Fusion of Multi-stage CNN Features for ECG Classification

Zahra Golrizkhatami

Submitted to the
Institute of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Computer Engineering

Eastern Mediterranean University
December 2018
Gazimağusa, North Cyprus

Approval of the Institute of Graduate Studies and Research

Assoc. Prof. Dr. Ali Hakan Ulusoy
Acting Director

I certify that this thesis satisfies the requirements of thesis for the degree of Doctor of Philosophy in Computer Engineering.

Prof. Dr. Hadi Işık Aybay
Chair, Department of Computer Engineering

We certify that we have read this thesis and that in our opinion it is fully adequate in scope and quality as a thesis for the degree of Doctor of Philosophy in Computer Engineering.

Assoc Prof. Dr. Adnan Acan
Supervisor

Examining Committee

1. Prof. Dr. Aydın Alatan _____

2. Prof. Dr. Tolga Çiloglu _____

3. Prof. Dr. Marifi Güler _____

4. Assoc. Prof. Dr. Adnan Acan _____

5. Assoc. Prof. Dr. Mehmet Bodur _____

ABSTRACT

Detecting and classifying cardiac arrhythmias is critical to the diagnosis of patients with cardiac abnormalities. Identification and classification of abnormalities are time consuming because it often requires analysing each heartbeat of the ECG recording. Moreover, computerized ECG classification can also be very useful in shortening hospital waiting lists and saving the life by discovering heart diseases at early stages. Therefore, automatic classification of the arrhythmias using machine-learning technologies can bring various benefits. In this thesis, novel and high-performance approaches based on deep learning techniques are proposed for the automatic classification of electrocardiogram (ECG) signals. In this research work, two fully automatic systems have been presented which are shown to have high efficiency and low computational cost.

In one of the proposed systems, a novel decision-level fusion of features is presented by three different approaches; the first one uses normalized feature-level fusion of handcrafted global statistical and local temporal features by uniting these features into one set, the second one uses the morphological feature subset, and the third one combines features extracted from multiple layers of a Convolutional Neural Network (CNN) through using a score-level based refinement procedure.

The second proposed system utilized a new architecture of deep neural networks, Directed Acyclic Graph Convolutional Neural Networks (DAG-CNNs). DAG-CNNs fuse the feature extraction and classification stages of the ECG classification into a

single automated learning procedure and utilize the multi-scale features and perform the score-level fusion of multiple classifiers automatically.

The results over the MIT-BIH arrhythmia benchmark database exhibited that the proposed systems achieve superior classification accuracy compared to all of the state-of-the-art ECG classification methods.

Keywords: electrocardiogram, convolutional neural networks, directed acyclic graph CNN, morphological feature, statistical feature, temporal features, multi-stage CNN-based features, feature-level fusion, score-level fusion, decision-level fusion.

ÖZ

Kalp ritim düzensizliklerinin tespit ve sınıflandırılması ritim bozukluğu yaşayan hastaların tanı ve tedavisi için kritik öneme sahiptir. Kalp ritim bozukluklarının tanınması ve sınıflandırılması her bir elekrokardiyografi (ECG) kaydının tek tek incelenmesini gerektirdiğinden zaman alan bir işlemlerdir. Dolayısıyla, ECG kayıtlarının bilgisayar yardımlı sınıflandırılması hastanede bekleme sıralarının kısaltılmasında ve kalp rahatsızlıklarının erken teşhis edilerek hayat kurtarılmasında çok kullanışlı olabilir. Bu anlamda, makina öğrenme teknolojileri kalp ritim bozukluklarının otomatik sınıflandırılması için çeşitli yararlar sağlarlar. Bu tezde, ECG sinyallerinin otomatik sınıflandırılmasına yönelik olarak derin öğrenme tekniklerine dayalı yeni ve yükek başarımlı yaklaşımlar önerilmiştir. Araştırma çalışmaları sonucunda, yüksek verim ve düşük hesaplama maliyetine sahip tam otomatik iki sistem önerilmiştir.

Önerilen sitemlerin birinde yenilik olarak özelliklerin karar verme düzeyinde birleştirilmesi üç farklı yaklaşımla gerçekleştirilmiştir. Birinci yaklaşım tümel istatistiksel özellikler ile zaman tabanlı yerel özelliklerin bir normallenmiş küme içerisinde birleştirilmesinden oluşur. İkinci yaklaşım biçimsel özelliklerin bir alt kümesini kullanır, üçüncü yaklaşım ise çok katmanlı evrişimli bir sinir ağından çıkarılan özellikleri bir puan tabanlı işlev kullanarak arıtıp birleştirir.

Önerilen ikinci sistem derin sinir ağları için yönlü döngüsüz çizge gösteriminde yeni bir evrişimli sinir ağı mimarisi (DAG-CNNs) sunar. Bu mimaride ECG sinyalleri için özellik çıkarımı ve sınıflandırma aşamaları tek bir otomatik öğrenme yöntemine

indirgenir ve puan tabanlı birleştirme çok ölçekli özelliklerin çoklu sınıflandırıcılar tarafından kullanılmasıyla elde edilir.

MIT-BIH ritim bozukluğu veri tabanı kullanılarak yapılan deneysel çalışmaların sonuçları göstermiştir ki, önerilen sistemler mevcut gelişmiş ECG sınıflandırma yöntemlerine göre üstün sınıflandırma başarımına ulaşmışlardır.

Anahtar Kelimeler: Electrokardiogram, evrişimli sinir ağları, yönlü döngüsüz çizge CNN, biçimsel özellikler, istatistiksel özellikler, zaman tabanlı özellikler, çok etaplı CNN tabanlı özellikler, özellik seviyeli birleştirme, puan seviyeli birleştirme, karar seviyeli birleştirme.

ACKNOWLEDGMENT

I would like to express my special appreciation and thanks to my supervisor Assoc. Prof. Dr. Adnan Acan, you have been a tremendous mentor for me. I would like to thank you for encouraging my research and for helping me to grow as a research scientist. Your advice on both research as well as on my career have been priceless.

At this moment of accomplishment, special heart-felt gratitude goes to my husband, Dr. Shahram Taheri. His constant love and support have made life and Ph.D. study a wonderful experience. I am blessed to have such an amazing partner for life and the eternities.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AAMI	American Association of Medical Instrumentation
AI	Artificial Intelligence
ANN	Artificial Neural Networks
BP	Back Propagation
BT	Breaking-Ties
CAD	Computer-Aided Diagnosis
CNN	Convolutional Neural Network
CPNN	Conditional Probability Neural Network
CS	Cumulative Score
DAG	Directed Acyclic Graph
DBN	Deep Belief Networks
DCT	Discrete Cosine Transform
DIP	Digital Signal Processing
DWT	Discrete Wavelet Transforms
ECG	Electrocardiogram
FIR	Finite Impulse Response
GCNN	Generic Convolutional Neural Network
ICA	Independent Component Analysis
ILSVRC	ImageNet Large Scale Visual Recognition Challenge
KNN	K-Nearest Neighbour
KPLS	Kernel Partial Least Square
LCNN	Lead Convolutional Neural Network

LDA	Linear Discriminant Analysis
MLP	Multi Layer Perceptron
NLP	Natural Language Processing
NN	Neural Networks
NN	Nearest Neighbour
ODFL	Ordinal Deep Feature Learning
OH	Ordinal Hyperplanes
PCA	Principal Component Analysis
ReLU	Rectified Linear Units
RBM	Restricted Boltzmann Machine
SDAE	Stacked Denoising Auto-Encoders
SFFS	Sequential Forward Floating Search
SReLU	S-shaped Rectified Linear Unit
SVEB	Supraventricular Ectopic Beats
SVM	Support Vector Machine
SVR	Support Vector Regressor
ST	S-Transform
TDCNN	Tuned Dedicated CNN
VEB	Ventricular Ectopic Beats
WT	Wavelet Transform

Chapter 1

INTRODUCTION

1.1 Problem Description and Motivation

Heart's electrical activities are captured by using some electrodes on chest's surface which can record the electrical signal that appears by cumulative heart cells action. Electrocardiogram (ECG) classification is one of the most challenging tasks in heartbeat analysis. Medical centres are using ECGs in order to detect various cardiovascular diseases. By monitoring a patient's ECG tape; expert cardiologists are able to recognize various cardiac arrhythmias that can cause serious heart diseases. During the last decade, researchers have proposed different pattern recognition systems in order to detect such arrhythmias automatically, which have been very helpful for cardiologists and clinicians in hospitals. Although collecting the ECG data is easy, challenges on extracting the most useful information from the ECG signal still exists. Also, due to limited accuracy of visual and manual interpretation of ECGs, researchers proposed the use of computer-aided diagnosis (CAD) systems for the analysis and interpretation of these signals automatically.

Electrocardiogram provides health information for patients. Cardiologists can detect various heart abnormalities by checking the ECG waveform. Electrocardiogram was invented by W. Einthoven in 20^{th} century. Since nowadays heart diseases are a common death reason of people in developed countries, many researchers are working on ECG analysis.

Electrical and mechanical heart actions are joining together. Electrocardiography is an essential device to estimate the heart's activity and it provides the information of normal and abnormalities of the heart. An ECG record consists of the repeatedly heart beats. Each single heart beat includes many waves and inter-waves.

Recently, the classification of heartbeat signal by using deep learning approaches is one of the promising techniques in detection of arrhythmias at early stages. Moreover, neuroscience problems, such as ECG signal classification, present a particular set of challenges that require cutting edge in machine learning.

Deep Learning and especially CNN is one of the best choices in many well-known artificial intelligence (AI) applications such as speech recognition, signal and image processing and natural language processing (NLP) [1]. CNN tries to mimic human brain functionality by using large data. Collecting a large annotated dataset for a particular problem is one of the CNN main challenges. Other CNN challenges are hyper-parameter optimization, high-performance hardware requirement and overfitting problem. These CNN issues lie in ECG classification as well. The number of annotated heartbeats in the public datasets is limited and strongly unbalanced in different arrhythmia types. In consequence, comparing to deep CNN architecture in image classification fields, the number of CNN layers in ECG classification systems are limited. However, the computational complexity is lower than CNN image-based applications.

Based on the demonstrated success of CNNs for biomedical signal and image processing, the research work presented in this thesis proposes a novel system that utilizes multi-stage CNN-based features that are combined with hand-crafted features

for the automated diagnosis of heartbeat signals. The set of handcrafted features consists of three subsets namely, wavelet transform based morphological features representing localized signal behaviour, statistical features exhibiting overall variational characteristics of the signal and temporal features representing the signal's behaviour on the time axis.

As illustrated in Figure 1.1, the generated waves distribute among the body and we record ECG motion and its wave components, such as P-wave, Q-wave, R-peak and S-wave. P-wave shows depolarisation of the atria, therefore the blood current moves from atria to ventricles. P-Q interval illustrates the generated wave from atria to ventricles. QRS complex indicates the depolarization of ventricles so blood is exited from right ventricle to arteria pulmonalis and also from left ventricle to aorta. Repolarization of atria cannot be observed during the recording since the QRS section covers it. Repolarization of ventricles is known as T-wave.

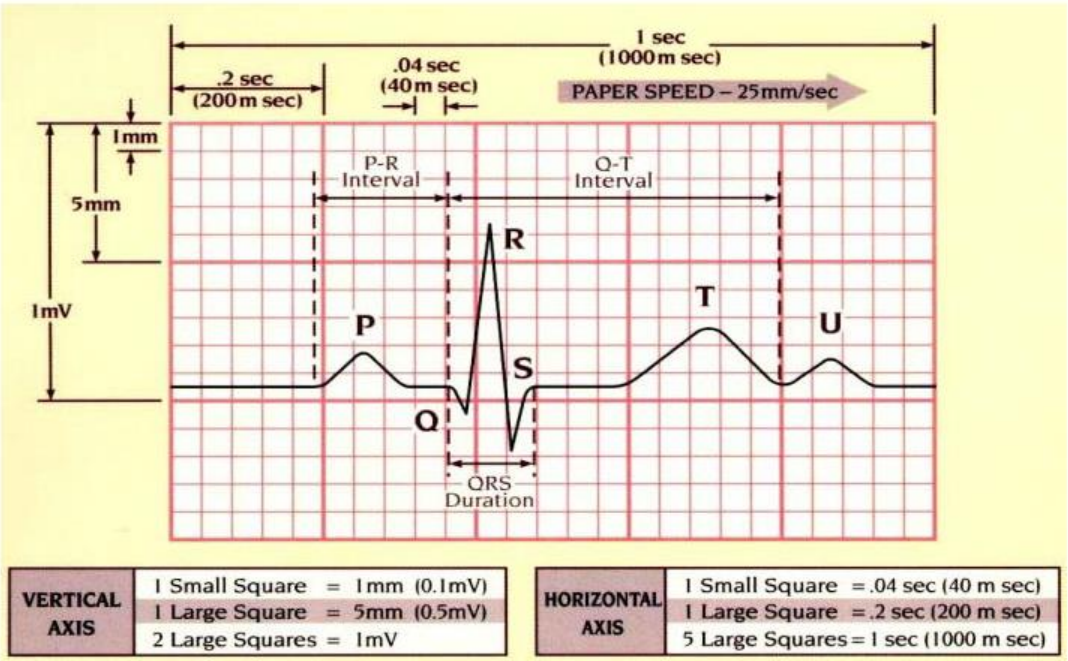


Figure 1.1: An ECG Waveform with the Standard ECG Intervals [2]

The muscles of heart can be affected by Cardiac arrhythmias which are the reason of disorder rhythms. This problem can be an obstacle to pumping the blood. When the blood pumping is not sufficient, it will increase the risk of death. Common clinical arrhythmia detection is based on an expert's decision. Since it is critical to assess and monitor a patient heart's situation, various automatic methods have been proposed, but majority of them have heavy computational cost to extract proper features, and they are able to classify limited disease types. Current available systems are very sensitive to noise and insufficient robustness is one of their weaknesses. Therefore, it is obvious that the systems need to improve their classifier's ability in order to classify overlapped classes and incomplete or noisy input samples.

In this thesis we have investigated the success of handcrafted and learned features for ECG classification. Additionally, various extension of deep learning architectures have been proposed for classifying the heartbeat signals.

The rest of this thesis is organized as follows: previous studies on ECG classification have been reviewed in Chapter 1. In Chapter 2, an introduction to ECG and signal processing are given. Mathematical methods are explained in Chapter 3, Background Materials have been stated in Chapter 4 and the data set used in this thesis has been introduced in Chapter 5. Pre-processing has been presented in Chapter 6. Proposed methods and experimental results are demonstrated in Chapter 7, Chapter 8, respectively. Finally, the Conclusion has been stated in Chapter 9.

Chapter 2

THE STATE OF THE ART

Recently many physicians use automated interpretation of ECGs for supporting their decisions. The performance and accuracy of some ECG analysers are approximately as well as expert physicians. There are various types of digital signal processing (DSP) procedures, varying from simple to complex, which are used for analysing heart activity and electrocardiography. These methods can be categorized into 3 classes: time-domain approach, frequency-domain approach, and time-frequency domain approach.

Time domain and frequency domain approaches are common methods and have good performance in QRS detection and recognition its onset-offset positions. Recent methods are motivated to use combination of time and frequency domain in time-frequency approach and use the prior methods benefits. These methods use frequency analysis and combine its results with time domain features extraction.

Time-domain methods don't have efficient results due to their low sensitivity. The main reason is that the amplitude of the signal has small changes in time domain. On the other hand, frequency domain approaches have more sensitivity to changes of the signal amplitude, but they can't determine the exact location of changes.

Wavelet transform (WT) is a common approach due to its easy implementation. Since it is very similar to one of the famous frequency method, Fourier transform, interpretation of its results can be done in the same way. There are various models of wavelet transform, so that they can be used in different applications. Choosing a specific kind of wavelet transform depends on the problem which can be varied from noise removal, detecting time and frequency elements, recognizing the essential peaks and etc.

For ECG classification different methods are introduced by the researchers but still none of them is completely successful. The most important part of classification is choosing proper discriminative features from raw ECG signal. Different features types have been used in order to recognize the abnormalities of ECG, such as Bayesian [3] and heuristic approaches, template matching, expert systems [4], hidden Markov models [5], artificial neural networks (ANNs) [6][7][8][9].

Most common methods are based on statistical pattern recognition approaches. They use different morphological features of ECG [10], such as interval length and amplitude of QRS complex, R-R interval, QRS component area, etc. [11]. Main disadvantage of these approaches is that they have limited ability when the morphology of ECG signal changes [12]. Despite that these methods have good accuracy, they have some drawbacks. They focused on finding some fiducial points on ECG signal which are sensible to changes of signal morphology that may occur among inter-class variation of different patient samples or even within intra-class variation of the same patient in different time. Therefore a few types of waveforms can completely capture these features.

Some researchers have used only the QRS complex features, while others added morphological features as well, which are extracted from the P-wave and T-wave [10][13]. The main limitation of this method is its accuracy directly depends on the correct detection of P and T-wave and also QRS complex component. These kinds of features are not suitable for analysing special types of arrhythmias, such as ventricular fibrillation [12].

Other approaches used Hermite functions [6], cumulate features [14], wavelets [15][16] correction waveform analysis [17], complexity measures [18], a total least squares-based Prony modelling algorithm [19], autoregressive modelling, non-linear measures and cluster analysis, etc. There are different approaches which use ANN and their combination with other approaches in order to classify ECG signal such as Fourier transform NNs [19], re-current NNs [20] and back propagation (BP) NNs [21] and etc.

In [22], the authors proposed a methodology for the classification of single-lead (ECG) signals. They exhibit the application of the Restricted Boltzmann Machine (RBM) and deep belief networks (DBN) for ECG classification of ventricular and supraventricular heartbeats. The effectiveness of this proposed algorithm is evaluated on the MIT-BIH database heartbeat signals. Simulation results showed that with a suitable selection of parameters, RBM and DBN can achieve high average recognition accuracies of ventricular ectopic beats (93.63%) and of supraventricular ectopic beats (95.57%).

A recent methodology is investigated by Yazhao et al. to classify patient-specific ECG heartbeats. The Generic Convolutional Neural Network (GCNN) is

trained first by using a large number of heartbeats without distinguishing patients. Based on the GCNN, fine-tuning technique is applied to modify the GCNN to a Tuned Dedicated CNN (TDCNN) for the corresponding individual. In order to accelerate the ECG classification, only the original ECG heartbeat is input to the CNN without other extended information from the neighbour heartbeats or FFT representation. A deeper CNN architecture with small-scale convolutional kernels is adopted to improve the speed and accuracy for classification. Accuracy of the proposed method is evaluated over MIT-BIH dataset.

A high performance CNN based arrhythmia classification system is presented in [24]. This research work proposed an automated ECG Classification by using dual heartbeat coupling based on convolutional neural network. In this study, the single channel ECG signal was segmented into heartbeats in accordance with the changing heartbeat rate. The beats were transformed into dual beat coupling matrix as 2-D inputs to the CNN classifier, which captured both beat morphology and beat-to-beat correlation in ECG. A systematic training beat selection procedure was also proposed which automatically include the most representative beats into the training set to improve classification performance. The classification system was evaluated for the detection of supraventricular ectopic beats (SVEB or S beats) and VEB using the MIT-BIH arrhythmia database. The classifier is also a personalized one by combining training set from a common pool and a subject-specific set of ECG data.

Linpeng et al. [25], present a new method based on lead convolutional neural network (LCNN) and rule inference for classification of normal and abnormal ECG records with short duration of normal and abnormal ECG records. First, two different LCNN models are obtained through different filtering methods and different training

methods, and then the multipoint-prediction technology and the Bayesian fusion method are successively applied to them. Finally, the utilization of the bias-average method is performed. The experiments are done over the Chinese Cardiovascular Disease Database with more than 150,000 ECG and achieved the accuracy of 86.22% and 0.9322 AUC (Area under ROC curve).

The authors in [26], proposed a 16-layer ECG classification problem skip connections were used to improve the rate of information transfer through the network. Skip connections led to a significant increase in the feature learning capabilities of the CNN as well as speeding up the training time. , this CNN based method identified normal rhythm, AF and other rhythms with an accuracy of 90%, 82% and 75%, respectively.

In [27], the authors proposed a deep learning approach for the classification ECG signals. In this paper, feature representation is performed in an unsupervised way by using stacked denoising auto-encoders (SDAEs) with sparsity constraint and adding a softmax regression layer on the top of the resulting hidden layer yielding the so-called deep neural network (DNN). Their method relies on the DNN posterior probabilities to associate confidence measures such as entropy and Breaking-Ties (BT) to each test sample.

Considering the fusion of features extracted by different approaches, several authors tried to integrate features from different descriptors to improve their system performance. Ye et al. [28] proposed a system for heartbeat classification based on a feature-level fusion of morphological and dynamic features. In their proposed study, the authors used wavelet transform (WT) and independent component analysis (ICA)

to extract the morphological features, while, RR interval information is computed to obtain the dynamic features. In [29], the authors employed a fast feature-fusion method for ECG heartbeat classification based on multi-linear subspace learning. They used wavelet-packet decomposition for feature extraction and generalized N dimensional ICA (GND-ICA) for feature fusion [30]. They proposed a sequential forward floating search (SFFS) algorithm in order to analyse a comprehensive feature set and find the most suitable feature subset. Sambhu et al. [31] showed that combination of the temporal, statistical and wavelet features is an effective method in the classification of cardiac abnormalities. Das and Ari [32] constructed two different systems that utilized feature-level fusion for ECG classification. The first system uses S-transform (ST) based features along with temporal features and the second one applies mixture of ST and WT based features along with the temporal ones. In all of the aforementioned studies, the information feature fusion is performed by integrating features from different hand-crafted descriptors. In this respect, the novel approach presented in this article takes the advantages of both automatically learned CNN features and hand-crafted features and applies fusion of features in a multi-level manner. Details of the presented method are illustrated in the following sections.

2.1 Pattern Recognition

Pattern Recognition is the task of classifying objects into predefined categories or classes. Pattern recognition systems can perform pattern identification and classify the objects. They perform it either by using some forms of prior information about the object distributions or some statistical knowledge which are embedded in the data. The objects could be assumed as sets of features or a series of experimental results which define the points in features space [33].

Pattern recognition systems consist of several subsystems:

- Data acquisition section which is responsible to measure or record the raw intended data.
- Feature extraction section which is responsible to extract distinctive information from the raw data.
- Feature selection section which selects the optimal subset of extracted features.
- Classification section which is the main part of the system and by using the features information classifies the input data into predefined classes.

The pattern recognition systems can be applied either supervised or unsupervised. In the supervised method, the system will be trained by using the data which has already been classified by an expert. On the other hand, unsupervised learning is referred to an algorithm which tries to find the distinctive patterns of the data by learning [34].

Various methods of supervised and unsupervised models have been implemented in literature, such as statistical pattern recognition, syntactic pattern recognition and AI approaches. Selection of these models is crucially depends on the characteristics of the problem [35].

In statistical pattern recognition approach which is based on statistical modelling of the data, we assume that patterns are produced by a stochastic system with some distribution probability. There are different type of methods such as Bayes linear classifier, the k-nearest neighbour and the polynomial classifier. Other important issues in these methods are the procedure for selecting discriminative features,

number of necessary features, and adjusting the model parameters. All of this setting is done by the classifier designer [35].

Syntactic or structural pattern recognition is an approach in which each pattern can be represented by a set of symbolic features. In this method, instead of dealing with numeric features, more complex multiple relationships between particular features are present. It is possible to use a sort of formal language in order to describe these features and uses some grammar syntax codes for discriminant [35] [36].

Artificial neural network (ANN) is one of the famous examples of artificial intelligence (AI) methods. ANN has been used in analysing non-linear signal, classification and clustering, and optimization problem. Selecting the type of topology, size of the network and number of neurons are completely problem dependent.

Chapter 3

ELECTROCARDIOGRAPHY AND SIGNAL PROCESSING

In this chapter, the anatomy of the heart and its functionality and the configuration of the ECG signal have been reviewed. The figures used in this chapter are reproduced from [37], which are freely available for public use. General descriptions of anatomy and condition system of human heart are presented in Appendix A.

3.1 Generation and recording of ECG

Human body is a good electrical conductor; hence electrical activity of the heart can be measured using surface electrodes. Electrodes record the projection of resultant vectors, which describe the main direction of electrical impulses in the heart. The overall projection is named as electrocardiogram. Different placement of electrodes provides spatiotemporal variations of the cardiac electrical field. The difference between a pair of electrodes is referred to as a lead. A large amount of possible lead systems has been invented; depending on a diagnostic purpose, a lead system is chosen and electrodes placed on accurate positions. The most commonly used system is standard 12-lead ECG system defined by Einthoven [38]: Three bipolar limb leads (I, II, III) - electrodes are placed to the triangle (left arm, right arm and left leg) with heart in the center (Figure 3.2). This placement is called the Einthoven's triangle.

The augmented unipolar limb leads (aVF, aVL, aVF) - electrodes are placed on same positions as in case of leads I, II and III. The difference is in the definition of leads.

Leads are calculated as the difference between potential of one edge of the triangle and the average of remaining two electrodes (Figure 3.3).

Unipolar precordial leads (V1-6) - leads are defined as the difference between potential of electrode on chest and central Wilson terminal (constant during cardiac cycle and is computed as average of limb leads). For details see Figure 3.4.

