Heart Anomaly Detection System For Ambulatory Electrocardiogram

by

Hong Zu Li

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Computing Science

University of Alberta

Abstract

Cardiovascular diseases are the leading cause of death globally, causing nearly 17.9 million deaths per year. Early detection and treatment are critical for improving this situation. Today's wearable medical devices are becoming popular because of their price and ease of use. Many manufacturers have developed products to continuously monitor patients' heart conditions as they perform their daily activities. However, one major challenge of collecting and analyzing heart data using mobile ECG is baseline wander and motion artifacts created by the patient's motion, resulting in false diagnoses. In addition, very few monitor devices can diagnose complex heart anomalies beyond detecting rhythm fluctuation. The research reported in this thesis proposes an anomaly detection system that could compensate for the motion noise and give a reliable diagnosis based on a single lead ECG signal. The noise removal algorithm could automatically remove the baseline wander and suppresses most motion artifacts in mobile ECG recordings. This algorithm shows a significant improvement over conventional noise removal methods. Two signal quality metrics are used to compare a reference ECG with its noisy version: correlation coefficients and mean squared error. For both metrics, the experimental results demonstrate that the signal filtered by our algorithm can improve the signal-to-noise ratio by ten.

In addition, this thesis describes a new algorithm that combines a Short-Time Fourier Transform (STFT) spectrogram of the ECG signal with hand-crafted features that outperform commercial product capabilities. The new al-

gorithm can detect 16 different rhythm anomalies with an accuracy of 99.79% with a 0.15% false alarm rate and a 99.74% sensitivity. In addition, the same algorithm can also detect 13 heartbeat anomalies achieving 99.18% accuracy, 0.45% false alarm rate, and 98.80% sensitivity.

Preface

This thesis is an original work by Hongzu Li.

Chapter 4 of this thesis has been published as Li, H.; Boulanger, P. "An Automatic Method to Reduce Baseline Wander and Motion Artifacts on Ambulatory Electrocardiogram Signals". Sensors 2021, 21, 8169.

Chapter 5 of this thesis has been published as Li, Hongzu, and Pierre Boulanger. "Structural Anomalies Detection from Electrocardiogram (ECG) with Spectrogram and Handcrafted Features." Sensors 22.7 (2022): 2467.

Even when walking in a party of no more than three, I can always be certain of learning from those I am with.

– Confucius, 475–221 BC (Translater: Arthur Waley, 1938)

Acknowledgements

I want to acknowledge and give my warmest thanks to my supervisor Dr. Pierre Boulanger who made this work possible. His guidance and advice carried me through all the stages of writing my project. I would also like to thank my committee members for letting my defense be an enjoyable moment and for your thoughtful comments and suggestions.

I would also like to give special thanks to my parents and my wife for their continuous support and understanding when undertaking my research and writing my project.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation

Cardiovascular diseases (CVDs) are the number one cause of death globally. An estimated 17.9 million people died from CVDs in 2019, representing 32% of all global deaths [66]. An electrocardiogram (ECG) device can record a patient's heart electrical signals activities over a long period [111] by measuring voltages from electrodes attached to the patient's chest, arms, and legs. ECGs are a quick, safe, and painless way to check for heart rate, heart rhythm, and signs of potential heart disease. Twelve leads ECG is today's standard tool used by cardiologists for detecting various cardiovascular abnormalities. However, a standard 12-lead ECG recording may not capture the heart problems in 10 seconds. Therefore, long-term ECG monitoring that tracks the patient's heart condition at all times and under any conditions has become possible with new sensing technologies. Recording devices such as Apple Watch [37], QardioMD [39], Vivalink [40], and the Astroskin Smart Shirt [38] are revolutionizing cardiac diagnostics by measuring 24/7 patient's cardiac activities and transmitting this information to a cloud service to be stored and processed remotely. In this context, there is a strong need for new automatic and reliable heart anomalies detection algorithms that can assist doctors in coping with this massive data set to perform diagnostics.

This thesis aims to develop automated heart anomaly detection algorithms to analyze a patient's entire ECG recording, compensate for motion artifacts and baseline wander during recording, and highlight events that correspond

to abnormal heart activities. Using these automated detection algorithms, cardiologists can then analyze them in context and proceed with their normal diagnostic process. This thesis makes two main contributions to solving these problems:

- Reduction of Motion Artifacts and Baseline Wander Using Sensor Fusion and Adaptive Filtering;

- Novel Neural Network Architecture to Detect Mobile ECG Signal Anomalies.

1.1.1 Reduction of Motion Artifacts and Baseline Wander Using Sensor Fusion and Adaptive Filtering

Personal health care devices possess many advantages, such as ease of use, long-term body monitoring, and easy access to personal health data. However, it also has some critical limitations. The massive data set produced by these sensors contains noise that does not exist when the patient is at rest. Most mobile healthcare devices use a rubber band to press the sensor electrodes onto the patient's chest. Therefore, the sensor is more likely to be affected by body motion during the patient's daily activities. These motion noises have two main categories: baseline wander (Figure 1.1) and motion artifacts (Figure 1.2). These artifacts distort the ECG signals that have nothing to do with the heart functions creating false alarms and making it challenging to automate classifications. In some cases, artifacts could even mimic true arrhythmia, leading to false diagnostics [32]. This thesis proposes a new noise removal method that automatically suppresses the baseline wander and motion artifact. The new algorithm consists of an adaptive empirical mode decomposition and reconstruction algorithm capable of removing baseline wander and some motion artifacts, followed by an adaptive filter algorithm that uses an accelerometer to reduce further the motion artifacts (see Figure 1.3). This algorithm shows a significant improvement compared to the conventional noise removal method. Two signal quality metrics are used to compare a reference ECG with its noisy version: correlation coefficients and mean squared error.
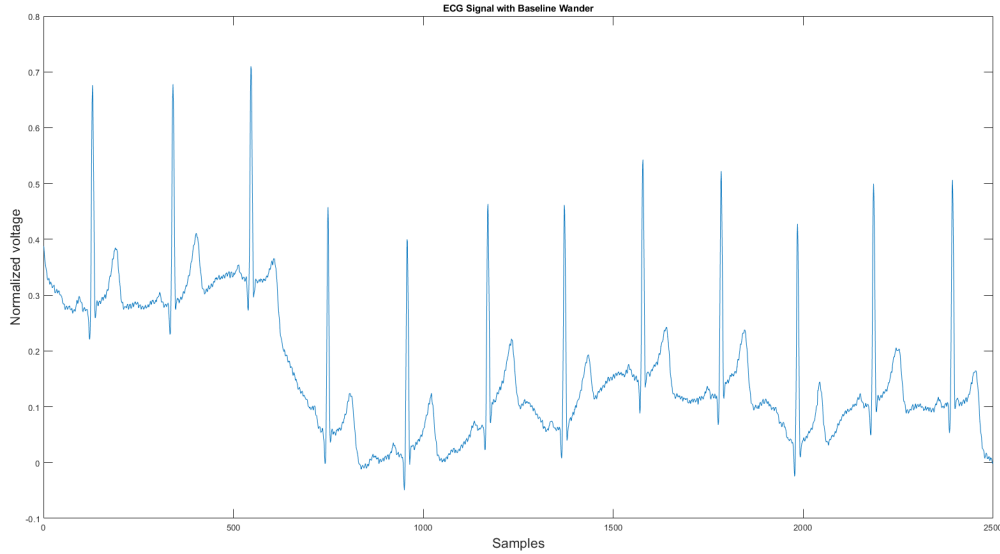
Figure 1.1: Baseline Wander affecting ECG Signal

For both metrics, the experimental results demonstrate that the noisy signal filtered by our algorithm can improve the signal-to-noise ratio by a factor of ten without significant signal distortion.

1.1.2 Novel Neural Network Architecture to Detect Mobile ECG Signal Anomalies

One of the main objectives of this thesis is to develop a cardiac monitoring system that could reduce motion artifacts and baseline wander and classify heart anomalies accurately during everyday activities (not only at rest). This patient-specific monitoring system will start by learning the user's ECG in a normal cardiac state, creating a neural network-based patient-specific model. Data acquired and compensated from artifacts created by body motion can then detect in real-time departures from normality.

The anomaly detection process from ECG signals contains rhythm and heartbeat classifications. Rhythm classification focuses on finding abnormal rhythms among normal rhythms. To find a rhythm anomaly, one must process multiple heartbeats. The heartbeat classification focuses on finding the pattern of one heartbeat signal [51]. This thesis proposes a new method that combines a Short-Time Fourier Transform (STFT) spectrogram of the ECG signal with
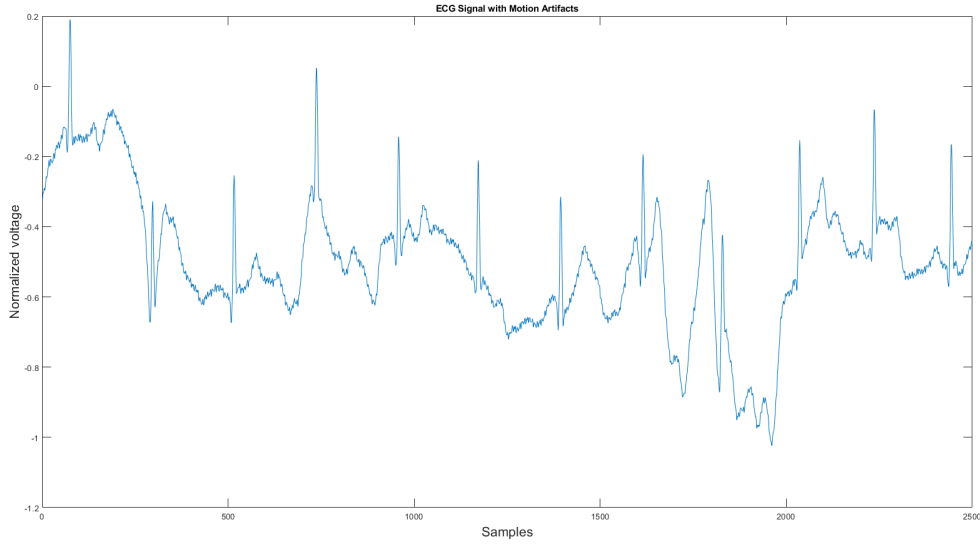
Figure 1.2: Motion artifact affecting ECG Signal

handcrafted features to detect heart anomalies beyond commercial product capabilities. Using the proposed Convolutional Neural Network, the algorithm can detect 16 different rhythm anomalies with an accuracy of 99.79% with 0.15% false alarm rate and 99.74% sensitivity. The same algorithm can also detect 13 heartbeat anomalies with 99.18% accuracy with 0.45% false alarm rate and 98.80% sensitivity. One can see in Figure 1.4 a block diagram of the new algorithm.

1.2   Thesis Contributions

This work makes the following novel contributions to the field:

- It proposed a new method to automatically reduce the baseline wander and suppress some motion artifacts in mobile ECG recording using sensor fusion and adaptive filtering;

- We demonstrate that an improvement of signal-to-noise-ration of ten is possible for ECG that is contaminated by motion artifacts;

- It proposed a novel neural network architecture using a Short-term Fourier Transform spectrogram and handcrafted features to detect mobile ECG

4
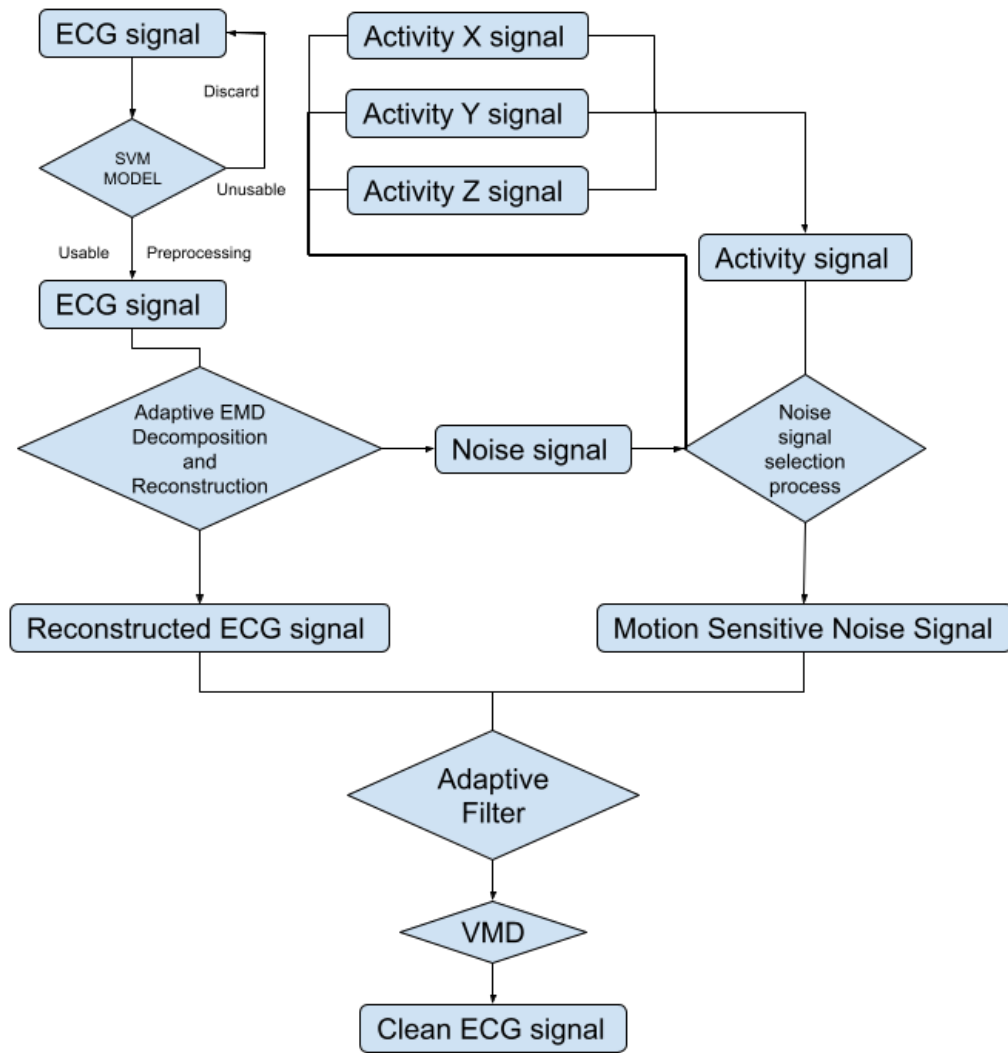
Figure 1.3: Block diagram of the proposed motion artifacts reduction algorithm

Proposed CNN
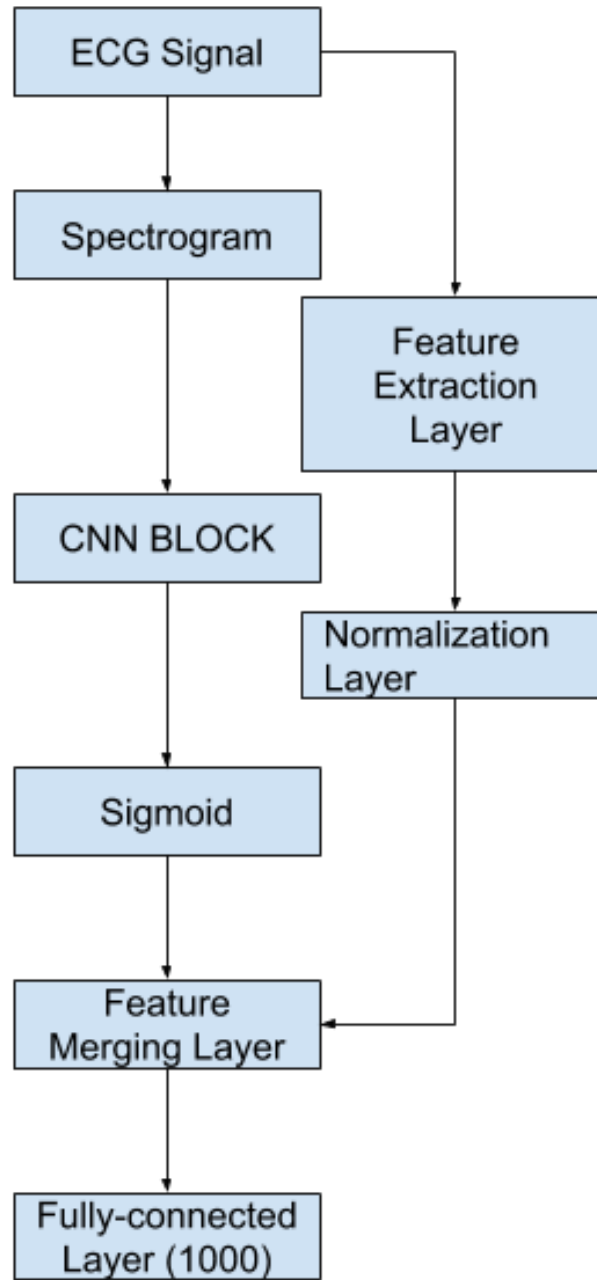
Figure 1.4: Block diagram of the proposed anomaly detection algorithm

signals with an accuracy of 99%;

- It shown the Experimental proof of the efficiency of the algorithms on real signals from commercial sensors such as Atroskin and VIVALNK.

The work of the thesis was published in top peer-review journals:

- Li, Hongzu, and Pierre Boulanger. "A survey of heart anomaly detection using ambulatory Electrocardiogram (ECG)." Sensors 20.5 (2020): 1461. https://www.mdpi.com/1424-8220/20/5/1461 Impact Factor: 3.576.

- Li, Hongzu, and Pierre Boulanger. An Automatic Method to Reduce Baseline Wander and Motion Artifacts on Ambulatory Electrocardiogram Signals. Sensors. 2021; 21(24):8169. https://www.mdpi.com/1424-8220/21/24/8169 Impact Factor: 3.576.

- Li, Hongzu, and Pierre Boulanger. Anomaly Detection from Electrocardiogram (ECG) with Spectrogram and Handcrafted Features.Sensors 2022, 22(7), 2467. https://www.mdpi.com/1424-8220/22/7/2467 Impact Factor: 3.576.

1.3   Dissertation Outline

Chapters 2 and 3 discuss the background knowledge of the ECG signal analysis and prior research relevant to this project. Chapter 4 describes and evaluates the new algorithm to reduce the effects of motion artifact and baseline wander. Chapter 5 explains and evaluates the heart anomaly detection algorithm. Finally, chapter 6 summarizes the findings of the thesis, describing the limitations of the methods, and suggesting paths for further research.

Chapter 2

Review of ECG Analysis

2.1 Normal ECG Signals

To detect anomalies in ECG signals, one must first know how a normal heartbeat looks. In [31], a regular rhythm (see Figure 2.1) is defined as the result of an electrical impulse that starts from the sinoatrial (SA) node that propagates through the heart muscles and then to the patient's chest. A normal rhythm is composed of the following segments: P wave generated by the atrial depolarization, the QRS complex generated by the ventricular depolarization, T wave and U wave generated by ventricular re-polarization. The P wave, QRS complex, and T wave should be similar over time at a frequency ranging from 50 to 100 bpm in normal ECG signals. A normal ECG signal should have PR intervals within 0.12-0.2 seconds and QT intervals less than half of the corresponding RR interval. In addition, the variation between the shortest PP/RR interval and the longest PP/RR interval should be less than 0.04 seconds (see Figure 2.2).
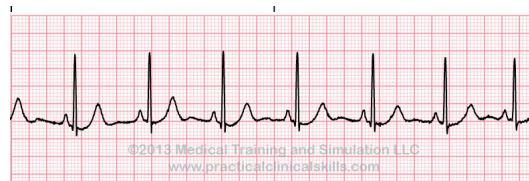
Figure 2.1: Normal Sinus Rhythm (NSR)

8
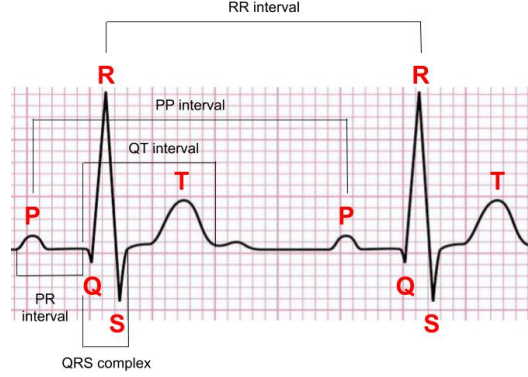
Figure 2.2: Normal ECG signal and corresponding notation

2.2 Abnormal ECG Signals

There are two kinds of anomalies in ECG signals: irregular rhythm and irregular heartbeat. Irregular heart rate indicates that the PP/RR intervals are larger than 0.04 seconds, or the PP/RR intervals are too short or too long. If the heartbeats start from the SA node, but the PP/RR intervals are too long, this may indicate *Sinus Bradycardia* (Figure 2.3(a)), which means that the heart is pumping too slow. And when PP/RR intervals are too short, this may be the sign of *Sinus Tachycardia* (Figure 2.3(b)). Moreover, if the variation between PP/RR intervals are too large, it may indicates *Sinus Arrhythmia*, *Sinus Block*, and *Sinus Arrest* (Figure 2.3(c,d,e)). These ECG anomalies may indicate the patient's current conditions. For instance, *Sinus Bradycardia* may be associated with hypothyroidism, hyperkalemia, sick sinus syndrome, sleep apnea syndromes, carotid sinus hypersensitivity syndrome, and vasovagal reactions. *Sinus Tachycardia* is commonly associated with anxiety, excitement, pain, drugs reaction, fever, congestive heart failure, pulmonary embolism, acute myocardial infarction, hyperthyroidism, pheochromocytoma, intravascular volume loss, and alcohol intoxication or withdrawal. *Sinus Block*, and *Sinus Arrest* can be caused by Hypoxemia, Myocardial ischemia or infarction, digitalis toxicity, and Toxic response to drugs [29].

Ectopic rhythms start from a source other than the sinus node. For example, *Atrial Rhythm* begins in the atria, and In this case, the P wave is shaped differently from the P wave beginning in the SA node. There are several abnor-

9
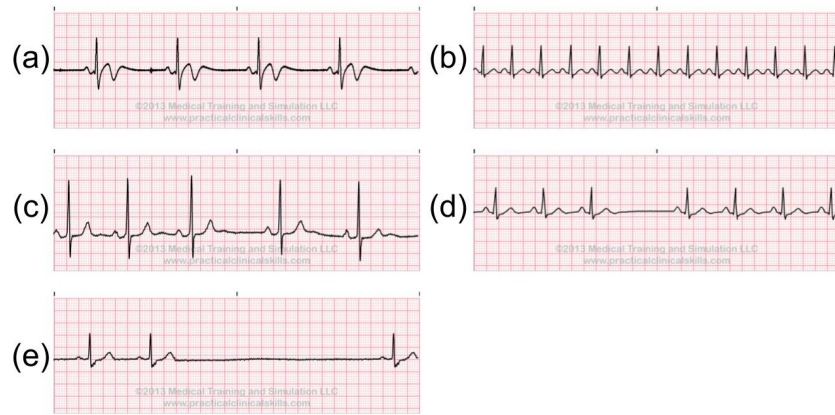
Figure 2.3: Abnormal Sinus Rhythms: (a) Sinus Bradycardia (b) Sinus Tachycardia (c) Sinus Arrhythmia (d) Sinus Block (e) Sinus Arrest

mal rhythms could occurs when the Atria is firing the heartbeat: *Premature Atrial Contraction*, *Wandering Atrial Pacemaker*, *Atrial Tachycardia*, *Atrial Flutter*, *Atrial Fibrillation*. Examples are shown in Figure 2.4. The *Premature Atrial Contraction* is a ubiquitous beat caused by emotional stress, excessive intake of caffeine, and hyperthyroidism. If *Premature Atrial Contraction* consecutively occurs three or more times, the rhythm is *Atrial Tachycardia*. It may cause light-headiness or even fainting. *Atrial Flutter* and *Atrial Fibrillation* are two distinct but closely related tachyarrhythmias. They could lead to many symptoms such as palpitations, light-headiness, fainting, angina, and congestive heart failure.

Junctional Rhythms happens when the atrioventricular (AV) junction paces the heart. In such a case, the P wave on the ECG signal may disappear or become negative. There are several anomaly examples showed at Figure 2.5: *Premature Junctional Complex*, *Junctional Escape Rhythm*, *Junctional Tachycardia*. The *Premature Junctional Complex* usually has the same cause as the *Premature Atrial Contraction* stated above. *Junctional Escape Rhythm* could be triggered by sick sinus syndrome, digitalis toxicity, excessive effects of beta-blockers or calcium channel blockers, acute myocardial infarction, hypoxemia, and hyperkalemia. One of the most common anomalies in the *Junctional Tachycardia* is the *Atrioventricular nodal re-entrant tachycardia (AVNRT)*. It is an arrhythmia that results from a rapidly recirculating impulse in the nodal
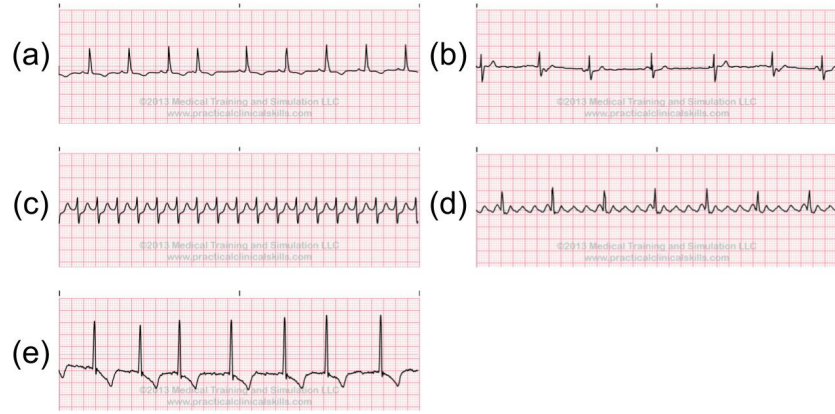
Figure 2.4: Abnormal Atrial Rhythms: (a) Premature Atrial Contraction (b) Wandering Atrial Pacemaker (c) Atrial Tachycardia (d) Atrial Flutter (e) Atrial Fibrillation
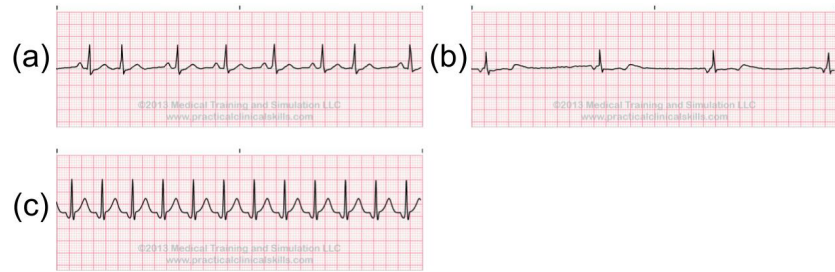
Figure 2.5: Abnormal Junctional Rhythms: (a) Premature Junctional Contraction (b) Junctional Escaped Rhythm (c) Junctional Tachycardia [92]

part of the AV junction, and it is caused by digitalis toxicity [29].

Ventricular Rhythms occurs when an ectopic site within a ventricle assumes responsibility for pacing the heart. As a result, the ventricular beats and rhythms usually have QRS complexes with abnormal shapes and longer lengths. The following are the examples of abnormal ventricular rhythm: *Premature Ventricular Contraction*, *Ventricular Escaped Rhythm*, *Accelerated Idioventricular Rhythm*, *Ventricular Tachycardia*, and *Ventricular Fibrillation*, *Ventricular Asystole*. One can see in Figure 2.6 the ECG signals. Individuals with *Premature Ventricular Contraction* may be the marker of severe organic heart disease associated with an increased risk of cardiac arrest and sudden death from *Ventricular Fibrillation*. *Ventricular Tachycardia* consists of three or more consecutive *Premature Ventricular Contraction*, and it could lead to

11
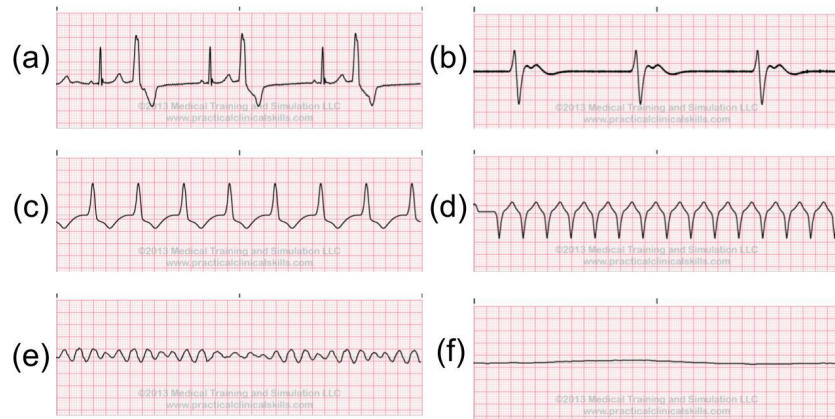
Figure 2.6: Abnormal Ventricular Rhythms: (a) Premature Ventricular Contraction (b) Ventricular Escaped Rhythm (c) Accelerated Idioventricular Rhythm (d) Ventricular Tachycardia (e) Ventricular Fibrillation (f) Ventricular Asystole

more life-threatening *Ventricular Fibrillation*. With *Ventricular Fibrillation*, the ventricles do not beat in any coordinated fashion, but instead, they fibrillate or quiver asynchronously. It will cause the patient to become unconscious immediately [29].

Since depolarization and re-polarization are slow in the atrioventricular (AV) node, this area is vulnerable to blocks in conduction. Therefore, when a delay or interruption happens during impulse conduction from atria to the ventricle, *AV blocks* may occurs. *AV blocks*, also call *Heart blocks*, is classified into: *First-degree AV blocks* ; *Second-degree AV blocks(types I and II)*; *Third-degree AV blocks (complete)* see in Figure 2.7. Among the heart blocks, the lower degree heart blocks could lead to *Third-degree AV blocks*, also called *Complete Heart block*, which is the most severe heart anomaly. With the *Complete Heart block*, the atria and ventricle are pacing independently which could slow down the ventricular rate, and eventually lead to fainting [29].

2.3 Difference Between 12-lead ECG vs 3-lead ECG

A standard 12-lead electrocardiogram provides views of the heart in both the frontal and horizontal planes and views the surfaces of the left ventricle
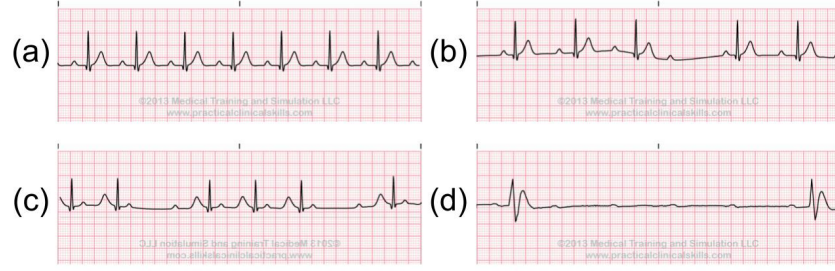
Figure 2.7: AV Blocks: (a) First-degree AV blocks (b) Second-degree AV blocks types I (c) Second-degree AV blocks types II (d) Third-degree AV blocks

from 12 different angles. 12 lead ECG have six limb leads: I, II, III, aVF, aVL, aVR, and six chest leads: V1-V6. The standard 12-lead ECG is used as a clinical dysrhythmia analysis tool for chest pain or discomfort, electrical injuries, electrolyte imbalances, medication overdoses, ventricular failure, stroke, syncope, and unstable patient. It is widely used in clinics and hospitals for heart disease diagnosis [31]. However, the 12-lead ECG is impractical when monitoring the patients continuously for 24 hours since the patient needs to be attached to 10 electrodes to record a 12-lead ECG. Instead, a 3-lead ECG system is a replacement for 12-lead ECG in this situation. In addition, many researchers have shown that 3-lead ECG is useful for making a valid diagnosis. Antonicelli *et.al* [4] was able to validate the accuracy of 3-lead tele-ECG compared to 12-lead tele-ECG in older population . Their study shows a high concordance between the ECG diagnosis using a simple home telecardiology device (3-lead tele-ECG) and more complex instruments like the 12-lead tele-ECG and the standard 12-lead ECG. The study also demonstrated that a simple 3-lead tele-ECG could detect cardiac alteration, like arrhythmias, atrioventricular blocks, and re-polarization abnormalities with good agreement with observations measured by 12-lead tele-ECG and the standard 12-lead ECG. Kristensen *et.al* [50] also evaluated how well an inexpensive portable three-lead ECG monitor can detect patients with atrial fibrillation (A.Fib) compared to a standard 12-lead ECG. Their study shows that the sensitivity of diagnosing A.Fib using PEM recordings was 86.7% and the specificity was 98.7% compared to a 12-lead ECG. According to cardiologists, the misclas-

sification of three PEM recordings was due to interpretation errors and not related to the PEM recording. Their article concluded that portable PEM devices could be used to diagnose A.Fib.

Furthermore, many researchers have tried to reconstruct a 12-lead ECG signal from a 3-lead signal. Atoui *et.al* [5] have introduced a neural network-based model that could synthesize standard 12-lead ECG from serial 3-lead ECG. As a result, the synthesized 12-lead ECG from the ANN model has an average of 0.93 correlation coefficient compared to the actual 12-lead ECG. In addition, Nelwan *et.al* [64] and Drew *et.al* [23][24] have done several studies that shows it is possible to reconstruct standard 12-lead ECG from reduced lead set ECG. This research indicates that the reduced ECG lead set, such as 3-lead ECG, could contain enough information to detect most heart anomalies.

