3-26-2021

Development of a Real-Time Single-Lead Single-Beat Frequency-Independent Myocardial Infarction Detector

Harold Martin
Florida International University, hmart027@fiu.edu

FLORIDA INTERNATIONAL UNIVERSITY

Miami, Florida

DEVELOPMENT OF A REAL-TIME SINGLE-LEAD SINGLE-BEAT FREQUENCY-INDEPENDENT MYOCARDIAL INFARCTION DETECTOR

A dissertation submitted in partial fulfillment of the

requirements for the degree of

DOCTOR OF PHILOSOPHY

in

ELECTRICAL AND COMPUTER ENGINEERING

by

Harold Martin

2021

To: Dean John L. Volakis
College of Engineering and Computing

This dissertation, written by Harold Martin, and entitled Development of a Real-Time Single-Lead Single-Beat Frequency-Independent Myocardial Infarction Detector, having been approved in respect to style and intellectual content, is referred to you for judgment.

We have read this dissertation and recommend that it be approved.

Jean Andrian

Armando Barreto

Mercedes Cabrerizo

Naphtali Rishe

Sharan Ramaswamy

Malek Adjouadi, Major Professor

Date of Defense: March 26, 2021

The dissertation of Harold Martin is approved.

Dean John L. Volakis
College of Engineering and Computing

Andrés G. Gil
Vice President for Research and Economic Development
and Dean of the University Graduate School

Florida International University, 2021

DEDICATION

I dedicate this dissertation to my loving family. Their love, support, and patience, has made it all possible.

ACKNOWLEDGMENTS

I would like to express my most sincere gratitude to my major advisor, Dr. Malek Adjouadi for his encouragement, patience, and guidance in my research and for his kindness and consideration over the last six years. He has been an incredible advisor, mentor, and professor who has taught me a great deal of things during my time with the Center for Advanced Technology and Education-CATE and more broadly, FIU.

Special thanks to all individuals who provided advice and guidance through my PhD studies. I am grateful to all my committee members Dr. Jean Andrian, Dr. Armando Barreto, Dr. Naphtali Rishe, Dr. Mercedes Cabrerizo, and Dr Sharan Ramaswamy for their professional support and assistance.

This research is supported by the National Science Foundation under grants: CNS-1920182, CNS-1532061, CNS-1338922, CNS-2018611, and CNS-1551221, and by the National Institutes of Health through NIA/NIH grants 1R01AG055638-01A1, 5R01AG061106-02, 5R01AG047649-05, and the 1P30AG066506-01 with the 1Florida Alzheimers Disease Research Center (ADRC). Our thanks also go to the Ware Foundation.

ABSTRACT OF THE DISSERTATION

DEVELOPMENT OF A REAL-TIME SINGLE-LEAD SINGLE-BEAT FREQUENCY-INDEPENDENT MYOCARDIAL INFARCTION DETECTOR

by

Harold Martin

Florida International University, 2021

Miami, Florida

Professor Malek Adjouadi, Major Professor

The central aim of this research is the development and deployment of a novel multi-layer machine learning design with unique application for the diagnosis of myocardial infarctions (MIs) from individual heartbeats of single-lead electrocardiograms (EKGs) irrespective of their sampling frequencies over a given range. To the best of our knowledge, this design is the first to attempt inter-patient myocardial infarction detection from individual heartbeats of single-lead (lead II) electrocardiograms that achieves high accuracy and near real-time diagnosis. The processing time of 300 milliseconds to a diagnosis is just at the time range in between extremely fast heartbeats of around 300 milliseconds, or 200 beats per minute. The design achieves stable performance metrics over the frequency range of 202Hz to 2.8kHz with an accuracy of 77.12%, positive predictive value (PPV) of 75.85%, and a negative predictive value (NPV) of 83.02% over the entire PTB database; 85.07%, 81.54%, 87.31% over the PTB-XL (the largest EKG database available for research) validation set, and 84.17%, 78.37%, 87.55% over the PTB-XL test set.

Major design contributions and findings of this work reveal (1) a method for the real-time detection of ventricular depolarization events in the PQRST complex from 12-lead electrocardiograms using Independent Component Analysis (ICA), with a slightly different use of ICA proposed for electrocardiogram analysis and R-peak detection/localization; (2) a multilayer Long-Short Term Memory (LSTM) neural network design that identifies

infarcted patients from a single heartbeat of a single-lead (lead II) electrocardiogram; (3) and integrated LSTM neural network with an algorithm that detects the R-peaks in real time for instantaneous detection of myocardial infarctions and for effective monitoring of patients under cardiac stress and/or at risk of myocardial infarction; (4) a fully integrated 12-lead real-time classifier with even higher detection metrics and a deeper neural architecture, which could serve as a near real-time monitoring tool that could gauge disease progression and evaluate benefits gained from early intervention and treatment planning; (5) a real-time frequency-independent design based on a single-lead single-beat MI detector, which is of pivotal importance to deployment as there is no standard sampling frequency for EKGs, making them span a wider frequency spectrum.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

According to the Center for Disease Control and Prevention (CDC), every 40 seconds someone in the USA suffers a heart attack. There are roughly 790 thousand myocardial infarctions per year in the United States (although some sources list this number to be even higher) and about 20% of them occur without warning or symptoms [1]. Altogether, 49 percent of Americans have relevant risk factors of heart diseases that could lead to myocardial infarctions at an average medical cost of $11,664 [2, 3]. Therefore, as they remain a condition with significant possible complications [4], their accurate and early diagnosis is of utmost importance and urgent necessity.

Thankfully, physicians have a powerful tool in electrocardiograms (ECGs/EKGs) to diagnose and track the progress of MIs and other heart defects and abnormalities. These non-invasively recorded signals can help researchers and clinicians assess the health and fitness of the heart through the inherent characteristics of its electrical activity [5]. Therefore, numerous papers have used the information present therein to attempt and diagnose myocardial infarctions and other cardiac ailments [6–31]. As could be expected, the source of the information being processed (i.e., type and number of EKG leads considered), the classification method used , and the performance measures obtained vary from one study to another.

Throughout the literature, various types of signal processing algorithms and artificial neural networks have been used to identify myocardial infarctions [6, 13, 16, 17, 20–23]. The authors of [6] use deep convolutional neural networks to detect myocardial infarctions using single lead EKGs (lead II) producing an intra-patient classifier. Padmavathi Kora in [7] uses a Hybrid Firefly algorithm to perform feature extraction to be fed into

either a support vector machine, a K-nearest neighbor classifier, or an artificial neural network. They also report an intra-patient classifier that utilizes single-lead EKG. Sharma and Sunkaria [8] introduce a 3-lead EKG inferior myocardial infarction classifier for both intra- and inter-patient classification. They use stationary wavelet transforms to produce features later employed by k-nearest neighbor or support vector machine classifiers. Liu et al. [9] achieve myocardial infarction classification by using a multiple-feature-branch convolutional neural network on 12-lead electrocardiograms. They present results for intra-patient classification as well as a modified version of inter-patient classification where they use the first 32 beats of every patient used in the testing set for the training phase. Support vector machines are once more used in [10] on 12-lead EKGs to perform intra-patient classification. Support vector machines are also applied in [11] along with multiple instance learning as a type of supervised learning. In such study, the authors attempted to detect myocardial infarction from EKG-level topic vectors of a 12-lead EKGs and perhaps develop an intra-patient classifier (the most likely data split method used appears to be file split). The work in [12] introduces a bagging tree classifier that is used to detect myocardial infarctions and arrhythmia. In their study the authors explored which of the EKGs leads would be most appropriate to identify the targeted diseases and arrived at the conclusion that V4 was the most appropriate. However, although good testing metrics are reported, the type of classifier achieved is not clearly defined, as the authors did not mention whether they attempted to produce an intra- or inter-patient classifier, nor did they indicate how the data was split into training and testing sets.

Real-time classification of myocardial infarctions is attempted in [13] and [14]. Liu et al. [13] use a multi-lead convolutional neural network to perform the inter-patient classification from a four-lead electrocardiogram. However, they require the whole patient's EKGs to achieve a diagnosis. Meanwhile, Sopic et al. [14] introduce an event-driven clas-

sification technique to be used in wearable systems. They feed a random forest classifier with key features extracted from a single lead (Lead 5) electrocardiogram to generate an inter-patient diagnosis.

More recently, [15] implemented an algorithm to detect and localize myocardial infarctions from a single beat of 12-lead electrocardiograms. They used Dual-Q TQWT along with wavelet packet tensor decomposition and a decision tree to achieve intra-patient classification. In [16], interestingly, a multi-branch fusion network is implemented and trained on an EKG image dataset rather the standard EKG signals to achieve significant detection accuracy and yield an inter-patient classifier. While, finally, [17] trained a multi-lead residual neural network to detect and locate myocardial infarctions from 12-lead EKGs. However, although they set out to create an inter-patient classifier, they end up with something in between intra and inter, as they split their dataset along file lines and not patients, thereby contaminating their test sets.

However, as there is no standard sampling frequency that is widely used, one common drawback of the currently available MI detection methods is that they are dependent on the sampling frequency of the EKGs used to train or develop them. Therefore, there would be significant benefit from developing a myocardial infarction detector and/or classifier which is frequency independent. In other words, the detector's performance would not vary over a predetermined frequency range, making it amenable to existing cardiac monitoring equipment of varied sampling frequencies with no need to retrain or redevelop the underlying model. Furthermore, as lead II electrocardiograms are currently widely available from exercise equipment, activity trackers, and fitness monitoring devices, it

would be very practical and technically significant to make such detector work with single lead, specifically Lead II, electrocardiograms.

Moreover, several attempts have been made at automating the R-peak detection in EKG signals, and multiple strategies have been developed for effective noise removal [32–36]. R-peak detection is especially useful as a starting point for the decomposition of the cardiac signal into its essential subcomponents and to subsequently determine and diagnose conditions associated with abnormal cardiac rhythms [37, 38]. Therefore, real-time detection of R-peaks would be extremely beneficial for the effective analysis of EKG signals and for online diagnosis and detection of abnormalities associated with the cardiac rhythm.

Among the many methods that deal with multivariate analysis, the Independent Component Analysis (ICA) decomposition has played a historically leading role in signal processing and denoising, especially where the observed signals are assumed to be mixtures of an original set of independent sources [39–42]. This statistical technique is widely applied in other disciplines such as electroencephalogram (EEG) denoising, which by isolating unknown sources of noise, ICA manages to minimize the degrading effects such sources have on the original signals to produce highly accurate results [39, 41]. ICA is often used during the pre-processing phase to remove noise and external perturbations to the signal that degrade the detection process [41–43]. It is a very powerful tool that allows for the independent sources present in mixed signals to be effectively separated. However, due to the nature of ICA, the separation matrix is often randomly initialized, resulting in independent sources that are differently ordered every time the algorithm is performed.

It is important to note that all these signal processing algorithms are often performed offline and in the background over an entire record, where the clinicians do not have ac-

cess to the decision process, and what could have been perceived as noise and is therefore removed or ignored. This is a common drawback of filtering techniques that might lead to the loss of important information, which may result in a degraded classification or identification process.

Consequently, to expand on the literary body of research related to myocardial infarction detection, and with the intent to address the aforementioned design deficiencies and drawbacks for effective MI detection, this dissertation explores and develops various machine learning methods and applications for the detection of myocardial infarctions from electrocardiogram signals. The ultimate goal is to develop a single-lead (Lead II electrocardiograms), single-beat and frequency independent detector for the accurate and real-time detection of myocardial infarction.

Chapter 2 starts by introducing a method for the real-time detection of ventricular depolarization events in the PQRST complex from 12-lead electrocardiograms using Independent Component Analysis (ICA). A slightly different use of ICA is proposed for electrocardiogram analysis and R-peak detection/localization. As ICA benefits from signals recorded from different points of views and yields the same number of independent sources as there are mixed signal, a 12-lead EKG will yield 12 independent components, one of which is bound to contain the R-peaks (as they are the most distinctive feature of the EKG). This approach separates the components of the cardiac electrical signal and applies the detection process on them. It preserves all the leads recordings of the signal for observation and validation while still performing heart disease biomarker detection over the independent components. The potential benefits of such real-time implementation are far reaching, from the online diagnosis of disease and related abnormalities to its use in tracking heart functioning during the testing and development of cutting edge research

and treatments, such as transcranial magnetic stimulation (TMS) [44], and automated localization and segmentation of beats for MI classification.

Chapter 3 assesses the different types of classifiers present in the research body. It also explores the three main types of data splits (beat split, file or record split and patient split) reported therein, along with their drawbacks, benefits, and the undesired potential for overfitting that results when datasets are not properly engineered. Training versus testing empirical implementations are provided in this chapter to clearly demonstrate the behavior and training progression of each of these data splits over the training epochs.

Chapter 4 introduces a multilayer Long-Short Term Memory (LSTM) neural network design to identify infarcted patients from a single heartbeat of a single-lead (lead II) electrocardiogram. This task is particularly challenging as detecting generalized cases of MI from Lead II alone which focuses on the inferior part of the heart and therefore has a hard time detecting anterior, antero-septal, posterior, and lateral MIs. However, the choice made on this specific lead is that it is extensively used in everyday exercise equipment and fitness trackers, making the deployment of the proposed algorithm feasible and cost effective. Although various types of artificial neural networks have been used in the past to identify myocardial infarctions [6, 13, 16, 17, 20–23], we choose to use LSTMs because they have a long and proven track record at classifying time-varying signals such as text, speech, and video [45–52]. Hence, the genesis for using LSTMs to identify heart ailments from electrocardiograms. By combining the resulting neural network with an algorithm that detects the R-peaks in real time [53], nearly instantaneous detection of myocardial infarctions can be achieved. Through fast diagnosis, this integrated method (LSTM and real-time R-peak detection) could become extremely beneficial for monitoring patients under cardiac stress and/or at risk of myocardial infarctions, while still signifi-

cantly improving the prospects for both the correct diagnosis of the disease condition and the planning of early treatment.

Chapter 5 expands on 4 and introduces a fully integrated 12-lead real-time classifier with even higher detection metrics and a deeper neural architecture. It proposes a novel and complete pipeline that achieves 40 millisecond myocardial infarctions detection leveraging a deep-LSTM architecture with real-time beat detection and segmentation from Chapter 2. We once again use LSTM networks for their long track record of positive results at classifying time-varying signals. Herein, the resulting real-time classifier could be of great value to the population at risk of MI and could serve as a monitoring tool for gauging disease progression and benefits gained from early intervention and treatment planning.

Chapter 6 expands the work of Chapters 4 and 5 and explores the possibility of a real-time frequency-independent single-lead single-beat Myocardial Infarction detector. The key insight of this chapter is the frequency independence of the MI detector. This is of pivotal importance, as there is no standard sampling frequency for electrocardiograms, therefore making the available EKGs span many frequencies in the spectrum. Thereby, by making the detector frequency independent, it could be readily deployed to any available device that already collects electrocardiogram signals and has adequate processing power. Furthermore, it expands the reach and potential benefits of the proposed algorithm.

Finally, Chapter 7 provide a retrospective on the aims and contributions of this research endeavor towards the effective and accurate detection of myocardial infarctions. The incremental advances in time that were made through this work are highlighted in terms

of the drawbacks faced, remaining hurdles that need to be overcome, as well as in terms of real-world deployment and computational burdens that were faced in seeking real-time processing.

CHAPTER 2

REAL-TIME R-SPIKE DETECTION IN THE CARDIAC WAVEFORM THROUGH INDEPENDENT COMPONENT ANALYSIS

2.1 Goal

In this chapter, we propose a use of Independent Component Analysis (ICA) focusing on the R-peak detection as a relevant and distinct characteristic in the QRS complex in EKG. ICA benefits from signals recorded from different points of views, yielding the same number of independent sources as there are mixed signal. Therefore, a 12-lead EKG will yield 12 independent components, one of which is bound to be the R-peaks, as they are the most distinctive feature of the EKG. Our approach separates the components of the cardiac electrical signal and applies the detection process on them. This approach preserves all the leads recordings of the signal for observation and validation while still performing heart disease biomarker detection over the independent components. The potential benefits of such real-time implementation are far reaching, from the online diagnosis of MI and other heart ailments diseases to its use in heart monitoring during the testing and development of cutting edge research and treatments, such as in the use of transcranial magnetic stimulation (TMS) to assess any effects post treatment [54].

2.2 Materials and Methods

2.2.1 Data Acquisition

The data used for the development of this method was provided by The PTB Diagnostic EKG Database [55] [56]. It consists of a collection of 549 records from 290 subjects.

Ages ranged from 17 to 87 with a mean of 57.2. There are 209 men and 81 women. Each record contains 15 simultaneously measured signals: the conventional 12 leads (I, II, III, aVR, aVL, aVF, V1, V2, V3, V4, V5, V6) and 3 Frank lead EKGs (vX, vY, vZ). The signals were sampled with a 16-bit ADC at 1000 samples per second, over a range of ±16.384 mV. There is also a clinical diagnosis associated with most of the records. The diagnosis available are: Myocardial infarction (148), Cardiomyopathy/Heart failure (18), Bundle branch block (15), Dysrhythmia (14), Myocardial hypertrophy (7), Valvular heart disease (6), Myocarditis (4), Miscellaneous (4), Healthy controls(52).

2.2.2 Data Filtering

The best noise to have is none. However, sources of noise are always present and we must deal with their interference. Some noise sources (i.e. power line interference) can be relatively easy to remove, others present more of a challenge and can lead to unwanted consequences (i.e. muscle movements capture by equipment, eye blinks in EEGs), and some are unavoidable (i.e. signals emanating from an unknown source, or simply unwanted interference that we are unaware of).

The simplest noise sources to deal with are external in nature, i.e. the environment, AC power lines, lighting, and interference from various electronic equipment. When recording EKGs, we are interested in a very specific set of electrical signals generated by the functioning of the heart, all other signals that are not related to it are simply noise and hinder our ability to perform proper processing.

Several methods have been proposed to deal with noise removal over the past decades. The process of filtering can be performed either in the time or frequency domain and some of the most widely used techniques to do so are noise subtraction using linear regression, adaptive filtering, and data decomposition.

In this study, we use simple filters, a 60Hz stop-band Infinite Impulse Response (IIR) to remove power line interference and a 500 milliseconds rolling average filter to reduce low-frequency noise.

2.2.3 Independent Component Analysis

The key assumption of Blind Source Separation (BSS or independent component analysis) is that the observed signals are a mixture of an original set of independent sources. There are different approaches that differ in the algorithms and the information used to estimate the mixing matrix and the source signals.

Second order statistical methods are based on the assumption that the original signal sources are uncorrelated and aim to decompose the observed signals into several uncorrelated components. Principal Component Analysis (PCA), perhaps the most widely known method, decomposes a given time series into a number of orthogonal (uncorrelated) components of decreasing significance, such that most of the variance of the original signal is contained by a small subset of the principal components (the components with highest corresponding eigenvalues). PCA is most often used as a dimensionality reduction method.

When the original signal sources are assumed to be independent, methods based on higher order statistics can be used to decompose the observed signals. Various methods that make use of various measures of statistical independence for independent component analysis (ICA) have been developed [57, 58].

An independent component analysis method searches for a linear transformation that minimizes the statistical dependence between the components of a given signal. The expansion of mutual information is utilized as a function of cumulants of increasing order to define a suitable search criterion [59].

Suppose we have a random variable S, with probability density function $P_s(s)$, and the cumulative distribution function $F(s) = P(S \leq s)$ defined as:

$$F(s) = \int_{-\infty}^{s} P_s(t)\, dt \;\; or \;\; P_s(s) = F'(s) \tag{2.1}$$

Therefore, we can model the distribution of the variable S either by specifying its probability density functions or its cumulative distribution function.

Assuming the data comes from n original sources, meaning $S \epsilon \mathbb{R}^n$(n different parts of the heart), we define $s_k^{(i)}$ as the signal from the independent component k at time i.

What we observe in the EKG recording is an actual combination of the independent sources, assuming linearity and ignoring propagation speed differences as well as conductance disparities for simplicity sake:

$$X^{(i)} = A * S^{(i)} \tag{2.2}$$

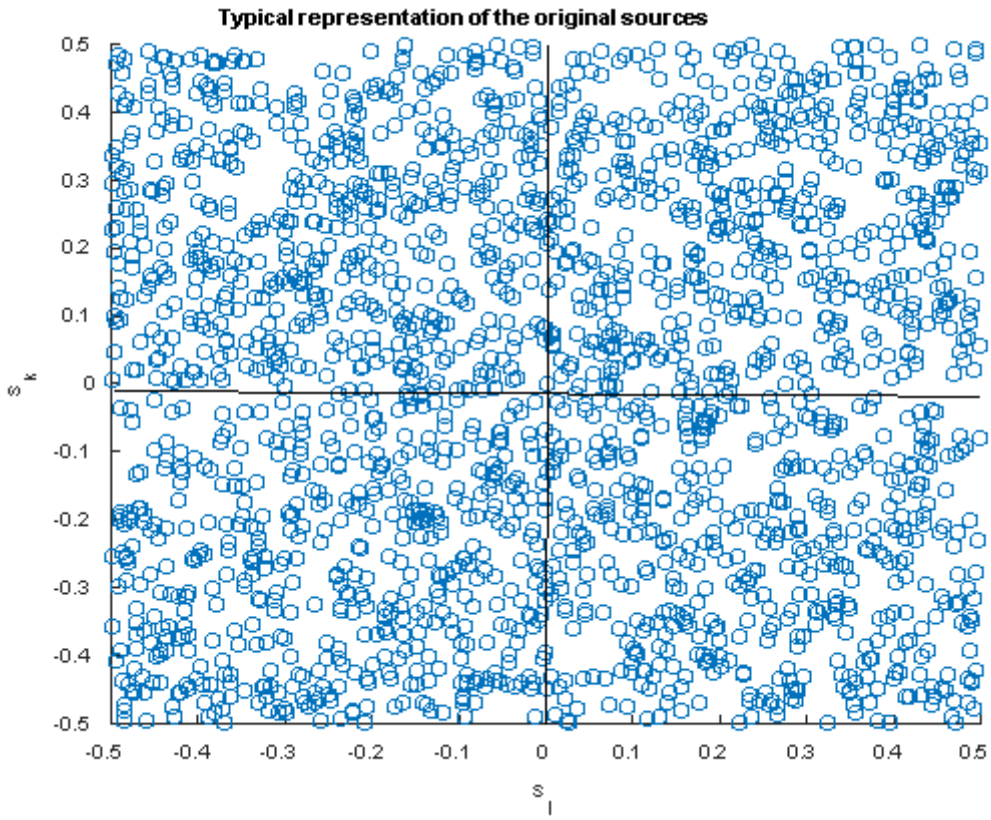

Figure 2.1: Independent Components.

where $X^{(i)} \epsilon \mathbb{R}^n$ corresponding to the n electrodes. Here the matrix A is known as the mixing matrix and what we observe at the $j^t h$ electrode at time i is:

$$x_j^{(i)} = kA_j ks_k^{(i)} \tag{2.3}$$

The goal of the ICA method is to find the optimal unmixing matrix $W = A^{-1}$ so that we can recover the original sources from the observable variables:

$$S^{(i)} = W * X^{(i)} \tag{2.4}$$

To further illustrate this, lets suppose each of the original sources (independent signals) are random white noise. This way, we can think of a 2D plot of any of the sources as the

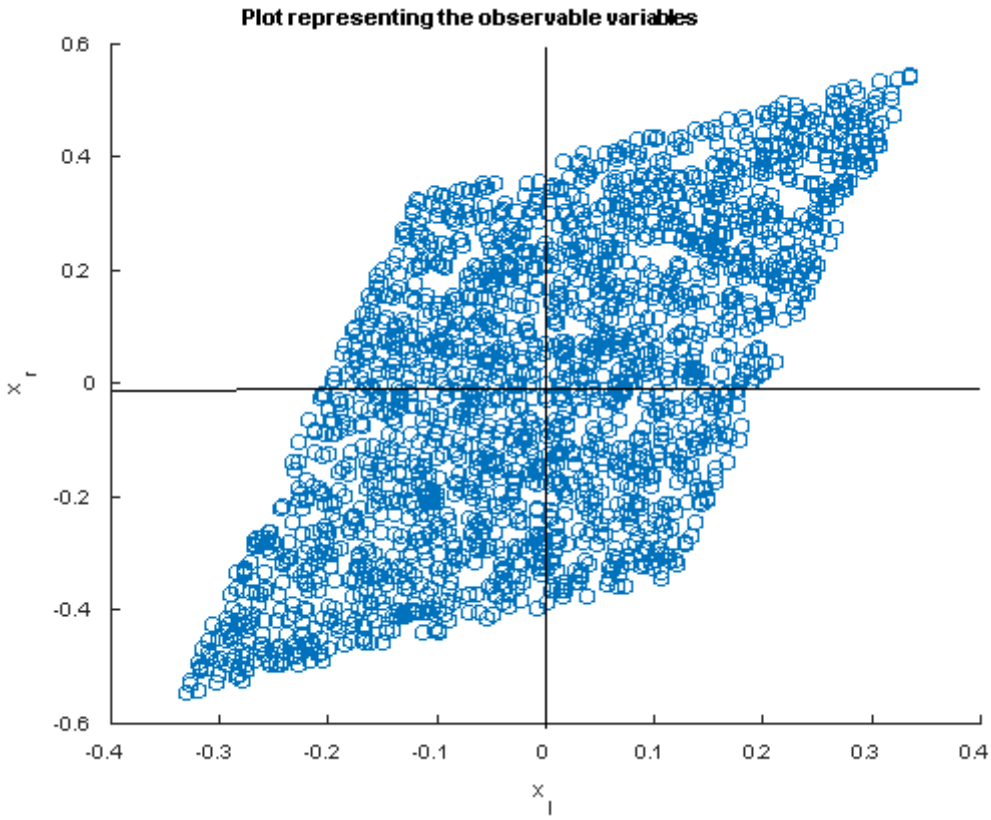

Figure 2.2: Mixture of Independent Components.

plot represented in Figure 1. On the other hand, a typical sample of what is recorded by the electrodes is represented in Figure 2.2.

The purpose of ICA is to find the transformation (*matrix* W) that would provide the results shown in Figure 2.1 from the mixture shown in Figure 2.2. Since the original sources are independent, the probability density function of s is going to be given by the product of the marginal probabilities of each s_i:

$$P(s) = \prod_{i=1}^{n} P_s(s_i) \tag{2.5}$$

Then the density for X will be:

$$P(x) = [\prod_{i=1}^{n} P_s(w_i^T x)] * |W| \tag{2.6}$$

Where w_i^T is the $i^t h$ row of the W matrix.

To complete the formulation of the ICA model, it is necessary to choose a probability density function for the sources s_i. What is usually selected is the cumulative distribution function as any signal that goes from zero to one. State-of-the-art ICA algorithms use a variety of functions. Some of them maximize one of the following: non-gaussianity, independence, or complexity.

The complete ICA model then follows as: Given my training set $\{x^1, x^2, , x^n\}$, we can derive the log likelihood of the parameters as:

$$l(W) = \sum_i log(\prod_j P_s(w_i^T x)) * |W| \tag{2.7}$$

The stochastic gradient descent algorithm is used to arrive at the optimal mixing matrix W:

$$W_{k+1} = W_k + \alpha \nabla_W l(W) \tag{2.8}$$

where α is the learning rate, and $\nabla_W l(W)$ is the partial derivative of the log likelihood function with respect to W.

The proposed approach uses FastICA [17], a reputed ICA algorithm that relies on the maximization of non-gaussianity of the independent sources. Traditionally, the mixing matrix is randomly initialized but through experimentation we have observed that using the identity matrix as a starting point tends to result in the first independent component being the most distinctive feature of the electrocardiogram, defining the R wave in this case.

