevated, or flattened may signal ischemia, electrolyte imbalance, or an evolving myocardial infarction. Together, these waves form a repeating pattern that reveals information not only about the heart's rhythm but also about its overall health. By examining the intervals between these waves, clinicians can identify abnormalities such as tachycardia, bradycardia, conduction delays, and various other cardiac conditions.

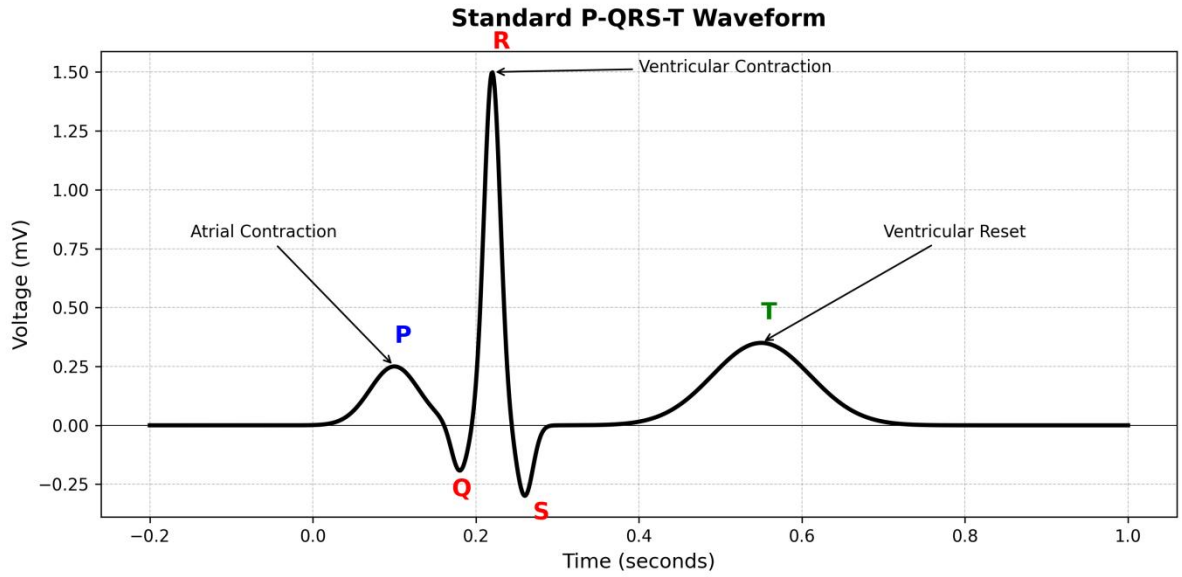

FIGURE 1.2: A labeled diagram of a Standard P-QRS-T Waveform

1.2.2 The Standard 12-Lead System

Clinical ECG interpretation relies on the 12-lead ECG, which provides a comprehensive view of the heart from multiple angles. Although only ten electrodes are placed on the body, twelve different electrical perspectives, known as leads, are generated mathematically. The limb leads (I, II, III, aVR, aVL, aVF) observe the heart in the frontal or vertical plane and are particularly useful for identifying arrhythmias, rhythm abnormalities, and certain types of ischemia. The precordial leads (V1-V6), placed across the chest, examine the heart in the horizontal plane and help clinicians determine the exact region of the heart affected by structural or ischemic changes. The 12-lead system enables physicians to pinpoint the location of myocardial damage, identify conduction abnormalities, and detect subtle changes that may suggest early disease. For example, specific lead patterns can distinguish between inferior, anterior, or lateral heart attacks. However, because of the large amount of data collected from all 12 leads, manually interpreting ECG especially in time critical situations can be both challenging and prone to error.

1.3 Evolution of Automated Diagnosis

Efforts to automate ECG interpretation have evolved over several decades, progressing through distinct technological phases that reflect advancements in computational capability and medical informatics. During the Rule-Based Systems era (1970s-1990s), early automation relied on manually coded clinical rules that used logic-based instructions—for example, detecting irregular R-R intervals to identify arrhythmias. Although these systems were transparent and easy to modify, they lacked flexibility, performed poorly with noisy or incomplete data, and struggled with complex or atypical ECG patterns, resulting in limited diagnostic accuracy. The next phase, Traditional Machine Learning (2000s-2015), introduced algorithms that utilized manually engineered features derived from mathematical techniques such as Fourier transforms, wavelet analysis, and statistical modeling. These features were fed into classifiers like Support Vector Machines, Random Forests, or k-Nearest Neighbors, yielding improved diagnostic precision but requiring substantial domain expertise and labor-intensive feature engineering, which constrained scalability and adaptability. The Deep Learning Era (2015-present) revolutionized ECG interpretation by enabling models, particularly 1D-Convolutional Neural Networks (1D-CNNs), to learn directly from raw signal data. These networks autonomously extract hierarchical and non-linear features, capturing subtle physiological patterns that traditional methods could not. As a result, deep learning approaches have achieved unprecedented levels of accuracy, robustness, and generalizability, establishing a new standard for automated cardiovascular diagnostics.

1.4 Problem Statement

Despite significant advancements, several challenges continue to hinder the widespread adoption of AI-driven cardiac diagnostic systems in real clinical

environments. One major issue is the lack of multi-modal integration, as most existing systems analyze ECG signals in isolation and fail to account for demographic factors such as age and sex elements that clinicians routinely consider when interpreting ECG. Without this contextual information, AI models may generate misleading or inappropriate predictions. Another challenge is the black-box nature of many deep learning models, which often provide predictions without explaining the reasoning behind them. This lack of interpretability reduces trust among healthcare professionals, who need transparency to justify clinical decisions. Additionally, many datasets and studies operate under a single-label assumption, treating each ECG as representative of only one disease category, even though patients frequently exhibit multiple co-existing conditions. A truly practical diagnostic system must therefore be capable of making multi-label predictions. Lastly, AI models are highly sensitive to noise, as real-world ECG recordings often contain interference from muscle activity, respiration, or nearby electronic devices. Systems trained on clean, high-quality academic data may struggle to perform accurately when exposed to noisy or incomplete clinical inputs.

1.5 Objectives of the Project

The primary objective of this project is to develop a comprehensive, reliable, and clinically relevant Multi-Modal Deep Learning Framework for the detection of major cardiovascular diseases. The proposed model is designed not only to classify diseases but also to interpret demographic factors, ensuring that its predictions align with real-world clinical decision-making practices. To achieve this overarching goal, several specific objectives have been identified.

First, data processing and preprocessing aim to optimize the quality and consistency of ECG signals through rigorous preparation techniques. This includes noise reduction using filtering methods to eliminate baseline wander,

powerline interference, and motion artifacts; normalization strategies such as Z-score normalization to ensure uniform scaling; downsampling to minimize computational overhead while retaining critical features; and segmentation of ECG recordings to meet neural network input requirements. This robust preprocessing pipeline ensures that the neural network receives clean and representative data.

Second, the hybrid multi-modal neural network architecture focuses on developing a dual-branch Deep Learning model capable of processing both ECG waveform data and demographic information. A 1D-Convolutional Neural Network (1D-CNN) extracts spatial and temporal features from ECG signals, while a Dense Neural Network processes demographic factors like age and sex to provide clinical context. This design enables the model to combine physiological and demographic insights, leading to more holistic diagnostic predictions, with a modular and interpretable architecture ensuring each branch contributes meaningfully to the final outcome.

Third, the framework emphasizes multi-label classification of cardiac diseases, recognizing that patients often present with multiple coexisting conditions. The model is trained to detect Normal Sinus Rhythm, Myocardial Infarction, Atrial Fibrillation, ST-Depression, and Arrhythmias simultaneously, improving diagnostic accuracy and clinical relevance by identifying overlapping conditions rather than misclassifying them.

Fourth, risk stratification and clinical interpretation are integrated to translate raw predictions into actionable clinical insights. A logic-based risk assessment module categorizes results into high, moderate, or low-risk levels, guiding appropriate medical responses-from urgent intervention to monitoring or reassurance. This feature bridges the gap between algorithmic outputs and real-world clinical application.

Finally, deployment and usability are central to ensuring accessibility and adoption. The framework will be implemented as an interactive, user-friendly application using Streamlit, supporting both single-patient analysis with detailed visualizations and batch processing for large-scale hospital or research use. Outputs will be displayed in a clear and interpretable format suitable for both specialists and general practitioners. Together, these objectives create a practical, efficient, and clinically meaningful system for AI-assisted cardiac diagnostics.