Chapter 3

Review of ECG Processing and Classification

As mentioned previously, this thesis contains two major parts, one is motion noise compensation on ECG signal, and the other one is anomaly detection. Motion artifacts compensation aims to remove the artifacts and baseline wander generated by body movement on the ECG signal. On the other hand, anomaly detection on ECG signals involves finding the irregular heart rate, heartbeat, and rhythm from the motion-compensated ECG signal. This thesis has separated anomaly detection into heartbeat detection, rhythm classification, and heartbeat classification. The heartbeat detection algorithm finds all heartbeat's R peak locations on the ECG signal to calculate the RR interval. Then the heart rate and heart rate variability are calculated using RR intervals. The rhythm classification shows the anomaly type of the heart rhythm. Heartbeat classification detects all the abnormal heartbeats on the ECG signal. This chapter reviews and discusses the related research and conventional methods found in the literature.
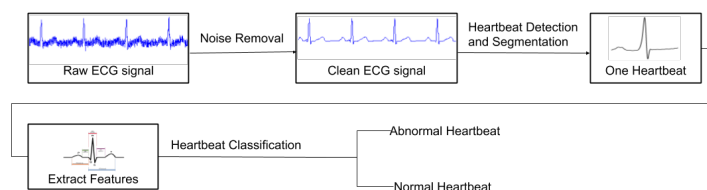
Figure 3.1: Typical Heartbeat Anomaly Detection

3.1 Motion Artifact and Noise Reduction

There are five conventional methods for reducing the effects of baseline line wander and motion artifacts [51] that one can find in the literature. In the following sub-sections, we explain each technique and its implementations. Moreover, we will discuss the advantage and disadvantages of each method.

3.1.1 Digital Filters

Digital filters could remove the unwanted frequency components in signals. The combination of different digital filters is called a filter bank. It combines various digital filters such as low-pass filters, high-pass filters, band-pass filters, notch filters, differentiators and normalizers. The ECG signal frequency component without noise range from 0.1 to 100 Hz [48]. Therefore, low bound and high bound filters must be applied according to the based ECG signal's frequency range to extract the pure ECG signal information. One of the filter bank advantages is that it is straightforward to implement and is highly efficient. However, its drawback is that fixed digital filters could introduce nonlinear phase distortion and key point displacement, which could cause the deformation of essential waves of the ECG signal, such as the QRS complex [55]. Digital filter parameters are explicitly tuned for one dataset. Therefore, it may not work well for the data collected from other devices. Also, digital filters do not perform well at removing the noises similar to the ECG signal.

3.1.2 Discrete Wavelet Transform

Mallat first introduced Discrete Wavelet Transform (DWT) in 1989 [59]. It can be used to analyze non-stationary signals, such as ECG [95]. In [74], Poungponsri combines the DWT with an artificial neural network to remove a wide range of noise. The commonly used mother wavelet basis functions are Daubechies filters (Db), Symmlet filters (Sym), Coiflet filters (C), Battle-Lemarie filters (Bt), Beylkin filters (Bl), and Vaidyanathan filter (Vd)[85]. According to studies in [76][85][54], the Daubechies filters of order 4 and 8 (Figure 3.2), and the Symmlet filters of order 5 and 6 (Figure 3.3) are the best

16
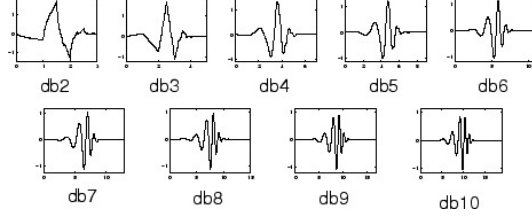
Figure 3.2: Daubechies wavelets
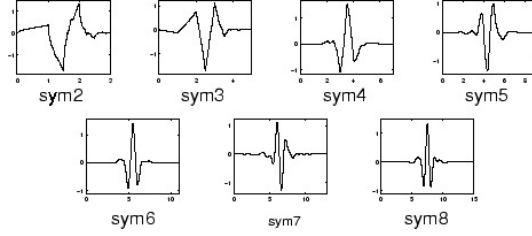
Figure 3.3: Symlets wavelets

wavelet functions for ECG signal analysis due to their similar signal structure to the QRS complex. After decomposing the ECG signal, a threshold method is applied to the DWT coefficients. A clean ECG signal could be reconstructed from the thresholded DWT coefficients. DWT decomposes the signal into detail coefficients to approximate the signal using a wavelet mother function. Each coefficient corresponds to a different frequency range. Therefore, by observing the extracted approximation and detail coefficients, one can easily distinguish the coefficients that contain the ECG signal and the coefficients that contain the signal noise. One could only apply noise removal methods such as a digital filter to the noisy coefficients as the coefficients with the ECG signal will not be affected. Despite the significant advantages of DWT, some drawbacks make it hard to use in practice. The first DWT requires setting and selecting the number of coefficients manually. Second, it is hard to find the most suitable wavelet mother function for different ECG signals. These limitations make it hard to automate such an algorithm.

3.1.3 Empirical Mode Decomposition

Empirical Mode Decomposition (EMD) is an adaptive and entirely data-driven technique that obtains the oscillatory modes present in the data [36]. Huang first introduced it in 1998. Similar to DWT, EMD decomposes the

17

ECG signal into many sub-band signals. The high-frequency components of the EMD decomposition are called Intrinsic Mode Functions (IMF), and the low-frequency components are called residual. Unlike DWT, which relies on the mother wavelet functions, EMD does not need to find the best wavelet function or set the number of IMFs since this method is entirely data-driven. However, the IMFs must satisfy two conditions:

- The number of extremas and zero-crossings must be equal or differ at most by one;

- All local maxima and minima must be symmetric to zero.

In 2006, Weng applied the EMD method to the ECG signal to remove high-frequency noise with minimum signal distortion [96]. In 2008, Blanco proved that the EMD method could be used for baseline wander reduction [9]. However, the number of IMFs may differ for different ECG signals, implying that the IMFs do not have a fixed frequency range. Therefore, human intervention is often required to separate clean and noisy IMFs for signal reconstruction. The EMD method decomposes the signal to n IMFs and one residual signal. Let $c(t)$ be IMFs, we will have $c_1(t)$ to $c_n(t)$ from higher frequency components to lower frequency components. Then digital filters or thresholds are applied to the IMFs that contain the noise. After processing, the signal is reconstructed using the following equation:

$$x(t) = \sum_{i=1}^{n} c_i(t) + r(t) \tag{3.1}$$

where $x(t)$ is the reconstructed signal, $c(t)$ is the IMFs, $r(t)$ is the residual. In [96] [9] [12], the authors performed EMD on MIT-BIH database to suppress the high frequency noise and the baseline wander. Ensemble Empirical Mode Decomposition (EEMD) [100] fixed EMD's shortcoming of mode mixing. The mode mixing can cause serious aliasing in the time-frequency distribution and make the physical meaning of individual IMF unclear. The EEMD adds one extra step compared to the EMD. By adding white noise to the original signal before decomposing the signal into IMFs using EMD. Many noise removal works can be found using the EEMD, such as [43] [107] [86].

3.1.4 Variational Mode Decomposition

Dragoniretskiy and Zosso proposed the Variational Mode Decomposition (VMD) in 2013 [22]. It is an alternative approach to the empirical mode decomposition. Unlike the EMD method, the VMD method decomposes the signal from low frequency to high frequency, where the residual contains the highest frequency component. Therefore, when dealing with the high-frequency noise in the signal, the VMD method performs better than the EMD method [58]. In 2016, Mohan applied the VMD method to reduce the effects of powerline interference on the ECG signal. However, since the VMD method is a variant of the EMD algorithm, human intervention is also required for clean IMF selection.

3.1.5 Adaptive Filter

Adaptive Filtering (AF) takes the original noisy signal and a reference noise signal as input. It then automatically adjusts the filter weights based on the reference noise signal to improve the signal reconstruction. There are three types of AF: Least Mean Square (LMS) filter, Normalized Least Mean Square (NLMS) filter, and Recursive Least Squares (RLS) filter. In 1991, Thakor *et al.* [89] first introduced the LMS adaptive filter to remove the baseline wander, 60 Hz power line noise, muscle noise, and motion wander. In their approach, they proposed two adaptive filter structures. The first one is that the primary input is $s_1 + n_1$, while the reference input is noise n_2, which is recorded from another generator that is correlated with n_1. The second one is that an ECG is recorded from several electrode leads, the primary input is $s_1 + n_1$ from one of the leads, and the reference input is S_2 from another lead that is noise-free. The signal s_1 can be extracted by recursively minimizing the Mean-squared Error (MSE) between the primary and the reference inputs. The MSE is calculated as:

$$E[\epsilon^2] = E[(s_1 - y)^2] + E[N_1^2]. \tag{3.2}$$

The LMS algorithm was used to minimize the MSE. The LMS algorithm is written as:

$$W_{k+1} = W_k + 2\mu\epsilon_k X_k \qquad (3.3)$$

where W_k is a set of filter weights at time k, X_k is the input vector at time k of the samples from the reference signal, $\epsilon =$ primary input d_k- filter output y, parameter μ is empirically selected to produce convergence at a desired rate. The error ϵ_k can be calculated as:

$$\epsilon_k = d_k - y_k \qquad (3.4)$$

where d_k is the desired primary input from the ECG to be filtered and y_k is the filter output that is the best least-squares estimate of d_k.

However, one major drawback of the LMS adaptive filters is that they are sensitive to the scaling of their input. In order to overcome the drawback, a power normalized least mean squares(NLMS) was introduced [33]. The NLMS filter solves this problem by normalizing the input.

Another convention adaptive filter type is the RLS adaptive filter. The RLS algorithm has excellent performance when working in time-varying environments but at the cost of increased computational complexity and some stability issues [90]. Adaptive filters work well when the noise is limited to a fixed frequency range. However, if the noise's frequency changes, such as those created by motion wander, these methods may not work. Using a better reference noise signal sensitive to motion, we will show that one can design an adaptive filter that can automatically cancel the noise for different ECG frequencies. An adaptive filtering approach could remove baseline wander, motion artifacts, powerline interference, and muscle noise; however, it requires a reference input correlated to the original noisy input. Obtaining a clean ECG signal is very difficult to acquire. Due to the added complexity of the data collection, many researchers have considered using an accelerometer as the reference noise signal for the adaptive filter. Raya *et. al* [81] explored the possibility of using both signal axis and dual-axis accelerometer signal as the noise reference input to an LMS adaptive filter and an RLS adaptive filter. The RLS algorithm recursively finds the filter coefficients that minimize the
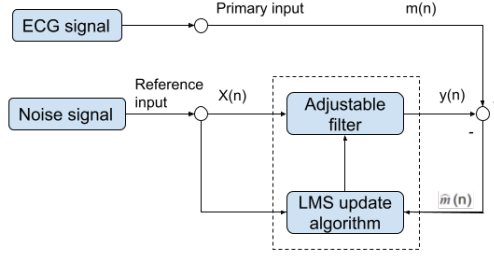
Figure 3.4: Block diagram of a generic adaptive filter

input signal's weighted least-squares cost function. It is known for its excellent performance when working in time-varying environments but at the cost of increased computational complexity and some stability problems [90]. The algorithm updates the filter weight vector using the following equations:

$$w(n) = \overline{w}^T(n-1) + k(n)\overline{e}_{n-1}(n), \tag{3.5}$$

$$u(n) = \overline{w_\lambda^{-1}}(n-1)x(n), \tag{3.6}$$

$$k(n) = u(n)/(\lambda + x^T(n)u(n)), \tag{3.7}$$

where $w(n)$ is the weights vector of iteration n, $x(n)$ is the input signal, λ is a small positive constant very close to but smaller than 1.

The filter output $\overline{y}_{n-1}(n)$ and the error signal $\overline{e}_{n-1}$ is calculated using the filter tap weights of the previous iteration and the current input vector as in the following equations:

$$\overline{y}_{n-1}(n) = \overline{w}^T(n-1)x(n), \tag{3.8}$$

$$\overline{e}_{n-1} = d(n) - \overline{y}_{n-1}(n). \tag{3.9}$$

As a result, the RLS adaptive filter outperforms the LMS adaptive filter. Using an accelerometer signal shows better results than using a dual-axis accelerometer signal. We will demonstrate that using one-axis reference input, particularly the y-axis, is sufficient to minimize the motion artifacts.

3.1.6 Heartbeat Detection and Segmentation

Heartbeat detection is often related to detecting irregular heart rate and inconsistent RR-intervals explained in Chapter 2. Heartbeat detection is also

21
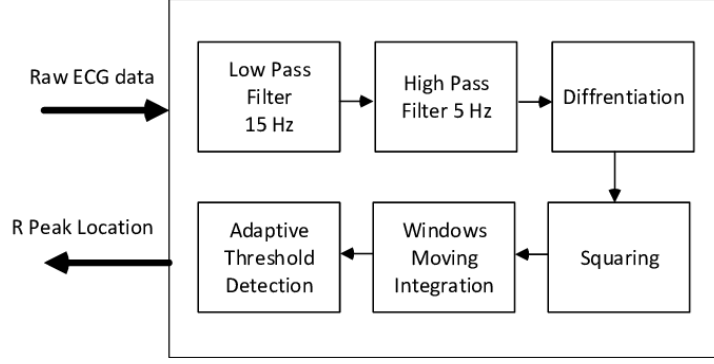
Figure 3.5: Pan Tompkins Algorithm

critical in extracting the heartbeats from the ECG signal for classification. Heartbeat detection consists of three main parts: P wave detection, QRS complex detection, and T wave detection. Therefore, it is usually related to heartbeat segmentation. Heartbeat segmentation usually means segmenting a heartbeat from its start point (onsite) of the P wave to its endpoint (offsite) of the T wave. However, the P wave and T wave may not be detectable in certain types of abnormal heartbeat, and the QRS complex is the most apparent waveform. Thus the location of the QRS complex is often used to locate the origin of the heartbeat. Many algorithms detect the R peak location in the QRS complex.

The Pan-Tompkins algorithm [69] is one of the most popular and earliest algorithms implemented (Figure 3.5). It is widely used in many applications because of its robustness and computational efficiency. The algorithm uses a filter bank that consists of band-pass filters, differentiator, squaring filter, and moving window integrator to reduce signal noise so that only R wave information is present. Inspired by the Pan-Tompkins algorithm, many researchers such as [2] [15] [16] [104] developed their own filter banks to improve the accuracy of the detection. In order to reduce the false positives, used predefined amplitude proposed by [15] and [104], and predefined RR interval length proposed by [16] and [104] are used as the threshold to remove unwanted detection.

Zidelmal *et.al* [109] introduced a QRS detection method based on wavelet decomposition. In the algorithm, the authors decompose the raw ECG signal
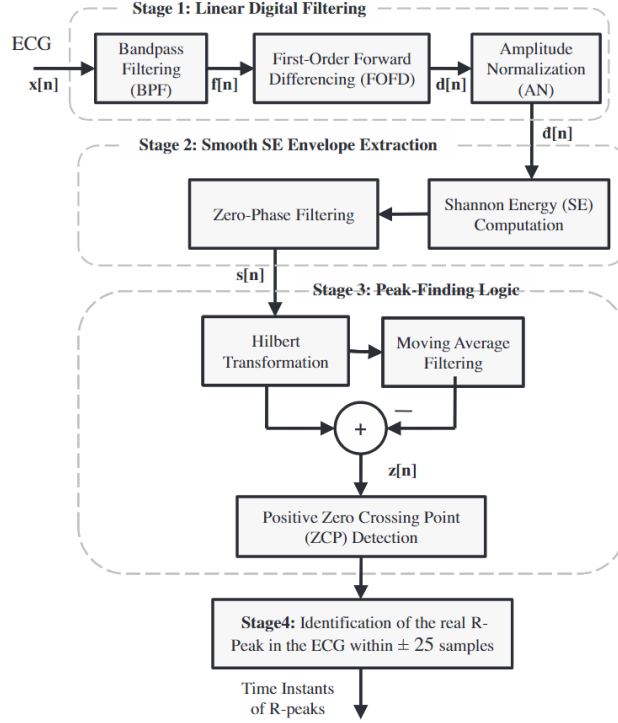
Figure 3.6: SEEHT

using discrete wavelet transform, then reconstruct the signal by selecting only the ECG information's sub-signals. A threshold is set to select the large amplitude peaks to detect a QRS complex. Similar works could be found in [10].

Manikandan *et.al* [60] introduces a new algorithm that uses Shannon Energy Envelop and Hilbert-Transform (SEEHT) (see Figure 3.6) to detect the QRS complex location . A band-pass filter is applied to the raw ECG signal to remove the baseline wander and high-frequency noise in the preprocessing stage. Then, a differentiator and normalizer were applied to clean the signal to highlight the QRS complex components. The Shannon energy of the processed signal is calculated using the following equation:

$$s[n] = -d^2[n]log(d^2[n]), \tag{3.10}$$

where d[n] is the processed signal. A zero-phase filter then processes the calculated Shannon energy sequence to preserve sharp peaks around the QRS complex and smooth out the noisy peaks. A Hilbert transform is applied on

23

all candidate R peaks in the peak finding algorithm to get the R wave envelope. For each R wave envelope, the zero-crossing locations indicate an R peak. Inspired by SEEHT, [108] introduced PSEE that improves the computational inefficiency of the Hilbert transform by using both predefined amplitude threshold and predefined RR interval length threshold. ISEE [80] improved further the SEEHT and PSEE algorithms by using a filter bank consisting of a moving average filter, a differentiator, a normalizer and a squaring filter to eliminate the noisy peaks. The filter bank computational costs are less than the Hilbert transform and do not use a predefined threshold. Most recently, Park [70] combined discrete wavelet transform and ISEE to detect R peaks on the ECG signals.

The P and T waves represent important information, and heartbeat segmentation depends on the P and T wave detection. Therefore, a good detection of the P and T waves is critical for diagnosis. Pal and Mitra [68] proposed an algorithm that could detect the PQRST peak points. The algorithm uses discrete wavelet decomposition. It reconstructs the signal from selected wavelet coefficients related to R, QS, and PT. For example, when the algorithm detects R peaks, a signal is reconstructed with d3, d4, and d5 coefficients and preserves the information for R peaks but diminishes other peaks. A few years later, Banerjee [6] also developed a T wave and QRS complex detection algorithm based on discrete wavelet decomposition and adaptive thresholding. Karimipour [45] uses discrete wavelet transform and adaptive thresholding to detect the QRS complex location and give an estimate of the P-wave and T-wave locations. In practice, many researchers such as [19] [20] used 'ecgpuwave' detector from PhysioNet for the heartbeat segmentation[27]. However, the P and T wave detection works well with normal heartbeats but not for many abnormal heartbeat types. Many researchers choose manual annotation such as [103] or a fixed window such as [102] [56] [103] [91] [78] for their heartbeat segmentation. In Table 3.1, one can compare the performance of some of the heartbeat detection algorithms that have been tested on the MIT-BIH Arrhythmia database [63]. The metrics used to compare each algorithm are as following:

Table 3.1: Heartbeat detection performance on MIT-BIH dataset

Method	Year	Total heartbeats	TP	FP	FN	SEN	+P	DER	ACC
Pan Tompkins [69]	1985	116137	115860	507	277	99.76%	99.56%	0.68%	99.33%
FBBBD [2]	1999	91283	90909	406	374	99.59%	99.56%	0.86%	99.15%
S.W.Chen [15]	2006	102654	102195	529	459	99.55%	99.49%	0.97%	99.04%
DOM [104]	2008	116137	115971	58	166	99.86%	99.95%	0.19%	99.81%
S.Choi [16]	2010	109494	109118	218	376	99.66%	99.80%	0.54%	99.46%
Z.Zidelmal [109]	2012	109494	109101	193	393	99.64%	99.82%	0.54%	99.47%
SEEHT [60]	2012	109496	109417	140	79	99.93%	99.87%	0.2%	99.80%
S.Banerjee [6]	2012	19140	19126	20	20	99.90%	99.90%	0.21%	99.79%
PSEE [108]	2013	109494	109401	91	93	99.92%	99.92%	0.17%	99.83%
F.Bouaziz [10]	2014	109494	109354	232	140	99.87%	99.79%	0.34%	99.66%
A.Karimipour [45]	2014	116137	115945	308	192	99.83%	99.74%	0.43%	99.57%
ISEE [80]	2016	109532	109474	116	58	99.95%	99.89%	0.16%	99.84%
WTSEE [70]	2017	109494	109415	99	79	99.93%	99.91%	0.16%	99.84%

- TP: Number of correctly detected heartbeat;

- FP: Number of incorrectly detected heartbeat;

- FN: Number of missed heartbeat;

- Sensitivity (SEN) = TP / (TP+FN);

- Positive Detection (+P) = TP / (TP+FP);

- Detection Error Rate(DER) = (FP+FN) / TP;

- Accuracy (ACC) = TP / (TP+FP+FN);

3.1.7 Heartbeat Classification

Heartbeat classification takes an ECG signal that only includes one heartbeat as the input. Therefore, it outputs the heartbeat type such as normal heartbeat, left/right bundle branch block beats, premature beats, etc. The heartbeat classification consists of two parts: feature extraction and classification tools. The feature extraction aims to extract useful information from the raw ECG signal and discard the rest, and the classification tools learn the pattern in the extracted information and predict.

Feature Extraction

According to the literature, there are two main conventional features: derived and morphological [51]. The derived features are calculated from the

Table 3.2: Conventional Morphological Features of Heartbeat

Features	Description	Reference
QRS complex duration	The time interval between the onsite of Q wave and offsite of S wave	[17][20][56] [42][106][82]
QRS velociy left	QRS slope velocity calculated for the time-interval between the QRS complex onset and the first peak	[17][42]
QRS velociy right	QRS slope velocity calculated for the time-interval between the first peak and the second peaks	[17][42]
QRS complex area	The sum of positive area and absolute negative area in QRS complex	[17][42]
QRS complex morphology	Sample points from the QRS onsite to the QRS offsite	[20]
QRS complex AC power	The total power content of the QRS complex signal	[91]
QRS complex Kurtosis	The kurtosis indicates the peakedness of the QRS complex	[91]
QRS complex Skewness	The skewness measure the symmetry of the distribution of the QRS complex	[91]
Q wave valley	The valley value of Q wave	[82]
S wave valley	The valley value of S wave	[82]
T wave peak	The peak value of T wave	[82]
T wave duration	The duration from QRS offsite to T wave offsite	[20]
T wave morphology	Sample points from the QRS offsite to the T wave offsite	[20]
P wave flag	A Boolean value indicates the presence or absence of the P wave	[20]
P wave duration	The duration from P wave onsite to P wave offsite	[106]
P wave morphology	Sample points from P wave onsite to P wave offsite	[106][87]
PR interval duration	The duration from P wave onsite to QRS complex onsite	[106]
PR interval morphology	Sample points from P wave onsite to QRS complex onsite	[87]
QT interval duration	The duration from QRS complex onsite to T wave offsite	[106]
QT interval morphology	Sample points from QRS complex onsite to T wave offsite	[87][82]
ST interval morphology	Sample points from S wave valley to T wave offsite	[82]
Max peak(R peak) value	The maximum amplitude of the heartbeat	[17][42][82]
Min peak value	The minimum amplitude of the heartbeat	[17][42]
Positive QRS complex area	The area of positive sample points in QRS complex	[17][42][106]
Negative QRS complex area	The area of negative sample points in QRS complex	[17][42][106]
Positive P wave area	The area of positive sample points in P wave	[106]
Negative P wave area	The area of negative sample points in P wave	[106]
Positive T wave area	The area of positive sample points in T wave	[106]
Negative T wave area	The area of negative sample points in T wave	[106]
Absolute velocity sum	Sum of the absolute velocities in the pattern interval	[17][42]
Ima	Time-interval from the QRS complex onset to the maximal peak	[17][42]
Imi	Time-interval from the QRS complex onset to the minimal peak	[17][42]
Pre-RR interval	The RR interval between the heartbeat and its previous heartbeat	[20][103][106]
Post-RR interval	The RR interval between the heartbeat and its following heartbeat	[20][103][106]
Post-PP interval	The PP interval between the heartbeat and its following heartbeat	[106]
Average-RR interval	The average value of all valid RR intervals in the ECG record	[20][103][106] [91] [82]
Local Average-RR interval	The average value of ten valid RR intervals surrounding the heartbeat	[20][103][106]
Normalized signal	The heartbeat sample points are normalized and down-sampled to have mean of zero and standard deviation of one	[18][53][52]
Raw/downsampled ECG signal	The unprocessed ECG signal or the only processing on the signal is downsampling	[14][48]

ECG signal using feature detectors. The common derived features found in the literature are: Auto-regressive (AR) coefficients [26]; Discrete Wavelet Transform (DWT) coefficients [67] [82]; Eigen Vectors [93]; Dual-Tree Complex Wavelet-Transform coefficients [91]; Intrinsic Mode Functions from Empirical Mode Decomposition [78]. The other morphological features deal with the morphological information found in the ECG signal. The most used morphological features are: QRS complex duration and peak values, P wave duration and peak value, T wave duration and peak value, PR interval length, QT interval length, and so on [106] [103]. In [51] one can see a detailed list of the most common morphological features used for ECG signal processing. There are many morphological features (see in Table. ??) that can be found in various research. Other features that are calculated directly from the ECG signal are:

- Vectorcardiography (VCG) vector;

- DWT coefficients produced by Discrete Wavelet Transform (DWT);

- Independent components from Independent Component Analysis (ICA);

- PCA components generated from Principal Component Analysis (PCA);

- IMFs from Empirical Mode Decomposition(EMD)/Ensemble EMD (EEMD);

- DTCWT coefficients from Dual Tree Complex Wavelet Transform;

- Eigenvector methods;

- Dynamic Time Warping (DTW) distance.

Vectorcardiography (VCG) is a part of the ECG analysis tools. It displays the various complexes of the ECG. In addition, it provides the possibility to use vector analysis on the cardiac electric potentials [11].

DWT decomposed the signal into many sub-signals(detail coefficients) with different frequency ranges than the one used for noise removal. The DWT method could remove unwanted noises and find features for the heartbeats since the heartbeat waves are much clearer in specific detail coefficients such as D4 and D5. Therefore, much research, such as [56] and [103], uses features from the detail coefficients to classify the heartbeat.

The conventional DWT technique lacks shift-invariance property due to the downsampling operation at each stage. Hence, the energy of the wavelet coefficient changes significantly for a small-time shift in the input pattern. The Dual-Tree Complex Wavelet Transform(DTCWT) [47] is a simple technique that overcomes the DWT shortcomings. The DTCWT uses two filters: one for level 1 decomposition and the other for the higher levels. In the first level decomposition, the original signal is decomposed into two trees, and each tree contains two sub-band signals. One tree could be interpreted as the real part of a complex wavelet, and the other three could be the imaginary part. For each tree, the conventional DWT is applied for further decomposition [91].

Thomas used the DTCWT method to extract heartbeat features to classify heartbeat variations [91].

Like DWT and DWCWT, the ICA, PCA, and EMD/EEMD decompose the signal into many sub-signals. The difference is that the ICA and PCA aim to reduce the input size to minimize the computation speed. EMD/EEMD features do not require knowledge of the level of scale and the basic function needed in DWT. The ICA method has been used in [103] to produce the independent components to be part of the heartbeat feature set. The PCA method used in [97] reduces the input size for higher efficiency. Rajesh *et. al* [78] computed heartbeat features from IMFs by applying the EMD/EEMD method to the ECG signal.

Eigenvector methods could estimate frequencies and powers of noise-corrupted signals. These methods used the eigen-decomposition of the correlation matrix of the noise-corrupted signal [93]. Ubeyli *et.al* [93] used three kinds of eigenvector methods to generate the feature set: Pisarenko uses Multiple Signal Classification (MUSIC), and Minimum-Norm (MN). The Pisarenko method is useful for estimating PSD, which contains sharp peaks at the expected frequencies. The MUSIC method is a noise subspace frequency estimator and could eliminate the effects of spurious zero on the noise subspace. The MN method aims to differentiate spurious zeros from real zeros, and it uses a linear combination of all noise subspace eigenvectors.

Dynamic Time Warping (DTW) measures the similarity between two heartbeat segments and computes the distance between these two heartbeat segments. Therefore, if one of the heartbeat segments is the sample heartbeat of a specific type and the other is the test heartbeat. DTW distance indicates the similarity score between the test and sample heartbeat. The similarity score could represent the heartbeat waveform features, such as in[106] and [18]. A detailed description of the features of each method can be seen in Table 3.3.

There are some advantages and drawbacks to both features. The derived features usually represent the ECG signal in the frequency domain. Therefore the noisy components in the ECG signal can be easily discarded during the process even though it may lose some essential information in the time domain,

Table 3.3: Conventional Derived Features of Heartbeat

Features	Method	Description	Reference
VCG amplitude	VCG	Maximal amplitude of the VCG vector	[17][56]
VCG sine angle	VCG	Sine component of the angle of the maximal amplitude vector	[17][56]
VCG cosine angle	VCG	Cosine component of the angle of the maximal amplitude vector	[17][56]
DTW distance	DTW	The Dynamic Time Warping distance between a heartbeat segment and the median heartbeat segment of the recording	[106] [18]
Positive peak of QRS complex	DWT	The positive peak amplitude of QRS complex on the 4th scale of the DWT	[56]
Negative peak of QRS complex	DWT	The absolute negative peak amplitude of QRS complex on the 4th scale of the DWT	[56]
Positive peak of T wave	DWT	The positive peak amplitude of T wave on the 4th scale of the DWT	[56]
Absolute T wave offsite	DWT	The absolute amplitude of T wave offsite on the 4th scale of the DWT	[56]
R-S interval distance	DWT	The relative distance between R peak and S valley on the 4th scale of the DWT	[56]
S-T interval distance 1	DWT	The relative distance between S valley to the T wave peak on the 4th scale of the DWT	[56]
S-T interval distance 2	DWT	The relative distance between S valley to the T wave offsite on the 4th scale of the DWT	[56]
Absolute maximum	DWT	The absolute maximum value and location on the 4th scale of DWT signal	[56]
Zero crossing	DWT	The zero crossing location on the 4th scale of DWT signal	[56]
Wavelet scale	DWT	Calculate which scale the QRS complex is centered on	[56]
DWT coefficients	DWT	The down-sampled 3rd and 4th detail coefficients and the 4th approximation coefficients	[103]
Independent Components	ICA	Independent components calculated with a fast fixed point algorithm	[103]
Fourier spectrum	DTCWT	Compute the absolute value of 4th and 5th scale DTCWT detail coefficients(dc). Then 1D FFT is applied to the selected dc to get the Fourier spectrum. And then take logarithm value of the Fourier spectrum	[91]
IMF sample entropy	EMD/EEMD	The sample entropy is measured of regularity of a time series used to quantify the complexity of heartbeat dynamics	[78]
IMF variation coefficient	EMD/EEMD	The coefficient of variation is a statistical parameter defined as σ^2/μ^2.[1]	[78]
IMF singular values	EMD/EEMD	The singular value decomposition	[78]
IMF band power values	EMD/EEMD	The band power is the average power of each IMF	[78]
PCA components	PCA	PCA components for size reduction	[97]
Pisarenko PSD	Eigenvector	Power spectral density estimates generated with Pisarenko method	[93]
MUSCI PSD	Eigenvector	Power spectral density estimates generated with Multiple signal classification method	[93]
Minimum-Norm PSD	Eigenvector	Power spectral density estimates generated with Minimum-Norm methods	[93]

[1] σ is the standard variation of the selected IMF, μ is the mean of the selected IMF

such as peak values, interval duration, etc. Moreover, the derived features may not distinguish between different rhythms or heartbeats. On the other hand, morphological features focus on the time domain even though they are easily contaminated by noise. They also represent the same information that cardiologists use to interpret ECG recordings.

Other than the standard derived and morphological features, the use of spectrogram of the ECG signal was proposed by [110] [84] for rhythm and heartbeat classification. A spectrogram contains more information from the time and frequency domains than the standard features. The drawback is that the spectrogram is a 2-D image and requires more complex classification algorithms.

Classification Tool

Once the feature vectors are extracted from the ECG signal, a classification algorithm is needed to detect each anomaly. The Euclidean distance is one of the most popular metrics to compare two vectors' similarities. It computes the distance between two same length vectors, and one could cluster the input vectors based on the calculated distance. Chuah [18] shows an example of using Euclidean distance for anomaly heartbeat detection. K-means clustering is the most common method of using Euclidean distance to cluster data entries. In [94], the K-means clustering method is used to classify heartbeat into normal and abnormal classes. The Euclidean distance and corresponding clustering method are fast and straightforward in distinguishing different categories. However, it has a high requirement on the extract features sufficiently. The difference between feature vectors must be significant to be assigned to other clusters. In addition, it is easily affected by data outliers and noise.

Clustering aims to find the similarity between the two groups (heartbeat segments) by computing the distance between the two groups. The conventional distance for ECG signal are Euclidean and DTW Distance [7]. Euclidean distance is the most common when comparing two groups with the same dimension. An example of using Euclidean distance for abnormal heartbeat detection can be found in Chuah and Fu's [18]. They introduce an Adaptive

Window Discord Discovery (AWDD) algorithm to detect the anomaly in ECG recordings. It was developed from a Brute Force Discord Discovery (BFDD) algorithm [73]. When comparing the heartbeats, the algorithm finds candidates' abnormal heartbeats by selecting the largest Euclidean distance. Also, they have set a threshold for the Euclidean distance to reduce the false alarm rate. The Euclidean distance only works when both heartbeat segments are the same length.

K-mean clustering is a popular clustering method that uses Euclidean distance. The K-mean clustering could cluster the heartbeat segments into many different clusters. Veeravalli *et.al* [94] developed an algorithm for real-time and personalized anomaly detection from wearable health care ECG devices. Then the K-means cluster algorithm is used to cluster all the heartbeat classes. To avoid calibration of the technique for individual users, they assigned the most frequent heartbeat segments as normal. The authors tested their algorithm on the MIT-BIH and European ST-T databases. They were able to achieve 97.1% sensitivity and 99.5% specificity.

Sivarake and Ratanamahatana [87] proposed a robust and accurate anomaly detection (RAAD) algorithm that reduced the false alarm detection rate on ECG anomaly detection. They extracted heartbeat morphological features to be their input feature vectors. Then, they calculated the dynamic time warping distance to measure the similarity between two variable-length heartbeats. In their experiment, they have tested their algorithm on INCARTDB01-05 [28], MIT-BIH arrhythmia database [63] [28], and MIT-BIH long term database [28]. Overall, their algorithm can achieve 94.35% accuracy.

Another primary method is traditional machine learning classification algorithms: Kth Nearest Neighbour (KNN), Linear Discriminant Analysis (LDA), Quadratic Discriminant Functions (QDF), Support Vector Machine (SVM), and Multilayer Perceptron Neural Network (MLPNN). These algorithms build a mathematical model based on the provided training data, and the trained model can correlate the input data with its corresponding label.

Ivaylo Christov *et. al* [17] used both the ECG morphology and VCG features to represent the heartbeat and then trained the feature vectors and their

labels with Kth nearest neighbour. As a result, the classification performance on both feature sets is over 96% for five heartbeat types (N, PVC, LBBB, RBBB, P).

