CLASSIFICATION OF ECG SIGNALS CONVERTED TO 2D IMAGES BASED ON EXTRACTED FEATURES FOR CARDIAC DIAGNOSIS

by

RANDALL ISAAC KAKAIRE
URN: 6605258

A dissertation submitted in partial fulfilment of the
requirements for the award of

MASTER OF SCIENCE IN DATA SCIENCE

September 2020

Department of Computer Science
University of Surrey
Guildford GU2 7XH

Supervised by: Dr Lillian Tang

I declare that this dissertation is my own work and that the work of others is acknowledged and indicated by explicit references.

Randall Isaac Kakaire
September 2020

© Copyright Randall Isaac Kakaire, September 2020

Abstract

Cardiovascular disuses (CVD) is one of the worst diseases in today's world, accounting for 17 million deaths in 2015. There is a need for early indication of CVD, survival rates increase when prevention methods are started early. In primary care there is a need for virtual assistant in helping with the task of CVD diagnosis via ECG. Reading ECG tests requires skillful experience, and there sometimes can be a shortage of physician to do this task, hence leading to misdiagnosis when ECG test conducted by a less experienced person. I have proposed a model which is to able to classify ECG signals based on image representation of the features within the ECG signal such as RR interval. The model achieved a loss of 0.6 and accuracy of 0.3 on the PTB-XL dataset which is a huge data of labeled ECG signals for cardiac diagnosis. This led to poor scores on the test set in term of Recall, Precision and Accuracy. This downfall of this model pointed me in a new direction on how to achieve diagnosis via images by focusing less on the temporal features and more on the statistical ,scalar features.

Acknowledgements

I would like to give thanks to my supervisor Dr. Lillian Tang, for her insightful input on this project, which pushed me in the right direction, to reach completion.

Most importantly I would like to thank my parents, for the countless sacrifices, who've given everything they have to support me, so I could reach this point in my academic career.

Lastly, I give thanks to my brothers and friends, their support through my degree, made everything more fun, and kept me driven throughout.

Contents

Chapter 1

Introduction

The human heart is one of the most important organs, each and everyone one of us is born with. It is responsible for pumping blood through our circulatory system and carrying oxygen and nutrients to the body, while taking metabolic waste such as carbon dioxide to the lungs (Hall 2010). Given the importance the heart has on normal bodily functions, the heart is also susceptible to damage and failure, just like the other organs. Cardiovascular disease (CVD) is a huge problem in today's world, especially given the increase in sedentary lifestyles amongst people, hence the increased risk of life ending damage to our hearts. Currently CVD is the number one cause of death globally.

Back in 2015, 17 millions deaths were considered premature (under the age of 70), and CVD was the cause for about 37 percent of these deaths. The CVD death rate can be curbed if there is change and a response to our current lifestyles and diet, but for some it may be too late, hence why early detection of those at risk for CVD can help prevent more deaths, with early use of medicine and counselling (WHO 2020). Here in the UK it is estimated that 115,000 lives could be saved over the next 10 years from CVD, if we match the level at which Canada is screening and providing earlier diagnosis of CVD (BHF 2018).

For a physician there are several indicators such as blood pressure, cholesterol or alcohol consumption that he/she may look into, to assess the health of the patient's heart and risk of CVD, along with some key tests to perform a diagnosis, for example an electrocardiogram (ECG) of the patient's heart. It has been shown that frequent screening for indicators regarding CVD, has a positive influence on the physicians and patients behaviour towards, earlier use of preventative measures, and changes in the patients behaviour, to help increase the rate of primary care prevention of CVD (Hopkins et al. 2010).

The focus of this thesis is to automate the classification of ECG signals, for primary care intervention of CVD, as this is one of the key methods for identifying early cases of CVD. There is a growing paramount importance in being able to detect CVD early especially within primary care. There is a need for devices which are able to classify a variety of diseases, which physicians can use, as a secondary assistant. Due to the complexity of the heart, and high level of experience needed for evaluation, patients are usually forwarded to the cardiology department for accurate assessment. The downside here is that there may be more patients which out number the amount of qualified cardiologists for this task, hence leading to a number of healthy patients in the waiting pool for an assessment, who may not actually need a thorough assessment and could have been consulted at the primary care stage. So by providing better technology at the primary care intervention stage, we can detect and remove more healthy patients, and allow for more frequent appointments for real patients, reducing the waiting time needed to see a cardiologist, and start prevention methods earlier than usual (Phoenix 2020).