The configuration of the electrodes is not subject to change during the collection of the data, and as such, the separation matrix of the signal should not change either. Therefore, we perform the ICA algorithm in the first 10 seconds of data and use the resulting separation matrix for the remainder of the collection. This allows us to perform real-time detection of the R peaks.

2.2.4 Modeling background interference and selection of R-peak candidates

The data used for this study comes from the recording of 12 distinct electrodes signals from different anatomical locations (which are transformed in to 12 signals). However, there can be many more co-occurring processes in the heart ($n \gg 2$). Since ICA inputs are these 12 signals, we can only recover 12 independent components out of the algorithm. Our assumption is that the R spikes are strong enough in the mixed signals in terms of frequency and amplitude that they will be captured completely by a single component of the ICA transformation. We will call it the R-component. In other words, the R spikes are going to be represented and clearly observed in one of the independent components of the ICA algorithm outputs. In the ideal case, with 12 ongoing recordings in the heart, the R-component will only capture the R spikes with only minimal noise. The real case scenario is that there still exist n-12 independent relevant processes that have no option but to be distributed among the 12 independent components being extracted. This leads to non-minimal noise in each of the output signals, including the R-component. Proving that this noise is Gaussian might be tedious in terms of the applicability of the central limit theorem and the assumption on how this electrode signals are measured. Moreover,

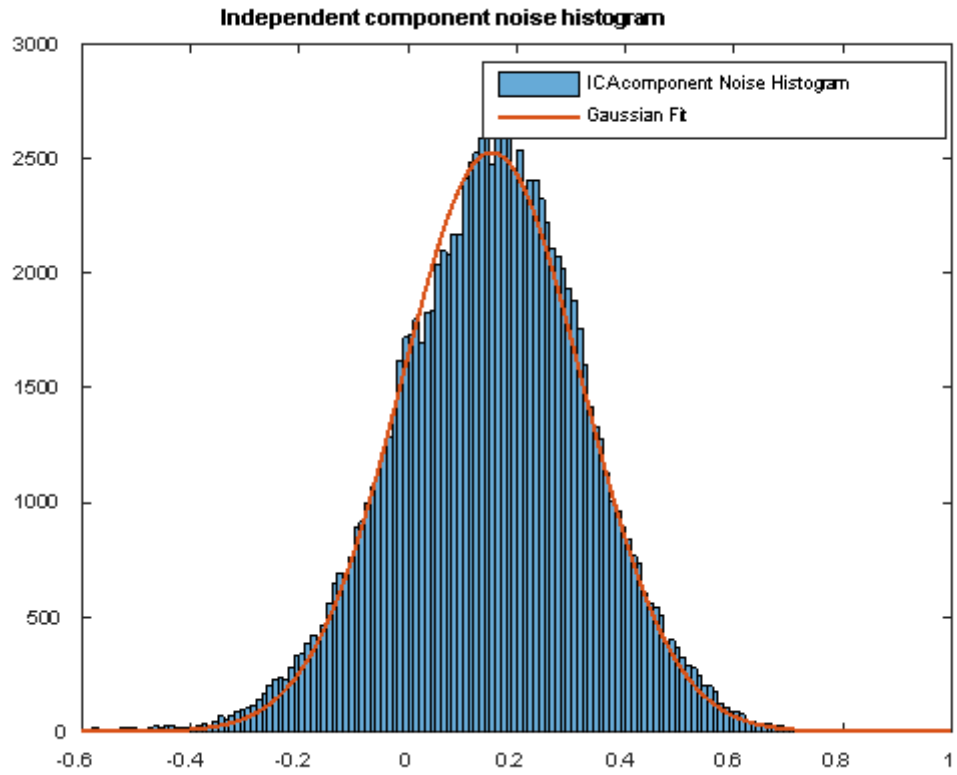

Figure 2.3: Histogram of the First Independent Component.

we do not plan to do statistical inferences based on normality assumptions. However, a near normal behavior could still be beneficial to somehow model and deal with this noise.

One empirical way to corroborate the normality assumption is to look at the histogram of the signal. We start by picking a random patient and obtaining the corresponding R-component signal. We then remove the extreme values that might correspond to the R-spikes or some other artifacts to leave only the supposed Gaussian noise behind. Figure 2.3 shows the histogram of this signal.

It can be seen from the figure, that the normality assumption is not a bad approximation. The Q-Q plot for this signal is also provided in Figure 2.4.

It can be observed that there is a slightly deviation from normality at the lower quartiles resulting in a heavier left tale. This is likely due to the presence of noisy artifacts or the presence of any other independent source in this component. Either way, the normality assumption is still valid for our practical purposes.

This behavior is representative of all the patients in our study. We did not proceed any further with normality testing, as it would have not been very useful.

The normality assumption allows us to model this independent component as a Gaussian-distributed background signal with the R-peaks superimposed. As a consequence, whatever fits inside the Gaussian distribution is considered background noise, and those events not explained by the model (3 standard deviations or more away from the mean) are considered R-peak candidates.

Taking advantage of the Normal distribution properties shown in the graph below, we can setup thresholds that would identify the occurrence of R-spikes:

Therefore, we proceed by computing the mean μ_j and the standard deviation σ_j of a 500 milliseconds rolling window in the selected component.

Following the model of the Gaussian distribution visualized in the figure, we can conclude that 99.7% of the noise will be within the $[\mu_j - 3 * \sigma_j, \mu_j + 3 * \sigma_j]$ range and any point that lays outside of this range is a possible R-peak candidate. Therefore, the algorithm is to test for the following condition:

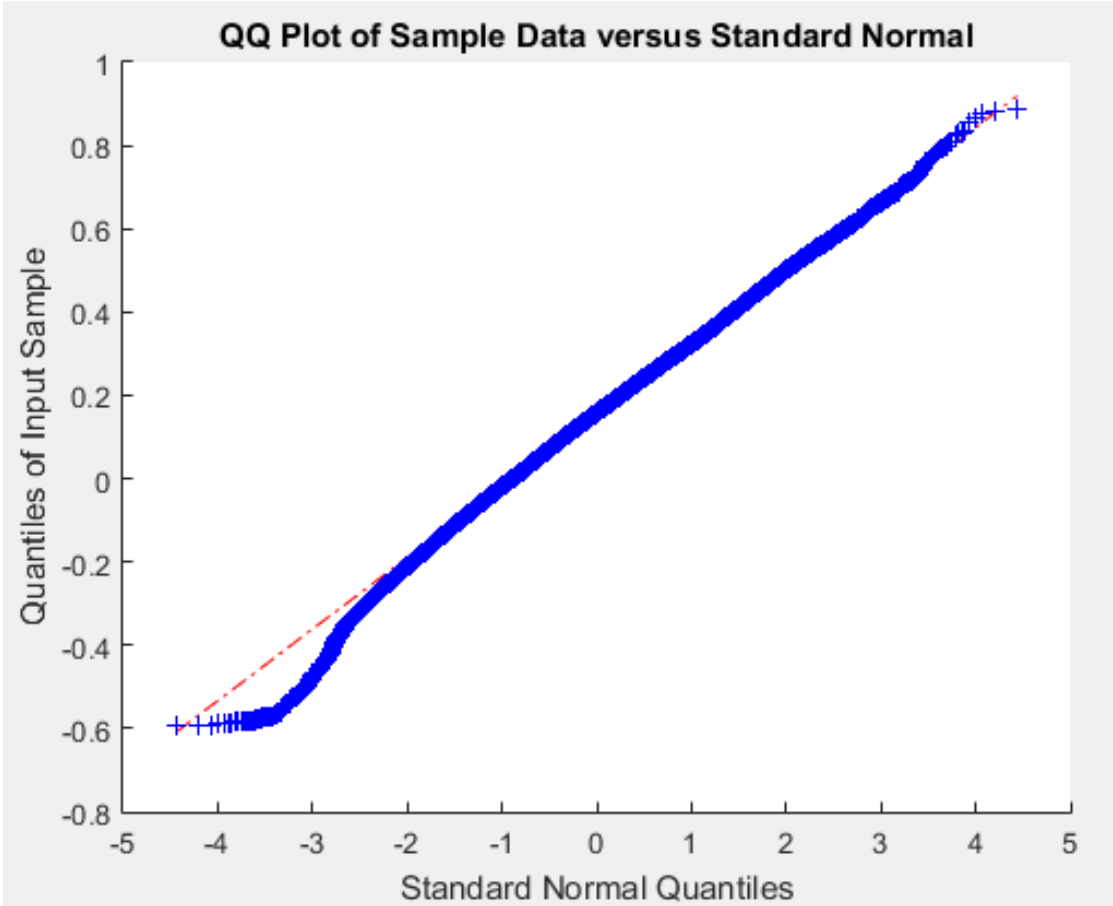

Figure 2.4: Q-Q plot of First Independent Component vs Standard Normal distribution.

$$if(\mu_j - 3 * \sigma_j \leq s_j^{(i)} \leq \mu_j + 3 * \sigma_j)$$
$$\{the\ i^{th}\ sample\ of\ the\ j^{th}\ signal\ is\ a\ spike\ candidate\} \quad (2.9)$$

When this condition is met, a set of possible R-peak candidates are expected.

The last detection step is to check whether the width of the proposed peaks is within 10 and 60 milliseconds. Figure 2.6 shows the first independent component (in black) alongside the rolling threshold (blue, as defined by the condition in equation 2.9).

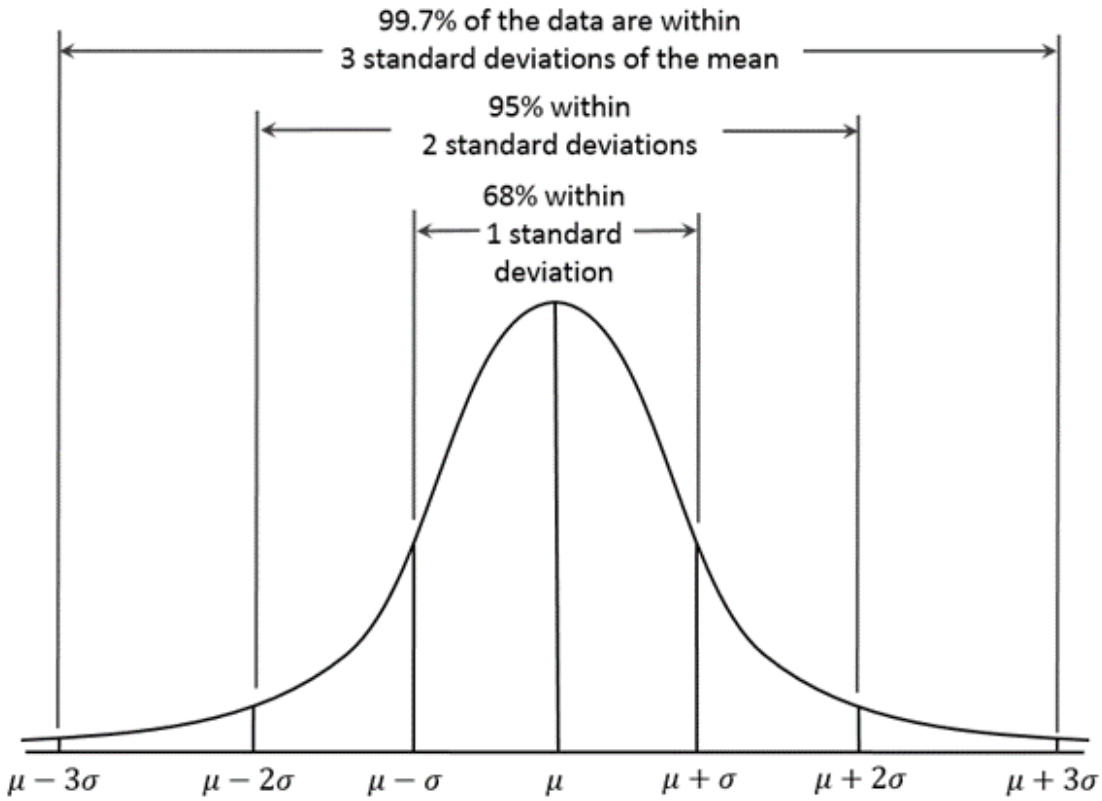

Figure 2.5: For the normal distribution, the values less than one standard deviation away from the mean account for 68.27% of the set; while two standard deviations from the mean account for 95.45%; and three standard deviations account for 99.73%. https://en.wikipedia.org/wiki/Normal_distribution.

2.3 Results

Initial experimental results and analysis can be reported for a subset of 80 randomly selected recording from the 549 available. Said sample is composed of 40 healthy individuals and 40 unhealthy ones with one or more diagnosis (myocardial infarction, cardiomyopathy/heart failure, bundle branch block, dysrhythmia, myocardial hypertrophy, valvular heart disease, myocarditis, miscellaneous). These records collectively contain over ten

(10) thousand R waves. Table 2.1 contains the summary statistics for the proposed method over the processed records.

Table 2.1: Summary Statistics for the ICA R-Spike Detection Method

Subjects	True Positives	False Positive	False Negatives	Detection Rate (Sensitivity)
Controls	5411	6	112	97.9%
Others	5188	3	90	98.29%
Totals	10599	9	202	98.13%

We are unable to report accuracy readings from one of the files as it contains numerous artifact that make the proper labeling of the R-peaks (by humans) impossible. Figure 2.7 depicts a portion of such recording along with the attempts of the proposed algorithm to detect the R-peaks (red vertical bars). Figure 2.8 shows the computed threshold (blue line) and the independent component being looked at. Finally, Figure 2.9 show how the artifacts affect all the independent components extracted from the EKG.

For all other 79 files, the proposed method performed accordingly, properly identifying over 98 percent of the R-peaks. Most of the files containing false negatives misdetected only the first, or first two, beats while the rolling threshold was being initialized. Only 8 EKG records had more than three (3) false negatives.

Figures 2.10 through 2.12 show how the algorithm detects the R-peaks. Figure 2.10 contains the 12-lead input EKG with the R-peaks labeled (red vertical bars), figure 2.11 is the results of applying independent component analysis to the input signal in 2.10, and Figure 2.12 demonstrates how the proposed rolling threshold works when identifying potential R-peak candidates in the identified independent component.

Another observation that can be made is how the algorithm ignores signal artifacts, as they would tend to belong to a different component than the one containing the R-peaks. Figure 2.13 is such a case, where unrecognized artifact, that could hinder the accuracy of R-peak detection algorithms, are present and not detected as R-peaks. From Figure 2.15, it is apparent that the artifacts affect the first three independent components but are rather attenuated in the fourth, where the R-peaks are. Figure 2.14, as in previous instances, shows the independent component containing the R-peaks (black) along with the moving threshold being computed (blue).

It should be noted that When the EKG waveform contains significant perturbations, the first component could capture a signal other than the R-peaks. This is the case for some records of patients with myocardial infarction and some controls that contained significant artifacts were the first component would alternate between the Q, R, and S spikes, and the component containing the R-peak could be as far back as the third one. Table 2.2 is a summary of the location of the R-peaks within the independent components.

Table 2.2: Number of EKGs containing the R-peak in the given Independent Component

	IC1	IC2	IC3	IC4
# of EKGs	63	15	1	1

Although the R-peaks are contained within the first independent component in most instances, a procedure to automatically identify its location, or force them to appear in the first IC at all times, remains a challenge to be met as future work.

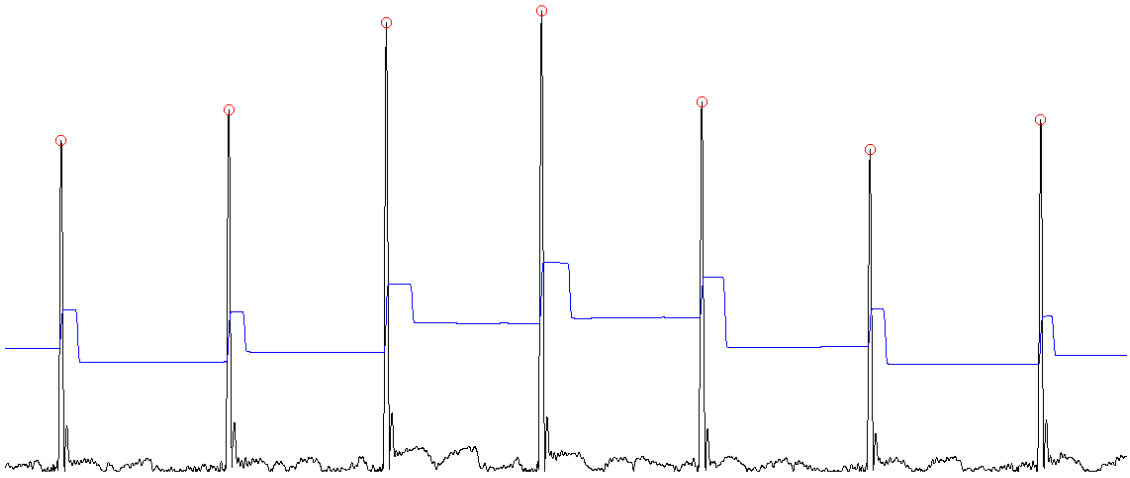

Figure 2.6: R-peak detection in First Independent Component (black) and Detection Threshold (blue).

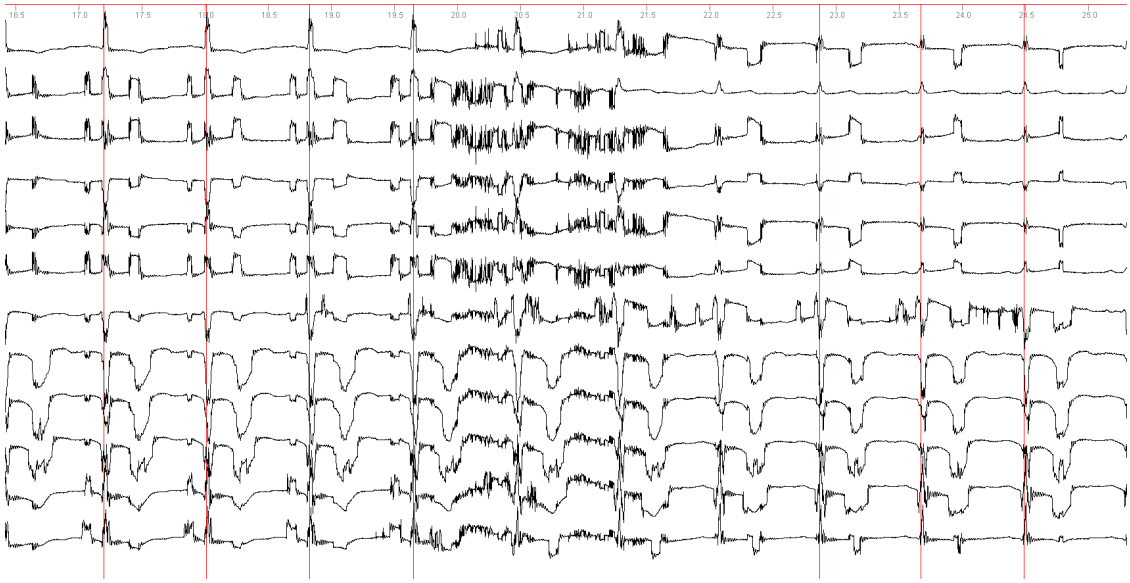

Figure 2.7: EKG Record 557 for patient 293. It contains multiple artifacts that make proper R-peak labeling impossible.

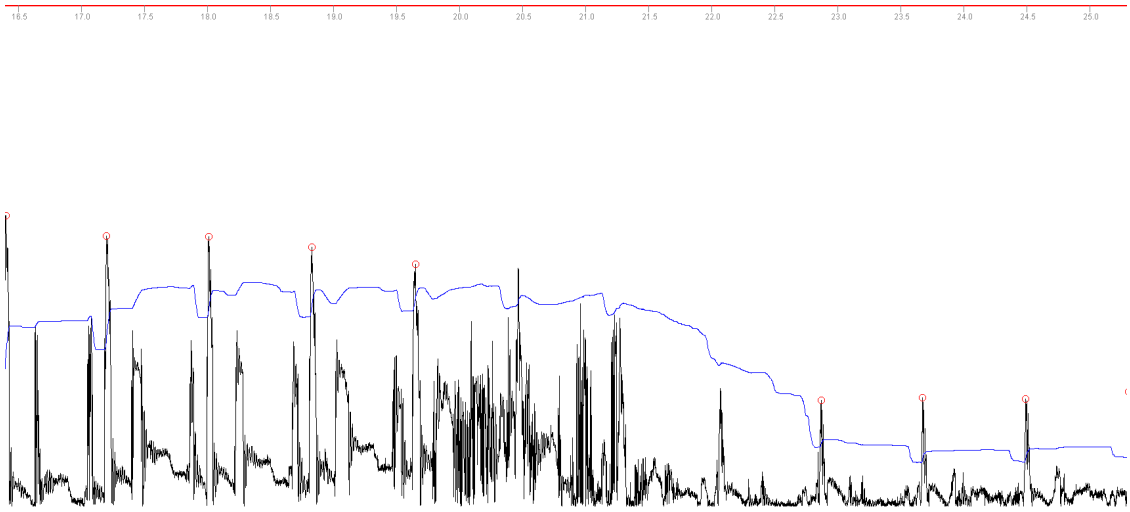

Figure 2.8: Second Independent Component (black) along with the computed rolling threshold (blue) for Figure 2.7.

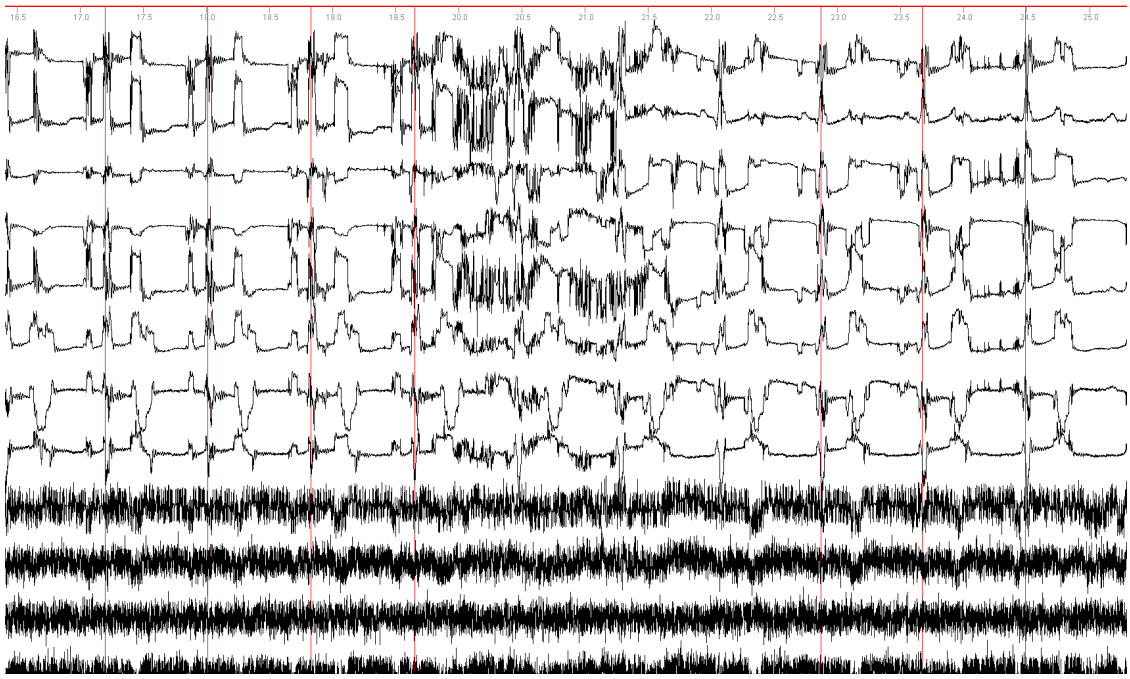

Figure 2.9: Independent Components of EKG Record 557 for patient 293. It shows how the artifacts from the EKG affect all components effectively not allowing the proposed method to extract or identify the R-peaks when they are present.

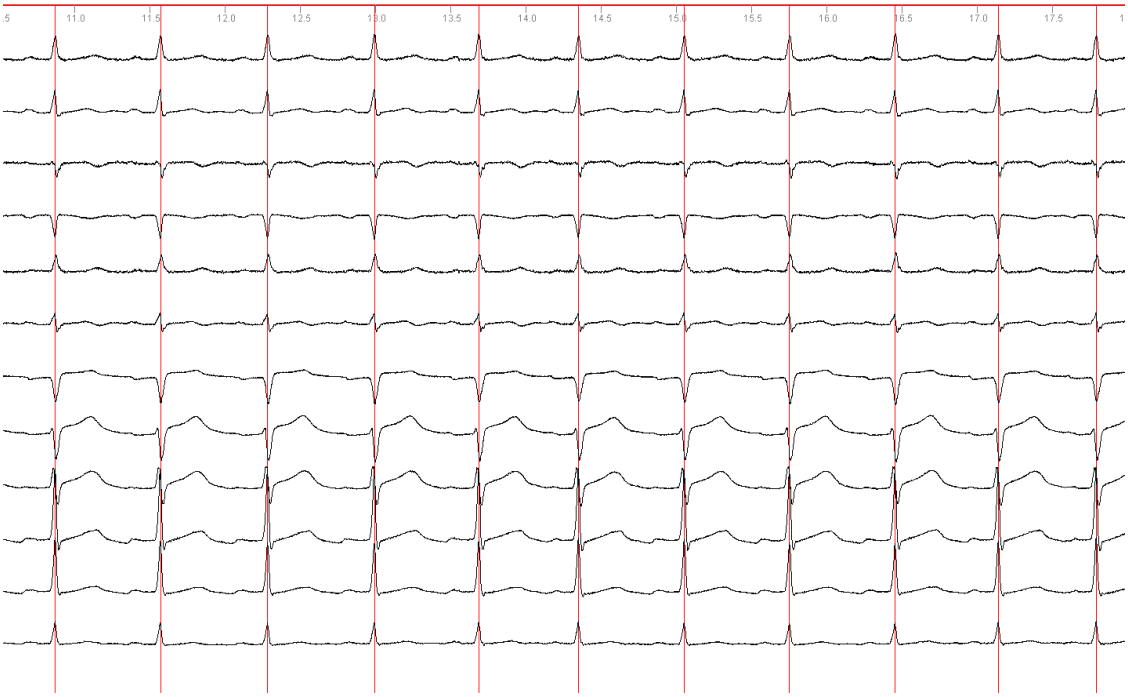

Figure 2.10: EKG Record 242 for patient 75. It is an EKG from a patient with a Myocardial Infarction.