Philip de Chazal *et. al* [20] used Linear Discriminant Analysis as a classification algorithm. The input feature vectors are ECG morphology features. As a result, this algorithm could perform around 97% accuracy on the MIT-BIH database with five heartbeat types (N, S, V, F, Q) classification.

Mariano Llamedo *et. al* [56] validated a heartbeat classification method for Normal, Supra-ventricular, and Ventricular heartbeats based on ECG interval features, morphological features, and DWT features. The feature vectors are trained with quadratic discriminant functions. The model had a 94% overall classification accuracy on the test dataset.

Li *et al.* [52] uses the concept of transductive transfer learning to detect the abnormal instance on an ECG signal. They trained a model to learn from labelled data sets to detect irregular heartbeat. Then they use a kernel mean matching (KMM) algorithm [30] to enable knowledge transfer between labelled data set and unlabeled data set. The model they used was a weighted transductive one-class support vector machine, which could solve the imbalanced data set problem. The authors performed experiment on records 100, 101, 103, 105, 109, 115, 121, 210, 215, 232 from the MIT-BIH database. They have achieved 87.89% average accuracy.

Ye *et. al* [103] classified 16 heartbeat types by using both morphological and dynamic features of ECG signals. Then, the support vector machine trained both morphological and dynamic features for the classification. Two channels of the ECG signal in the database were trained separately and generated two models. Both models contribute to the final classification part. Finally, the authors introduced two ways of making a final decision: one is rejection, which requires both models to make the same decision, and the other is Bayesian, which fuses the results of both models.

Zhang *et.al* [106] built 46 feature vectors to represent the heartbeat to classify the abnormal heartbeat shape on the MIT-BIH database. In the study, the authors apply the ecgpuwave tool from PhysioNet [27] to detect the boundaries

of P wave, QRS complex, ST waves. Then they collected five types of features: five inter-heartbeat intervals, five intra-heartbeat intervals, 29 morphological amplitudes, six morphological areas, and morphological distance. The five features could generate a feature vector with 46 morphological features. In the classification step, the author used the support vector machine to learn the patterns of the feature vectors. Both channels of the ECG signal were trained using a support vector machine model. The results of both models contribute to the final classification results. The result table of the paper shows that the algorithm has nearly 90% accuracy for four heartbeat types(N, F, V, S) classification.

Thomas *et. al* [91] introduced an automatic ECG arrhythmia classification idea using dual-tree complex wavelet-based features to detect normal, paced, RBBB, LBBB, PVC heartbeats. The authors proposed a feature extraction technique based on a Dual-Tree Complex Wavelet Transform (DTCWT) algorithm. Then the feature vectors were input to a multilayer perceptron neural network for abnormal heartbeats detection. The experiment result of this research is compared in Table 3.4.

Kandala Rajesh *et.al* [78] used Ensemble Empirical Mode Decomposition (EEMD) features to classify normal PVC, PAC, LBBB, RBBB heartbeats. As for the classification tool, a sequential minimal optimization SVM is used to train and classify the different heartbeat types. The experiment result of this research is compared in Table 3.4.

Wess *et.al* [97] implemented a multi-layer perceptron (MLP) classifier to detect anomalies in ECG signal . To reduce the size of feature vectors, the author applied PCA to the extracted heartbeat. Finally, the processed feature vectors are used to train an MLP neural network. Furthermore, the trained model could be used for classifying the anomalies in the ECG signal. The authors were able to test their model on the MIT-BIH database with an overall accuracy of 99.82%.

Most researchers have used traditional methods to solve the problem. Traditional machine learning classification methods do not require a considerable amount of training data, and they do not need much computational power.
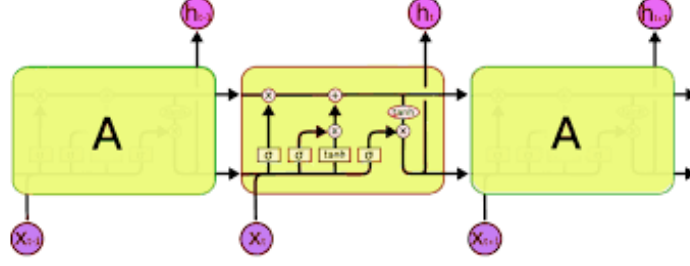
Figure 3.7: Long-Short Term Memory Layer

Due to the development of GPUs, deep learning has been proven reliable and fast for classification problems. Compared to traditional algorithms, deep learning does not require cardiology experts to extract features since the network can extract the features automatically. Instead, a deep learning model needs many labelled data for training. Luckily, public data sets can be easily found on the Internet. Therefore, many algorithms using various deep learning architectures have been published to classify heartbeats.

Ubeyli's algorithm [93] uses Eigenvector as the feature vectors and recurrent neural network as the classification tool. Normal, congestive heart failure, VT, and AFIB rhythms were trained and tested in the experiment. The experiment result of this research is compared in Table 3.4.

Chauhan and Vig [14] developed a predictive algorithm that could detect normal, PVC, PAC, paced heartbeats via a deep LSTM (Long-Short Term Memory) neural network (Figure 3.7). However, the feature extraction/selection step is neglected in their algorithm. Instead, raw ECG data and corresponding labels are used as inputs to the stacked (2-layer) LSTM neural network. In the experiment, they split the MIT-BIH database into four sets: non-anomalous training set (S_N), non-anomalous validation set (V_N), a mixture of both abnormal and normal validation sets (S_{N+A}) and the test sets (t_{N+A}). First, the LSTM network was trained on S_N, and used V_N for early stopping. The trained LSTM network was then applied to S_{N+A} to find the threshold for detecting abnormal heartbeats. Finally, the chosen threshold is used on t_{N+A} to discriminate between regular and anomalous heartbeat while predicting. The presented model achieved a 97.5% precision with a 46.47%

34

recall on the test set.

Kiranyaz *et. al* [48] presented a fast and accurate patient-specific ECG classification and monitoring system. Their experiment setup has picked five heartbeat types, N, V, S, F, Q, from 20 ECG records (100-124) from the MIT-BIH database as the training samples. The raw heartbeat segments were submitted to a 1-D adaptive CNN for pattern recognition. The 1-D convolutional neural network acts as a feature extraction tool and a classification tool. The classification times for this model are 0.58 and 0.74 msec for 64 and 128 sample heartbeat resolutions, respectively. The speed is more than 1000x faster than the real-time requirement. The experiment result of this research is compared in Table 3.4.

Sahoo *et. al* [82] improved Rai's algorithm [77] by using multi-resolution wavelet transform and machine learning to detect Normal, LBBB, RBBB, and Paced heartbeats. The authors used Q-peak, R-peak, S-peak, T-peak, QR-interval, ST-interval, RR-interval, and QRS duration as the input feature vector and used an MLP and an SVM classifier as the classification tool. In their experiment result, the overall classification accuracy of normal, LBBB, RBBB, and Paced heartbeats is 96.67% and 98.39%, respectively. The algorithm was tested on the MIT-BIH database [63].

In addition to training with a public data set, some researchers used a patient-specific approach to train the model. The first step of a patient-specific approach is to train an initial classifier with the public data set. Then the second step requires a local cardiologist to review and correct the produced labels by the initial classifier. The final step consists of training the initial classifier with corrected labels to produce the final classifier for this specific patient. The patient-specific approach could eliminate the inter-patient variations of the ECG signals. Biel *et al.* [8] shows that the variance in different human heartbeats can be very high. A number of research works [35], [20], [44], [41], [49], and [57] has proven that by using a patient-specific model, detection algorithms have a higher accuracy than the traditional system in practical cases.

3.1.8 Rhythm Classification

Rhythm classification determines the type of the rhythm anomalies such as Normal Sinus Rhythm (NSR), Atrial Fibrillation (AF), Ventricular Flutter (VF), etc. The ECG segments for rhythm classification are usually contained in several heartbeats, which could be normal or abnormal.

Similar to the heartbeat classification, the rhythm classification also could be divided into feature extraction and classification. However, the features and classification techniques used in heartbeat and rhythm classifications are very similar. The difference is that the algorithm needs to process more than one heartbeat to find a rhythm anomaly.

Ge *et.al* [26] use an auto-regressive (AR) modeling technique to classify the Normal, PAC, PVC, SVT, VT, VF rhythms. The algorithm uses Burg's algorithm to compute the AR coefficients X. The authors have tied two ways to classify the AR coefficients X: Generalized Linear Model (GLM) and multilayer feed-forward neural network. GLM equation is:

$$Y = X\beta + \varepsilon, \tag{3.11}$$

where $Y = [y_1, y_2, ..., y_N]$ is an N-dimensional vector of observed responses, X is N * P matrix of AR coefficients, β is a P-dimensional vector, ε is an N-dimensional error vector. The GLM outputs, y_1 to y_N, compared to predefined conditions to classify various heartbeat types. An artificial neural network with the AR coefficients as inputs was used for training and classification. Their experimental results show that artificial neural networks perform better than GLM.

Ozbay *et.al* [67] integrate a type-2 fuzzy clustering and discrete wavelet transform in order to build a neural network-based ECG classifier to detect Normal, Br, VT, SA, PAC, P, RBBB, LBBB, AF, AFI. The proposed diagnostic algorithm can distinguish ten different rhythm types. The system contains a fuzzy clustering layer, a feature extraction layer, and a classifier. The fuzzy clustering layer select segments represent the arrhythmia class in the ECG. A wavelet transformation was applied to the ECG segments to generate features. The authors have trained three Type-2 Fuzzy Clustering Neural

Network models (T2FCWNN-1, T2FCWNN-2, and T2FCWNN-3) with three different training data sets. The three training data sets have the same amount of ECG segments. However, the length of each ECG segment is 101 sample points, 52 sample points, and 27 sample points correspondingly. As a result, the T2FCWNN-3 had the lowest training time, 4.86 seconds, and the test error rate, which was 0.23% among all three models.

Patel *et.al* [71] used a thresholding technique to detect arrhythmia on ECG collected from a mobile platform. In the paper, the first uses the Pan Tompkins [69] algorithm to detect R peaks on the ECG recordings. Then they characterized SB, ST, PVC, PAC, and Sleep Apnea using a predefined threshold to classify different rhythms. Their system could reach a 97.3% detection accuracy.

Rajpurkar *et.al* [79] have developed an algorithm that could out-perform board-certified cardiologist in detection of 12 types of arrhythmia using a 34-layer CNN. The network took a 30-second raw ECG signal recording as input, and the output was a sequence of label predictions. The model output a new prediction every second. The training data set contains 64121 ECG records from 29163 patients, and the testing data set contains 336 records from 328 patients. The model performs at 80.9% precision, 82.7% sensitivity, and 0.809 F1-score.

Acharya *et.al* [1] use two 11-layers CNN to detect AFIB, AFL, VF(VFL) from normal heartbeat rhythm. The two networks, Net A and Net B used a 2-second raw ECG recording and a 5-second raw ECG recording as input, and the output is its corresponding label. The ECG segments are Z-sored normalized and then submitted to the 1-D deep CNN. The result of Net A and Net B are compared in Table 3.5.

3.1.9 Heartbeat/Rhythm Classification Algorithm Comparison

In the previous sections, we reviewed many algorithms that classify ECG in various categories. One can see in Table 3.4 the classification results performed using the MIT-BIH database. In addition, some algorithms' perfor-

Table 3.4: Heartbeats classification performance on MIT-BIH dataset

Method	Year	Abnormal/Normal	Heartbeat types	TP	FP	TN	FN	Sensitivity	False Alarm	Accuracy
Christov *et.al*[17]-morphology	2006	18378/47239	5	18042	1604	45635	336	98.17%	3.40%	97.04%
Christov *et.al*[17]-frequency	2006	18378/47239	5	17590	1459	45780	788	95.71%	3.09%	96.58%
Chazal *et.al*[20]-frequency	2006	4317/34394	5	4108	1962	32432	209	95.16%	5.70%	94.39%
Ubeyli *et.al*[93]	2009	269/90	4	268	2	88	2	99.26%	2.22%	99.89%
Llamedo *et.al*[56]	2010	5441/44188	3	4752	2238	41950	689	87.34%	5.06%	94.10%
Ye *et.al*[103]-rejection	2012	19913/64042	16	19815	93	63949	98	99.51%	0.15%	99.77%
Ye *et.al*[103]-bayesian	2012	20745/65264	16	20557	286	64978	188	99.09%	0.44%	99.45%
Zhang *et.al*[106]	2014	5653/44011	4	5248	4869	39142	405	92.84%	11.06%	89.38%
Thomas *et.al*[91]	2015	26626/67268	5	22900	1300	65968	3726	86.01%	1.93%	94.65%
Kiranyaz *et.al*[48]	2015	7366/42191	5	6539	1228	40963	827	88.77%	2.97%	95.85%
Rajesh *et.al*[78]	2017	8000/2000	5	7677	33	1967	323	95.96%	1.65%	96.44%
Sahoo *et.al*[82]	2017	807/244	4	798	5	239	9	98.88%	2.04%	98.67%

Table 3.5: Rhythm classification performance on MIT-BIH dataset

Method	Year	Abnormal/Normal	Rhythm types	Rhythm Length	TP	FP	TN	FN	Sensitivity	False Alarm	Accuracy
Ge *et.al*[26]	2002	713/143	6	1.2 s	706	10	133	7	88.77%	6.99%	98.01%
U. Acharya Net A[1]	2017	20807/902	4	2 s	19160	62	840	1647	92.08%	6.87%	92.13%
U. Acharya Net B[1]	2017	8322/361	4	5 s	7946	376	294	67	95.48%	18.56%	94.9%

mance metrics were converted to binary classification to detect normal and abnormal heartbeats. The reason is that computer diagnoses are not 100% accurate. We still need doctors to make the final diagnosis as they only know the context. Therefore, the methods should focus on a binary classification that classifies all abnormal heartbeats.

Similarly, Table 3.5 compares all methods that classify the rhythms on the MIT-BIH database. In addition, the Table has only shown the algorithms that have provided enough information to compute our metrics.

3.2 Discussion

3.2.1 Challenges for Heart Anomaly Detection on Ambulatory Electrocardiogram

There are still several challenges in heart anomaly detection:

- ECG signals may be contaminated with motion artifacts since the patient is constantly moving. The noisy signal may have similar morphology to abnormal cardiac signals resulting in a false positive. It is easy for the human eye to identify these conditions; however, it is much harder for computers to separate noise from the signal;

- The model training requires a labelled ECG signal. Trained personnel

is required to label the ECG data set. In addition, the labelling process is very time-consuming. For example, a 10-second ECG signal has 2500 data points, and continuous monitoring usually takes 24-48 hours;

- The ECG heartbeat data is highly imbalanced. Over 99% of the heartbeat data is the typical case, and only 1% of the heartbeat data present 16 abnormal points. Therefore, the highly imbalanced dataset makes it more challenging to adjust the learning step. Several options could be explored to reduce the effect of imbalanced data, such as database re-sampling or using a cost-sensitive method, kernel-based method, or active learning; [34].

3.2.2 Conclusion

This chapter introduces the definition of anomaly detection on ambulatory electrocardiograms (ECG) and its importance. In addition, since the ECG signal collected from a wearable device is frequently contaminated with electrical noise and motion artifacts, noise reduction (Section 3.1) is critical for anomaly detection. Furthermore, most electrocardiogram anomalies fall into two major categories: irregular heart rate and irregular heart rhythm. The irregular heart rate on ECG could indicate bradycardia, tachycardia, heart block, arrhythmia, etc. The irregular heart rhythm could be ectopic heartbeat when checking a period of ECG signal. Therefore, based on the different irregularities on the ECG, anomaly detection could be divided into several categories: heartbeat detection (Section 3.1.6) for detecting the location of each heartbeat; heartbeat segmentation (Section 3.1.6) for segmenting the heartbeats from the entire ECG signal; heartbeat classification (Section 3.1.7) for classifying the type of one heartbeat; rhythm classification (Section 3.1.8) for classifying the type of a period of ECG signal.

From the literature, we have reviewed the conventional methods for each part. Many researchers have used fixed digital filters, discrete wavelet transform, empirical mode decomposition, and adaptive filters been used by many researchers. For heartbeat detection, much research used fixed digital filters,

discrete wavelet transform, and Shannon energy envelope to remove the noise and unwanted waves while preserving R peak information. They then use the R peak location to compute the heartbeat. For heartbeat segmentation, the most common method is to use a predefined window to segment the heartbeat signal from the entire signal. The authors used morphological features and derived features to represent the heartbeat signal for the literature on heartbeat classification. The morphological features are calculated from the ECG signal, and derived features are computed using other methods such as discrete wavelet transform, independent component analysis, empirical mode decomposition, etc. Both morphological and derived features are then used for training to generate a mathematical model of the heartbeat signal. The most popular model uses k nearest neighbour, linear discriminant analysis, support vector machine, multilayer perceptron neural network, and deep neural networks such as CNN and RNN. Similarly, algorithms take a period of ECG signal as the input to the model for the rhythm classification.

There are three significant challenges that this thesis addresses:

- The reduction of motion artifacts on the ECG signal interference with anomaly detection.

- Model training needs a massive amount of labelled data that are hard to come by.

- ECG databases have very imbalanced data making it difficult for Deep Learning model training.

Chapter 4

An Automatic Method to Reduce the Effects of Baseline Wander and Motion Artifact on Ambulatory Electrocardiogram

4.1 Introduction

Commercial heart monitoring devices possess many advantages, such as ease of use, long-term body monitoring, and easy access to personal health data. However, it also has some important limitations. One major limitation is that the collected ECG is frequently contaminated with noise, making the ECG interpretation impossible. For example, the standard 12-lead ECG device in clinics requires measuring the patient at rest, but most mobile healthcare devices use a rubber band to press the sensor onto the patient's chest; therefore, the sensor is more likely to be affected by body motion. In addition, body motion introduces two major noise types: baseline wanders and motion artifacts. ECG signals are easily distorted by artifacts that have nothing to do with heart functions. In some cases, artifacts could even mimic true arrhythmia, leading to false diagnostics [32]. Therefore, noise removal is necessary for any ambulatory ECG monitoring applications. This thesis proposes a new noise removal method that automatically suppresses the baseline wander and motion artifact. The new algorithm consists of an adaptive empirical mode decomposition and reconstruction algorithm capable of removing baseline wander

and some motion artifacts, followed by an adaptive filter algorithm that uses an accelerometer to reduce further the motion artifacts. Section 4.2 describes the proposed algorithm. Then in Section 4.3 the experimental testing setup and the results. We then conclude in Section 4.5 by reviewing the advantages of our approach.

4.2 Proposed Algorithm

We have developed a new algorithm to reduce the baseline wander and motion artifact automatically. The new algorithm consists of a preprocessing step and a filter step. The preprocessing step contains ECG signal preprocessing and unusable ECG signal detection. The filter step consists of an adaptive empirical mode decomposition and reconstruction step that could automatically decompose the signal without human input to reduce motion wander, a motion-sensitive adaptive filter that uses a 3-Axis accelerometer that automatically selects the best reference noise signal to remove motion artifacts, and a variational mode decomposition and reconstruction method to remove high-frequency noise. The complete block diagram of the algorithm can be seen in Figure 4.1.

Data Acquisition

In this research, we have used the Astroskin Smart Shirt [38] and VIVALNK wearable ECG device [40] for ECG collection. The Astroskin Smart Shirt can collect a 3-lead ECG at 250 Hz and 3-Axis accelerometer data at 50 Hz. The VIVALNK wearable ECG device can collect a single lead ECG at 128 Hz and 3-Axis accelerometer data at 5 Hz. The ECG signals collected by the VIVALNK ECG device are up-sampled to 250 Hz. The devices are shown at Figure 4.2 and Figure 4.3. Both devices collect lead II ECG signals and accelerometer signals. The test subject was a healthy male adult age 32 with no heart problem. Each device was worn by the subject for one hour consecutively. During the one hour, the test subject was free to do his daily activities, including resting, sitting, walking, and cycling. During resting time, the subject was sitting but

42
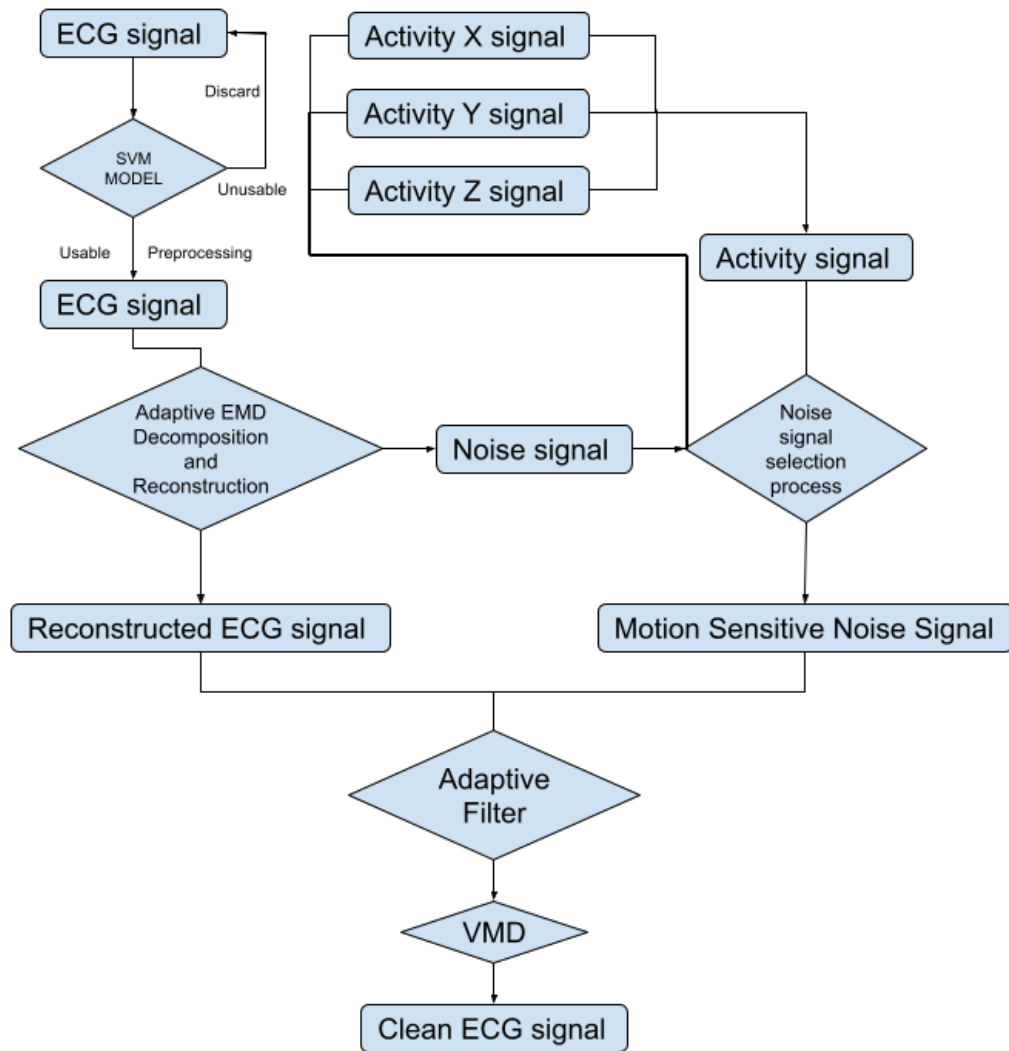
Figure 4.1: Block diagram of the proposed algorithm

Figure 4.2: Astroskin smart shirt bio-signals sensors [38]

not moving his upper body for 5 minutes during the resting time. The resting signals were used as the reference signal.

Since the device could not record the noise source or a clean signal for the algorithm to compare, the reference signal is used as the ground truth signal to compare the ECG signal before and after applying the proposed noise reduction algorithm. Figure 4.4 is the reference signal collected from the Astroskin Smart Vest. And Figure 4.5 is the reference signal collected from Vivalnk patch. The reference signal in Figure 4.4 is 10 second long. It contains 12 heartbeats, and its average RR interval is 0.824 seconds. The reference signal in Figure 4.5 is 10 seconds long with 11 heartbeats. The average RR interval is 0.922 seconds. The sitting time mainly includes working with computers and eating during the data collection. Active time includes walking on flat ground, going up and downstairs, and cycling on a stationary bike. The total sitting time was around 25 minutes, the total walking time was 15 minutes, and the total cycling time was 15 minutes.

4.2.1 ECG Signal Preprocessing

Once the data is collected, a preprocessing step is needed. First, the ECG signal is normalized using the z-score normalization method [72], and the 3-axis accelerometer is also normalized using a min-max normalization method.

44

Figure 4.3: VIVALNK wearable ECG device [40]
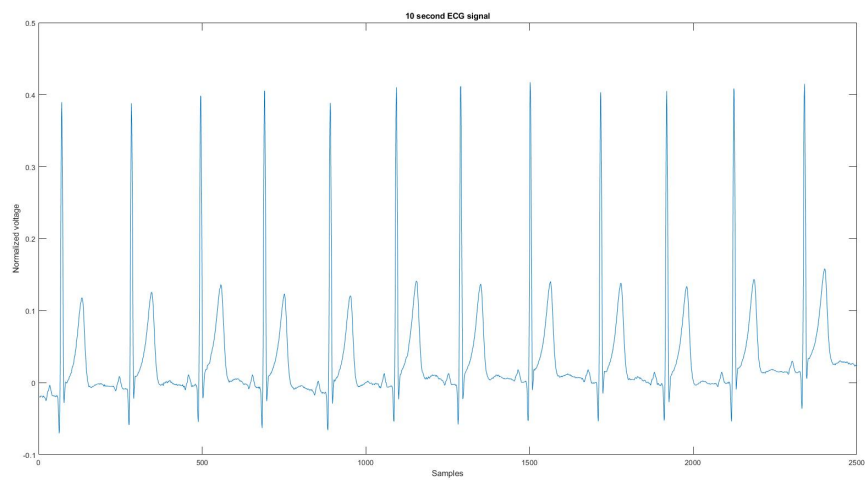
Figure 4.4: The reference signal from Astroskin Smart Skirt
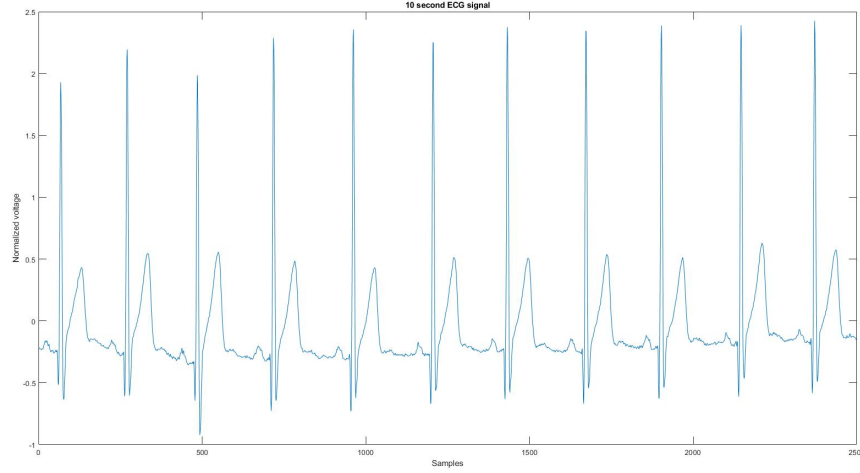
Figure 4.5: The reference signal from Vivalnk ECG device

The z-score normalization method transforms the original data x relative to its mean value μ and standard deviation σ using the following equation:

$$\hat{x} = \frac{x - \mu}{\sigma}.\tag{4.1}$$

The min-max normalization method normalized the signal between $[0, 1]$ using the following equation:

$$\hat{x} = \frac{x - Min(x)}{Max(x) - Min(x)}\tag{4.2}$$

where $\hat{x}$ is the normalized signal, x is the original signal, $Min(x)$ is the minimum value of the signal, and $Max(x)$ is the maximum value of the signal. After the normalization step, the accelerometer signal is over-sampled using the following algorithm.

- The process first inserts p (for example $p = 250$) zeros to up-sample the signal.;

- Then, the new signal is filtered by an FIR anti-aliasing filter to match the shape of the original signal. In this part, a Kaiser window method was used to approximate the ideal anti-aliasing filter;

46

- Finally, q (for example, $q = 50$) samples in the up-sampled signal are discarded to get the final signal.

The p is the target sample frequency, and the q is the original sample frequency. Then with the same length for the ECG signal and the accelerometer signal, they are segmented into 10-second segments where the noise reduction algorithm is applied.

4.2.2 Detection of Unusable Signal with Support Vector Machine

Following the preprocessing step, some sections of the signals were unusable. This may be because the patient's movement was too pronounced during the ECG collection step, resulting in lost heartbeat information. Therefore, this step aims to detect if the signal segments are too noisy. The standard to distinguish a usable and unusable signal is to find if the number of QRS complexes on the ECG signal is the same as the expected number of QRS complexes. If the number of QRS complexes on the ECG signal is less than 60% of the expected number, the segments will be considered unusable. A Support Vector Machine (SVM) model was trained using eight features to achieve this objective. SVM is a computer algorithm that learns a classifier function from labelled data that can then be used to classify new unlabeled data [65]. The SVM algorithm has been proven to be very accurate, and efficient for binary classification in many applications such as [65] [98] [25]. Another major reason for using SVM is that the SVM does not require a lot of training data. During the training, the kernel function used for the SVM is the Gaussian function:

$$G(x_j, x_k) = exp(-||x_j - x_k||^2) \qquad (4.3)$$

where $G(x_j, x_k)$ is element (j, k) of the Gram matrix, and x_j, x_k are the vectors representing observations j and k in x.

In Figure 4.6 and Figure 4.7, the frequency plots show the frequency range for a usable 10-second long ECG signal segment and an unusable 10-second long ECG signal segment. By comparison, one can see that the difference
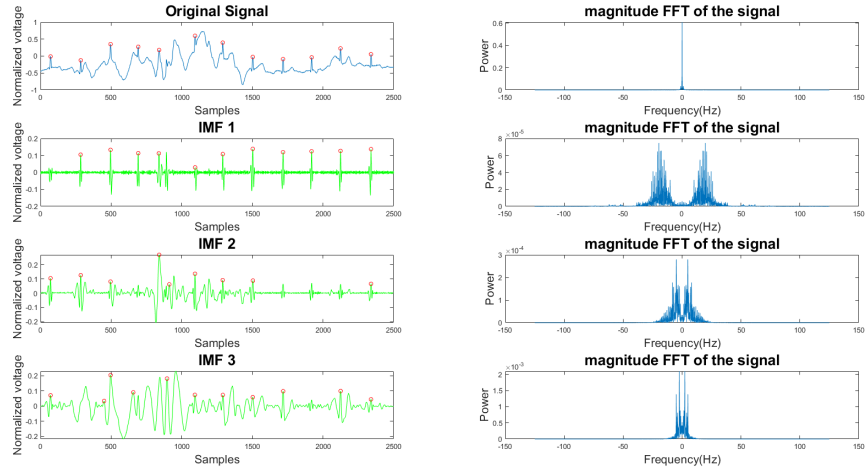
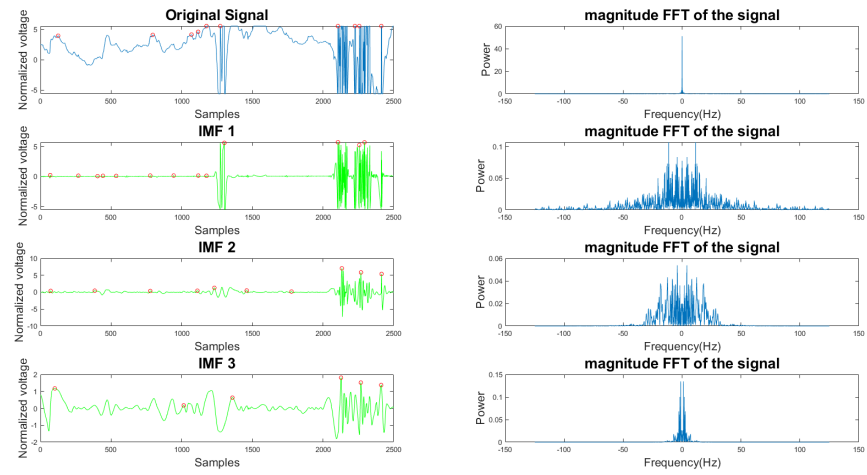Figure 4.6: Time and frequency form of an usable ECG signal and its IMFs

Figure 4.7: Time and frequency form of an unusable ECG signal and its IMFs

Table 4.1: Confusion matrix and performance

TP	FP	TN	FN	SEN	SPC	ACC
99	2	48	1	99%	96%	98%

between the two ECG signals is significant. The EMD algorithm was used for the ECG signal to extract features from the ECG signal. During the EMD algorithm, the signal is decomposed into many IMFs, and the first three IMFs contain the most information on the ECG signal. Based on the frequency spectrum, 90% of the frequency components were chosen to represent each signal. These frequency components count from 0 to a frequency i are performed by summation of each frequency power component ($power_0$ to $power_i$) until they correspond to 90% of the total signal power. The 90% frequency concentration of the original signal (f_0) and its corresponding IMF1 (f_1), IMF2 (f_2),and IMF3 (f_3) were calculated. In addition, the 90% frequency components of the reconstructed signal (rf_0) get from Adaptive Empirical Mode Decomposition and Reconstruction (AEMDR) described in Section 4.2.3 and its corresponding IMF1 (rf_1), IMF2 (rf_2),and IMF3 (rf_3) were calculated. Finally, the following feature vector f_0, f_1, f_2, f_3, rf_0, rf_1, rf_2, rf_3 is used to represent the ECG signal and to train the SVM model. Figure 4.8 describes in detail of the feature extraction and training process.