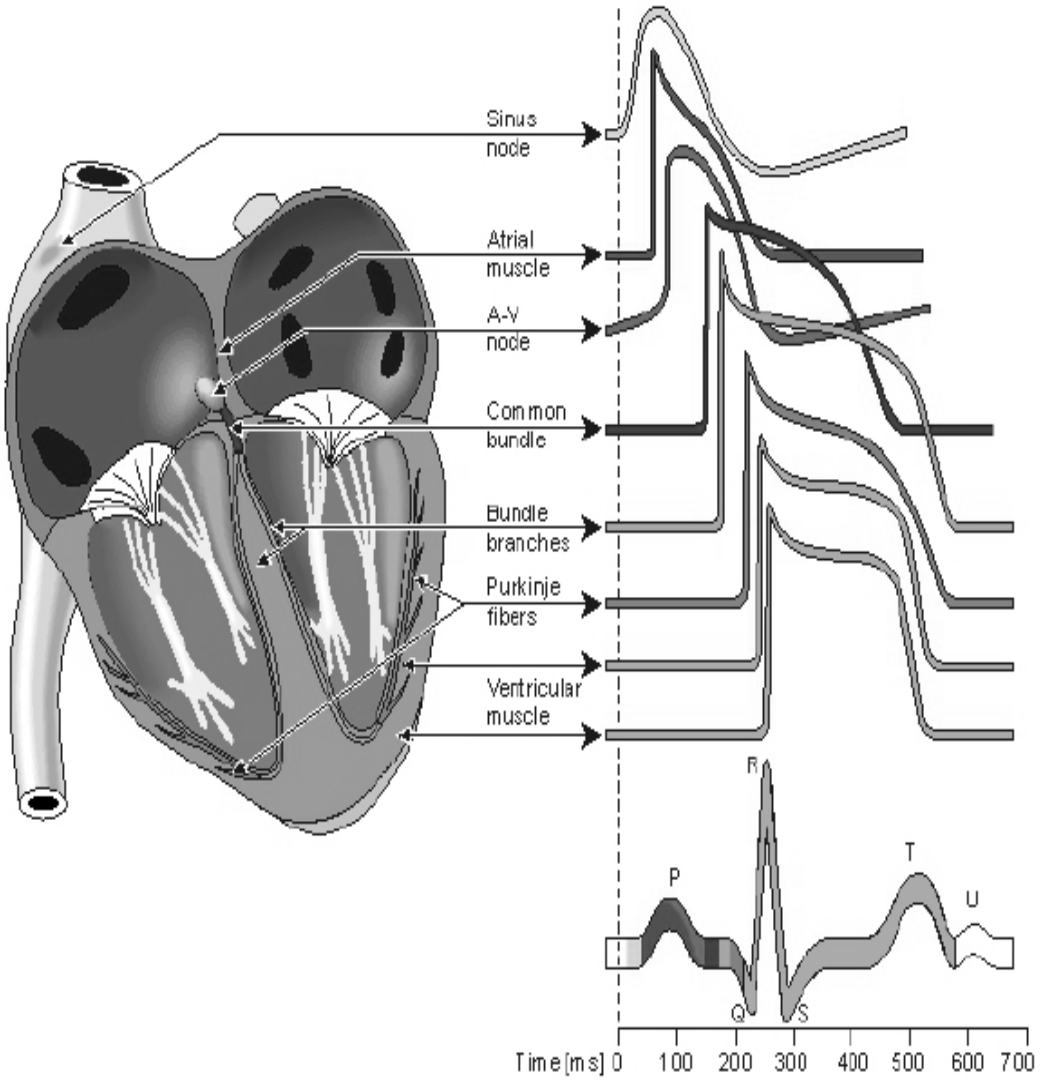


Figure 3.1: Schematic Representation of ECG Waveform Generation by Summing of Different Action Potentials [37]

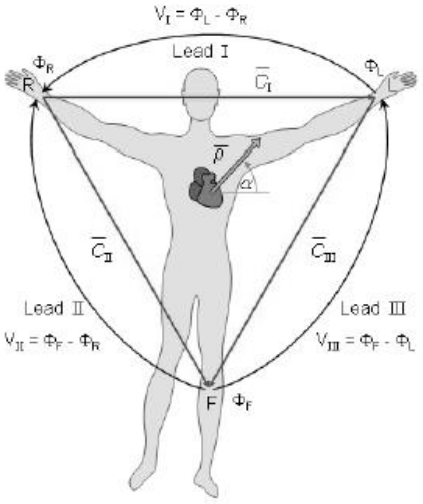


Figure 3.2: Schematic Representation of Einthoven Triangle Electrode Placement[37]

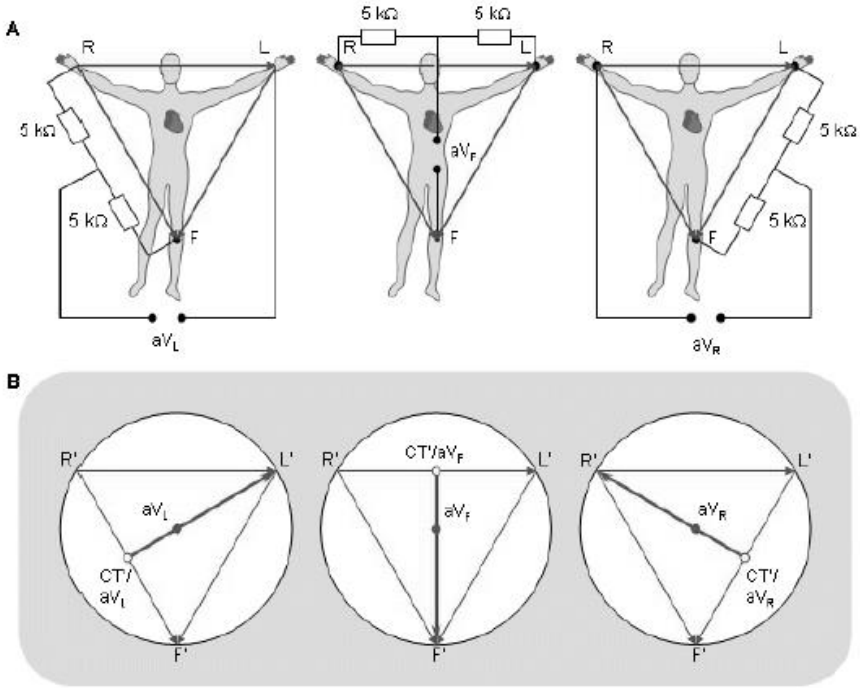


Figure 3.3: Schematic Representation of Augmented Limb Leads Calculation [37]

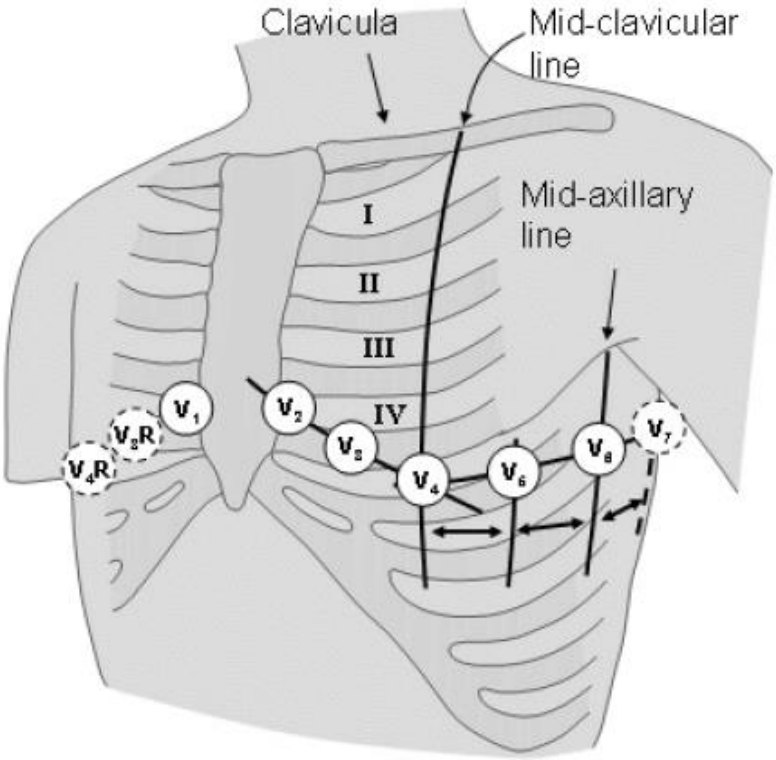


Figure 3.4: Precordial Leads Electrodes Positions [37]

3.1.1 ECG Waveform Description

As mentioned earlier, ECG wave is formed as a projection of summarized potential vectors of the heart. ECG wave has several peaks and "formations", which is useful for its diagnosis (Figure 3.5). These are:

- P-wave - indicates the depolarized wave that distributes from the SA node to the atria, and its duration is between 80 to 100 milliseconds.
- P-R interval - indicates the amount of time that the electrical impulse passing from the sinus node to the AV node and entering the ventricles and is between 120 to 200 milliseconds.
- P-R segment - Corresponds to the time between the ends of atrial depolarization to the onset of ventricular depolarization. Last about 100ms.

- QRS complex - Represents ventricular depolarization. The duration of the QRS complex is normally 0.06 to 0.1 seconds.
- Q-wave - Represents the normal left-to-right depolarization of the inter ventricular septum.
- R-wave - Represents early depolarization of the ventricles.
- S-wave - Represents late depolarization of the ventricles.
- S-T segment – it appears after QRS and indicates that the entire ventricle is depolarized.
- Q-T interval - indicates the total time that need for both repolarization and ventricular depolarization to happen, so it is the estimation for the duration of the average ventricular action. This time can vary from 0.2 to 0.4 seconds corresponding to heart rate.
- T-wave - indicates ventricular repolarization and its time is larger than depolarization.

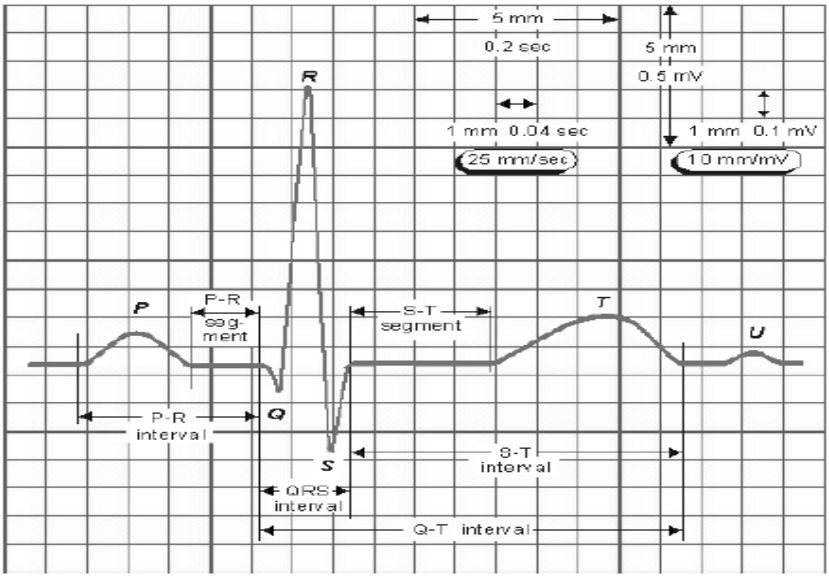


Figure 3.5: Normal ECG Waveform [37]

Chapter 4

BACKGROUND MATERIALS

4.1 Introduction

In the following sections, the mathematical background of used methods, namely, wavelet transform, neural networks, convolutional neural networks and support vector machine have been reviewed. The wavelet transform is a remarkable mathematical method with the ability to examine the signal simultaneously in time and frequency domains, in a different way from Fourier transform method.

4.2 Wavelets

Fundamental of Wavelet transform (WT) is on the use of a series of computational analysing elements called "wavelets". By applying the WT to a specific signal, its features are store in the wavelet coefficients. Each resulting wavelet coefficient corresponds to measurement in the signal in a given time instant and a given frequency band. Most of the following sections are based on the Illustrated Wavelet Transform Handbook [39]. All given information refers to this source, unless stated otherwise.

Wavelet analysis has been applied in a wide range of applications, from climate analysis to signal compression and medical signal analysis. The application of wavelet transform analysis in science and engineering began to increase in the beginning of the 1990s, directly reflecting the interest of the scientific community [31].

Some of the more commonly used wavelets are depicted in Figure 4.1. We can notice that they have the shape of a small wave, localized on the time axis. Depending both on the signal we need to analyse and what characteristic we are analysing, one wavelet can be better suited than others.

There are some important constraints for wavelet function $\psi(t)$:

- Limited finite energy:

$$E = \int_{-\infty}^{\infty} |\psi(t)|^2 \, dt < \infty \tag{4.1}$$

- It shouldn't have any zero frequency components ($\hat{\psi}(t)(0) = 0$), or if $\hat{\psi}$is the Fourier transform of $\psi(t)$:

$$\hat{\psi}(\mathrm{f}) = \int_{-\infty}^{\infty} \psi(t)\, e^{-i(2\pi f t)} dt = \infty \tag{4.2}$$

- It must hold the following constraint:

$$C_g = \int_0^{\infty} \frac{|\hat{\psi}(f)|^2}{f} df < \infty \tag{4.3}$$

  The above equation is known as admissibility condition and Cg is known as admissibility constant and is dependent on the chosen wavelet.

- An extra criterion, on complex wavelets, is that the Fourier transform must vanish for negative frequencies and also must be real.

4.2.1 Wavelet Transform

The wavelet can be used as the location or the scale. If the wavelet and the signal's appearance are close, then the convolution has a high value; otherwise, the transform results in a low value. The wavelet transform is performed at various locations and scales of the signal: this transformation can be applied as the continuous wavelet transforms (CWT) or the discrete wavelet transforms (DWT). In this thesis, DWT has been used as one part of feature extraction.

The wavelet can be manipulated in two ways: it can change its location or its scale (Figure 4.2). If, at a point, the wavelet matches the shape of the signal, then the convolution has a high value. Similarly, if the wavelet and the signal do not correlate well, the transform results in a low value. The wavelet transform is computed at various locations of the signal and for various scales of the wavelet: this is done in a continuous way for the continuous wavelet transform (CWT) or in discrete steps for the discrete wavelet transforms (DWT).

The operations over the wavelet are defined by the parameters a (for dilation) and b (for translation). The shifted and dilated versions of the wavelet are denoted as follows:

$$\psi[[t-b]/a] \tag{4.4}$$

For sake of simplicity, let us take the Mexican hat wavelet:

$$\psi(t) = (1-t^2)e^{-t^2/2} \tag{4.5}$$

The shifted and dilated equation for this type of the wavelet function would be:

$$\psi\left(\frac{t-b}{a}\right) = [1-\left(\frac{t-b}{a}\right)^2]e^{-\frac{1}{2}[(t-b)/a^2]} \tag{4.6}$$

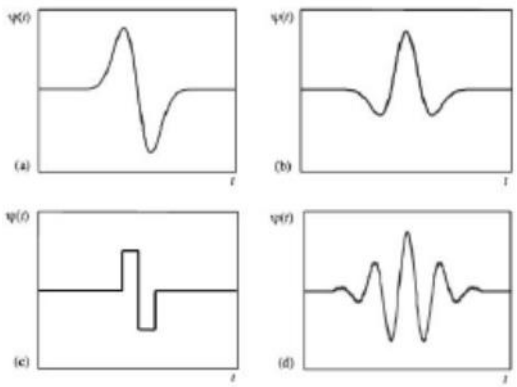


Figure 4.1: Example of wavelets: a) Gaussian Wave (first derivative of a Gaussian), b) Mexican Hat (second derivative of a Gaussian). c) Real part of Morlet [30]

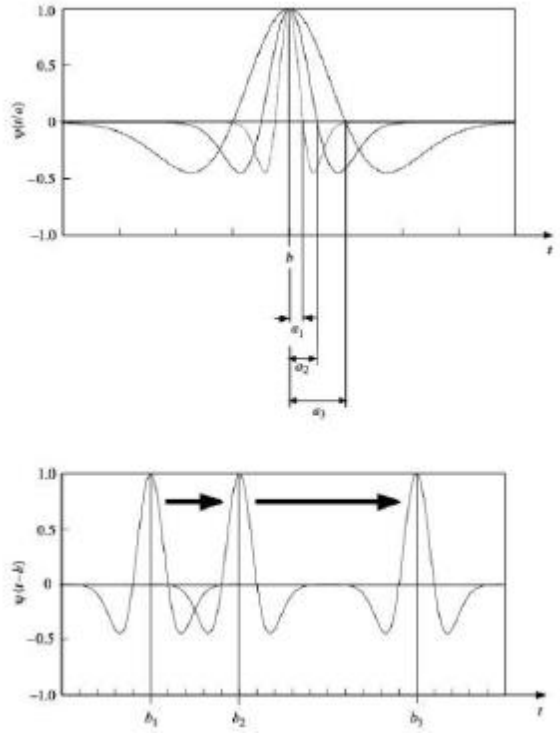


Figure 4.2: Two Possible Manipulations with Wavelets: a) Translation, b) Scale [31]

4.2.2 Discrete Wavelet Transform

Discrete Wavelet Transform (DWT) is a popular approach for representing the signal by a linear combination of its basis functions (Teolis and Benedetto 1988). DWT decomposes the signal into various frequency bands by different resolutions called detail (D) and approximation (A) components. The main advantage of DWT is its ability to provide representational information in time and frequency domains. In order to compute the time-frequency components of an ECG signal, a wavelet basis function, called the mother wavelet, is defined as follows:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right) \tag{4.7}$$

where the parameters a and b represent the scale and the shift, respectively. The DWT of an ECG signal $x(t)$, can be defined as follows:

$$T_{a,b} = \int_{-\infty}^{\infty} x(t)\,\psi_{a,b}(t)dt \tag{4.8}$$

In order to reconstruct the input signal from its wavelet coefficients, the inverse DWT is computed as,

$$x(t) = \sum_{a=-\infty}^{\infty} \sum_{b=-\infty}^{\infty} T_{a,b}\psi_{a,b}(t) \tag{4.9}$$

Figure 4.3, shows the DWT decomposition tree of an input signal and its approximation (A) and detail (D) coefficients up to fourth level. The input signal is represented by its corresponding wavelet coefficients as $(A_4 + D_4 + D_3 + D_2 + D_1)$.

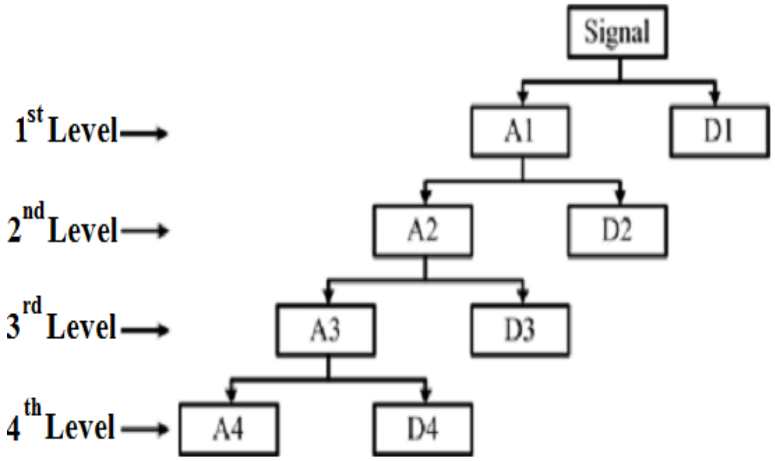


Figure 4.3: DWT Decomposition Tree [40]

4.3 Neural Networks

Artificial Neural Networks (ANNs) or short Neural Networks (NNs) have been modelled to handle complex tasks by following patterns that work like the human brain. NNs contain interconnected units with very simple functions. When these simple units combined together, it can build a complex classification function.

The first known structure was built in 1958 by Rosenblatt [41]. As illustrated in Figure 4.4, his network was very simple. The network consisted of one neuron that

had multiple inputs and one output. The output is the sum of all the weighted input x_i and the bias b as described in Equation (4.8),

$$f(x) = \varphi(b + \sum_i^n x_{i=1}.w_i) \tag{4.8}$$

where φ the Heavyside step function is defined as,

$$\varphi(x) = f(x) = \begin{cases} 1 & if\ x \geq 0 \\ 0 & else \end{cases} \tag{4.9}$$

This network was used to classify two classes: if the result of the summation is greater than or equal to zero, than the network votes for the first class. Otherwise, if the result is less than zero, then the network votes for the other class.

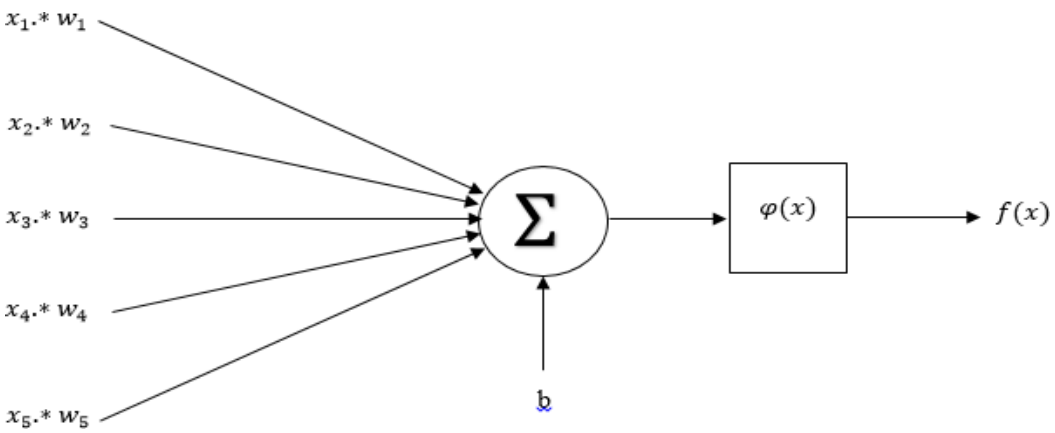


Figure 4.4: The Neural Network was Proposed by Rosenblatt [41]

When the data are not linearly separable, this simple architecture cannot handle the tasks. This weakness motivates the using of Multi-Layer Perceptron. Multi-Layer Perceptron (MLP) is a feed forward artificial neural network architecture that has one or more than one hidden layer. Figure 4.5, illustrates a MLP with one hidden layer containing neurons where each neuron in each hidden layer represents one perceptron.

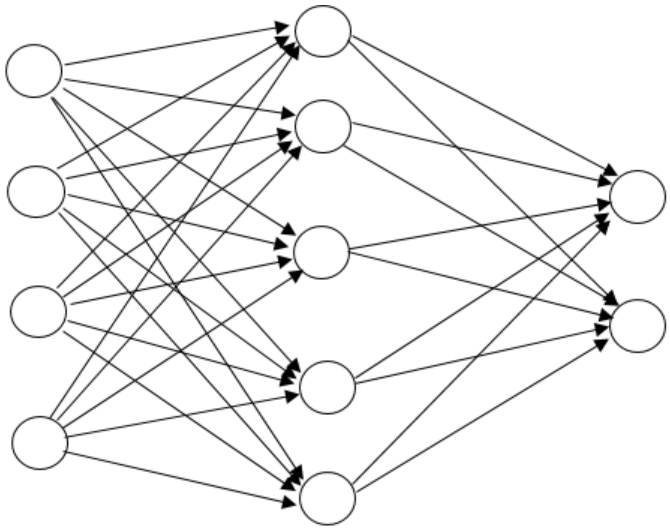


Figure 4.5: Multi-Layer Perceptron with a Single Hidden Layer

The hidden layer receives data from the input layer and then maps it to the output layer. The weights in the MLP layers are generated from the training data.

4.4 Convolutional Neural Network

Deep learning is a subfield of machine learning that is inspired by artificial neural networks, which in turn are inspired by biological neural networks. A specific kind of such a deep neural network is the convolutional network, which is commonly referred to as CNN or ConvNet. It's a deep, feed-forward artificial neural network. The models are called "feed-forward" because there are no feedback connections in which outputs of the model are fed back into itself.

CNNs specifically are inspired by the biological visual cortex. The cortex has small regions of cells that are sensitive to the specific areas of the visual field. This idea was expanded by a captivating experiment done by Hubel and Wiesel in 1962 [42]. In this experiment, the researchers showed that some individual neurons in the brain activated or fired only in the presence of edges of a particular orientation like vertical

or horizontal edges. For example, some neurons fired when exposed to vertical sides and some when shown a horizontal edge. Hubel and Wiesel found that all of these neurons were well ordered in a columnar fashion and that together they were able to produce visual perception. This idea of specialized components inside of a system having specific tasks is one that machines use as well and one that you can also find back in CNNs.

Convolutional neural networks have been one of the most influential innovations in the field of computer vision. They have performed a lot better than traditional computer vision and have produced state-of-the-art results. These neural networks have proven to be successful in many different real-life case studies and applications, such as: Image and signal classification, object detection, segmentation, face recognition; Self-driving cars that leverage CNN based vision systems, Classification of crystal structure using a convolutional neural network; and many more. Alex Krizhevsky in 2012 [43] used convolutional neural networks to win that year's ImageNet Competition, reducing the classification error from 26% to 15%.

ImageNet Large Scale Visual Recognition Challenge (ILSVRC) which began in the year 2010 is an annual competition where research teams assess their algorithms on the given data set and compete to achieve higher accuracy on several visual recognition tasks. This was the time when neural networks regained prominence after quite some time. This is often called the "third wave of neural networks". The other two waves were in the 1940s until the 1960s and in the 1970s to 1980s.

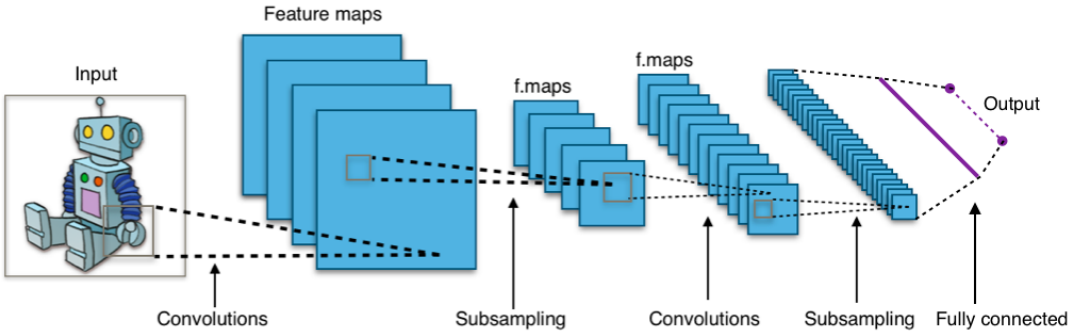


Figure 4.6: Convolutional Neural Network [44]

ConvNets contain one or more of each of the following layers:

- convolution layer
- ReLU (rectified linear units) layer
- pooling layer
- fully connected layer
- loss layer (during the training process)

Figure 4.6, shows that the input image is fed as an input to the network, which goes through multiple convolutions, subsampling, a fully connected layer and finally outputs something.

4.4.1 Convolution Layer

The convolution layer computes the output of neurons that are connected to local regions or receptive fields in the input, each computing a dot product between their weights and a small receptive field to which they are connected to in the input volume. Each computation leads to extraction of a feature map from the input image.