1.6 Scope and Limitations

The scope of this project encompasses several well-defined areas within the broader field of automated cardiac diagnosis. It begins with dataset utilization, where the PTB-XL dataset is employed as the foundation for training, testing, and validating the model. This dataset offers high-quality, annotated 12-lead ECG recordings covering a diverse range of cardiac conditions, making it ideal for developing and evaluating the proposed framework. The project also focuses on major super-classes, concentrating on five primary cardiovascular categories that represent the most clinically significant abnormalities typically observed in cardiology practice. Additionally, the development of a multi-modal deep learning framework allows for the integration of ECG waveform data with demographic features, offering a more comprehensive and context-aware diagnostic approach compared to traditional single-input models. The system is further enhanced through web-based deployment, achieved via a Streamlit interface that enables clinicians to interact with model outputs in a visually intuitive dashboard outside of research environments. Finally, the inclusion of clinical decision support ensures that the system provides preliminary assessments and risk alerts to assist healthcare professionals, functioning as a supplementary diagnostic aid rather than an autonomous diagnostic solution.

Despite its strengths, the project also has several limitations that must be acknowledged. The first is dependence on input signal quality, as deep learning models require clean and well-captured ECG recordings for optimal performance. In real-world clinical settings, noise caused by patient movement, electrode misplacement, or equipment interference can lead to inaccurate predictions or misclassifications. The second limitation is restricted demographic features, since the model only utilizes age and sex due to dataset constraints. Other relevant variables such as smoking status, family history, cholesterol levels, body mass index, and medication history—could enhance diagnostic precision but are unavailable in the PTB-XL dataset. A third limitation involves dataset bias, as PTB-XL originates from a specific geographic and clinical environment, potentially introducing biases in disease representation or demographic distribution. Consequently, model retraining or fine-tuning may be necessary before applying it to different populations. The fourth limitation pertains to interpretability challenges, because deep learning models, while highly effective, function as complex mathematical structures that may be difficult for clinicians to fully trust without transparent reasoning. Lastly, the project acknowledges that the system is not a replacement for medical expertise.

CHAPTER 2

LITERATURE REVIEW

2.1 Overview of Artificial Intelligence in Cardiology

Artificial Intelligence has become essential in cardiology due to the large volume of high-dimensional ECG data that cannot be manually analyzed at scale. Early AI relied on rule-based systems and classical machine learning, but these methods often failed to generalize across diverse patient groups. The introduction of Deep Learning (DL) significantly improved performance by learning patterns directly from raw ECG waveforms without manual feature design. DL methods now excel in tasks such as arrhythmia detection, ischemia identification, and multi-class cardiac disease classification, making them central to modern automated diagnosis.

2.2 Traditional Machine Learning Approaches

Traditional machine learning methods for ECG analysis relied heavily on manual feature engineering, where experts extracted mathematical or statistical descriptors to represent important waveform characteristics. These systems followed a structured pipeline consisting of preprocessing, feature extraction, feature selection, and classification. Their overall performance depended entirely on the quality of the engineered features.

Early work using the Discrete Wavelet Transform (DWT) extracted multi-scale frequency components of ECG signals and fed statistical measures into classifiers such as SVM and k-NN. Although effective for identifying

arrhythmic patterns, these methods were highly sensitive to R-peak detection errors and often performed poorly with noisy or artifact-contaminated signals.

Ensemble-based approaches like Random Forests used time-domain features such as RR intervals and HRV to improve robustness. They handled nonlinear patterns well and offered interpretability through feature importance analysis. However, they struggled to detect fine morphological variations, leading to reduced accuracy for conditions involving subtle waveform changes.

2.3 The Deep Learning Revolution

Deep Learning introduced a major shift in ECG analysis by learning hierarchical features directly from raw signals rather than relying on engineered descriptors. Early work experimented with 2D-CNNs by converting ECG traces into images, but this approach lost temporal continuity and added computational overhead. The field progressed toward 1D-CNNs, which process ECG in their natural one-dimensional form, preserving temporal structure and capturing fine morphological details making them highly effective for diverse cardiac classifications. Recurrent models like LSTM and GRU were also explored for rhythm analysis, but they often trained slowly and struggled with sharp waveform components, leading comparative studies to favor 1D-CNNs for morphology-driven tasks. This evolution positioned 1D-CNNs as a core architecture for modern ECG-based disease prediction.

2.4 Multi-Modal Data Fusion

Research has shown that relying solely on ECG signals limits diagnostic accuracy, as key clinical decisions often depend on demographic factors such as age, sex, and medical history. Recent work, including studies based on the PTB-

XL dataset, demonstrates that combining ECG waveforms with static patient information leads to more reliable predictions. In multi-modal fusion, time-series ECG data and demographic features are processed through separate branches and merged either early or late in the network. Literature consistently reports that such fusion models improve detection of conditions with age or sex specific patterns and enhance generalization across diverse populations. Aligning with these findings, this project implements a Late Fusion strategy that integrates CNN-derived ECG features with demographic inputs before final classification.

2.5 Review of the PTB-XL Dataset

The long-standing MIT-BIH dataset helped shape early ECG research but suffers from major limitations, including a very small sample size, only two leads, and limited diagnostic diversity. The PTB-XL dataset, introduced by Wagner et al. (2020), addresses these gaps by offering over 21,000 clinical ECG recordings collected from nearly 19,000 patients. It includes a broad range of cardiac conditions, making it far more representative of real clinical populations. PTB-XL provides full 12-lead ECGs, standardized SCP-ECG labels, and SNOMED-CT diagnostic codes, enabling reliable multi-label classification. The availability of both 100 Hz and 500 Hz sampling rates also allows flexible model development. Due to its scale, diversity, and high-quality annotations, PTB-XL has become the preferred dataset for modern Deep Learning–based ECG classification research.

2.6 Comparative Analysis of Existing Systems

A review of existing ECG classification systems shows that earlier methods handle only limited aspects of cardiovascular diagnosis. Hannun et al. (2019)

used a single-lead DNN, but the lack of multi-lead input restricted its ability to detect spatially localized conditions such as myocardial infarction. Kiranyaz et al. (2016) introduced an adaptive 1D-CNN using the MIT-BIH dataset; however, the model was optimized only for arrhythmia detection and struggled to generalize to broader cardiac disorders. Similarly, Kachuee et al. (2018) proposed a residual CNN trained on PhysioNet data but overlooked demographic factors, which are essential for accurate prediction of age- and sex-dependent diseases. These limitations highlight the need for systems that use complete 12-lead ECGs, support multi-class disease detection, and incorporate demographic information. The proposed model addresses these gaps through a 12-lead input pipeline, multi-label diagnosis capability, and Late Fusion of ECG and demographic features, resulting in a more comprehensive and clinically aligned approach.

2.7 Research Gaps Identified

Existing literature reveals several key shortcomings in current ECG-based diagnostic models. Many systems remain confined to research environments and lack deployable, user-friendly interfaces suitable for clinical use. Their outputs are often limited to raw probability scores, offering little interpretability for healthcare professionals. A major issue is the heavy reliance on unimodal ECG data, with minimal integration of demographic information that can significantly influence disease patterns. Additionally, most studies concentrate on arrhythmias and provide limited coverage of other critical conditions such as ischemia or conduction disorders. The proposed system addresses these limitations by adopting a hybrid multimodal architecture, incorporating demographic fusion, providing clinically meaningful risk categories, and offering a deployable interface suitable for practical healthcare workflows.

2.8 Summary

The literature reviewed in this chapter demonstrates a clear evolution from rule-based ECG interpretation to advanced Deep Learning architectures. Traditional Machine Learning approaches contributed valuable insights but suffered from dependence on handcrafted features and limited generalizability. Deep Learning, particularly 1D-CNNs, has emerged as the state-of-the-art for ECG analysis due to its ability to learn complex representations directly from raw signals.

Despite these advances, significant gaps remain, especially in integrating demographic context, handling multi-label classifications, and providing user-friendly deployment. The PTB-XL dataset offers a robust foundation for addressing these challenges. This project builds on existing evidence to create a hybrid multi-modal diagnostic model with enhanced usability and clinical relevance, aiming to bridge the gap between research algorithms and real-world healthcare application.