The data for training and testing were collected from both the Astroskin Smart Shirt and the VIVALNK wearable device. There were 81 unusable ECG segments and 162 usable ECG segments in the training data set. There were 100 unusable ECG segments and 50 usable ECG segments in the testing data set. All ECG signal segments were randomly picked from the 1-hour-long ECG record. The confusion matrix and the performance are shown in Table 4.1. The TP and FP are the corresponding numbers of correctly detected unusable signals and usable signals.
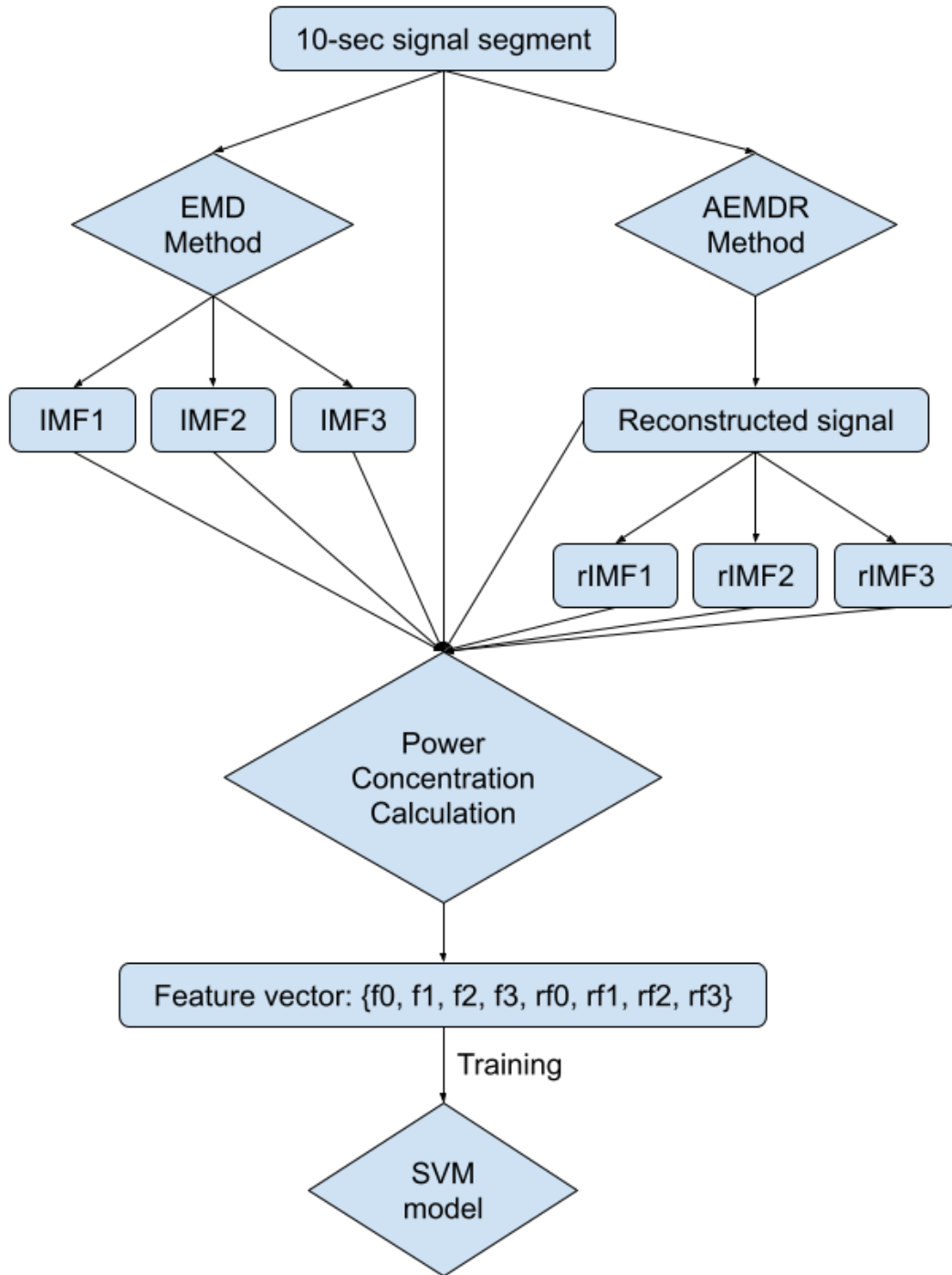
Figure 4.8: Block diagram of feature extraction and SVM model training process

4.2.3   Adaptive Empirical Mode Decomposition and Reconstruction

The first step of the processing algorithm is the Adaptive Empirical Mode Decomposition and Reconstruction (AEMDR). This step aims to remove as much low-frequency noise as possible. Still, in the meantime, the algorithm is also trying to preserve the QRS complex. The AEMDR is divided into two parts. The first part deals with signal decomposition, and the second part deals with signal reconstruction. A standard Empirical Mode Decomposition (EMD) algorithm is used for the signal decomposition part. The IMF and the lowest frequency component are decomposed using the following rules:

- The number of extremes and zero-crossings must be equal or differ at most by one;

- All local maxima and minima must be symmetric to zero.

After signal decomposition, one wants to reconstruct the high-frequency components as the clean signal and the low-frequency as an estimate of the signal noise. The original EMD algorithm decomposes the signal into x IMFs, but the x is not fixed. Thus, human intervention is often needed to separate clean from noisy IMFs. We propose an automatic method to select the clean and noisy IMFs for signal and noise reconstruction to solve this problem. To achieve our goal, we used the Pan-Tompkins algorithm [69] to detect the QRS complex's number on the original signal and all the IMFs. The Pan-Tompkins algorithm uses a filter bank that consists of a band-pass filter, a differentiator, a squaring filter, and a moving window integrator to reduce the signal noise so that only R wave information is present [51].

The statistical terms are explained as follows:

- The total QRS complex is the number of QRS complexes detected from the original signal with the Pan-Tompkins algorithm;

- The true positive (TP) is the number of QRS complexes detected on the IMFs and the original signal;
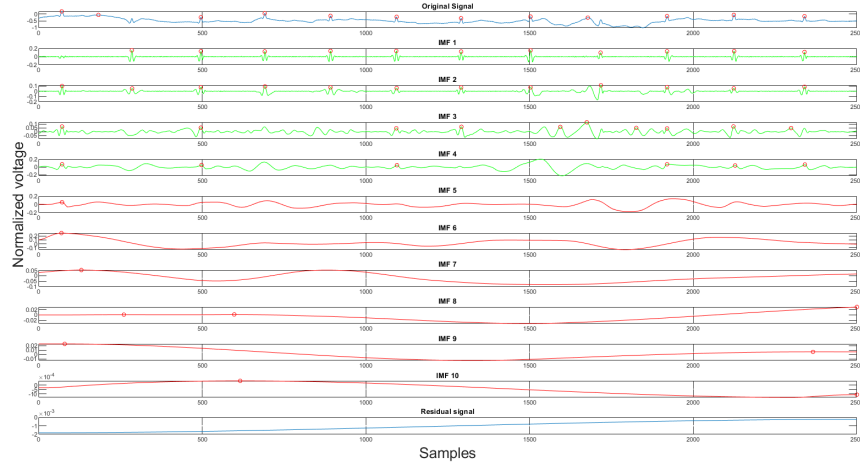
Figure 4.9: The original signal, clean IMFs, noisy IMFs and residual

- The false positive (FP) is the number of QRS complexes detected on the IMFs, but not on the original signal;

Then, we have applied the following rules to all IMFs to select the clean ones:

- The IMF is selected when its TP is larger than 50% of the total QRS complex;

- The IMF that satisfies the first condition but has FP that is larger than 50% of the total QRS complex will be applied to the AEMDR method again;

- For the signal reconstruction, if the IMF's meet two conditions: TP is larger than 50% of the total QRS complex; FP is less than 50% of the total QRS complex, then it is considered a clean IMF, or otherwise it is considered as a noisy IMF.

Once the clean IMFs and noisy IMFs are selected, the algorithm will proceed with the reconstruction. The decomposition and the IMF selection result can be seen in Figure 4.9. In the figure, the green IMFs are selected on the clean signals, and the red IMFs are the noise signals, where the x-axis and y-axis are sample numbers and normalized voltage correspondingly. This algorithm produces a clean signal and a noise signal using the following equations:

52

$$x(n) = \sum_{i=1}^{k} c_i(n), \qquad (4.4)$$

$$s(n) = \sum_{j=1}^{k} c_j(n) + r(n) \qquad (4.5)$$

where $x(n)$ is the clean signal, $s(n)$ is the noise signal, $c_i(n)$ and $c_j(n)$ the clean IMFs and the noisy IMFs respectively, m is the number of clean IMFs, and the k is the number of noisy IMFs, $r(n)$ is the residual signal. The acquired clean and noise signals are the desired input and reference input processed by the adaptive filter.

4.2.4 Motion Sensitive Noise Signal Generation

The motion-sensitive noise signal generation step is a preparation step for the following adaptive filter described in section 4.2.5. Finding the best reference noise input is always the most challenging task for an adaptive filter application. Removing baseline wander and motion artifact is harder since it is almost impossible to get the noise source from body motion. Some researchers [81] [101] explore the possibility of using accelerometer data as the reference noise signal. The accelerometer is standard equipment for many new commercial mobile healthcare devices. It can capture the user's body movement and is used mainly for fall detection. Even though the accelerometer data is related to the motion artifact, it does not perform well when directly applied to the adaptive filter. Therefore, we have proposed an algorithm to combine the noise signal extracted from AEMDR analysis with the collected accelerometer signals. We could simulate the actual noise on the ECG signal by doing so.

The motion noise consists of baseline wander and motion artifacts. The baseline wanders noise is a low-frequency noise that is caused by respiration and body motion. The motion artifacts are in the same frequency range as the ECG waves and are caused by the movement of the ECG leads. Our Adaptive EMD decomposition and reconstruction algorithm preserves the high frequency of the original ECG signal and discards the low-frequency noise signal. The discarded noise signal contains the baseline wander information.

Then by using the low-frequency noise signal from the AEMDR analysis, the motion artifacts, and the activity signal collected by the accelerometer, one can create a reference signal that the adaptive filter can use.

The next step is to select the activity signal with the highest positive correlation coefficient and the lowest negative correlation coefficient to discriminate against unrelated activity signals that interfere with the adaptive filter. The reason for choosing the highest positive correlation coefficient and the lowest negative correlation coefficient is that the two signals are unrelated when the correlation coefficients are closer to zero. Therefore, we want the reference noise signal to be correlated to the original signal as much as possible. The correlation coefficient is computed using the following equation:

$$\rho(A, B) = \frac{1}{N-1} \sum_{i=1}^{N} (\frac{A_i - \mu_A}{\sigma_A})(\frac{B_i - \mu_B}{\sigma_B}) \tag{4.6}$$

where A and B represent the two signals, and μ and σ are the mean and standard deviation of the corresponding signals, N is the total number of sample of the signal.

The activity signals consist of an activity signal on the x-axis (A_x), an activity signal on the y-axis (A_y) and an activity signal on the z-axis (A_z). The 3-axis combined activity signal A_{xyz} is calculated using the following equation:

$$A_{xyz} = \sqrt{A_x^2 + A_y^2 + A_z^2}. \tag{4.7}$$

The selected activity signals are combined with the noise signal generated by the EMD algorithm to generate a new noise signal as follows:

$$X(n) = s(n) + A_{xyz/x/y/z}(n) \tag{4.8}$$

where $X(n)$ is the final reference noise signal, $s(n)$ is the noise generated by the EMD algorithm, and $A_{xyz/x/y/z}(n)$ is the activity signals in the x, y, and z direction, and the combined activity signal A_{xyz}. One can see in Figure 4.10 an example of the combined noise signal.
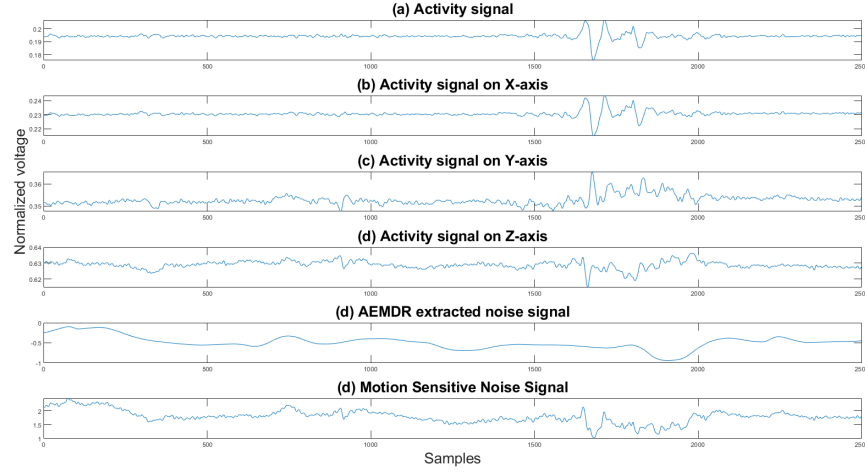
54

Figure 4.10: Activity signals, EMD noise signal, and the motion sensitive reference signal

4.2.5 Adaptive Filter

Once the motion-sensitive noise signal is generated, we can proceed to the next step. This step processes the original signal using an adaptive filter. This step aims to suppress the noise that has a similar frequency as the QRS complex. An adaptive filter is a linear filter where the transfer function is controlled by variable weights that are adjusted according to an optimization algorithm. The closed-loop nature of an adaptive filter uses feedback as an error signal to refine its transfer function, in our case, the combination of the reference noise signal that is sensitive to motion. Generally speaking, the closed-loop adaptive process involves using a cost function, which is a criterion for optimum performance of the filter. One can see in Figure 4.11 the architecture of our adaptive filter. The measured ECG signal $m(n)$ is composed of the desired signal $d(n)$, which is contaminated by an additive noise $\hat{y}(n)$ and is defined by:

$$m(n) = d(n) + \hat{y}(n). \tag{4.9}$$

The main goal of the adaptive filter is to find a noise signal $y(n)$ from a modified reference noise signal $X(n)$ that can minimize the effect of the additive noise

55

$\hat{y}(n)$ is expressed by:

$$\hat{m}(n) = d(n) + \hat{y}(n) - y(n). \tag{4.10}$$

Taking expectation on both sides and realizing that the measured ECG $m(n)$ is uncorrelated with $\hat{y}(n)$ and $y(n)$ is expressed by:

$$E[\hat{m}^2] = E[d^2] + E[(\hat{y} - y)^2] + 2E[d(\hat{y} - y)] \tag{4.11}$$

where $E[d(\hat{y} - y)] = 0$. The signal power $E[d^2]$ is unaffected as the filter is adjusted to minimize $E[\hat{m}^2]$:

$$\min[\hat{m}^2] = E[d^2] + \min E[(\hat{y} - y)^2]. \tag{4.12}$$

When the filter is adjusted to minimize the output noise power $E[\hat{m}^2]$, the output noise power $E[(\hat{y} - y)^2]$ is also minimized. Since the output signal remains constant, minimizing the total output power maximizes the output signal-to-noise ratio. In our implementation, we used the well-known LMS algorithm [99] to optimize the filter parameters for the following reasons. Compared to LMS, NLMS solves the limitation of the LMS with input scaling. In our experiment, all signals are scaled, and therefore, the limitations of the LMS filter will not affect the result. We will also compare the performance of LMS and RLS algorithms. Both produced a similar result, but because the LMS algorithm complexity $2N + 1$ is much lower than the RLS algorithm $4N^2$ [21], and the fact that the LMS algorithm converges much faster than the RLS algorithm, we decided to use the LMS algorithm [99]. The LMS algorithm is based on the steepest descent algorithm that adapts the coefficient sample by sample toward the optimum vector on the performance surface [13]. The LMS algorithm [99] can be described as:

$$\mathbf{W}(k + 1) = \mathbf{W}(k) + 2\mu\mathbf{X}(k)[m(k) - \mathbf{X}(k)^T\mathbf{W}(k)] \tag{4.13}$$

where $\mathbf{W}(k)$ is the weights of the filter of size M at iteration k, where M is the filter order. $\mathbf{X}(k)$ is an input vector of size M at iteration k of the corresponding samples from the reference noise signal $X(n)$, $m(k)$ is the value
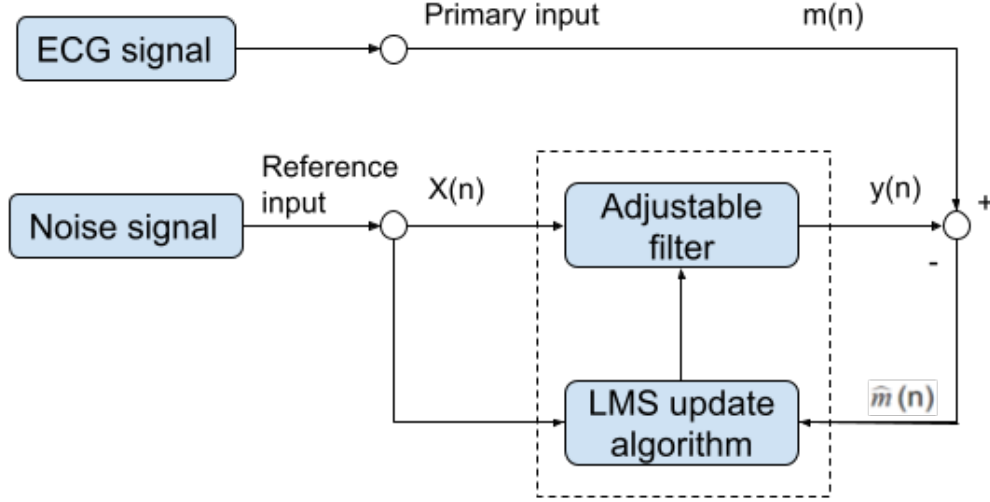
Figure 4.11: Block diagram of our noise removal adaptive filter

of the input signal at iteration k, the parameter μ is the learning rate (step size). Since $\hat{m}(n) - d(n) = \hat{y}(n) - y(n)$ this is equivalent to causing the output $\hat{m}(n)$ to be a best least squares estimate of the signal $d(n)$.

The LMS update algorithm will update the filter weights by minimizing the power of the error signal $(m(n) - y(n))^2$. Using the LMS algorithm, the adjustable filter learns the noise signatures from the original ECG signal. Then, it converts the input noise signal to a new noise signal that can be subtracted from the original ECG signal.

An adaptive filter needs to set two main parameters: learning rate (step size) and the filter order. We have done a series of experiments to test the best value or range for these two parameters. The learning rate we have tested ranges from 0.005 to 0.05, and the filter order range from 1 to 28. Based on the result of our experiments, the learning rate (step size) is set to 0.01, and the filter order is set to 10.
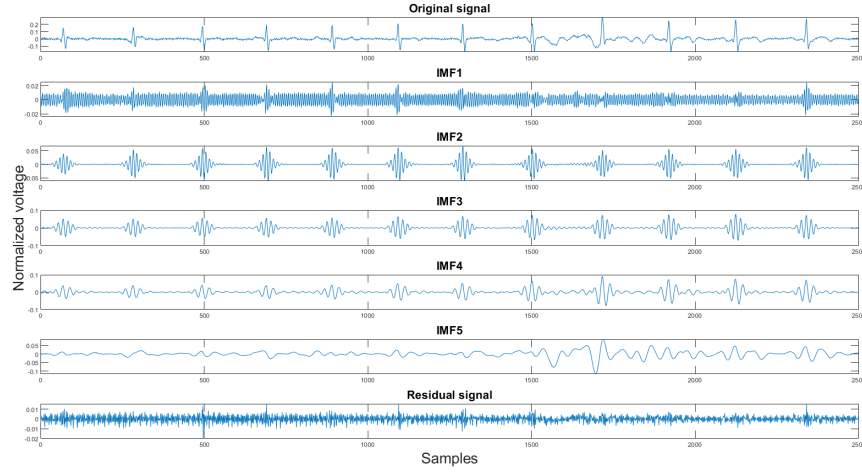
Figure 4.12: The original signal, IMFs, and residual signal by VMD method

4.2.6 Variational Mode Decomposition and Reconstruction

Finally, the last step of the proposed algorithm is the Variational Mode Decomposition and Reconstruction (VMDR) [22]. The objective of this process is to remove the high-frequency noise on the filtered signal to make the signal smoother. The variational mode decomposition and reconstruction are divided into two parts: standard variational mode decomposition and the other is selecting clean IMFs for reconstruction. In the decomposition part, the standard VMD method is applied. The Variational Mode Decomposition (VMD) method is similar to the EMD method. It decomposes the signal into K different IMFs. The second part is finding the clean IMFs for reconstruction. The VMD method produces a high-frequency residual signal. This is also proven in our experiment shown in Figure 4.12. The x-axis is the sample number in the figures, and the y-axis is the normalized voltage. Therefore, for the signal reconstruction, the high-frequency residual signal was discarded.

The reason for choosing this method to filter out the high-frequency noise is that, unlike the EMD method, the VMD method is more sensitive toward the high-frequency components [58].
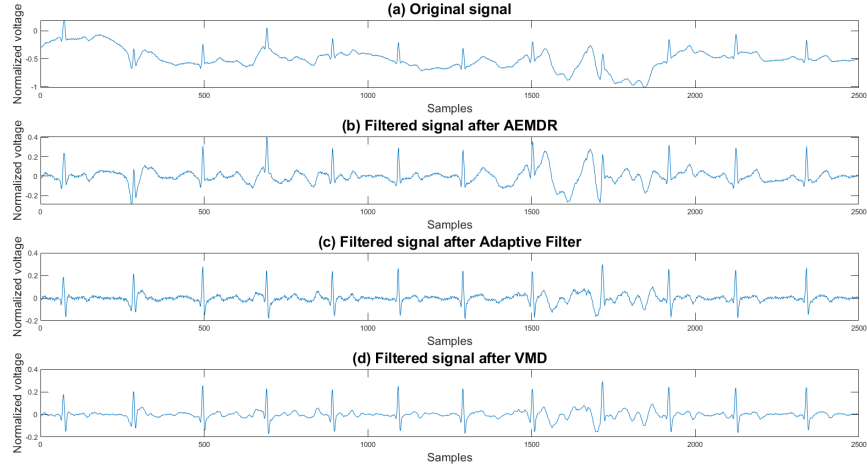
58

Figure 4.13: (a) Original signal (b) Signal after AEMDR step (c) Signal after adaptive filter step (d) Signal after VMD step

4.2.7 Example results of the algorithm

In Figure 4.13, one can view the result after each step. Figure 4.13(a) shows the original signal. The signal contains baseline wander and motion artifacts. Figure 4.13(b) is the filter result after the AEMDR step. One can see that the baseline wander was significantly reduced, and almost all QRS complexes were in the same baseline. But the motion artifacts remain in the signal. Then, after applying the adaptive filter step, int the 4.13(c), most of the motion artifacts were compensated. Finally, the VMD method removes the high frequency noise components in Figure 4.13(c), and produces the final result 4.13(d).

4.3 Experimental Results

As described in Section 4.2, the data used in this experiment were collected using Astroskin Smart Shirt and VIVALNK wearable devices. The raw 3-lead ECG signal and its corresponding 3-axis accelerometer signal were acquired for 8-hours. However, the ECG signals are split into 10-second segments during the noise removal process. There were 372 10-second segments obtained from the Astroskin Smark Vest and 330 10-second segments acquired from the

59

VivaLnk. After the prediction of the SVM model, 328 segments from the Astroksin and 320 segments from VivaLnk were labelled as usable. There were many unusable ECG segments from Astroskin Smart Vest because its sensors were not attached to the human body resulting in a loss of contact. There was a total of 648 ECG segments of 10-second to be analyzed. However, it is unrealistic to present all of them in this paper. Therefore, we have selected one clean segment and three noisy segments from both devices. The following sections will explain the metrics we used to show how the proposed algorithm improves signal quality. We use three metrics to illustrate how the signal was enhanced objectively. These are the histogram of the difference, the correlation coefficients, and the mean square error between the test signal and the reference signal. To compare the proposed algorithm to the conventional algorithm for motion artifact removal, we have also implemented several noise removal algorithms based on [3] [81]. The reference noise removal algorithms are: a zero-filtering Butterworth high pass filter (IIR) with a cut-off frequency of 0.5 Hz and a filter order of 2, a moving average filter (MA) with window length set as $\frac{sample frequency}{2*0.5}$, a DWT with Daubechies wavelet of order 8 and 9 levels decomposition, and the last detail coefficient was discarded, an EMD with the previous two intrinsic mode functions discarded, a variational mode composition with k equals to the number of IMFs from the EMD method and with the last two IMF discarded as well, and an adaptive filter with using only activity signal.

4.3.1 Visual Result and the Bivariate Histogram

This section compares the visual result and the bivariate histogram of the original and filtered ECG signals. The visual result shows the ECG signal after all the noise removal steps. And the bivariate histogram shows the difference between the reference signal and the test signal.

To set a standard for the bivariate histogram, we have chosen another clean signal as an example shown in Figure 4.14(1). And the Figure 4.14(2) demonstrates the bivariate histogram between two clean signals. One can see that the shared bin values form a diagonal line. With this, one can compare
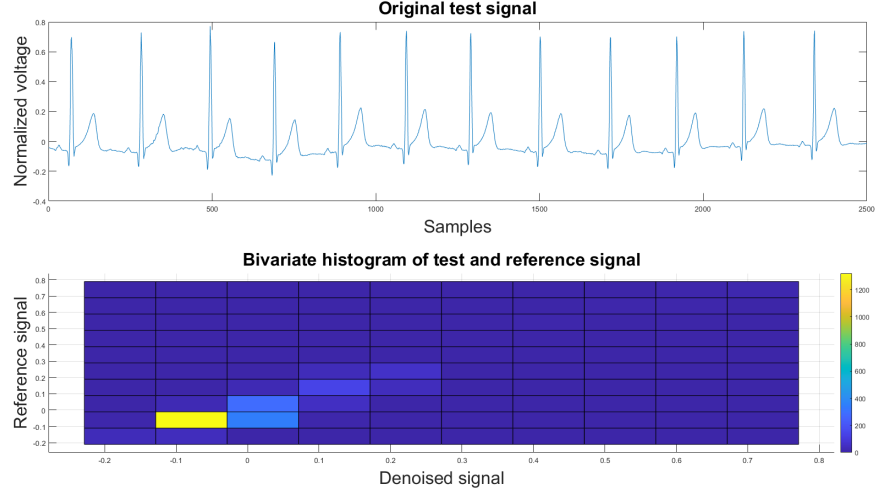
Figure 4.14: (1) clean example signal (2) bivariate histogram between test and reference signal

the histograms of all denoised signals to the figure. If the histogram is denser on a diagonal line, it shows a better denoising result; otherwise, it will show a divergent pattern.

During the testing, to make the test signal aligned with the reference signal, we have annotated all R peaks on both the reference signal and the test signal. And then, all RR intervals on the test signal were re-sampled to have the same length as the RR intervals in the reference signal.

Figure 4.15 to 4.20 compares 6 examples of the original signal and the filter result using the proposed algorithm. In all six examples, the bivariate histogram of the filtered ECG signal shows a more dense pattern than the original signal. To compare our algorithm to other traditional algorithms, Figures A.1 to Figure A.36 in the Appendix shows the filter result using the Butterworth high pass filter, the moving average filter, the discrete wavelet transform, the empirical mode decomposition, the variational mode decomposition and the adaptive filter with accelerometer data.

4.3.2 Correlation Coefficient

In addition to visual results and signal-to-noise ratio, we have introduced another metric to validate the algorithms. This metric computes the corre-
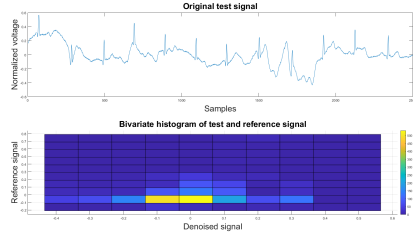
(a) Original signal
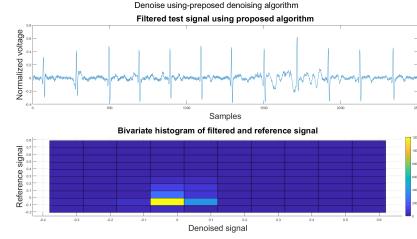
(b) Filtered signal

Figure 4.15: Example 1: (1) test signal (2) bivariate histogram

(a) Original signal

(b) Filtered signal

Figure 4.16: Example 2: (1) test signal (2) bivariate histogram
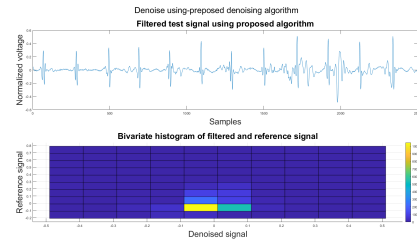
(a) Original signal

(b) Filtered signal

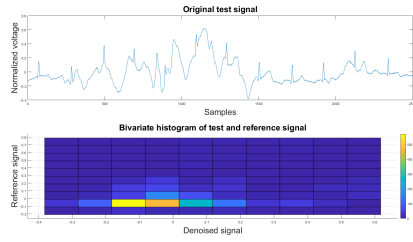Figure 4.17: Example 3: (1) test signal (2) bivariate histogram
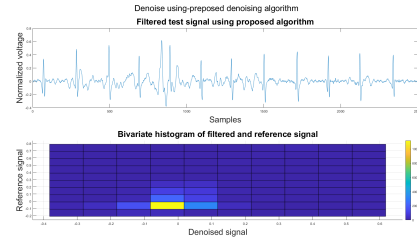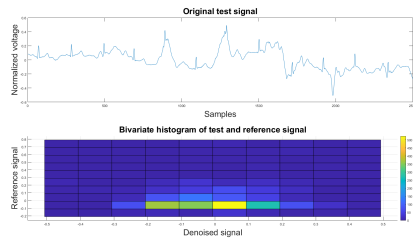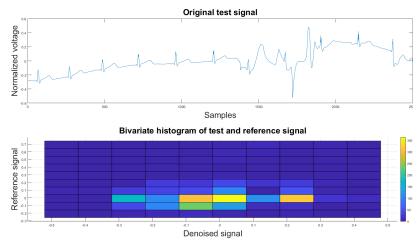

(a) Original signal
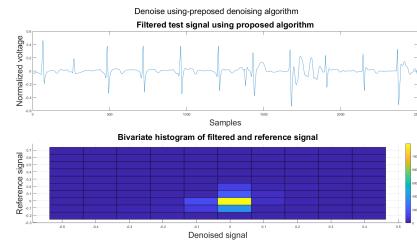
(b) Filtered signal

Figure 4.18: Example 4: (1) test signal (2) bivariate histogram

(a) Original signal

(b) Filtered signal
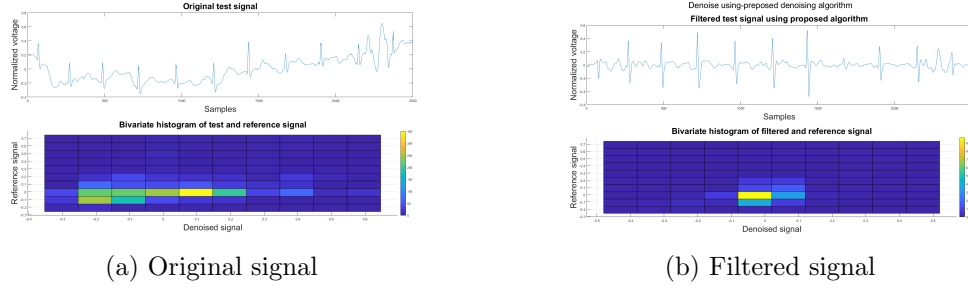
Figure 4.19: Example 5: (1) test signal (2) bivariate histogram

(a) Original signal

(b) Filtered signal
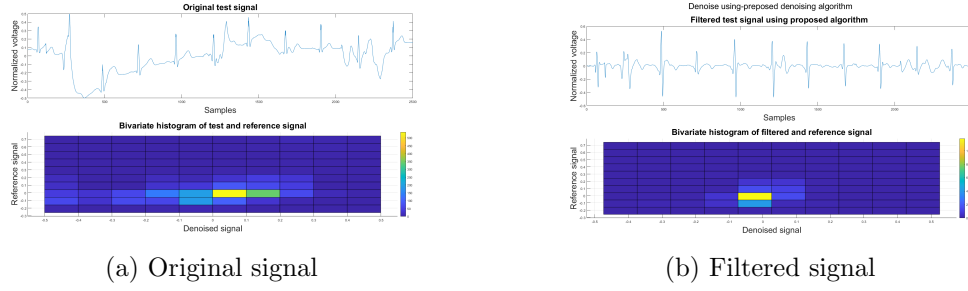
Figure 4.20: Example 6: (1) test signal (2) bivariate histogram

lation coefficient between the reference and the test signal. The correlation coefficient ranges from -1 to 1. As the coefficient is closer to 1, the two signals are more similar. On the other hand, as the coefficient is close to -1, the two signals have a negative relationship. If the coefficient is closer to 0, the two signals are not similar. The correlation coefficient is calculated using the following equation:

$$\rho(R,T) = \frac{1}{N-1} \sum_{i=1}^{N} \frac{R_i - \mu_R}{\sigma_R} \frac{T_i - \mu_T}{\sigma_T} \qquad (4.14)$$

where R, T refer to reference signal and test signal, N is the number of samples of the signal (the two signal have the same length), μ is the mean of the signal, and σ is the standard deviation of the signal.

4.3.3   Mean Squared Error

Another metric we have used to compare the algorithms is the mean squared error which measures the difference between two vectors. If the two

Table 4.2: Comparison between different noise removal methods

Figure	Metric	Original	IIR	MA	DWT	EMD	VMD	AF	Proposed
4.15	CORR	-0.0337	0.2962	0.3169	0.3303	0.4540	0.4395	0.4594	**0.5369**
4.15	MSE	0.0419	0.0326	0.0307	0.1357	0.0090	0.0260	0.0086	**0.0047**
4.16	CORR	0.0651	0.3072	0.3442	0.2507	0.4981	0.4712	0.4988	**0.5902**
4.16	MSE	0.0163	0.0129	0.0123	0.3053	0.0066	0.0092	0.0065	**0.0040**
4.17	CORR	0.1316	0.1846	0.2186	0.1307	0.2719	0.3047	0.3660	**0.4243**
4.17	MSE	0.1091	0.0425	0.0369	0.1000	0.0412	0.0302	0.0092	**0.0054**
4.18	CORR	0.1557	0.3438	0.3181	0.1638	0.1881	0.3031	0.2929	**0.4409**
4.18	MSE	11.1281	1.3549	10.2285	8.403	3.0299	0.9603	0.5916	**0.1647**
4.19	CORR	0.2992	0.4227	0.4393	0.3129	0.4220	0.4759	0.4867	**0.6011**
4.19	MSE	4.0803	0.7239	0.7073	3.7537	0.9917	0.3016	0.3188	**0.1144**
4.20	CORR	0.2875	0.4844	0.4523	0.051	0.3839	**0.5292**	0.4867	0.5242
4.20	MSE	5.8771	0.3676	0.4371	5.6547	0.3931	0.4822	0.3188	**0.1304**

vectors are more similar, the value is close to 0. The mean squared error can be calculated using this equation:

$$MSE = \frac{1}{n} \sum_{i=1}^{N} (R - T)^2 \tag{4.15}$$

where R, T refer to reference signal and test signal and N is the number of samples of the signal.