4.4.2 ReLU Layer

The rectified linear units (ReLU) layer commonly follows the convolution layer. The addition of the ReLU layer allows the neural network to account for non-linear relationships, i.e. the ReLU layer allows the ConvNet to account for situations in

which the relationship between the pixel value inputs and the ConvNet output is not linear. Note that the convolution operation is a linear one. The output in the feature map is just the result of multiplying the weights of a given filter by the pixel values of the input and adding them up:

$$y = w_1x_1 + w_2x_2 + w_3x_3 + \cdots \tag{4.10}$$

where w is a weight value and x is a pixel value.

The ReLU function takes a value x and returns 0 if x is negative and x, if x is positive.

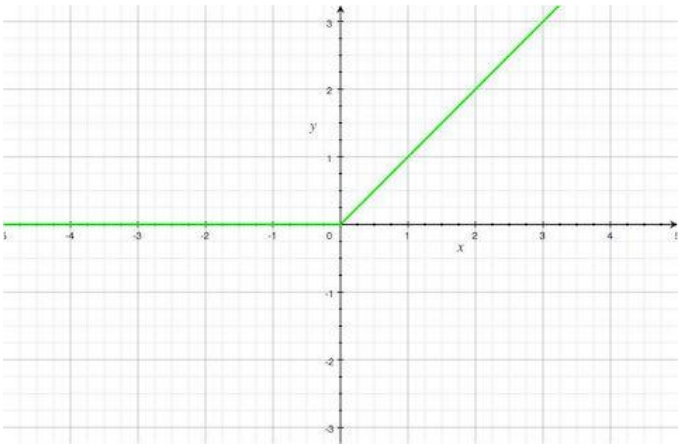


Figure 4.7: ReLU , f(x) = max(0,x)

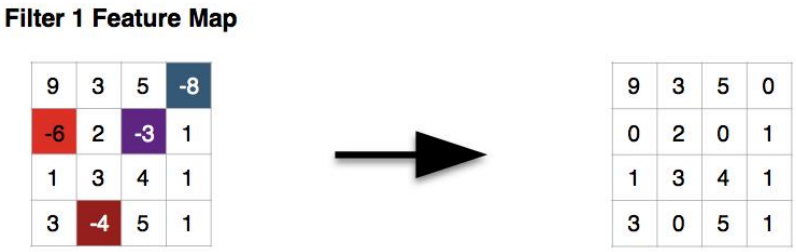


Figure 4.8: Applying ReLU on a Feature Map [45]

4.4.3 Pooling Layer

The pooling layer also contributes towards the ability of the ConvNet to locate features regardless of where they are in the image. In particular, the pooling layer

makes the ConvNet less sensitive to small changes in the location of a feature, i.e. it gives the ConvNet the property of translational invariance in that the output of the pooling layer remains the same even when a feature is moved a little. Pooling also reduces the size of the feature map, thus simplifying computation in later layers.

One of the techniques of subsampling is max pooling that takes the largest value from the window of the image currently covered by the kernel. For example in Figure 4.9, a max-pooling layer of size 2 x 2 is applied on a 4x4l image.

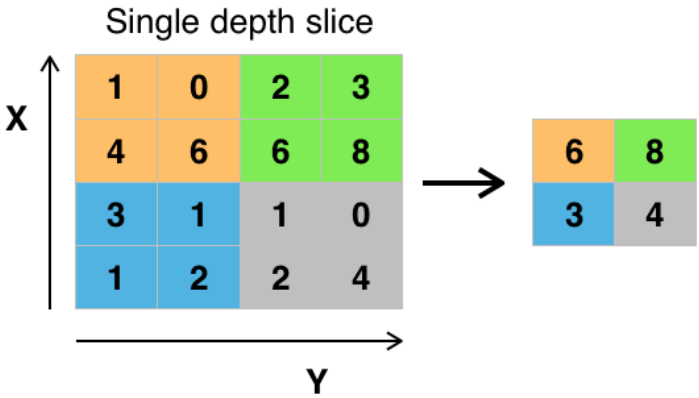


Figure 4.9: Max-Pooling [45]

The objective of the fully connected layer is to flatten the high-level features that are learned by convolutional layers and combining all the features. It passes the flattened output to the output layer where you use a softmax classifier or a sigmoid to predict the input class label.

4.4.4 The Fully-Connected and Loss Layers

The fully-connected layer is where the final "decision" is made. At this layer, the ConvNet returns the probability that an object in a photo is of a certain type. The fully-connected layer has at least 3 parts - an input layer, a hidden layer, and an

output layer. The input layer is the output of the preceding layer, which is just an array of values.

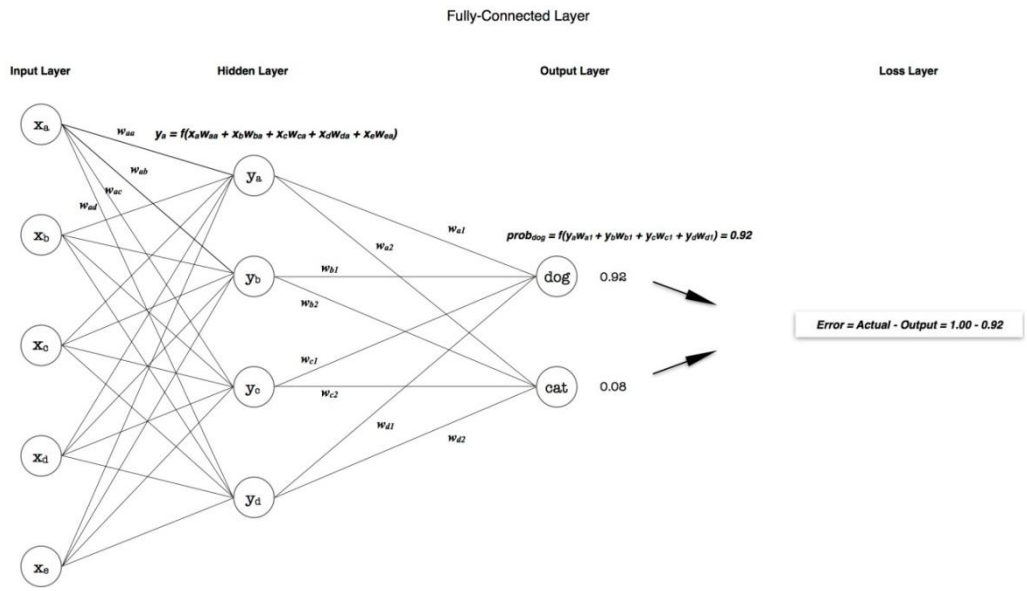


Figure 4.10: Fully-Connected and Loss Layers [45]

Figure 4.10, shows a fully-connected network for classifying the input image into two classes, dog or cat. In this example, the fully-connected layer might return an output like "0.92 dog, 0.08 cat" for a specific image, indicating that the image likely contains a dog.

Following the fully-connected layer is the loss layer, which manages the adjustments of weights across the network. Before the training of the network begins, the weights in the convolution and fully-connected layers are given random values. Then during training, the loss layer continually checks the fully-connected layer's guesses against the actual values with the goal of minimizing the difference between the guess and the real value as much as possible. The loss layer does this by adjusting the weights in both the convolution and fully-connected layers.

4.5 Support Vector Machine (SVM)

Support Vector Machine (SVM) developed by Boser, Guyon, and Vapnik in 1992. SVM is a supervised learning algorithm which can be used for different applications, from pattern classification to regression analysis [46]. In other words, SVM is a tool which uses a training dataset in order to create maximum prediction accuracy classifier while it avoids over-fitting to training data. The first application which made SVM so popular was a task for classification of handwriting. The SVM results are comparable to large NNs with complicated features [47]. Nowadays SVM is used in various areas like face recognition, text classification, signal classification and etc. [48]. Generalization is defined as the ability of a classifier to correctly classify an unseen data [50]. One fundamental objective of the machine learning algorithms is to learn the behaviours of the target functions. In other words, machine learning algorithms aim to generate a hypothesis that correctly classify the training data without over fitting to the data; however in the early algorithms they didn't pay attention to this important point [49].

Neural networks (NNs) show a good performance in both unsupervised and supervised classification task. One of the famous architecture for such learning task is Multilayer perceptron (MLP) which can be used for general function approximation. In MLP we can design multiple inputs and outputs neuron. The learning process and finding the proper weight connection can be done with input-output patterns [51].

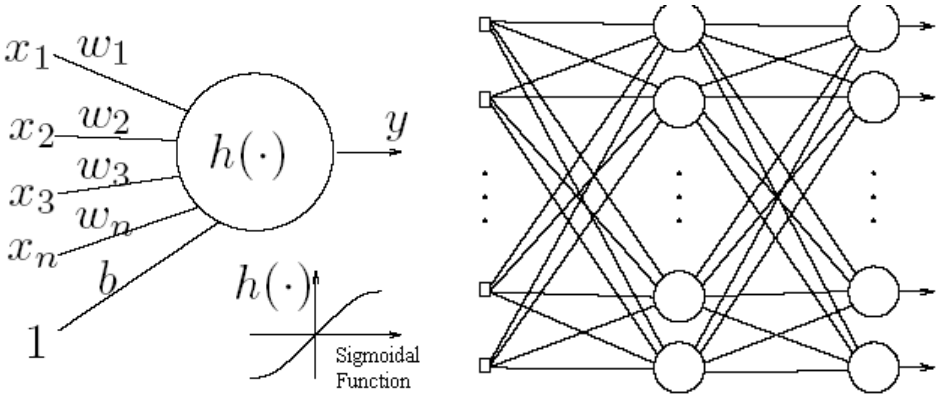


Figure 4.11: left) Simple Neural Network, right) Multilayer Perceptron [52]

But NNs have some drawbacks: they may convergence to local minima. Another disadvantage of NNs is that there are many tuning parameters such as number of neurons, learning rate and etc. which is need to correctly selected for a specific task. In order to understand the necessity of SVM, in figure 4.12, we plot some sample data and try to find a linear classifier for them. Figure4.12 shows multiple hyper planes that can correctly classify the data.

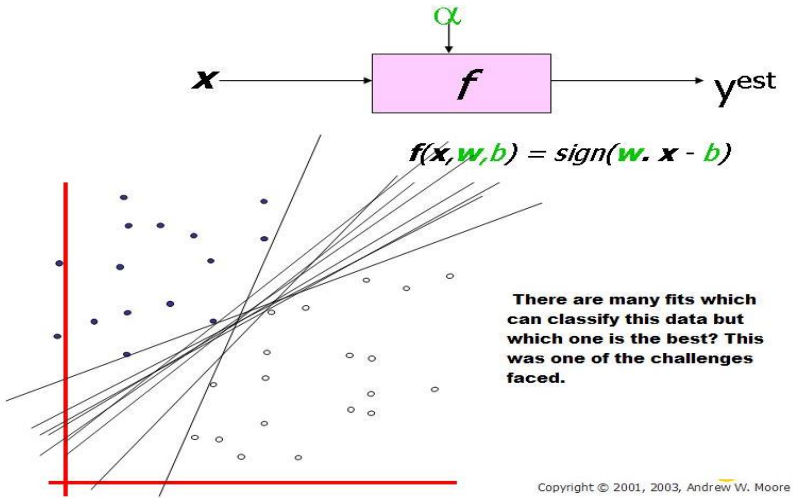


Figure 4.12: Multiple Possible Linear Classifiers for a Certain Data Set [47]

According to prior explanation, different linear classifier can be found to classify these data although some of them have better separation. It is important to have maximum margin separator since if we select a hyper plane for classification, it is

probable to be quite close to some of the samples in respect to the others. Then when an unseen test data entered to the system it is more likely to classify correctly. Figure 4.13, shows an example of maximum margin classifier and how it solves this problem [53].

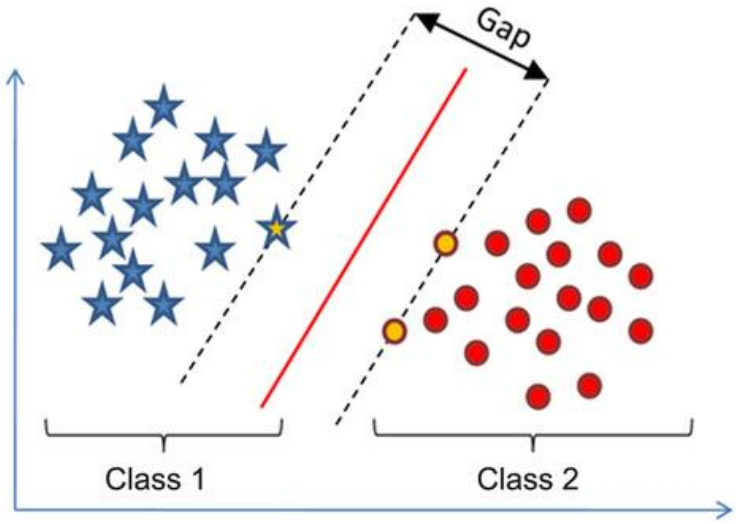


Figure 4.13: Example of Linear SVM [49]

The equation for obtaining Maximum margins is [49] [53]:

$$margin \equiv arg_{x \in D} \min d(x) = arg_{x \in D} \frac{|x.w+b|}{\sqrt{\sum_{i-1}^{d} w_i^2}} \quad (4.11)$$

In the previous example maximum distance is achieved by linear classifier. One of the reason is that the maximum margin classifier provide better result than the other classifiers since if a little error occurred in estimating the location and direction of classifier hyperplane, we still have chance to classify test data accurately.

The aim of SVM is to find a decision boundary to separate different classes of the training data. If it is not possible to do it by a linear hyperplane then SVM map the training data into a higher dimensional feature space by using some predefined kernel functions [53]. This fundamental can be written as the following formulas:

a) If Y_i= +1 or $\mathbf{x_i}$ belongs to class 1 then $\mathbf{w}.\mathbf{x_i} + b \geq 1$ (4.12)

b) If Y_i= -1 or $\mathbf{x_i}$ belongs to class 2 then $\mathbf{w}.\mathbf{x_i} + b \leq -1$ (4.13)

or we can combine this equations in the following one:

$$\forall \mathbf{x_i} : Y_{i*} (\mathbf{w}.\mathbf{x_i} + b) \geq 1 \quad (4.14)$$

In these equations $\mathbf{x_i}$ is a pattern vector and $\mathbf{w}$ is learned weight vectors. There may be multiple hyperplane in feature space that satisfy this constraint, support vector machine chooses the hyper plane where its distances to the closest sample of each classes are as far as possible.

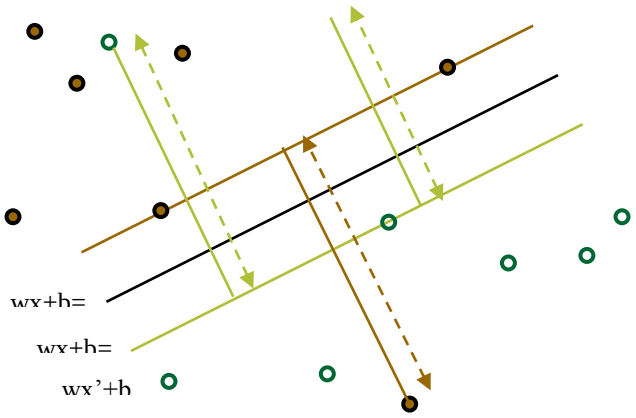


Figure 4.14: SVM hyper planes [54]

4.5.1 Kernel Trick

If the classes can be separated linearly, data can be discriminated by a linear decision boundary. But in practical situation, classes cannot separate linearly and the decision boundary is a curve with the degree higher than 1. For solving this problem we can uses kernels which are functions that map the input data feature vector to a higher dimensional space. The mapped data in new space can be separate linearly [46]. As an example we can define a simple mapping kernel as shown in figure 4.15, [54].

The Kernel formula is:

$$K(x, y) = \varphi(x).\varphi(y) \tag{4.15}$$

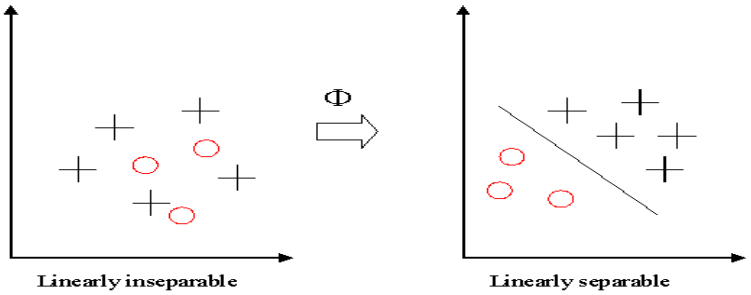


Figure 4.15: Kernels approach [54]

4.5.2 Expanding Feature Space

Increasing the dimension of feature space give us a higher chance to classify the data which is not linearly separable [46].

$$< x_1.x_2 > \leftarrow k(x_1, x_2) =< \emptyset(x_1).\emptyset(x_2) > \tag{4.16}$$

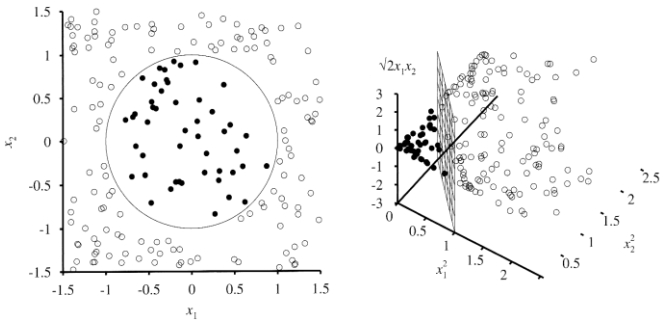


Figure 4.16: Changing the feature space dimensions from 2 into 3 [54]

4.5.3 Popular Kernel Functions

- Polynomial:

$$K(x, x') =< x, x' >^d \tag{4.17}$$

$$K(x, x') = (< x, x' > +1)^d \tag{4.18}$$

- Gaussian Radial Basis Function:

$$K(x, x') = exp\left(-\frac{\|x-x'\|^2}{2\sigma^2}\right) \quad (4.19)$$

- Exponential Radial Basis Function:

$$k(x, x') = exp\left(-\frac{\|x-x'\|}{2\sigma^2}\right) \quad (4.20)$$

Chapter 5

MIT-BIH ARRHYTHMIA DATABASE

The Massachusetts Institute of Technology Beth Israel Hospital (MIT-BIH) arrhythmia database [55] is a well-known source which provides various biomedical datasets, such as multiple ECG datasets, EEG datasets and etc. the ECG dataset which has been used in this study, contains the ECG signals which had been collected from patients in the Arrhythmia hospital laboratory at the Beth Israel Hospital between 1975 and 1979.

From 4000 Holter tape records, 48 records have been annotated by experts and divided into two partitions. The first partition (labelled 1xx) includes 23 records which are more common and routine arrhythmia types. The second partition consists of 25 records (labelled 2xx) which are more complex and uncommon arrhythmias that ends to more difficulty to detect.

The samples were collected from 25 male in age ranges of 32 to 89 years old, and 22 female with age ranges of 22 to 89 years old. Each record is about 30 minutes in length. The signals were sampled at a frequency of 360 Hertz, but not necessarily at the same gain due to different capturing equipment with variant electrical gains for digitization of the records. Moreover, the digital amplitude values range between [0, 2047], where 1024 represents 0 volts. Hence, the normalization process must be done before the signals being used.

Annotation file of each record contains useful information such as the occurrence time of R-peak locations or class of the corresponding heartbeat classes. In MIT-BIH database, there are 15 different heartbeat categories as summarized in Table 1. The patient's age diversity and the standard physical conditions make the MIT-BIH database, the best choice for ECG analysis.

Various ECG types in the MIT-BIH database are as follows:

1- Normal Sinus Rhythm (N): this is the term for the normal condition (Figure 5.1).

2- Left Bundle Branch Block Beat (L): this arrhythmia is caused by a problem in conduction in the His bundle in the left side ventricle. This is seen as a widening of the QRS complex. This ECG type is invariably an indication of heart disease [56]. Figure 5.2, indicates that the QRS complex is notably wider than that shown in Figure 5.1, this is due to the extra time taken for depolarization caused by poor electrical conduction (block).

3- Right Bundle Branch Block Beat (R): the cause of this arrhythmia is similar to **(L).** However, the conduction problem now occurs on the right side of the His bundle branch and the ECG indicates a problem in the heart but also can be seen in a healthy heart. This type of arrhythmia is identified by a wide bimodal QRS complex (Figure 5.3).

4- Paced Beat (P): this problem arises in patients that have been fitted with an artificial pacemaker. Pacemakers are used when a person has bradycardia (a very slow heart rhythm), which causes poor circulation and cannot be corrected by treatment with drugs. Pacemakers stimulate the heart muscle. This type of arrhythmia

is indicated by the occasional missing of the P-wave and the presence of a spike representing the stimulus from the pacemaker, followed by a wide QRS complex (Figure 5.4).

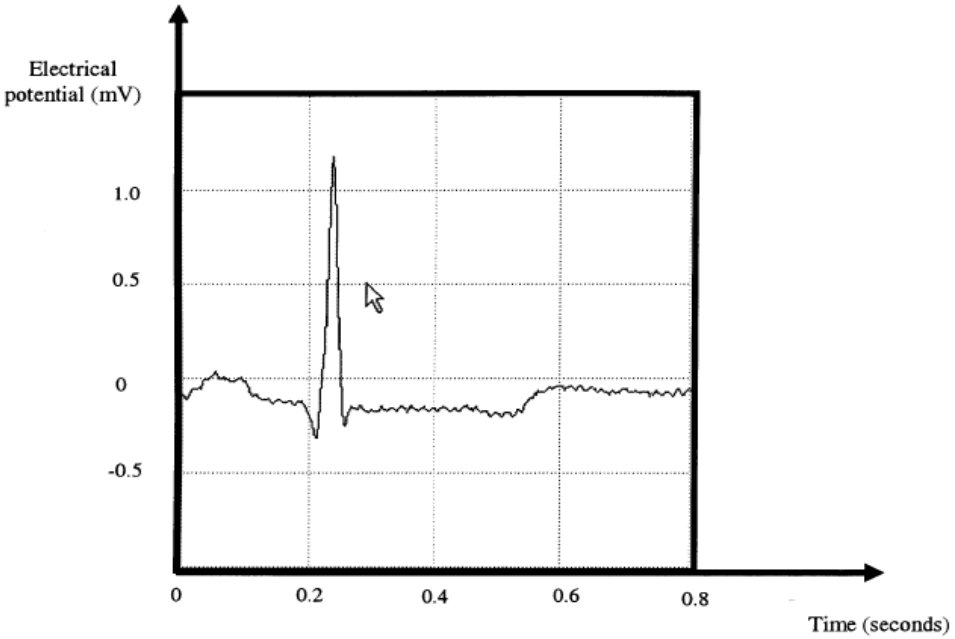


Figure 5.1: Normal Sinus Rhythm (N) Type (MIT-BIH Database, Record 100)

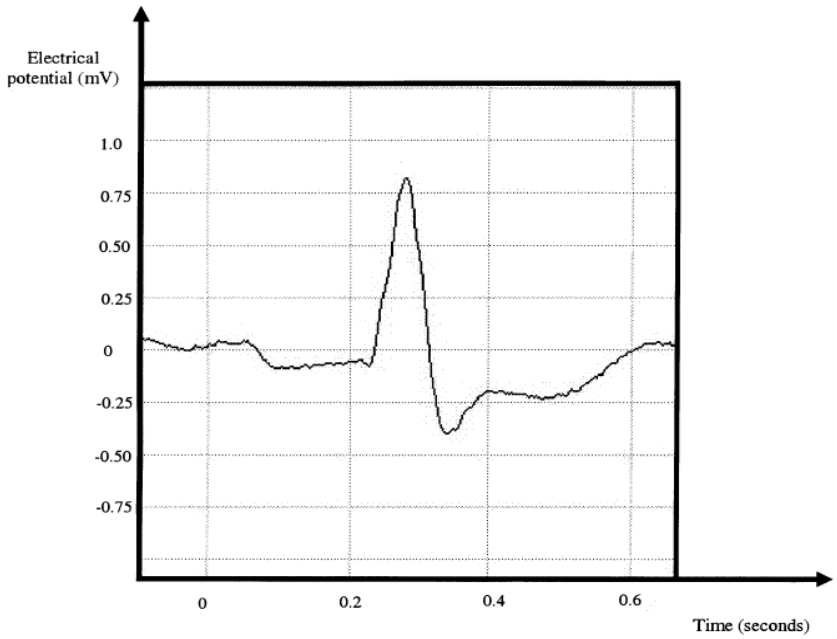


Figure 5.2: Left Bundle Branch Block (L) type (MIT-BIH Database, Record 109)

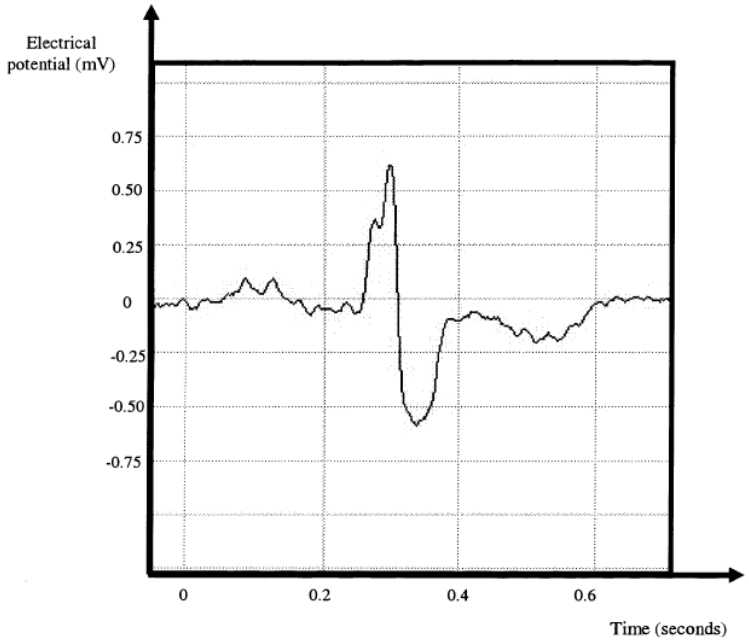


Figure 5.3: Right Bundle Branch Block (R) type (MIT-BIH Database, Record 118)

5- Premature Ventricular Contraction (V): this arrhythmia occurs when the heartbeats earlier than it should. This is because of the abnormal electrical activity of the ventricles which causes premature contraction of the lower chambers of heart, the ventricles. The premature contraction is followed by a pause as the heart's electrical

system “resets” itself. The contraction following the pause is usually more forceful than normal. With this type, the QRS complex is misshapen and prolonged representing ventricular contraction without earlier atrial stimulation (Figure 5.5).