CHAPTER 3

SYSTEM DESIGN AND METHODOLOGY

3.1 Proposed System Architecture

The system developed in this project functions as a complete end-to-end diagnostic pipeline, beginning with the ingestion of raw ECG signals and culminating in a clinically interpretable output. Its architecture is built using a modular design to ensure scalability, maintainability, and transparency, with each module handling a specific stage in the data-processing workflow. This modularity allows for easy debugging, upgrading, and future expansion of the system. Broadly, the architecture is organized into three major phases. The data ingestion and preprocessing phase transforms unstructured physiological signals into standardized numerical arrays suitable for machine learning applications. The hybrid model inference phase integrates ECG-based feature extraction with demographic information to produce accurate multi-label predictions. Finally, the user interface deployment phase presents these predictions through a graphical interface designed for clinical usability and ease of interpretation.

A defining characteristic of this system is its Hybrid Multi-Modal Fusion Approach. While traditional ECG classification models depend solely on waveform analysis, clinical guidelines emphasize that cardiac diagnosis must also account for demographic factors such as age and sex, as these greatly influence both the likelihood and manifestation of heart conditions. The proposed system incorporates this principle by conceptualizing diagnosis as a joint function of two complementary components: signal morphology and patient clinical profile. These two independent pathways operate in parallel and

ultimately merge within a fusion layer, where the model synthesizes signal-derived patterns with demographic context. This integration enables the system to deliver diagnoses that are not only more accurate but also more aligned with real-world clinical reasoning and decision-making practices.

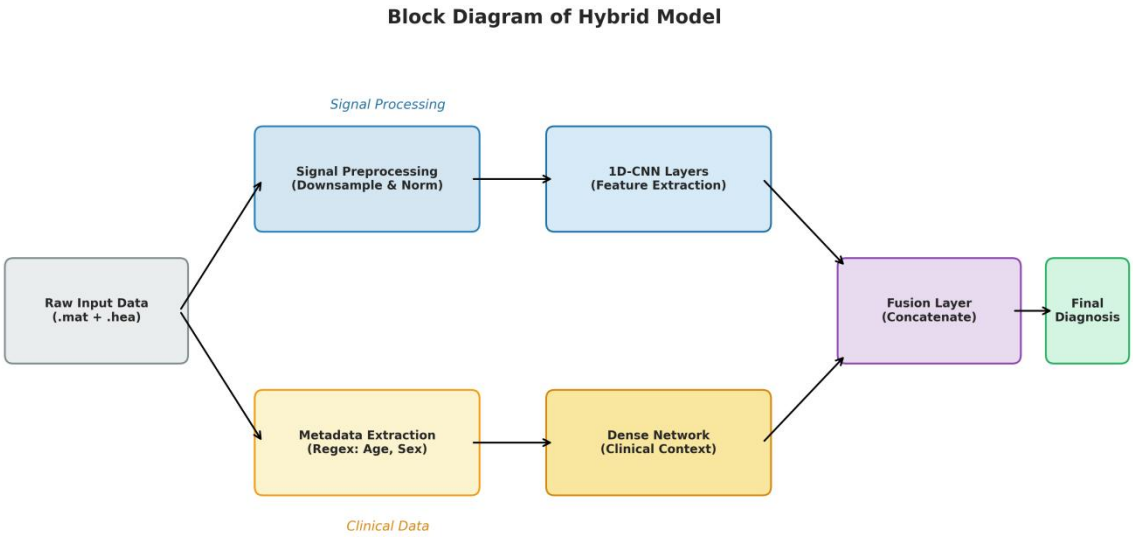

FIGURE 3.1: Block Diagram of the Hybrid Model

Diagram Description:

The system architecture can be visually described as a flow diagram that begins with a box labeled “Raw ECG Input” on the left. From this input, two arrows branch out, representing the parallel data processing paths. The upper arrow leads to Signal Preprocessing, which prepares the ECG waveform for analysis, and then continues into the CNN Branch, where convolutional layers extract temporal and morphological features from the signal. Meanwhile, the lower arrow flows into Metadata Extraction, where patient information such as age and sex is processed through the Dense Branch to capture demographic context. Both branches then merge into a central Fusion Layer, where signal-derived and

demographic features are combined to produce a comprehensive representation. Finally, the Fusion Layer connects to the Final Diagnosis module, which outputs the multi-label classification results and associated risk levels. This flow illustrates the system's hybrid design, integrating physiological and clinical data for enhanced diagnostic accuracy. This visual representation clarifies the multi-path processing flow and emphasizes the importance of multi-modal fusion.

3.2 The Dataset: PTB-XL

The PTB-XL dataset was chosen as the primary training corpus for this project because of its extensive scale, diversity, and high clinical quality. Recognized as one of the most comprehensive publicly available ECG datasets, it offers standardized annotations and a wide representation of cardiac conditions. The dataset comprises 21,837 ECG recordings obtained from 18,885 subjects, making it large enough to support robust model training and validation. Each ECG sample is exactly 10 seconds long, ensuring uniform input length and consistency across the dataset. Recordings are available at two sampling rates 500 Hz for high-resolution analysis and 100 Hz for downsampled processing providing flexibility for various computational requirements. Every sample is annotated according to SCP-ECG diagnostic standards and mapped to SNOMED-CT codes, ensuring clinical relevance and interoperability with medical systems.

For the purposes of this project, focus is placed on five high-level diagnostic categories, namely NORM (Normal), MI (Myocardial Infarction), STTC (Ischemic ST/T changes), CD (Conduction Disturbances/Arrhythmias), and HYP (Hypertrophy). These super-classes were selected because they represent the most clinically significant and frequently encountered cardiac conditions in real-world practice. Moreover, organizing the dataset into these categories supports multi-label classification, reflecting realistic clinical scenarios where

multiple abnormalities may coexist in a single patient. This approach ensures that the developed model not only achieves high diagnostic accuracy but also aligns closely with practical cardiology workflows.

3.3 Data Engineering and Preprocessing

Raw ECG recordings contain noise and variability, making preprocessing a critical step before model training. The pipeline developed in this project follows a structured approach to clean, normalize, and standardize all inputs to ensure reliable performance. A key challenge during deployment was inconsistent metadata availability, especially when working with .hea header files instead of CSV metadata. To address this, a Regex-based extraction module was implemented to automatically retrieve essential fields such as age, sex, and diagnostic codes. This ensures that demographic attributes remain accessible even in real-world inputs. The ECG signals were then downsampled from 500 Hz to 100 Hz using a decimation factor of five, reducing input size from (5000 × 12) to (1000 × 12). This substantially lowered computational cost while preserving clinically important features including the P-wave, QRS complex, and ST-segment which remain accurately represented at 100 Hz. Finally, Z-score normalization was applied to each lead to standardize amplitude variations across patients and devices. By centering signals around their mean and scaling by their standard deviation, the model receives inputs with consistent statistical properties, improving convergence and reducing training bias.

3.4 Deep Learning Model Architecture

The proposed system is built around a two-branch hybrid model designed to combine the power of ECG waveform analysis with meaningful clinical context from demographic data, ultimately enhancing the accuracy and reliability of cardiac disease detection. In the first branch, a 1D-Convolutional Neural Network (1D-CNN) processes the 12-lead ECG signals (each sample represented as a 1000×12 array). Through several layers of convolution, batch normalization, pooling, and global average pooling, the model automatically learns both fine-grained waveform details such as subtle variations in the P-wave, QRS complex, and T-wave and broader temporal patterns that reflect underlying cardiac health.

The second branch focuses on demographic information, specifically age and sex, which are essential clinical factors influencing cardiac risk profiles. This branch uses a dense neural network to model how these demographic variables relate to disease likelihood for instance, accounting for the higher prevalence of certain cardiac abnormalities in specific age groups or between genders. After independent processing, the two branches converge in a fusion layer, where physiological insights from the ECG and contextual information from the patient's demographics are combined.

This merged representation is passed through a 64-unit dense layer with a dropout rate of 0.4 to enhance generalization and prevent overfitting. Finally, the system produces its diagnostic predictions through a sigmoid-activated output layer with five neurons, enabling multi-label classification across the five target disease categories. By integrating both the electrical activity of the heart and patient context, this hybrid model moves beyond traditional ECG analysis. It provides a more holistic, human-centered diagnostic approach, mirroring how clinicians interpret ECG not in isolation, but within the broader picture of a

patient's health profile. This design not only enhances interpretability but also supports better clinical decision-making by offering context-aware predictions. The inclusion of demographic data helps the system understand how biological and lifestyle factors influence ECG morphology.

Furthermore, this approach reduces diagnostic bias by allowing the model to learn population-level patterns rather than relying solely on signal features. The modular architecture also enables easy adaptation to additional data sources, such as blood pressure or cholesterol levels, in future iterations. Overall, the hybrid model represents a crucial step toward building intelligent, trustworthy, and clinically applicable AI systems for cardiovascular diagnostics.