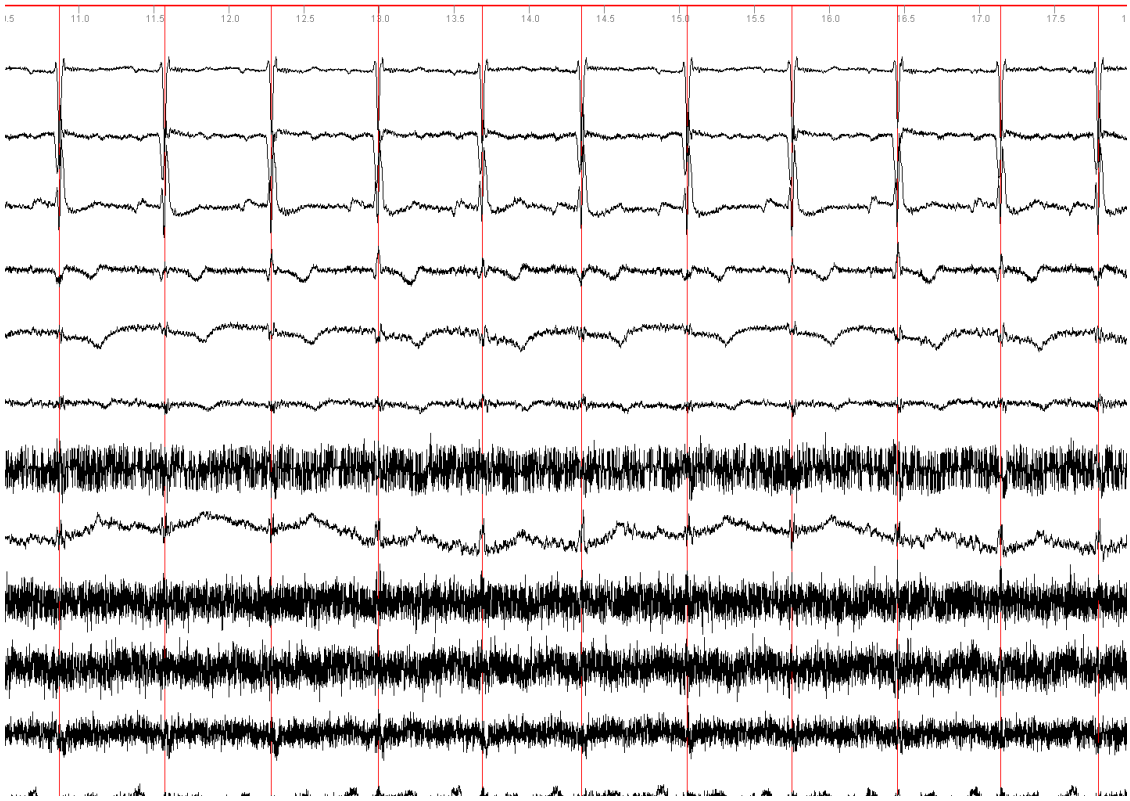

Figure 2.11: Independent Components of EKG Record 242 for patient 75.

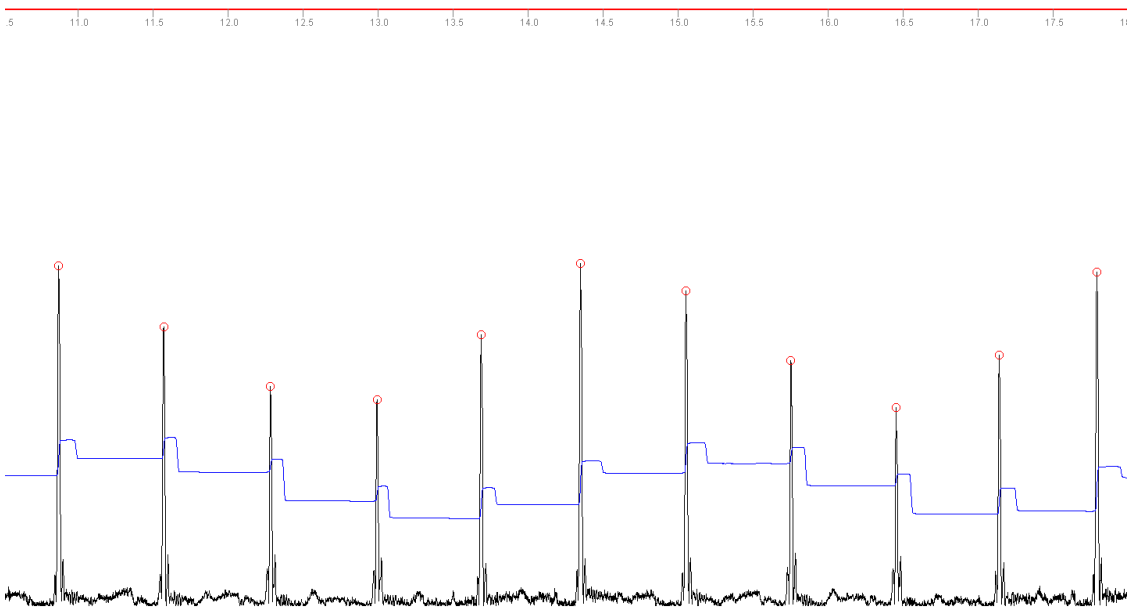

Figure 2.12: Second Independent Component (black) along with the computed rolling threshold (blue) forFigure 2.10.

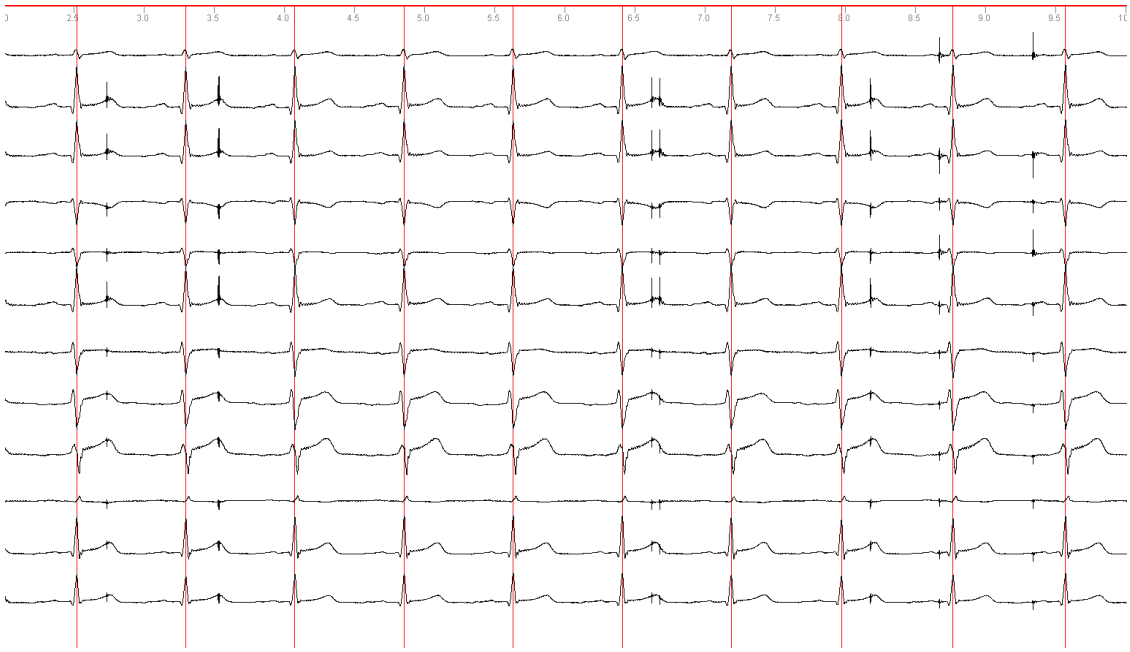

Figure 2.13: EKG Record 543 for patient 284. An EKG containing multiple artifacts that could hinder the performance of R-peak detection algorithms.

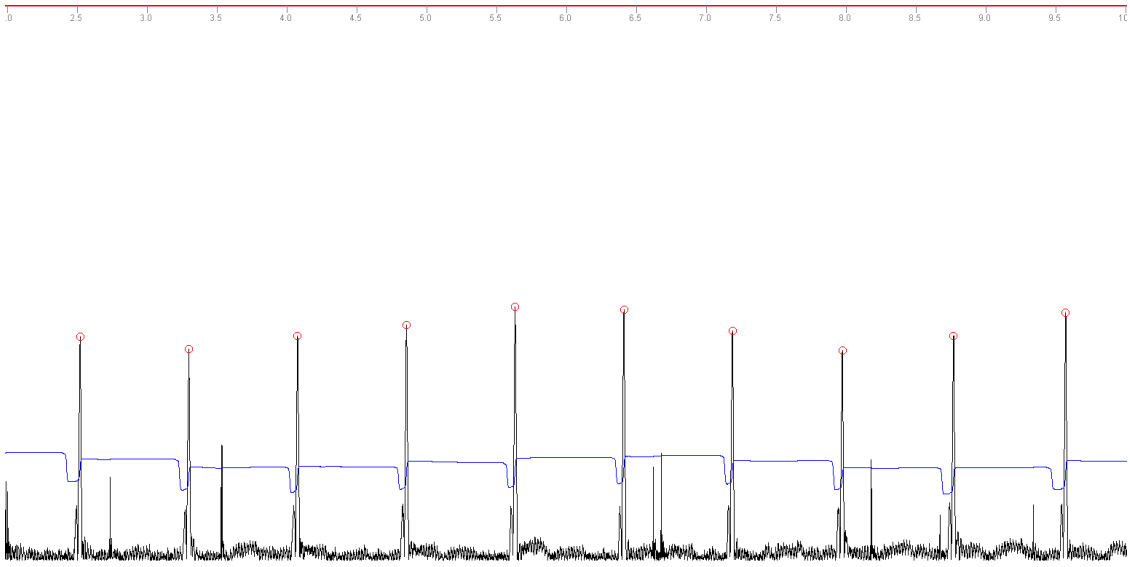

Figure 2.14: Forth Independent Component (black) along with the computed rolling threshold (blue) for Figure 2.13.

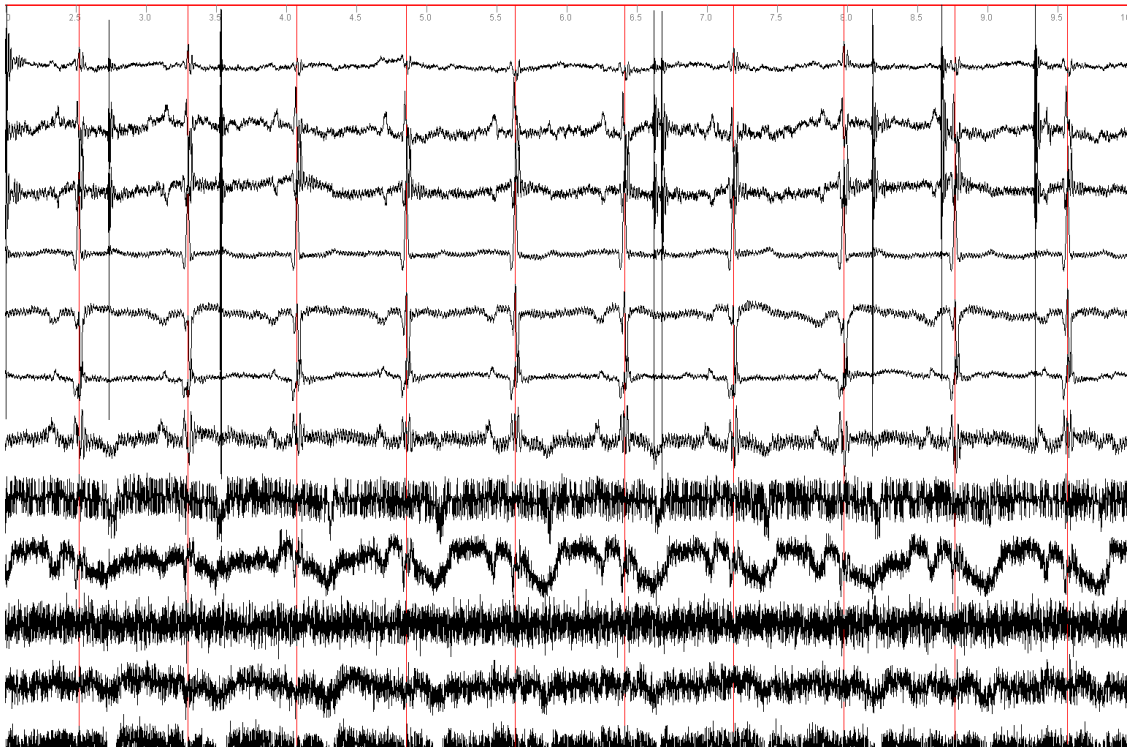

Figure 2.15: Independent Components of EKG Record 543 for patient 284.

CHAPTER 3

DATA SPLITS FOR MACHINE LEARNING TRAINING

3.1 Goal

This chapter examines the different types of classifiers present in the research body and explores the merits and drawbacks of the three main types of data splits which are beat split, file or record split and patient split. Measures of precaution are taken to asses the potential for overfitting that is present when datasets are not properly engineered. Empirical evaluations are performed to explore the testing and training performance of these different data split methods, highlighting their overall behavior in the training and testing phases and tracking their progression over the different epochs.

3.2 Materials and Methods

3.2.1 Data

The data used in this study was obtained from The Physikalisch-Technische Bundesanstalt (PTB) diagnostic EKG Database [60, 61]. This database consists of 549 12-lead individual EKG records collected from 290 subjects (209 men and 81 women) with ages ranging from 17 to 87 years old, of which only 148 Myocardial Infarction (MI) and 52 Healthy Control (HC) patients are used for this study. However, although 12 leads are available, we only consider Lead II for this study as it is widely used in wearable/portable devices and exercise equipment; thereby positioning the proposed network to be readily deployed in these existing devices and thus eliminating the need to create custom ones.

Each record has its associated diagnosis out of which we are interested in the ones containing acute myocardial infarction, old myocardial infarction, and normal control.

A 60Hz stop-band and a 500-millisecond moving average filters are applied to remove powerline and low-frequency noise. Once the data is filtered, we apply an Independent Component Analysis (ICA) based algorithm [62] to identify the locations of the ventricular depolarization events (also known as R peaks) and thereafter separate each individual heartbeat to be used in the training phase of the proposed multilayer LSTM network. Each training sample is one second long (500 milliseconds before and 500 milliseconds after the ventricular depolarization event) of the EKGs Lead II data (the information from all other leads is discarded).

3.2.2 Structure

In the PTB DB data is structured in a hierarchical inclusive manner as depicted in Figure 3.1. There are multiple patients, each patient can have multiple records, and each record has multiple heartbeats or sample segments.

Therefore, if we choose the size of any given training/testing sample to be a heartbeat, there are three primary ways in which they are separated into sets for training, testing, and/or validation:

Beat-Split: The electrocardiograms in the datasets are split into its constituent heart beats and these are further randomly added to one of the sets (i.e. training, testing, validation, or cross validation sets). In this particular data-split practice we only care that

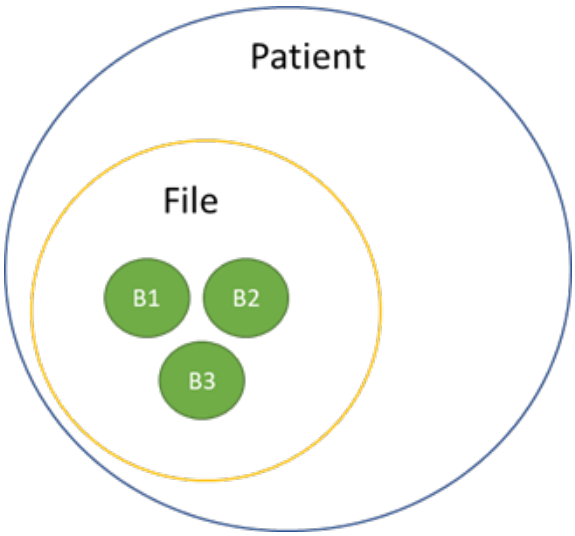

Figure 3.1: Data Structure.

independent heart beats seen during training are not also used for testing. However, individual patients could have a set of heartbeats in the training and another set of heartbeats, albeit different, in the testing set; making it the least restrictive of all methods. This technique is used in the literature to perform **intra-patient** classification [6–10, 12, 15, 24–28].

File-Split: Also known as Record-Split. As each patient in a given dataset may have multiple recording sessions (record), the individual heart beats associated with each of the individual records can only be present in one of the derived sets. That is, any given record cannot have simultaneous representation in the testing and training sets [15, 16, 23].

Patient-Split: Each patient (with all of its available records) can only be present in one of the following derived sets (i.e. training, testing, validation, or one of the cross validation sets). Here, data from patients used for training cannot be used for testing, and vice versa. This technique better ensures that the results achieved during testing will more adequately

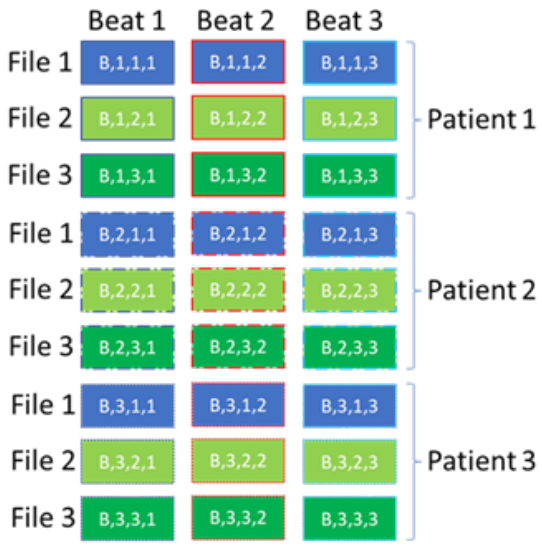

Figure 3.2: Patient Data Sample.

represent those of new unseen patients. This practice is used in the literature to perform **inter-patient** classification and is the most restrictive data-split method [8, 9, 13, 14, 16].

Figures 3.2 and 3.3 give a more visual example of how these splitting techniques would work in practice. Figure 3.2 introduces the data from 3 patients and figure 3.3 brakes them down into training and testing sets following the aforementioned split criteria.

Overarchingly, there are two main type of classifiers/detectors that can be achieved and are present in the literature:

Intra-patient classifiers would be those that take advantage of the reduced intra-patient variability (heartbeats of the same patient are more alike than heartbeats from different patients) to achieve a classification. These classifiers produce a classification result for new heartbeats of patients seen during training. Although they can be an effective tool and help speed up the learning process, special care must be taken to ensure that there is

Figure 3.3: Dataset Spliting Methods Examples.

some variability in the heartbeat/sample label, as if they are homogeneous, it will most likely tend to prompt the classifier to learn to distinguish the patients rather than the disease. Perhaps taking advantage of the lower intra-patient variability rather than the larger inter-sample variability of the disease itself.

Inter-patient classifiers, on the other hand, would be those that, once trained, can be effectively used to classify novel patients not seen during training. These classifiers are especially useful for cases were sample labels are homogeneous (only one label exists for all samples of the same subject). In this instance the testing performance would more realistically follow that of a deployed system, as the classifier would be exposed to novel data point during testing and final deployment.

3.2.3 Methods

To test how the different dataset generation methods affect the learning process we will make advanced used of the LSTM network introduced in Chapter 4. As a brief introduction to the model (depicted in Figure 3.4), we use a shallow LSTM architecture to minimize the occurrence of overfitting from possible memorization of samples that could occur with very deep architectures, along with RMSProp for weight updates, gradient clipping, and L2 regularization.

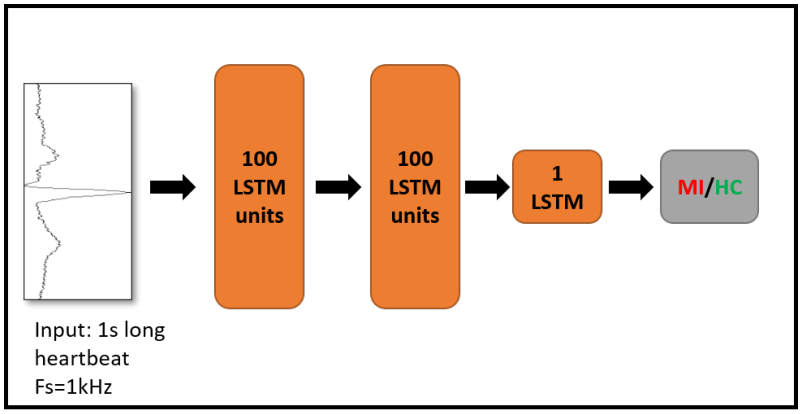

Figure 3.4: Simplified Network Architecture.

This model was implemented using DeepLearning4J version 0.9.1, a deep learning library for Java, and was trained and tested in a 64-bit Windows 10 PC with an AMD FX-8350 Eight-Core Processor, 32 GB of DDR3 RAM, and an NVIDIA GeForce GTX1070. Each training epoch ran for approximately 5 hours and covered roughly 60 thousand samples.

3.3 Results and Discussion

3.3.1 Effects of the Different Data Splits

We first focus on the beat-split method (used to generate intra-patient classifiers) and use 10-fold cross-validation. As with the beat-split technique, we only care that independent heart beats seen during training are not also used for testing. This has a side effect of allowing patients used for training to also be used for testing (since the distinction is made only on the individual heart beats). Figure 3.5 shows the training evolution of this particular test case.

Figure 3.5: Training Progress for Beat-Split Case.

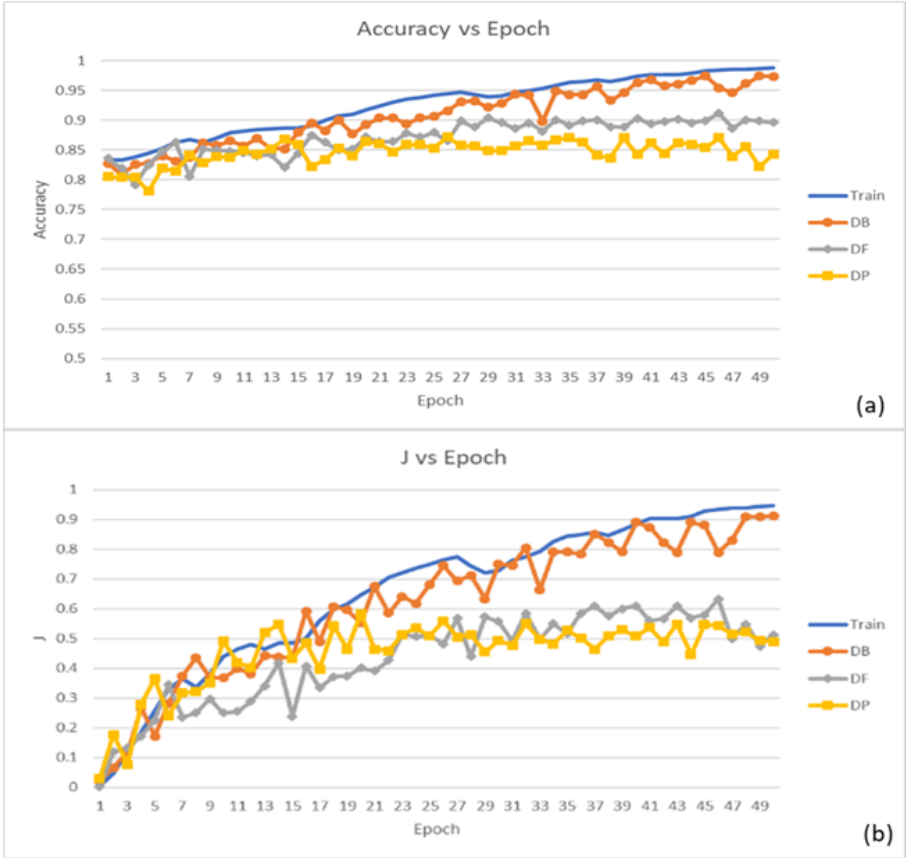

Figure 3.6: Training Progression of Different Split Methods.

These results show that the networks performance on the testing data continuously increases and closely tracks the training metrics. Therefore, it is important to review all testing results in combined plots to better understand the model as it is being trained. Figure 3.6 shows the testing results for the different types of data splits including beat-split (DB), file-split (DF), and patient-split (DP).

It is evident from these plots that the network first starts to show decreased performance on testing sets that use patient-split followed by test sets generated with file-split. That is, the model seems to start learning to classify the disease but eventually learns something else (perhaps patient level characteristics that help it distinguish hard heartbeats, taking

advantage of the lower intra-patient variability). As its performance starts to drop for patients never seen before (as it is the case in DP) and then from files (records) not seen during training (DF) it appears to learn features that are shared across files and eventually features shared across heart beats of the same patient; effectively memorizing patients by learning to identify which beats belong to patients seen during training and assigning them the appropriate diagnosis based on prior knowledge.

It is also important to point out that minor epoch-level differences among the three split modes (DB, DF, and DP) are due to the nature of training; where the classifier focuses on different features at different timepoints as it descends down the multidimensional gradient in its attempt to minimize the classification error. What should be taken from the plots is the general direction/trend of the plotlines themselves and not the local differences.

3.3.2 Randomizing labels of myocardial infarction patients and normal controls

Because the network could be learning many things, and not specifically memorizing patients, we take it one step further and randomize the labels (MI/Controls) of each patient before retraining the model using beat-split. This action effectively eliminates any existing link(s) between biomarkers and disease label. Therefore, if the network model was never about memorizing the patients but just about learning the nature of the disease, we shouldnt be able to attain a mean accuracy greater than 50%, given that the new labels are uniformly and randomly distributed. This time around, we trained the network for 90

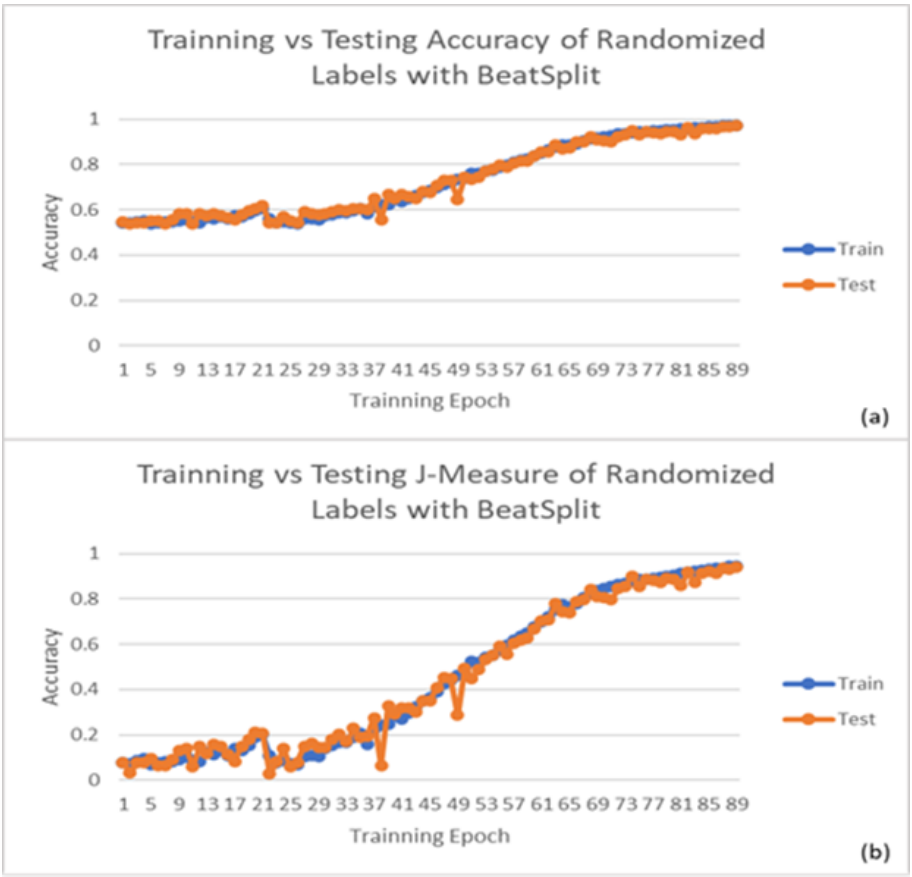

Figure 3.7: Training Over Randomized Labels.

epochs to identify any distinctive patterns that may arise. Figure 3.7 shows the results of the experiment.

These results clearly show that the proposed network is capable of learning features that help it identify heart beats from previously seen patients, effectively memorizing them, if adequate time is given. Therefore, it is of paramount importance that an adequate data-splitting method be used when creating the training and testing datasets, as choosing an inappropriate one will mask overfitting and lead to erroneously high testing metrics that are not representative of the models performance on the general population.