6- Atrial Premature Beat (A): this arrhythmia is associated with early depolarization of atrium this type can be identified by a premature, small and distorted P-wave (Figure 5.6).

7- Aberrated Atrial Premature Beat (a): early depolarization of atria. These manifest itself as an abnormal P-wave (wide prolonged), narrow R-wave, and distorted QRS complex (Figure 5.7).

8- Nodal (junctional) Escape Beat (j): the cause of this arrhythmia is that the region around the AV node takes over as the focus of the depolarization; the rhythm is called “nodal” or ‘junctional’ escape. Figure 5.8 shows one beat cycle of this arrhythmia which has no Q- and S-waves. Also, the P-wave has an inverse polarity compared to that of the normal sinus rhythm (Figure 5.8).

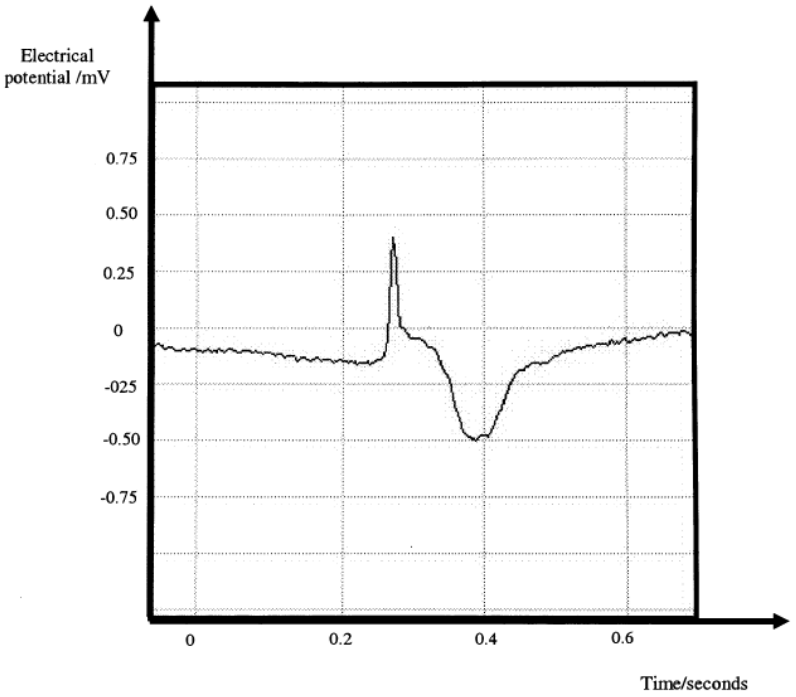


Figure 5.4: Beat Stimulated by an Artificial Pacemaker (‘Pace’) type (MIT-BEH Database, Record 104)

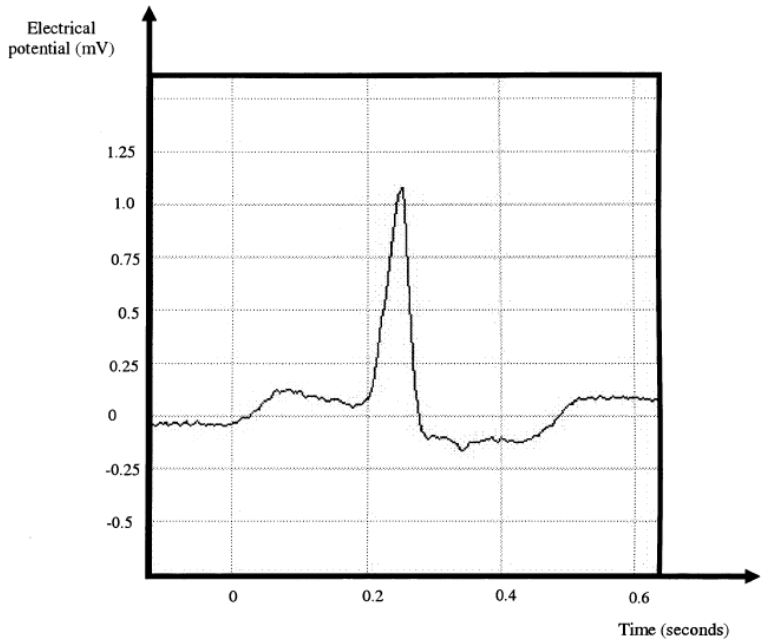


Figure 5.5: Premature Ventricular Contraction (V) type (MIT-BEH Database, Record 105)

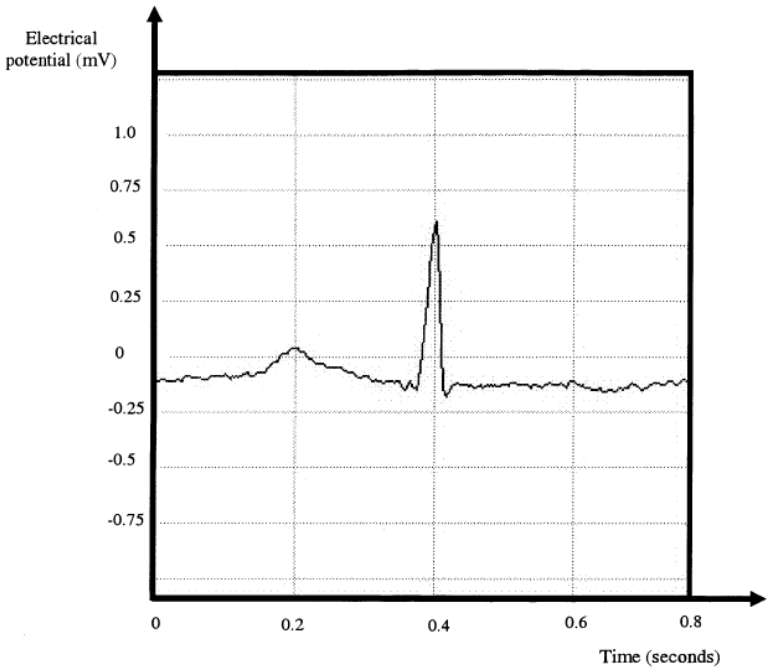


Figure 5.6: Atrial Premature Beat (A) type (MIT-BIH Database, Record 100)

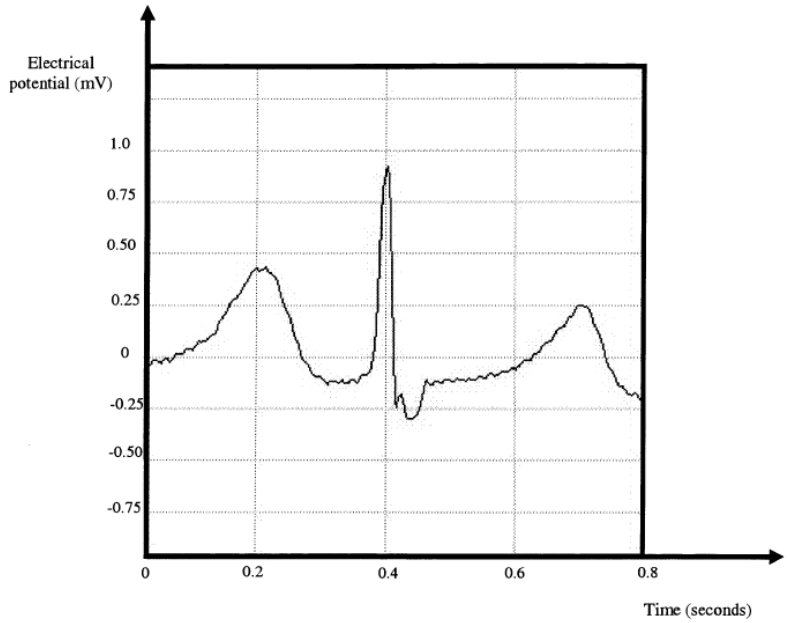


Figure 5.7: Aberrated Atrial Premature Beat (a) type (MIT-BIH Database, Record 105)

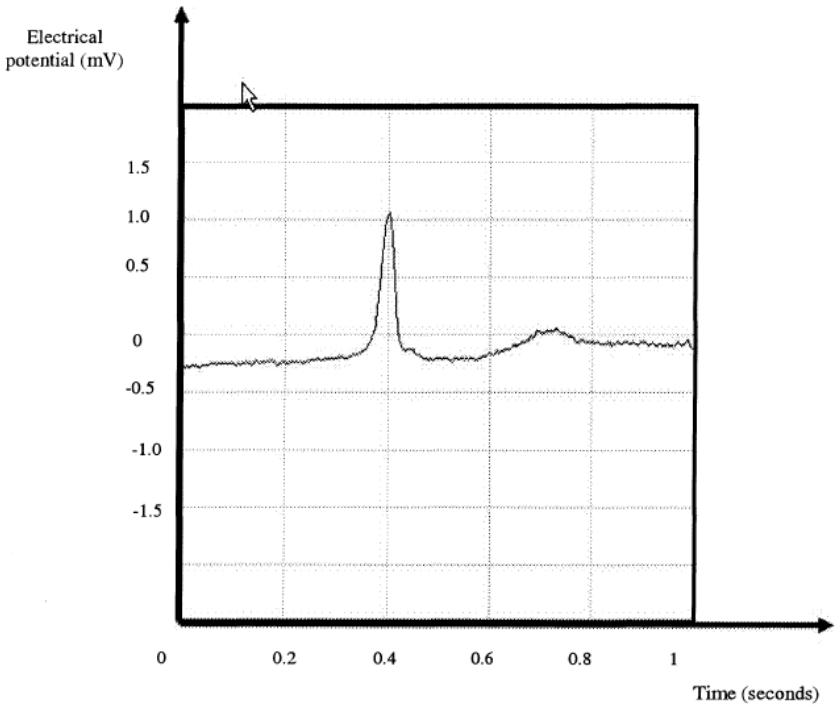


Figure 5.8: Nodal (junctional) Escape Beat (j) type (MIT-BIH Database, Record 201)

9- Ventricular Escape Beat (E): this most commonly occurs when the ventricle contracts without nodal stimulation. This is classically associated with complete heart blockage. The QRS complexes are wide whereas the P-waves are occasionally absent as demonstrated in Figure 5.9.

10- Fusion of paced and normal beats (f): this type of arrhythmia is a mixture of paced and normal beats. The P-waves have large amplitudes and are wide, and the QRS complexes are distorted, especially in the S-waves portion (Figure 5.10).

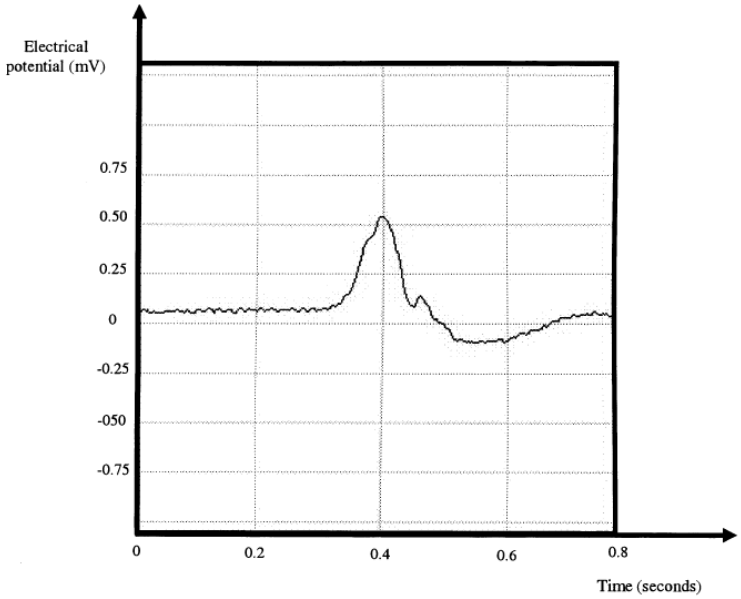


Figure 5.9: Ventricular Escape Beat (E) type (MIT-BIH Database, Record 207)

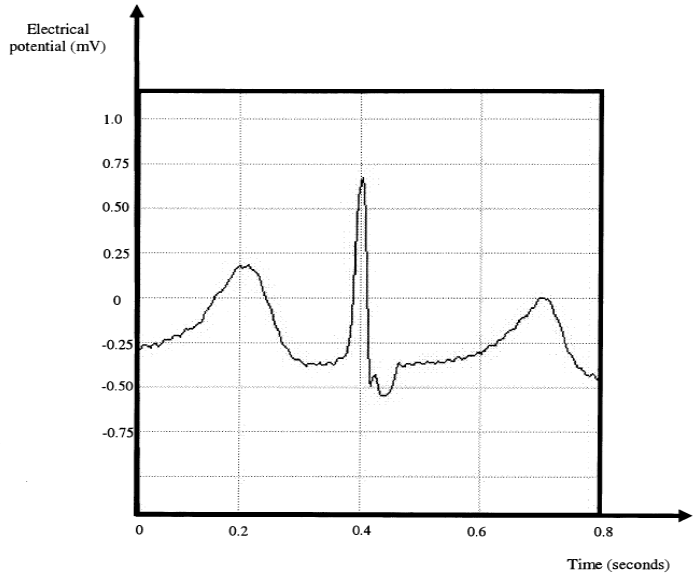


Figure 5.10: Fusion of Paced and Normal Beats (f) (MIT-BIH Database, record 113)

Table 5.1: Statistical overview of different beat types in the MIT−BIH Arrhythmia Database [46]

Record	Normal beat	LBBB	RBBB	Atrial premature beat	Abberated atrial premature beat	Nodal premature beat	Supraventricular premature beat	Ventricular premature beat	Fusion of ventricular or normal beat	Ventricular flutter wave	Atrial escape beat	Nodal escape beat	Ventricular escape beat	Paced rhythm	Fusion of paced or normal beat	Pause	Unclassified beat
100	2239	-	-	33	-	-	-	1	-	-	-	-	-	-	-	-	-
101	1860	-	-	3	-	-	-	-	-	-	-	-	-	-	-	-	2
102	99	-	-	-	-	-	-	4	-	-	-	-	-	2028	56	-	-
103	2082	-	-	2	-	-	-	-	-	-	-	-	-	-	-	-	-
104	163	-	-	-	-	-	-	2	-	-	-	-	-	1380	666	-	18
105	2526	-	-	-	-	-	-	41	-	-	-	-	-	-	-	-	5
106	1507	-	-	-	-	-	-	520	-	-	-	-	-	-	-	-	-
107	-	-	-	-	-	-	-	59	-	-	-	-	-	2078	-	-	-
108	1739	-	-	4	-	-	-	17	2	-	-	1	-	-	-	11	-
109	-	2492	-	-	-	-	-	38	2	-	-	-	-	-	-	-	-
111	-	2123	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-
112	2537	-	-	2	-	-	-	-	-	-	-	-	-	-	-	-	-
113	1789	-	-	-	6	-	-	-	-	-	-	-	-	-	-	-	-
114	1820	-	-	10	-	2	-	43	4	-	-	-	-	-	-	-	-
115	1953	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
116	2302	-	-	1	-	-	-	109	-	-	-	-	-	-	-	-	-
117	1534	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-
118	-	-	2166	96	-	-	-	16	-	-	-	-	-	-	-	10	-
119	1543	-	-	-	-	-	-	444	-	-	-	-	-	-	-	-	-
121	1861	-	-	1	-	-	-	1	-	-	-	-	-	-	-	-	-
122	2476	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
123	1515	-	-	-	-	-	-	3	-	-	-	-	-	-	-	-	-
124	-	-	1531	2	-	29	-	47	5	-	-	5	-	-	-	-	-
200	1743	-	-	30	-	-	-	826	2	-	-	-	-	-	-	-	-
201	1625	-	-	30	97	1	-	198	2	-	-	10	-	-	-	37	-
202	2061	-	-	36	19	-	-	19	1	-	-	-	-	-	-	-	-
203	2529	-	-	-	2	-	-	444	1	-	-	-	-	-	-	-	4
205	2571	-	-	3	-	-	-	71	11	-	-	-	-	-	-	-	-
207	-	1457	86	107	-	-	-	105	-	472	-	-	105	-	-	-	-
208	1586	-	-	-	-	-	2	992	373	-	-	-	-	-	-	-	2
209	2621	-	-	383	-	-	-	1	-	-	-	-	-	-	-	-	-
210	2423	-	-	-	22	-	-	194	10	-	-	-	1	-	-	-	-
212	923	-	1825	-	-	-	-	-	-	-	-	-	-	-	-	-	-
213	2641	-	-	25	3	-	-	220	362	-	-	-	-	-	-	-	-
214	-	2003	-	-	-	-	-	256	1	-	-	-	-	-	-	-	2
215	3195	-	-	3	-	-	-	164	1	-	-	-	-	-	-	-	-
217	244	-	-	-	-	-	-	162	-	-	-	-	-	1542	260	-	-
219	2082	-	-	7	-	-	-	64	1	-	-	-	-	-	-	133	-
220	1954	-	-	94	-	-	-	-	-	-	-	-	-	-	-	-	-
221	2031	-	-	-	-	-	-	396	-	-	-	-	-	-	-	-	-
222	2062	-	-	208	-	1	-	-	-	-	-	212	-	-	-	-	-
223	2029	-	-	72	1	-	-	473	14	-	16	-	-	-	-	-	-
228	1688	-	-	3	-	-	-	362	-	-	-	-	-	-	-	-	-
230	2255	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-
231	314	-	1254	1	-	-	-	2	-	-	-	-	-	-	-	2	-
232	-	-	397	1382	-	-	-	-	-	-	-	1	-	-	-	-	-
233	2230	-	-	7	-	-	-	831	11	-	-	-	-	-	-	-	-
234	2700	-	-	-	-	50	-	3	-	-	-	-	-	-	-	-	-

Examples of the above arrhythmias and normal ECGs were extracted from records100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 111, 112, 113, 114, 115, 116, 117,118, 119, 121, 122, 123, 124, 200, 201, 202, 203, 205, 207, 208, 209, 210, 212, 213,214, 215, 217, 219, 220, 221, 222, 223, 228, 230, 231, 232, 233, 234. Table 5.1 provides an overview of the different beat types in the MIT−BIH database.

There are mainly two different approaches to assess the performance of a machine learning algorithm in ECG domain: class based and subject-based. Class-based methods are applied based on the selection of various heartbeats and their class labels associated with the disease categories. The subject based methods, that are more widely studied in literature, are based on the use of AAMI standard (AAMI 1987) which breaks down the 15 heartbeat classes into 5 sub-classes, namely, non-ectopic (*N*), supraventricular ectopic (*S*), ventricular ectopic (*V*), fusion (*F*), and unknown (*Q*) [57][58][59][60][61][62]. Table 5.2 illustrates the 15 heartbeat classes and the 5 aforementioned AAMI classes together with the association between MIT-BIH arrhythmia annotation and AAMI standard.

Table 5.2: Different heartbeat classes provided by the MIT-BIH database and corresponding regrouping by AAMI standard

MIT-BIH class	Annotation	AAMI groups
Normal beat	N	*N*:beats not found in the classes S, V, F, and Q
Left bundle branch block beat	L	
Right bundle branch block beat	R	
Atrial escape beats	e	
Nodal (junctional) escape beat	j	
Atrial premature beats	A	*S*: supraventricular ectopic beats
Aberrated atrial premature beats	a	
Nodal (junctional) premature beats	J	
Non-conducted P-wave (blocked APB)	x	
Premature ventricular contraction	V	*V*: ventricular ectopic beats
Ventricular escape beat	E	
Ventricular flutter wave	!	
Fusion of ventricular and normal beat	F	*F*: Fusion of ventricular and normal beat
Fusion of paced and normal beat	f	*Q*:paced beats or unclassified beats
Unclassified beat	Q	

Chapter 6

PREPROCESSING

6.1 Introduction

In order to classify the ECG signal, we need to process the raw signal. This process can be generally divided into three main parts: preprocessing, feature extraction and classification. The preprocessing stage eliminates the noise and baseline drift from the raw ECG signal. The goal of feature extraction is to form distinctive properties and diagnostic information for each subject. For this purpose, selection of the best distinctive and relevant features is crucial.

6.2 Preprocessing of ECG Signals

ECG signal inherently contains various types of unwanted noise and artefact effects, such as baseline drift, noise of electrode contact, polarization noise, the internal amplifier noise, noise due to muscle movement, and motor artefacts. The movements of electrodes induced artefacts noise. Therefore, in order to make the ECG signal ready for feature extraction step, we must remove baseline wander and eliminate above noise.

We used the wavelet filtering to filter the ECG signal, since this technique is suitable for computing the R-peak locations without change of the shape or position of the original signal. According to the previous experimental knowledge, in order to optimize the signal filtering, we must consider these two criteria: the signal sampling frequency and the knowledge that most of the noises are located outside of the

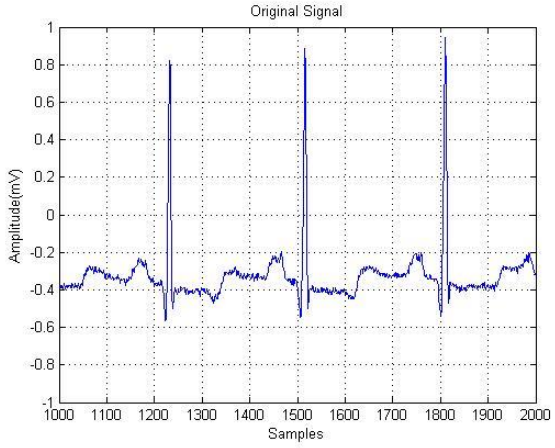
frequency interval between 1.5 Hz to 50 Hz [63]. For this purpose, we use a band pass filter which is constructed by a high pass filter with cut-off frequency 1.5 Hz. This filter eliminates baseline variations. The output of this filter is cascade with a low pass filter with cut-off frequency 50 Hz. This filter removes high frequency noise.

The scale and type of the mother function parameters are specific to each filter. Thus, the automatic compute of optimal scale for high pass filtering when the sampling frequency is 256 is equal to order 6. The optimal scale order for the low pass filtering is equal to order 2.

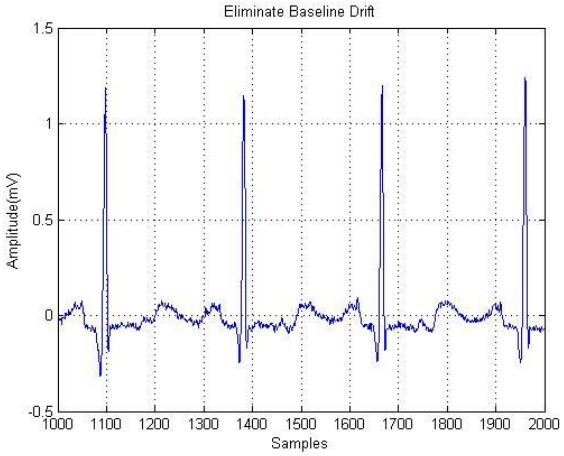
Finally, Z-score normalization is applied on all sampled values in order to convert them to a common scale with an average of zero and standard deviation of one by using the Equation (6.1).

$$Z = \frac{X - \mu}{\sigma} \tag{6.1}$$

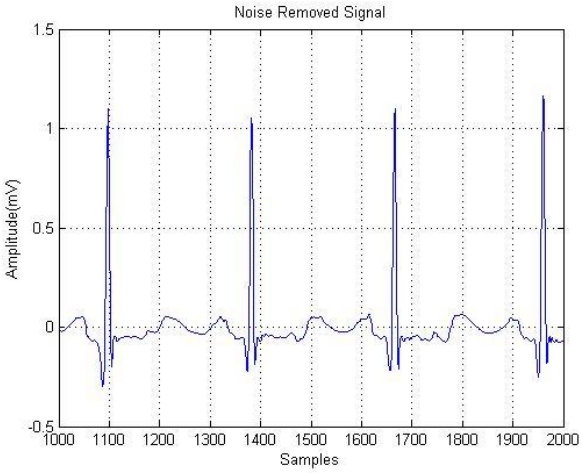
where X, μ and σ are signal, signal's mean and signal's standard deviation respectively.The results of above steps are shown in the figure 6.1.



(a)



(b)



(c)

Figure 6.1: Implementation Results of Preprocessing on Record [100] from MIT-BIH Arrhythmia Database: (a) Original Wave, (b) Eliminated Baseline (c) Noise Removal

6.3 QRS Detection

Each ECG cycle is consists of a P-wave which is corresponding to the atrial depolarization, a QRS complex which is corresponding to the ventricular depolarization and a T wave which is point to the rapid repolarization of the ventricles. A normal ECG signal and its time intervals are shown in Figure 6.2.

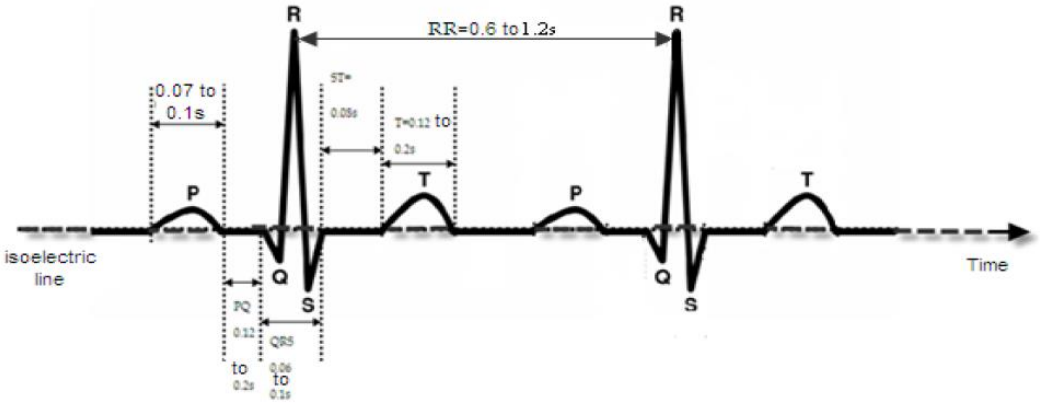


Figure 6.2: Standard Waves of a Normal Electrocardiogram [74]

Most of the clinically features which are useful for diagnostic the disease can be found in the time interval between components of ECG and the value of the signal amplitude. For example, the Q-T feature is used to recognition one dangerous disease, the Long Q-T Syndrome (LQTS), which is responsible of thousand deaths each year [64]. The shape of T wave is a critical factor and it is essential to identify it correctly since inverted T waves can be caused as an effect of a serious disease named coronary ischemia [65].