4.4   Result Discussion

In Table 4.2, the first column refers to the figures shown in Section 4.3.1. The remaining columns show the corresponding mean square error (MSE) and the correlation values between the reference and filtered noisy signals.

According to Sections 4.3.1, 4.3.2 and 4.3.3, both visual results and metrics in Table 4.2 show that the proposed algorithm can reduce the baseline wander and suppress motion artifacts to a limit. The histogram of the proposed algorithm always shows a normal Gaussian distribution which is more similar to the clean signal example in Figure 4.14.

Therefore, the proposed algorithm can preserve the QRS complex information and perform best when the QRS complex is complete in the ECG signal. The algorithm may not reconstruct the P wave and T wave completely when the noise signal contaminates these two waves, such as in Figure 4.15. However, if the body motion causes the entire heartbeat to move, this algorithm can put the moved heartbeat back to its location, such as the result examples

shown in Section 4.3.1. According to Table 4.2, one can see that the signal filtered by the proposed algorithm has a lower MSE and a higher correlation coefficient compared to the algorithms found in the literature. Moreover, based on over 600 segments of 10 seconds each, the average processing time is 1.5201 seconds per segment using an average computer programmed in Matlab.

4.5   Conclusion

This chapter presents a new method that can automatically remove the baseline wander and reduce the motion artifacts. The algorithm contains five major steps. The first one is the usable ECG signal detection step (Section 4.2.2). This step keeps the ECG signal that can be read and discards unreadable signals using a trained SVM model. The second step is an adaptive empirical mode decomposition and reconstruction algorithm (AEMDR) (Section 4.2.3). It can separate the signal components that contain QRS complex and the components that include low-frequency noise, such as baseline wander. This step produces an ECG signal and a noise signal. The third is the motion-sensitive signal generation, which combines accelerometer signals and AEMDR extracted noise signals together to form a motion-sensitive signal (Section 4.2.4). This signal is used as the reference signal for the adaptive filter in the fourth step (Section 4.2.5). This step removes motion artifacts with a similar frequency range as the QRS complex and produces a cleaned signal. The last step is the variational mode decomposition and reconstruction (VMDR). The method can remove the high-frequency components that are left in the ECG signal to produce a smoother signal (Section 4.2.6). The results in Sections 4.3.1–4.3.3 clearly show improvements in the quality of the test ECG signals compared to various algorithms found in the literature. In addition, the ECG recordings were collected from two different ECG devices: Astroskin Smart Vest and the VivaLnk ECG device. Therefore, the proposed noise removal method is robust and can be used for various medical-grade sensors.

Chapter 5

Anomaly Detection from Electrocardiogram(ECG) Using Spectrogram and Handcrafted Features

5.1 Training Dataset and Methodology

The following section describe the public ECG dataset used to test and train the proposed algorithm (see Table 5.1 and 5.2 third columns).

5.1.1 Data Description

The test ECG datasets are from the MIT-BIH database [63] and the European ST-T database [88] both located in the PhysioNet database [28]. The MIT-BIH data contains 48 half-hour excerpts of two-channel ambulatory ECG recordings obtained from 47 subjects. The data is fully annotated with 15 different rhythm classifications. The European ST-T database consists of 92 annotated 2 hours excerpts of two-channel ambulatory ECG recordings from 79 subjects with ten different rhythm classifications, and a complete description of the rhythms is shown in Table 5.1. Similarly, a complete description of the heartbeat is shown in Table 5.2. In these tables, M is short for the MIT-BIH database, and E is short for the European ST-T Database.

Table 5.1: Complete Description of Rhythms Types

Symbol	Meaning	Database
NSR	Normal sinus rhythm	M, E
AB	Atrial bigeminy	M, E
AFIB	Atrial fibrillation	M, E
AFL	Atrial flutter	M
B	Ventricular bigeminy	M, E
BII	2° heart block	M
B3	3° heart block	E
IVR	Idioventricular rhythm	M
NOD	Nodal (A-V junctional) rhythm	M
P	Paced rhythm	M
PREX	Pre-excitation (WPW)	M
SAB	Sino-atrial block	E
SBR	Sinus bradycardia	M, E
SVTA	Supraventricular tachyarrhythmia	M, E
T	Ventricular trigeminy	M, E
VFL	Ventricular flutter	M
VT	Ventricular tachycardia	M, E

5.1.2 Methodology

As described previously, anomaly detection can be divided into an abnormal rhythm and heartbeat detection. First, abnormal rhythm detection will find the abnormal rhythms in the ECG record and label them. However, some rhythms may be labelled as normal rhythms, and they could contain irregular heartbeat (An example one can see in Figure 5.1). Therefore, the second part of the algorithm will go through all heartbeats in the normal rhythm and label the irregular heartbeats. This section describes the steps of the algorithm and the features used to model the rhythms and heartbeat signal.

Signal Pre-processing

The signal pre-processing steps consist of re-sampling, normalization, data extraction, and balancing. The ECG signals in the MIH-BIH arrhythmia dataset have a 360 Hz sampling frequency, and the signals in the European ST-T dataset have a 250 Hz sampling frequency. To make the sampling frequency

Table 5.2: Complete Description of Heartbeat Types

Symbol	Meaning	Database
N	Normal	M, E
LBBB	Left bundle branch block beat	M
RBBB	Right bundle branch block beat	M
PAC	Atrial premature beat	M
a	Aberrated atrial rremature beat	M, E
J	Nodal(junctional) premature beat	M, E
S	Supraventricular premature beat	M, E
PVC	Premature ventricular contraction	M, E
F	Fusion of ventricular and normal beat	M, E
e	Atrial escape beat	M
j	Nodal (junctional) escape beat	M, E
E	Ventricular escape beat	M, E
P	Paced beat	M
f	Fusion of paced and normal beat	M, E
Q	Unclassified beat	M

at the same sampling rate, both signals were down-sampled at a rate of 250 Hz, with an antialiasing lowpass Filter provided by Matlab[61]. The algorithm starts by inserting 250 zeros to the original signal $S(t)_{orig}$ to get the up-sampled signal $S(t)_{up}$. Then $S(t)_{up}$ is filtered by an FIR antialiasing filter that uses a Kaiser window method to maintain the original shape of $S(t)_{orig}$. Finally, 250 samples in the $S(t)_{up}$ are discarded to get the down-sampled signal $S(t)_{down}$. Next, the signals $S(t)$ from both datasets are normalized between 0 and 1 using the min-max normalization formula.

Rhythm and Heartbeat Signal Extraction

The ECG signal is split into 3-second segments using a sliding window method to extract the rhythm data. The window size is chosen to be $3 \times$ sample frequency, and the step size is chosen to be $1 \times$ sample frequency. The heartbeat data are extracted using the provided heartbeat annotation location. The R peak locations of the last heartbeat and the following R peak will be the beginning and ending points of the current heartbeat. After rhythm and heartbeat extraction, the extracted data is used to compute the spectrogram
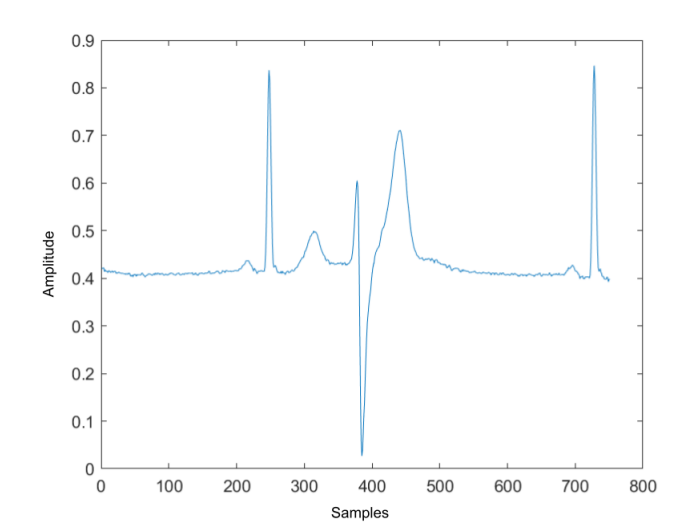
Figure 5.1: Normal rhythm contains abnormal heartbeat

and features for data modelling.

Spectrogram Using Short-Time Fourier Transform (STFT)

Traditional signal processing approaches assume that the signal is stationary, which is not true, especially for ECG. Hence, time or frequency descriptions alone are insufficient to provide a comprehensive analysis of the ECG signals [83]. Like speech analysis, time-frequency analysis is much more suitable for non-stationary signals. Using STFT, one can determine the frequency and phase content of local sections of a signal as it changes over time [83]. STFT is a sequence of Fourier transforms of a windowed signal. The STFT matrix is the combination of the Discrete Fourier Transform (DFT) of each windowed segments which can be expressed as $X(f) = [X_1(f), X_2(f), X_3(f), ..., X_m(f), ..., X_k(f)]$ [62], where k is the number columns of the spectrogram.

To compute the STFT spectrum, one must choose the window function and size. The window function selected must be adapted to the signal frequencies. The window size affects the time and frequency of information presented in the spectrogram. If the spectrogram is computed using a long window, then the spectrogram has higher frequency accuracy but less time resolution. On the other hand, if the spectrogram is calculated using a short window, then
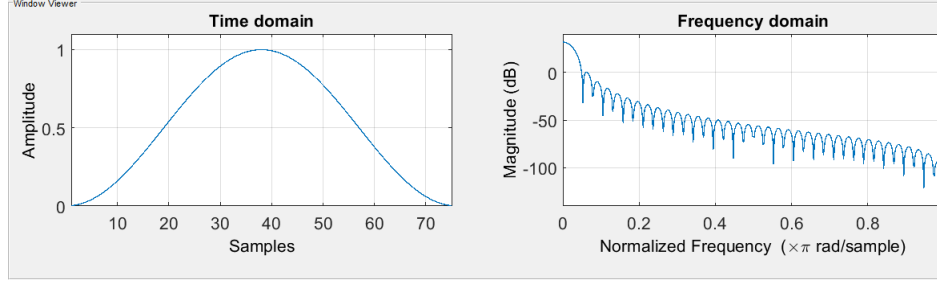
69

Figure 5.2: Hanning window function in the time domain and the frequency domain

the spectrogram has lower frequency accuracy but better time resolution. The window size from 0.1 seconds to 1 second was tested to find a balance between frequency and time resolutions. In this search, based on many experiments, a window length L of 167 samples was chosen for the rhythm data, and 28 samples were chosen for the heartbeat data, which is 0.668 seconds and 0.11 seconds correspondingly. The best window function for the STFT is the Hanning window for both data (shown in Figure 5.2). The Hanning window was selected because it ranged from 0 to 1 in amplitude, matching the normalized ECG signal range. And the shape of the Hanning window function is similar to the waves of the ECG signal.
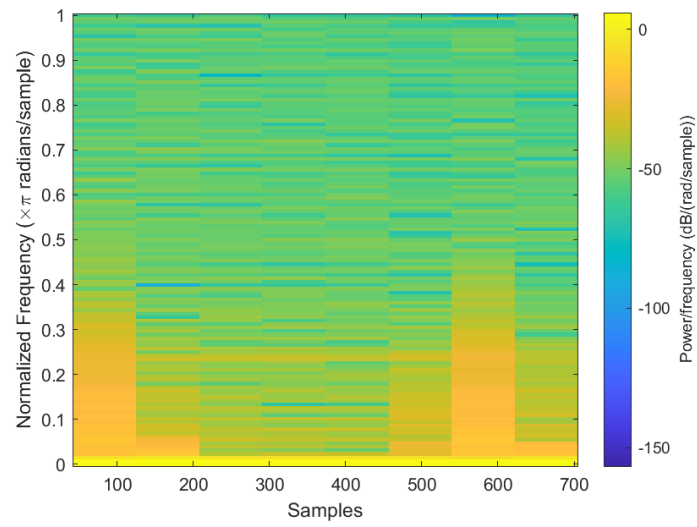
Examples of the spectrogram and its corresponding ECG rhythm and heartbeat can be seen in Figure 5.3, Figure 5.4, Figure 5.5 and Figure 5.6. The y axis for the ECG signal were set to normalized voltage.

Handcrafted Features Rhythm Classification

Sometimes, humans can quickly identify relationships between two objects where a computer cannot. In the proposed research, in addition to the extracted features from the Convolutional Neural Network, nine extra handcrafted features were added to improve the recognition performance. The feature vector for rhythm classification contains the number of R peaks detected from two algorithms, skewness, kurtosis, variance, average interval length, average QRS complex length, average PR interval length, and average QT interval length. The features will be explained in the following sections.

70
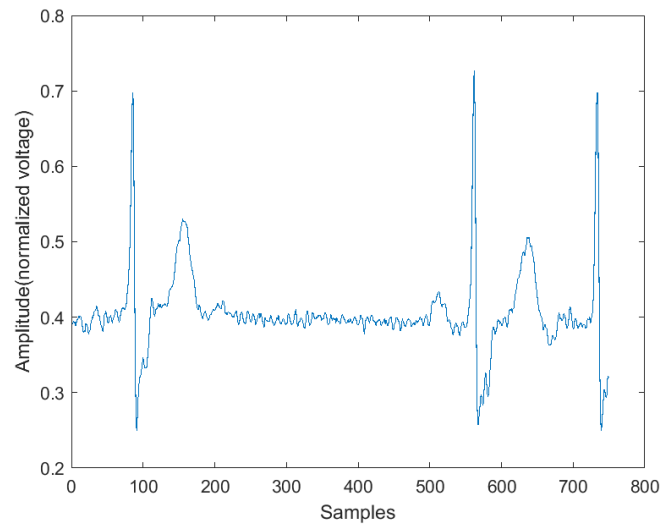
(a) abnormal rhythm plot

(b) abnormal rhythm spectrogram

Figure 5.3: Abnormal ECG rhythm and its corresponding spectrogram

71
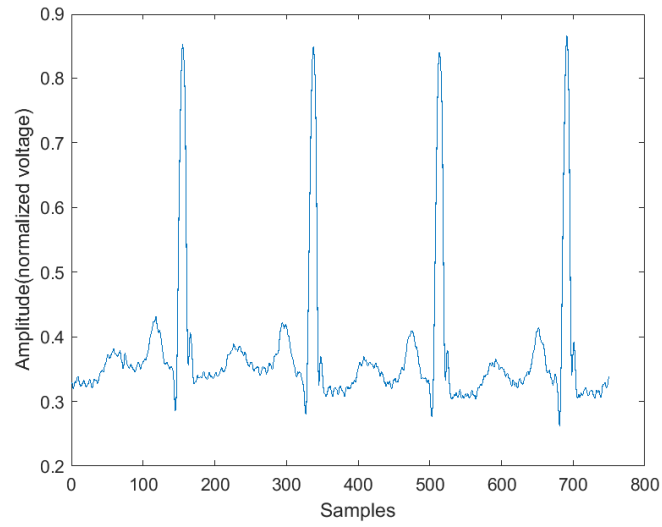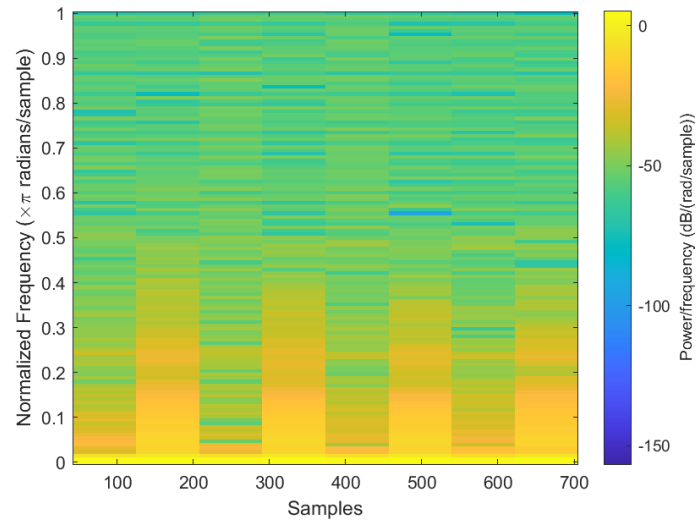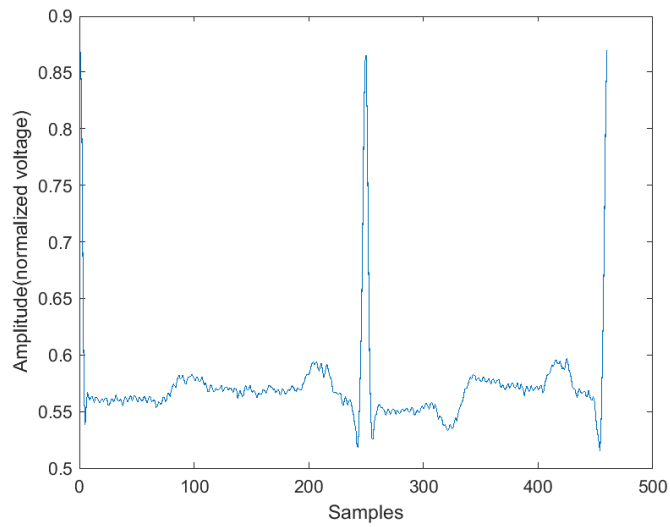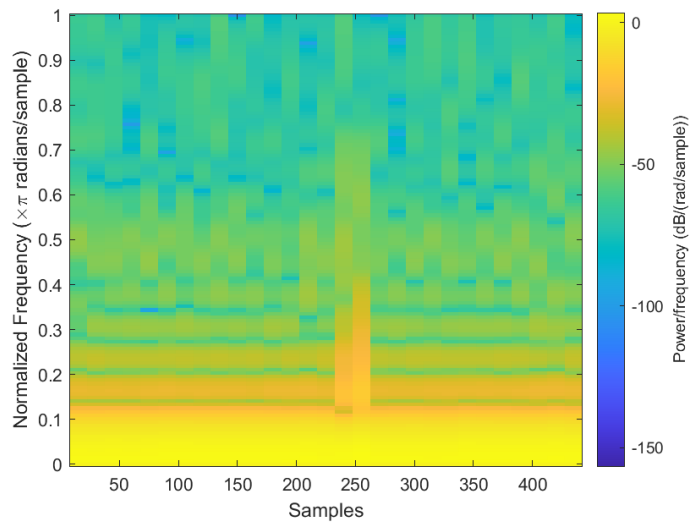
(a) Normal rhythm plot

(b) Normal rhythm spectrogram

Figure 5.4: Normal ECG rhythm and its corresponding spectrogram
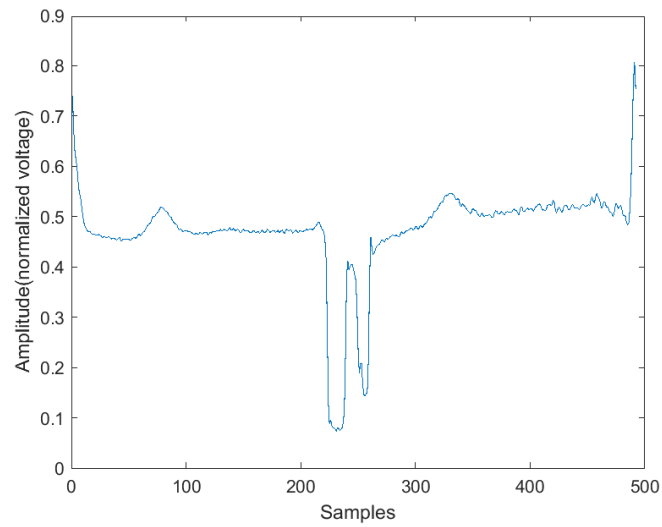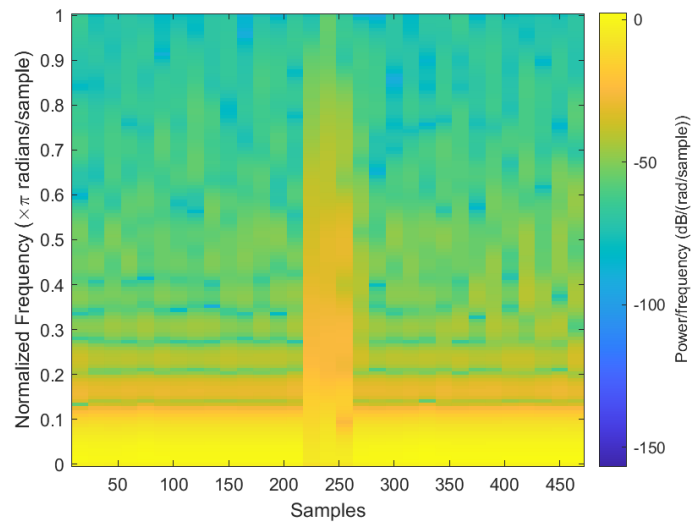
(a) Normal heartbeat plot

(b) Normal heartbeat spectrogram

Figure 5.5: The normal heartbeat plot and spectrogram

(a) Abnormal heartbeat plot

(b) Abnormal heartbeat spectrogram

Figure 5.6: The abnormal heartbeat plot and spectrogram

74

Detecting the Number of R Peaks

For this feature, two algorithms were used to detect the R peaks. One is the Pan-Tompkins algorithm [69], and the other one is the Shannon Energy Envelope Hilbert transform (SEEHT) [60]. The detected R peaks were included in the final feature vectors. The reason for using two R peaks detection algorithms is that they perform differently for different ECG signals. The Pan-Tompkins uses a filter bank that only keeps the R peaks on the ECG signal. The SEEHT algorithm calculates the Shannon energy envelope of each heartbeat on the ECG signal and then applies Hilbert transform to the Shannon energy signal. By finding the zero-crossing of the converted signal, one can detect the R peak locations. By comparison, the SEEHT has higher overall detection accuracy, and the Pan-Tompkins has higher abnormal heartbeat detection accuracy. For example, if one selects a record from the MIT-BIH database with 472 ventricular flutter (VF) beats, the Pan-Tompkins algorithm will detect 363 VF beats, and the SEEHT will only detect 177 VF beats. Therefore, using both algorithms as features can improve detection accuracy.

Statistical Features

Another handcrafted is a skewness measurement that computes the asymmetry of the ECG signal about its mean. Negative skewness indicates the signal is leaning right, and positive skewness indicates the signal is leaning left. The skewness estimate can be computed using the following equation:

$$skewness = \frac{\mu^3}{\sigma^3},$$ (5.1)

where the μ is the mean, and the σ is the standard deviation.

The kurtosis describes the 'peak' of a signal. The kurtosis should be zero for a perfectly normal distribution. A signal with positive excess kurtosis indicates a high peak, and a signal with negative kurtosis indicates a flat-topped curve [46]. The kurtosis can be computed using the following equation:

$$kurtosis = \frac{\mu^4}{\sigma^4}, \qquad (5.2)$$

where the μ is the mean, and the σ is the standard deviation.

The variance of a signal computes how far the sample points are away from their mean. It can be computed using the following equation:

$$Variance = \frac{1}{N-1} \sum_{i=1}^{N} |X_i - \mu|^2, \qquad (5.3)$$

where X_i is the signal, N is the length of the signal, and μ is the mean of the signal.

ECG Intervals Measurement

Four ECG interval measurements are also used as features. They are average RR interval length, average QRS complex length, average PR interval length, and average QT interval length. The detection tool used for detecting the Q wave and T wave is the 'ecgpuwave' program, which is provided by PhysioNet [28]. The 'ecgpuwave' uses Pan-Tompkins algorithm [69] to detect the QRS complex. Then the detector detects the P peak and T peak based on the detected R peaks. For P and T peak detection, the detector defined search windows before and after the R peak. The detector searches the maximum and the minimum values in the defined window. The P and T wave peak is assumed to occur at the zero-crossing between the maximum and the minimum values in the search windows. Figure 5.7 (a) and Figure 5.7 (b) show the examples of the P-QRS-T detection using the 'ecgpuwave' detector. These two examples represent the most common situation for detecting the P, R and T peaks on the heartbeats with the 'ecgpuwave' detector, which are the peaks on abnormal heartbeats that cannot be fully detected and the detection on normal heartbeats are perfectly detected. Therefore, for undetected peaks, their corresponding interval length and peak value will be set to 0.
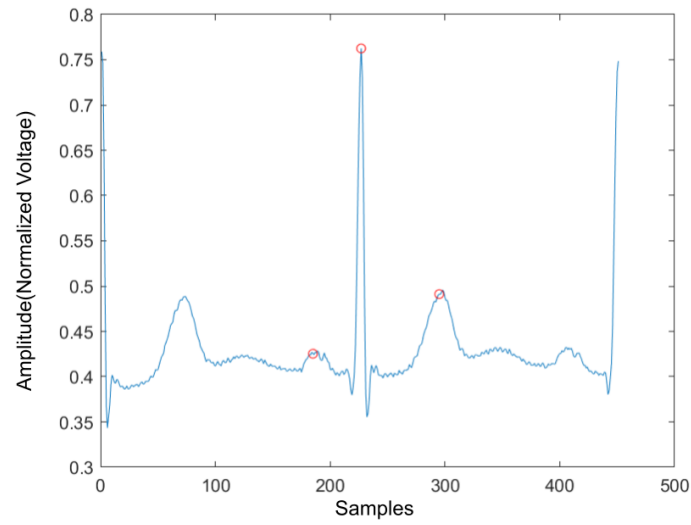
(a) P,R,T peak detection on normal heartbeat
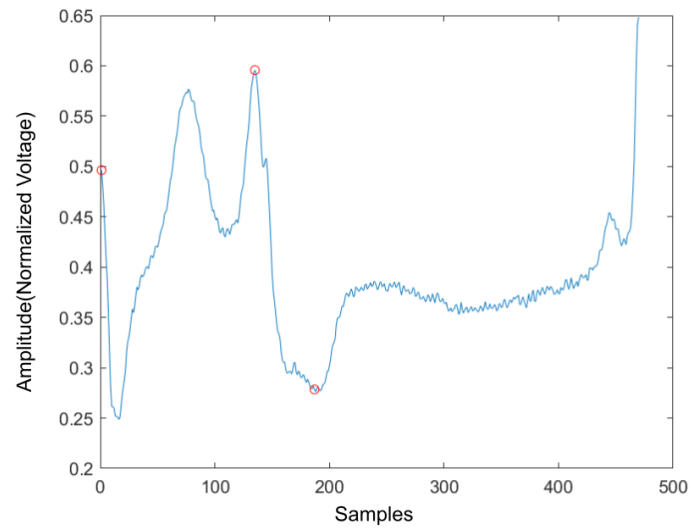
(b) P,R,T peak detection on abnormal heartbeat

Figure 5.7: The P,R,T peak detection on both normal and abnormal heartbeat

Handcrafted Features for Heartbeat Classification

The previous statistical features for the rhythm are also used for the heartbeat signal: skewness, kurtosis, and variance. But there are still some features that are different from the rhythm. Since there is only one heartbeat in the heartbeat signal, the number of R peaks, average RR interval length, average QRS complex length, average PR interval length, and average QT interval length no longer exists. They are replaced with normalized R peak value, left RR interval length, right RR interval length, PR interval length, QRS complex length, QT interval length, In addition, as shown in Figure 5.7, the 'ecgpuwave' detector could detect the peak location of the heartbeat signal, and the P wave and T wave peak value could also be retrieved. Therefore, the final two features added to the feature vector are normalized P wave peak and T wave peak values.

To sum up, there are a total of 11 features in the handcrafted feature vector, which are normalized R peak value, left RR interval length, right RR interval length, skewness, kurtosis, variance, QRS complex length, PR interval length, QR interval length, normalized P wave peak value, and normalized T wave peak value.

Network Architecture

The proposed neural network architecture consists of a feature extraction layer, a normalization layer, a CNN block, and a feature merging layer. The proposed CNN structure is shown in Figure 5.8. According to the figure, the original input is the raw ECG signal. From the original input, the program could produce the STFT spectrogram and the handcrafted features in Section 5.1.2. Once the handcrafted features are obtained, the normalization layer will normalize those features to a range between [0,1] so that each feature contributes the same weight to the classification. The STFT spectrogram is sent to the CNN block and a sigmoid layer. The CNN block could be any popular CNN structure such as VGG-19 CNN, plain 34 layers CNN, 34-layers Residual Net(shown in Figure 5.9). The primary purpose of the CNN block
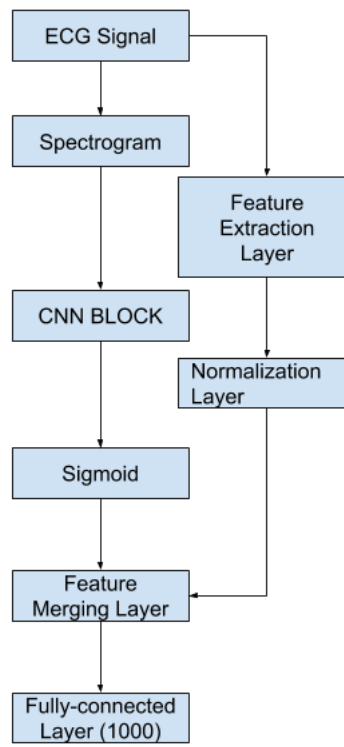
Figure 5.8: The proposed CNN

was to extract the features from the spectrogram using convolutional layers. The sigmoid layer acts as a normalization layer converting CNN features to be in the range [0,1] so that the CNN features will have the same weight as the handcrafted features at the final layer. Once the CNN and handcrafted features are obtained, the feature merging layer then combines the normalized handcrafted features with the extracted features from the spectrogram. Finally, the merged feature vector will input a fully connected layer to produce the final classification. The output value of the fully connected layer is in the range of 0 and 1 where the value of the last classification label will be decided by:

$$
label = \begin{cases} normal & \text{if x} <= 0.5 \\ abnormal & \text{if x} > 0.5 \end{cases} \tag{5.4}
$$

where x is the output value of the fully connected layer.

5.2 Experimental Setup and Result Discussion

5.2.1 Dataset Setup

As described previously, there are 17 different rhythm types in the data sets. However, in the experiment, the data were classified as 1 for abnormal and 0 for normal.

There are 26,532 abnormal ECG rhythms for the rhythm dataset and 1,199,065 normal ECG rhythms. Since there were much more normal rhythm data than anomalous rhythm data, dataset balancing was necessary. The normal rhythm dataset was constructed by randomly selecting 26,532 samples from the original normal rhythm dataset. As a result, the reconstructed dataset has 26,532 abnormal ECG rhythms and 26,532 normal rhythms. For the final dataset, 53,064 ECG rhythms were generated. A total of 90% of the data were used as the training dataset, and 10% were used as the testing dataset. Therefore, the training set contains 47,758 ECG rhythms, and the testing dataset contains 5,306 ECG rhythms.

For generating the heartbeat dataset, there is a slight difference. According
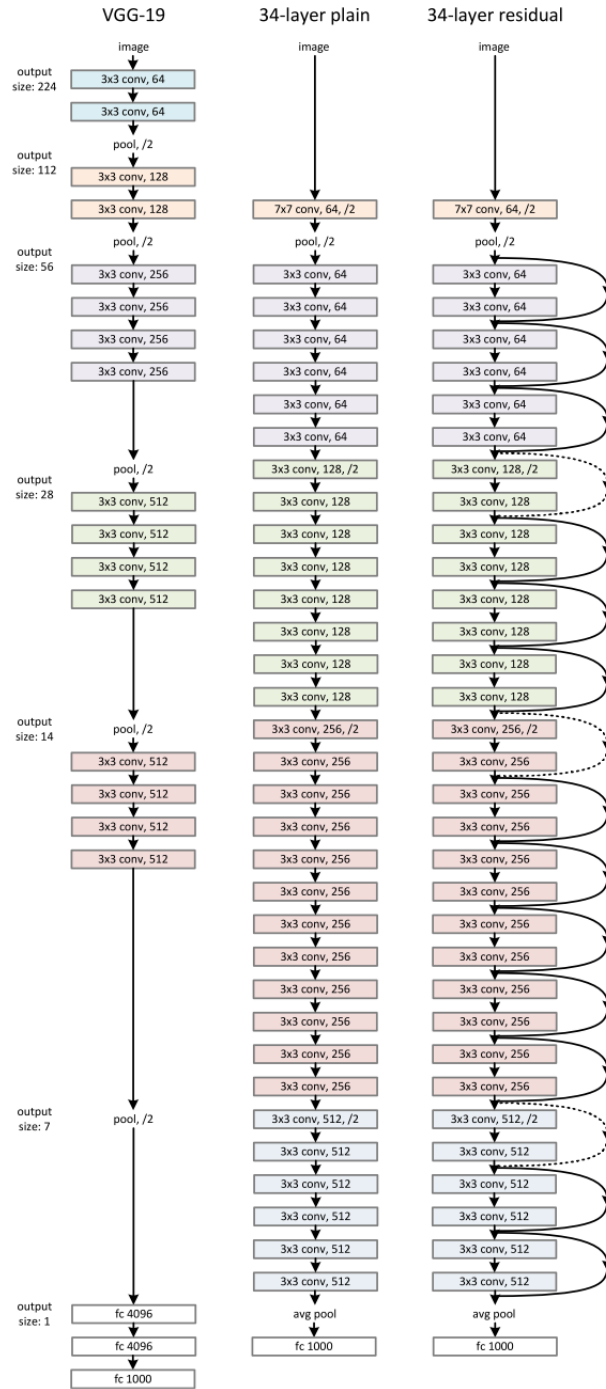
80

Figure 5.9: VGG-19 CNN, 34-layer plain CNN, 34-layer ResNet

to Section 5.1.2, the only abnormal heartbeats that occur in normal rhythm are used for training and testing. There are no Right Bundle Block Beat in these two databases, and Pace Beat occurs in the normal rhythm. Therefore, there were 14 heartbeat types in the training and testing dataset. Similar to the generation of the rhythm dataset, the training and testing dataset split follows a 9:1 ratio. Finally, the generated training dataset contains 43,516 heartbeats, and the testing dataset contains 4,850 heartbeats. Moreover, the normal and abnormal heartbeats in both datasets are 50-50 split so that the testing result will be non-biased.

The spectrograms were generated using STFT described in Section 5.1.2. The dimension of the output RGB spectrograms is $875 \times 656 \times 3$. The spectrogram power values are encoded as an RGB colour image to emphasize the difference in one version. The power spectrum power values are coded as a greyscale image where an integer represents each pixel in a second version.