CHAPTER 4

NEAR REAL-TIME SINGLE-BEAT MYOCARDIAL INFARCTION DETECTION FROM SINGLE-LEAD ELECTROCARDIOGRAM USING LONG-SHORT TERM MEMORY NEURAL NETWORK

4.1 Goal

In this chapter, a multilayer Long-Short Term Memory (LSTM) neural network is developed to identify infarcted patients from a single heartbeat of a single-lead (lead II) electrocardiogram. This task for detecting generalized cases of MI from Lead II alone is challenging since Lead II focuses on the inferior part of the heart and therefore has a hard time detecting anterior, antero-septal, posterior, and lateral MIs. However, we choose to do so as this lead is extensively available in everyday exercise equipment and fitness trackers, making the deployment of the proposed algorithm feasible and cost effective. Although various types of artificial neural networks have been used in the past to identify myocardial infarctions [6, 13, 16, 17, 20–23], we choose to use LSTMs because they have a long and proven track record at classifying time-varying signals such as text, speech, and video [45–52]. Hence, the genesis of their use here to identify heart defects and other cardiac conditions from electrocardiograms. By combining the resulting neural network with an algorithm that detects the R-peaks in real time, the results indicate near-instantaneous detection of myocardial infarctions. Through this expedited diagnosis, this integrated method (LSTM with real-time detection of R peaks) could become extremely beneficial for monitoring patients under cardiac stress and/or at risk of myocardial infarctions while significantly improving the prospects for both the correct diagnosis of the disease condition and for the planning of early treatment.

4.2 Materials and Methods

4.2.1 Data

The data used in this study was obtained from The Physikalisch-Technische Bundesanstalt (PTB) diagnostic EKG Database [60, 61] as was the case for the study performed in Chapter III. Same database, same amount of patients, and same type of filtering techniques are used here as described earlier in Section 3.2.1.

After the data is segmented, it must be separated into groups to be used for training and testing. There are three primary ways in which electrocardiogram heartbeats are separated in the literature when used to detect myocardial infarctions as described in the previous chapter: Beat-Split, File-Split, and Patient-Split. However, as we aim to train a classifier that can be used in real life applications, we use the most restrictive data split available, namely Patient-Split, to ensure that the testing dataset best represents the population at large. As such, once all the independent heart beats have been segmented, each patient (along with all its segmented heartbeats) is assigned to one of 10 distinct and non-overlapping groups, to be used for ten-fold cross-validation. Imperatively, no one patients data is present in more than one group at any given time. This is important, as having individual patients simultaneously represented in both the training and testing phases results in overfitting that could yield to misleading high testing metrics, and hence overestimating the performance of the algorithm. Also, this type of data splitting is more amenable to real-world applications, as the network would be trained on a dataset of recorded patient data but will be used and deployed on patients not seen during the training phase.

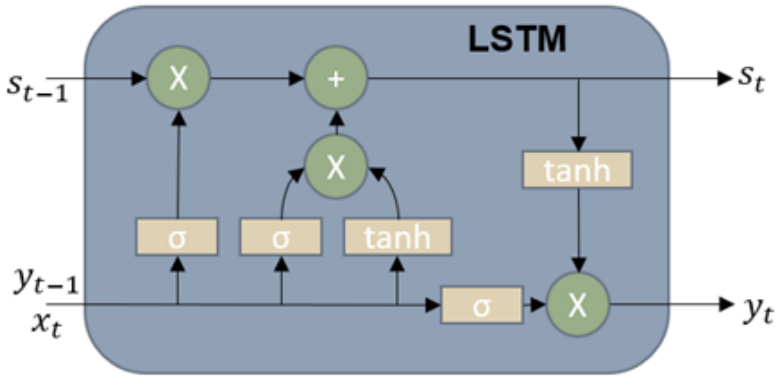

Figure 4.1: LSTM unit structure.

4.2.2 Network Architecture

We use a three-layer LSTM network to differentiate between subjects with myocardial infarctions (MI) and healthy controls (HC). Layer one and two consist of 100 LSTM units with hyperbolic tangent activation, while the third and last layer is a single LSTM unit with sigmoid activation function. Each training and testing sample is one second long and centered on the ventricular depolarization event (500 milliseconds before and after the R-peak) of the EKGs Lead II data. The design construct of the LSTM model is shown in Figure 4.1.

LSTMs, shown in Figure 4.2, are a type of recurrent neural networks (RNNs) that see the input signals as time-varying and can therefore make temporal relation inferences about the signals being categorized. Whereas, other common neural networks architectures used in the literature such as convolutional neural networks (CNNs) can only make spatial relations and are therefore limited by the size of their kernels. This is not the case for RNNs as they can learn to remember events indefinitely long in the past of the sequence and make decisions based on it. Also, LSTMs dampen the effect of vanishing gradients

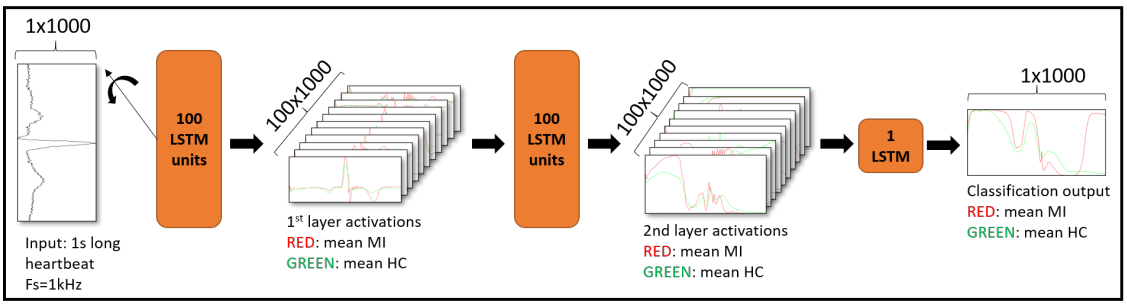

Figure 4.2: Network Architecture.

present in the original RNNs where events long in the past have diminishing influence in current decisions.

It can be seen from Figure 4.1 that an LSTM unit requires as an input its previous state s_{t-1}, its last output y_{t-1}, and the current input vector x_t to generate a new internal state s_t and a new output y_t. Therefore, an LSTM neuron can produce a classification prediction for every input timestep thereby generating time-varying classifications that becomes more accurate and final as more sample points of the signal are processed. This is especially evident in the output activations of the network in Figure 4.2, where the final classification label is produced during the last 200ms of the signal once the PQRST complex has been processed.

This architecture was particularly chosen as it requires no previous expert knowledge of the signal at hand, being able to come up with the features it deems important on its own from the training data. It is also narrower in terms of the required processing steps and hence less complex than other architectures reported in the literature for myocardial infarction classification, making it less prone to overfitting over a limited amount of data as a more complex model gives the network more flexibility to overfit.

The network itself was trained using stochastic gradient decent while implementing L2 regularization [62] with λ = 0.001, RMSProp updater [63] (learning rate = 0.1, weight decay = 0.95, and epsilon = 10e-6) with weight decay, and gradient clipping (5.0). The weights are initialized using the method proposed by Xavier Glorot [64]. The specific LSTM units used in our study were first described by Alex Graves in [65] to label sequential data such as speech and hand-written text.

The network described herein, was trained and deployed in a 64-bit Windows 10 PC with an AMD FX-8350 Eight-Core Processor, 32 GB of DDR3 RAM, and an NVIDIA GeForce GTX1070. Each training iteration (epoch) took approximately 5 hours and covered around 60K samples, depending on the cross-validation split being used. The network architecture itself was implemented using DeepLearning4J version 0.9.1, a deep learning library for Java.

4.3 Results

The network is trained using early stopping, for which after 10 epochs of no performance improvement we stop training and back up to the best set of weights. We repeat this process for each of the cross-validation folds. The mean training time is 34.4 epochs ($\pm$19.4 $epochs$ 95% CI) when optimizing for accuracy and 27.4 ($\pm$12.64 95% CI) epochs when the best J-Measure is the target. The maximum number of epochs required to train any given fold was 75 and the minimum was 4. Classification of a single 1-second sample takes around 40 milliseconds, which would be appropriate for online classification

as the time between fast heartbeats is around 300 milliseconds, well above the required processing time.

To measure performance, we have used Accuracy (4.1), F1-Score (4.2), Precision (4.3), Recall (4.4), Specificity (4.5), and Youdens J statistic (J-Measure) as defined in (4.6).

$$Acc = (TP + TN)/(TP + TN + FP + FN) \tag{4.1}$$

$$F1 = (2 * TP)/(2 * TP + FP + FN) \tag{4.2}$$

$$Prec = TP/(TP + FP) \tag{4.3}$$

$$Recall = TP/(TP + FN) \tag{4.4}$$

$$Spec = TN/(TN + FP) \tag{4.5}$$

$$J = Recall + Spec - 1 \tag{4.6}$$

Where TP stands for true positives, TN for true negatives, FP for false positives, and FN as false negatives.

Table 4.1 shows the detailed metrics obtained for both optimization methods (accuracy and J measure) on the testing datasets in the 10-fold cross-validation using the Patient-Split

approach, and Figure 4.3 shows the training progression vs the testing results of the 80 epochs during the training phase.

It can be observed from these figures that the network starts to stabilize in terms of testing accuracy early in the training, somewhere in between epoch 20 and epoch 25. At that point the network keeps on learning the training dataset, eventually reaching the high nineties before being stopped at epoch 80. This could be due to a lack of significant variability in the training set that would best represent the population at large, a problem which could be addressed by adding more patients into the training dataset.

Another significant fact to appreciate from Table 4.1 is that the Recall measure is significantly larger than Specificity for either optimizing condition. This is the case mainly

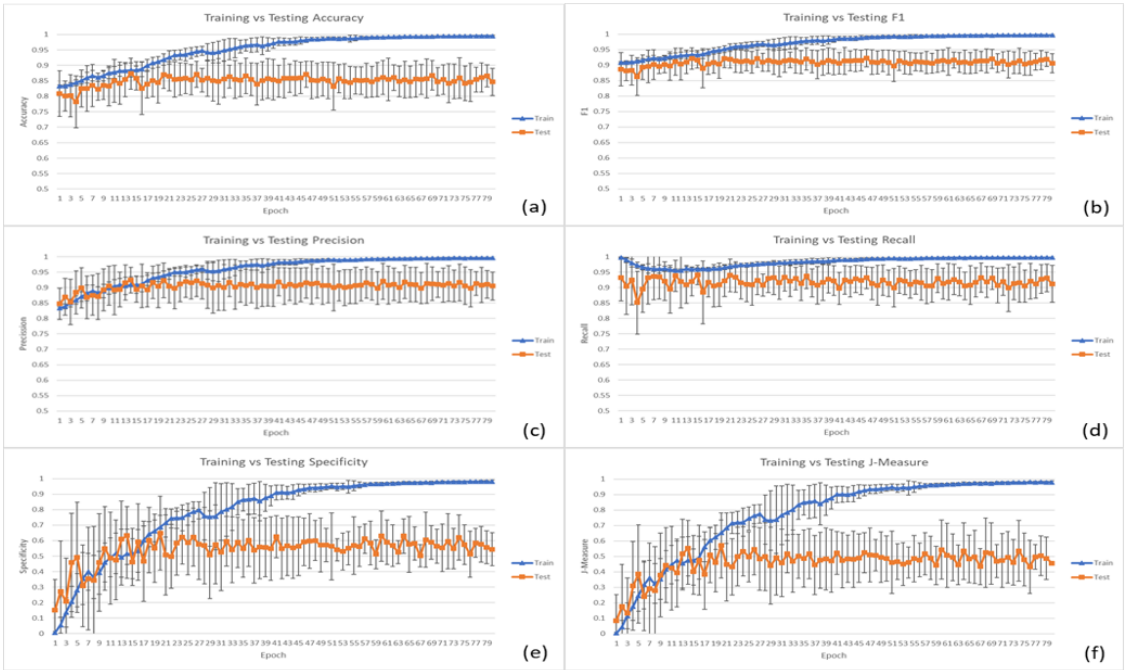

Figure 4.3: Evolution of performance metrics for the training and testing phase for the proposed LSTM model. (a) Accuracy, (b) F1 Score, (c) Precision, (d) Recall, (e) Specificity, (f) J-Measure

because the dataset itself is greatly unbalanced, being largely composed of myocardial infarction patients with significantly less control subjects. To improve performance more controls should be added to the dataset to increase data variability and hence allow the model to better generalize.

Table 4.1: Testing Results of the Proposed LSTM Method

	Acc	F1	Precission	Recall	Specificity	J	# Epochs
Acc	91.36 (±2.88%)	94.71 (±1.97%)	93.54 (±2.46%)	96.00 (±2.45%)	69.28 (±8.41%)	65.28 (±8.41%)	34.40 (±19.40%)
J	89.56 (±2.79%)	93.45 (±1.94%)	95.30 (±2.86%)	91.88 (±3.13%)	80.81 (±9.62%)	72.69 (±8.98%)	27.40 (±12.64%)

4.3.1 Influence of MI Location on Detection Rate

The location of the myocardial infarction has a great potential to influence the detection rate of the proposed classifier. As we choose to use a single ECG lead, the model proposed herein has a limited view of the heart's electrical activity and certain MI types could be obscured from such a view.

In practice, we can encounter various types of MIs which themselves affect the electrical activity of different sections of the heart. In the database used for this study, there are four main types of infarction that are described as follows:

*Anterior MI*s happen when the left anterior descending (LAD) coronary artery is obstructed, which could lead to changes in leads V1 through V6 but might or might not show in Lead II, depending on where the occlusion happens. *Antero-Lateral* MIs (a combination of Anterior and Lateral) show up in leads V3 through V6, while *Antero-Septal* MIs can

be seen in leads V1 through V4. All along, these types of Anterior MIs could produce changes in Lead II depending on the location of the LAD occlusion [66].

Inferior MIs, account for about 40% of all MIs and generally involve a blockage of the right coronary artery. When the Inferior region is the main location of the infarction, leads II, III, aVF, and aVL show changes in their respective electrical activity [67].

Lateral MIs are rare in their pure form; they generally occur as part of larger infarctions involving multiple areas, because the left ventricular lateral wall is perfused by the left anterior descending artery and the left circumflex artery. Leads I, III, aVL, aVF, V5, and V6 can all show electrical changes associated with Lateral MIs [68].

Posterior MIs present subtle changes in ECGs, therefore making them challenging to diagnose when they occur in isolation and often lead to misdiagnosis. They commonly take place in combination with Inferior and Infero-Lateral MIs. Purely posterior MIs can be hard to observe in ECGs and might require the addition of extra leads not present in the typical 12-lead ECGs (V7-V9) [69].

Figure 4.4 shows the detection rates of the proposed MI detector according to the location of the myocardial infarction as well as the availability of samples for each of the cases.

It is evident from this figure that the cases where the model underperforms (*Lateral* and *Posterior*) are ones where the electrical changes associated with them are hard to observe from the lead used in this study [68, 69]. On the other hand, we must also be cautious

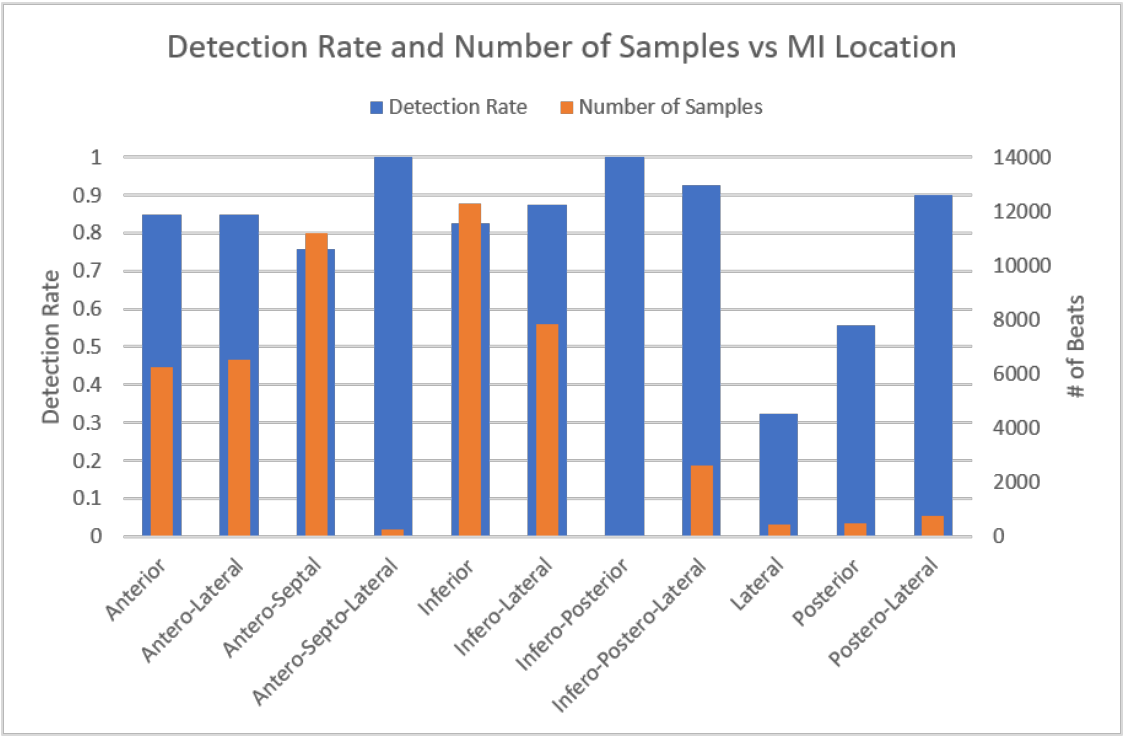

Figure 4.4: Detection Rate According to MI Location

of cases where the model performs exceedingly well, especially *Antero-Septo-Lateral* and *Infero-Posterior*, as the number of available samples for training and testing is minuscule.

4.3.2 Behavior when Encountering Other Pathologies

Although various pathologies are present in the PTB database, we only used the data available form patients diagnosed with myocardial infarctions and that from healthy controls for both training and testing. However, much like in real life, the proposed model will encounter samples from classes other than MI or HC. Therefore, we would benefit

from understanding or at least viewing how the classifier would respond when presented with samples from these other classes.

In this section, we run all samples from every patient from the PTB not used in this study through the already trained classifiers for each fold of training and report the average detection rate along with the respective standard deviation.

Figure 4.5 shows the distribution of the classifier's output for heartbeats from patients with a main diagnosis other than myocardial infarction or healthy control. In this figure the columns are associated with a given reason for admission and each of them has two colors; orange represents MI and blue represents HC. If a column is mostly composed of orange that means that most of the heartbeats associated with that diagnosis are classified as MI by the proposed model, while if mostly blue is present the classifier saw them as HCs.

In this figure, we can see the primary reasons for admission present in the PTB header files associated with each record. Going even deeper, we can further break down how within each reason for admission the classifier proposed herein treats each subject.

Bundle Branch Block: There are 20 independent records, belonging to 18 different patients, that list bundle branch block as a diagnosis. Of them, 17 (from 15 different patients), this condition was listed as the reason for the ECG and admission to the hospital and only 5 records (from 4 distinct patients) had more than fifty percent (50%) of its heartbeats classified as healthy controls.

Cardiomyopathy: There are 17 independent records (from 15 different patients) for which this condition was listed as the reason for the ECG and admission to the hospital.

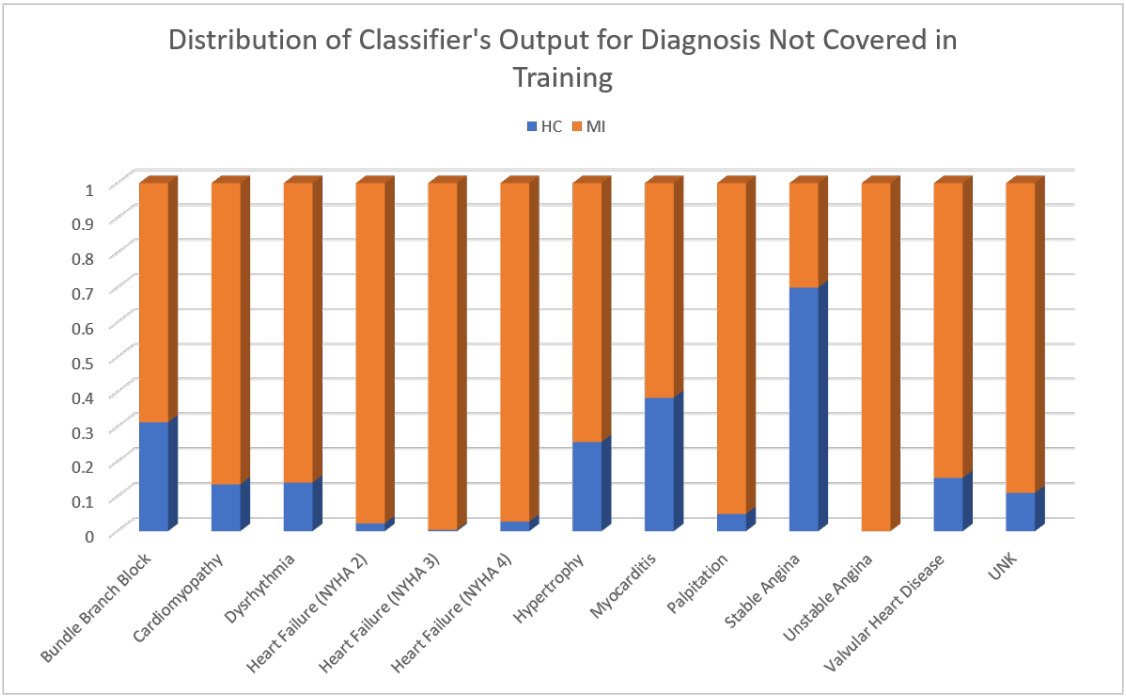

Figure 4.5: Distribution of Classifier's Output for Diagnosis Not Covered in Training

Of these only 2 records had more than fifty percent (50%) of its heartbeats classified as healthy controls and they each belonged to different patients.

Dysrhythmia: There are 16 different files from 14 different patients for which the main reason for admission was the condition of dysrhythmia. However, dysrhythmia condition rarely appeared alone and "Atrial Fibrillation" (AFib) was also evident in 8 (50%) of them where more than fifty percent (50%) of the heartbeats in each of those records were classified as belonging to the myocardial infarction class. "Coronary Artery Disease" (CAD) also accompanied dysrhythmia. Only one record from a single patient containing the main diagnosis of dysrhythmia had more than fifty percent (50%) of its heartbeats classified as healthy controls.

Heart Failure (Types 2,3, and 4): There are very limited number of records exhibiting heart failure; only one of each type from separate subjects exist in the PTB. Each of this records had a very high percentage (> 96%) of its constituent heartbeats classified as MI.

Hypertrophy: is present as the main reason for admission in seven (7) ECGs from seven patients. Hypertrophy was never diagnosed alone and only one patient's record (being additionally diagnosed with hypertension) had 93% of its heartbeats classified as healthy. For all other records, at least 61% of the heartbeats where considered infarcted.

Myocarditis: is observed in four (4) records form four (4) distinct patients. Only one of these records had most (> 94%) of its heartbeats classified as healthy. The record with most beats classified as MI (> 84%) has an additional diagnosis of "Bundle Branch Block".

Palpitation: are only present in one single record where over 95% of the heartbeats are classified as infarcted. This record also had an additional diagnosis of Coronary Heart Disease.

Stable Angina: is the main reason for the admission of two separate patients (for a total of two ECGs). One of them only has additional cold and Hyperlipoproteinemia as a diagnosis and received a healthy classification for over 90% of its heartbeats. The other has CAD and hypertension, resulting in over 95% of its heartbeats being classified as MI.

Unstable Angina: is only present in one record and 100% of the beats are classified as MI.

Valvular heart disease: is listed as the main reason for the admission of six (6) patients, for a total of six (6) ECGs. Every one of these records has multiple associated diagnoses and over 75% of the heartbeats in each record were classified as MI.

Atrial Fibrillation (AFib): is never listed as the main reason for the admission of any patient but is listed as an associated diagnosis in fourteen (14) separate records from eleven (11) patients. Every time AFib is listed as a diagnosis for a given record, over 60% of the heartbeats therein were classified as MI.

Unknown (UNK): The reason for the admission of 22 patients is listed as unknown. From these, 27 separate records were recorded. Most of them have over 90% of their heartbeats classified as MIs, about 22 records. Of the remaining five (5), four (4) have over 50% of the heartbeats classified as MI and only one has over 90% of them classified as healthy.

It seems to be evident from these broad statistics that the model proposed in this study, although only trained to differentiate between myocardial infarctions and healthy controls, appears to be detecting life threatening heart ailments. This possibility would have to be further explored in future studies that encompass broader datasets, as no conclusive finding can be done over these limited samples.

4.3.3 Visualizing the Errors

Although high classification metrics were achieved using the proposed approach, it is important to understand where the proposed model falls short and does not achieve

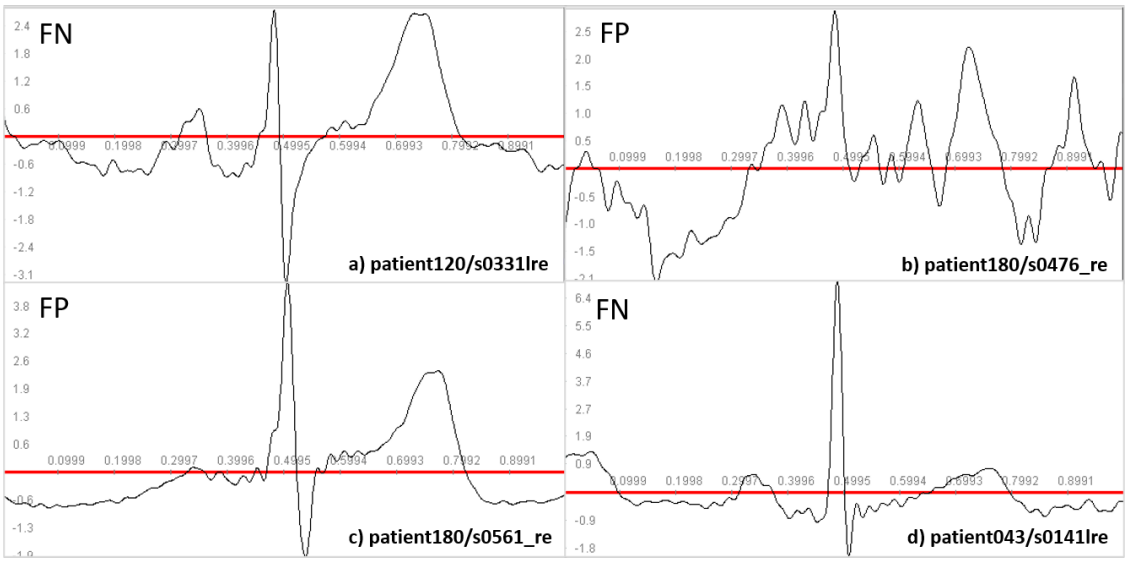

Figure 4.6: Examples of Missclassified Heartbeats

the best performance. To start, figure 4.6 depicts interestingly different heartbeats from specific records were the classification accuracy was particularly low (<10%).

The heartbeat present in Figure 4.6a belongs to a patient originally admitted for an *Anterior MI* but the ECG was recorded over a year after the infarction took place, to be more specific 396 days after. This heartbeat was classified by the model as "Healthy Control" although the truth value from the diagnosis was "Myocardial Infarction" making it a false negative. Only eleven (11) out of the one-hundred and forty-two (142) heartbeats for patient 120 were correctly labeled as MI instead of HC.