Designing an algorithm in order to extract the ECG features automatically is very hard since ECG signal has a time-variant behaviour. As a result of these signal properties, we face with multiple physiological constraints and the existence of noise.

In recent years, several algorithms have been proposed for detection those features. In [65] they introduced a method to extract wavelet features and used SVM for classification. In their purposed method, the classification is done without completely identify ECG components. Castro et al. introduced a method that used wavelet based features and classify various form of abnormal heartbeats [66]. Tadejko and Rakowski proposed an algorithm which is based on computational morphology [67]. Their main goal is the assessment of various automatic classifiers for detection of disorder in the ECG. In [68], authors proposed a method to extract feature from ECG based on a multi resolution wavelet transform. First, they remove noise from ECG signal by discarding the coefficient which caused noise. In next step, they detect QRS complexes and by using them the start and end of each wave part is determined. They assess proposed method on some records from MIT-BIH Arrhythmia Database.

In this thesis, we propose a method for recognition of time interval and amplitude of various wave parts of ECG. In the first stage of our approach, the R-peak is detected accurately. For this purpose we used wavelet. In the second stage, the other ECG components are identified by using a local search around the detected R-peak. We can summarize this approach:

- The location of the R-wave has been identified by using wavelet transform.
- Each R-R interval from ECG signal is segmented as follow:
    - Within an interval, finding the maximum and minimum of the wave which are corresponding to the Q and S waves
    - Since P-wave and T-wave are dependent to other factors; we must provide some deterministic points in order to find their location.

These points are including the end point of the S-wave or Soff, the start point of T-wave or Ton, and the start point of Q-wave or Qon.

6.4 R-Peaks Detection

The detection of R-peak is the first step of feature extraction. For this purpose, we used DWT due to its ability to recognize different locations of the waves accurately. Similarly to the preprocessing, we apply the same steps in order to compute the scale and choose the mother function. We have the QRS complex signal as an input which has the frequencies between 5Hz and 15Hz, so we select scale of order 4 and choose the Db4 mother wavelet. The Db4 wavelet is very popular for the detection and location of R peaks due to the strong similarity of its shape to the ECG signal.

By performing wavelet decomposition the input has been down sampled. Therefore, the amounts of unnecessary information are reduced but the components of QRS are not changed. In order to find the location of R-peak, first we choose the locations which their amplitudes are greater than 60% of the max value of the whole input signal. Since we remove the noise from the signal in the previous step, it is useful for R-peak detection.

Since we decompose the signal into 4^{th} level, the R-peak location in this modified signal is at least 0.25 of the R-peak location in the original signal. So in order to find the actual location of R-peal we must convert the founded positions by multiplying them with 4.

Another important point is that R-peak location in modified signal is not exactly on the original signal at a scale of 4. Position of the signal changes during the down

sampling, so we applied local search around the R-peaks which calculated in previous part. The interval of this search can be limited to a window of ± 20 samples.

6.5 P, Q and S Detection Algorithms

The accuracy of detecting R-peak completely affected on P, Q and S detection parts, since their location is determined relatively to R-peak. In the other hand, detect the location of R-peaks are corresponding to recognize the heart beat interval.

One of the most popular features in ECG signal processing is the R-R interval which can be computed by the following formula:

$$R\_R(i) = R(i+1) - R(i) \tag{6.2}$$

where R(i) and R(i + 1) are the indexes of the current and next R wave peak respectively.

6.5.1 S-wave Detection

The S-wave is located on the end of the QRS complex so in order to find its location we started from R-peak location plus 6 units because range of the shortest length of it is between 0.016 and 0.036 seconds. This range is corresponding to 6 and 13 samples. The stop point of search interval is related to the value of R-R interval. However the maximum length of the RS intervals is recorded is around 0.27 seconds where its R-R interval was 1.41 seconds [69].

6.5.2 Q-wave Detection

The Q-wave indicates the start point of the QRS complex section. It is reported that Q-wave peak location can be found in the range between 0.02 to 0.06 seconds from R-peak. In the other hand, this interval is equal to 8 and 22 samples. But this interval must be relevant to the value heart beat length. Therefore Q-R interval varies from one patient to another, for example a patient with a R-R equal to 235 can have a Q-R

interval equal to 19 and another one can have Q-R equal to 8 while has a R-R equal to 292. As a result, the range for search becomes larger for longer R-R interval. The process of Q-wave detection is illustrated in Figure 6.3.

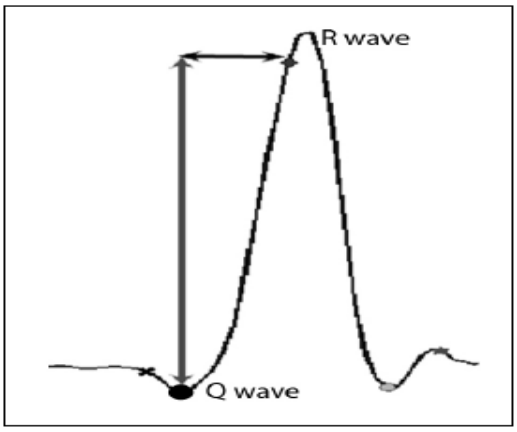


Figure 6.3: Q-wave Identification (150 samples in this case)

6.5.2.1 Q-wave Onset Detection

The start point of Q and Qon can be indicated as the points with the maximum amplitude near the negative peak location of the Q-wave. Therefore the search interval for finding the Qon started from Q-12 and continues till Q-5. Since some points before Qon may have larger amplitude, its index may need some corrections. For this purpose, it is necessary to:

- Compute the amplitude difference between Qon and Q by the following equation

$$level = y_{Qon}(i) - y_Q(i) \tag{6.3}$$

- if amplitude(P)-level>0.25 then threshold=0.90 else it sets to 0.87.
- Start searching from Qon in order to find the first value which

$$\text{Point\_ amplitude} < \text{thanlevel(i)} * \text{threshold.}$$

6.5.3 P-wave detection

Since P-wave can be located far or near from Q-wave, it is necessary for its interval to be relative to the R-R interval value. It is reported that duration of the P-R interval is between 0.09 and 0.19 seconds and this interval also depends on the R-R interval. This interval is equal to 19 and 38 samples. From the point of view of proportional, the limits are 14% to 22% of the respective RR range. One of the benefits of this approach is that we can detect P-waves with low amplitude, therefore according to the search area interval, we have two cases:

- Case 1: search_interval set to $0.81 * \text{R\_R(i)} - 7$ to $\text{Q(i)} - 18$. This interval works for most of the patients but have some problem with records numbers 111,215 and 218.
- Case 2: search_interval=$0.71 * \text{R\_R(i)} - 7$ to $\text{Q(i)} - 18$. This search interval solves the above problem but now when we can't find P-wave and the S-T segment is depressed, we must start searching the P-wave from the start point in its equation.

6.6 T-wave Detection

Finding T-wave in ECG signal is the most complicated task. Designing a procedure for detecting T-wave is difficult since it has a time variant behaviour. By checking the ECG waveform someone can see that the T- wave is located at the interval which has largest amplitude between S and the middle of the R-R interval. Therefore, the search interval for T, starts from S-wave and finishes at the middle point of the R-R interval.

6.6.1 T-wave Onset Detection

Another important point in analysing the ECG signal is the start point of T-wave or Ton, because it is used as a support point for determining the polarity of T which can

be positive, negative or flat. Existence of negative or flat T waves in ECG shows a serious disease, the cardiac ischemia. The searching area is start from the S-wave plus small offset till T-wave and Ton is a point with minimum value in this interval.

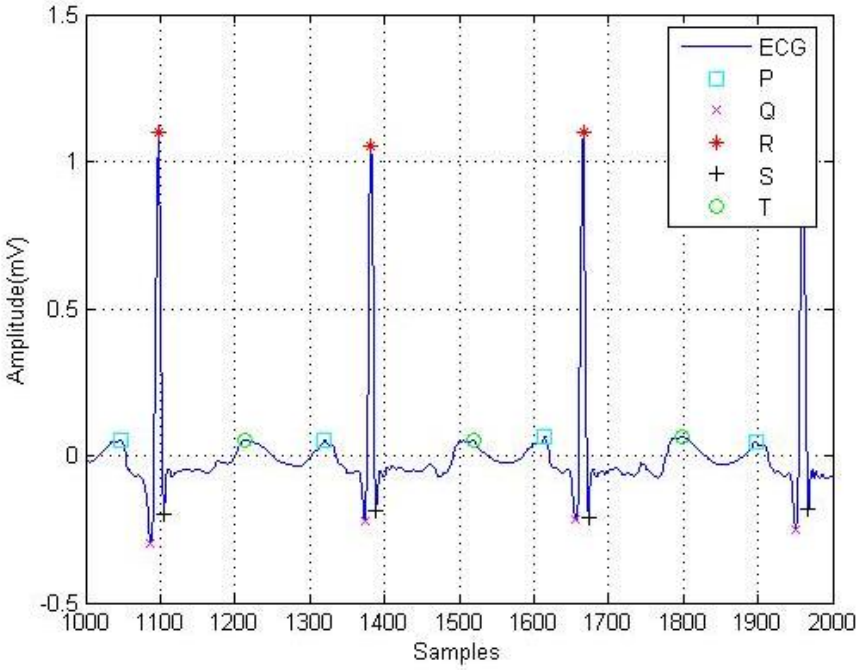
6.6.2 T-wave End Detection

Detection of the end point of T-wave, Toff, is another difficulty in this domain since there is still discussion between specialists about it. The best properties for detect Toff is finding the point which has the lowest amplitude after T within a limited range. For this purpose, it is necessary to make the signal smooth. We can do it by adding previous values of the signal to it. If point's amplitude be larger than the amplitude's of all the previous 3 samples, it can be consider as a Toff point. Table 6.1 summarized the search intervals used to find ECG components. It shows the indexes of the start and the end ranges of the search. The output of this local search algorithm is the index which satisfies corresponding constraint type [69].

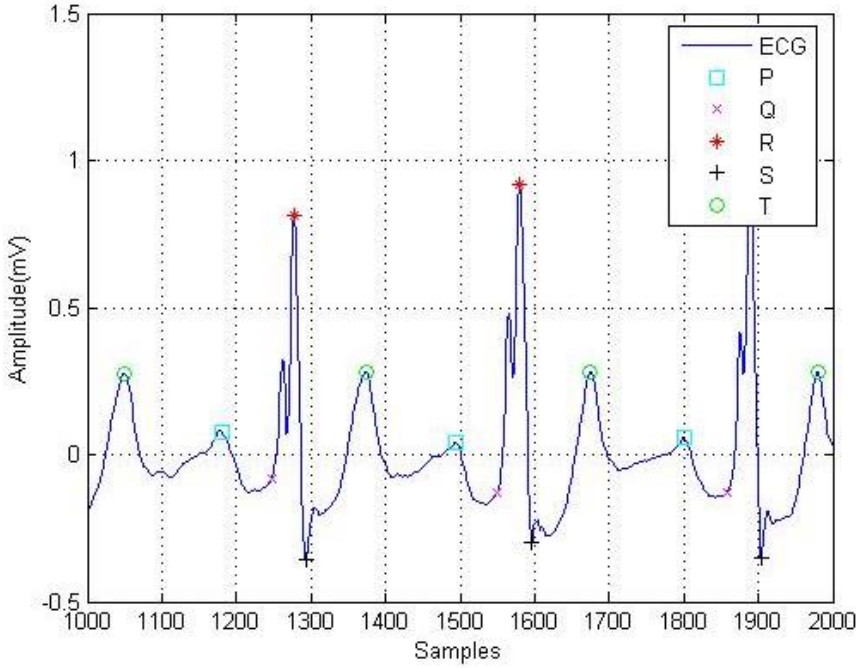
Table 6.1: Search Intervals [69]

Wave	Beginning	End	Type
P	$0.71 * \text{R\_R(i)} - 7$	$\text{Qon(i)} - 12$	max
Q	$\text{R\_R(i)} - 25$	$\text{R\_R(i)} - 7$	min
Qon	$\text{Q(i)} - 12$	$\text{Q(i)} - 5$	max
S	$\text{R(i)} + 6$	$\text{R\_R(i)}/5 - 10$	min
Ton	$0.7 * ((\text{T(i)} - \text{S(i)})$	$\text{T(i)} - 10$	min
T	$\text{S(i)} + 15$	$\text{R\_R(i)}/2$	max

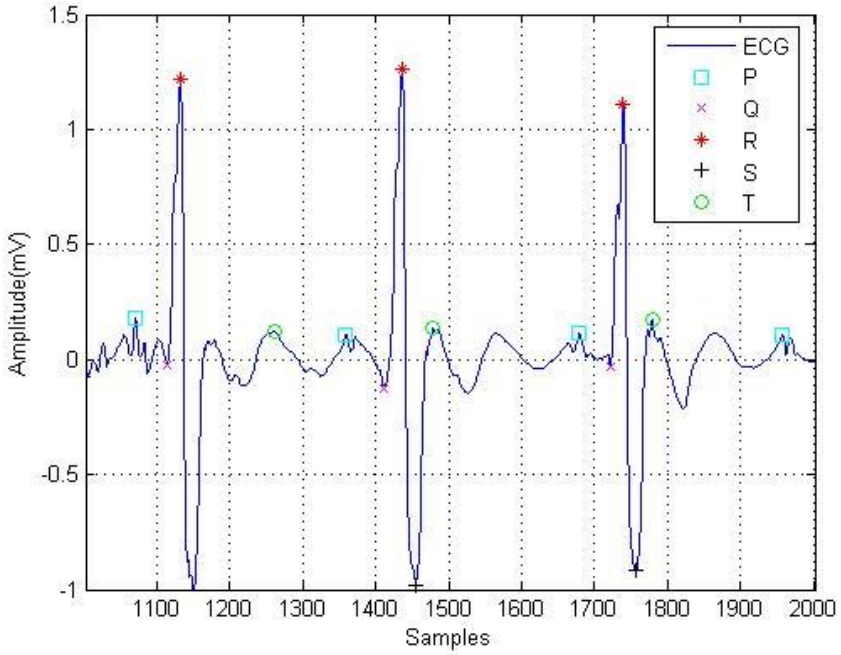
The results of implementation of PQRS detection algorithm for some ECG records from MIT-BIH database are shown in the Figure 6.4.


ECG
P
Q
R
S
T
Amplitude(mV)
Samples
1.5
1
0.5
0
-0.5
1000
1100
1200
1300
1400
1500
1600
1700
1800
1900
2000

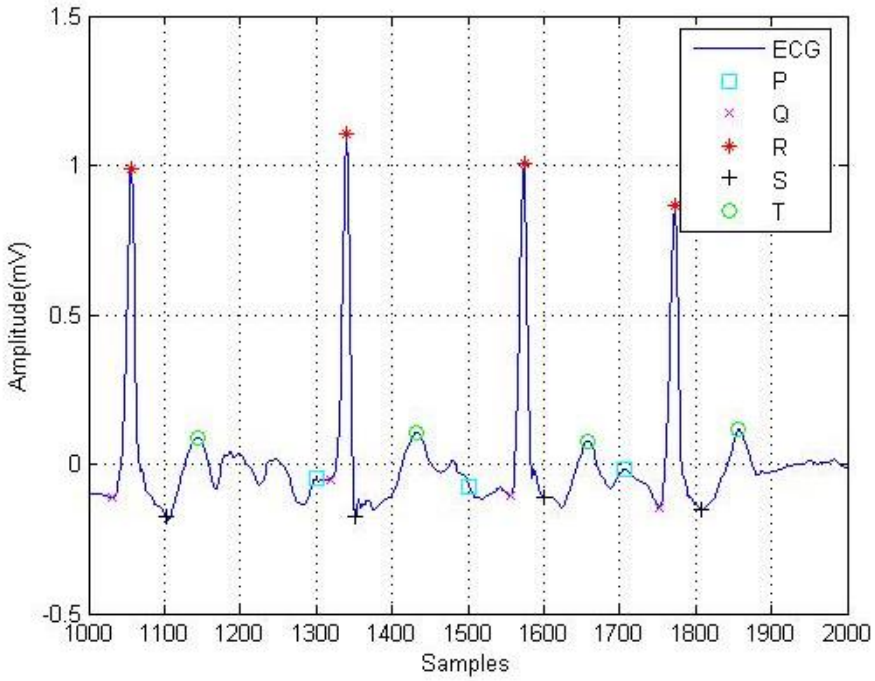
(a)


ECG
P
Q
R
S
T
Amplitude(mV)
Samples
1.5
1
0.5
0
-0.5
1000
1100
1200
1300
1400
1500
1600
1700
1800
1900
2000

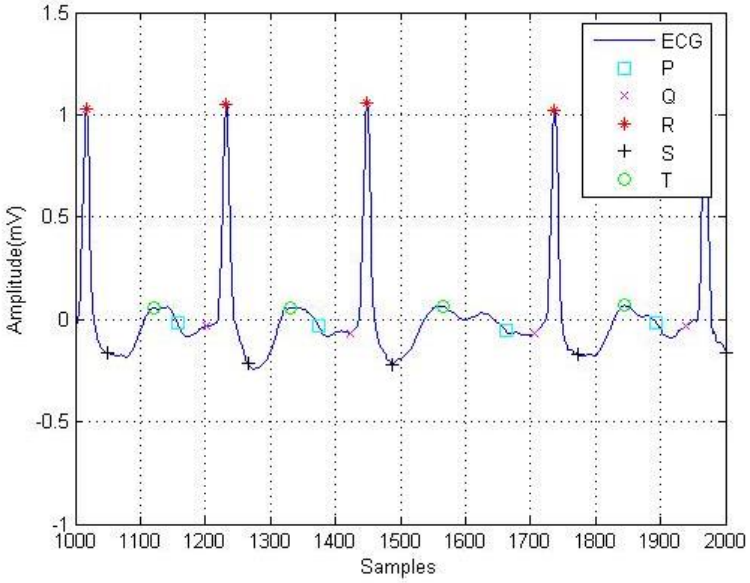
(b)


ECG
P
Q
R
S
T
Amplitude(mV)
Samples
1.5
1
0.5
0
-0.5
-1
1100
1200
1300
1400
1500
1600
1700
1800
1900
2000

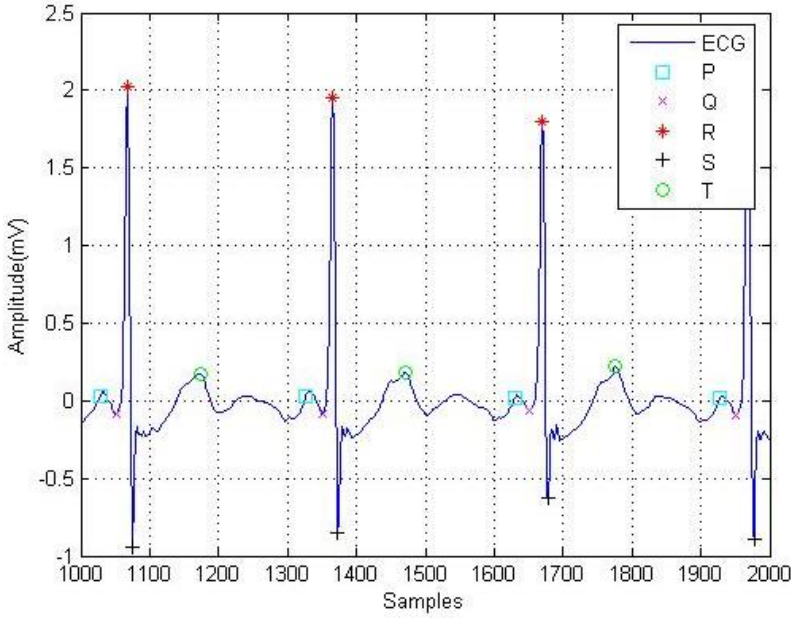
(c)


ECG
P
Q
R
S
T
Amplitude(mV)
Samples
1.5
1
0.5
0
-0.5
1000
1100
1200
1300
1400
1500
1600
1700
1800
1900
2000

(d)



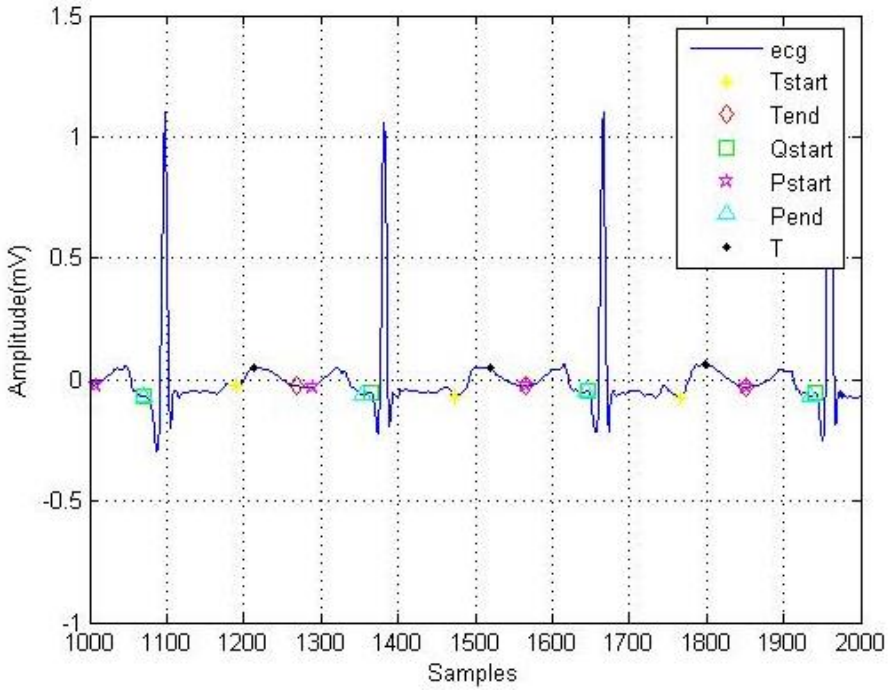
(e)



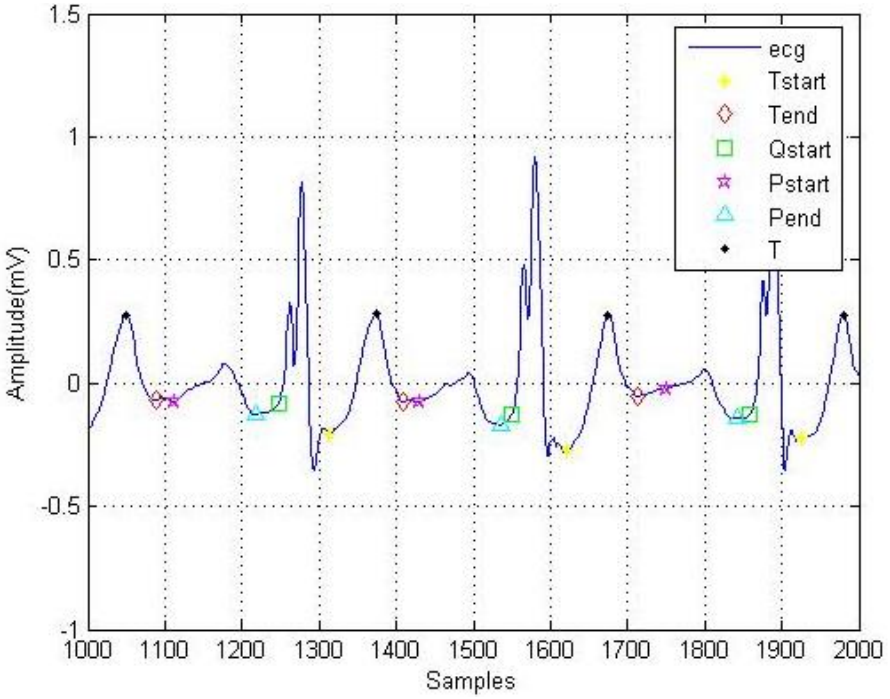
(f)

Figure 6.4: PQRST detected in (a) record [100], (b) record [111], (c) record [118], (d) record [201], (e) record [210], (f) record [220] from ECG MIT-BIH arrhythmia database

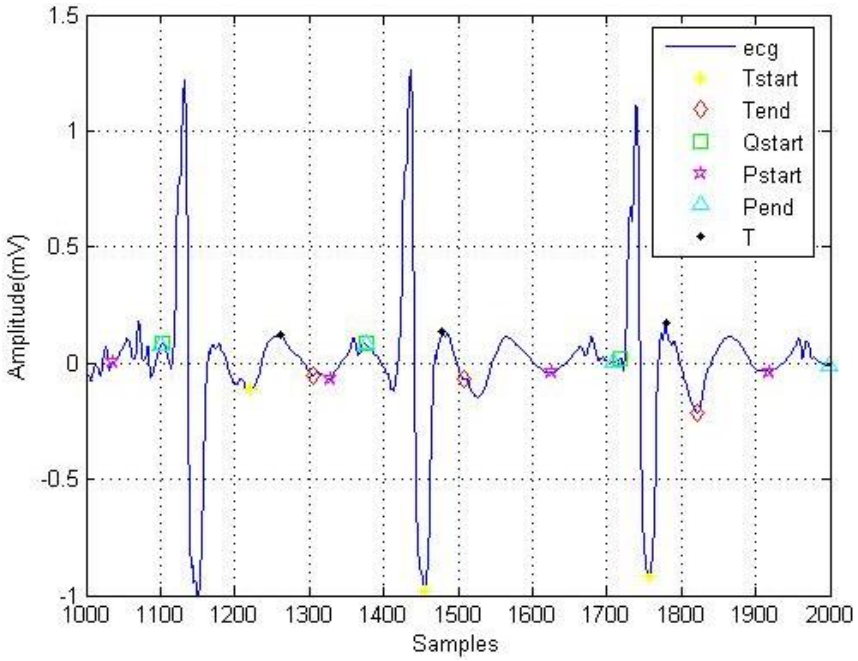
The results of implementation of $Onset\ and\ Offset$ detection algorithm for previous ECG records are shown in Figure 6.5.