5.2.2   Model Training and Testing

A total of 80% data is used as training during the training step, and 20% data is used as validation. The learning rate decay technique is used during the training to determine the best learning rate. The learning rate decay adopts an initially large learning rate and then decays it by a certain factor after pre-defined epochs. According to You's study, an initially large learning rate suppresses the memorization of noisy data while decaying the learning rate improves the learning of complex patterns [105]. In the experiment, the start learning rate is 0.1, and the decay rate is 50%. The neural network architectures used in the CNN block are VGG-16, VGG19, ResNet-18 and ResNet-34. The ResNet-50, ResNet-101 and ResNet-150 architectures are not tested here because these two models have too many layers, so the models' prediction speed will be slow. Therefore the models won't be able to give a diagnosis in real-time. Moreover, to show how the feature merging layer could increase the prediction accuracy, regular VGG 16,19 and Residual Neural Network 18 and 34 were trained as well (Figure 5.9). The results are compared in Table 5.3.

Multi-lead Utilization

As described in Section 5.1.1, the MIT-BIH database and the European ST database are both collected with two-channel ECG devices. Therefore, a DNN model for each lead signal was trained to utilize the multi-lead signal. The models for each lead were all used for the final classification. As a result, the final prediction is calculated using the following equations:

$$Y = Y_1 + Y_2 - 0.5 \tag{5.5}$$

where Y_1 is the prediction from lead one model and Y_2 is the prediction from lead two model. The label is determined using the following conditions:

$$Y = \begin{cases} 1, & \text{if } Y \geq 0.5 \\ 0, & \text{otherwise} \end{cases} \tag{5.6}$$

The models could also be used for individual prediction for another dataset with lead-II channel data. Therefore, the models can be used for further learning with other ECG datasets such as 12-lead ECG database.

Over-fitting Prevention

Most deep neural network architectures are prone to over-fitting, meaning that if the accuracy of the network on the training set is increased, the accuracy of the network on the actual test set may not improve. This situation indicates that the model is overfitting the training set and cannot generalize all data. During the training of the CNN model, the following two methods were used to prevent over-fitting. The first one consists of adding a dropout layer after the fully connected layer, and the second consists of using an early stopping technique, as shown by [75]. After each training epoch, the model predicts a validation set to determine accuracy. The training process stops if the validation accuracy drops two times in two consecutive epochs. The training and testing were computed on a workstation with Intel Core i5-8400 CPU and Nvidia Geforce GTX 1080 GPU. One can see in Figure 5.10 the training and validation of rhythm classification loss function value vs. epochs. Both training loss and validating loss converge around 20 epochs.
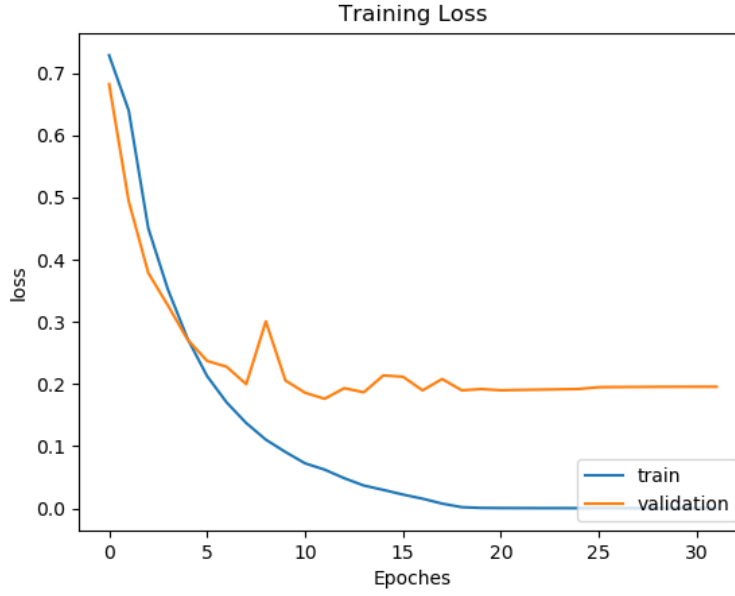
Figure 5.10: Training and validation loss curve for rhythm classification

5.2.3 Evaluation Metrics

The evaluation metrics used for comparison with ground truth are the following:

- TP: The number of successfully detected abnormal rhythms/heartbeats;

- FP: The number of wrongly detected abnormal rhythms/heartbeats;

- TN: The number of successfully detected normal rhythms/heartbeats;

- FN: The number of wrongly detected normal rhythms/heartbeats;

- Sensitivity(SEN) = TP/(TP+FN);

- False Alarm Rate(FAR)= 1 - Specificity = FP/(FP+TN);

- Positive Predictive Value(PPV) = TP/(TP+FP);

- Accuracy(ACC) = (TP+TN)/(TP+FP+TN+FN).

The main goal is to find the most abnormal rhythms and heartbeats on the ECG signal. The chosen metrics could show whether the algorithm detects

abnormal cases. The sensitivity is also called the true positive rate, which shows the detection rate among all abnormal cases. The false alarm shows how many normal cases are detected as abnormal cases. The positive predictive value indicates the number of real abnormal cases among all detected abnormal cases. The accuracy is the overall accuracy for both normal and abnormal detection.

5.2.4 Comparison Results and Discussion

Rhythm Classification Result Discussion

In Table 5.3, one can compare the result of the proposed algorithm and its comparison to the state-of-arts. In the Table, 'RGB' indicates that the model was trained with an RGB spectrogram which contains three channels, and 'Gray' suggests that the model was trained with a grayscale spectrogram which only includes one channel. Regular ResNet corresponds to the results obtained by the typical Residual neutral network, shown in Figure 5.9, that are trained with only the spectrograms. Proposed ResNet corresponds to the new Residual Neural Network architecture, shown in Figure 5.8. The 'Time(s)' column shows how many seconds the models used to predict the test dataset.

Compared to results shown in the Table 5.3, the proposed models achieved better accuracy and could detect more various rhythm types. Furthermore, the proposed model with feature merging layers performs better than the regular ResNet models. In addition, when comparing the proposed model to Zihlmann's model [110] which is only trained with a spectrogram, one can see that our approach had much higher accuracy. This demonstrates that prior knowledge like handcrafted features can increase the accuracy of heart rhythm classification.

The accuracy difference between the RGB and Grayscale spectrogram trained models should be minimal. Since the RGB spectrogram uses different colours to express power density, the grayscale spectrogram expresses the power density using its pixel intensities. However, as the effect of model training, the RGB spectrograms increase the learning capacity of the model since it

contains more channels. The result in Table 5.3 supports this finding as well. The average accuracy difference between VGG networks is around 3%. The accuracy difference between regular ResNet is around 1%. In addition, the accuracy difference between ResNet and with feature merging layer is none. For a regular residual neural network, the models trained with RGB spectrograms have a lower prediction accuracy (about 1%) than those trained with Grayscale spectrograms. This indicates that the model could be overwhelmed by irrelevant information containing three RGB spectrogram channels. However, there is no significant difference between the models trained with RGB spectrogram and those trained with grayscale spectrogram after adding the feature merging layer to the regular residual neural network. This shows that the handcrafted features have a heavier weight than the CNN layer extracted features for modelling the rhythm signal. From the time aspect, the Grayscale spectrogram trained models took about half the time as the RGB spectrogram trained models.

For the VGG Neural Network architecture, the models trained with grayscale spectrograms are better than those trained with RGB spectrograms. The accuracy for models trained with both spectrogram and handcrafted features is slightly better than those trained with spectrograms only. The VGG19 models have the best overall accuracy compared to the VGG16 models. For the Residual Neural Network architecture, the ResNet18, ResNet34, and ResNet50 have different layers, where the ResNet18 has the least layers, and the ResNet50 has the most layers. The number of layers also indicates the number of trainable variables. More trainable variables mean that the model is harder to learn, in other words, the model takes a long time to converge. Also, during the prediction step, the larger network takes a longer time than the smaller network. The 'Time' column in the Table also shows that the ResNet18 models used less time than ResNet34 and ResNet50 models. In addition, the ResNet50 models did not have the best performance in the experiment. The reason could be that more trainable variables in the network caused the model to be easier to overfit to the training dataset than the other models; therefore, it performed worse on the test dataset. By comparing the VGG and ResNet architectures, it is

clear that the ResNet architecture performs better than the VGG architecture on accuracy and takes less time for a prediction.

The proposed ResNet18-Gray held the highest accuracy with the third-lowest false alarm rate by comparing all the models. The other proposed ResNet-18 and ResNet-34 network models also have similar performances. However, between these models, the proposed ResNet-18 Gray only took 55 seconds to complete the prediction of the test dataset, which is the fastest model compared to other models. Therefore, the system should produce a high anomaly detection accuracy and low false alarm rate in clinical heart anomaly detection. Also, the abnormality detection system should predict as fast as possible. Therefore the proposed ResNet18-Gray model should be the best choice for clinical applications.

Table 5.3: Classification Result of Rhythms

Method	A/N	Types	TP	FP	TN	FN	SEN	FAR	PPV	ACC	Time(s)
AR modeling[26]	713/143	6	706	10	133	7	99.02%	6.99%	98.60%	98.01%	N/A
Acharya Net A[1]	20807/902	4	19160	62	840	1647	92.08%	6.87%	99.68%	92.13%	N/A
Acharya Net B[1]	8322/361	4	7946	376	294	67	99.16%	56.12%	95.48%	94.90%	N/A
Zihlmann[110]	-	4	-	-	-	-	-	-	-	82.3%	N/A
Regular VGG16-RGB	2653/2653	17	2522	185	2468	131	95.06%	6.97%	93.17%	94.04%	433
Regular VGG19-RGB	2653/2653	17	2436	235	2418	217	91.82%	8.86%	91.20%	91.48%	452
Regular VGG16-Gray	2653/2653	17	2554	215	2438	99	96.27%	8.10%	92.24%	94.08%	332
Regular VGG19-Gray	2653/2653	17	2544	116	2537	109	95.89%	4.37%	95.64%	95.76%	349
Regular ResNet18-RGB	2653/2653	17	2536	121	2532	117	95.59%	4.56%	95.45%	95.51%	117
Regular ResNet34-RGB	2653/2653	17	2550	131	2522	103	96.12%	4.94%	95.11%	95.59%	152
Regular ResNet18-Gray	2653/2653	17	2599	155	2498	54	97.96%	5.84%	94.37%	96.06%	61
Regular ResNet34-Gray	2653/2653	17	2635	150	2503	18	99.32%	5.65%	94.61%	96.83%	96
Proposed VGG16-RGB	2653/2653	17	2433	173	2480	220	91.71%	6.52%	93.36%	92.59%	425
Proposed VGG19-RGB	2653/2653	17	2476	199	2454	177	93.33%	7.50%	92.56%	92.91%	446
Proposed VGG16-Gray	2653/2653	17	2567	172	2481	86	96.76%	6.48%	93.72%	95.14%	319
Proposed VGG19-Gray	2653/2653	17	2572	134	2519	81	96.95%	5.05%	95.05%	95.95%	345
Proposed ResNet18-RGB	2653/2653	17	2648	15	2638	5	99.81%	0.57%	99.44%	99.62%	110
Proposed ResNet34-RGB	2653/2653	17	2646	10	2643	7	99.74%	0.38%	99.62%	99.68%	144
Proposed ResNet18-Gray	2653/2653	17	2646	4	2649	7	99.74%	0.15%	99.85%	99.79%	55
Proposed ResNet34-Gray	2653/2653	17	2649	8	2645	4	99.85%	0.30%	99.70%	99.77%	90

Heartbeat Classification Result Comparison

In Table 5.4, one can compare the result of the proposed method with state-of-art algorithms. The naming of the proposed model is similar to the one described in Section 5.2.4.

Similar to the rhythm classification, three different datasets were tested using both proposed ResNet and regular ResNet. As one can see in Table

5.4, the proposed ResNet models are better than the standard ResNet models in all metrics. Moreover, the proposed ResNet Models also performed better than Sen's algorirhm [84] which is only using spectrogram and CNN. Sen's algorithm achieved 99.57% sensitivity, 2.99% false alarm rate, 96.72% positive predictive value and 98.21% accuracy in detecting three heartbeats. The performance metrics were very similar to the regular ResNet-34 model and are 1% less accurate than the proposed ResNet-34 model. This shows that the handcrafted feature could increase the accuracy of ECG heartbeat signal recognition. The overall performance between the models that used the Grayscale spectrogram dataset and the RGB spectrogram dataset is similar for the rhythm classification. The accuracy difference for all metrics was about 1%. However, the 34-layers ResNet models had better performance than the 18-layers ResNet models on the heartbeat datasets for the number of ResNet layers. This is caused by the fact that there is less information on the heartbeat spectrogram since there is only one heartbeat. In addition, as shown in Table 5.4, the ResNet models were faster and more accurate than the VGG models. This result indicates that the residual convolutional neural network architecture is more suitable for processing the ECG STFT spectrogram.

Compared to results found in the literature, the proposed models generally performed better except when compared to Ubeyli's and Ye's algorithms. However, Ubeyli's model has only been tested on 359 heartbeats. Therefore the result may not be suitable for all different heartbeat types. Compared to Ye's algorithm, the performance is quite similar. Ye's algorithm is 0.18% better overall accuracy, but the proposed model has a better positive predictive value, which means the proposed algorithm could detect anomalies in ECG signals more accurately. One possible reason for this situation may be caused by Ye's test data sets containing more normal heartbeats than abnormal heartbeats. Therefore, correctly detecting a normal heartbeat could increase the overall detection accuracy. Another factor that could affect the model accuracy performance is that the training and testing data sets for the proposed model only contain the abnormal heartbeat found in normal rhythms. Therefore, the sample is less than the irregular heartbeat on all ECG signals. Still,

Table 5.4: Classification Result of Heartbeats

Method	A/N	Types	TP	FP	TN	FN	SEN	FAR	PPV	Accuracy	Time(s)
Christov [17]-morphology	18378/47239	5	18042	1604	45635	336	98.17%	3.40%	91.84%	97.04%	N/A
Christov [17]-frequency	18378/47239	5	17590	1459	45780	788	95.71%	3.09%	92.34%	96.58%	N/A
Chazal [20]-frequency	4317/34394	4	4108	1962	32432	209	95.16%	5.70%	67.68%	94.39%	N/A
Ubeyli [93]	269/90	4	268	2	88	2	99.26%	2.22%	99.26%	99.89%	N/A
Llamedo [56]	5441/44188	3	4752	2238	41950	689	87.34%	5.06%	67.98%	94.10%	N/A
Ye [103]	20745/65264	16	20557	286	64978	188	99.09%	0.44%	98.63%	99.32%	N/A
Zhang [106]	5653/44011	4	5248	4869	39142	405	92.84%	11.06%	51.87%	89.38%	N/A
Thomas [91]	26626/67268	5	22900	1300	65968	3726	86.01%	1.93%	94.63%	94.65%	N/A
Kiranyaz [48]	7366/42191	5	6539	1228	40963	827	88.77%	2.97%	84.19%	95.85%	N/A
Rajesh [78]	8000/2000	5	7677	33	1967	323	95.96%	1.65%	99.57%	96.44%	N/A
Sahoo [82]	807/244	4	798	5	239	9	98.88%	2.04%	99.38%	98.67%	N/A
Sen [84]	1154/1305	3	1149	39	1266	5	99.57%	2.99%	96.72%	98.21%	N/A
Regular VGG16-RGB	2425/2425	14	2196	75	2350	229	90.56%	3.09%	96.70%	93.73%	784
Regular VGG19-RGB	2425/2425	14	2171	58	2367	254	89.53%	2.39%	97.40%	93.57%	862
Regular VGG16-Gray	2425/2425	14	2276	45	2380	149	93.86%	1.86%	98.06%	96%	613
Regular VGG19-Gray	2425/2425	14	2359	22	2403	66	97.28%	0.91%	99.08%	98.19%	701
Proposed VGG16-RGB	2425/2425	14	2324	49	2376	101	95.84%	2.02%	97.94%	96.91%	774
Proposed VGG19-RGB	2425/2425	14	2188	57	2368	237	90.23%	2.35%	97.46%	93.94%	861
Proposed VGG16-Gray	2425/2425	14	2366	28	2397	59	97.57%	1.15%	98.83%	98.21%	612
Proposed VGG19-Gray	2425/2425	14	2367	28	2397	58	97.61%	1.15%	98.83%	98.23%	707
Regular ResNet18-RGB	2425/2425	14	2338	53	2372	87	96.41%	2.19%	97.78%	97.11%	388
Regular ResNet34-RGB	2425/2425	14	2354	60	2365	71	97.07%	2.47%	97.51%	97.30%	453
Regular ResNet18-Gray	2425/2425	14	2373	22	2403	52	97.66%	0.91%	98.47%	99.08%	212
Regular ResNet34-Gray	2425/2425	14	2377	19	2406	48	98.02%	0.78%	99.21%	98.62%	272
Proposed ResNet18-RGB	2425/2425	14	2390	26	2399	35	98.56%	1.07%	98.92%	98.74%	381
Proposed ResNet34-RGB	2425/2425	14	2394	32	2393	31	98.72%	1.32%	98.68%	98.70%	444
Proposed ResNet18-Gray	2425/2425	14	2402	18	2407	23	99.05%	0.74%	99.26%	99.15%	209
Proposed ResNet34-Gray	2425/2425	14	2396	11	2414	29	98.80%	0.45%	99.54%	99.18%	279

some rhythm classes are harder to classify since some heartbeat types are easy to detect but do not exist on normal rhythms, such as paced and Right Bundle block beats.

By comparison, the proposed ResNet34-Gray performs the best on the heartbeat classification. It has 98.80% sensitivity, 0.45% false alarm rate, 99.54% positive prediction value, and 99.18% accuracy. It took 279 seconds to complete the prediction on the test dataset, which is very fast compared to other models. Compared to the existing state-of-art algorithm, the proposed algorithm has a better predictive value which is higher by 0.9%, and the other three metrics are only lower from 0.01% to 0.3%. Therefore, the proposed algorithm has better performance in detecting ECG signal anomalies.

Discussion

Based on the comparison in Section 5.2.4 and Section 5.2.4, The proposed neural network models perform better than other algorithms in the Table 5.3 and 5.4. The novelty of the proposed neural network structure is that it combines the features extracted from the STFT spectrogram using the CNN layer with handcrafted features. Therefore, it utilizes information from both

time and frequency domains. In addition, the proposed algorithm can be easily adapted to different ECG databases with variable ECG signal lengths and detect both abnormal rhythms and abnormal heartbeats from the ECG signal. Most detection algorithms found in the literature are designed only to detect abnormal rhythms or heartbeats.

5.3 Conclusion

This chapter proposes a new CNN architecture that uses STFT spectrogram and handcrafted features to model the ECG signals. The proposed architecture could adapt the latest CNN architectures to process the ECG spectrogram and combine the extracted features with handcrafted features. The proposed algorithm could be used on rhythm and heartbeat classifications with corresponding handcrafted features. The ResNet architecture is the best architecture for ECG spectrogram. The proposed ResNet contains a feature merging layer that combines the CNN extract features from spectrograms and handcrafted features. A series of experiments have been conducted to show that the feature merging layers can improve the ResNet model's accuracy. According to the experimental result, the new algorithm has superior performance to the current stat-of-arts algorithms. The best model for rhythm classification in the experiment is the proposed ResNet18-Gray. It can classify 17 different rhythm types into normal and abnormal classes with an exceptional accuracy of 99.79%, a sensitivity of 99.74%, a false alarm rate of 0.15%, and a positive predictive value of 99.85% which makes this new algorithm the most practical algorithm for rhythm classification in the literature. The heartbeat classification is used only on normal rhythms to detect abnormal heartbeats in the proposed anomaly detection algorithm. Therefore, the proposed algorithm is better at detecting state-of-art heartbeat classification algorithms. However, the proposed algorithm has a decent detection rate. The best model for heartbeat classification in the experiment is the ResNet34-Gray. It could classify 14 different heartbeat types into normal and abnormal classes with an accuracy of 99.18%, a sensitivity of 98.80%, a false alarm rate of 0.45%,

and a 99.54% positive predictive value. The result shows that the proposed model has the best detection rate for abnormal heartbeat compared to other algorithms found in the literature. The proposed approach also illustrates that in some situations, the introduction of handcrafted features can significantly improve the performance of neural networks.

Chapter 6

Conclusion

6.1 Result Discussion

In this research, an automatic ECG anomaly detection system was proposed. The detection system was designed for processing ECG signals from wearable devices. The ECG signal collected by wearable devices can be easily contaminated with noises generated by the patient's body movement. Therefore, the proposed system must compensate for the artifacts created by the motion. Moreover, the proposed method uses deep learning models to detect abnormal rhythms and heartbeats to detect the anomaly in ECG signals.

The proposed noise removal algorithm could automatically reduce the effects of the baseline wander and reduce the motion artifacts. It contains five significant steps. The first one is the usable ECG signal detection which classifies the noise level of the signals using a trained SVM model. The second step is the AEMDR algorithm. It separates the signal components that contain QRS complex and the components that include low-frequency noise, such as baseline wander. This step produces an ECG signal and a noise signal. The third is the motion-sensitive signal generation, which combines accelerometer signals and AEMDR extracted noise signals to form a motion-sensitive noise signal. The adaptive filter used the noise signal as the reference input in the fourth step. The filter removes motion artifacts with a similar frequency range as the QRS complex and produces a cleaned signal. The last step is the VMDR algorithm. The method can remove the high-frequency components left in the ECG signal to produce a smoother signal. The results clearly show improve-

ments in the quality of the test ECG signals compared to various algorithms found in the literature. In addition, the ECG recordings were collected from two different ECG devices: the Astroskin Smart vest and the VivaLnk ECG device. Therefore, the proposed noise removal method is robust and suitable for various medical-grade sensors.

The proposed abnormality detection model is an improved Convolutional Neural Network(CNN) architecture. The model could process the Short-time Fourier Transform spectrogram of the ECG signal and handcrafted features derived from the ECG signal with a feature merging layer. Several famous CNN architectures are tested during the study, and the ResNet models have the highest anomaly detection rate in the experiments. In addition, based on the tested results, the feature merging layers could improve the accuracy of the CNN models. Moreover, according to the experimental results, the proposed algorithm had superior performance to the current stat-of-arts algorithms. The best model for rhythm classification in the experiment is the proposed ResNet18-Gray. It could classify 17 different rhythm types into normal and abnormal classes with an exceptional accuracy of 99.79%, a sensitivity of 99.74%, a false alarm rate of 0.15%, and a positive predictive value of 99.85% which makes this new algorithm the most practical algorithm for rhythm classification in the literature. The heartbeat classification is used only on normal rhythms to detect abnormal heartbeats in the proposed anomaly detection algorithm. Therefore, the proposed algorithm is better at detecting state-of-art heartbeat classification algorithms. The best model for heartbeat classification in the experiment is the ResNet34-Gray. It could classify 14 different heartbeat types into normal and abnormal classes with an accuracy of 99.18%, a sensitivity of 98.80%, a false alarm rate of 0.45%, and a 99.54% positive predictive value. The result shows that the proposed algorithm has the best detection rate for abnormal heartbeat compared to other algorithms found in the literature. The proposed approach also illustrates that in some situations, the introduction of handcrafted features can significantly improve the performance of neural networks.

6.2 Future Work

In this research, the interference that affects the ECG signals measurement and anomaly detection is not discussed since the MIT-BIH and European ST databases reduced noise. However, high noise levels could affect the ECG signal measurement and anomaly detection. We built a dataset with noise-contaminated ECG signals from actual cardiac patients in the original plan. However, due to the consequence of the COVID, we could not conduct any medical research. With everything back to normal, we are currently working with Dr. Becher to collect the initially planned database. The study can be expanded to research how noise will affect anomaly detection on ECG signals and how well the accuracy is after applying our noise reduction algorithm.

References

[1] U. R. Acharya, H. Fujita, O. S. Lih, Y. Hagiwara, J. H. Tan, and M. Adam, "Automated detection of arrhythmias using different intervals of tachycardia ecg segments with convolutional neural network," *Information sciences*, vol. 405, pp. 81–90, 2017.

[2] V. X. Afonso, W. J. Tompkins, T. Q. Nguyen, and S. Luo, "Ecg beat detection using filter banks," *IEEE transactions on biomedical engineering*, vol. 46, no. 2, pp. 192–202, 1999.

[3] X. An and G. K Stylios, "Comparison of motion artefact reduction methods and the implementation of adaptive motion artefact reduction in wearable electrocardiogram monitoring," *Sensors*, vol. 20, no. 5, p. 1468, 2020.

[4] R. Antonicelli, C. Ripa, A. M. Abbatecola, C. A. Capparuccia, L. Ferrara, and L. Spazzafumo, "Validation of the 3-lead tele-ecg versus the 12-lead tele-ecg and the conventional 12-lead ecg method in older people," *Journal of telemedicine and telecare*, vol. 18, no. 2, pp. 104–108, 2012.

[5] H. Atoui, J. Fayn, and P. Rubel, "A novel neural-network model for deriving standard 12-lead ecgs from serial three-lead ecgs: Application to self-care," *IEEE transactions on information technology in biomedicine*, vol. 14, no. 3, pp. 883–890, 2010.

[6] S. Banerjee, R. Gupta, and M. Mitra, "Delineation of ecg characteristic features using multiresolution wavelet analysis method," *Measurement*, vol. 45, no. 3, pp. 474–487, 2012.

[7] D. J. Berndt and J. Clifford, "Using dynamic time warping to find patterns in time series.," in *KDD workshop*, Seattle, WA, vol. 10, 1994, pp. 359–370.

[8] L. Biel, O. Pettersson, L. Philipson, and P. Wide, "Ecg analysis: A new approach in human identification," *IEEE Transactions on Instrumentation and Measurement*, vol. 50, no. 3, pp. 808–812, 2001.

[9] M. Blanco-Velasco, B. Weng, and K. E. Barner, "Ecg signal denoising and baseline wander correction based on the empirical mode decomposition," *Computers in biology and medicine*, vol. 38, no. 1, pp. 1–13, 2008.

[10] F. Bouaziz, D. Boutana, and M. Benidir, "Multiresolution wavelet-based qrs complex detection algorithm suited to several abnormal morphologies," *IET Signal Processing*, vol. 8, no. 7, pp. 774–782, 2014.

[11] G. E. Burch, "The history of vectorcardiography," *Medical History*, vol. 29, no. S5, pp. 103–131, 1985.

[12] K.-M. Chang, "Arrhythmia ecg noise reduction by ensemble empirical mode decomposition," *Sensors*, vol. 10, no. 6, pp. 6063–6080, 2010.

[13] A. Chaturvedi, K. Raj, and A. Kumar, "A comparative analysis of lms and nlms algorithms for adaptive filtration of compressed ecg signal," in *2012 2nd International Conference on Power, Control and Embedded Systems*, IEEE, 2012, pp. 1–6.

[14] S. Chauhan and L. Vig, "Anomaly detection in ecg time signals via deep long short-term memory networks," in *Data Science and Advanced Analytics (DSAA), 2015. 36678 2015. IEEE International Conference on*, IEEE, 2015, pp. 1–7.

[15] S.-W. Chen, H.-C. Chen, and H.-L. Chan, "A real-time qrs detection method based on moving-averaging incorporating with wavelet denoising," *Computer methods and programs in biomedicine*, vol. 82, no. 3, pp. 187–195, 2006.

[16] S. Choi, M. Adnane, G.-J. Lee, H. Jang, Z. Jiang, and H.-K. Park, "Development of ecg beat segmentation method by combining lowpass filter and irregular r–r interval checkup strategy," *Expert Systems with Applications*, vol. 37, no. 7, pp. 5208–5218, 2010.

[17] I. Christov, G. Gómez-Herrero, V. Krasteva, I. Jekova, A. Gotchev, and K. Egiazarian, "Comparative study of morphological and time-frequency ecg descriptors for heartbeat classification," *Medical engineering & physics*, vol. 28, no. 9, pp. 876–887, 2006.

[18] M. C. Chuah and F. Fu, "Ecg anomaly detection via time series analysis," in *International Symposium on Parallel and Distributed Processing and Applications*, Springer, 2007, pp. 123–135.

[19] P. De Chazal, M. O'Dwyer, and R. B. Reilly, "Automatic classification of heartbeats using ecg morphology and heartbeat interval features," *IEEE transactions on biomedical engineering*, vol. 51, no. 7, pp. 1196–1206, 2004.

[20] P. De Chazal and R. B. Reilly, "A patient-adapting heartbeat classifier using ecg morphology and heartbeat interval features," *IEEE transactions on biomedical engineering*, vol. 53, no. 12, pp. 2535–2543, 2006.

[21] J. Dhiman, S. Ahmad, K. Gulia, *et al.*, "Comparison between adaptive filter algorithms (lms, nlms and rls)," *International journal of science, engineering and technology research (IJSETR)*, vol. 2, no. 5, pp. 1100–1103, 2013.

[22] K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," *IEEE transactions on signal processing*, vol. 62, no. 3, pp. 531–544, 2013.

[23] B. J. Drew, M. G. Adams, M. M. Pelter, S.-F. Wung, and M. A. Caldwell, "Comparison of standard and derived 12-lead electrocardiograms for diagnosis of coronary angioplasty-induced myocardial ischemia," *The American journal of cardiology*, vol. 79, no. 5, pp. 639–644, 1997.

[24] B. J. Drew, M. M. Pelter, D. E. Brodnick, A. V. Yadav, *et al.*, "Comparison of a new reduced lead set ecg with the standard ecg for diagnosing cardiac arrhythmias and myocardial ischemia," *Journal of electrocardiology*, vol. 35, p. 13, 2002.

[25] K. S. Durgesh and B. Lekha, "Data classification using support vector machine," *Journal of theoretical and applied information technology*, vol. 12, no. 1, pp. 1–7, 2010.

[26] D. Ge, N. Srinivasan, and S. M. Krishnan, "Cardiac arrhythmia classification using autoregressive modeling," *Biomedical engineering online*, vol. 1, no. 1, p. 5, 2002.

[27] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, e215–e220, 2000 (June 13), Circulation Electronic Pages: http://circ.ahajournals.org/content/101/23/e215.full PMID:1085218; doi: 10.1161/01.CIR.101.23.e215.

[28] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, e215–e220, 2000.

[29] A. L. Goldberger, Z. D. Goldberger, and A. Shvilkin, *Clinical Electrocardiography: A Simplified Approach E-Book: A Simplified Approach.* Elsevier Health Sciences, 2017.

[30] A. Gretton, A. Smola, J. Huang, M. Schmittfull, K. Borgwardt, and B. Schölkopf, "Covariate shift by kernel mean matching," *Dataset shift in machine learning*, vol. 3, no. 4, p. 5, 2009.

[31] J. Hampton, *The ECG Made Easy E-Book.* Elsevier Health Sciences, 2013.

[32] R. A. Harrigan, T. C. Chan, and W. J. Brady, "Electrocardiographic electrode misplacement, misconnection, and artifact," *The Journal of emergency medicine*, vol. 43, no. 6, pp. 1038–1044, 2012.

[33] S. S. Haykin, B. Widrow, and B. Widrow, *Least-mean-square adaptive filters*. Wiley Online Library, 2003, vol. 31.

[34] H. He and E. A. Garcia, "Learning from imbalanced data," *IEEE Transactions on knowledge and data engineering*, vol. 21, no. 9, pp. 1263–1284, 2009.

[35] Y. H. Hu, S. Palreddy, and W. J. Tompkins, "A patient-adaptable ecg beat classifier using a mixture of experts approach," *IEEE transactions on biomedical engineering*, vol. 44, no. 9, pp. 891–900, 1997.

[36] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of the Royal Society of London. Series A: mathematical, physical and engineering sciences*, vol. 454, no. 1971, pp. 903–995, 1998.

[37] A. Inc. (2018). "Apple watch series 4," [Online]. Available: `https://www.apple.com/ca/apple-watch-series-4/health/` (visited on 02/25/2019).

[38] C. T. Inc. (2012). "Hexoskin smart shirt," [Online]. Available: `https://www.hexoskin.com/` (visited on 02/25/2019).

[39] Q. Inc. (2018). "Qardiomd," [Online]. Available: `https://www.getqardio.com/qardiomd-ecg/` (visited on 02/25/2019).

[40] V. Inc. (2014). "Vivalnk." url: https://www.vivalnk.com/products/medical-wearable-sensors/continuous-ecg-monitor.

[41] T. Ince, S. Kiranyaz, and M. Gabbouj, "A generic and robust system for automated patient-specific classification of ecg signals," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 5, pp. 1415–1426, 2009.

[42] I. Jekova, G. Bortolan, and I. Christov, "Assessment and comparison of different methods for heartbeat classification," *Medical Engineering & Physics*, vol. 30, no. 2, pp. 248–257, 2008.

[43] J. Jenitta and A. Rajeswari, "Denoising of ecg signal based on improved adaptive filter with emd and eemd," in *2013 IEEE Conference on Information & Communication Technologies*, IEEE, 2013, pp. 957–962.

[44] W. Jiang and S. G. Kong, "Block-based neural networks for personalized ecg signal classification," *IEEE Transactions on Neural Networks*, vol. 18, no. 6, pp. 1750–1761, 2007.

[45] A. Karimipour and M. R. Homaeinezhad, "Real-time electrocardiogram p-qrs-t detection–delineation algorithm based on quality-supported analysis of characteristic templates," *Computers in biology and medicine*, vol. 52, pp. 153–165, 2014.

[46] H.-Y. Kim, "Statistical notes for clinical researchers: Assessing normal distribution (2) using skewness and kurtosis," *Restorative dentistry & endodontics*, vol. 38, no. 1, pp. 52–54, 2013.

[47] N. Kingsbury, "Complex wavelets for shift invariant analysis and filtering of signals," *Applied and computational harmonic analysis*, vol. 10, no. 3, pp. 234–253, 2001.

[48] S. Kiranyaz, T. Ince, and M. Gabbouj, "Real-time patient-specific ecg classification by 1-d convolutional neural networks," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 3, pp. 664–675, 2015.

[49] S. Kiranyaz, T. Ince, R. Hamila, and M. Gabbouj, "Convolutional neural networks for patient-specific ecg classification," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, IEEE, 2015, pp. 2608–2611.

[50] A. N. Kristensen, B. Jeyam, S. Riahi, and M. B. Jensen, "The use of a portable three-lead ecg monitor to detect atrial fibrillation in general practice," *Scandinavian journal of primary health care*, vol. 34, no. 3, pp. 304–308, 2016.