Given the rise in popularity and power of machine/deep learning techniques for classification tasks, ECG signal classification tasks have received a lot of attention, due to the complex nature of the ECG signal, as the signal can vary on a patient by patient basis, even between healthy patients the signal can appear different. There are many researchers who have proposed novel techniques for ECG classification, including different methods of signal preprocessing, feature extraction and feature selection. I will be taking a new approach to this problem, by converting this problem from the time series domain to the image domain and frame this problem as an image classification problem to take advantage of the ability of Convolutional Neural Networks to understand and extract high level features.

1.1 Background

1.1.1 History of The Heart

The heart has been a well studied and documented organ over many centuries, there is evidence dating back as early as the 19th century B.C, where Egyptians, had used anatomical terms to somewhat describe the heart as we know it today, where they believed it was the recorder of deeds, as they believed a person had a heavy heart if they committed bad deeds or a light heart if they committed good deeds.

Fast forwarding to the 5th century B.C, where Hippocrates first introduced the concepts of medical thinking and ethics, came to be known as the “Father of Medicine”, he postulated that there was blood circulating cyclically from the body through the heart to the lungs. If we look back to 3rd century B.C, Aristotle one of the greatest philosophers and scientists to ever lived, he described the heart as a three chambered organ which was the center of vitality for the body, although his statement of the heart containing the 3 chambers came to be wrong, he did manage to successfully conclude that blood vessels came from the heart. During the latter part of the 3rd century came Herophilus, who also made great strides in anatomy coming to be known as the “Father of Anatomy”, as he was one of the first to recognise the heart has four chambers. Then came Galen in the 2nd century B.C, who was also a prominent physician in Greek-Roman medicine, whose most prominent discovery about the anatomy of the heart was arteries also carry blood.

Transitioning into the Middle Ages (5th - 15th centuries) this is a time where progress in the west was slowed down due to religious beliefs, whereas Arabs in the east, translated and continued the work of notable physicians I had mentioned above. One of the most notable arabic physicians, Haly Abbas, was one of the first physicians to reject some of the work by Galen and Aristotle, and to establish the relationship between the veins and arteries, also describing the difference between them (Roberts et al. 2019).

Going forward, there are several other key contributors who have helped form our understanding of the heart, but looking back at history it has shown how different people have interpreted the heart differently, henceforth different types of progress was made in deepening our understanding of the anatomy of the heart. All this work has helped shape our modern day understanding of the heart, which I shall explain in the next section, as it is pivotal knowledge in understanding how an ECG test is conducted on the heart. The study of the anatomy of the heart is still of great interest, as heart disease surpassed infectious diseases as the leading cause of death worldwide, midway through the 20th century (Neubauer 2007).

1.1.2 Structure of The Heart

Understanding an ECG analysis requires some relevant knowledge about the structure of the heart. The heart as we know it today is a 4 chambered muscular organ, it consists of two