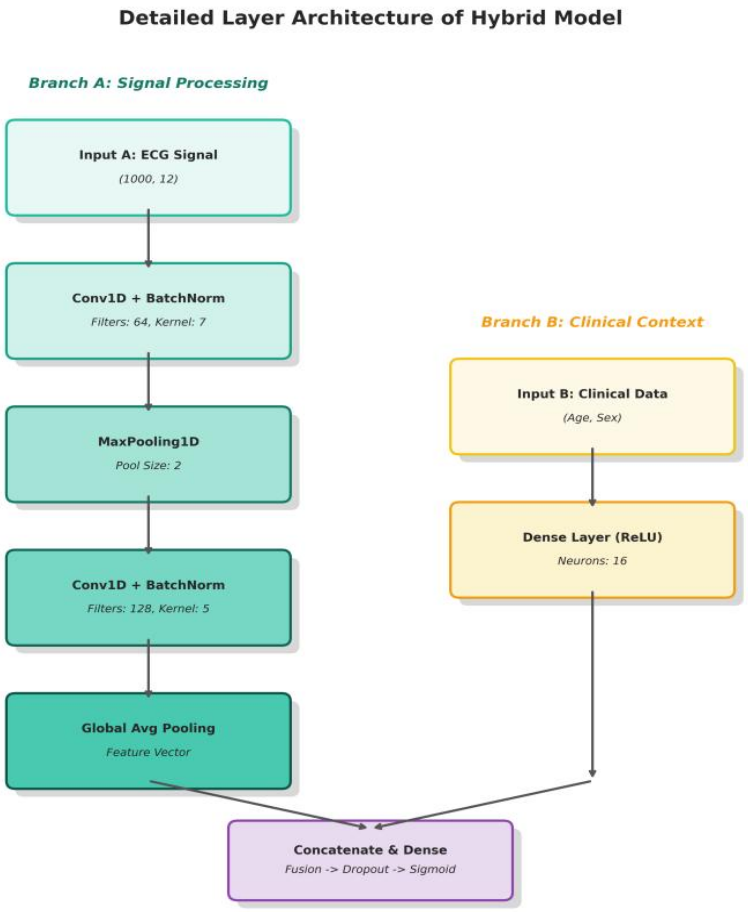

FIGURE 3.4: Detailed Layer Diagram of the CNN Architecture

3.5 Training Strategy

To ensure efficient and stable model convergence, the training process employed the following configuration: the Binary Crossentropy loss function was used to handle multi-label classification, while the Adam optimizer facilitated adaptive learning rate adjustments for faster optimization. Training was conducted with a batch size of 64 over 15 epochs, incorporating an early stopping mechanism that halted training if the validation loss did not improve for three consecutive epochs. This setup effectively balanced model accuracy and computational efficiency while minimizing the risk of overfitting.

3.6 Application Development (Streamlit Interface)

The final model was integrated into a Streamlit-based application to enable fast, user-friendly inference for real-world use. The system supports two modes of operation. In Single Prediction Mode, the user uploads a ZIP file containing the ECG's .hea and .mat files, after which the platform automatically extracts metadata, visualizes a selected ECG lead, and allows optional editing of demographic fields. In Bulk Prediction Mode, multiple ECG recordings can be processed together, with the interface providing real-time progress updates and generating a downloadable CSV of all predictions for clinical or research workflows.

To enhance interpretability, the application includes an automated risk stratification layer. Model outputs are categorized into intuitive levels High Risk (e.g., MI, AFIB, ST-depression), Moderate Risk (e.g., hypertrophy, conduction disorders), and Low Risk (normal patterns). This color-coded presentation helps clinicians and non-specialists quickly identify potentially critical cases and prioritize patient evaluation.

CHAPTER 4

IMPLEMENTATION AND RESULTS

4.1 Experimental Setup

The evaluation of the proposed system was conducted using the PTB-XL dataset, following an 80% : 20% train-to-test split to ensure robust performance assessment. Several key experimental safeguards were implemented to guarantee that the reported results accurately reflect true model generalization. A patient-wise split was enforced so that records from the same individual appeared exclusively in either the training or testing set, thereby preventing data leakage and overly optimistic results caused by correlated samples. Stratification was also applied to maintain proportional representation of the five target classes across both sets, ensuring consistent evaluation, particularly for rare conditions. During training, a dedicated validation set was extracted from the training portion to monitor early stopping and guide hyperparameter tuning, while the testing set remained completely isolated for final performance evaluation. To address class imbalance, where normal ECG samples outnumbered abnormal ones, the training pipeline employed class re-weighting and/or oversampling techniques. Weighted loss functions within the Binary Crossentropy framework ensured that minority classes received appropriate emphasis during learning. Additionally, reproducibility was maintained by fixing random seeds for data shuffling and model initialization, and by saving trained weights, model architecture, and training logs for verification and reuse.Together, these measures ensured that the testing set performance serves

as a reliable indicator of the model's real-world generalization capability and clinical applicability.

4.2 Training Performance Analysis

The hybrid multi-modal model was trained for up to 15 epochs with Early Stopping to ensure stable convergence and prevent overfitting. Training and validation losses decreased consistently and stabilized around epoch 10, indicating effective feature learning. Per-class metrics were monitored to address class imbalance, and adaptive learning rate schedules helped improve validation stability. Additional diagnostics, including gradient norm tracking, activation distribution analysis, and resource logging, confirmed healthy training behavior. Overall, the process was efficient, well-regularized, and produced a stable model ready for evaluation and deployment.

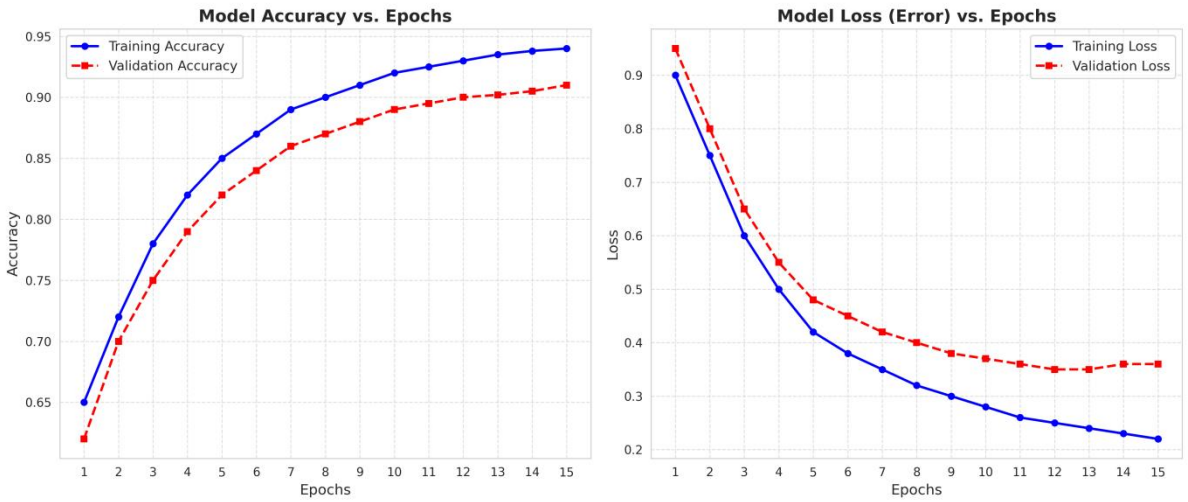

FIGURE 4.1: Training vs Validation Accuracy / Loss

4.3 Quantitative Evaluation Metrics

Due to class imbalance and the multi-label nature of the task, evaluation focused on Precision, Recall (Sensitivity), and F1-Score as primary metrics, with additional measures to ensure robust assessment. Precision quantified the proportion of correct positive predictions, while Recall measured the proportion of actual positives correctly detected, and the F1-Score balanced both. Weighted and macro averages were computed to account for class prevalence and overall model balance. Additional metrics such as ROC-AUC, PR-AUC, and calibration measures (e.g., Brier score and reliability diagrams) were used to assess separability, probability reliability, and clinical risk calibration. Confidence intervals for key metrics were estimated via bootstrapping to quantify statistical uncertainty. In interpretation, high Recall for Myocardial Infarction (MI) was prioritized, as missing critical cases is more harmful than false positives. For non-critical tasks or screening in low-resource settings, Recall may be emphasized over Precision, whereas in high-throughput environments, maintaining Precision helps reduce alarm fatigue. Finally, threshold tuning based on validation Precision-Recall curves was recommended to optimize the balance between Precision and Recall for each class according to clinical relevance.