ecg
Tstart
Tend
Qstart
Pstart
Pend
T
Amplitude(mV)
Samples
1.5
1
0.5
0
-0.5
-1
1000
1100
1200
1300
1400
1500
1600
1700
1800
1900
2000

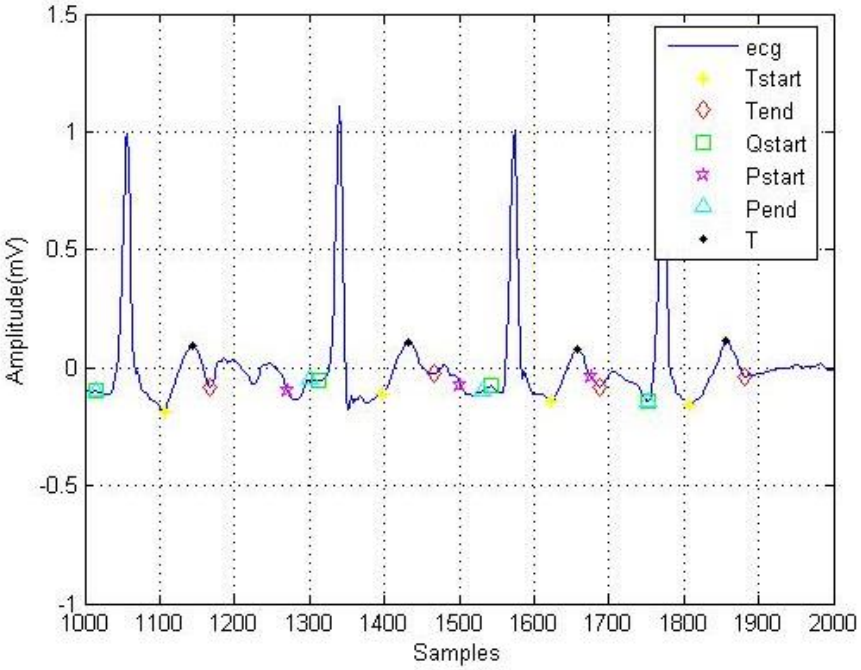
(a)


ecg
Tstart
Tend
Qstart
Pstart
Pend
T
Amplitude(mV)
Samples
1.5
1
0.5
0
-0.5
-1
1000
1100
1200
1300
1400
1500
1600
1700
1800
1900
2000

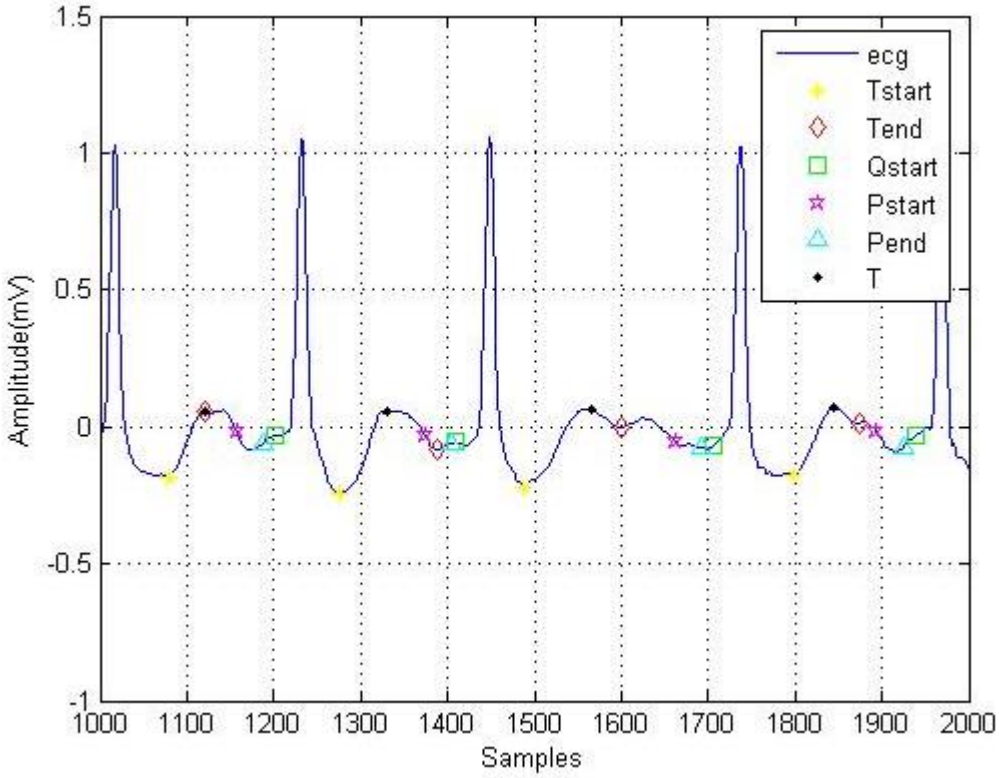
(b)


ecg
Tstart
Tend
Qstart
Pstart
Pend
T
Amplitude(mV)
Samples
1.5
1
0.5
0
-0.5
-1
1000
1100
1200
1300
1400
1500
1600
1700
1800
1900
2000

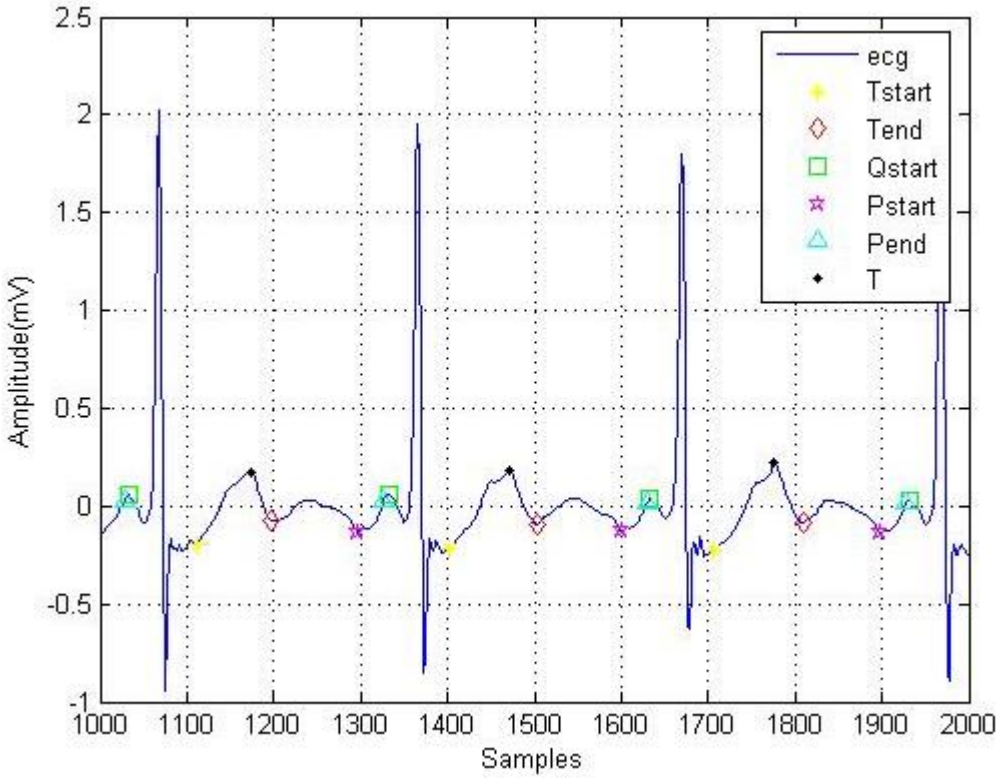
(c)


ecg
Tstart
Tend
Qstart
Pstart
Pend
T
Amplitude(mV)
Samples
1.5
1
0.5
0
-0.5
-1
1000
1100
1200
1300
1400
1500
1600
1700
1800
1900
2000

(d)



(e)



(f)

Figure 6.5: Onset-Offset of waves detected in (a) record [100], (b) record [111], (c) record [118], (d) record [201], (e) record [210], (f) record [220] from ECG MIT-BIH arrhythmia database

Chapter 7

PROPOSED METHODS

7.1 Introduction

Automatic classification of the arrhythmias using machine-learning technologies can bring various benefits. In this thesis, novel and high-performance approaches based on deep learning techniques are proposed for the automatic classification of electrocardiogram (ECG) signals. In this thesis, two fully automatic systems have been presented which shown to have high efficiency and low computational cost which achieve superior classification accuracy compared to the other methods.

7.2 Proposed 3-Level Feature Fusion Approach

In this section, we propose a new ECG classification system consisting of three subsystems for the classification of a heartbeat into five different AAMI class types by exploiting multi-stage learned features from a trained CNN and hand-crafted features extracted using well-known algorithms [70]. Hand-crafted feature sets constitute the P-QRS-T temporal features, morphological wavelet features and various statistical features. The overall schematic of the proposed method is illustrated in Figure 7.1.

This is motivated by our observation that the distributions of the measures of the diseased group are often skewed, heavy-tailed, or multimodal, whose features cannot be well captured by using a single feature extractor. It turns out that the performance of utilizing several feature descriptors is better than the single one. Secondly, we

combine commonly used hand-crafted feature descriptors on ECGs with learned features and use decision-level fusion method to identify the heartbeat type.

Previous studies in multimodal classification systems have shown that information fusion in both feature-level and score-level can improve the classification performance [71]. Considering the large intra-class similarity as well as many inter-class variability between different ECG signal class types, the feature-level fusion is performed over the hand-crafted local descriptors while score-level fusion is applied on the global learned features. Finally, the decisions obtained from three subsystems are consolidated by a decision-level fusion approach to predict the heartbeat classes.

In the first subsystem, the temporal features and statistical features extracted for each ECG beat signal are combined by the feature-level fusion method. The resulting feature vector is fed into a SVM-based classifier to predict the input heartbeat's class-label. As explained in the previous section, signal characteristics such as heart rate, ST-Interval, duration of QRS complex, TT-Interval, PR-Interval, QT-interval, PP-Interval and RR-Interval between consecutive and RR interval averaged over the last ten beats are considered as the temporal features, while features such as mean, variance, sum, root mean square, mean absolute deviation, skewness (γ_3), kurtosis (γ_4) and the 5-th moment of the input ECG signal are constituting the statistical features. These features are first normalized by the z-score method and then concatenated in order to perform feature-level fusion of P-QRS-T-based and statistical features.

The second subsystem uses DWT in order to extract morphological features. For this purpose, the Mayer's wavelet is selected and decomposed the ECG signal up to

fourth-level. Afterwards, morphological features include the maximum and minimum values of detail coefficients and approximation coefficients of the decomposed signal are extracted. Also, the amplitude values of P-wave, QRS-wave and T-wave considered as part of the morphological features. Consequently, the size of morphological features is reduced by using PCA for both computational efficiency and improved classification performance. The first 10% of the principal components associated with the morphological features is selected and submitted to a SVM-based classifier with RBF kernel to predict the class label of the input heartbeat.

In order to employ the strength of CNNs as automatic feature extractor, the proposed system adds a third subsystem, that is not used in other CNN-based classifier architectures published so far, which utilizes cross-layer heterogeneity features within the decision making process. In this approach, the cost-free mid-level features extracted from intermediate layers of a CNN are combined to the output layer features by score-level fusion method. Resultantly, it enhances the discriminative power of the proposed classifier system in classifying different patterns of varying complexities. This score-level fusion process is explained in detail by Algorithm 1. This algorithm can be expanded for more than two feature sets. In order to find the optimal subset, different combinations of layers' features are investigated by using a greedy approach. The early layers are ignored by this greedy heuristic since their learned features are too simple and general, while the intermediate layers are chosen for fusion with the last layer. At the final stage of the proposed system, decision-level fusion is performed on the outputs of the three individual classifiers to determine the input heartbeat sample's class-label. This decision-level fusion is implemented by the majority-voting technique. In case of a tie, one of the subsystem's results will randomly be selected and considered as the final decision.

7.2.1 Handcrafted Feature Descriptors

The choice of features to classify the ECG signal is crucial. Existing features can be roughly divided into 2 categories: the hand-crafted and the learned ones. By hand-crafted features we understand those which are extracted from separate signals according to a certain manually predefined algorithm based on the expert knowledge. Contrary to hand-crafted features, the learned ones are derived from a signal dataset by a training procedure in order to full fill a certain task (e.g. ECG classification). Convolutional Neural Networks (CNNs) [43] are examples of deep neural networks which can be used to extract learned features.

7.2.1.1 Temporal Features

Temporal features are among the most relevant features that have already been used in many applications. They include heart rate and interval-type features such as ST-Interval, the duration of QRS-complex, TT-Interval, PR-Interval, QT-Interval and PP- Interval. Figure 7.2 shows a typical ECG cycle and its temporal features where the RR-Interval is also a temporal feature between consecutive heartbeats. Its discrete and mean values averaged over a number of heartbeats are included within the extracted temporal features.

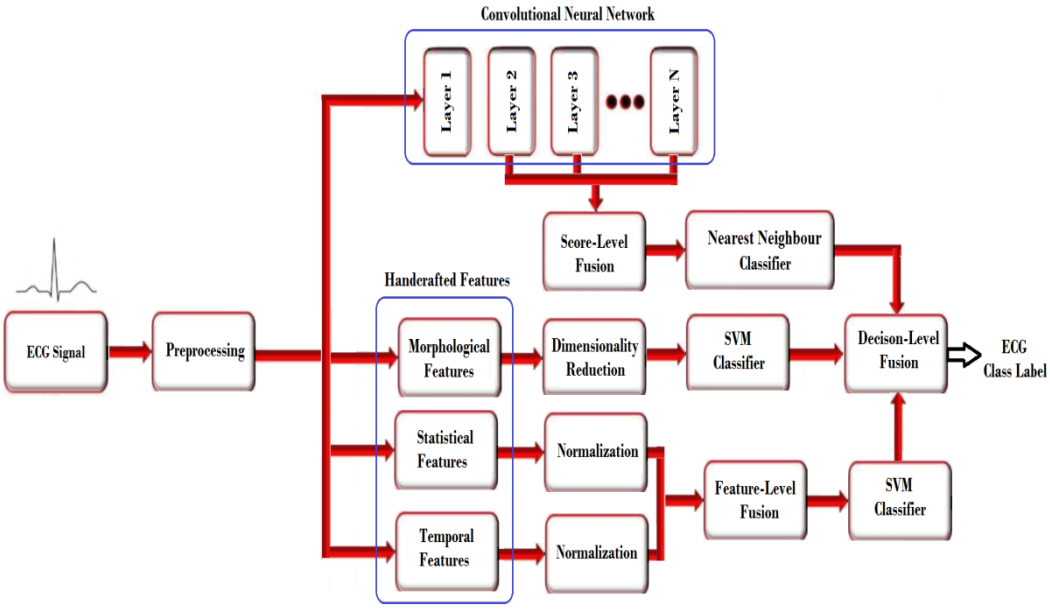


Figure 7.1: The overall schematic of the proposed method [70]

Morphological features: The shape-based characteristics, called the morphological features, of ECG signals can be computed through using the detail and approximation coefficients from a wavelet transformation. For this purpose, the finite impulse response (FIR) approximation of Mayer's wavelet is selected and performed on input ECG signals to decompose it, up to the fourth-level. Consequently, maximum and minimum values of detail and approximation coefficients of the decomposed signal together with the amplitudes of P-wave, QRS-wave and T-wave constitute the vector of morphological features.

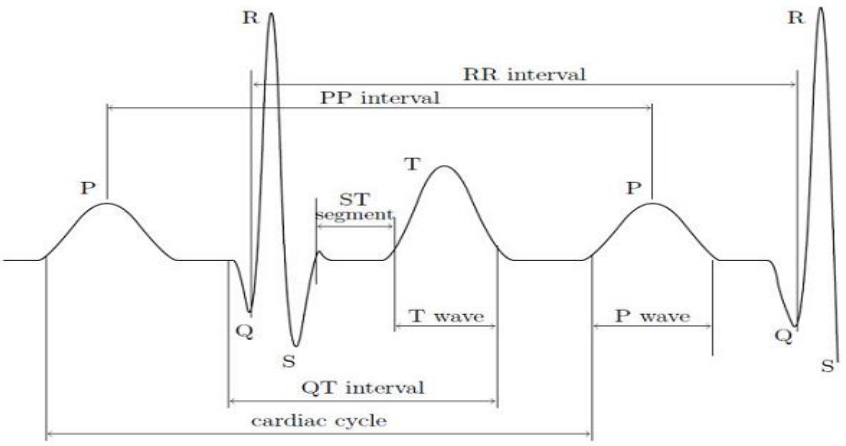


Figure 7.2: A Normal ECG Signal and It's Temporal Features

Statistical features: Statistical features which exhibit the overall variational characteristics of the signal. Statistical features complement the hand-crafted ones. Common statistical features include mean, variance, sum, root mean square, mean absolute deviation and 5-Th moment of the original signal. These are extracted to delineate the ECG signal for an accurate classification. The statistical features extracted from the given input ECG signal are as follows:

Mean: It is the average of a set of values and defined as the ratio between the summation of these values and the number of elements in the set.

$$\bar{X} = \frac{\sum X}{N} \tag{7.1}$$

The normalized 5^{th} central moment:

$$\frac{\bar{X}_5}{\sigma^5} = \frac{E\left|(X-\bar{X})^5\right|}{\sigma^5} \tag{7.2}$$

Standard Deviation: The amount of dispersion or variation of a given dataset is quantified by using this measure.

$$\sigma = \sqrt{\frac{\sum|x-\bar{X}|^2}{N}} \tag{7.3}$$

Variance: Let us consider a set of random numbers X and its mean $\bar{X}$. Variance measures how far X are spread out from$\bar{X}$. Also, it is computed from the squared deviation of X from it's $\bar{X}$.

$$VAR(X) = \frac{\sum(x - \bar{X})^2}{N - 1} \tag{7.4}$$

Additionally, skewness and kurtosis are two important higher order moments that give characteristic information on the shape of the probability distribution of a set of observations. In this respect, they are of the statistical features used in the proposed methods. Definitions of skewness (γ_3) and kurtosis (γ_4) are as follows:

$$\gamma_3 = \frac{E\{[x - E(x)]^3\}}{(E\{[x - E(x)]^2\})^{3/2}} \tag{7.5}$$

$$\gamma_4 = \frac{\mathrm{E}\{[\mathrm{x} - \mathrm{E}(\mathrm{x})]^4\}}{(\mathrm{E}\{[\mathrm{x} - \mathrm{E}(\mathrm{x})]^2\})^2} - 3 \tag{7.6}$$

where x denotes the observed random variable and E denotes its expected value.

7.2.2 Learned Feature Descriptors

Convolutional neural networks can be used as automatic feature extractors such that the learned features provided by CNNs can be fed directly to classifiers like SVMs to predict the output class labels. Conventional implementations that use CNNs for classification use only the features at the output layer. However, mid-level features at intermediate layers of a CNN can also be discriminative for classifying different patterns with varying complexities. From computational point of view, these mid-level features are already computed when the system is trained to extract high-level features, and hence, their usage does not bring any extra computational burden.

Recent studies [71][72] demonstrated the usage of learned features extracted from different layers of CNN is useful for the improvement of the classifier performance. It is also shown that the samples with lower complexity are better discriminated by

using features of the early layers while the complex patterns are well represented by features of the late layers. Therefore, as proposed in this research work, multi-stage fusion of features is expected to improve the accuracy of the implemented classifier system.

There are two ways in using the pre-computed mid-level features of CNN: feature-level fusion and score-level fusion. In feature-level fusion approach, features from different layers are concatenated to create a unified feature vector which is fed into a classifier for a decision associated with the test input. One of the common problems with feature-level fusion is the large size of the feature vectors. In CNNs, the size of the learned features in intermediate layers can be very large and combining these features simply by concatenation may cause the curse of dimensionality problem. Dimensionality reduction methods such as PCA or Discrete Cosine Transform (DCT) can be used to overcome this problem with a significant computational cost.

Another approach for overcoming the dimensionality problem is score-level fusion. More precisely, as illustrated in Algorithm 1, two different layers' feature sets FS^{m} and FS^n are considered for the score-level fusion. These feature sets are extracted for all the training set samples and stored in $F_m^{X_{tr}}$ and $F_n^{X_{tr}}$, respectively. Afterwards, in each FS^{m} and FS^n feature space, the distance between a test sample and all of the training set samples are computed and stored in the test sample's score-vectors, namely, $\boldsymbol{score^m}$ and $\boldsymbol{score^n}$. These score vectors should be normalized before being used in score-level fusion. For this purpose, Min-Max normalization method [71] is used as follows:

$$\mathrm{S}_{\mathrm{i}}' = \frac{\mathrm{S}_{\mathrm{i}} - \mathrm{Min(S)}}{\mathrm{Max(S)} - \mathrm{Min(S)}} \tag{7.7}$$

where S_i stands for i[th] entry in the original score vector, $Max(\boldsymbol{S})$ and $Min(\boldsymbol{S})$ are the maximum and minimum values of the original score vector $\boldsymbol{S}$ respectively and S_i' indicates the i[th] entry of the normalized score. Finally, the normalized score vectors are used by one of the different fusion methods such as Sum-rule or Product-rule [71]. This fused score vector is used to make the final classification decision through using a minimum distance-based classifier algorithm.

Algorithm-1 Score-level fusion of two feature sets for ECG classification

Input:

Trainset samples $X_{train} = \{(X_{tr}^i, L_{tr}^i) \mid X_{tr}^i: beat^i's\ signal, L_{tr}^i: beat^{i'}s\ class\ label, i = 1, \dots, N_{train}\}$

Testset samples $X_{test} = \{(X_{te}^i, L_{te}^i) \mid X_{te}^i: beat^i's\ signal, L_{te}^i: beat^{i'}s\ class\ label, i = 1, \dots, N_{test}\}$

Output:

Beat_Class_Labels

For i= 1 To N_{train}

Compute FS^m and FS^n

$F_m{}^{X_{tr}{}^i} = \langle feat\_m_1^i, \dots, feat\_m_{Sm}^i \rangle, Sm: size(Feat^m)$

$F_n{}^{X_{tr}{}^i} = \langle feat\_n_1^i, \dots, feat\_n_{Sn}^i \rangle, Sn: size(Feat^n)$

End For

For j= 1 To N_{test}

$F_m^{X\_te^j} = \langle feat\_m_1^j, \dots, feat\_m_{Sm}^j \rangle$

$F_n^{X\_te^j} = \langle feat\_n_1^j, \dots, feat\_n_{Sn}^j \rangle$

For i = 1 To N_{train}

$score_i^m$= *Compute_Distance* ($F_m{}^{X_{tr}{}^i}$, $F_m{}^{X_{te}{}^j}$)

$score_i^n$= *Compute_Distance* ($F_n{}^{X_{tr}{}^i}$, $F_n{}^{X_{te}{}^j}$)

End For

For i = 1 To N_{train}

$score_i^m$ is normalized according to Eq. (7.7)

$score_i^n$ is normalized according to Eq. (7.7)

End For

$\boldsymbol{fusion}$ = *Score_Level_Fusion*($\boldsymbol{score^m}, \boldsymbol{score^n}$)

$minIndex$=*Find_Min_Index*($\boldsymbol{fusion}$)

$Beat\_Class\_Labels_i = L_{tr}^{minindex}$

End For

7.3 Proposed DAG-CNN Model

In this section, DAG-CNN architecture is proposed to improve the discrimination capability of a deep neural network by allowing its layers to share their learned features and work collaboratively for classification [73]. The proposed multi-scale CNN topology applies learned features with different level of complexity in order to predict the output label with high precision.

This is motivated by the fact that hand-crafted features need a lot of domain expertise, human intervention only capable of what they're designed for, are computationally intensive due to high dimensions. On the other hand, the DAG-CNN tries to learn high-level features from data in an incremental manner. This eliminates the need of domain expertise and hard core feature extraction.

CNNs can be used as automatic feature extractors and the learned features can be fed to classifiers such as, SVMs or NNs to predict the output labels. Mid-level features at intermediate layers of a CNN can be discriminative for classifying different patterns with varying complexities. However, in CNN architectures used in literature so far, these cross-layer heterogeneity features are ignored. It is obvious that these mid-level features are already computed when the system is trained to extract high-level features, and hence, their usage does not bring any extra computational burden within our proposed model. Instead of performing feature level fusion and feeding the results to a classifier, we propose a multi-scale system by using a CNN with directed acyclic graph topology. Our proposed model can automatically learn different level of features, combine them and predict the output label.

One of the common problems with feature level fusion is the size of the feature vectors. In CNNs, the size of the learned features in intermediate layers can be very large and combining these features may cause the curse of dimensionality problem. To overcome this problem, we compute marginal activations by performing average pooling on the learned features of some layers which are used for feature level fusion.

The directed acyclic graph (DAG) networks can represent more complex network architectures compared to ones consisting of a linear chain of layers. The main advantage of DAG-structured networks is that their forward layers can have multiple input parameters from several backward layers. This way, they can achieve different levels of signal representations.

Very deep CNNs which are proposed by the researchers in the recently published work can be found in [74][75][76]. A fundamental feature of this deep architecture is the use of connections between their layers, called “skip connection”, that is similar to DAG-CNNs main idea, and it is shown that these skip connections can improve the accuracy of the classification tasks significantly.

DAG-CNN was proposed by Yang and Ramanan [77] to learn a set of multi-scale image features that are successfully used for classification of three standard scene benchmarks. They showed that the multi-scale model can be implemented as a DAG-structured feed forward CNN. By this approach, it is possible to use an end–to-end gradient-based learning for automatically extracting multi-scale features using generalized back propagation algorithm over the layers that have more than one input. In fact, all the required equations for training the network are standard CNN

equations except for the Add and ReLU layers since they have multiple inputs or outputs. Let us consider the i[th] ReLU layer in Figure 7.3, Let α_i be its input, $\beta_i^{(j)}$ be the output for its j[th] output branch (its j[th] child in the DAG), and assume that z is the final output of the softmax layer. The gradient of z with respect to the input of the i[th] ReLU layer can be computed as Equation (7.8):

$$\frac{\partial_z}{\partial_{\alpha_i}} = \sum_{j=1}^{C} \frac{\partial_z}{\alpha\beta_i^{(j)}} \frac{\alpha\beta_i^{(j)}}{\partial\alpha_i} \tag{7.8}$$

where C is the number of output edge of the i[th] ReLU.