[51] H. Li and P. Boulanger, "A survey of heart anomaly detection using ambulatory electrocardiogram (ecg)," *Sensors*, vol. 20, no. 5, p. 1461, 2020.

[52] K. Li, N. Du, and A. Zhang, "Detecting ecg abnormalities via transductive transfer learning," in *Proceedings of the ACM Conference on Bioinformatics, Computational Biology and Biomedicine*, ACM, 2012, pp. 210–217.

[53] P. Li, K. L. Chan, S. Fu, and S. M. Krishnan, "An abnormal ecg beat detection approach for long-term monitoring of heart patients based on hybrid kernel machine ensemble," in *International Workshop on Multiple Classifier Systems*, Springer, 2005, pp. 346–355.

[54] H.-Y. Lin, S.-Y. Liang, Y.-L. Ho, Y.-H. Lin, and H.-P. Ma, "Discrete-wavelet-transform-based noise removal and feature extraction for ecg signals," *Irbm*, vol. 35, no. 6, pp. 351–361, 2014.

[55] X. Lisheng, W. Kuanquan, D. Zhang, and S. Cheng, "Adaptive baseline wander removal in the pulse waveform," in *Proceedings of 15th IEEE Symposium on Computer-Based Medical Systems (CBMS 2002)*, IEEE, 2002, pp. 143–148.

[56] M. Llamedo and J. P. Martınez, "Heartbeat classification using feature selection driven by database generalization criteria," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 3, pp. 616–625, 2010.

[57] K. Luo, J. Li, Z. Wang, and A. Cuschieri, "Patient-specific deep architectural model for ecg classification," *Journal of healthcare engineering*, vol. 2017, 2017.

[58]   U. Maji and S. Pal, "Empirical mode decomposition vs. variational mode decomposition on ecg signal processing: A comparative study," in *2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, IEEE, 2016, pp. 1129–1134.

[59]   S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," in *Fundamental Papers in Wavelet Theory*, Princeton University Press, 2009, pp. 494–513.

[60]   M. S. Manikandan and K. Soman, "A novel method for detecting r-peaks in electrocardiogram (ecg) signal," *Biomedical Signal Processing and Control*, vol. 7, no. 2, pp. 118–128, 2012.

[61]   MATLAB, *version 9.11.0 (R2021b)*. Natick, Massachusetts: The Math-Works Inc., 2021.

[62]   S. K. Mitra and Y. Kuo, *Digital signal processing: a computer-based approach*. McGraw-Hill New York, 2006, vol. 2.

[63]   G. B. Moody and R. G. Mark, "The impact of the mit-bih arrhythmia database," *IEEE Engineering in Medicine and Biology Magazine*, vol. 20, no. 3, pp. 45–50, 2001.

[64]   S. P. Nelwan, J. A. Kors, S. H. Meij, J. H. van Bemmel, and M. L. Simoons, "Reconstruction of the 12-lead electrocardiogram from reduced lead sets," *Journal of electrocardiology*, vol. 37, no. 1, pp. 11–18, 2004.

[65]   W. S. Noble, "What is a support vector machine?" *Nature biotechnology*, vol. 24, no. 12, pp. 1565–1567, 2006.

[66]   W. H. Oraganization. (2018). "Cardiovascular diseases (cvds)," [Online]. Available: `https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)` (visited on 02/25/2019).

[67]   Y. Özbay, R. Ceylan, and B. Karlik, "Integration of type-2 fuzzy clustering and wavelet transform in a neural network based ecg classifier," *Expert Systems with Applications*, vol. 38, no. 1, pp. 1004–1010, 2011.

[68]   S. Pal and M. Mitra, "Detection of ecg characteristic points using multiresolution wavelet analysis based selective coefficient method," *Measurement*, vol. 43, no. 2, pp. 255–261, 2010.

[69]   J. Pan and W. J. Tompkins, "A real-time qrs detection algorithm," *IEEE Trans. Biomed. Eng*, vol. 32, no. 3, pp. 230–236, 1985.

[70]   J.-S. Park, S.-W. Lee, and U. Park, "R peak detection method using wavelet transform and modified shannon energy envelope," *Journal of healthcare engineering*, vol. 2017, 2017.

[71]   A. M. Patel, P. K. Gakare, and A. Cheeran, "Real time ecg feature extraction and arrhythmia detection on a mobile platform," *Int. J. Comput. Appl*, vol. 44, no. 23, pp. 40–45, 2012.

[72] S. Patro and K. K. Sahu, "Normalization: A preprocessing stage," *arXiv preprint arXiv:1503.06462*, 2015.

[73] N. D. Pham, Q. L. Le, and T. K. Dang, "Hot asax: A novel adaptive symbolic representation for time series discords discovery," in *Asian Conference on Intelligent Information and Database Systems*, Springer, 2010, pp. 113–121.

[74] S. Poungponsri and X.-H. Yu, "An adaptive filtering approach for electrocardiogram (ecg) signal noise reduction using neural networks," *Neurocomputing*, vol. 117, pp. 206–213, 2013.

[75] L. Prechelt, "Early stopping-but when?" In *Neural Networks: Tricks of the trade*, Springer, 1998, pp. 55–69.

[76] H. M. Rai and A. Trivedi, "De-noising of ecg waveforms based on multiresolution wavelet transform," *International Journal of Computer Applications*, vol. 45, no. 18, pp. 25–30, 2012.

[77] H. M. Rai, A. Trivedi, and S. Shukla, "Ecg signal processing for abnormalities detection using multi-resolution wavelet transform and artificial neural network classifier," *Measurement*, vol. 46, no. 9, pp. 3238–3246, 2013.

[78] K. N. Rajesh and R. Dhuli, "Classification of ecg heartbeats using nonlinear decomposition methods and support vector machine," *Computers in biology and medicine*, vol. 87, pp. 271–284, 2017.

[79] P. Rajpurkar, A. Y. Hannun, M. Haghpanahi, C. Bourn, and A. Y. Ng, "Cardiologist-level arrhythmia detection with convolutional neural networks," *arXiv preprint arXiv:1707.01836*, 2017.

[80] M. Rakshit, D. Panigrahy, and P. Sahu, "An improved method for r-peak detection by using shannon energy envelope," *Sādhanā*, vol. 41, no. 5, pp. 469–477, 2016.

[81] M. A. D. Raya and L. G. Sison, "Adaptive noise cancelling of motion artifact in stress ecg signals using accelerometer," in *Proceedings of the Second Joint 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society][Engineering in Medicine and Biology*, IEEE, vol. 2, 2002, pp. 1756–1757.

[82] S. Sahoo, B. Kanungo, S. Behera, and S. Sabut, "Multiresolution wavelet transform based feature extraction and ecg classification to detect cardiac abnormalities," *Measurement*, vol. 108, pp. 55–66, 2017.

[83] E. Sejdić, I. Djurović, and J. Jiang, "Time–frequency feature representation using energy concentration: An overview of recent advances," *Digital signal processing*, vol. 19, no. 1, pp. 153–183, 2009.

[84] S. Y. ŞEN and N. ÖZKURT, "Ecg arrhythmia classification by using convolutional neural network and spectrogram," in *2019 Innovations in Intelligent Systems and Applications Conference (ASYU)*, IEEE, 2019, pp. 1–6.

[85] B. N. Singh and A. K. Tiwari, "Optimal selection of wavelet basis function applied to ecg signal denoising," *Digital signal processing*, vol. 16, no. 3, pp. 275–287, 2006.

[86] G. Singh, G. Kaur, and V. Kumar, "Ecg denoising using adaptive selection of imfs through emd and eemd," in *2014 International Conference on Data Science & Engineering (ICDSE)*, IEEE, 2014, pp. 228–231.

[87] H. Sivaraks and C. A. Ratanamahatana, "Robust and accurate anomaly detection in ecg artifacts using time series motif discovery," *Computational and mathematical methods in medicine*, vol. 2015, 2015.

[88] A. Taddei, G. Distante, M. Emdin, P. Pisani, G. Moody, C. Zeelenberg, and C. Marchesi, "The european st-t database: Standard for evaluating systems for the analysis of st-t changes in ambulatory electrocardiography," *European heart journal*, vol. 13, no. 9, pp. 1164–1172, 1992.

[89] N. V. Thakor and Y.-S. Zhu, "Applications of adaptive filtering to ecg analysis: Noise cancellation and arrhythmia detection," *IEEE transactions on biomedical engineering*, vol. 38, no. 8, pp. 785–794, 1991.

[90] R. K. Thenua and S. Agarwal, "Simulation and performance analysis of adaptive filter in noise cancellation," *International Journal of Engineering Science and Technology*, vol. 2, no. 9, pp. 4373–4378, 2010.

[91] M. Thomas, M. K. Das, and S. Ari, "Automatic ecg arrhythmia classification using dual tree complex wavelet based features," *AEU-International Journal of Electronics and Communications*, vol. 69, no. 4, pp. 715–721, 2015.

[92] M. Training and S. LLC. (2019). "Ekg reference," [Online]. Available: https://www.practicalclinicalskills.com/ekg-reference (visited on 07/18/2019).

[93] E. D. Übeyli, "Combining recurrent neural networks with eigenvector methods for classification of ecg beats," *Digital Signal Processing*, vol. 19, no. 2, pp. 320–329, 2009.

[94] B. Veeravalli, C. J. Deepu, and D. Ngo, "Real-time, personalized anomaly detection in streaming data for wearable healthcare devices," in *Handbook of Large-Scale Distributed Computing in Smart Healthcare*, Springer, 2017, pp. 403–426.

[95] A. Velayudhan and S. Peter, "Noise analysis and different denoising techniques of ecg signal-a survey," *IOSR Journal of Electronics and Communication Engineering (IOSR-JECE)*, eISSN–2278, 2016.

[96] B. Weng, M. Blanco-Velasco, and K. E. Barner, "Ecg denoising based on the empirical mode decomposition," in *2006 International Conference of the IEEE Engineering in Medicine and Biology Society*, IEEE, 2006, pp. 1–4.

[97] M. Wess, P. S. Manoj, and A. Jantsch, "Neural network based ecg anomaly detection on fpga and trade-off analysis," in *2017 IEEE International Symposium on Circuits and Systems (ISCAS)*, IEEE, 2017, pp. 1–4.

[98] A. Widodo and B.-S. Yang, "Support vector machine in machine condition monitoring and fault diagnosis," *Mechanical systems and signal processing*, vol. 21, no. 6, pp. 2560–2574, 2007.

[99] B. Widrow, J. McCool, and M. Ball, "The complex lms algorithm," *Proceedings of the IEEE*, vol. 63, no. 4, pp. 719–720, 1975.

[100] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: A noise-assisted data analysis method," *Advances in adaptive data analysis*, vol. 1, no. 01, pp. 1–41, 2009.

[101] F. Xiong, D. Chen, and M. Huang, "A wavelet adaptive cancellation algorithm based on multi-inertial sensors for the reduction of motion artifacts in ambulatory ecgs," *Sensors*, vol. 20, no. 4, p. 970, 2020.

[102] S. Yang and H. Shen, "Heartbeat classification using discrete wavelet transform and kernel principal component analysis," in *IEEE 2013 Tencon-Spring*, IEEE, 2013, pp. 34–38.

[103] C. Ye, B. V. Kumar, and M. T. Coimbra, "Heartbeat classification using morphological and dynamic features of ecg signals," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 10, pp. 2930–2941, 2012.

[104] Y.-C. Yeh and W.-J. Wang, "Qrs complexes detection for ecg signal: The difference operation method," *Computer methods and programs in biomedicine*, vol. 91, no. 3, pp. 245–254, 2008.

[105] K. You, M. Long, J. Wang, and M. I. Jordan, "How does learning rate decay help modern neural networks?" *arXiv preprint arXiv:1908.01878*, 2019.

[106] Z. Zhang, J. Dong, X. Luo, K.-S. Choi, and X. Wu, "Heartbeat classification using disease-specific feature selection," *Computers in biology and medicine*, vol. 46, pp. 79–89, 2014.

[107] Z. Zhidong, L. Yi, and L. Qing, "Adaptive noise removal of ecg signal based on ensemble empirical mode decomposition," in *Adaptive filtering applications*, IntechOpen, 2011.

[108] H. Zhu and J. Dong, "An r-peak detection method based on peaks of shannon energy envelope," *Biomedical Signal Processing and Control*, vol. 8, no. 5, pp. 466–474, 2013.

[109] Z. Zidelmal, A. Amirou, M. Adnane, and A. Belouchrani, "Qrs detection based on wavelet coefficients," *Computer methods and programs in biomedicine*, vol. 107, no. 3, pp. 490–496, 2012.

[110] M. Zihlmann, D. Perekrestenko, and M. Tschannen, "Convolutional recurrent neural networks for electrocardiogram classification," in *2017 Computing in Cardiology (CinC)*, IEEE, 2017, pp. 1–4.

[111] D. P. Zipes, P. Libby, R. O. Bonow, D. L. Mann, and G. F. Tomaselli, *Braunwald's Heart Disease E-Book: A Textbook of Cardiovascular Medicine*. Elsevier Health Sciences, 2018.