Figure 4.6b shows an example of improper lead placement or contact that yielded a very noisy ECG recording. This heartbeat belongs to patient 180, a healthy control subject, and is part of record s0476_re. In this record, there is a vast amount of noise present in Lead II, while leads 1 through 6 are fairly clean. However, as this classifier only uses the information from Lead II it is hindered by the noise and is challenged to properly

classify the detected heartbeats. Only one (1) out of the two-hundred and forty-one (241) detected heartbeats is properly classified as HC, while all others are improperly labeled as MI.

A truly interesting false positive case is present in Figure 4.6c. This heartbeat also belongs to patient 180, but this time the record is s0561_re and there is no significant noise to contend with. In this instance, the classifier was unable to properly classify around 63% of all detected heartbeats, assigning them an MI label instead of the correct HC.

Figure 4.6d shows another instance of a false negative outcome. This heartbeat belongs to records s0141lr of patient 43, a subject with a lateral infarction. In this instance there are a couple of factors that made the classification difficult: 1) the electrocardiogram was recorded over eight (8) months after the infarction took place and a catheterization was performed, and 2) the type of infarction in question was a "Lateral MI", which are difficult to diagnose from Lead II alone as electrical activity changes associated with it might not be revealed in such a lead [68].

4.4 Discussion

As the type of data split used for training is seen to greatly affect the performance metrics, we must restrict our comparative assessment contrasting the proposed method to related studies that use the same data split method in order to avoid unfair comparisons. Some relevant studies covered in the literature are presented in Table 4.2. Herein, we pro-

vide a simplified evaluation of their dataset generation techniques along with the number of leads used and the sample size required to make a classification.

Table 4.2: Simplified Analysis of Relevant Studies

Study	# Leads	Sample Length	Beat Split	File Split	Patient Split
[6]	1	0.65s	x		
[7]	1	RR interval	x		
[8]	3	3s	x		x
[9]	12	1s	x		x
[10]	12	10s	x		
[12]	1	2 beats	x		
[13]	4	Whole record			x
[14]	1	0.65s			x
[15]	12	0.65s	x	x	
[16]	12	~7s		x	x
[17]	12	4s		x	

Given the unique way of assuming only a single heart beat and single-lead electrocardiograms together with the adoption of patient split method of analyzing data, to the best of our knowledge, there are no other studies of myocardial infarction detection and classification that can be completely and fairly compared to ours, as types of MI classified vary and the datasets and dataset subsets are not equal to ours. However, in Table 4.3 we provide an in depth comparison of some of the available studies that generate inter-patient detection of MI. In this table, we cover and provide the numbers of leads used, number of patients and heartbeats evaluate, the sample length required, the method used, and the general performance metrics of each study.

Of all the studies presented herein, [8] and [14] are somewhat close, although much more restrictive (they use subsets of the available data), as they use electrocardiograms available from the PTB diagnostic ECG Database [60, 61].

Sharma and Sunkaria [8] focused on acute myocardial infarctions in the inferior portion of the heart (IMI) leading them to a rather restricted dataset of only 30 MI subjects and 52 HC. Moreover, they use three leads of the available twelve, specifically leads II, III and aVF, as better representatives of electrical activity in the inferior portion of the heart. Sopic et al. [14] use random forest feed by expert features extracted from Lead 5 of the available electrocardiograms using discrete wavelet transforms. They use 52 MI subjects as they want to keep a balanced set for training and testing due to the limited availability of healthy controls (52). Just as in [8], this study is not directly comparable to ours as they use a different ECG lead and a more restrictive, yet balanced dataset.

Although our results are not directly comparable, we achieve the highest accuracy and recall, when comparing against either [8] or [14], for either of the two optimization cases (best Accuracy or best balance between Specificity and Recall). However, our specificity and J-Measure oscillate, providing the best number only when we train our network to achieve the best balance between Specificity and Recall. The lower specificity is likely due to the fact that we have a heavily unbalanced dataset, as there are about three times more myocardial infarction subjects than healthy controls in our datasets, and a higher accuracy will tend to lead the network to prioritize properly classifying MI records over HCs. We overcome this in this study, by optimizing, or targeting the point in time at which the network is having the best balance between specificity and recall instead of using the common practice of just targeting the best accuracy. We also only require one lead of the electrocardiogram as does [14] but unlike them our lead is more commonly

sampled and simpler to record that lead V5. Neither [8] nor [14] provides the diagnosis time (or processing time), but study [14] provides a hardware implementation which is commendable and could be used as means to reduce the processing time.

From Table 4.3 we can also see that the detection accuracy is positively correlated with the number of leads used, that is, more leads equals greater accuracy. This is to be expected, as more electrocardiogram's leads equates to more views of the heart and more information is gathered. The extra information collected from other leads is especially useful to detect infarctions that are not visible or hard to see from a single lead, as it is our case [66, 68, 69]. By using 12-lead ECGs, [9] and [16] are able to achieve higher performance metrics than any other single lead method.

Longer sample lengths seem to also influence the performance of the different models, but the relevance of this pattern is not very clear. In the case of [13], by using the whole record to produce a classification, the authors achieve higher performance metrics than those reported by [16] who used more leads (12 in [16] vs 4 in [13]). However, [16] is detecting MI from ECG images, not digital signals like all other studies in Table 4.3 are, and this could explain the difference in performance. The authors of [13] also tailored their lead selection to the type of MI they were trying to detect (Generalized Anterior MI) which makes it less clear whether or not the sample length is the key factor yielding the performance improvement. Therefore, it is hard to gauge the influence of sample length in classifier performance due to the fact that such vary from study to study along with classification method, number of leads, and even dataset used. To the best of our knowledge there is no study that explores the influence of sample length on MI detection rate.

One more significant factor to point out is that we are not able to provide a statistical comparison of our method to other present in the literature or those presented in this section, as they only provide the average performance metrics of their validation approaches and do not provide the standard deviations of such.

Overall, the results obtained are highly competitive in comparison to state-of-the-art algorithms, although stringent conditions are set up in the training phase to overcome overfitting, data leakage and bias from features seen in the training phase. The proposed method would have performed even better if a balance between the normal controls and MI subjects could have been reached through the collection of more data from normal controls.

Table 4.3: Comparative Results of Methods Using Patient Split Method

Study	Leads Used	#Patients	# Beats	Sample Length	Method	Acc	Recall	Specificity	J-Measure
[8]	II, III, aVF	30MI, 52HC	3240 MI, 3037 HC	3s	SWT & SVM	81.71	79.01	79.26	58.27
[9]	12	128MI, 52HC	48690MI, 10646HC	1s	MFB-CNN	98.79	98.73	99.35	98.08
[13]	aVL,V2, V3,V5	Records: 167MI, 80HC	Not Specified	Whole Record	ML-CNN	96.00	95.40	97.37	92.77
[14]	V5	52MI, 52HC	Not Specified	0.65s	Random Forest	83.26	87.95	78.82	66.77
[16]	12	Images: 483MI, 474HC	Not Specified	~7s	MBFN-CNN	94.73	96.41	95.94	92.35
Proposed (Accuracy Optimization)	II	148MI, 52HC	50732MI, 10123HC	1s	LSTM	91.36 (±2.88%)	96.00 (±2.45%)	69.28 (±8.41%)	65.28 (±8.41%)
Proposed (J-Measure Optimization)	II	148MI, 52HC	50732MI, 10123HC	1s	LSTM	89.56 (±2.79%)	91.88 (±3.13%)	80.81 (±9.62%)	72.69 (±8.98%)

CHAPTER 5

12-LEAD REAL-TIME MYOCARDIAL INFARCTION DETECTION

5.1 Goal

In this study, we propose a novel and complete pipeline for the online real-time classification (40 milliseconds) of myocardial infarctions. We do so by using the Independent Component Analysis (ICA) R-spike detection method described in Chapter 2, to identify and localize the occurrence of ventricular depolarization events, alongside a multilayer LSTM network, to detect and classify infarcted heartbeats. We particularly use LSTM neurons for their long track record of positive results at classifying time-varying signals such as speech, text, and video. The resulting real-time classifier could be of great value to the population at risk of MI and as a monitoring tool for gauging disease progression and evaluating merits of treatment or therapeutic intervention.

5.2 Material and Methods

5.2.1 Data and Preprocessing

We use the popular Physikalisch-Technische Bundesanstalt (PTB) diagnostic EKG database [60, 61]. This dataset is composed of 549 12-lead EKG records from 209 men and 81 women. There are multiple different diagnoses in the dataset, but for our study we only use patients with myocardial infarctions (MI) and the healthy controls (HC), this brings the patient count to 200 patients from the original 290. Furthermore, we only use records from those patients that were recorded no more than 5 days after the infarction

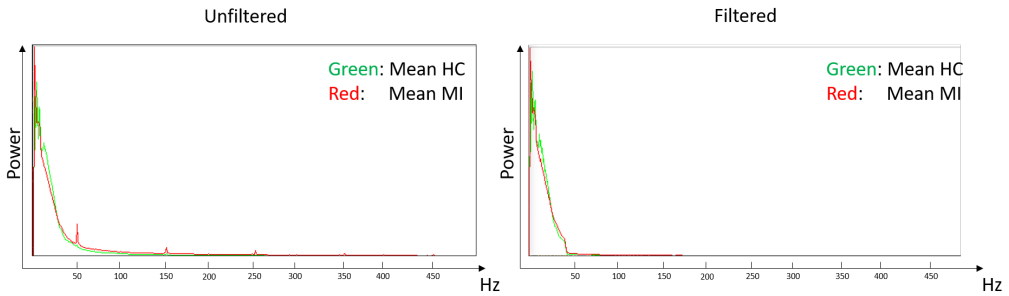

Figure 5.1: Frequency Power Spectrum of unfiltered (left) and filtered (right) data

date, except for the first electrocardiogram taken at admission, as after admission to the hospital, patients are given treatment and the hearts electrical activity responds to it.

The data is filtered using a 50Hz band-stop cascade IIR filter to remove any powerline interference that might be left over after the initial recording, and a 500 milliseconds moving average filter to remove baseline wander as an inherent EKG artifact. Frequency spectrums of the signal are shown in Figure 5.1 before and after the filtering step. After filtering, we apply the real-time R-peak detection algorithm described in [70] to produce near instantaneous ventricular depolarization detection. Heartbeats are segmented into 1-second samples, centered at the R-peaks, and separated into training and testing datasets. We have followed the Patient-Split dataset generation procedures as the performance metrics obtained from it on the testing dataset will most closely resemble those of the systems deployment in real-unseen data. That is, we have taken special care to ensure that any data from patients seen during training is not used for testing. A segmented sample of a processed heartbeat can be seen in Figure 5.2

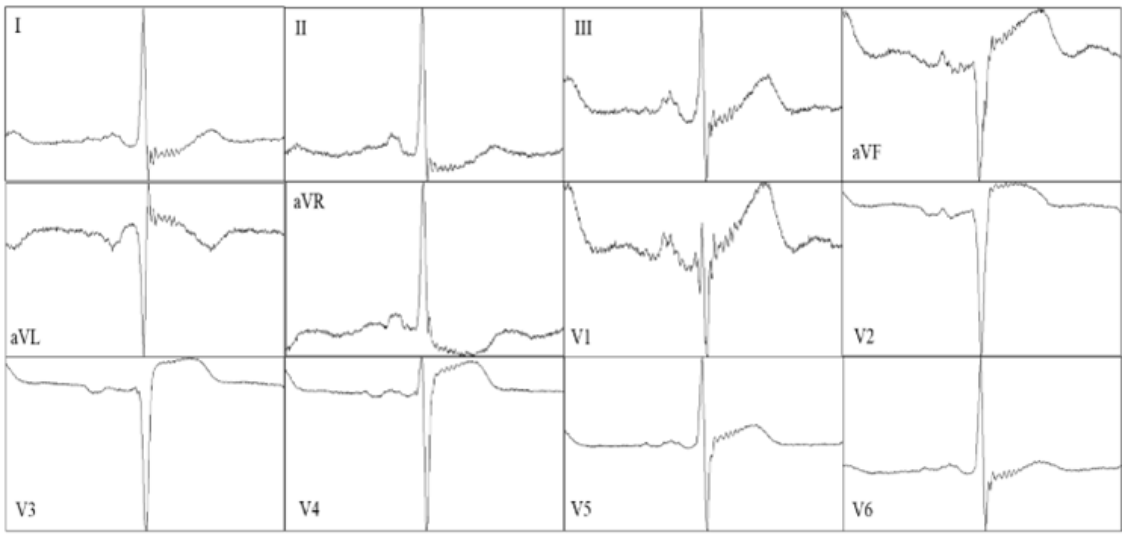

Figure 5.2: Sample Filtered and Centered Heartbeat

5.2.2 Methods

The simplified system architecture of the proposed system for online and real-time diagnosis of myocardial infarctions is shown in Figure 5.3. The first step is to filter the incoming data to remove undesired frequency components and noise. Subsequently, the ICA R-Spike detector is used to identify the center of the ventricular depolarization event and to use it as a reference for segmenting and centering samples to be passed to the MI classifier.

The proposed architecture continuously generates classification outputs for each detected heartbeat. The filter and the R-Peak detection stage run non-stop seeking out new samples to be processed. While the classifier produces a classification label (MI/HC) for every point in the input sample, only the last 80ms are used. This is because the classification does not become valid until the classifier has had a chance to look at the complete heartbeat (after the T wave).

For our classifier we use a deep LSTM neural network with five layers. The size of the network is chosen so that the network is big enough to accommodate a significant number of features and dropout is used during training on every other layer to minimize overfitting. A more detailed view of the network architecture can be seen in Figure 5.4. This neural network architecture is particularly suited for this problem as it can handle time-varying data, does not require prior expert knowledge of the signal at hand, and is simpler than other previously proposed architectures. Furthermore, its simplicity makes it less prone to overfitting and yields low inference times.

The proposed model is built with Graves LSTM [65] units and initialized using Glorot and Bengios proposed initialization [64]. It uses stochastic gradient decent with L2 regularization [62], RMSProp [63] as the optimizer, and gradient clipping to avoid exploding gradients due to excessive weight updates. A more detailed view of the networks training behavior and the influence of dropout can be appreciated in Figure 5.5, where the first and third layer implement different dropout rates. By using dropout during training, we force the network to prioritize and create redundant copies of important latent features in large layers. We do so by randomly zeroing out, or dropping, the outputs of 80% of the neurons in the first layer and 50% of the neurons in the third one, thereby effectively reducing the amount of data available to the subsequent layers.

5.3 Results

The network is trained using early stopping, where after 10 epochs of no performance improvements we stop training and back up to the best set weights. A 10-fold cross validation approach is performed to avoid reporting on a particularly beneficial or detrimental

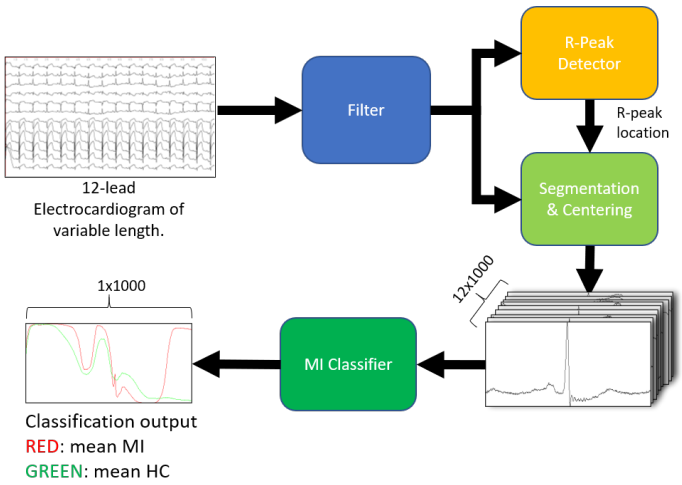

Figure 5.3: Simplified System Architecture

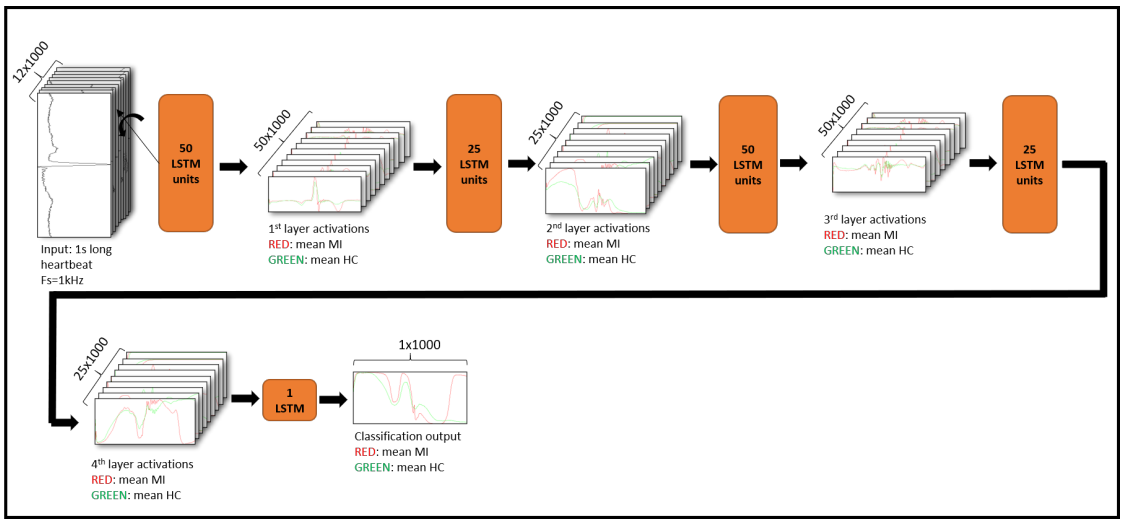

Figure 5.4: 12-Lead Classifier Architecture

dataset split due to their small sizes. We train the model in a 64-bit Windows 10 PC with an AMD FX-8350 Eight-Core Processor, 32 GB of DDR3 RAM, and an NVIDIA GeForce GTX1070 graphics card. The system proposed herein was implemented and deployed using Java and DeepLearning4J version 0.9.1 as the machine learning library.

Performance was measured by using the aforementioned standard metrics of Accuracy, F1-Score, Precision, Recall, Specificity, and Youdens J statistic (J-Score/J-Measure) as defined earlier in (4.1) through (4.6) in Chapter 4. Recall that true positives (TP) define patients with MI diagnosed as MI, true negatives (TN) define healthy controls (HC) diagnosed as HC, false positives (FP) define HC diagnosed as MI, and false negatives (FN) define MI diagnosed as HC. We trained our system seeking two important measures and report these values in Table 5.1: highest classification accuracy and highest J-Score.

The forward inference time, the time required for the network to process a sample, is 12.3 milliseconds on average, while the whole system requires around 40 milliseconds to process a sample when accounting for preprocessing time. Training took on average 5 hours and covered around 30 thousand 1-second samples per epoch (a single run through all training instances). From Table 5.1, the maximum number of training epochs was 50 for the slowest converging training fold and 3 for the fastest, with an average of 19.44 epochs when looking for the maximum accuracy and 21.33 when looking for the best

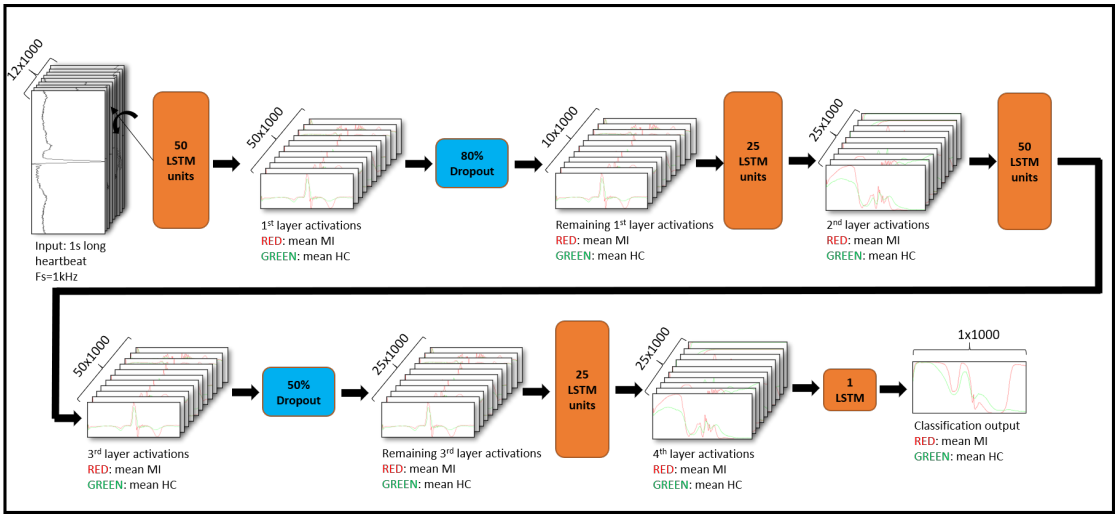

Figure 5.5: 12-Lead Classifier Architecture during Training

balance between specificity and sensitivity. The average accuracy across the 10-fold cross validation is 95.76% with a 95% confidence interval (CI) from 93.36% to 98.16% and the average J-Score is 90.31% with a 95% CI from 84.11% to 96.51%.

Table 5.1: Performance Results of Two Training Scenarios

	Acc	F1	Precission	Recall	Specificity	J	# Epochs
Acc	95.76 (±2.4%)	96.73 (±2.1%)	96.86 (±2.7%)	96.67 (±2.4%)	93.64 (±5.7%)	90.31 (±6.2%)	19.44 (±18.9%)
J	95.66 (±2.5%)	96.64 (±2.2%)	97.13 (±2.8%)	96.27 (±2.6%)	94.38 (±5.5%)	90.64 (±6.0%)	21.33 (±17.5%)

5.4 Comparison and Discussion

Although the literature on myocardial infarction detection and classification is abound, the way in which they create their training and testing datasets are not uniformly consistent, resulting in intra- and inter- patient classifiers that are not directly comparable. Therefore, in this section we will attempt to compare our results to those from other methods that most closely resemble the data splitting method we use, that is inter-patient classifiers. We will also explain why we believe that the proposed method is not directly comparable to some of the references we consider in this section.

As can be appreciated from Table 5.2, our proposed method has some of the highest metrics in comparison to other approaches. These are methods that use the same number of Electrocardiogram leads (12) and are trained on the same database (although not specifically the same dataset). However, not all the methods covered in this table are directly comparable as detailed below.

Table 5.2: Comparative Results of the Proposed Methods

Study	Dataset	Sample Length	Method	Accuracy	Recall	Specificity	J
[71]	128MI, 52HC	0.6 seconds	MFB-CNN	98.79	98.73	99.35	98.08
[72]	148MI, 52HC	0.6 seconds	CNN and BLSTM	93.08	94.42	86.29	80.71
[17]	113MI, 52HC	4 seconds	ML-ResNet	95.49	94.85	97.37	92.22
[73]	148MI, 52HC	4 seconds	PCA and SVM	92.69	80.96	80.96	61.92
[74]	148MI, 52HC	5 seconds	DWT-PCA-ANN	98.21	99.40	98.22	97.62
[74]	148MI, 52HC	5 seconds	Deep Residual CNN	100	100	100	100
Proposed	148MI, 52HC	1 second	Deep-LSTM	95.66	96.27	94.38	90.64

Liu et al., in [71], introduce a Multiple-feature-branch CNN classifier, that produces a particularly good inter-patient MI classifier, but they use data from the patients in the testing set during training (the first 32 beats of each patient), effectively contaminating it and no longer producing a pure inter-patient classifier and rather something that is somewhat in between intra- and inter- patient.

In [72], Liu et al. expand on their previous work on multiple-feature-branch CNN classifiers and add bidirectional LSTMs, producing an MFB-CBRNN. This time around, they avoided contamination of their training sets and did produce a truly inter-patient MI classifier. Their approach classifies single heartbeats, just like ours, but produces less accuracy, 93.08% compared to 95.66% for ours, and lower J-Score, 80.71% vs. ours at 90.64%.

Han and Shi produced two good classifiers in [17] and [73], however, even though they claim to produce inter-patient classifiers, it is unclear from their dataset generation descriptions that they took special care to ensure that data from patients seen during training was not used for testing. It appears that their dataset generation method resembles that of a File-Split instead. They also require particularly long sample sizes that encompass more than a single heartbeat.

Perhaps the best classifier in terms of accuracy is reported by Jafarian et al. [74], achieving 100% accuracy in their end-to-end deep neural network model, but their method requires five (5) second samples and is perhaps resource intensive to deploy. Nevertheless, they do take particularly good care to avoid cross-contamination of their training and testing sets, yielding truly inter-patient classifiers.

CHAPTER 6

REAL-TIME FREQUENCY-INDEPENDENT SINGLE-LEAD AND SINGLE-BEAT MYOCARDIAL INFARCTION DETECTION

6.1 Goal

This chapter presents a single-lead frequency independent myocardial infarction detection algorithm. It expands on the architecture and knowledge of previous Chapters to produce a readily deployable system able to detect myocardial infarctions from single-lead electrocardiogram. Wherein, perhaps the most relevant feature that it attempts to achieve is the frequency independence of the algorithm, as electrocardiograms come in no standard sampling frequencies and therefore vary form device to device. This inter-device variation will force frequency dependent classifiers to be retrained with signals collected at the specified sample frequency, thereby undermining deployability and increasing cost and difficulty. This chapter will attempt to overcome these drawbacks and provide readiness for deployment, cost effectiveness and increased usability.

6.2 Material and Methods

6.2.1 Data

This study makes use of the newly released PTB-XL database, the largest currently publicly available electrocardiography dataset [61, 75, 76]. It contains 10-second long 12-lead electrocardiograms from 18885 distinct patients, with a total of 21837 combined records. Of these, there are 9528 healthy control (HC) records and 5486 myocardial infarction (MI) records. The records themselves are stored in WaveForm DataBase (WFDB)

format with a resolution of 1uV/LSB and are provided in two different files with different sampling frequencies of 500 and 100 Hz, where the 100 Hz files are downsampled from the 500 Hz ones.

Furthermore, the PTB-XL database is broken down by its publishers into 10 non-patient overlapping sets to be used for algorithm training, validation, and test. These sets are obtained using stratified sampling and ensure that all records associated with the same patient are only present in one of the sets. They recommend that sets 1-8 be used for training, set 9 for validation, and set 10 for testing. We followed this recommendation to ensure that future algorithms can be fairly compared to ours. The PTB DB [60], used in previous studies, is also used to validate the trained classifier, as it is composed of different patients and is sampled at a different frequency. In both these datasets we are only interested in using Lead II, as it is widely employed by wearable/portable devices and exercise equipment, and therefore start by removing all other lead information. We also only keep records with diagnosis of acute myocardial infarction, old myocardial infarction, and/or normal control.