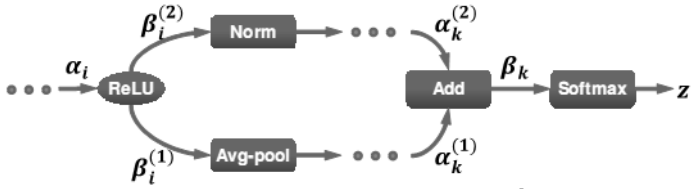


Figure 7.3: Visualization of the Parameter Setup at i^{th} ReLU and k^{th}ADD [77]

For the Add layer, let $\beta_k = g(\alpha_k^{(1)}, \cdots, \alpha_k^{(N)})$ represents the output of an Add layer with multiple inputs. We can compute the gradient along the layer by applying the chain rule as Equation (7.9):

$$\frac{\partial_z}{\partial\alpha_i} = \frac{\partial_z}{\partial\beta_k}\frac{\partial\beta_k}{\partial\alpha_i} = \frac{\partial_z}{\partial\beta_k}\sum_{j=1}^{C}\frac{\partial\beta_k}{\partial\alpha_k^{(j)}}\frac{\partial\alpha_k^{(j)}}{\partial\alpha_i} \tag{7.9}$$

In the convolutional layers (layer L1, L4, L7 and L10 in Figure 7.4), the convolution operation is computed by the Equation (7.10):

$$X_n = \sum_{k=0}^{N-1} y_k f_{n-k} \tag{7.10}$$

where y and f are the ECG signal and the applied filter, respectively, and N is the number of elements in the ECG signal y. The convolution layer's output is represented by vector$X$. For all layers of the DAG-CNN architecture, except ReLU

and ADD layers, the Equations (7.11) and (7.12) are used to update biases and weights as follows:

$$\Delta W_t(t+1) = -\frac{x_\lambda}{r} W_l - \frac{x}{n}\frac{\partial C}{\partial W_l} + m\Delta W_l(t) \tag{7.11}$$

$$\Delta B_l(t+1) = -\frac{x}{n}\frac{\partial C}{\partial B_l} + m\Delta B_l(t) \tag{7.12}$$

where $W, B, l,\ \lambda, x, n, m, t,$ and C denote the weight, bias, layer number, regularization parameter, learning rate, total number of training samples, momentum, updating step, and cost function respectively.

In DAG-CNNs, since lower layers are directly connected to the output layer through multi-scale connections, it is guaranteed that these layers' neurons receive a strong gradient signal during learning and do not suffer from the problem of vanishing gradients.

In this thesis, DAG-CNNs are used for automatically extracting and combining discriminative features and classifying the ECG data into different heartbeat classes. We assumed that the output of each component (convolution, ReLU, pooling, normalization, fully connected and ADD) of the proposed model is treated as a separate layer. Therefore our model has 22 layers. Our proposed model consists of a typical CNN as its base structure and some links from the intermediate and last ReLU layers. These links are connected to an average pooling layer to reduce their dimensionality, then are normalized and given to a separate fully connected MLP layers. Each of these fully connected layers have the same number of neurons in their last layer and is equal to number of class labels (five different AAMI classes) and generate a score vector for each samples. These score-vectors are added with each

other, element by element and fed into the final decision layer with softmax activation function to predict the class label.

In this study, we have shown that combining different level features can improve the classification accuracy significantly. Particularly, the classification accuracy is improved when we add features learned by intermediate layers, with the exception of the low-level features of early layers that cause a decrease in classification accuracy. For the purpose of testing different combinations of feature layers and finding the best one experimentally, features of the last layer are considered as of necessary and intermediate layer features are added layer-by-layer, one at a time, in a backward fashion until no improvement observed in classification accuracy. This greedy approach_ignores the features of layers closer to the input layer. Experimental evaluations as illustrated within the next section exhibited that the proposed system's capability of fusion of multi-scale features improves the accuracy of classification tasks.

The details of the DAG-CNN model have been summarized in Table 7.1. S-shaped rectified linear unit (SReLU) [78] has been used in our networks. Compared to other activation functions, SReLU is able to learn both convex and non-convex functions. SReLU is defined with the following formulation:

$$h(x_i) = \begin{cases} t_i^r + a_i^r(x_i - t_i^r), & x_i \geq t_i^r \\ x_i, & t_i^r > x_i > t_i^l \\ t_i^l + a_i^l(x_i - t_i^l), & x_i \leq t_i^l \end{cases} \tag{7.13}$$

where are four learnable parameters used to model an individual SReLU activation unit. a_i^r is the slope of the right line when the inputs exceed the threshold t_i^r, t_i^l is used to represent threshold in the negative direction.

All of the max-pooling and average-pooling layers kernel size and stride are set to 2. All of the 3 fully-connected layers (FC) in Figure 7.7, are multi-layer perceptron (MLP) with 3 layers consist of 25, 15 and 5 neurons respectively.

In the last layer, the softmax function is used to generate the final decision of the system which can be one of the output classes namely N, S, V, F, and Q. The overall schematic of the proposed method is illustrated in Figure 7.4.

Table 7.1: The details of back-bone CNN architecture of DAG-CNN model

Layers	**Type**	**No. of Neurons**	**Kernel size**	**Stride**
1	Convolution	276 × 5	5	1
3	Max-pooling	138 × 5	2	2
4	Convolution	135 × 10	4	1
6	Max-pooling	67 × 10	2	2
7	Convolution	65 × 15	3	1
9	Max-pooling	32 × 10	2	2
10	Convolution	30 × 20	3	1
12	Max-pooling	15 × 20	2	2
13	Average-pooling	15 × 10	2	2
14	Average-pooling	15 × 10	2	2
18-19-20	Fully Connected	25 × 15 × 5	-	-

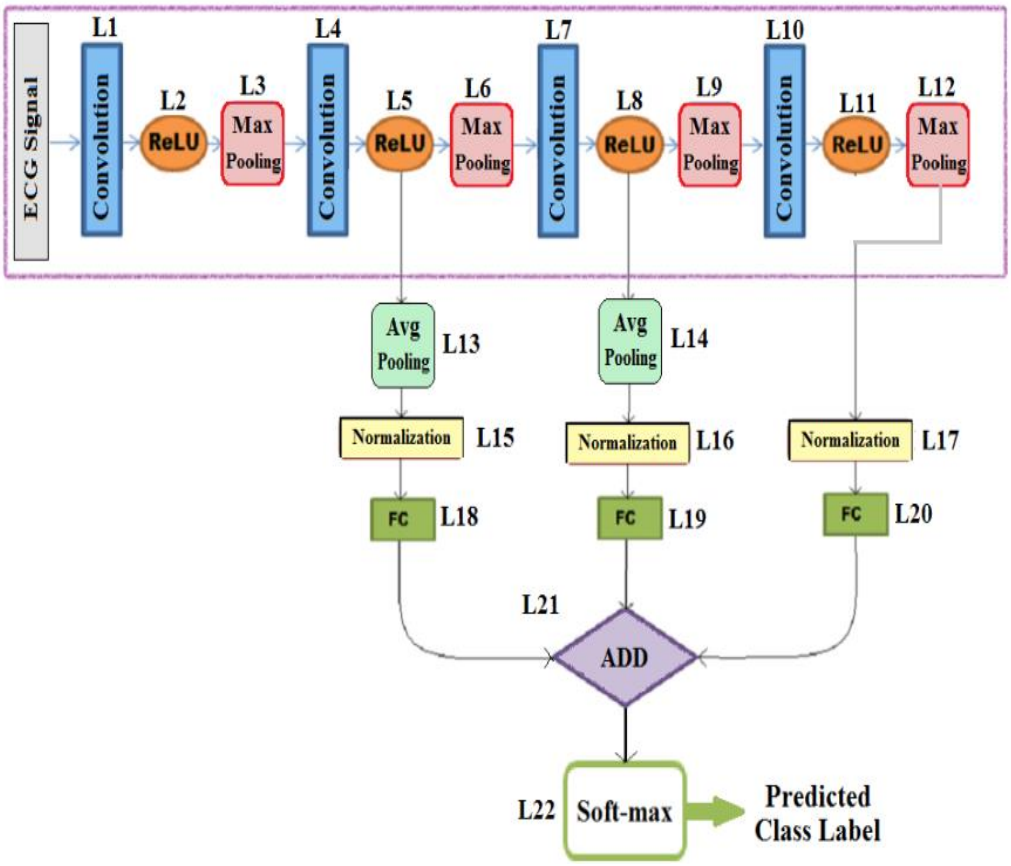


Figure 7.4: Overview of Proposed DAG-CNN Method [73]

Chapter 8

EXPERIMENTAL RESULTS

8.1 Introduction

In our experiments, following the AAMI standard as illustrated in Table 8.1, four records labelled with (102,104,107 and 217) containing paced beats with low signal quality were removed. Moreover, in order to follow AAMI recommendations, the remaining 44 records are divided into two disjoint sets, namely DS1 and DS2. Training set (DS1) contains records labelled with 101, 106, 108, 109, 112, 114, 115, 116, 118, 119, 122, 124, 201, 203, 205, 207, 208, 209, 215, 220, 223, and 230; and the test set (DS2) contains records labelled with 100, 103, 105, 111, 113, 117, 121, 123, 200, 202, 210, 212, 213, 214, 219, 221, 222, 228, 231, 232, 233, and 234. In all of the experiments, the patient-specific assessment strategy has been followed. Therefore, DS1 and the first 300 heartbeat samples of DS2 records are included within the training set while the remaining heartbeat samples of DS2 are used for testing.

Table 8.1: Summary of the training and testing heartbeat samples

AAMI heartbeat type	Total # Samples	Training set	Testing set
N	89695	45653	44042
S	2946	983	1963
V	7459	4252	3207
F	811	423	388
Q	15	8	7

We implemented the proposed system on a Pc with 2 Intel Xeon E5 processors-2.00 GHz with 8-GB of memory, running MATLAB 2017b without GPU. The total time taken by a single beat during all sub-systems propagation was 0.67 msec which shows the significantly low computational cost of the proposed model.

Many experiments are conducted to evaluate the performance of the proposed method over the MIT-BIH data set. The data set, metrics, parameters and experimental setup details are given in the following subsections.

8.1.1 Metrics

In order to evaluate different methods compared to each other, the following standard metrics are computed: classification accuracy (Acc), sensitivity (Sen), specificity (Spe), and positive predictive ratio (Ppr). The equations associated with these popular metrics are as follows: Accuracy is the ratio of the number of correctly classified patterns to the total number of patterns classified:

$$Acc = \frac{(TP + TN)}{TP + TN + FP + FN} \tag{8.1}$$

Sensitivity is the proportion of correctly classified events among all events:

$$Sen = \frac{TP}{TP + FN} \tag{8.2}$$

and Positive predictive ratio is the proportion of correctly classified events in all detected events:

$$Ppr = \frac{TP}{TP + FP} \tag{8.3}$$

where TP,TN,FP and FN stand for true positives, true negatives, false positives and false negatives, respectively.

8.2 Experimental results of proposed 3-Level Feature Fusion

8.2.1 Handcrafted Feature Results

For each heartbeat signal in the training and test sets, the above described temporal features are extracted. In this respect, the RR interval between the current beat and the next one, RR interval between the current beat and the previous one and RR interval averaged over the last ten beats are considered as the additional temporal features. Consequently, statistical features that are explained in Section 4.1 are computed for each ECG beat signal. These features are normalized by Z-score normalization method and then concatenated one after the other to implement their feature-level fusion. The fused feature vector is fed to a SVM classifier with RBF kernel to predict the class label for an input heartbeat. Table 8.2 shows the confusion matrix of this sub-classifier system.

Table 8.2: Confusion matrix of statistical and temporal feature fusion subsystem

		Predicted Label			
		N	**S**	**V**	**F**
True Label	**N**	36978	218	398	390
	S	427	1168	4	2
	V	152	6	2554	123
	F	96	4	16	237

For the extraction of morphological features, each beat segment is passed through the FIR approximation of Mayer's wavelet and the corresponding wavelet 4-th level approximation and details coefficients are computed. In [78], the FIR approximation of Mayer's wavelet is claimed to achieve the highest classification accuracy after a detailed evaluation several different wavelet basis functions independently for the decomposition the ECG signals. After reducing the dimensionality of the extracted wavelet features with PCA, the first 10% of the principal components were selected and applied to a SVM classifier with RBF kernel

to predict the class label of the input heartbeat. Table 8.3, illustrates the confusion matrix associated with the evaluations using this sub-classifier system.

In training and testing procedures of these sub-classifier systems, a five-fold cross validation approach is followed. Accordingly, heartbeat samples are divided into five separate subsets such that uniform distributions of different heartbeat classes exist in each subset. Each experiment is repeated five times and the average values of the computed scores are considered as the subsystem's performance accuracy.

Table 8.3: Confusion matrix of morphological feature based subsystem

		Predicted Label			
		N	**S**	**V**	**F**
True Label	**N**	36428	356	691	509
	S	469	1110	15	7
	V	61	0	2720	54
	F	58	2	11	282

8.2.2 CNN-based Learned Features

Architectural properties of the proposed CNN for ECG classification are summarized in Table 5. This architecture contains five convolutional layers and three fully-connected layers. For all the five subsequent convolutional layers, the kernel sizes changed from large to small as mentioned as illustrated in Table 8.4. Each convolutional layer followed by a S-shaped rectified linear unit (SReLU) [77] and a max pooling layer with kernel size equal to 2, strides equal 2 and a local response normalization layer. Following the 5 convolutional layers, there are 3 fully connected layers containing 40, 30 and 5 neurons, respectively. In the last layer, the softmax function is used to generate the final decision of the system which can be one of the output classes namely *N, S, V, F,* and *Q*.

Table 8.4: The details of proposed CNN architecture

Layers	Type	No. of Neurons	Kernel size	Stride
L1	Convolution	274 × 10	7	1
L2	Max-pooling	137 × 10	2	2
L3	Convolution	133 × 15	5	1
L4	Max-pooling	66 × 15	2	2
L5	Convolution	62 × 20	5	1
L6	Max-pooling	31 × 20	2	2
L7	Convolution	29 × 25	3	1
L8	Max-pooling	14 × 25	2	2
L9	Convolution	12 × 30	3	1
L10	Max-pooling	6 × 30	2	2
L11	Fully connected	40	-	-
L12	Fully connected	30	-	-
L13	Fully connected	5	-	-

The proposed CNN architecture was trained through the standard backpropagation technique with a batch size of 32. In order to obtain the optimum performance, the other learning parameters are set as follows: to prevent overfitting of training data, the regularization parameter (λ) is set to 0.1, momentum parameter which adjusts the speed of learning during training is set to 0.9 and learning rate that controls the speed of convergence is set to 0.001 and linearly changed according to the mean-square error values in each ten iterations. The training was performed over 50 epoch rounds.

8.2.2.1 Score-level Fusion of Multi-stage CNN Learned Features

For the purpose of testing different combinations of feature layers and finding the best one experimentally, features of the last layer are considered as necessary and intermediate layer features are added, one at a time, in a backward fashion until no improvement is observed in classification accuracy. This greedy approach ignores the features of layers close to the input layer. After a number of experimental evaluations, selection of features from layers L5, L7 and L9 resulted in the best classification performance. The selected layers' features are input to Algorithm 1 to carry out the score-level fusion.

In order to investigate the effect of score-level fusion on the performance of ECG classification system, learned features from different layers of a trained CNN are combined together using the score-level fusion approach. Experimental evaluations exhibited that score-level fusion of multilayer CNN features together with the last layer features achieved significant improvement in classification performance. Table 8.5 presents the confusion matrix of the score-level fusion of multi-stage CNN features.

Table 8.5: Confusion matrix of multi-stage CNN features fusion subsystem

		Predicted Label			
		N	**S**	**V**	**F**
True Label	**N**	37517	94	215	158
	S	499	1079	16	7
	V	147	1	2572	115
	F	82	1	18	252

8.2.3 Decision-level Fusion

In order to utilize both handcrafted and learned features, the classification results of the above described three sub-classifier systems are fused together to predict the final class label by the proposed system. This decision-level fusion is performed by the simple majority-voting technique. In case of tie, one of the subsystem's results is randomly selected and considered as the final decision. Table 8.6 presents the confusion matrix of the decision-level fusion and Table 8.7 illustrates the classification accuracy of all the three subsystems and the proposed method.

Table 8.6: Proposed method confusion matrix

		Predicted Label			
		N	**S**	**V**	**F**
True Label	**N**	37743	40	124	77
	S	384	1210	5	2
	V	97	0	2659	79
	F	42	0	8	303

The achieved improvement in the classification accuracy shows that hand-crafted features play a complementary role when used together with the learned features.

Table 8.7: Classification metrics of the proposed method and its subsystems (%)

System	Accuracy	N		S		V		F	
		Sen	Ppr	Sen	Ppr	Sen	Ppr	Sen	Ppr
Morphological-based	94.8	95.9	98.4	69.3	75.6	95.9	79.1	79.9	33.1
Fusion of Statistical and Temporal	95.7	97.4	98.2	72.9	83.6	90.0	86.0	67.1	31.5
Multi-stage CNN	96.8	98.8	98.1	64.4	91.8	90.7	91.2	71.4	47.4
Proposed 3-Level Feature Fusion	98.00	99.4	98.6	75.6	96.8	93.8	95.1	85.8	65.7

8.2.4 Statistical Analysis of Experimental Results

The last step of experimental evaluations is the Friedman aligned ranks test that is implemented over all average sensitivity (Sen) scores achieved by the 7 algorithms and the proposed algorithm. The objective of this test is to check the statistical similarity of our results to those of others listed in Table 8.8.

This test is carried out using the corresponding function in RTools and Table 8.8 presents the pairwise p values obtained through the Benjaminyi-Hochberg post-hoc procedure. The computed p values indicate the significant difference between the compared pairs of algorithms. It is well that the smaller the p value, the more statistically different is the corresponding methods. As illustrated in Table 8.8, due to a small number of test cases for each data sample, namely, N, S, V and F, a great number of pairs of methods seem statistically similar to each other as illustrated with the corresponding p values larger than or equal 0.5. However, considering our proposed method, it seems statistically similar to only Luo et al. [59] and Chazal et al.[10], hence it is statistically different from 5 of its 7 competitors. Particularly, except Martis et al. [78], the proposed method's statistical similarity to its

competitors is comparably very low as illustrated by the corresponding p values. As such, it is one of the best algorithms that are statistically different from their competitors.

Table 8.8: Friedman aligned ranks test results

	Proposed	Luo et al.	Chazal et al.	Jiang et al.	Ince et al.	Martis et al.	Ye et al.
Luo et al.[59]	0.684						
Chazal et al.[10]	0.714	0.917					
Jiang et al.[58]	0.185	0.268	0.268				
Ince et al.[62]	0.185	0.268	0.268	0.942			
Martis et al.[78]	0.429	0.657	0.602	0.575	0.550		
Ye et al.[28]	0.183	0.185	0.185	0.657	0.684	0.268	
Zhang et al.[61]	0.268	0.470	0.429	0.713	0.684	0.714	0.449

8.3 Experimental Results of the Proposed DAG-CNN Model

The proposed DAG-CNN architecture is trained through the standard backpropagation technique with a batch size of 8. In order to obtain optimum performance, the other learning parameters are set as follows: in order to prevent overfitting of training data, the regularization (λ) is set to 0.2, momentum parameters which adjust the speed of learning during training is set to 0.8, and learning rate that control the convergence of the training data are set to 0.0002 and linearly changed according to the mean-squared error values in each five iteration. The training was performed 40 epochs rounds. The goal of our DAG-CNN model is to classify heartbeats into AAMI five classes. Table 8.9 summarizes beat-by-beat classification results of ECG heartbeat patterns for all test records.

For each of the four classes N, S, V and F, we compare the classification performance of our system with the state-of-the-art approaches in Table 8.11. For this comparison we compute the four standard metrics: classification accuracy (Acc), sensitivity (Sen), specificity (Spe), and positive predictive ratio (Ppr).

For clinical application, the sensitivity, specificity, and positive predictively measurement are more relevant performance criteria because there is a large difference in the number of beats from different classes in the training/testing data.

Clinically, supraventricular ectopic beats (SVEB) and ventricular ectopic beats (VEB) are two critically abnormal and serious heartbeats. So for performance evaluation, we also present the results in terms of VEB (V class versus [N, S, and F]) and SVEB (S class versus [N, V, and F]). The VEB and SVEB classification results of the proposed technique over all DS2 records are summarized in Table 8.12. It is observed that overall, the performance of the proposed method in VEB and SVEB detection is significantly better than most of the state-of-the-art methods and comparable with the best method [59].

Table 8.9: The details of back-bone CNN architecture of DAG-CNN model

Layers	**Type**	**No. of Neurons**	**Kernel size**	**Stride**
1	Convolution	276 × 5	5	1
3	Max-pooling	138 × 5	2	2
4	Convolution	135 × 10	4	1
6	Max-pooling	67 × 10	2	2
7	Convolution	65 × 15	3	1
9	Max-pooling	32 × 10	2	2
10	Convolution	30 × 20	3	1
12	Max-pooling	15 × 20	2	2
13	Average-pooling	15 × 10	2	2
14	Average-pooling	15 × 10	2	2
18-19-20	Fully Connected	25 × 15 × 5	-	-

Table 8.10: Different heartbeat types classification result

		N	S	V	F	Q
Grand truth	**N**	37571	80	192	133	8
	S	427	1156	11	5	2
	V	88	2	2671	74	0
	F	60	1	2	281	0
	Q	6	0	0	1	0

8.4 Comparison with the State-of-the-art Methods

As illustrated in Table 8.11, classification performance of the proposed system is compared with the state-of-the-art approaches for each of the four AAMI classes N, S, V and F. The same experimental settings are used for all algorithms. Clinically, supraventricular ectopic beats (SVEB) and ventricular ectopic beats (VEB) are two critically abnormal and serious heartbeats. Accordingly, experimental results associated with VEB (V class versus [N, S, and F]) and SVEB (S class versus [N, V, and F]) heartbeats are used for comparative performance evaluations. Different evaluation metrics of the proposed methods for VEB and SVEB classification are mentioned in Table 8.12. The robustness and effectiveness of the proposed method are studied in terms of accuracy, sensitivity and positive predictive ratio in Table 8.11 and Table 8.12. Consequently, the experimental results show that our proposed method outperforms all of the state-of-the-art methods. Also, it is observed that the proposed system recognizes VEB and SVEB more accurate than all of the state-of-the-art methods. The results demonstrate that combining different features in different fusion levels (feature-level and score-level) enhances the performance of the ECG classification system.

Table 8.11: Classification metrics compared to the state-of-the-art (percentage)

Reference	Accuracy	N		S		V		F	
		Sen	Ppr	Sen	Ppr	Sen	Ppr	Sen	Ppr
Luo [59]	97.5	99.0	98.4	71.4	94.4	93.3	93.3	82.7	58.5
Chazal[10]	93.9	94.3	99.4	87.7	47.0	94.3	96.2	74.0	29.1
Jiang [58]	94.5	98.7	96.2	50.6	68.0	86.6	89.4	35.8	84.2
Ince [62]	93.6	97.0	97.0	62.1	56.7	83.4	86.5	61.4	73.4
Martis[78]	89.0	94.2	99.2	86.2	56.7	92.4	93.4	66.4	17.7
Ye [28]	88.2	90.0	98.2	56.4	55.1	84.7	59.5	35.8	5.8
Zhang[61]	88.3	88.9	99.0	79.1	36.0	85.5	92.8	93.8	13.7
Proposed 3-Level Feature Fusion	**98.00**	**99.4**	**98.6**	**75.6**	**96.8**	**93.8**	**95.1**	**85.8**	**65.7**
Proposed DAG-CNN model	**97.15**	**98.8**	**98.3**	**72.3**	**93.0**	**92.0**	**92.8**	**81.7**	**57.00**

Table 8.12: Comparison of SVEB-VEB classification with the state-of-the-art

Reference	SVEB				VEB			
	Acc	Sen	Ppr	Spe	Acc	Sen	Ppr	Spe
Luo [59]	98.8	71.4	94.4	99.8	99.1	93.3	93.3	99.5
Kiranyaz [60]	96.4	64.6	62.1	98.6	98.6	95	89.5	98.1
Chazal [10]	95.9	87.7	47.0	96.2	99.4	94.3	96.2	99.7
Jiang [58]	96.6	50.6	68.0	98.8	97.7	86.6	89.4	98.9
Ince [62]	97.3	63.5	53.7	98.3	98.0	84.6	86.7	99.0
Martis [78]	93.3	83.2	33.5	93.7	97.4	86.8	75.9	98.1
Ye [28]	97.4	56.4	55.1	98.6	94.6	84.7	59.5	95.4
Zhang [61]	93.3	79.1	36.0	93.9	98.6	85.5	92.7	99.5
Proposed 3-level multistage CNN	**99.0**	**75.6**	**96.8**	**99.9**	**99.3**	**93.8**	**95.1**	**99.7**
Proposed DAG-CNN model	**98.8**	**72.2**	**93.3**	**99.8**	**99.1**	**94.2**	**92.8**	**99.5**

Chapter 9

CONCLUSION

Two novel ECG classification methods based on deep learning approaches have proposed in this thesis. The proposed methods are fully automatic, non-invasive and require minimal interaction with the clinician and can be applied to other heartbeat abnormalities as well. The only limitation of this extension is collecting a large annotated database including other types of abnormal cardiac rhythms, such as idioventricular rhythms and asystole.

By comprehensive experiments on MIT-BIH dataset, the following conclusions are proposed:

- multi-stage deep feature based method can archive better classification accuracy than methods based on the hand-crafted feature fusion.
- methods based on multi-stage CNN feature fusion such as DAG-CNN, can achieve better performance than CNN based method.
- shallow features from early layers of CNN can be complementary to deep features of the last layer of CNN and fusion of these features improve the discriminative capability of the feature set.