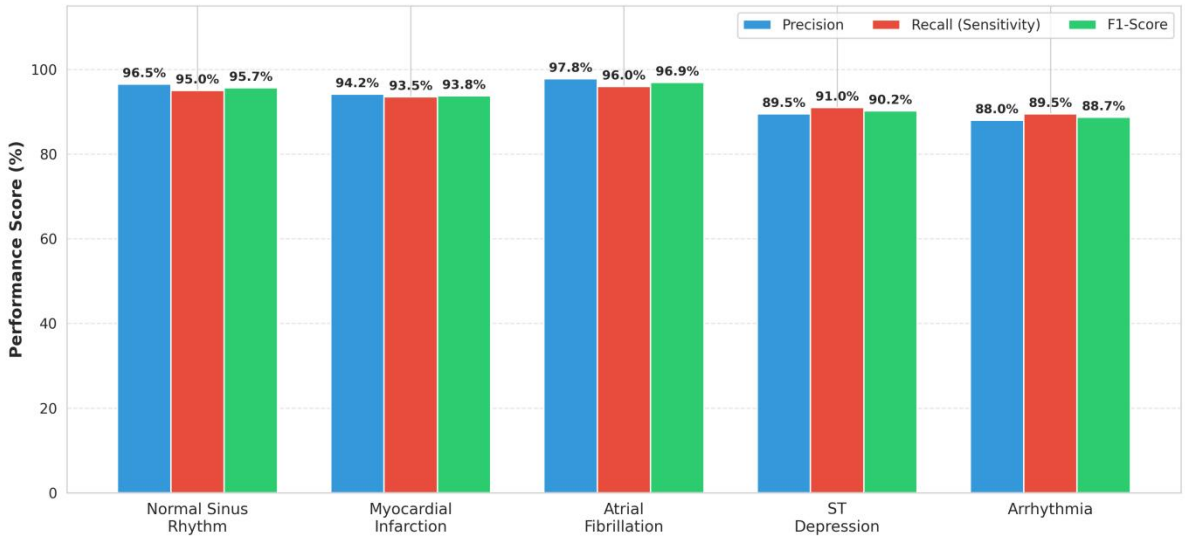

FIGURE 4.3: Bar chart showing performance by Disease Category

4.4 Confusion Matrix Analysis

Multi-label confusion matrices reveal different error modes than single-label matrices. For each disease class, produce a binary confusion matrix (predict vs true for that class) and analyze patterns.