A low-pass filter with a cutoff frequency of 45Hz and a 500-millisecond moving average filter are applied to each record to remove powerline and low-frequency noise, respectively. Special care must be taken to ensure that the powerline interference filter (either low pass or stop-band) removes 50Hz frequencies, as these datasets were collected in Europe, where 50Hz power lines are standard, and not in America where 60Hz power distribution is the case. However, this would not generally be a problem if the noise were present in both MI and HC records, but Figure 6.1 makes it clear that this 50Hz spike is only present in MI patient and not in HC ones in the PTB-DB, thereby making it a clear distinguishing feature between the two. Therefore, if a 60Hz stop-band filter were to be applied, it is possible that the deep-learning classifier built in this design could latch onto

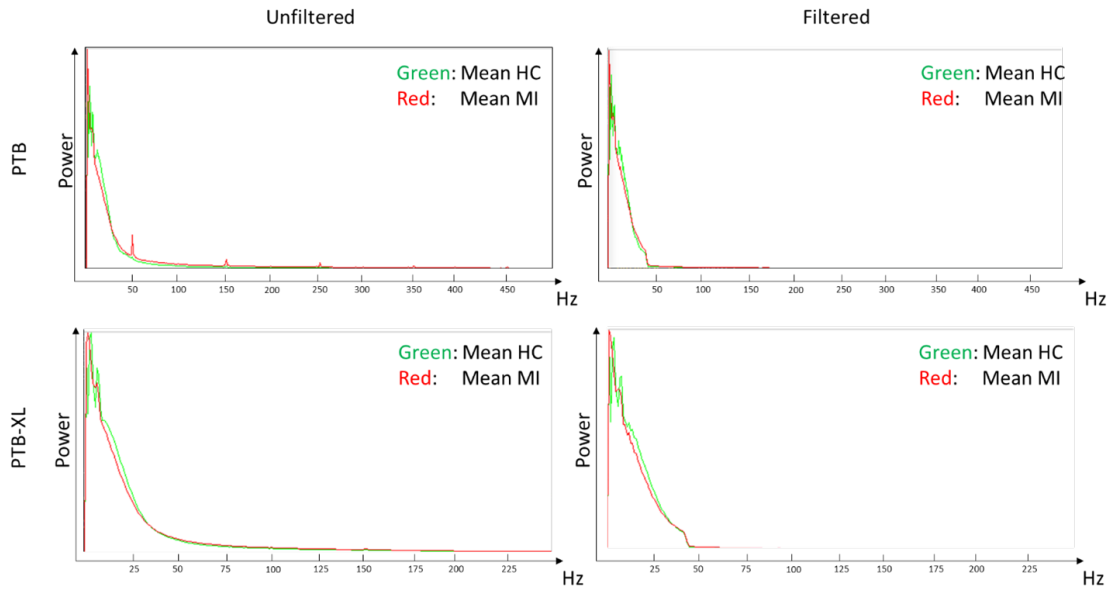

Figure 6.1: Frequency Spectrum of pre- and post- filtered data for PTB and PTB-XL datasets

the characteristic 50Hz spike present mostly in the MI cases to incorrectly provide a diagnosis based on the frequency content of the signal rather than on the disease biomarkers.

Once the data is filtered, a design variation of the Pam and Tompkins algorithm described by Chen and Chen in [53] is applied to identify the locations of the ventricular depolarization events (also known as R peaks), and to separate thereafter each individual heartbeat to be used in the training and testing of the proposed multilayer LSTM network. Each sample is one second long (500 milliseconds before and 500 milliseconds after the ventricular depolarization event) of the EKGs Lead II data (information from all other leads is discarded).

6.2.2 Network Architecture

The complete pipeline used to process the electrocardiograms is presented in Figure 6.2. The first step is to filter the incoming data to remove undesired frequency components and noise. Subsequently, the R-Spike detector from [53] is used to identify the center of the ventricular depolarization events and to use them as references for segmenting and centering of the samples to be passed to the MI/HC classifier.

The proposed architecture continuously generates classification outputs for each detected heartbeat. The filter and the R-Peak detection stage run non-stop seeking out new samples to be processed. While the classifier produces a classification label (MI/HC) for every point in the input sample, only the last 80ms are used. The classification does not become valid until the classifier has had a chance to look at the complete heartbeat (after the T wave).

The classifier itself is presented in Figure 6.3. It is four (4) layers deep and consist solely of LSTM units. The first, second, and third layers are made up of 50, 25, and 10 units respectively with hyperbolic tangent activation functions, while the last layer has a single neuron with a sigmoid activation function.

LSTM neurons can produce a classification prediction for every input timestep thereby generating time-varying classifications that ultimately become more accurate and final as more sample points of the signal are processed. This is especially evident in the output activation of the network in Figure 6.3, where the final classification label is produced during the last 200ms of the signal once the PQRST complex has been processed.

Similar to [6, 13, 16, 17, 77–79] and others, deep neural networks are used as they require no expert knowledge of the signal or the data being processed and are able to interpret which underlying features are important on their own from the training data.

The model is trained using stochastic gradient descent and implements L2 regularization [43] with $\lambda = 0.001$, RMSProp updater [44] (learning rate = 0.1, weight decay = 0.95, and epsilon = 10e-6) with weight decay for parameter updates, and gradient clipping (5.0). Herein, the weights are initialized using the method proposed by Xavier Glorot [45] and the specific LSTM units used were first described by Alex Graves in [46] to label sequential data.

The network was trained and deployed in a 64-bit Windows 10 PC with an AMD FX-8350 Eight-Core Processor, 32 GB of DDR3 RAM, and an NVIDIA GeForce GTX1070. Each training iteration (epoch) took approximately 3.5 days and covered around 435K samples. The network architecture itself was implemented using DeepLearning4J version 0.9.1, a deep learning library for Java.

6.2.3 Multifrequency Scaling

To make the classifier frequency independent we expose it to signals of different sampling frequencies at training time. To do so, we down- and up- sample the electrocardiograms in the continuous range between 250Hz and 1000Hz, given that the training dataset is sampled at a set frequency of 500Hz and only a downsampled version of it is provided at 100Hz.

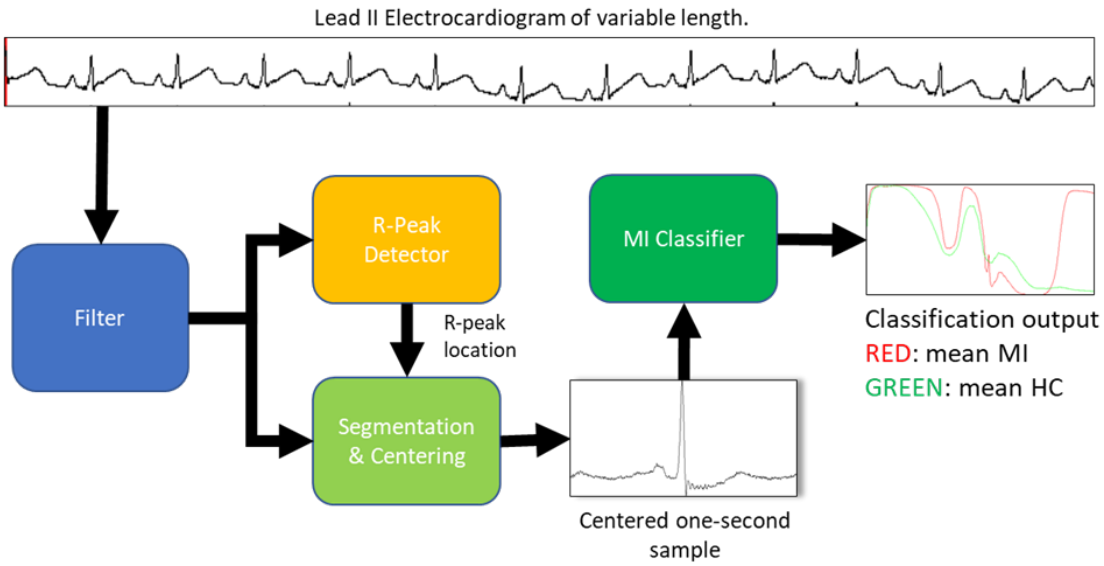

Figure 6.2: End-to-End MI Detection Pipeline

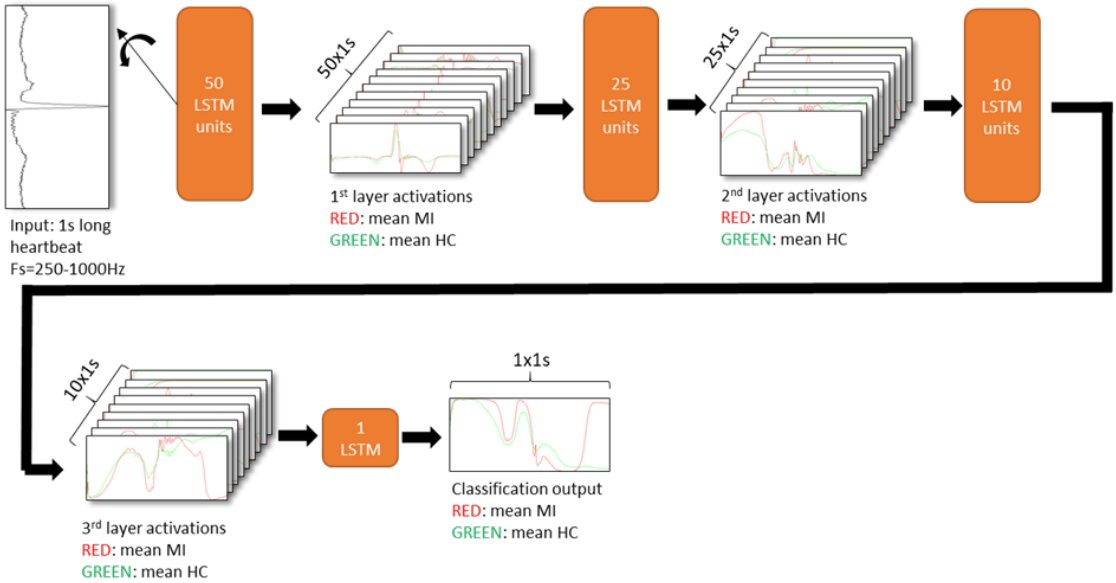

Figure 6.3: MI/HC Classifier Architecture

However, there are infinitely many frequencies between 250Hz and 1000Hz and we cannot reasonably generate and make available all possible frequencies for every single epoch. To overcome this hurdle, we randomly pick ten (10) different frequencies in this

range for every file at every epoch. Thereby generating a ten-fold increase in the available data and covering a large percentage of the frequency spectrum, which will further increase with successive epochs.

The up- and down- sampling is achieved using Lanczos resampling [80] with a 20-point kernel as described by equation 6.1.

$$S(x) = \sum_{i=floor(x)-a+1}^{floor(x)+a} s_i L(x-i)$$

$$L(x) = \begin{cases} 1 & if\ x = 0 \\ \dfrac{a * sin(\pi x) \sin\left(\frac{\pi x}{a}\right)}{\pi^2 x^2} & if\ -a \leq x < a\ and\ x \neq 0 \\ 0 & otherwise \end{cases} \tag{6.1}$$

Where a is the kernel size, S is the re-sampled signal, and s is the original one.

By using these equations, we are able to re-sample all the available EKGs to any of the frequencies of interest. This step is only critical for training, and not required for deployment or real-time application, as we need to expose the classifier during training to the wide range of possible frequencies that could be encountered when deployed.

6.2.4 Dataset Balancing

Both the PTB and the PTB-XL datasets are highly unbalanced, albeit in different directions. The PTB dataset is MI heavy, while the PTB-XL dataset is HC dominant. Left unbalanced, this has the effect of causing the classifier to lean, or have higher classification

metrics, in the direction of the subset with a larger number of subjects. This is evident in [77], where the proposed model showed increased recall at the expense of specificity, as there are nearly three times as many MI subjects as there are HCs in the dataset used.

Therefore, to address the lack of balance, we put forth a subexperiment where we will test whether balancing the dataset has any effect on the classifiers performance. In this particular case, because we will be using the PTB-XL, the unbalanced scenario will use all available training samples and will have a ratio of around 1.7:1 in favor of healthy controls.

The dataset rebalancing takes effect every epoch. This will ensure that no training sample is wasted, and that enough variability remains, while at the same time keeping each of the minibatches balanced.

To balance the datasets, we start by identifying which subgroup (either HC or MI) has the least training samples and use it as a starting point for the training set. Once a set has been identified, we randomly subsample the same number of instances from the other and merge them into the set of samples to be used for the current epoch. The pseudo code used to achieve the dataset balancing can be seen in Figure 6.4.

6.3 Results and Discussion

The network is trained using early stopping, for which after 10 epochs of no performance improvement we stop training and back up to the best set of weights. As recommended by the PTB-XL publishers, we train the model on sets 1 through 8, use set 9 for validation, and test on fold 10. We also test the performance of the model on the entire

```
trainingSamples.clear()
if (balanced) {
        shuffle(posSamples)
        shuffle(negSamples)
        if (posSamples.size > negSamples.size) {
                nSamples = negSamples.size * 2
        } else {
                nSamples = posSamples.size * 2
        }
        samples.addAll(posSamples.subList(0, nSamples/2))
        samples.addAll(negSamples.subList(0, nSamples/2))
} else {
        nSamples = posSamples.size + negSamples.size
        samples.addAll(posSamples)
        samples.addAll(negSamples)
}
shuffle(samples)
```

Figure 6.4: Pseudo Code for Dataset Balancing

PTBDB, as the sampling frequency is different from that of the training set and different set of patients are acquired under different conditions.

Training took 35K parameter updates, or 30.6 days, and the best performance was obtained at 29K parameter updates for both accuracy and J-Measure. Processing of a single 1-second sample is achieved in 300 milliseconds, on average by the classifier at a sampling frequency of 500Hz, and there is negligible difference over the entire pipeline. This fast processing time would be appropriate for online classification as the time between fast heartbeats is around 300 milliseconds. However, this processing time is sampling frequency dependent, where it ranges from 100 ms for 202Hz to 1.2 seconds for 2.8kHz. Therefore, for sampling frequencies larger than 500Hz, we recommend an extra downsam-

pling step be added to the front of the pipeline, bringing the sampling frequency below 500Hz, in order to attain real-time processing.

Table 6.1 shows the final testing metrics in the different datasets, while Figure 6.5 depict the evolution of the validation metrics during training. Figure 6.6 shows the Receiver Operating Characteristics (ROC) curve of the final model over the different datasets.

To measure performance, we have used Accuracy (4.1), F1-Score (4.2), Precision (4.3), Recall (4.4), Specificity (4.5), and Youdens J statistic (J-Measure) as defined in (4.6).

Table 6.1: Testing and Validation Results for the Proposed Algorithm.

Dataset (fold)	Accuracy	F1	Precision	Recall	Specificity	J	AOC
Validation (9)	85.07%	80.02%	78.52%	81.54%	87.31%	68.66%	0.918
Test (10)	84.17%	78.45%	78.54%	78.37%	87.55%	65.92%	0.903
PTB	77.12%	84.5%	95.38%	75.85%	83.02%	58.87%	0.880
Combined	79.69%	83.42%	91.59%	76.59%	85.89%	62.48%	0.893

6.3.1 Balanced vs Unbalanced Training

As both datasets used in the study suffer from data imbalance, we have performed an experiment to determine the impact of such, if any, on training performance and speed. Figure 6.7 shows the evolution of the validation metrics during training using an unbalanced set while Figure 6.8 visually compares the progression of the validation metrics for both balanced and unbalanced training sets.

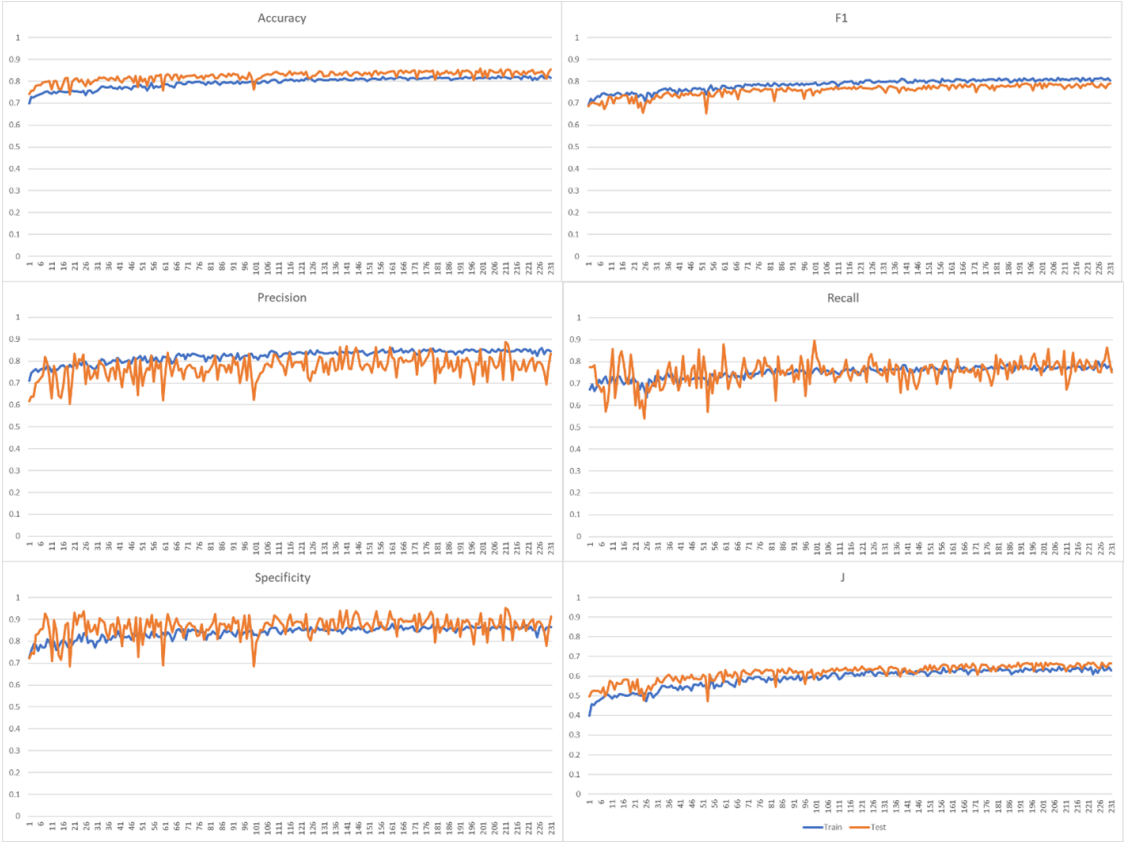

Figure 6.5: Evolution of Evaluation Metrics at Training with Balanced Training Set

It is important to point out the validation sets for balanced and unbalanced training are identical, there are no differences between them, the only difference in these models is whether their respective training sets were balanced.

It is evident, from Figure 6.8, that while there is no difference in the fitted trend line in accuracy progression, there are marked differences in the J measure, where the balanced set outperforms the unbalanced one. Therefore, as it is evident that the balanced dataset performance is superior when training our particular model, we continued training on a

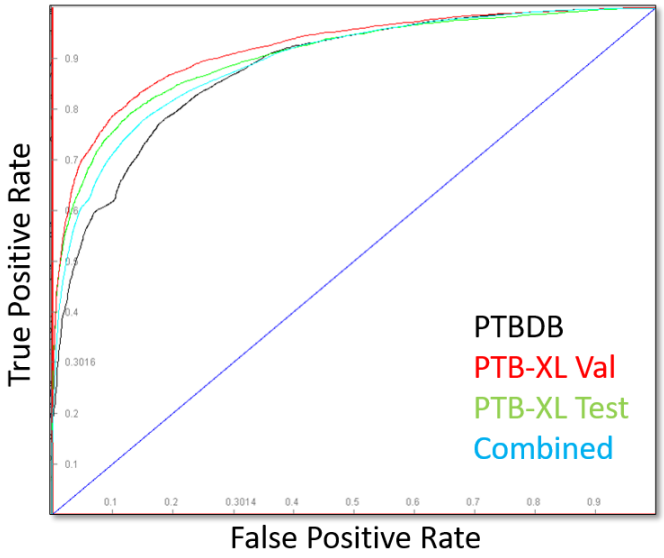

Figure 6.6: ROC plot of the Proposed Classifier

balanced training set for further iterations to attain better performance. The numbers and findings reported hereafter belong to the model trained on the balanced dataset.

6.3.2 Influence of Myocardial Infarction Location on Detection

The heart is nourished by multiple arteries, and as such, myocardial infarctions can occur in one or many of its different regions. When the blood supply to one of those regions is cut, the electrical activity of the muscle fiber therein starts to change. The morphology of the changes visible in the electrocardiogram recordings depends on the leads used.

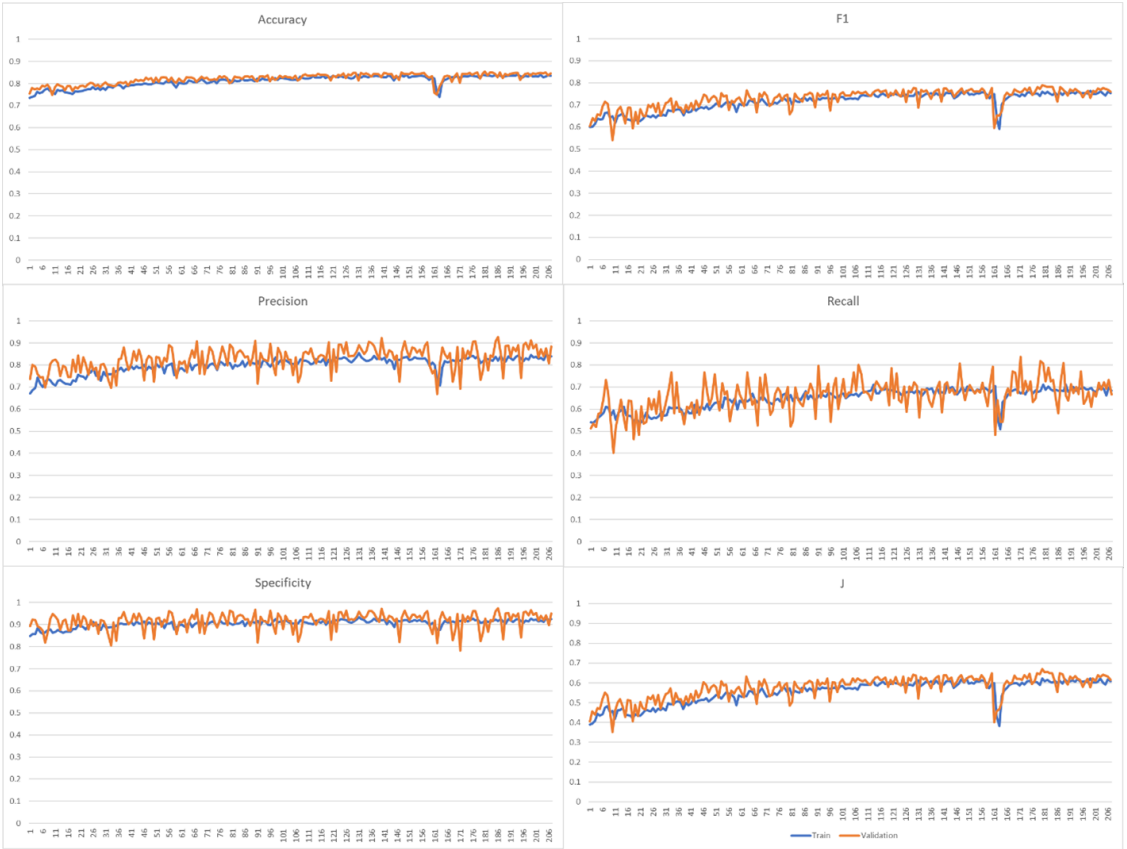

Figure 6.7: Evolution of Evaluation Metrics at Training with Unbalanced Training Set

Therefore, in this section we break down the dataset and explore how the proposed detector fares for the different kinds of MIs present therein. Generally speaking, we will be looking at the following types of myocardial infarctions.

Anterior MI produces electrical changes in leads V1 through V6 but might or might not produce changes in Lead II, depending on where the left anterior descending (LAD) coronary artery is occluded. *Antero-Lateral* MI can be observed in leads V3 through V6, while *Antero-Septal* MI can be observed in leads V1 through V4. Similarly to Anterior MIs, changes in Lead II for either case depend on the location of the LAD occlusion. [66]

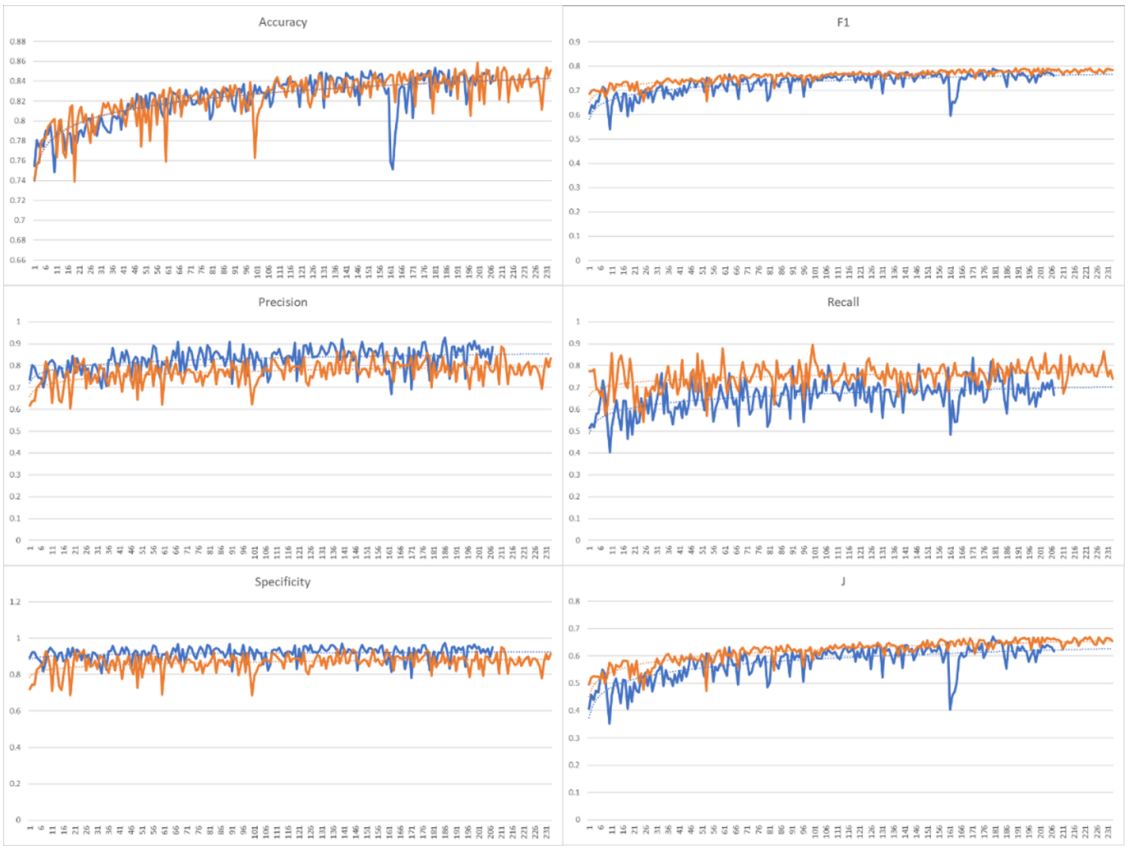

Figure 6.8: Comparison of Balanced VS Unbalanced Training

Inferior MIs, in most cases, involve a blockage of the right coronary artery and they account for about 40% of all MIs. When an MI occurs in this region, the resulting electrical activity changes can be observed in leads II, III, aVF, and aVL. [67]

Lateral MIs are rare in isolated form; they usually occur as part of larger infarctions involving multiple areas, as the left ventricular lateral wall is perfused by the left anterior descending artery and the left circumflex artery. In EKGs, lateral MIs can be observed in leads I, III, aVL, aVF, V5, and V6. [68]

Posterior MIs are challenging to diagnose when they occur in isolation due to subtle changes in EKGs that often lead to misdiagnosis. They commonly occur in conjunction to Inferior and Infero-Lateral MIs. EKG changes associated with solely posterior MIs may be hard to observe and might require the addition of extra leads not included in the typical 12-lead EKGs (V7-V9). [69]

Both the PTB and PTB-XL databases have broadly the same types of MI classifications, but the PTB-XL further assign probabilities to the location of the infarctions. MI beats within the PTBDB have only one infarction location label.

Table 6.2 brakes down the total number of heart beats for every type of MI in the dataset, as well as healthy controls, and gives a detection rate metric. Figure 6.9 visually represents these values as a plot for ease of understanding as well as for the availability of samples for each of the cases.