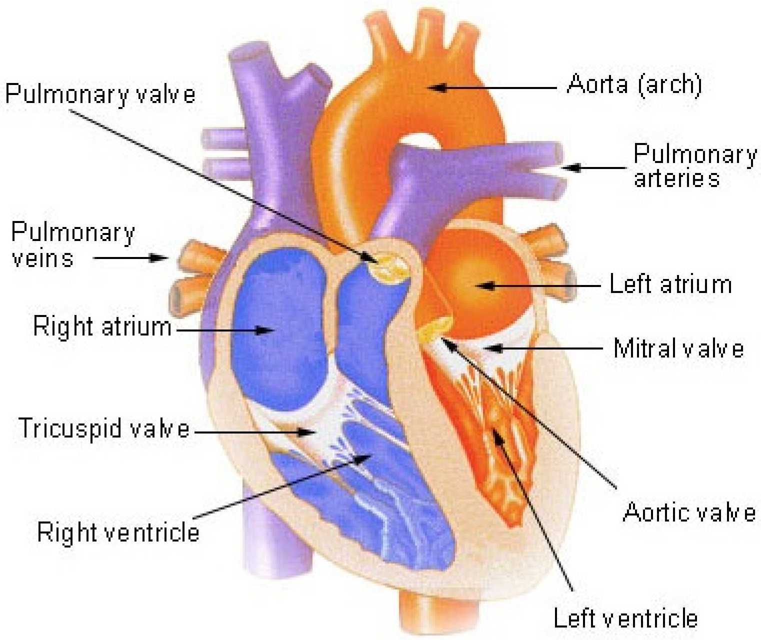


Figure 1.1: Internal View of The Heart

Atria, which are the upper chambers and two Ventricles, which are the lower chambers; the Atria receives blood from the veins, while the Ventricles pump blood out of the heart through the arteries. The left and right side of the heart is separated by a wall of muscle known as the Septum. The walls of the heart are made up of these following layers, the Epicardium, which is the protective layer is mostly made of connective tissue, the Myocardium, the muscles of the heart and lastly the Endocardium, which lines the inside of the heart and protects the valves and chambers. The atria and ventricles contract and relax in turn, producing a rhythmic heartbeat (NIH 2020).

1.1.3 Electrocardiogram

Referring back to the overall aim of automating the classification of electrocardiogram (ECG) signals for cardiac diagnosis by taking the ECG signal which is in the time series domain and converting it to the Image domain, it is pertinent that we look into, how ECG tests are conducted, the scientific process behind the analysis of the heart using ECG signals, current contributions to ECG analysis and lastly current issues with ECG analysis.

An electrocardiogram is a medical test, which is used to assess the cardiac rhythm of your heart, by measuring the electrical activity which occurs across your heart using electrodes. ECG, provides a non-invasive method of evaluating patients cardiac health. Due to its non-invasive nature it has become one of the most used tools for diagnostic purposes.

Modern day ECG tests use 10 leads which are electrical cables attached to the body, one to each limb and six placed across the chest. These 10 cables allow you to produce 12 different electrical signals covering different orientations of the heart. as some heart irregularities can only be noticed at certain orientations of the cables (Ashley 2004). If you refer to the figure 1.2 below, you can see the different signals produced by the 10 leads.

Figure 1.2: ECG lead signals

A normal heart rhythm will produce a characteristic waveform shape, which can be seen in figure 1.3. You can see there are positive and negative deflections from the baseline occurring in the ECG. The ECG waveform as we know it is split into different segments corresponding to different waves, which represent phases of the heart during one cardiac cycle . A positive deflection refers to a wave of depolarisation of cells heading towards the positive electrode or a wave of repolarization heading away from the positive electrode. A negative deflection refers to a wave of depolarisation away from the positive electrode or a wave of repolarization towards a positive electrode. Any cardiac health defects detected from the heart will cause the waveform to deviate from what is the norm (Ashley 2004).

P waves represent atrial depolarisation, this occurs just before atrial contraction is occurring. The QRS complex of ECG signals is composed of the Q wave , R wave and S wave. In most ECG signals, the Q wave will always be negative, followed by a large positive R wave, followed by a negative deflection from the S wave. The QRS complex represents ventricular depolarization, which also occurs just before ventricular contraction. Following on from the QRS complex is the T wave, this represents ventricular repolarization, when ventricular relaxation of the heart is occurring (ACLS 2020).

An important part of an ECG signal is the ST segment which refers to the part of the signal, beginning at the end of the QRS complex, to the beginning of the T wave. ECG specialists are able to determine whether the ST segment is raised above the baseline by at least 2mm, this is known as ST segment elevation myocardial infarction, which is one of the indicators the patient is at risk for a heart attack. There is also the RR interval which is the time between two successive cardiac cycles, it is measured from the peaks of the QRS complex waves, the R points. The RR interval is used to calculate the heart rate of the patient. There are many more diagnoses which can be determined from an ECG, hence why it is such a powerful tool cardiac health assessment (Ashley 2004).