Fusion of feature descriptors extracted from a signal through different methods is an important issue for the exploitation of representational power of each descriptor. As a part of this research work, a novel system which exploits multi-stage features from a

trained convolutional neural network (CNN) and precisely combines these features with a selection of handcrafted features is proposed. The set of handcrafted features consists of three subsets namely, wavelet transform based morphological features representing localized signal behaviour, statistical features exhibiting overall variational characteristics of the signal and temporal features representing the signal's behaviour on the time axis. Different levels of information fusion are proposed in this thesis.

Various hand-crafted features such as temporal features and statistical features are combined during the first level of information fusion (feature-level) procedures. Additionally, morphological features are extracted by using DWT method and used to classify ECG signals separately. Multi-stage learned features from different layers of a trained CNN for ECG classification are combined together by the score-level fusion method. Finally, the obtained results of these three approaches are aggregated by using a decision-level fusion method. The aggregation result shows that generic features extracted from CNN are enhanced by combining them with domain-specific features.

The proposed DAG-CNN architecture, presents a multi-scale system that can automatically learn different level of features, combine them and predict the output label. The proposed system is very efficient and instead of performing feature level fusion manually and feeding the results into a classifier, the system works in fully automatic manner.

Compared with the state-of-the-art methods, our proposed approaches obtained significant improvement in classification accuracy on MIT-BIH database. The

proposed system results showed that the information fusion of handcrafted features in feature-level and fusion of global learned features in score-level provide a higher accuracy than the other algorithms.

Based on the obtained experimental results, it is evident that the proposed systems have the potential to be used in real clinical environments and can be implemented on portable device for the long-term monitoring of cardiac arrhythmia.

As the future works we intend to conduct studies on the employing modern deep CNN architectures such as GoogLeNet and ResNet for the other biomedical signal classification problems to further verify the multi-stage feature fusion idea according to their characteristics.

REFERENCES

[1] Druzhkov, P.N. & Kustikova, V.D. (2016). A survey of deep learning methods and software tools for image classification and object detection. *Pattern Recognition and Image Analysis*, 26(1), 9–15.

[2] Malmivuo, J., & Plonsey, R. (1995). Bioelectromagnetism: principles and applications of bioelectric and biomagnetic fields. Oxford University Press, USA, 25-32.

[3] MacKay, D. J. (1992). A practical Bayesian framework for backpropagation networks. *Neural computation*, 4(3), 448-472.

[4] Hu, Y. H., Palreddy, S., & Tompkins, W. J. (1997). A patient-adaptable ECG beat classifier using a mixture of experts approach. *IEEE transactions on biomedical engineering*, 44(9), 891-900.

[5] Coast, D. A., Stern, R. M., Cano, G. G., & Briller, S. A. (1990). An approach to cardiac arrhythmia analysis using hidden Markov models. *IEEE Transactions on biomedical Engineering*, 37(9), 826-836.

[6] Lagerholm, M., Peterson, C., Braccini, G., Edenbrandt, L., & Sornmo, L. (2000). Clustering ECG complexes using Hermite functions and self-organizing maps. IEEE *Transactions on Biomedical Engineering*, 47(7), 838-848.

[7] Silipo, R., & Marchesi, C. (1998). Artificial neural networks for automatic ECG analysis. *IEEE transactions on signal processing*, 46(5), 1417-1425.

[8] Thakor, N. V., & Zhu, Y. S. (1991). Applications of adaptive filtering to ECG analysis: noise cancellation and arrhythmia detection. *IEEE transactions on biomedical engineering*, 38(8), 785-794.

[9] Chang, W. H., Lin, K. P., & Tseng, S. Y. (1988). ECG analysis based on Hilbert transform descriptor. In Engineering in Medicine and Biology Society, 1988. *Proceedings of the Annual International Conference of the IEEE* (pp. 36-37).

[10] De Chazal, P., O'Dwyer, M., & Reilly, R. B. (2004). Automatic classification of heartbeats using ECG morphology and heartbeat interval features. *IEEE transactions on biomedical engineering*, 51(7), 1196-1206.

[11] Zhou, S. H., Rautaharju, P. M., & Calhoun, H. P. (1993). Selection of a reduced set of parameters for classification of ventricular conduction defects by cluster analysis. *Computers in Cardiology*, Proceedings. (pp. 879-882).

[12] Herrero, G. G., Gotchev, A., Christov, I., & Egiazarian, K. (2005, March). Feature extraction for heartbeat classification using independent component analysis and matching pursuits. *In Acoustics, Speech, and Signal Processing*, Proceedings. (ICASSP'05). IEEE International Conference on (Vol. 4, pp. iv-725). IEEE.

[13] Christov, I., & Bortolan, G. (2004). Ranking of pattern recognition parameters for premature ventricular contractions classification by neural networks. *Physiological Measurement*, 25(5), 1281.

[14] Osowski, S., & Linh, T. H. (2001). ECG beat recognition using fuzzy hybrid neural network. *IEEE Transactions on Biomedical Engineering*, 48(11), 1265-1271.

[15] Senhadji, L., Bellanger, J. J., Carrault, G., & Passariello, G. (1995). Comparing wavelet transforms for recognizing cardiac patterns. *IEEE Engineering in Medicine and Biology Magazine*, 14(2), 167-173.

[16] Afonso, V. X., & Tompkins, W. J. (1995). Detecting ventricular fibrillation. *IEEE Engineering in Medicine and Biology Magazine*, 14(2), 152-159.

[17] Caswell, S. A., Kluge, K. S., Chiang, C. M., Jenkins, J. M., & DiCarlo, L. A. (1993). Pattern recognition of cardiac arrhythmias using two intracardiac channels. In *Computers in Cardiology* , Proceedings. (pp. 181-184).

[18] Zhang, X. S., Zhu, Y. S., Thakor, N. V., & Wang, Z. Z. (1999). Detecting ventricular tachycardia and fibrillation by complexity measure. *IEEE Transactions on biomedical engineering*, 46(5), 548-555.

[19] Minami, K. I., Nakajima, H., & Toyoshima, T. (1999). Real-time discrimination of ventricular tachyarrhythmia with Fourier-transform neural network. *IEEE transactions on Biomedical Engineering*, 46(2), 179-185.

[20] Frasconi, P., Gori, M., Maggini, M., & Soda, G. (1995). Unified integration of explicit knowledge and learning by example in recurrent networks. *IEEE Transactions on Knowledge and Data Engineering*, 7(2), 340-346.

[21] Hu, Y. H., Tompkins, W. J., Urrusti, J. L., & Afonso, V. X. (1993). Applications of artificial neural networks for ECG signal detection and classification. *Journal of electrocardiology*, 26, 66-73.

[22] Mathews, S. M., Kambhamettu, C., & Barner, K. E. (2018). A novel application of deep learning for single-lead ECG classification. *Computers in biology and medicine*.

[23] Li, Y., Pang, Y., Wang, J., & Li, X. (2018). Patient-specific ECG classification by deeper CNN from generic to dedicated. *Neurocomputing*, 314, 336-346.

[24] Zhai, X., & Tin, C. (2018). Automated ECG Classification using Dual Heartbeat Coupling based on Convolutional Neural Network. *IEEE Access*.

[25] Linpeng,J.& Jun, D.(2018).Normal Versus Abnormal ECG Classification by the Aid of Deep Learning. *Artificial Intelligence*,1-22.

[26] Xiong, Z., Stiles, M. K., & Zhao, J. (2017). Robust ECG Signal Classification for Detection of Atrial Fibrillation Using a Novel Neural Network. *Computing*, 44, 1.

[27] Al Rahhal, M. M., Bazi, Y., AlHichri, H., Alajlan, N., Melgani, F., & Yager, R. R. (2016). Deep learning approach for active classification of electrocardiogram signals. *Information Sciences*, 345, 340-354.

[28] Ye, C., Kumar, B.V. & Coimbra, M.T. (2012). Heartbeat classification using morphological and dynamic features of ECG signals. *IEEE Transactions on Biomedical Engineering*, 59(10), 2930–2941.

[29] Ai, D., Yang, J., Wang, Z., Fan, J., Ai, C., & Wang, Y. (2015). Fast multi-scale feature fusion for ECG heartbeat classification. *EURASIP Journal on Advances in Signal Processing*, 2015(1), 46.

[30] Mar, T., Zaunseder, S., Martínez, J.P. & et al. (2011). Optimization of ECG classification by means of feature selection. *IEEE transactions on Biomedical Engineering*, 58(8), 2168-2177.

[31] Sambhu, D. & Umesh, A.C. (2013). Automatic classification of ECG signals with features extracted using wavelet transform and support vector machines. *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering*, 2(1), 235-241.

[32] Das, M. K., & Ari, S. (2014). ECG beats classification using mixture of features. *International scholarly research notices*.

[33] WIKIPEDIA, the free encyclopedia. *Pattern recognition* [online]. Available from: http://en.wikipedia.org/wiki/Pattern_recognition.

[34] Fu, K. S., Keidel, Levelt, W. D., & W. J. M. (1980). *Digital Pattern Recognition*. Springer-Verlag, Berlin Heidelberg New York.

[35] Dickhaus, H., & Heinrich, H. (1996). Classifying biosignals with wavelet networks [a method for noninvasive diagnosis]. *IEEE Engineering in Medicine and Biology Magazine*, 15(5), 103-111.

[36] Pandya, A. S., & Macy, R. B. (1995). *Pattern recognition with neural networks in C++*. CRC press.

[37] Malmivuo, J., & Plonsey, R. (1995). Bioelectromagnetism: principles and applications of bioelectric and biomagnetic fields. *Oxford University Press*, USA.

[38] Einthoven, W. (1908). Weiteres über das Elektrokardiogramm. *Archiv für die gesamte Physiologie des Menschen und der Tiere*, 122(12), 517-584.

[39] Addison, P. S. (2005). Wavelet transforms and the ECG: a review. *Physiological measurement*, 26(5), R155.

[40] Arvanaghi, R., Daneshvar, S., Seyedarabi, H. & Goshvarpour, A. (2017). Fusion of ECG and ABP signals based on wavelet transform for cardiac arrhythmias classification. *Computer methods and programs in biomedicine*, 151, 71-78.

[41] Sandler, J., & Rosenblatt, B. (1962). The concept of the representational world. *The psychoanalytic study of the child*, 17(1), 128-145.

[42] Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of physiology*, 160(1), 106-154.

[43] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *In Advances in neural information processing systems* (pp. 1097-1105).

[44] Advances in machine learning.(2018, November 10). Retrieved from http://adventuresinmachinelearning.com

[45] Convolutional neural network for visual recognition. (2018, November 10). Retrieved from http://cs231n.github.io/convolutional-networks

[46] Wikipedia, the free encyclopedia. Pattern recognition. (2018, November 10). Retrieved from http://en.wikipedia.org/wiki

[47] Tutorial slides by Andrew Moore. (2018, November 10). Retrieved from http://www.cs.cmu.edu/~awm

[48] Vapnik, V. (1995). The nature of statistical learning theory Springer, New York Google Scholar.

[49] Burges, C. J. (1998). A tutorial on support vector machines for pattern recognition. *Data mining and knowledge discovery*, 2(2), 121-167.

[50] Vapnik, V., & Mukherjee, S. (2000). Support vector method for multivariate density estimation. *In Advances in neural information processing systems* (pp. 659-665).

[51] Evgeniou, T., Pontil, M., & Poggio, T. (2000). Statistical learning theory: A primer. *International Journal of Computer Vision*, 38(1), 9-13.

[52] McKinney, B. A., Reif, D. M., Ritchie, M. D., & Moore, J. H. (2006). Machine learning for detecting gene-gene interactions. *Applied bioinformatics*, 5(2), 77-88.

[53] Cristianini, N., & Shawe-Taylor, J. (2000). An introduction to support vector machines and other kernel-based learning methods. *Cambridge university press*.

[54] Burges, C. J. (1998). A tutorial on support vector machines for pattern recognition. *Data mining and knowledge discovery*, 2(2), 121-167.

[55] Moody, J. & Darken, C., (1989). Fast learning in networks of locally tuned processing units. *Neural Computation*, 1, pp. 281-294.

[56] Hampton, J. R., (1998). The ECG Made Easy 5th edition. *Churchill Livingstone*, London.

[57] De Chazal, P.P. & Reilly, R.B. (2006). A patient-adapting heartbeat classifier using ECG morphology and heartbeat interval features, *IEEE Transactions on Biomedical Engineering*, 53(12), 2535–2543.

[58] Jiang, W. & Kong, S.G. (2007). Block-based neural networks for personalized ECG signal classification. *IEEE Transactions on Neural Networks*, 18(6), 1750–1761.

[59] Luo, K., Li, J., Wang, Z. & Cuschieri, A. (2017). Patient-Specific Deep Architectural Model for ECG Classification. *Journal of Healthcare Engineering*, 2017, Article ID 4108720, 13 pages

[60] Kiranyaz, S., Ince, T. & Gabbouj, M. (2016). Real-Time Patient-Specific ECG Classification by 1-D Convolutional Neural Networks. *IEEE Transactions on Biomedical Engineering*, 63(3), 664-675.

[61] Zhang, Z., Dong, J., Luo, X., Choi, K.S. & Wu, X. (2014). Heartbeat classification using disease-specific feature selection. *Computers in Biology and Medicine*, 46, 79– 89.

[62] Ince, T., Kiranyaz, S. & Gabbouj, M. (2009). A generic and robust system for automated patient-specific classification of electrocardiogram signals. *IEEE Transactions on Biomedical Engineering*, 56(5), 1415–1426.

[63] Gokhale, P. S. (2012). ECG Signal De-noising using Discrete Wavelet Transform for removal of 50Hz PLI noise. *International Journal of Emerging Technology and Advanced Engineering*, 2(5), 81-85.

[64] Thakor, N. V., Zhu, Y. S., & Pan, K. Y. (1990). Ventricular tachycardia and fibrillation detection by a sequential hypothesis testing algorithm. *IEEE Transactions on Biomedical Engineering*, 37(9), 837-843.

[65] Al-Qawasmi, A. R., & Daqrouq, K. (2010). ECG signal enhancement using wavelet transform. *WSEAS Transactions on Biology and Biomedicine*, (2), 62-71.

[66] Gautam, R., & Sharmar, A. (2010). Detection of QRS complexes of ECG recording based on wavelet transform using Matlab. *International Journal of Engineering Science and Technology*, 2(7), 3038-3044.

[67] Sumathi, S., & Sanavullah, M. Y. (2009). Comparative study of QRS complex detection in ECG based on discrete wavelet transform. *International Journal of Recent Trends in Engineering*, 2(5), 273.

[68] Priyadarshini, B., Ranjan, R. K., & Arya, R. (2012). Determining ECG characteristics using wavelet transforms. *International Journal of Engineering Research & Technology (IJERT)*, 1(6), 2278-018.

[69] Sasikala, P., & Wahidabanu, R. (2010). Robust r peak and qrs detection in electrocardiogram using wavelet transform. *International Journal of Advanced Computer Science and Applications-IJACSA*, 1(6), 48-53.

[70] Golrizkhatami, Z., & Acan, A. (2018). ECG classification using three-level fusion of different feature descriptors. Expert Systems with Applications, 114, 54-64.

[71] He, M., Horng, S.J., Fan, P., Run, R.S., Chen, R.J., Lai, J.L., Khan, M.K. & Sentosa, K.O. (2010). Performance evaluation of score level fusion in multimodal biometric systems. *Pattern Recognition*, 43(5), 1789-1800.

[72] Yim, J., Ju, J., Jung, H. & Kim, J. (2015). Image classification using convolutional neural networks with multi-stage feature. *Robot Intelligence Technology and Applications.* 3, 587–594.

[73] Golrizkhatami, Z., Taheri, S., & Acan, A. (2018). Multi-scale features for heartbeat classification using directed acyclic graph CNN. Applied Artificial Intelligence, 32(7-8), 613-628.

[74] Tang, P., Wang, H. & Kwong, S. (2017). G-MS2F: GoogLeNet based multi-stage feature fusion of deep CNN for scene recognition. *Neurocomputing*, 225, 188-197.

[75] Raiko, T., Valpola, H., & LeCun, Y. (2012, March). Deep learning made easier by linear transformations in perceptrons. *In Artificial Intelligence and Statistics* (pp. 924-932).

[76] Sermanet, P., Kavukcuoglu, K., Chintala, S., & LeCun, Y. (2013). Pedestrian detection with unsupervised multi-stage feature learning. *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 3626-3633).

[77] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D. & Rabinovich, A. (2015). Going deeper with convolutions. *In Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1-9).

[78] Yang, S., & Ramanan, D. (2015). Multi-scale recognition with DAG-CNNs. *In Proceedings of the IEEE International Conference on Computer Vision* (pp. 1215-1223).

[79] Jin, X., Xu, C., Feng, J., Wei, Y., Xiong, J., & Yan, S. (2016). Deep Learning with S-Shaped Rectified Linear Activation Units. *In AAAI* (Vol. 3, No. 2, pp. 3-2).

[80] Martis, R. J., Acharya, U. R., & Min, L. C. (2013). ECG beat classification using PCA, LDA, ICA and discrete wavelet transform. *Biomedical Signal Processing and Control*, 8(5), 437-448.

APPENDIX

Appendix A: Anatomy and Function of Human Heart

The human heart is located within the thoracic cavity, medially between the lungs in the space known as the mediastinum. The heart pumps blood through a closed system of blood vessels. Blood vessels allow blood to circulate to all parts of the body. Arteries usually colored red because oxygen rich, carry blood away from the heart to capillaries within the tissues. Veins usually colored blue because oxygen poor, carry blood to the heart from the capillaries. Capillaries are the smallest vessels within the tissues where gas exchange takes place. The function of the cardiovascular system is to deliver oxygen and nutrients to the body tissues and remove carbon dioxide and wastes products.

As illustrated in Figure A.1, the heart has 4 chambers, two champers in each left and right sides. The upper chambers are called the left and right atria, and the lower chambers are called the left and right ventricles. The muscles namely septum, divide the left and right atria and ventricles. Tricuspid, pulmonary, mitral and aortic valves are controlling the blood flow through all parts of the body. Atrium and ventricle are two heart chambers which are located at the left and right sides of the heart respectively. There are 4 valves in the charge of the blood flow namely, tricuspid, pulmonary, mitral and aortic.

The tricuspid and the mitral valves divide the right atriums and ventricles and left atriums and ventricles, respectively. The pulmonary valve controls the blood flow from heart to lungs. Finally, the aortic valve directs blood to the body circulation system. Walls of the heart are formed by cardiac muscle (myocardium). This muscle

is responsible for the mechanical work done by the heart. For controlling the pumping process specialized muscle cells that conduct electrical impulses evolved. These impulses are called action potential and they are responsible for forming the ECG waveform on the body surface.

In order to distribute oxygen to whole body, human's heart never stops. It works in periodic cycles. A cycle works as follows: Deoxygenated blood flows through superior vena cava to the right atrium. When the atrium is contracted, blood is pumped to the right ventricle. From the right ventricle the blood flows through pulmonary artery to the lungs. Lungs remove carbon dioxide from blood cells and replace it with oxygen.

Oxygenated blood returns to the left atrium and after another contraction it is pumped to the left ventricle. Finally the blood is forced out of the heart through aorta to the systemic circulation. The contraction period is called systole, during which the heart fills with blood. The relaxation period is called diastole. From electrical point of view the cycle has two stages - depolarization (activation) and repolarization (recovery).

A.1 The Conduction System of the Heart

To maintain the cardiac cycle the heart developed a special cell system for generating electrical impulses and by these impulses mechanical contraction of the heart muscle is ensured. This system is called conduction system (Figure A.2).

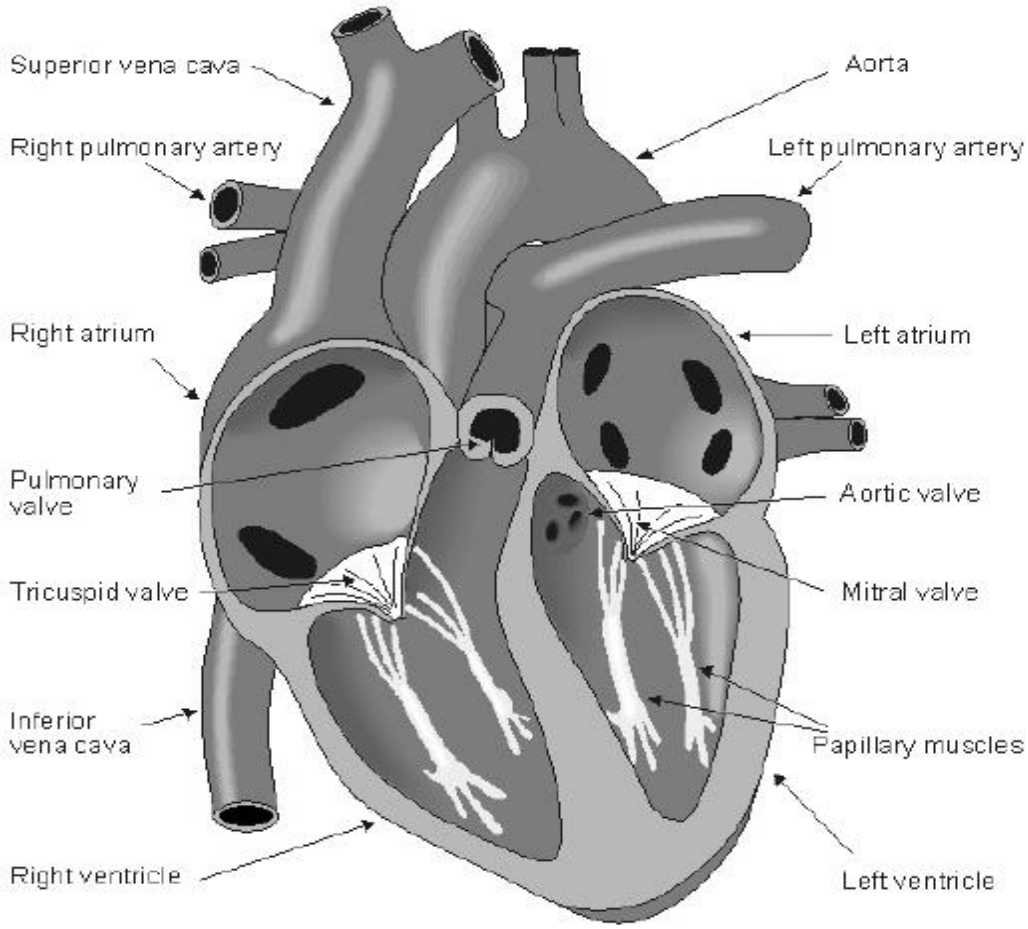


Figure A.1: Basic Heart Anatomy Schema - There are four chambers, two on the left (right heart) side responsible for pumping the blood to lungs and two on the right (left heart) responsible for pumping the blood to body[37]

It conveys impulses rapidly through the heart. Normal rhythmical impulse, which is responsible for contractions, is generated in the sinoatrial (SA) node. Then, it propagates to the right and left atrium and to the atrioventricular node (AV). The impulse is delayed in the AV node in order to allow proper contraction of the atria. Thus all blood volume in the atria is forced out to the ventricles before its contraction. Atrium and ventricles are electrically connected by bundle of His. From here, the impulse is conducted to the right and left ventricle. The pathway to the ventricles is divided to the left bundle branch and right bundle branch. Further, the bundles ramify into the Purkinje fibres that diverge to the inner sides of the ventricular walls.

The primary pacemaker of the heart is the sinoatrial node. However, other specialized cells in the heart (AV node, etc.) can also generate impulses but with lower frequency. If the connection from the atria to the atrioventricular node is broken, the AV node is considered as the main pacemaker. If the conduction system fails at the bundle of His, the ventricles will beat at the rate determined by their own region. All cardiac cell types have also different waveform of their action potentials (Figure A.2).

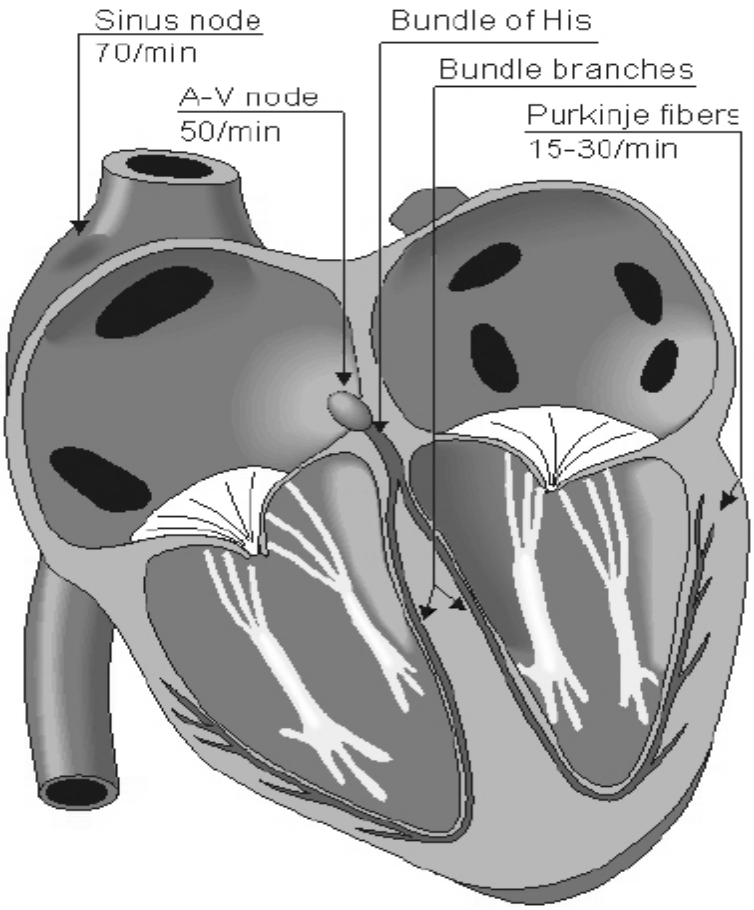


Figure A.2: Conduction System of the Heart Consists of Sinus Node, Atrioventricular node, Bundle of His, Bundle Branches and Purkinje Fibres [37]