From these data, it can be concluded that the proposed algorithm performs exceedingly well for the wide range of MIs present in the datasets, considering that this is a single lead detector using only Lead II EKGs.

Furthermore, it is not possible to conclude whether the diminished performance in Lateral MI detection is due to the reduced availability of samples in the dataset and at large (purely lateral infarctions are uncommon [68]) or to the inability of the model to detect them. A measure of caution is taken for the Antero-Septal-Lateral and Postero-Lateral MIs due to their limited presence in the available data, where their exceedingly good recognition rate might be due to chance.

Table 6.2: Influence of MI Location on Detection Rate

Type	*PTBDB (1kHz)*			*PTB-XL Val (500 Hz)*			*PTB-XL Test (500 Hz)*		
	Correct	*Total*	*Detection Rate (%)*	*Correct*	*Total*	*Detection Rate (%)*	*Correct*	*Total*	*Detection Rate (%)*
Anterior	4840	6261	**77.30**	96	167	**57.49**	51	124	**41.13**
Antero-Lateral	4423	6506	**67.98**	230	244	**94.26**	134	159	**84.28**
Antero-Septal	7737	11173	**69.25**	1504	1705	**88.21**	1323	1531	**86.41**
Antero-Septo-Lateral	270	273	**98.901**	**-**	**-**	-	**-**	**-**	-
Inferior	9219	12250	**75.26**	852	911	**93.52**	844	903	**93.47**
Infero-Lateral	6609	7866	**84.02**	323	335	**96.42**	318	329	**96.66**
Infero-Posterior	**-**	**-**	-	25	47	**53.19**	23	23	**100**
Infero-Postero-Lateral	2447	2615	**93.56**	32	38	**84.21**	38	38	**100**
Lateral	48	458	**10.48**	**-**	**-**	-	**-**	**-**	-
Posterior	344	467	**73.66**	18	18	**100**	12	12	**100**
Postero-Lateral	777	777	**100**	**-**	**-**	-	**-**	**-**	-
Healthy Control	9392	10550	**89.02**	8442	9606	**87.88**	8429	9661	**87.25**

6.3.3 Frequency Independence

Frequency independence is one of the key appeals and features of the proposed algorithm. By not being limited to a specific sampling frequency, the model proposed herein can be easily deployed to existing hardware and take advantage of the wide variety of data available, albeit recorded at different sampling frequencies, without the need of expensive and time-consuming retraining.

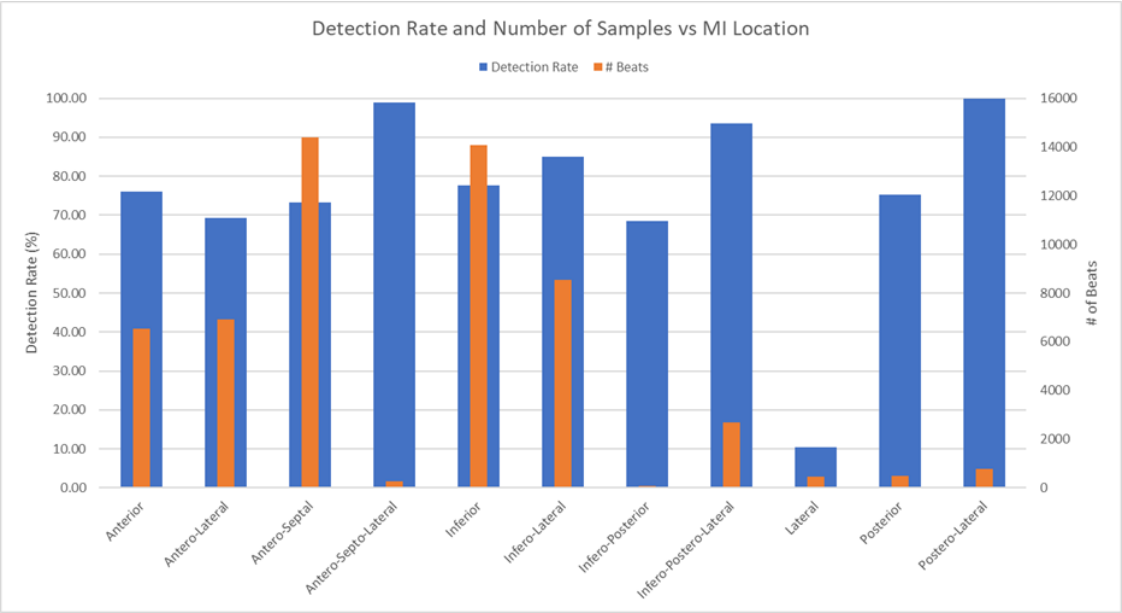

Figure 6.9: Detection Rate According to MI Location

To test the frequency independence of the proposed model, a series of tests are performed on three portions of the frequency spectrum of interest:

250-1000Hz: This is the training range of the detector and were the highest and most stable detection rate is expected to be found.

10-250Hz: Low range frequency response. As the sampling frequency decreases, the amount of data available to the model also decreases and as a result we expect a decrease in performance as less data is available.

1kHz-10KHz: High frequency range. As the sampling frequency increases more data is available. However, the time in between samples is also reduced, which might confuse

the detector. One significant drawback of large sampling rates is increased processing time and hardware cost.

Each of these ranges is subdivided into ten (10) distinct frequency sub-bands and the model is tested on thirty (30) randomly selected distinct frequencies from the continuous range within each bin boundary. For each frequency point within the bin, thirty (30) patients are randomly selected from the PTB-XL test set (fold 10) to obtain the performance metrics. Within each bin, no two different frequency points can share data from the same patient. Figure 6.10 showcases the accuracy distribution within each frequency band/bin.

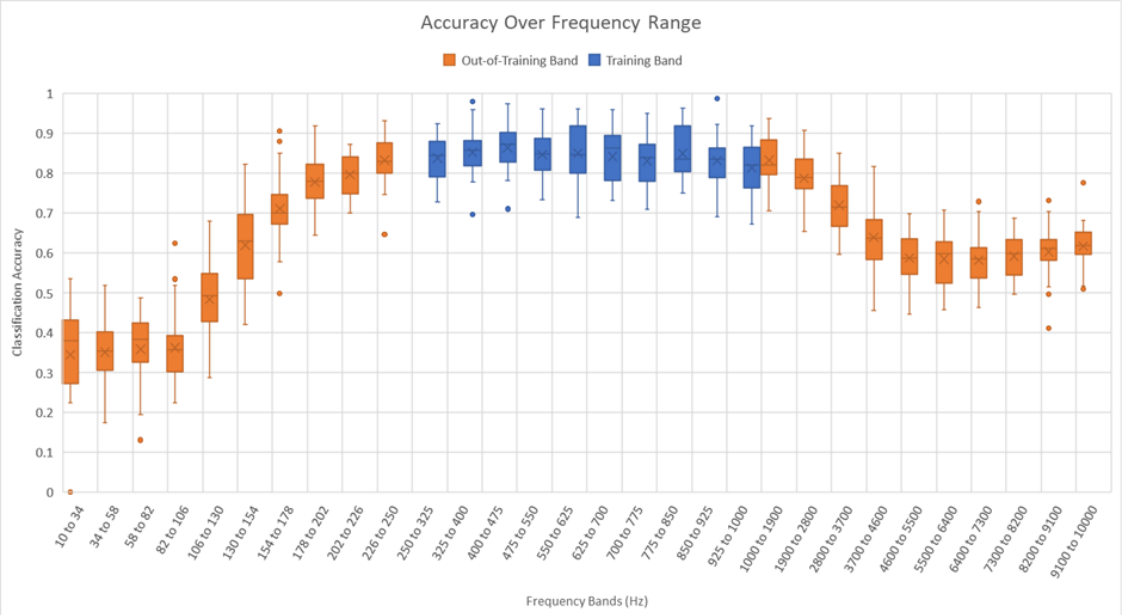

Figure 6.10: Accuracy vs Frequency

To assess the claim of frequency independence, we perform a multiple comparison test on the accuracy distributions for each of the frequency bands in the three relevant regions. We use a pairwise t-test with Bonferroni adjustment. The significance level is at $p < 0.05$.

Frequency Independence from 250Hz to 1kHz:Training Band. Within this frequency band, there is no statistical evidence that any of the bins is different from each other. That is, all p values for each of the pairwise t-tests are more than 0.05. See Table 6.3.

Table 6.3: Training Band Multiple Comparison Test

	250 to 325	325 to 400	400 to 475	475 to 550	550 to 625	625 to 700	700 to 775	775 to 850	850 to 925
325 to 400	1.00	Nd	Nd	Nd	Nd	Nd	Nd	Nd	Nd
400 to 475	1.00	1.00	Nd	Nd	Nd	Nd	Nd	Nd	Nd
475 to 550	1.00	1.00	1.00	Nd	Nd	Nd	Nd	Nd	Nd
550 to 625	1.00	1.00	1.00	1.00	Nd	Nd	Nd	Nd	Nd
625 to 700	1.00	1.00	1.00	1.00	1.00	Nd	Nd	Nd	Nd
700 to 775	1.00	1.00	1.00	1.00	1.00	1.00	Nd	Nd	Nd
775 to 850	1.00	1.00	1.00	1.00	1.00	1.00	1.00	Nd	Nd
850 to 925	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	Nd
925 to 1000	1.00	0.60	0.06	1.00	0.78	1.00	1.00	1.00	1.00

Frequency Independence from 10Hz to 250Hz: Low Frequencies. In this band there are multiple bins which are different from each other. The statistical analysis shows no evidence that there is any difference in accuracy from 10-106 Hz. While the same is true for 178-250Hz. See Table 6.4.

Table 6.4: Low Frequencies Multiple Comparison Test

	10 to 34	34 to 58	58 to 82	82 to 106	106 to 130	130 to 154	154 to 178	178 to 202	202 to 226
34 to 58	1.00	Nd	Nd	Nd	Nd	Nd	Nd	Nd	Nd
58 to 82	1.00	1.00	Nd	Nd	Nd	Nd	Nd	Nd	Nd
82 to 106	1.00	1.00	1.00	Nd	Nd	Nd	Nd	Nd	Nd
106 to 130	< 0.001	< 0.001	< 0.001	< 0.001	Nd	Nd	Nd	Nd	Nd
130 to 154	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	Nd	Nd	Nd	Nd
154 to 178	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	Nd	Nd	Nd
178 to 202	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	0.08	Nd	Nd
202 to 226	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	1.00	Nd
226 to 250	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	0.46	1.00

Frequency Independence from 1kHz to 10kHz: High Frequencies. This band performs similarly to the low band. There is an accuracy transition period right after 2800 Hz which

stabilizes after 3700 Hz. There is no statistical evidence of difference in accuracy from 3.7 kHz to 10 kHz. See Table 6.5.

Table 6.5: High Frequencies Multiple Comparison Test

	1000 to 1900	1900 to 2800	2800 to 3700	3700 to 4600	4600 to 5500	5500 to 6400	6400 to 7300	7300 to 8200	8200 to 9100
1900 to 2800	0.28	Nd	Nd	Nd	Nd	Nd	Nd	Nd	Nd
2800 to 3700	< 0.001	< 0.001	Nd	Nd	Nd	Nd	Nd	Nd	Nd
3700 to 4600	< 0.001	< 0.001	< 0.001	Nd	Nd	Nd	Nd	Nd	Nd
4600 to 5500	< 0.001	< 0.001	< 0.001	0.08	Nd	Nd	Nd	Nd	Nd
5500 to 6400	< 0.001	< 0.001	< 0.001	0.04	1.00	Nd	Nd	Nd	Nd
6400 to 7300	< 0.001	< 0.001	< 0.001	0.03	1.00	1.00	Nd	Nd	Nd
7300 to 8200	< 0.001	< 0.001	< 0.001	0.19	1.00	1.00	1.00	Nd	Nd
8200 to 9100	< 0.001	< 0.001	< 0.001	1.00	1.00	1.00	1.00	1.00	Nd
9100 to 10000	< 0.001	< 0.001	< 0.001	1.00	1.00	1.00	1.00	1.00	1.00

To explore the width of the frequency invariant region around the training area, we perform a new multiple comparison test with the mid bin accuracy of the training region and the low and high side frequency bins. See Table 6.6.

Table 6.6: MidBand Performance Range Test

	178 to 202	202 to 226	226 to 250	1000 to 1900	1900 to 2800	2800 to 3700
550 to 625	0.02	1.00	1.00	1.00	0.18	< 0.05

From these statistical analyses we can conclude that the proposed models performance is independent from the signals sampling frequency from around 202Hz to 2800Hz, a range considerably larger than that covered during training. However, although frequency independence is observed in the training region, it is relevant to explore the possible causes of the decrease in performance outside of it.

Figure 6.11 shows how the classifiers confidence (output) changes over the frequency spectrum for both MI and HC beats. It is observable that the model remains fairly confident (confidence $\geqslant 0.5$) on MI beats, but it becomes confused on healthy ones as the

sampling frequency decreases past the lowest and increases above the maximum training frequencies, thereby confusing healthy beats for infarcted ones.

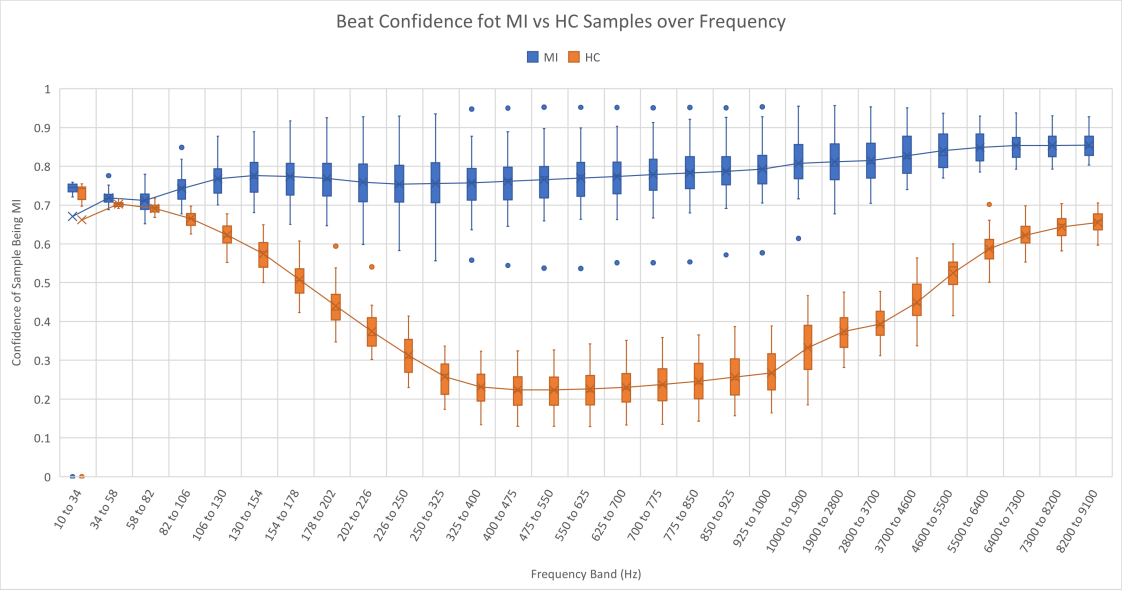

Figure 6.11: MI detection Confidence VS Frequency for MI and HC beats

Here, perhaps, it is important to recall how LSTMs see time varying signals and to remember that they produce a classification output for each input timestep. With this in mind, when the classifier is presented with a one second signal sampled at half the minimum sampling frequency (250 Hz) seen at training (that is fs=125Hz), it sees a sample that ends where it expects to find the R-peak and not sufficient data has been collected to make a determination from the models perspective. On the other hand, when the signal is sampled at a rate faster than the maximum seen during training, time apparently slows down, taking more samples for events to happen than what was previously experienced. When the sampling frequency is 2kHz, the model sees 2000 points for a one second sample. By the time the R-peak happens in the sample, at 1000 points, it is expecting to have seen the end of the longest previously experienced sample. However, the proposed

final trained classifier is able to remain frequency invariant up to frequencies two and a half (x2.5) times higher than the maximum seen at training.

Another important key variable in the classifiers performance is the ability of the beat detection algorithm to actually detect beats. Figure 6.12 plots the number of detected beats versus the signals sampling frequency. It is evident in this plot that the detectors performance decreases as the sampling frequency does, particularly below 106Hz, with the worst performing range being 10 to 34 Hz, where in some instances the detector fails to detect any beats whatsoever. This is to be expected due to the fact that at very low sampling frequencies, the EKG becomes unable to reliably sample its most important feature for beat detection and segmentation, the R-spike.

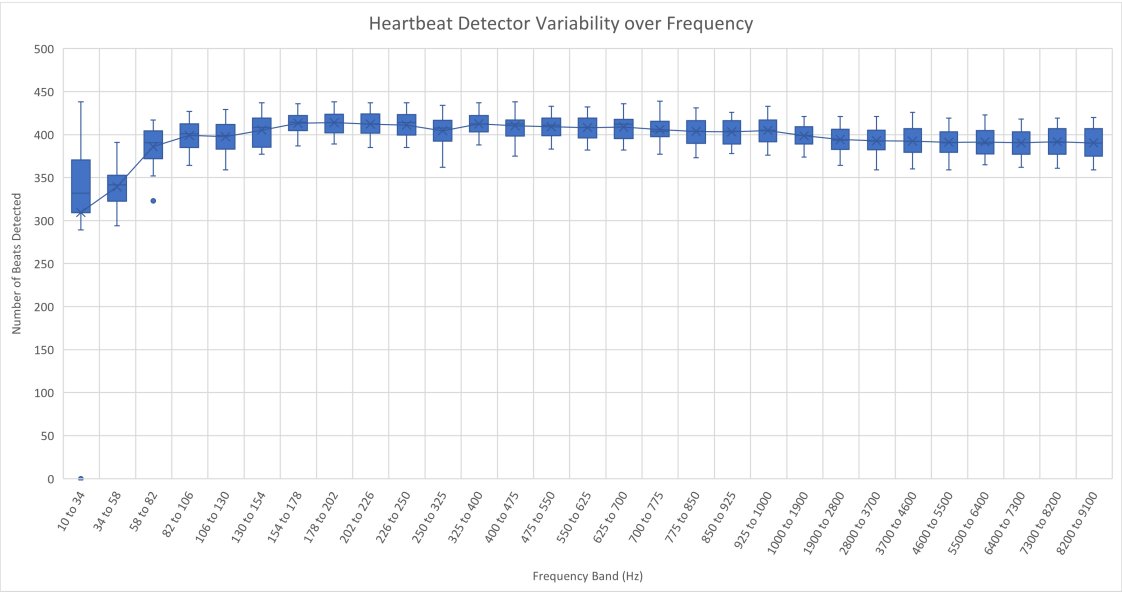

Figure 6.12: Heartbeat Detection Variability vs Frequency

Nonetheless, the detector and beat segmentation algorithm performs adequately for frequency bands above 58Hz where no significant difference in their performance can be observed.

Therefore, to improve the models performance at high frequencies, we simply need to expose it to higher frequency samples at training time, where the highest expected sampling frequency is taken into consideration. However, increasing performance on the lower side of the spectrum is a bit more challenging, as we are bound by the Nyquist criterion when setting a lower bound on a practical frequency at which we can still sample the frequency components associated with the ventricular depolarization events.

6.3.4 Comparison

Comparing the proposed method to existing ones is not very straight forward for the following reasons:

1) Type of dataset generation and classifier, inter- vs intra- patient.

2) Dataset used. The same classifier can have drastically different performance over different datasets.

3) Leads used. The number of leads used by the model affects the performance for the simple reason that some type MIs are more visible in some leads over others. More leads (i.e. more views of the heart) equals more information.

4) Frequency dependency. To our knowledge, there is no other frequency independent MI detector present in the literature.

However, we still present an encompassing picture of related studies and go ever their differences and similarities in the forthcoming tables. Table 6.7 goes over the different variables of importance associated with the different models proposed in the literature. It shows how there is large diversity of models/methods employed along with different sample sizes, number of EKG leads, dataset generation method, and even preprocessing steps (filtering).

Table 6.7: Myocardial Infarction Detection Methods Comparison

Study	Dataset	Fs	Leads	Filter	Split Method	Sample Length	Method
[66]	**PTB**	1kHz	12	None	**Patient**	0.6 seconds	MFB-CNN
[67]	**PTB**	250Hz	12	None	**Patient**	0.6 seconds	CNN and BLSTM
[17]	**PTB**	200Hz	12	45Hz LP	File (unclear)	4 seconds	ML-ResNet
[68]	**PTB**	200Hz	12	45Hz LP	File (unclear)	4 seconds	PCA, SVM
[69]	**PTB**	1kHz	12	60Hz BS	**Patient**	5 seconds	DWT-PCA-ANN
[69]	**PTB**	1kHz	12	60Hz BS	**Patient**	5 seconds	Deep Residual CNN
Chapter 5	**PTB**	1kHz	12	50Hz BS	**Patient**	1 second	Deep LSTM
[6]	**PTB**	1kHz	1	Wavelet	Beat	0.65s	DeepCNN
[7]	**MIT- BIH**	360Hz	1	LLSS	Beat	RR Interval	LMNN
[8]	**PTB**	250Hz	II, III, aVF	SG and Wavelet	**Patient**	3s	SWT & SVM
[12]	**PTB**	1kHz	1	0.5-45Hz BP	Beat	2 beats	Bagging Tree Classifier
[13]	**PTB**	250Hz	V2, V3, V5, aVL	FIG	**Patient**	Whole ECG	ML-CNN
[14]	**PTB**	1kHz	V5	0.05-40Hz BP	**Patient**	0.65s	Random Forest
Chapter 4	**PTB**	1kHz	II	60Hz BS	**Patient**	1 second	LSTM
Proposed	**PTB, PTB-XL**	250Hz-1kHz	II	45Hz LP	**Patient**	1 second	Deep LSTM

Table 6.8 shows a more detailed performance comparison of the methods that most closely resemble the type of data used by ours and those that can be more fairly compared

to each other. From it, we can appreciate that the proposed algorithm outperforms all others in terms of specificity over all tested sets. However, it does not achieve better metrics for Accuracy, Specificity, or J; the best performance in those is achieved by the model we originally proposed in Chapter 4. The increased performance from the latter might be due to the fact that said was trained over the PTB database and perhaps learned some intrinsic characteristics of such that help it at testing. On the other hand, it could also be latching on to the 50Hz underlaying signal present in the PTB database for only MI patients, something their preprocessing steps fail to remove, as they use a 60Hz band-stop (BS) filter instead of a 50Hz one. Furthermore, it could also be due to specialization, or in this models case generalization; perhaps learning how to detect MIs at various sampling frequencies causes the model to learn generalization procedures that decrease its performance at any given single frequency point.

Table 6.8: MI Detector Performance Comparison

Study	**Dataset**	**Leads**	**Sample Length**	**Method**	**Accuracy**	**Recall**	**Specificity**	**J**
[8]	**PTB** HC: 3,240 MI: 3,037	II, III, aVF	3s	SWT & SVM	81.71%	79.01%	79.26%	58.27%
[14]	**52MI, 52HC**	V5	0.65s	Random Forest	83.26%	87.95%	78.82%	66.77%
Chapter 4	**PTB** HC: 10,123 MI: 50,732	II	1s	LSTM	**89.56%**	**91.88%**	80.8%	**72.69%**
Proposed	**PTB** HC: 10,550 MI: 48,646	II	1s	Deep LSTM	77.12%	75.85%	83.02%	58.87%
	PTB-XL Val HC: 9,606 MI: 3,465				85.07%	81.54%	87.31%	68.66%
	PTB-XL Test HC: 9,661 MI: 3,119				84.17%	78.37%	**87.55%**	65.92%

Overall, the detection model and associated processing pipeline presented herein, to the best of our knowledge, are the first to attempt myocardial infarction detection independent of the signals sampling frequency. We are also the first to test a trained MI detector across datasets, that is, training in one dataset and testing its performance in another. The

proposed algorithm and training procedure presented herein, have not been previously employed in the literature and highlight a new possible research direction. Finally, the model itself shows great promise and has the potential and ability to be readily deployed on existing hardware devices that collect electrocardiogram signals from standard lead II with no need for retraining.

CHAPTER 7

CONCLUSION

7.1 Central Aim

This study aimed to develop a real-time single-beat frequency independent myocardial infarction detector from single-lead (Lead II) electrocardiograms for the early detection and treatment of the leading cause of dead in the United States prior to COVID-19, heart attacks. To fulfill this aim, we explored the literature for the latest detection algorithms and their deficiencies, while developing new methods and algorithms to overcome them.

7.2 Contributions

In Chapter 2, we introduced an algorithm for the real-time detection of the most distinct component of the cardiac electrical signal, the R-peak. It implemented independent component analysis (ICA) to isolate the source of this signal and a rolling threshold to select potential candidates. This approach could lead to real-time analysis and decomposition of the complete cardiac signal and the online diagnosis of cardiac abnormalities. The potential benefits of such real-time implementation are far reaching, from the online diagnosis of diseases and abnormalities to its use in tracking heart functioning during the testing and development of cutting-edge research and treatments, such as transcranial magnetic stimulation.

In Chapter 3, we explored the literature for the latest myocardial infarction classification and detection methods, as well as the available databases and their makeup. We further evaluate how training datasets are currently generated in the literature and ex-

perimentally highlighted the tendency of machine learning models to overfit and provide misleading testing metrics when evaluated over improperly engineered sets.

In Chapter 4, we proposed a novel multilayer LSTM neural network for near-real time and accurate infarction detection using one-second EKG samples of Lead II EKGs. 148 myocardial infarction patients and 52 healthy controls which were split into 10 non-patient-overlapping sets for 10-fold cross validation. The proposed algorithm, which uniquely relies on a single heartbeat of single-lead (lead II) electrocardiograms, achieved an accuracy of 89.56%, recall/sensitivity of 91.88%, a specificity of 80.81%, all with a 95% Confidence Interval (CI). It is emphasized that in the approach considered, while deploying the patient split method, care was taken that no heart beats or subjects (MIs and HCs) seen in the training phase are considered in the testing phase. Moreover, within the design construct of the model allowed for a processing time of only 40 milliseconds to diagnosis, which is well within the time in between two heartbeats of 300 milliseconds, assuming a fast heart rate. This fast-processing capability of the model allows for its deployment on existing wearable/portable devices and other test instruments which could have significant societal impact in the lives of at-risk patients and the population at large.

In Chapter 5, we put forward a novel pipeline for the real-time online detection of myocardial infarction from a single heartbeat of a 12-lead electrocardiogram. Our pipeline combines a real-time R-spike detector with a novel deep LSTM classifier to produce highly accurate and fast detection results (40 milliseconds when accounting for preprocessing time). The proposed system achieves state-of-the-art performance with an accuracy of 95.76% with a 95% confidence interval (CI) from 93.36% to 98.16% and the average J-Score is 90.31% with a 95% CI from 84.11% to 96.51%. The uses and benefits of the

proposed system are far reaching as they can have significant societal and clinical impacts in the lives of not only at-risk patients but also the population at large.

Finally, in Chapter 6 we combine previous knowledge from chapters 3, 4 and, 5 and present a novel and complete pipeline along with a deep-LSTM neural network for the detection of Myocardial Infarction. By combining a modified real-time heartbeat detector with frequency resampling, and dataset balancing, we achieve a real-time single-beat frequency independent myocardial infarction detector that could be readily deployed and integrated to existing exercise and cardiac monitoring equipment with sampling frequencies ranging from 202Hz to 2.8KHz with no need to modifications to any component of the pipeline. The proposed model achieves MI detection rates of 77.02% combined across the tested datasets (of different sampling frequencies) and a specificity of 88.08% with a combined accuracy of 80.90%.