Appendix A

Conventional methods noise reduction result
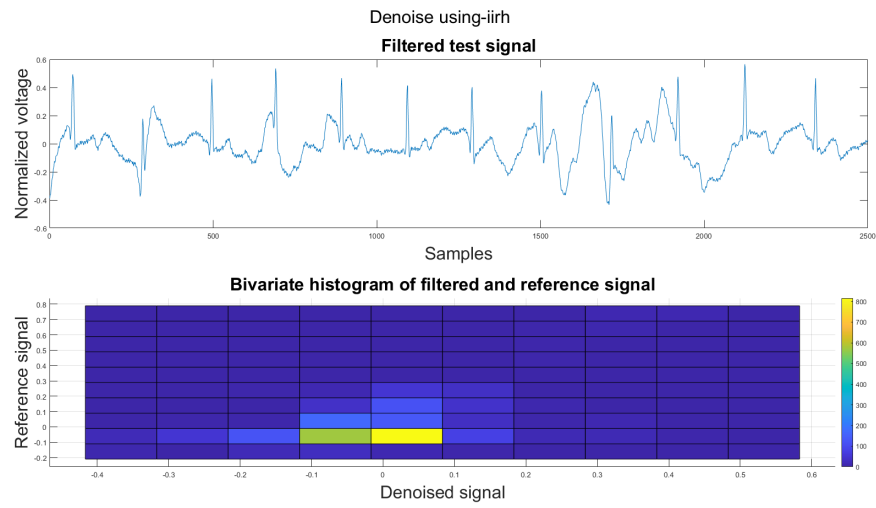
Figure A.1: (1) Test signal (2) bivariate histogram between test and reference signal
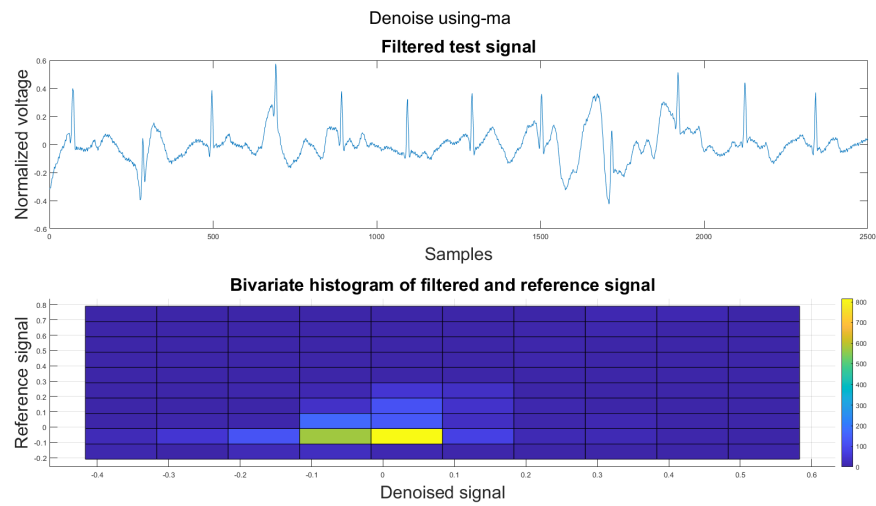
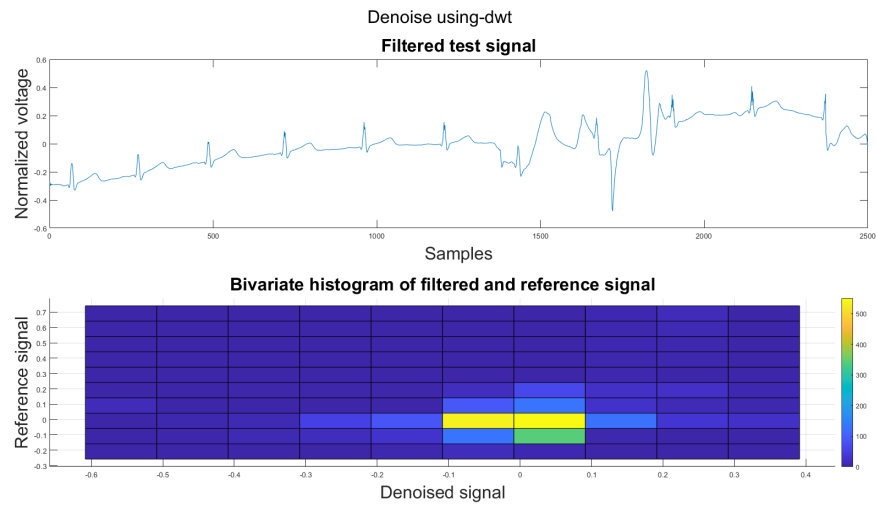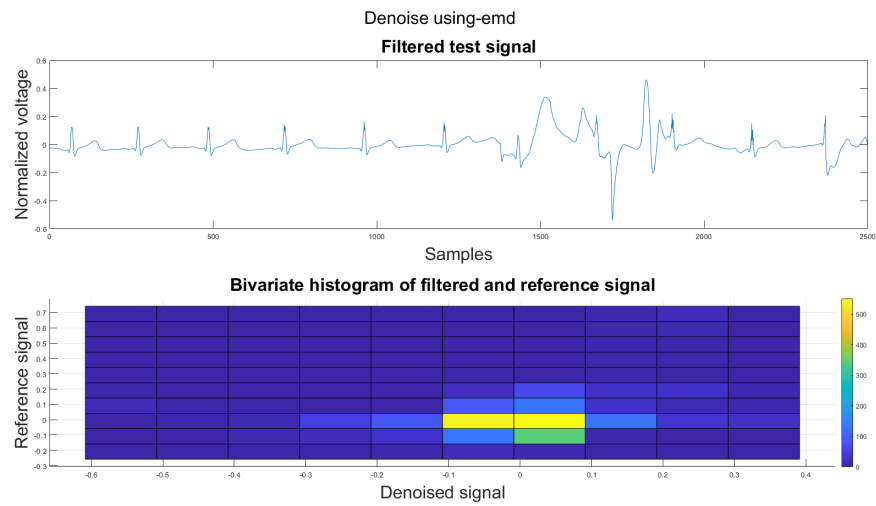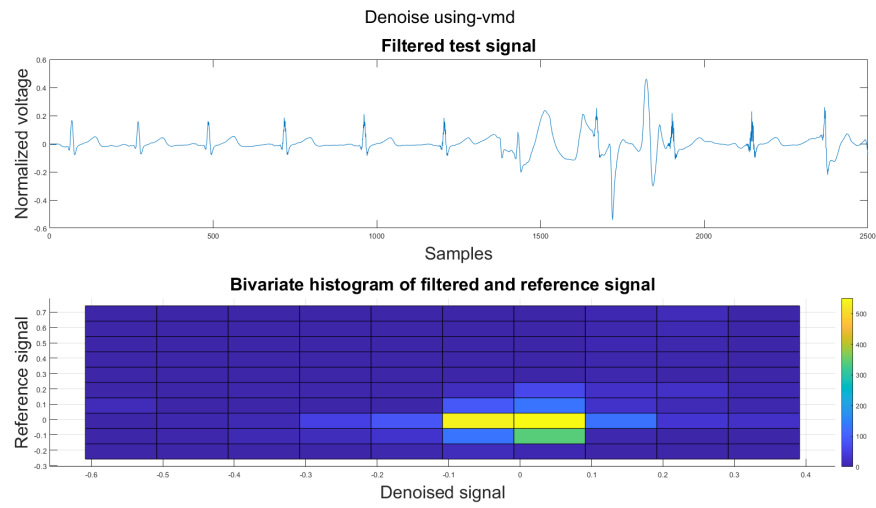Figure A.2: (1) Test signal (2) bivariate histogram between test and reference signal
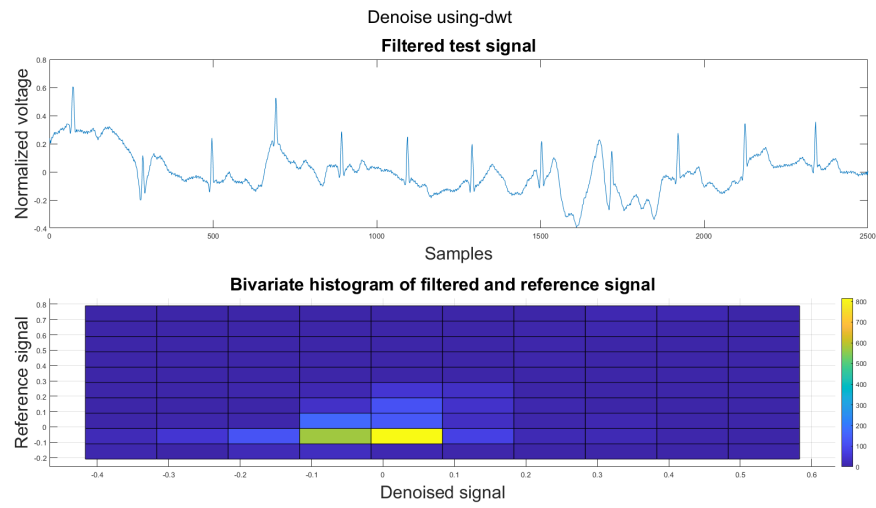
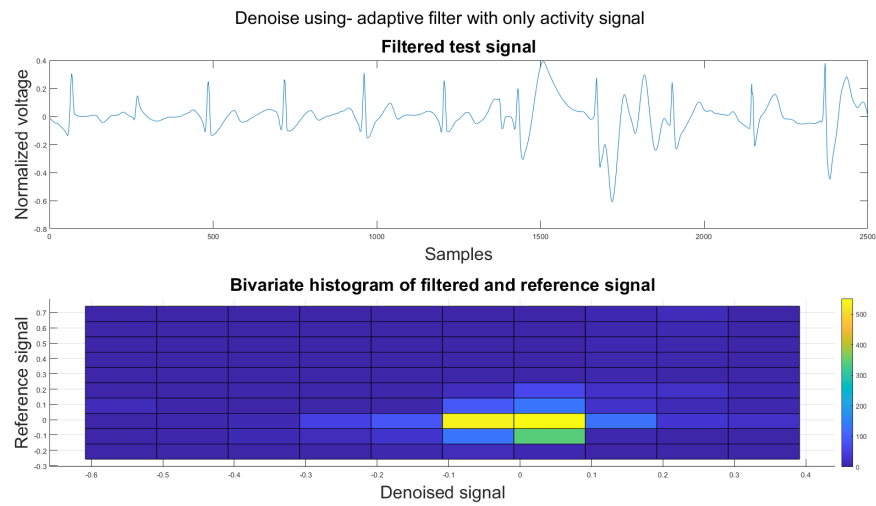Figure A.3: (1) Test signal (2) bivariate histogram between test and reference signal
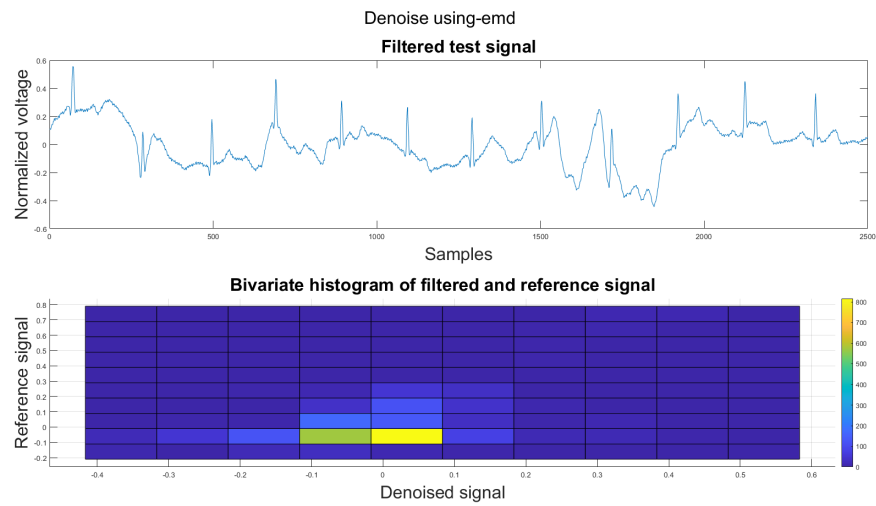
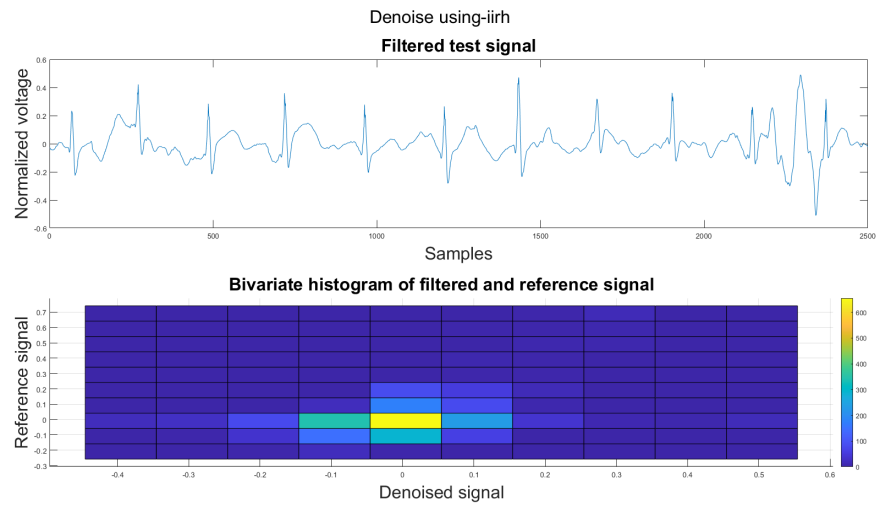Figure A.4: (1) Test signal (2) bivariate histogram between test and reference signal
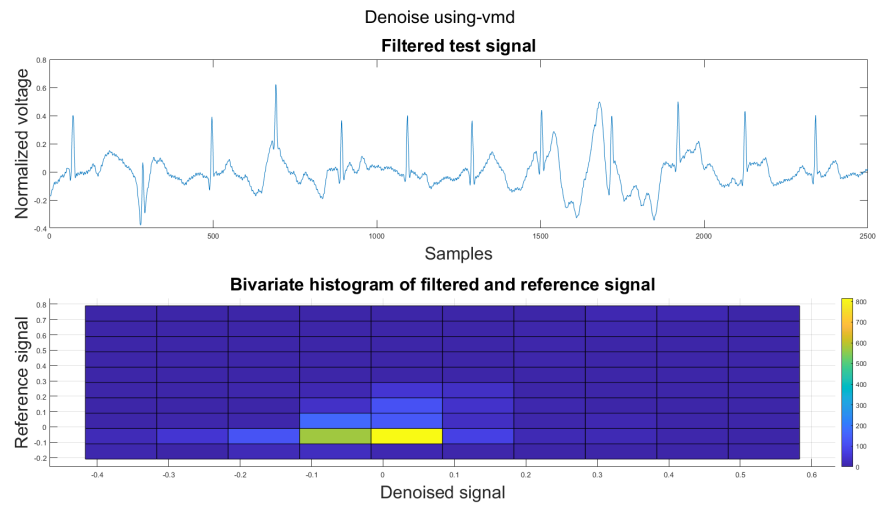
Figure A.5: (1) Test signal (2) bivariate histogram between test and reference signal
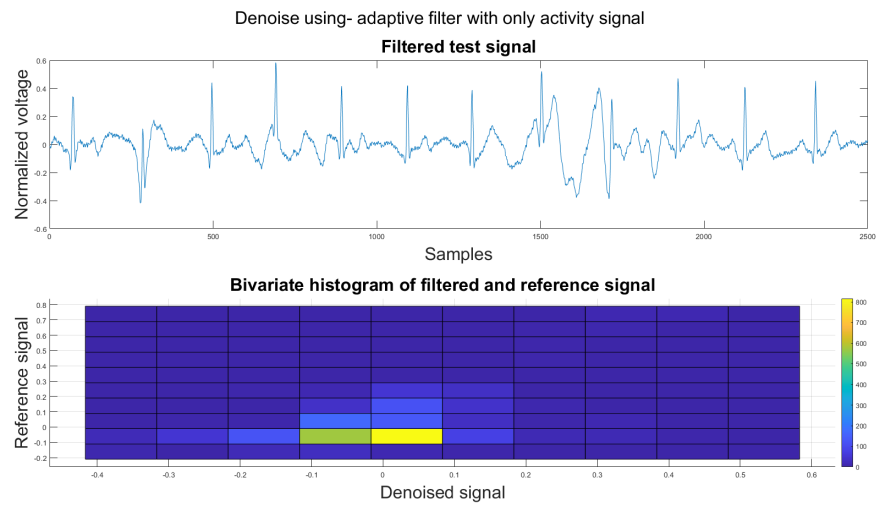
Figure A.6: (1) Test signal (2) bivariate histogram between test and reference signal
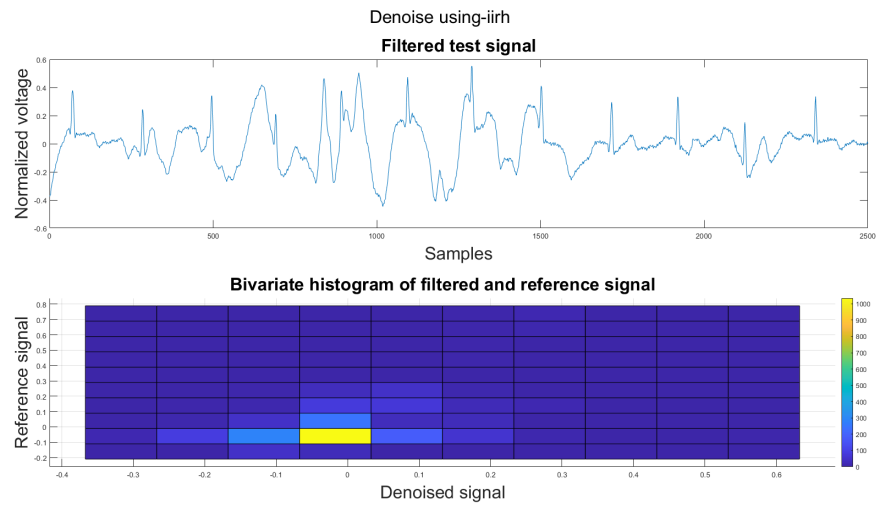
Figure A.7: (1) Test signal (2) bivariate histogram between test and reference signal
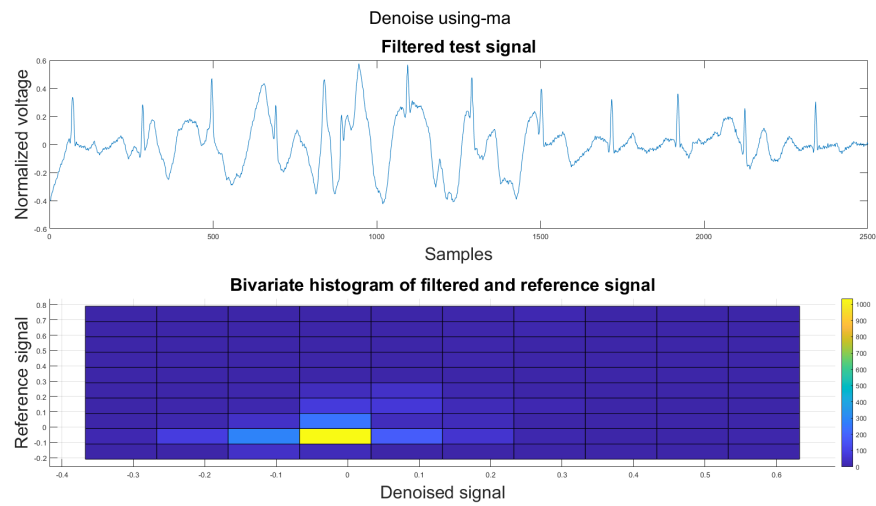
Figure A.8: (1) Test signal (2) bivariate histogram between test and reference signal
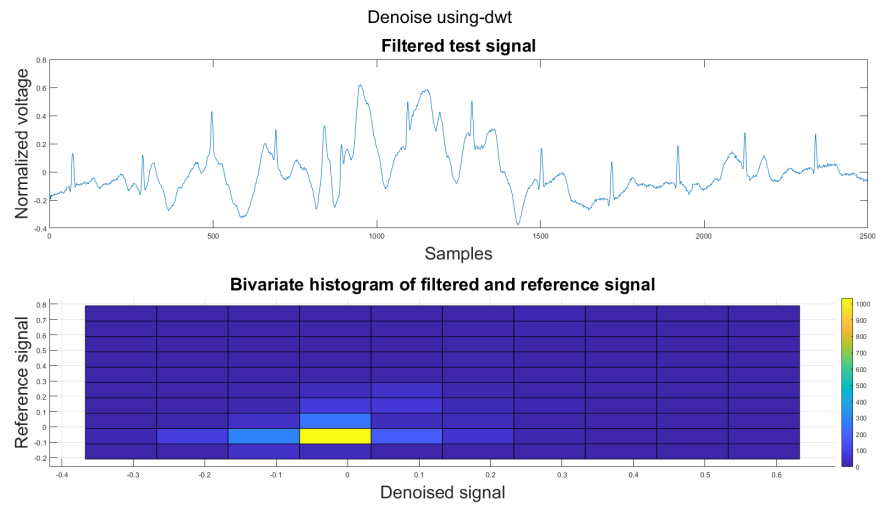
Figure A.9: (1) Test signal (2) bivariate histogram between test and reference signal
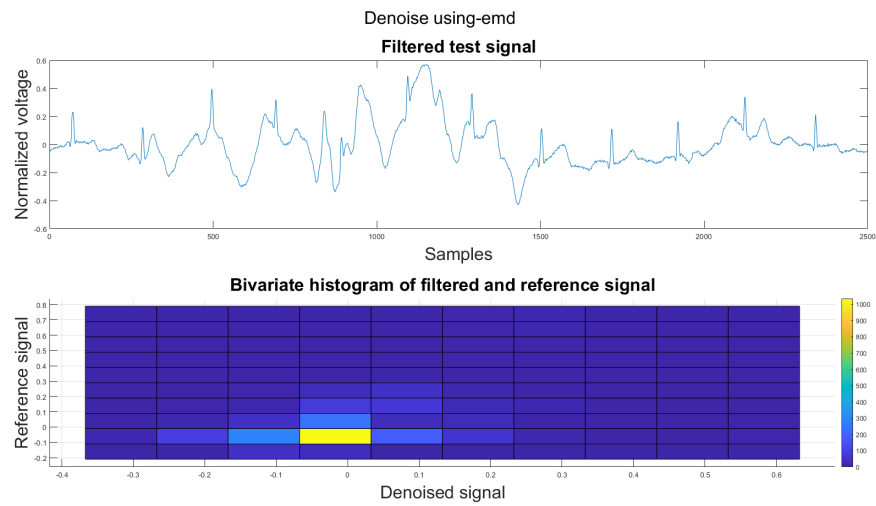
Figure A.10: (1) Test signal (2) bivariate histogram between test and reference signal
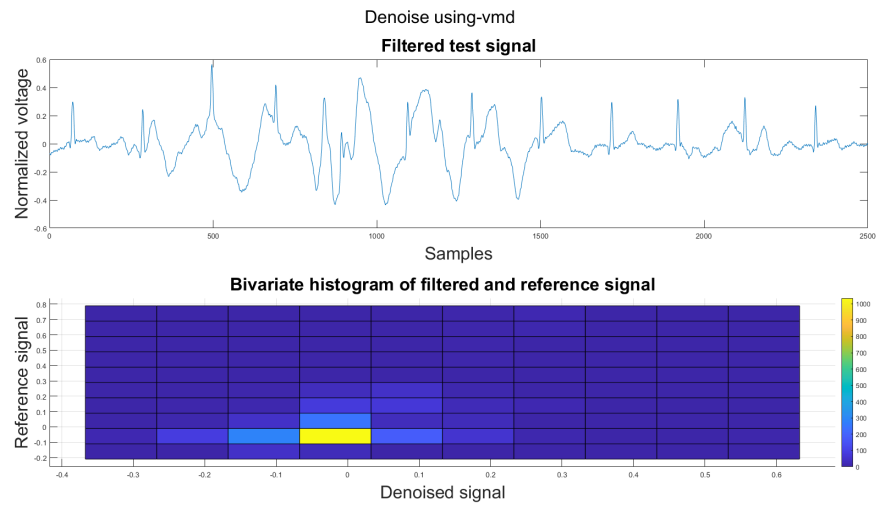
Figure A.11: (1) Test signal (2) bivariate histogram between test and reference signal
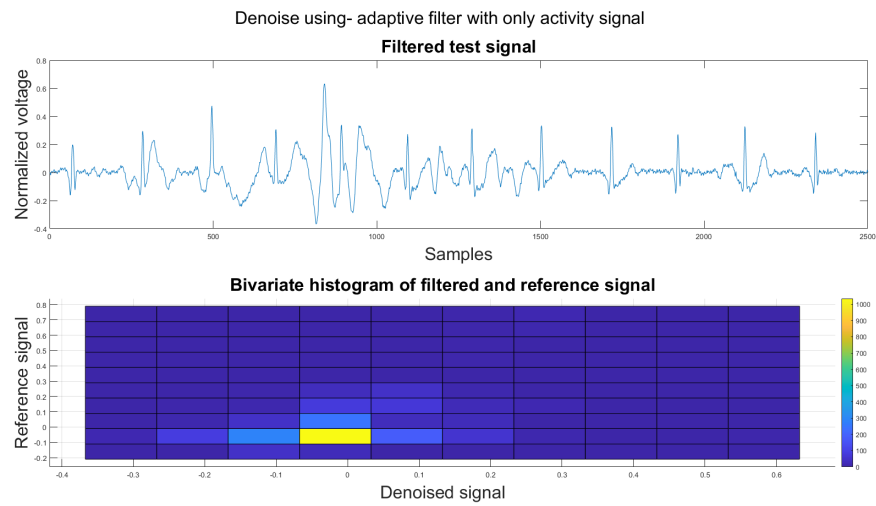
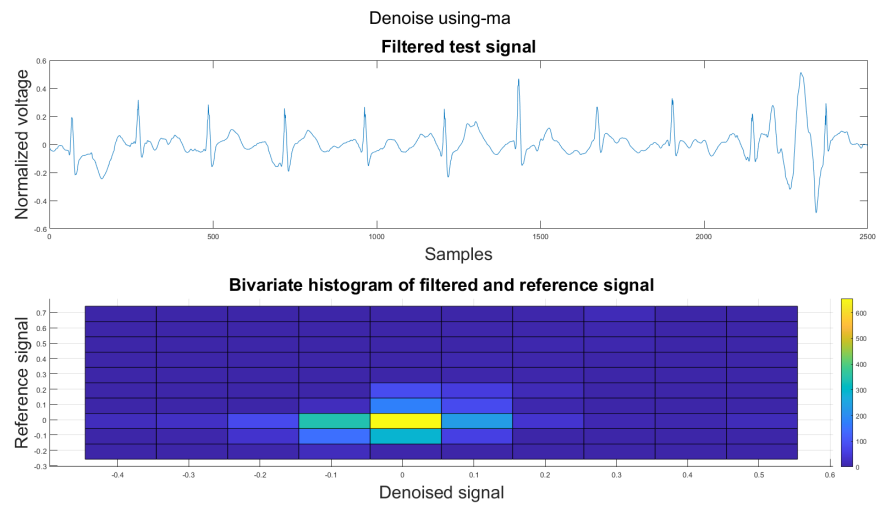Figure A.12: (1) Test signal (2) bivariate histogram between test and reference signal
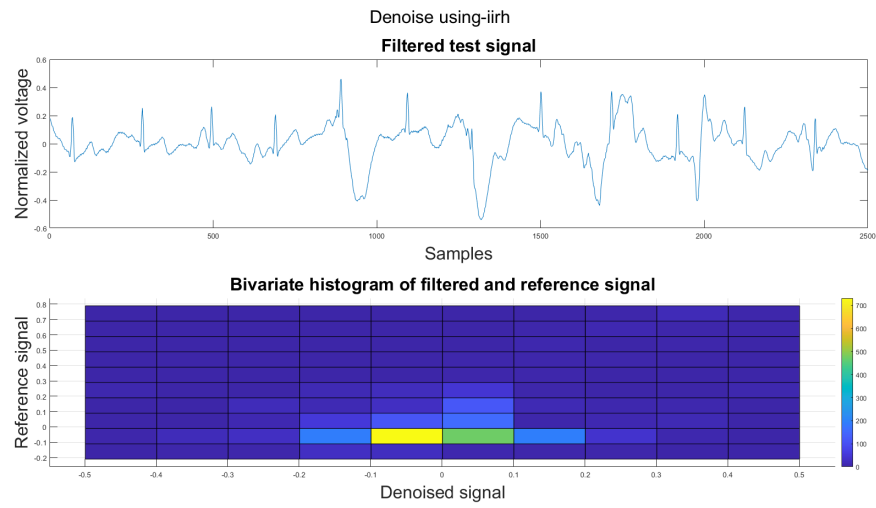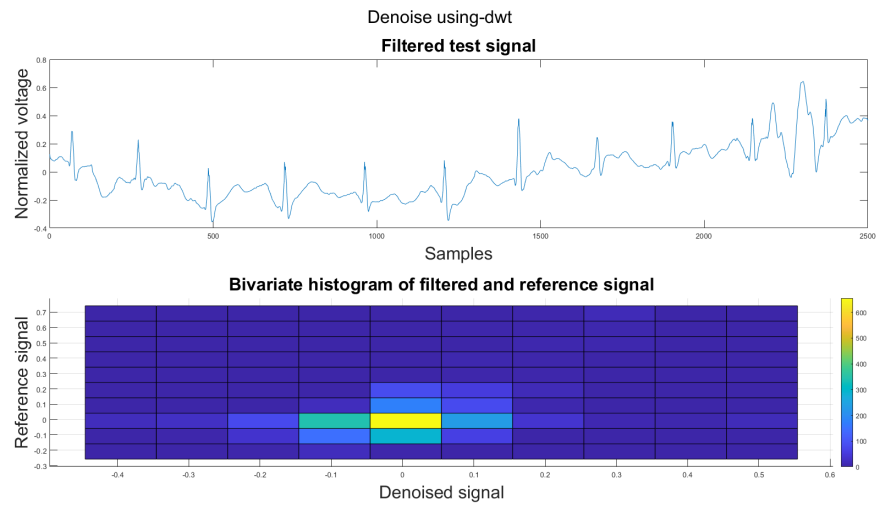
Figure A.13: (1) Test signal (2) bivariate histogram between test and reference signal
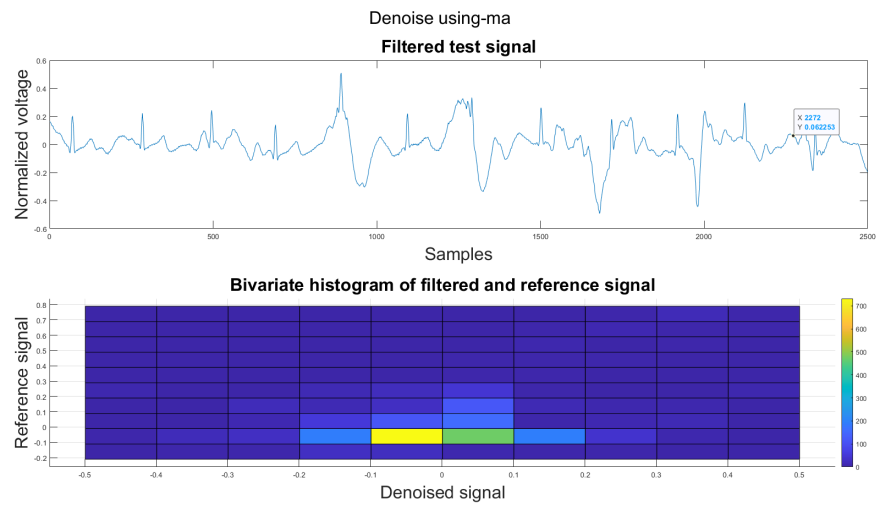
Figure A.14: (1) Test signal (2) bivariate histogram between test and reference signal
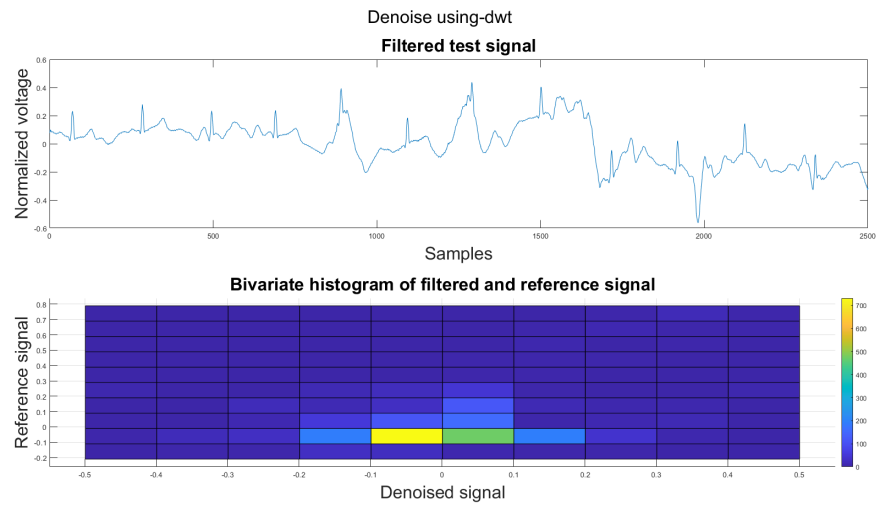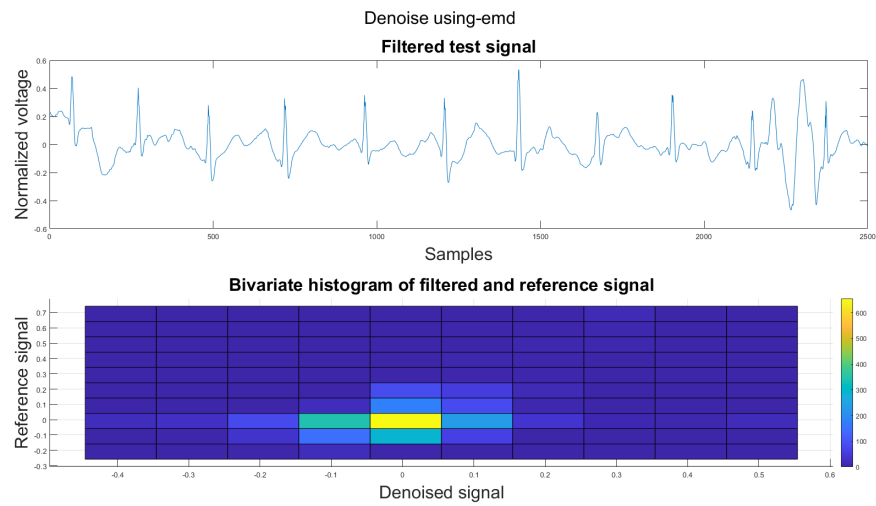
Figure A.15: (1) Test signal (2) bivariate histogram between test and reference signal
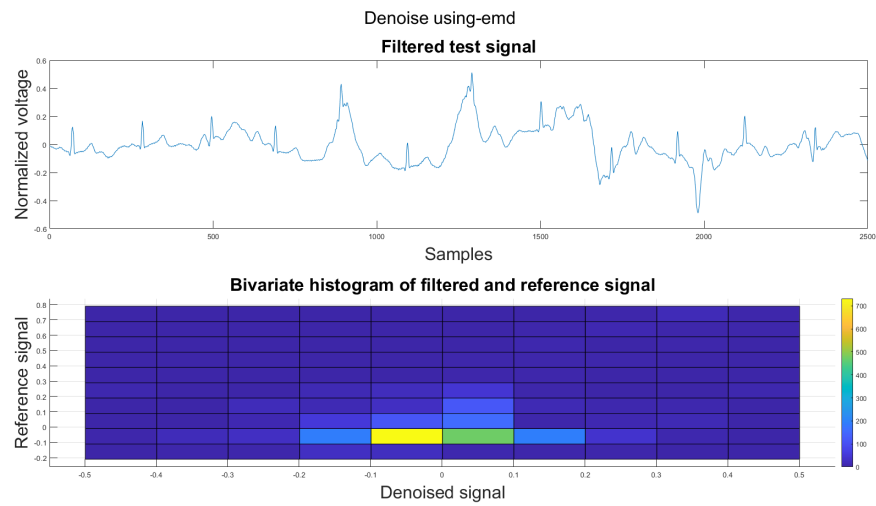
Figure A.16: (1) Test signal (2) bivariate histogram between test and reference signal
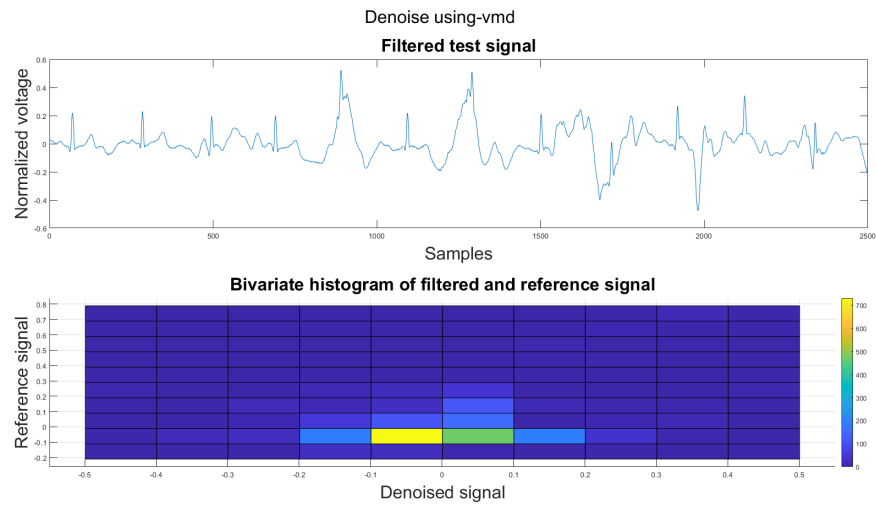
Figure A.17: (1) Test signal (2) bivariate histogram between test and reference signal
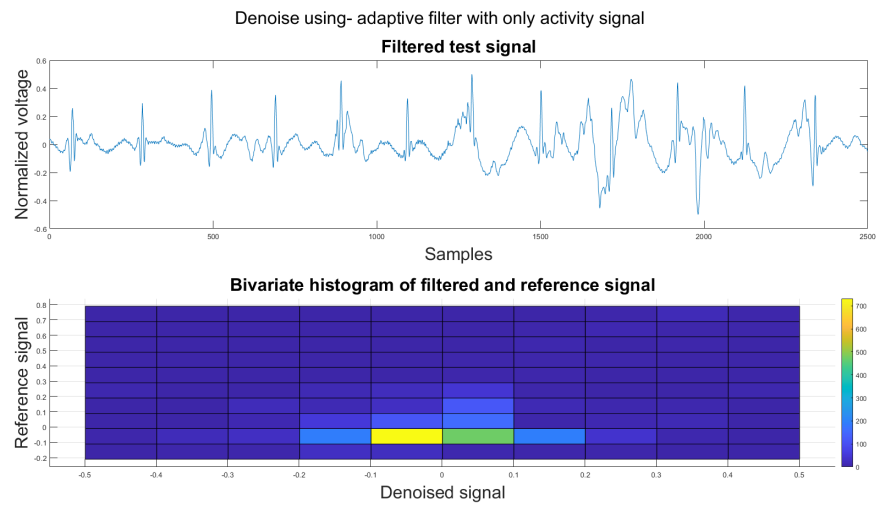
Figure A.18: (1) Test signal (2) bivariate histogram between test and reference signal
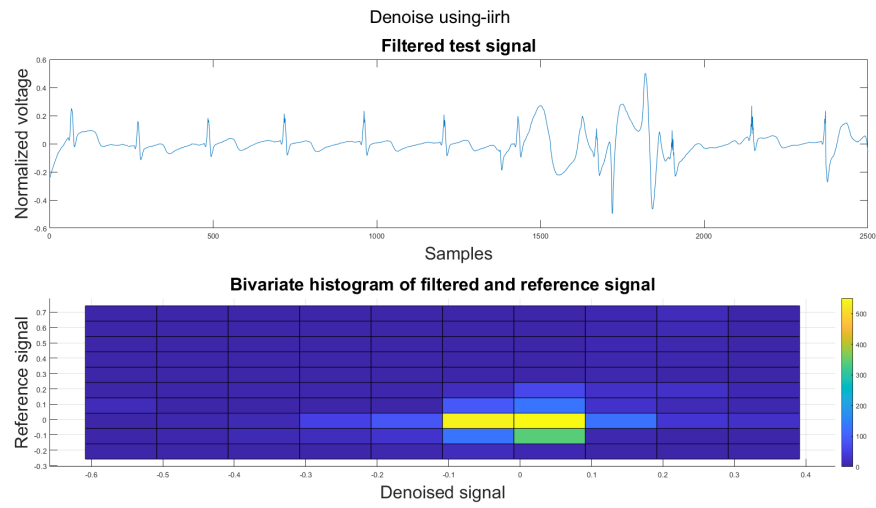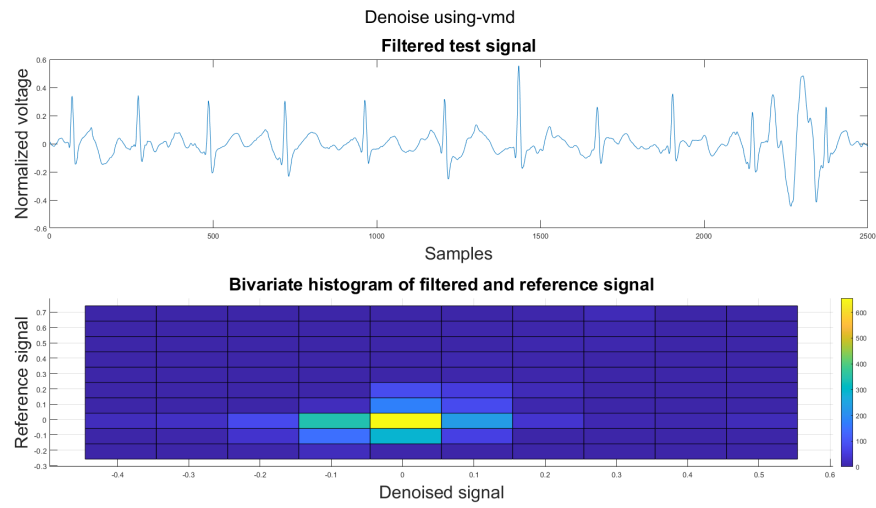
Figure A.19: (1) Test signal (2) bivariate histogram between test and reference signal
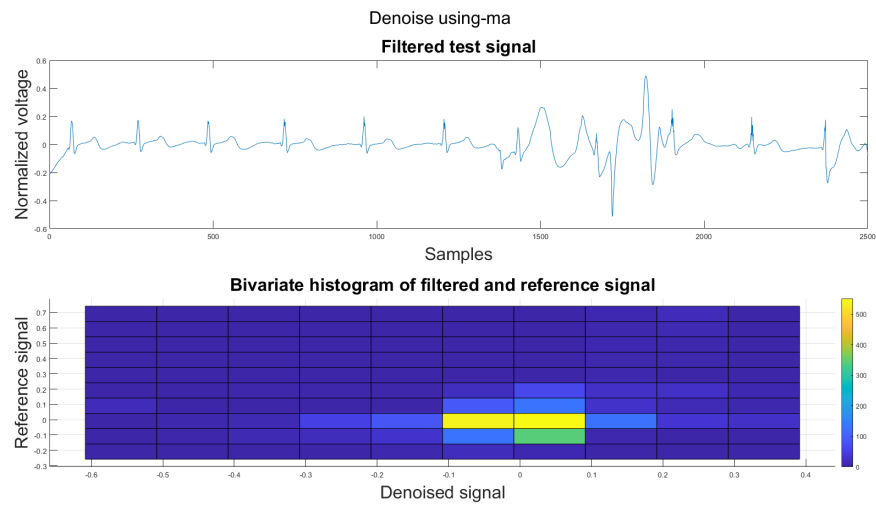
Figure A.20: (1) Test signal (2) bivariate histogram between test and reference signal

Figure A.21: (1) Test signal (2) bivariate histogram between test and reference signal
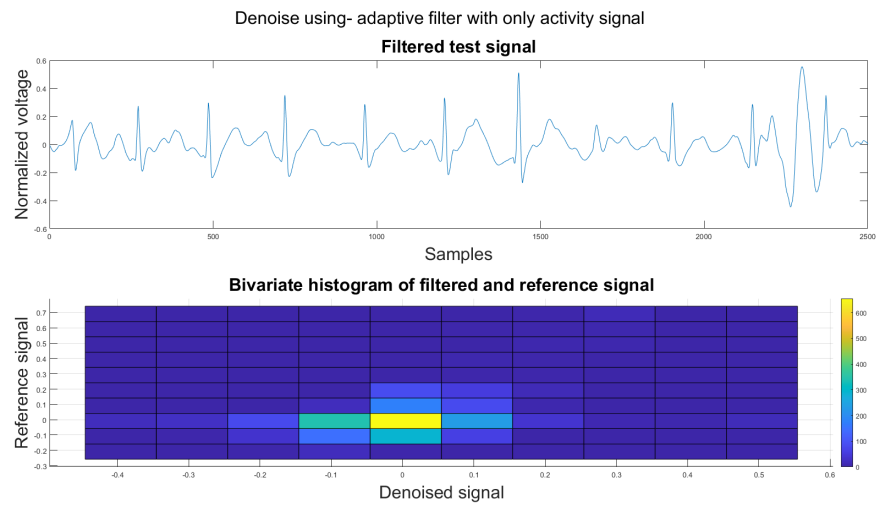

Figure A.22: (1) Test signal (2) bivariate histogram between test and reference signal

Figure A.23: (1) Test signal (2) bivariate histogram between test and reference signal
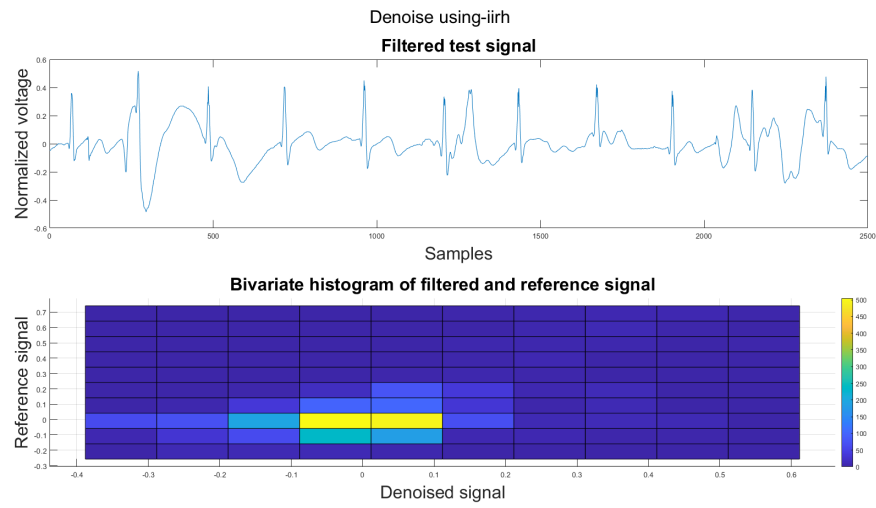

Figure A.24: (1) Test signal (2) bivariate histogram between test and reference signal

Figure A.25: (1) Test signal (2) bivariate histogram between test and reference signal

Figure A.26: (1) Test signal (2) bivariate histogram between test and reference signal
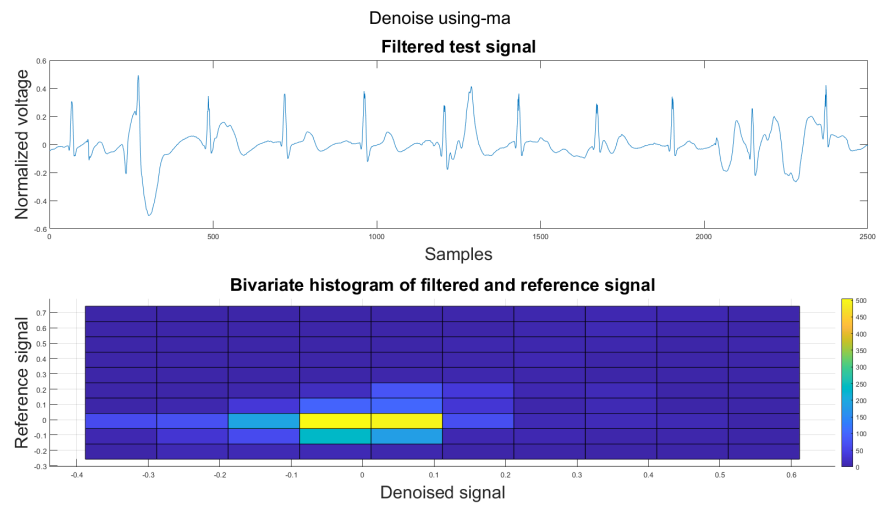
Figure A.27: (1) Test signal (2) bivariate histogram between test and reference signal
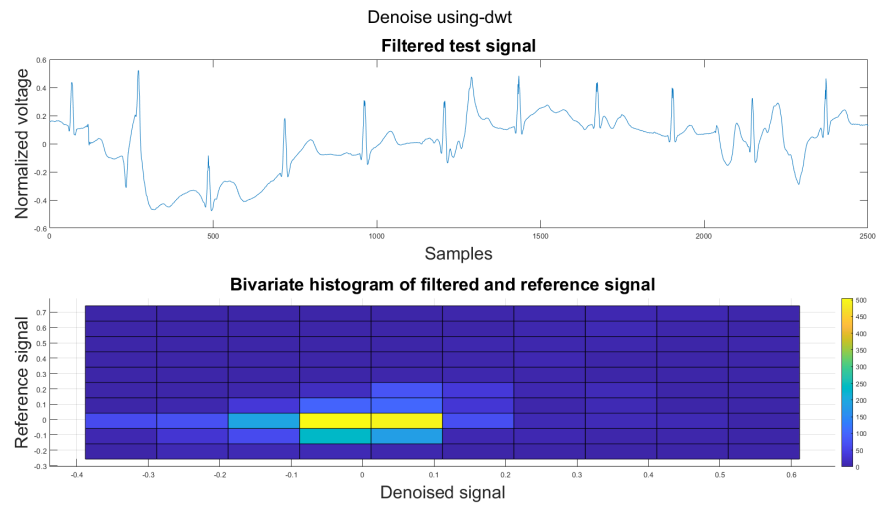

Figure A.28: (1) Test signal (2) bivariate histogram between test and reference signal

Figure A.29: (1) Test signal (2) bivariate histogram between test and reference signal

Figure A.30: (1) Test signal (2) bivariate histogram between test and reference signal
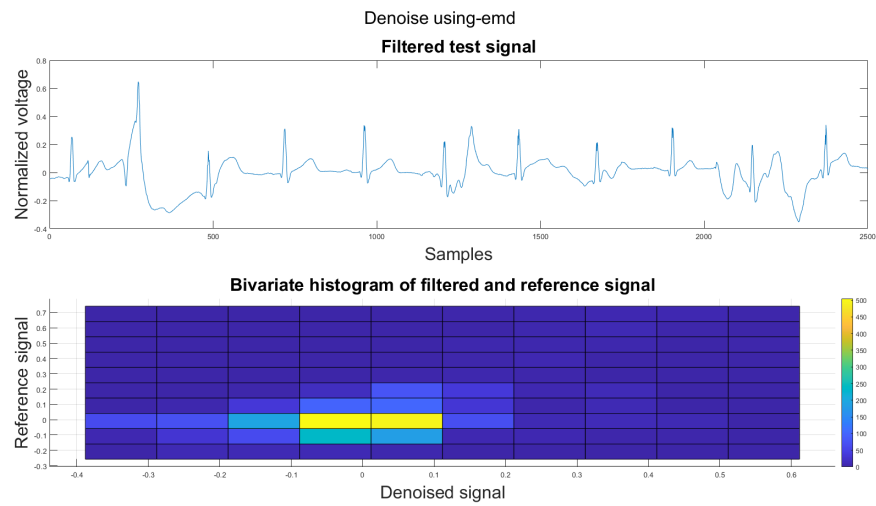
Figure A.31: (1) Test signal (2) bivariate histogram between test and reference signal

Figure A.32: (1) Test signal (2) bivariate histogram between test and reference signal

Figure A.33: (1) Test signal (2) bivariate histogram between test and reference signal

Figure A.34: (1) Test signal (2) bivariate histogram between test and reference signal
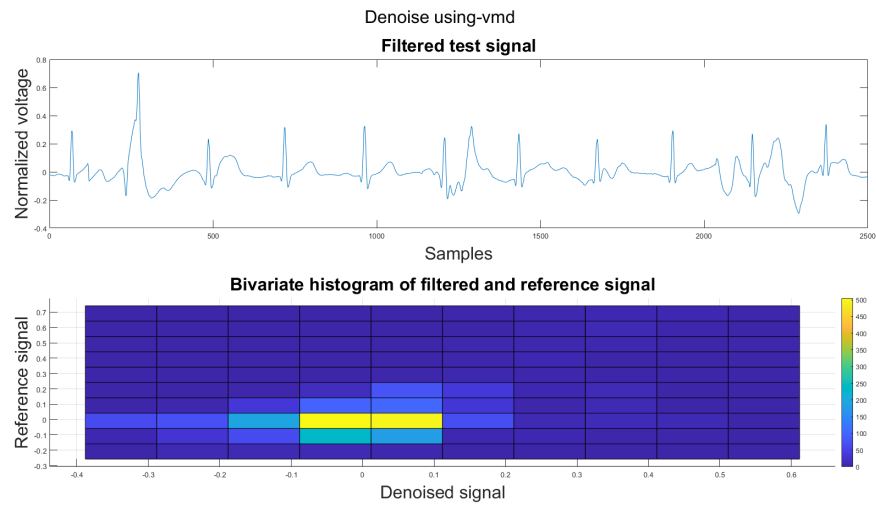
Figure A.35: (1) Test signal (2) bivariate histogram between test and reference signal
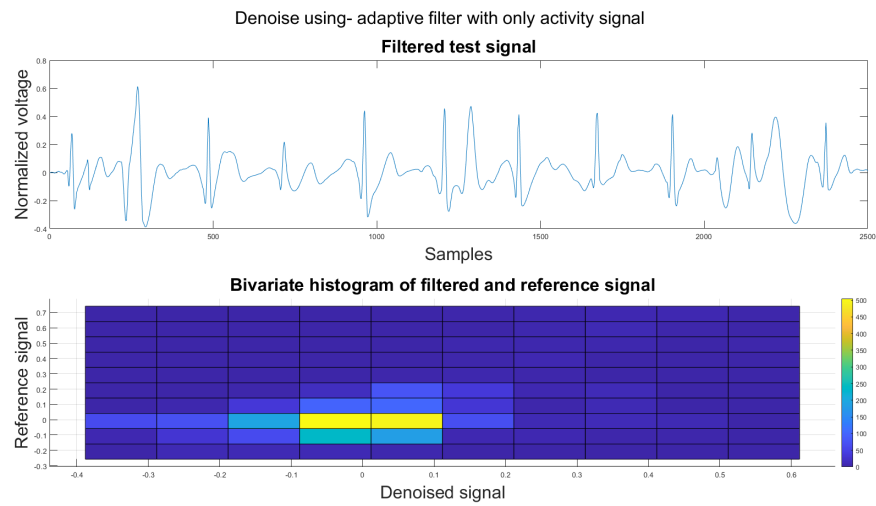

Figure A.36: (1) Test signal (2) bivariate histogram between test and reference signal