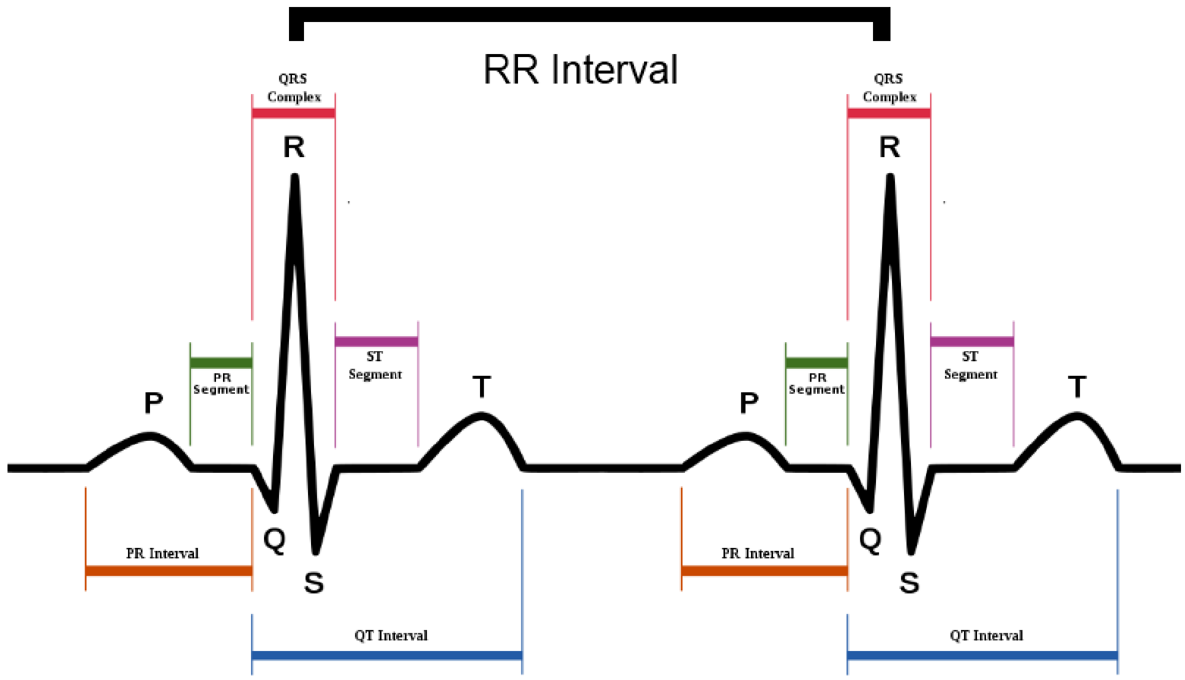


Figure 1.3: Normal ECG Waveform

From these different segments and waves, we are able to use the normal characteristic wave for comparison to determine any relevant changes. We have built up a database overtime, regarding the amplitudes and the periods for these different types of waves, so using this as reference we can see if there is any indication of deviation from the norm.

Conducting ECG tests, and reviewing the results to detect any changes, and conclude with the right possible diagnosis, is a specialised skill which takes many years of practise. So it is essential for those who conduct the test, to have an understanding of the disorders behind electrographic phenomena, to decrease the likelihood of error (AlGhatrif & Lindsay 2012).

From a machine learning perspective, extracting all these morphological features, the ECG presents and sampling the signal proves useful in building models, to discern between different cardiovascular diseases. As with most modelling tasks, feature selection has proven to be a very important part of the process.

1.2 Thesis Motivation

Using neural networks for pattern classification is a task which has been done for many years, and is an area of great interest for many different applications. Classifying ECG signals corresponding to different cardiovascular disease categories is a difficult process due to the complex nature of the signal. The ECG signals pick up noise due to power lines and other electrical interferences, hence requiring heavy preprocessing, before it can be used for the classification tasks. There is no standardization of ECG features, across the many tests which are conducted, the boundaries placed on the amplitude and time domains are not fixed, hence there has to be a way to account for this when extracting features.