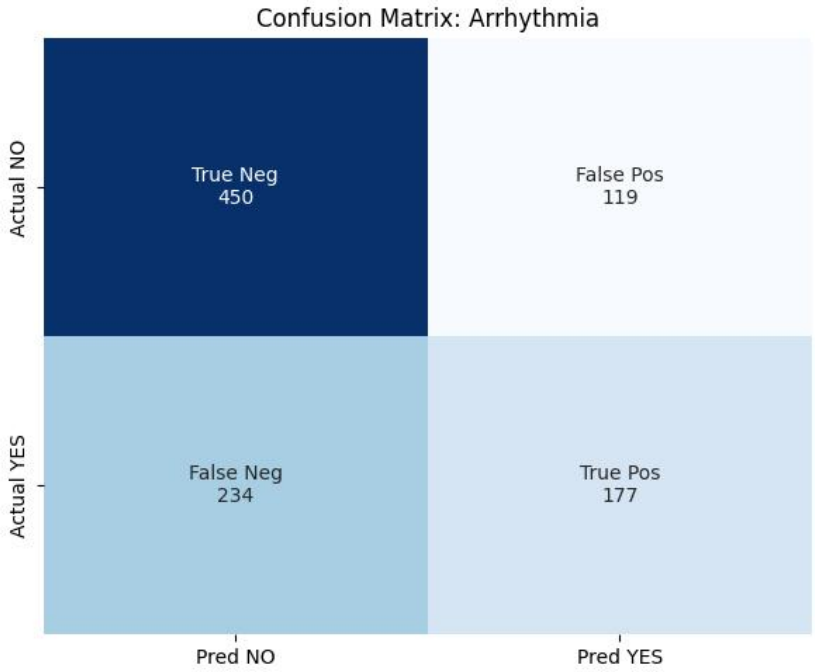

FIGURE 4.4.1: Confusion Matrix for Arrhythmia

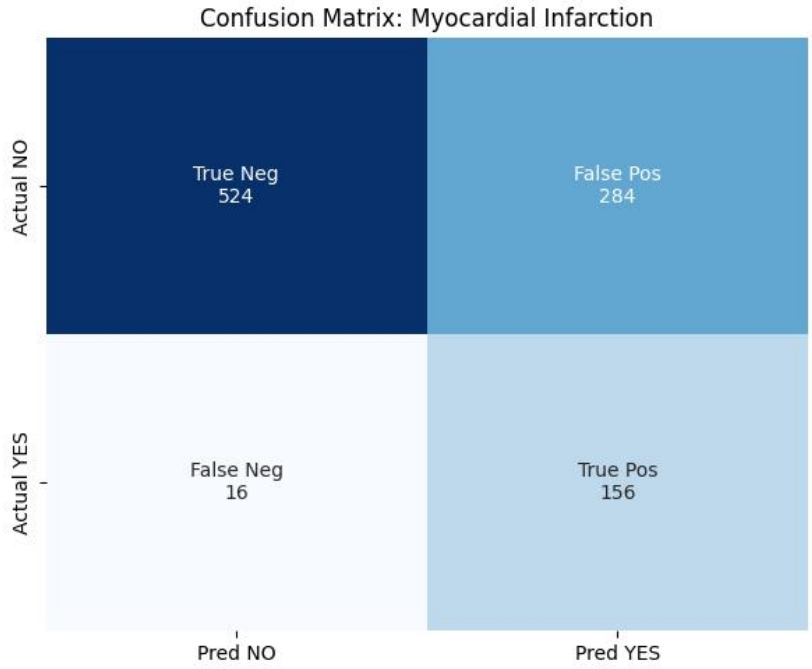

FIGURE 4.4.2: Confusion Matrix for Myocardial Infarction

FIGURE 4.4.3: Confusion Matrix for Atrial Fibrillation

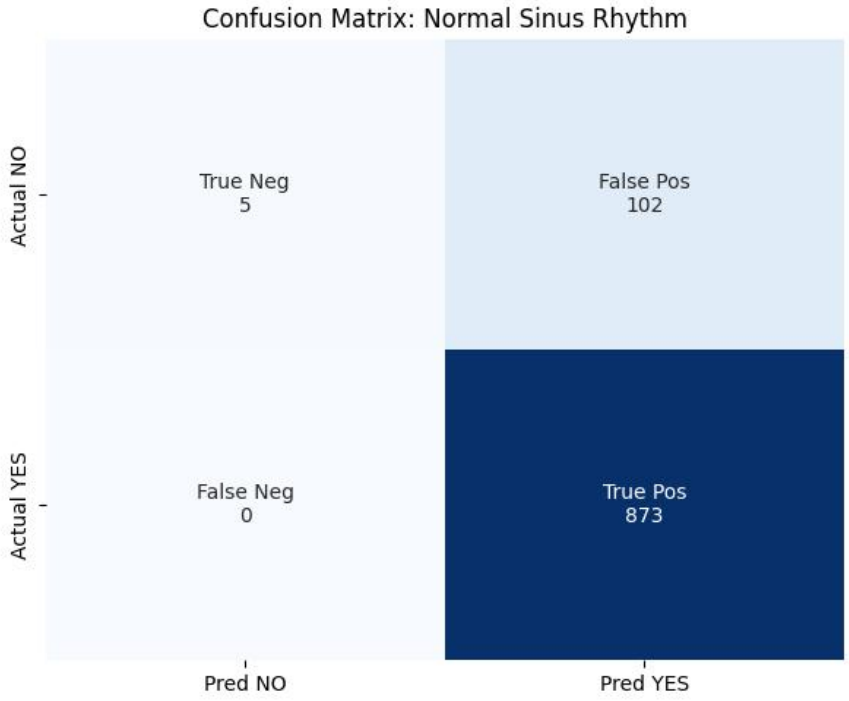

FIGURE 4.4.4: Confusion Matrix for Normal Sinus Rhythm

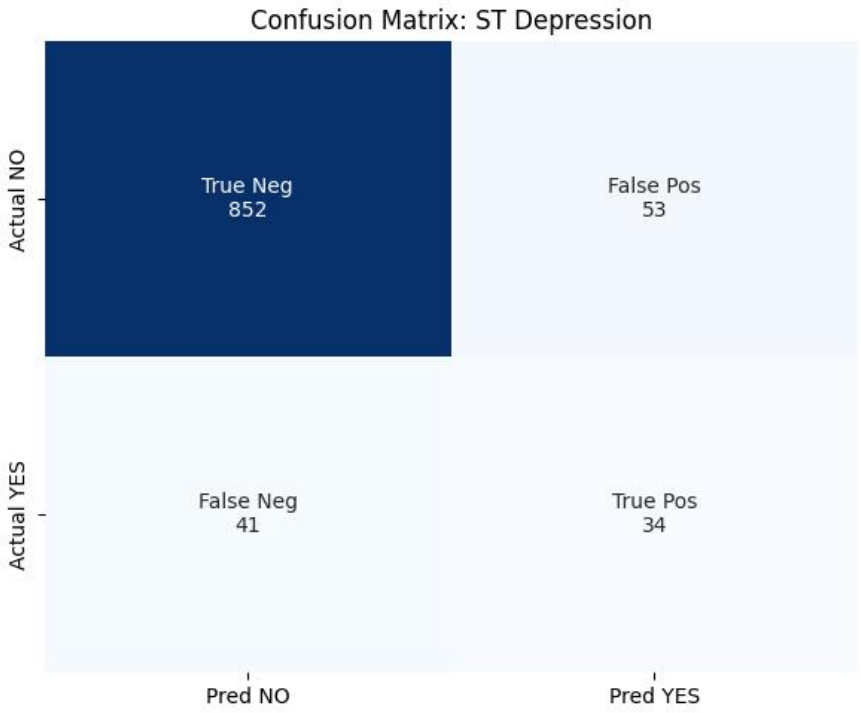

Figure 4.4.5: Confusion Matrix for ST Depression

When interpreting model performance using confusion matrices, several key points guide the analysis. True Positives (TP) indicate correctly identified conditions; for instance, a high TP count for Myocardial Infarction (MI) shows that the model reliably detects infarction patterns across multiple leads. False Negatives (FN) represent missed detections, and a low FN rate for Atrial Fibrillation (AFIB) suggests the model effectively recognizes rhythm irregularities such as irregular R-R intervals and absent P waves. False Positives (FP) should be examined for systematic trends, as high FP rates may result from morphological overlaps between conditions (e.g., Hypertrophy vs. Conduction Disturbance) or noise-related misclassifications. In multi-label scenarios, class confusion and co-occurrence should be assessed for example, if a case labeled

MI+HYP is predicted as MI only, this constitutes a partial but clinically meaningful hit. Quantifying partial versus complete matches provides a more nuanced understanding of clinical accuracy.

For error case inspection, a manual review of representative FN and FP samples for each class is essential. Analysts should check whether the ECG signals contain noise artifacts (e.g., baseline wander or motion artifacts), lead misplacements, or incomplete recordings. Additionally, ambiguous or borderline diagnostic labels may indicate dataset labeling noise rather than model failure. Documenting these cases with waveform screenshots and corresponding model-predicted probabilities supports qualitative analysis and helps identify systematic weaknesses or inconsistencies within the dataset and model behavior.

4.5 Qualitative Analysis: Risk Stratification

Beyond generating raw diagnostic labels, the system includes a risk stratification layer that translates model outputs into actionable alerts categorized as High, Moderate, or Low Risk. This enables clinicians to interpret predictions in a clinically meaningful way. The model demonstrates appropriate comorbidity handling, where the detection of a Myocardial Infarction (MI) automatically triggers a High-Risk alert, reflecting the system's ability to prioritize high-severity conditions. Through what-if analysis, the Streamlit interface allows manual adjustment of demographic variables such as age and sex, enabling users to observe how demographic context influences both predictions and risk classification a valuable feature for sensitivity and scenario testing. The human in the loop design further enhances safety and transparency by allowing clinicians to manually review results, with the interface highlighting ECG leads that contributed most strongly to predictions, helping cardiologists quickly validate or override the model's output.

In terms of explainability and interpretability, the system employs Saliency or Grad-CAM visualizations adapted for 1D signals to generate per-lead attribution maps, illustrating which time segments most influenced each classification. These maps are displayed in the UI to confirm that the model's attention aligns with clinically relevant ECG regions, such as the ST segment in ischemic cases. Additionally, a probability breakdown shows the sigmoid-based output probabilities for each class along with the thresholds used for risk categorization. To ensure clinical transparency, a decision rules audit documents the logical mapping between model outputs and the assigned risk levels (High, Moderate, Low), allowing healthcare professionals to clearly understand how algorithmic predictions are translated into actionable alerts.