However, regardless of what results we obtain in this study and as mentioned before in the literature, the true benefits of any MI detector would depend on the numbers of at-risk individuals. Because a significant portion of MIs mortality is due to the lack of awareness to the condition and therefore the lack of immediate medical attention. Consequently, such an approach could significantly improve the odds of detecting silent MIs by monitoring at-risk individuals and providing them with an early diagnosis. Determining the early sign of MI could help in the planning of early treatment and extending the time available for doctors to plan ahead on an individual basis in case of an emergency.

Furthermore, integrating the proposed approach to a standalone software/hardware platform that would monitor a patients cardiac activity, perhaps through the use of widely available fitness trackers and other currently commercially available wearable devices,

would be a significant portion of our research efforts for the next step in the development of this software-based design. Moreover, given the implemented architecture of our final LSTM model, we could also seek to identify at which point of the heartbeat the presence of myocardial infarctions becomes evident, as the last LSTM layer can be set up to provide a per-time-step diagnosis.

7.3 Retrospective

The genesis of this research is in the intent of developing and deploying a novel multilayer Long-Short Term Memory Neural Network (LSTM) architecture with unique application for the diagnosis of myocardial infarctions from individual heartbeats of single-lead electrocardiograms. To the best of our knowledge, this research work is the first to attempt single lead (lead II) electrocardiograms on individual heartbeats that achieve near real-time diagnosis with a processing time of 40 milliseconds. It is worthy to note that the processing time of 40 milliseconds is most appropriate for online classification as the time between fast heartbeats is around 300 milliseconds, which is well above the achieved processing time. In order to achieve this processing time and still have a high accuracy, a simplified LSTM-based model is proposed without compromising its flexibility and accuracy for myocardial infarction detection. The designed architecture takes into consideration all the different data splits methods in the training phase and cautions against data splits that could conceal the occurrence of overfitting to produce misleadingly high-testing metrics of the models performance. Moreover, the proposed approach with its fast processing time and optimal design architecture lends itself to a standalone integrated software/hardware platform that would monitor a patients cardiac activity, through the use

of widely available fitness trackers and other currently commercially available wearable devices.

The novelty of this research endeavor in the area of myocardial infarction detection and the contributions made are reflected through the following observations:

1. The depth (number of layers), size of layers, and architecture of the model (the model itself is not widely used, although the LSTM neurons are). In general, there is nothing special about similar architectures of decent size, as we can get the same performance out of a network with 100 neurons in the first layer and 50 in the second layer, as we did for ours. The main objective behind the proposed network is to minimize the risk of overfitting while allowing for flexibility and accuracy in the results. The number of neurons (100) is deemed large enough to accommodate the significant features but not too large as to prohibitively increase the training time. When added complexity is required of the model, then more layers could be added, but with caution that it could lead to overfitting if there is not enough variability in the training data.

2. This research work provides valuable insight into the importance of proper data splitting techniques when generating training and testing datasets and we highlight key findings on the issue of data splits and show how they can be ineffectual if not deployed properly.

3. Empirical evaluations demonstrate the propensity to overfitting that the powerful LSTMs architecture have through empirical data (randomized data labels and different splitting methods). Evidently, all machine learning methods, even KNNs, have a propensity to overfit, so what we tested was two-fold:(a) we show that if the data is not properly

split into training and testing sets, overfitting occurs in a non-apparent way, masked by the contamination of the testing set, and misleadingly yielding a high performance that will not match that of the deployed system in real-unseen data; (b) in the particular case of LSTMs, we show that they can overfit to such an extreme as to memorize data from a patient to effectively yield a biometric detector, while the only thing linking heartbeats from the testing and training sets is that the testing data came from patients seen during training (even though the testing heartbeats were not used for training). This is extremely important to consider when deploying such LSTM-based models. In retrospect, we do not claim that we have found a way to overcome overfitting, overfitting will always happen in the long run; what we did is made sure that we are detecting it and that we are not being blinded to it by ineffectual splitting techniques.

4. Proposed the use of the of independent component analysis as means to extract that independent component bound to contain the R-peaks in the QRS complex as the most distinct and important feature for analyzing EKG data. This approach separates the components of the cardiac electrical signal and applies the detection process on them. It preserves all the leads recordings of the signal for observation and validation while still performing heart disease biomarker detection over the independent components.

5. Developed and implemented an integrated method that combined the LSTM and the real-time R-peak detection as means to consolidate their potential for more accurate and faster detection of myocardial infarction condition. A method found to be extremely beneficial for monitoring patients under cardiac stress and/or at risk of myocardial infarctions, while still significantly improving the prospects for both the correct diagnosis of the disease condition and the subsequent planning of early intervention and treatment.

6. In MI research, since there is no common standard frequency range by which a detector could be designed to be stable over any prescribed frequency range, this dissertation also delves into designing an MI detector that is not only single lead but also frequency independent for its easy deployment in currently available exercise equipment and fitness monitoring devices. Recall that in this mission of seeking frequency independence of this single lead detector, the real-time processing issue for its practical deployment has already been resolved.

As a result of all these new developments, a novel myocardial infarction (MI) system design was created for inter-patient MI detection from individual heartbeats of single-lead (lead II) electrocardiograms not seen in the training phase. This detector assimilates high detection accuracy and near real-time processing of 300 milliseconds to a diagnosis. The design was further augmented to be frequency independent and stable performance metrics were still obtained over the frequency range of 202Hz to 2.8kHz. MI detection, applied over this wide frequency range and on the entire PTB database (the largest EKG database that is available to the research community), was achieved with an accuracy of 77.12%, a positive predictive value (PPV) of 75.85%, and a negative predictive value (NPV) of 83.02%. In achieving this final design model, initial design steps took into serious consideration the different data splits methods in the training phase to prevent data contamination and overfitting. The ultimate goal of this research was reached by establishing an MI detector that is accurate, capable of real-time processing, and amenable to real-world use on fitness trackers and other wearable heart monitoring devices.

REFERENCES

[1] Heart Attack, howpublished=https://www.cdc.gov/heartdisease/heart_attack.htm. Accessed: Aug.18, 2017.

[2] Heart Disease Fact Sheet, howpublished=https://www.cdc.gov/dhdsp/data_statistics/fact_sheets/fs_heart_disease.htm. Accessed: Aug.23, 2017.

[3] G. Nicholson, S. R. Gandra, R. J. Halbert, A. Richhariya, and R. J. Nordyke. Patient-level costs of major cardiovascular conditions: a review of the international literature. *ClinicoEconomics and Outcomes Research*, 8:495–506, September 2016.

[4] R. Barnett. Case histories acute myocardial infarction. *Lancet*, 393(10191):2580, June 2019.

[5] P. Kligfield, L.S. Gettes, J.J. Bailey, R. Childers, B.J. Deal, E.W. Hancock, G. van Herpen, J.A. Kors, P. Macfarlane, D.M. Mirvis, O. Pahlm, P. Rautaharju, and G.S. Wagner. Recommendations for the standardization and interpretation of the electrocardiogram - part i. *Circulation*, 25(10):1306–1324, March 2007.

[6] U. R. Acharya, H. Fujita, S. L. Oh, Y. Hagiwara, J. H. Tan, and M. Adam. Application of deep convolutional neural network for automated detection of myocardial infarction using ecg signals. *Information Sciences*, 415:190–198, November 2017.

[7] P Kora. Ecg based myocardial infarction detection using hybrid firefly algorithm. *Computer Methods and Programs in Biomedicine*, 152:141–148, December Dec. 2017.

[8] L. Sharma and R. Sunkaria. Inferior myocardial infarction detection using stationary wavelet transform and machine learning approach. *Signal Image and Video Processing*, 12(2):199–206, Feb 2018.

[9] W. Liu, Q. Huang, S. Chang, H. Wang, and J. He. Multiple-fracture-branch convolutional neural network for myocardial infarction diagnosis using electrocardiogram. *Biomedical Signal Processing and Control*, 2018.

[10] A. Dohare, V. Kumar, and R. Kumar. Detection of myocardial infarction in 12 lead ecg using support vector machine. *Applied Soft Computing*, 64:138–147, Mar. 2018.

[11] L. Sun, Y. Lu, K. Yang, and S. Li. Ecg analysis using multiple instance learning for myocardial infarction detection. *IEEE Transaction on Biomedical Engineering*, 59(12):3348–3356, Dec. 2010.

[12] S. Khatun and B. Morshed. Detection of myocardial infarction and arrhythmia from single-lead ecg data using bagging trees classifier. *IEEE International Conf. on Electro Information Technology (EIT)*, May 14-17, 2017.

[13] W. Liu, M. Zhang, Y. Zhang, Y. Liao, Q. Huang, S. Chang, H. Wang, and J. He. Real-time multilead convolutional neural network for myocardial infarction detection. *IEEE Journal of Biomedical and Health Informatics*, 22(5):1434–1444, Sep. 2018.

[14] D. Sopic, A. Aminifar, and D. Atienza. Real-time event-driven classification technique for early detection and prevention of myocardial infarction on wearable systems. *IEEE transactions on Biomedical Circuits and Systems*, 12(5):982–992, Oct. 2018.

[15] J. Liu, C. Zhang, Y. Zhu, T. Ristaniemi, T. Parviainen, and F. Cong. Automated detection and localization system of myocardial infarction in single-beat ecg using dual-q tqwt and wavelet packet tensor decomposition. *Computer Methods and Programs in Biomedicine*, 184:105120, Feb. 2020.

[16] P. Hao, X. Gao, Z. Li, J. Zhang, F. Wu, and C. Bai. Multi-branch fusion network for myocardial infarction screening from 12-lead ecg images. *Computer Methods and Programs in Biomedicine*, 184:105286, Feb. 2020.

[17] C. Han and L. Shi. Ml-resnet: A novel network to detect and locate myocardial infarction using 12 leads ecg. *Computer Methods and Programs in Biomedicine*, 185:105138, March 2020.

[18] D. Sadhukhan, S. Pal, and M. Mitra. Automated identification of myocardial infarction using harmonic phase distribution pattern of ecg data. *IEEE Trans. on Instrumentation and Measurement*, 67(10):2303–2313, Oct. 2018.

[19] S. Padhy and S. Dandapat. Third-order tensor based analysis of multilead ecg for classification of myocardial infarction. *Biomedical Signal Processing and Control*, 31:71–78, Jul. 2016.

[20] W. Baxt. Application of artificial neural networks to clinical medicine. *Lancet*, 346(8983):1135–1138, Oct. 1995.

[21] W. Baxt and J. Skora. Prospective validation of artificial neural network trained to identify acute myocardial infarction. *Lancet*, 347(8993):12–15, Jan. 1996.

[22] W. Baxt, F. Shofer, F. Sites, and J. Hollander. A neural computational aid to the diagnosis of acute myocardial infarction. *Annals of Emergency Medicine*, 39(4):366–373, Apr. 2002.

[23] B. Heden, H. Ohlin, R. Rittner, and L. Edenbrandt. Acute myocardial infarction detected in the 12-lead ecg by artificial neural networks. *Circulation*, 96(6):1798–1802, Sep. 1997.

[24] L. Sharma, R. Tripathy, and S. Dandapat. Multiscale energy and eigenspace approach to detection and localization of myocardial infarction. *IEEE Trans. on Biomedical Eng*, 62(7):1827–1837, Jul. 2015.

[25] P. Li, Y. Wang, J. He, L. Wang, Y. Tian, T. Zhou, T. Li, and J. Li. High-performance personalized heartbeat classification model for long-term ecg signal. *IEEE Trans. on Biomedical Eng*, 64(1):78–86, Jan. 2017.

[26] U. Acharya, H. Fujita, V. Sudarshan, S. Oh, M. Adam, J. Koh, J. Tann, D. Ghista, R. Martis, C. Chua, C. Poo, and R. Tan. Automated detection and localization of myocardial infarction using electrocardiogram: A comparative study of different leads. *Knowledge-Based Systems*, 99:146–156, May 2016.

[27] U. Acharya, H. Fujita, V. Sudarshan, S. Oh, M. Adam, J. Tann, J. Koo, A. Jain, C. Lim, and K. Chua. Automated characterization of coronary artery disease, myocardial infarction, and congestive heart failure using contourlet and shearlet transforms of electrocardiogram signal. *Knowledge-Based Systems*, 132:156–166, Sep 2017.

[28] H. He, Y. Tan, and J. Xing. Unsupervised classification of 12-lead ecg signals using wavelet tensor decomposition and two-dimensional gaussian spectral clustering. *Knowledge-Based Systems*, 163:392–403, Jan 2019.

[29] P. Gomis, D. Jones, P. Caminal, E. Berbari, and P. Lander. Analysis of abnormal signals within the qrs complex of the high-resolution electrocardiogram. *IEEE Trans. on Biomedical Eng*, 44(8):681–693, Aug. 1997.

[30] A. Spiegl, P. Steinbigler, I. Schmucking, A. Knez, and R. Haberl. Analysis of beat-to-beat variability of frequency contents in the electrocardiogram using two-dimensional fourier transforms. *IEEE Trans. on Biomedical Eng*, 45(2):235–241, Feb. 1998.

[31] M. Ghiasi, S. Zendehboudi, and A. Mohsenipour. Decision tree-based diagnosis of coronary artery disease: Cart model. *Computer Methods and Programs in Biomedicine*, 192:105400, Aug. 2020.

[32] J. Park, S. Lee, and U. Park. R peak detection method using wavelet transform and modified shannon energy envelope. *JOURNAL OF HEALTHCARE ENGINEERING*, 2017.

[33] Y. Ma, T. Li, Y. Ma, and K. Zhan. Novel real-time fpga-based r-wave detection using lifting wavelet. *CIRCUITS SYSTEMS AND SIGNAL PROCESSING*, 35(1):281–299, JAN 2016.

[34] S. Rekik and N. Ellouze. Enhanced and optimal algorithm for qrs detection. *IRBM*, 38(1):56–16, FEB 2017.

[35] V. Abromavicius and A. Serackis. Algorithm for real-time detection of heart rate from noisy ecg signals supported by continuous blood pressure analysis. *2015 OPEN CONFERENCE OF ELECTRICAL ELECTRONIC AND INFORMATION SCIENCES*, page 2015, 2015.

[36] S. Sahoo, P. Biswal, T. Das, and S. Sabut. De-noising of ecg signal and qrs detection using hilbert transform and adaptive thresholding. *Book Series: Procedia Technology*, 25:68–75, 2016.

[37] L. Sharma and R. Sunkaria. A robust qrs detection using novel pre-processing techniques and kurtosis based enhanced efficiency. *Measurement*, 87:194–204, 2016.

[38] Y. Yeh and W. Wang. Qrs complexes detection for ecg signal: the difference operation method. *Computer Methods and Programs in Biomedicine*, 91:245–254, 2008.

[39] Nian Zhang and Jing Nie. Independent component analysis based blind source separation algorithm and its application in the gravity and magnetic signal processing. *2015 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)*, pages 269–273, 2015.

[40] L. De Lathauwer, B. De Moor, and J. Vandewalle. Fetal electrocardiogram extraction by blind source subspace separation. *IEEE Transactions on Biomedical Engineering*, 47(5):567–572, May 2000.

[41] M. Hamaneh, N. Chitravas, K. Kaiboriboon, S. Lhatoo, and K. Loparo. Automated removal of ekg artifact from eeg data using independent component analysis and continuous wavelet transformation. *IEEE Transactions on Biomedical Engineering*, 61(6):1634–1641, June 2014.

[42] M. Chen, X. Zhang, X. Chen, and M. Zhu. Fastica peel-off for ecg interference removal from surface emg. *BIOMEDICAL ENGINEERING ONLINE*, 15(65), JUN 13 2016.

[43] U. Nelakuditi. Removal of ocular artifacts in eeg. *PROCEEDINGS OF THE 10TH INTERNATIONAL CONFERENCE ON INTELLIGENT SYSTEMS AND CONTROL (ISCO'16)*, 2016.

[44] M. Cabrerizo, A. Cabrera, J. Perez, J. De La Rua, N. Rojas, Q. Zhou, A. Pinzon-Ardilla, S. Gonzalez-Arias, and A. Malek. Induced effects of transcranial magnetic stimulation on the autonomic nervous system and the cardiac rhythm. *The Scientific World Journal*, 2014:349718, Jul 2014.

[45] H. Zen and H. Sak. Unidirectional long short-term memory recurrent neural network with recurrent output layer for low-latency speech synthesis. *2015 IEEE International Conference On Acoustics, Speech, And Signal Processing (ICASSP)*, pages 19–24, Apr. 2015.

[46] T. Breuel, A. Ul-Hasan, M. Azawi, and F. Shafait. High-performance ocr for printed english and fraktur using lstm networks. *12TH International Conference on Document Analysis And Recognition (ICDAR)*, pages 25–28, Aug. 2013.

[47] A. Ullah, J. Ahmad, K. Muhammad, M. Sajjad, and S. Baik. Action recognition in video sequences using deep bi-directional lstm with cnn features. *IEEE Access*, 6:1155–1166, 2018.

[48] D. Marino, K. Amarasinghe, and M. Manic. Building energy load forecasting using deep neural networks. *42nd Annual Conference of The IEEE Industrial Electronics Society*, pages 24–27, Oct. 2016.

[49] K. Zhang, J. Xu, M. Min, G. Jiang, and K. Pelechrinis. Automated it system failure prediction: A deep learning approach. *4th IEEE International Conference on Big Data*, pages 5–8, Dec. 2016.

[50] M. Yousefi, M. Soheili, T. Breuel, E. Kabir, and D. Stricker. Binarization-free ocr for historical documents using lstm networks. *13th IAPR International Conference on Document Analysis And Recognition (ICDAR)*, pages 23–26, Aug. 2015.

[51] Y. Luo. Recurrent neural networks for classifying relations in clinical notes. *Journal of Biomedical Informatics*, 72:85–95, Aug. 2017.

[52] Q. Zhang, L. Yang, Z. Chen, and P. Li. A survey on deep learning for big data. *Information Fusion*, 42:146–157, Jul. 2018.

[53] Hc Chen and Chen. A moving average based filtering system with its application to real-time qrs detection. *Computers in Cardiology*, 2003.

[54] M. Cabrerizo, A. Cabrera, J.O. Perez, J. de la Rua, N. Rojas, Q. Zhou, A. Pinzon-Ardilla, S.M Gonzalez-Arias, and A. Malek. Induced effects of transcranial magnetic stimulation on the autonomic nervous system and the cardiac rhythm. *The Scientific World Journal*, 2014:349718, 2014.

[55] R. Bousseljot, D. Kreiseler, and A. Schnabe. Nutzung der ekg-signaldatenbank. *CARDIODAT der PTB uber das Internet*, 1995.

[56] Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PCh, Mark RG, Mietus JE, Moody GB, Peng C-K, and Stanley HE. Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals, 2000.

[57] Nian Zhang and Jing Ni. Independent component analysis based blind source separation algorithm and its application in the gravity and magnetic signal processing. *12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)*, pages 269–273, 2015.

[58] G. Repovs. Dealing with noise in eeg recording and data analysis. *Informatica Medica Slovenica*, 15, 2010.

[59] Y.-C. Yeh and W.-J. Wang. Qrs complexes detection for ecg signal: the difference operation method. *Computer Methods and Programs in Biomedicine*, 91:245–254, 2008.

[60] R. Bousseljot, D. Kreiseler, and A. Schnabel. Ekg-signaldatenbank cardiodat der ptb über das internet. *Biomedizinische Technik*, 40:317, 1995.

[61] A. Goldberger, L. Amaral, L. Glass, J. Hausdorff, P. Ivanov, R. Mark, and H. Stanley. Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals. *Circulation*, 101(23):215–220, 2000.

[62] A. E. Hoerl and R. W. Kennard. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12:55–67, 1970.

[63] G. Hinton, N. Srivastava, and K. Swersky. Neural networks for machine learning. `http://www.cs.toronto.edu/~tijmen/csc321/slides/lecture_slides_lec6.pdf`. Accessed: Apr. 18, 2019.

[64] X. Glorot and Y. Bengio. Understanding the difficulty of training deep feedforward neural networks. *Journal of Machine Learning Research*, 9:249–256, 2010.

[65] A. Graves. *Supervised Sequence Labelling with Recurrent Neural Networks*. PhD thesis, Dept. of Computer Science, University of Toronto, 2012.

[66] K. Bansal, M. Gore, and P. Nalabothu. Anterior myocardial infarction. *StatPearls*, Aug. 31, 2020.

[67] M. Warner and V. Tivakaran. Inferior myocardial infarction. *StatPearls*, Aug. 8, 2020.

[68] Dipesh Ludhwani. Lateral wall myocardial infarction. *StatPearls*, Sep. 18, 2020.

[69] J. Lizzo and Y. Chowdhury. Posterior myocardial infarction. *StatPearls*, Nov. 29, 2020.

[70] H. Martin, W. Izquierdo, M. Cabrerizo, and M. Adjouadi. Real-time r-spike detection in the cardiac waveform through independent component analysis. *2017 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*, 2017.

[71] W. Liu, Q. Huang, S. Chang, H. Wang, and J. He. Multiple-feature-branch convolutional neural network for myocardial infarction diagnosis using electrocardiogram. *Biomedical Signal Processing and Control*, 45:22–32, August 2018.

[72] W. Liu, F. Wang, Q. Huang, S. Chang, H. Wang, and J. He. Mfb-cbrnn: A hybrid network for mi detection using 12-lead ecgs. *IEEE Journal of Biomedical and Health Informatics*, 24(2):503–514, Feb. 2020.

[73] C. Han and L. Shi. Automated interpretable detection of myocardial infarctions fusion energy entropy and morphological features. *Computer Methods and Programs in Biomedicine*, 175:9–23, Jul. 2019.

[74] K. Jafarian, V. Vahdat, S. Salehi, and M. Mobin. Automating detection and localization of myocardial infarction using shallow and end-to-end deep neural networks. *Applied Soft Computing Journal*, 93, August 2020.

[75] P. Wagner, N. Strodthoff, R. Bousseljot, W. Samek, and T. Schaeffter. Ptb-xl, a large publicly available electrocardiography dataset (version 1.0.1). *PhysioNet*, 2020.

[76] P. Wagner, N. Strodthoff, R. Bousseljot, D. Kreiseler, F. Lunze, W. Samek, and T. Schaeffter. Ptb-xl: A large publicly available ecg dataset. *Scientific Data*, 2020.

[77] H. Martin, W. Izquierdo, U. Morar, M. Cabrerizo, A. Cabrera, and M. Adjouadi. A fast and accurate myocardial infarction detector. *The 2020 International Conference on Computational Science and Computational Intelligence (CSCI20)*, December 2020.

[78] W. Liu, F. Wang, Q. Huang, S. Chang, H. Wang, and J. He. Mfb-cbrnn: A hybrid network for mi detection using 12-lead ecgs. *IEEE Journal of Biomedical and Health Informatics*, 24:503–514, February 2020.

[79] K. Jafarian, V. Vahdat, S. Salehi, and M. Mobin. Automating detection and localization of myocardial infarction using shallow and end-to-end deep neural networks. *Applied Soft Computing Journal*, August 2020.

[80] W. Burger and M. Burge. Principles of digital image processing: core algorithms. pages 231–232, 2009.

VITA

HAROLD MARTIN

EDUCATION

2016-2021	Ph.D., Electrical & Computer Engineering Florida International University Miami, Florida
2014-2016	M.S., Computer Engineering Florida International University Miami, Florida
2010-2014	B.S., Electrical Engineering Florida International University Miami, Florida

PUBLICATIONS AND PRESENTATIONS

H. Martin, U. Morar, W. Izquierdo, M. Cabrerizo, A. Cabrera, and M. Adjouadi, "Real-time Frequency-Independent Single-Lead and Single-Beat Myocardial Infarction Detection", *Artificail Intelligence in Medicine*, (under review), Feb. 2021

H. Martin, W. Izquierdo, M. Cabrerizo, A. Cabrera, and M. Adjouadi, "Near Real-Time Single-Beat Myocardial Infarction Detection from Single-Lead Electrocardiogram using Long Short-Term Memory Neural Network", *Biomedical Signal Processing and Control*, (second revision), Feb. 2021

M. Mafi, W. Izquierdo, H. Martin, M. Cabrerizo, M. Adjouadi, Deep Convolutional Neural Network for Mixed Random Impulse and Gaussian Noise Reduction in Digital Images, IET Image Processing Journal, Vol.14(15), pp. 3791-3801, December 2020.

H. Martin, W. Izquierdo, U. Morar, M. Cabrerizo, A. Cabrera, and M. Adjouadi, "A Fast and Accurate Myocardial Infarction Detector", *The 2020 International Conference on Computational Science and Computational Intelligence (CSCI'20)*, Las Vegas, Nevada, USA, Dec. 2020

U. Morar, H. Martin, W. Izquierdo, P. Forouzannezhad, E. Zarafshan, R. Curiel, M. Roselli, D. Loewenstein, R. Duara, and Malek Adjouadi, "A Deep-Learning Approach for the Prediction of Mini-Mental State Examination Scores in a Multimodal Longitudinal

Study", *The 2020 International Conference on Computational Science and Computational Intelligence (CSCI'20)*, Las Vegas, Nevada, USA, Dec. 2020

M. Mafi, H. Martin, M. Cabrerizo, J. Andrian, A. Barreto, and M. Adjouadi, "A Comprehensive Survey on Impulse and Gaussian Denoising Filters for Digital Images", *Signal Processing*, vol. 157, pp 236-260, Apr. 2019

H. Martin, W. Izquierdo, M. Cabrerizo, and M. Adjouadi, "Myocardial Infarction Detection Through Long-Short Term Memory Neural Network from Single-Lead Electrocardiogram", *HENAAC 2018*, Pasadena Convention Center, Passadena, California, 10/17/2018-10/21/2018, Poster Presentation

L. Galarza, H. Martin, and M. Adjouadi, "Time-of-Flight Sensor in a Book Reader System Design for Persons With Visual Impairment and Blindness", *IEEE Sensors Journal*, vol. 18, issue: 18, pp 7697-7707, Sep. 2018

L. Galarza, H. Martin, M. Adjouadi, "Integrating Low-resolution Depth Maps to High-Resolution Images in the Development of a Book Reader Design for Persons with Visual Impairment and Blindness", *International Journal of Innovative Computing, Information and Control*, vol. 14, Number 3, Jun. 2018

H. Martin, W. Izquierdo, M. Cabrerizo, M. Adjouadi, "Real-time R-spike detection in the cardiac waveform through independent component analysis", *2017 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*, Dec. 2017 Philadelphia, PA, USA

H. Martin, W. Izquierdo, M. Cabrerizo, M. Adjouadi "A Fast and Accurate Algorithm for the detection of the RR Interval in Cardiac Rhythm Waveform.", *HENAAC*, 2017

W. Izquierdo, H. Martin, M. Cabrerizo, W.W. Baker, D. Loewenstein, R. Duara, M. Adjouadi "Predicting Cognitive Test Scores in Alzheimers Patients Using Multimodal Longitudinal Data.", *Alzheimers & Dementia*, vol. 13, issue 7, supplement, pp. 796-797, Jul. 2017

M. H. Cintuglu, H. Martin and O. A. Mohammed, "An intelligent multi agent framework for active distribution networks based on IEC61850 and FIPA standards", *2015 18th International Conference on Intelligent System Application to Power Systems (ISAP)*, Porto,2015, pp. 1-6.

M. H. Cintuglu, H. Martin and O. A. Mohammed, "Real-Time Implementation of Multiagent-Based Game Theory Reverse Auction Model for Microgrid Market Operation", *IEEE Transactions on Smart Grid*, vol. 6, no. 2, pp. 1064-1072, March 2015.