Many researchers have used sampling techniques and a collection of morphological features such as the peak amplitudes of the waves which make up the ECG signals to provide features for the machine learning models. Most researchers have achieved very good scores across the metrics they have used, but one major issue concerning ECG classification is the amount of data which is available, most datasets tend to be on smaller side, and slightly unbalanced hence training of these datasets may not be representative of a larger population, and may lead to models which do not generalise as well to a larger population. The performance of the model is very dependent on the features extracted from the signal, and it needs to be as consistent as possible. Now in recent times there are new datasets, which are about ten times bigger than what was previously used most of the time, with the onset of bigger models trained on larger datasets, this can allow us to build better models, leading to more certainty about the performance of the models for this task.

My motivation for this thesis is to address the issue of producing models using smaller datasets, by using large datasets then evaluating how it generalises to smaller datasets, by fine tuning the models to a new smaller sample set. My other motivation for this thesis is to explore techniques used to convert time series data into images so we can leverage CNN ability to understand spatial awareness between pixels and and extract high level features, in hopes we can classify the images through the visual pattern they display.

1.3 Thesis Goals

The overall goal of the thesis is to be able to use unique imaging techniques which capture the essence of the ECG signal, essentially converting the ECG signal from the time series domain to the image domain. The image should capture key features of the ECG signal, in a format which allows an image classification model to extract the same type of knowledge, if we were to do feature extraction on the signal itself. These images should be able to capture the variations shown in the ECG signals, in such a way that a model can discern the images for different CVD classes. Second goal is to evaluate the techniques of transfer learning for ECG data. I will be investigating the ability of the model to learn from one dataset of patients ECG data, which includes a variety cardiovascular diseases, then applying and fine tuning this model to another dataset, with more specific issues regarding CVD which would be heart arrhythmias. So the following goals I want to investigate and achieve are:

- Convert electrocardiogram data to images, based on extracted features, to produce a cardiovascular disease image classification model.
- Apply transfer learning to the developed models and evaluate the models ability to transfer knowledge gained from one dataset of patients to another.

1.4 Thesis Outline

1.5 Dissertation Format

This template is provided to facilitate the process of writing up your dissertation while ensuring its format is consistent with requirements. In using this template, *do not*, under any circumstances, make any changes to the class file provided. In particular, do not attempt to make any changes to the title page, the statement of originality, or the copyright page.

Use of this template is not mandatory; if you choose not to use it, then you must make your final layout resemble this output as closely as possible. In particular, the textual content and layout of the title page, statement of originality, and copyright page must not be changed.

Chapter 2

Literature Review

2.1 Overview

In this section, I will explain all the relevant techniques and models, which have been applied for the task of ECG classification, while also providing critical analysis of previous works in this area. I will split the literature review into relevant sections which correspond to different parts of the developmental process of producing ECG classification models. The literature review has the following structure and sections, Dataset Collection -> Data Preprocessing -> Feature Extraction -> Feature Reduction -> Models.

2.2 Dataset Collection

In this section we will look at and evaluate some of the most prominent datasets which have been developed and used primarily for ECG analysis. As with most learning tasks, the dataset chosen is an important choice, as this is what the model will be learning from.

2.2.1 MIT-BIH Arrhythmia Database

The MIT-Arrhythmia Database is one of the oldest datasets, to be publicly available for the general public, so ECG analysis could be conducted by different research groups. The dataset was made in the mid 1970s and released to the public in 1980. The dataset was made to address the current issues in automated ECG arrhythmia classification. Progress in automated ECG arrhythmia classification was slow, due to the lack of public datasets. If you wanted to conduct research, you had to source your own dataset (Moody & Mark 2001).