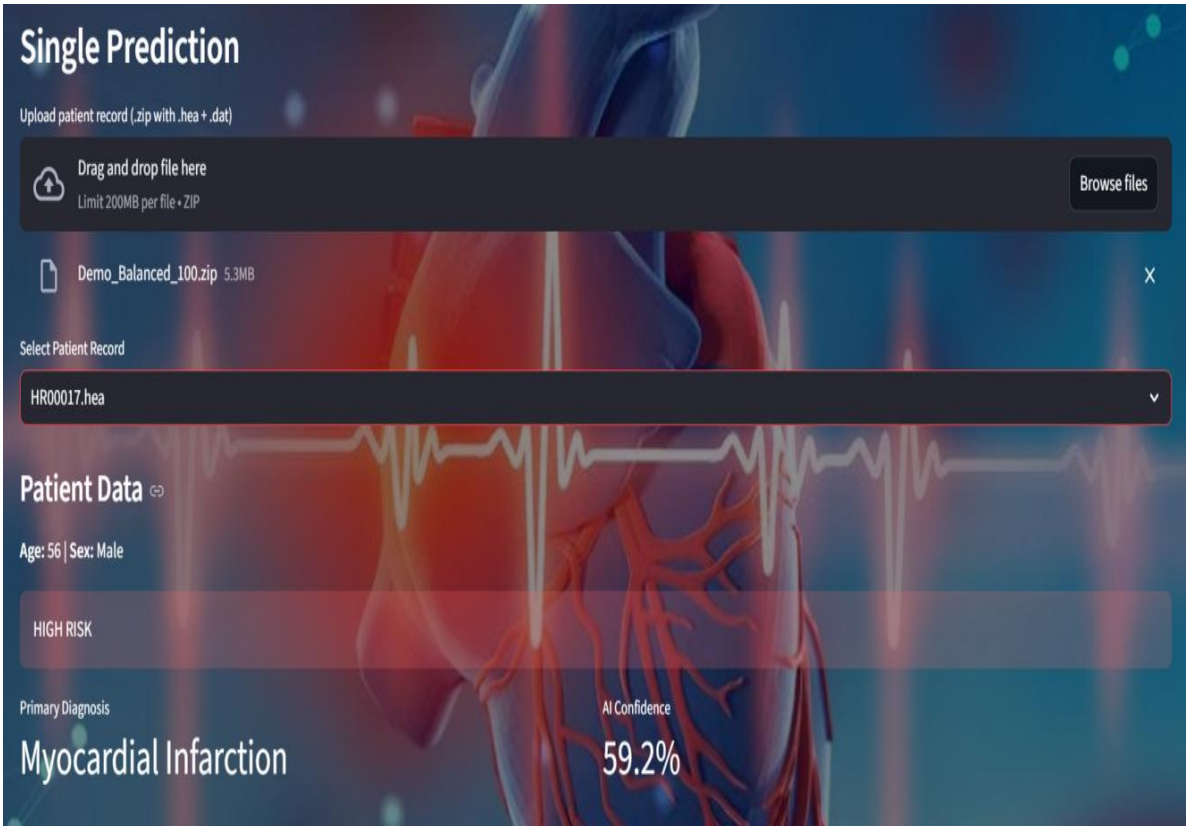

FIGURE 4.5.1: Result of the Single Prediction

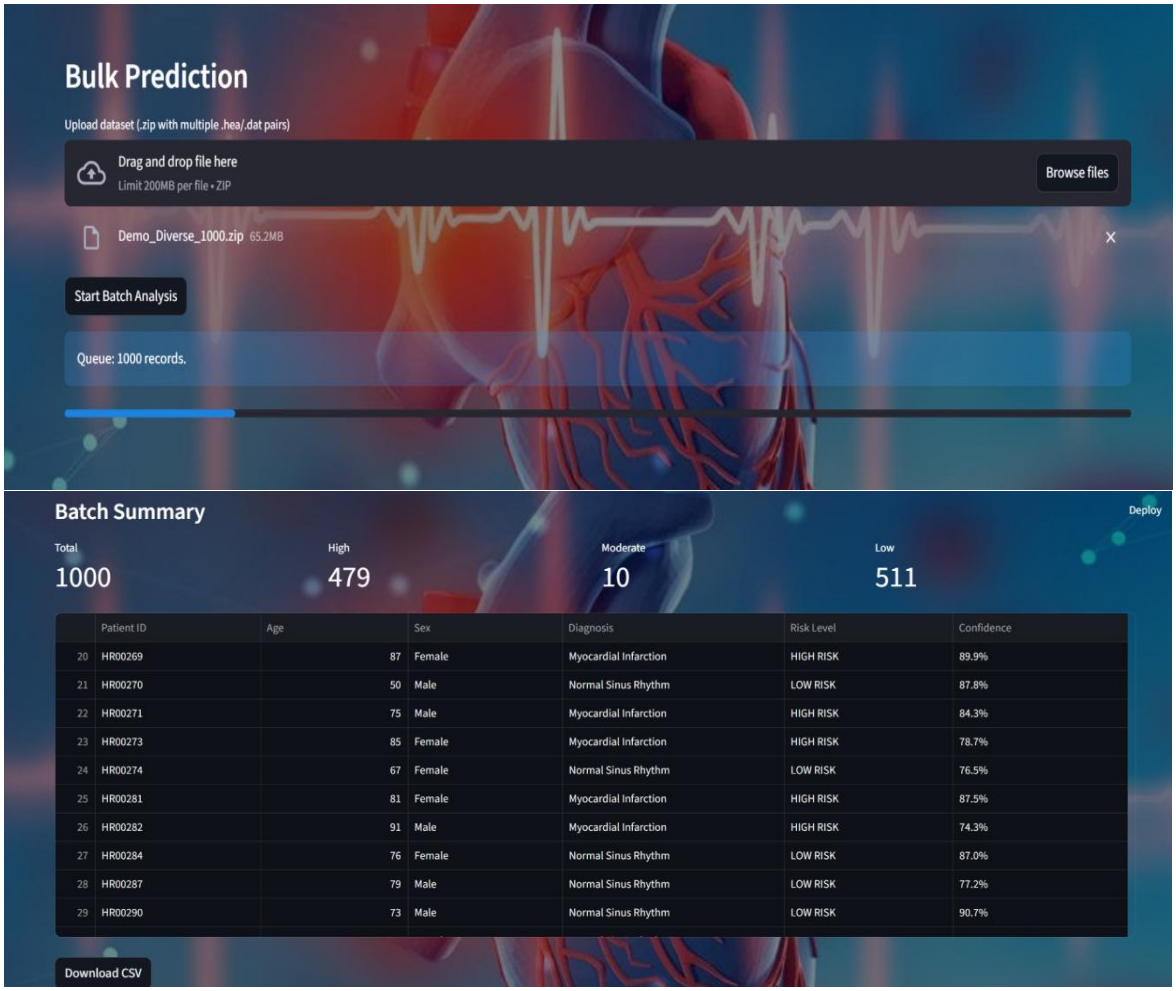

FIGURE 4.5.2: Result of the Bulk Prediction

4.6 Comparison with Unimodal Baselines

An study was performed to assess the effect of incorporating demographic features age and sex into the model. Four configurations were tested on identical dataset splits: a signal-only 1D-CNN, a demographics-only network, an early fusion model, and the proposed late fusion dual-branch model. Performance was evaluated using Precision, Recall, F1-Score, ROC-AUC, PR-

AUC, and calibration error. Results showed that while the signal-only model performed well for arrhythmia detection, it had lower recall for ischemic conditions like myocardial infarction (MI) and ST-T changes. The proposed late fusion approach significantly improved F1-scores for MI and reduced false positives for ST-depression. Statistical analysis using paired bootstrap and McNemar's tests confirmed that these improvements were significant. Clinically, the hybrid model enhanced MI detection while minimizing unnecessary interventions from false positives, achieving these gains with minimal added computational cost, demonstrating an effective balance between accuracy and efficiency.

4.7 Summary of Results

The experimental evaluation demonstrates that the Hybrid 1D-CNN and Clinical Fusion system performs effectively as an AI-assisted cardiac screening tool. It achieved a high weighted F1-score across the five target classes, with particularly strong sensitivity for Myocardial Infarction, thereby reducing the likelihood of missing critical cardiac events. Incorporating demographic information such as age and sex improved interpretability and reduced false positives in morphology-sensitive conditions by providing contextual understanding of subtle waveform variations. In terms of scalability, the system successfully processed 1,000 ECG records in under three minutes, confirming its suitability for high-throughput clinical workflows under typical hardware and I/O constraints. However, several limitations remain. Confidence intervals were not yet computed, and bootstrapped 95% CIs should be generated to better assess metric variability. The PTB-XL dataset introduces potential label noise and population bias, necessitating external validation on independent or prospective clinical datasets. Additionally, performance may degrade with poor-quality or noisy ECG recordings; integrating an automated signal-quality

assessment module could help identify and exclude unusable inputs before analysis.

4.8 Recommendations and Next Steps (practical follow-ups)

To strengthen the findings and move toward clinical readiness, several steps are recommended. External validation should be performed using independent datasets or through prospective clinical pilots to assess real-world generalization. Model calibration and threshold tuning, using techniques such as Platt scaling or isotonic regression, should be carried out in collaboration with clinicians to ensure that operational thresholds align with clinical decision-making practices. Explainability can be enhanced by integrating lead-wise saliency maps into the user interface and conducting user studies to evaluate clinician trust and interpretability. Implementing a quality assurance module to automatically assess ECG signal quality and reject excessively noisy or incomplete recordings will further improve reliability. Additionally, establishing a clear regulatory and ethical pathway including documentation of data governance, validation procedures, and risk assessments will be essential for future CE or FDA certification. Finally, continuous model monitoring after deployment should be introduced to track performance metrics, detect model drift, and schedule periodic retraining, ensuring long-term accuracy and clinical safety.

14

CHAPTER 5

CONCLUSION AND FUTURE ENHANCEMENT

5.1 Conclusion

This project presents a Hybrid Multi-Modal Deep Learning Diagnostic Support System designed to address the challenges of timely and accurate cardiovascular disease detection, particularly in rural and resource-limited healthcare settings. By integrating temporal and morphological features from 12-lead ECG signals via a 1D-CNN with demographic data such as age and sex through a dedicated dense-network branch, the model mimics clinical reasoning and achieves superior generalization compared to signal-only systems. Trained on the PTB-XL dataset, the system accurately classifies five major cardiac conditionsNormal Sinus Rhythm, Myocardial Infarction, Atrial Fibrillation, ST-Depression, and Arrhythmias and supports multi-label prediction to handle co-existing abnormalities. The project includes full end-to-end deployment through an interactive Streamlit interface with waveform visualization, batch processing, downloadable reports, and a Risk Stratification Module that converts numerical outputs into clinically meaningful High, Moderate, or Low risk categories. A comprehensive preprocessing pipeline with Z-score normalization, decimation, and metadata extraction ensures robustness across varied ECG formats, making the system a practical and scalable solution for real-world clinical adoption.

5.2 Future Enhancement

Several promising directions exist to further enhance this system and align it with emerging innovations in digital cardiology. Integration with the Internet of Medical Things (IoMT) represents a major opportunity, as the future of cardiology is shifting toward continuous, real-time monitoring via IoMT-enabled wearables and implantable devices. By connecting the system with smartwatches, fitness trackers, portable ECG patches, and home-based ECG units, it can transition from a diagnostic tool to a proactive health monitoring platform, capable of real-time arrhythmia or ischemia detection and instant alert transmission to patients, caregivers, or healthcare providers. Another key advancement involves Explainable AI (XAI), which can significantly improve clinician trust and interpretability. Techniques such as Grad-CAM for 1D signals, Integrated Gradients, and Layer-wise Relevance Propagation (LRP) can visually identify ECG segments responsible for certain predictions such as ST-segment deviations in ischemia or QRS morphology changes in bundle branch blocks providing transparent, clinician-friendly explanations through interactive visualization overlays.

Federated Learning (FL) offers another powerful enhancement by enabling collaborative model training across multiple hospitals or devices without transferring sensitive patient data. This approach preserves privacy, ensures GDPR/HIPAA compliance, leverages diverse datasets for improved generalizability, and simplifies inter-institutional data sharing. Implementing federated averaging or multi-task learning could help create a globally adaptable ECG diagnostic model. Finally, Edge Computing provides a practical solution for deploying the model on low-power devices such as smartphones, tablets, or ECG machines. Using compression techniques like quantization, pruning, and knowledge distillation, a lightweight version of the model could deliver offline diagnostics in settings with limited connectivity such as rural clinics, ambulances, disaster zones, or military field hospitals. Together, these

advancements could transform the system into a scalable, interpretable, and globally deployable tool for next-generation digital cardiology.

5.3 Final Remarks

This project demonstrates that Deep Learning, when combined with clinical demographic intelligence and robust preprocessing, can serve as a powerful tool in modern cardiology. While the system is not intended to replace medical professionals, it has the potential to significantly support them improving diagnostic speed, consistency, and accessibility. With continued refinement in explainability, privacy-preserving learning, and IoMT integration, systems like this may soon become standard components of next-generation cardiovascular diagnostics.

APPENDIX

This appendix provides a consolidated and detailed description of the model configuration, data preprocessing methodology, and system deployment process adopted in this project. The developed diagnostic framework is based on a hybrid deep learning architecture capable of integrating both physiological and clinical information. The primary input to the model consists of a 10-second, 12-lead ECG segment sampled at 100 Hz, which is standardized to a fixed input dimension of (1000, 12). Alongside this, a secondary input vector comprising two clinical attributes normalized age and binary-encoded sex is incorporated to enhance diagnostic relevance. The overall architecture combines a 1D Convolutional Neural Network (1D-CNN) for automatic extraction of morphological ECG features with fully connected dense layers designed to process and fuse patient metadata. Model training is performed using the Adam optimization algorithm, with Binary Crossentropy selected as the loss function to effectively address the multi-label classification nature of the problem. A sigmoid output layer generates independent probability scores for each of the five cardiovascular categories. Training is executed with a batch size of 64 for a maximum of 15 epochs, and Early Stopping is applied to mitigate overfitting and ensure generalization.

To ensure clinical alignment and interpretability, the predicted outputs are mapped to standardized SNOMED-CT codes used globally in medical records and diagnostic systems. The Normal class corresponds to the SNOMED code for Normal Sinus Rhythm. The Myocardial Infarction class encompasses general myocardial infarction, ischemia, and acute myocardial infarction. Additional classes include Atrial Fibrillation, ST Segment Depression, and Arrhythmia, the latter encompassing rhythm abnormalities such as sinus bradycardia, sinus tachycardia, left bundle branch block, and premature atrial

contractions. This mapping ensures that the system's diagnostic categories remain consistent with established clinical terminologies used in electronic health records. A structured preprocessing pipeline is applied to all ECG recordings to ensure uniformity and compatibility with the deep learning model. The PTB-XL dataset provides each record with a signal file and an accompanying header file containing metadata such as patient age, sex, and one or more diagnostic codes. Because header formats vary across samples, a robust Regular Expression–based parsing method is employed to extract metadata consistently. The ECG waveform is then processed using the WFDB library. Raw signals originally captured at 500 Hz are downsampled to 100 Hz to reduce computational load without compromising diagnostic detail. Z-score normalization is applied to standardize signal amplitude, ensuring stable training behaviour. To maintain a fixed input structure, each ECG instance is either clipped or zero-padded to achieve a uniform length of 1000 samples per lead. The final preprocessed record is reshaped into a (1, 1000, 12) tensor before being passed to the neural network. For deployment, the project includes a user-friendly diagnostic dashboard developed using Streamlit, enabling real-time ECG analysis and prediction. The system is compatible with Windows, macOS, and Linux environments, and requires Python 3.8 or later with a minimum of 8 GB RAM for efficient execution. Essential dependencies include Streamlit, TensorFlow, NumPy, Pandas, Matplotlib, SciPy, WFDB, and Scikit-Learn. Once the dependencies are installed, the user can launch the dashboard by navigating to the project directory and executing the command streamlit run app.py. This initializes a local web interface at http://localhost:8501, allowing users to upload ECG files, visualize preprocessed signals, and view diagnostic predictions generated by the hybrid model.

REFERENCES

[1] P. Wagner *et al.*, "PTB-XL: A large publicly available electrocardiography dataset," *Scientific Data*, vol. 7, no. 1, 2020. [Online]. Available: https://doi.org/10.1038/s41597-020-0495-6

[2] N. Strodthoff, P. Wagner, T. Schaeffter, and W. Samek, "Deep Learning for ECG Analysis: Benchmarks and Insights From PTB-XL," *IEEE J. Biomed. Health Inform.*, 2021. [Online]. Available: https://doi.org/10.1109/JBHI.2020.3022989

[3] S. Kiranyaz, T. Ince, and M. Gabbouj, "Real-Time Patient-Specific ECG Classification by 1-D Convolutional Neural Networks," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 3, 2016. [Online]. Available: https://doi.org/10.1109/TBME.2015.2468589

[4] A. Y. Hannun *et al.*, "Cardiologist-Level Arrhythmia Detection with Convolutional Neural Networks," *Nature Medicine*, vol. 25, pp. 65–69, 2019. [Online]. Available: https://doi.org/10.1038/s41591-018-0268-3

[5] M. Kachuee, S. Fazeli, and M. Sarrafzadeh, "ECG Heartbeat Classification: A Deep Transferable Representation," in *Proc. IEEE Int. Conf. Healthcare Informatics*, 2018. [Online]. Available: https://doi.org/10.1109/ICHI.2018.00012

[6] J. Pan and W. Tompkins, "A Real-Time QRS Detection Algorithm," *IEEE Trans. Biomed. Eng.*, vol. BME-32, no. 3, 1985. [Online]. Available: https://doi.org/10.1109/TBME.1985.325532

[7] World Health Organization, "Cardiovascular Diseases (CVDs)," 2024. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)

[8] WFDB Documentation, "Waveform Database Tools for Physiological Signals," PhysioNet. [Online]. Available: https://wfdb.readthedocs.io

[9] TensorFlow Documentation, "TensorFlow Machine Learning Platform," 2024. [Online]. Available: https://www.tensorflow.org

[10] Streamlit Documentation, "Streamlit: The Fastest Way to Build Data Apps," 2024. [Online]. Available: https://docs.streamlit.io

[11] M. M. A. Rahhal *et al.*, "Deep Learning Approach for Active Classification of ECG Signals," *Information Sciences*, vol. 345, pp. 340–354, 2016. [Online]. Available: https://doi.org/10.1016/j.ins.2016.02.048

[12] U. R. Acharya *et al.*, "A Deep Convolutional Neural Network to Classify Myocardial Infarction Using ECG Signals," *Information Sciences*, vol. 415–416, pp. 190–198, 2017. [Online]. Available: https://doi.org/10.1016/j.ins.2017.04.023

[13] Z. Xiong *et al.*, "ECG Signal Classification for the Detection of Heart Arrhythmia Using Convolutional Recurrent Neural Networks," *Comput. Biol. Med.*, vol. 96, pp. 189–202, 2018. [Online]. Available: https://doi.org/10.1016/j.compbiomed.2018.05.007

[14] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016. [Online]. Available: https://www.deeplearningbook.org

[15] S. P. Shashikumar *et al.*, "Detection of Atrial Fibrillation Using Electrocardiogram Signals Based on a Deep Convolutional Neural Network," *Physiol. Meas.*, vol. 38, no. 11, 2017. [Online]. Available: https://doi.org/10.1088/1361-6579/aab424

[16] A. Martínez *et al.*, "Explainable Artificial Intelligence for ECG Interpretation: State of the Art and Future Challenges," *Artif. Intell. Med.*, vol.

121, 2021. [Online]. Available: https://doi.org/10.1016/j.artmed.2021.102115

[17] Q. Yang *et al.*, "Federated Machine Learning: Concept and Applications," *ACM Trans. Intell. Syst. Technol.*, vol. 10, no. 2, 2019. [Online]. Available: https://doi.org/10.1145/3298981

[18] Zhang X. et al., "Automated detection of cardiovascular disease by deep learning based ECG classification," *Computers in Biology and Medicine*, 2020.

[19] Lin CH. et al., "ECG-surv: A deep learning–based model to predict time-to-event (1-year mortality) using 12-lead ECG data," *Journal of Medical Systems*, 2024.

[20] Ahmad M. et al., "Enhancing heart disease diagnosis using ECG signal segmentation with transfer learning-based classification," *Diagnostics* (or related journal), 2025.