The database offers researchers well defined and characterised ECG recordings, all complete with annotations. There are 48 half hour recordings included in this database, recorded using two leads. A well defined standard dataset is pertinent for progress in ECG arrhythmia classification, as there is now a standard dataset which all models can be tested on which allows for comparison between models, hence allowing us to see if there is really an improvement in automated classification. The dataset allows researchers to quantifiably evaluate whether their model offered improvements in ECG arrhythmia classification.

The database eliminates issues with models being produced on your own sourced dataset as there is no baseline to compare to. Now with a standard dataset you can discern that it was the algorithm, which is the sole reason for improvements in classification.

2.2.2 PTB-XL

The PTB-XL is the largest 12-lead ECG waveform public dataset to date, since its release in May 2020. The dataset contains 21837 clinical 12 lead-ECG records of 10 second length from 18885 patients (Wagner et al. 2020). Compared to the MIT-BIH dataset, this dataset has a variety of ECG recordings corresponding to a variety of cardiac diagnoses. The diagnostic labels for the ECG recordings are aggregated into superclasses and subclasses. This allows researchers to focus on a more general ECG classification method or a more precise ECG classification method.

2.2.3 Comparison Between the PTB-XL and MIT-BIH Database

Majority of research studies have used the MIT-BIH dataset, investigating automated ECG classification of heart Arrhythmias (Irregular heart beats). For many years this dataset has been the first choice, due to easy comparison between previous works, hence providing good benchmarks for performance. But there are some disadvantages using this dataset if we are going to continue progress in this area. The size of the dataset is very small, most models trained on this dataset are achieving high accuracy scores, this leaves a little bit of doubt due to the size of the dataset, as would it be beneficial to know how it performs on a larger sample size. Second issue with the MIT-BIH dataset, it only has ECG recordings using 2 leads, when this is compared to today's standard, this is not very reflective of current ECG examinations, as they are conducted with 12 leads. There are additional features which could be extracted from the other leads. PTB-XL, is relatively new, but given the size of the dataset, it is worth producing models developed on this dataset, as this should provide new insights into ECG classification. It features 12 leads meaning, hence there are more features which could be extracted. Lastly, the PTB-XL offers the opportunity to investigate other cardiovascular diseases (CVD). Going forward I believe models should be developed in conjunction with these two datasets. MIT-BIH can be used for comparative purposes and the PTB-XL used for scaling of performance purposes.

Dataset	Number of Recordings	Duration	Type	leads
MIT-BIH	48	30 mins	Arrhythmia	2
PTB-XL	21837	10 seconds	CVD	12

Table 2.1: Summary of ECG datasets

2.3 Data Preprocessing

This is one of the most important and mandatory steps for any workflow involving data. This is where we clean and reconstruct the data into the right format for the task which is going conducted. If not completed properly, the models developed will not learn correctly and produce wrong output (García & Luengo 2014). Essentially garbage in garbage out, this is what we want to avoid, for the whole data processing pipeline.

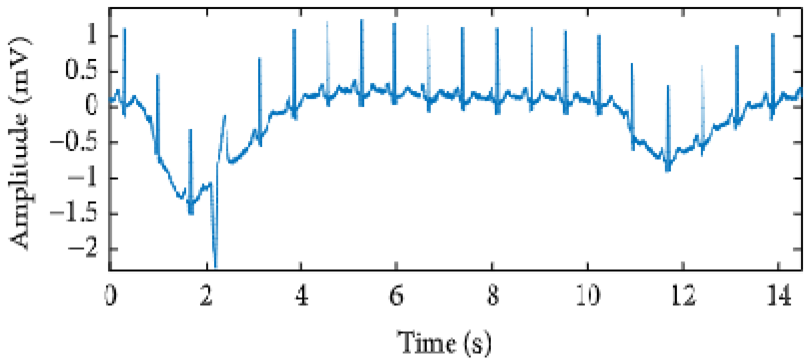
2.3.1 Issues with ECG data

The main issue regarding ECG data across the majority of research studies is the corruption of the signal due to noise, more specifically ECG signals experience the following types of noise, baseline wandering noise, alternating current power noise, muscular contraction noise, and/or electrode contact noise (Kannathal et al. 2003).

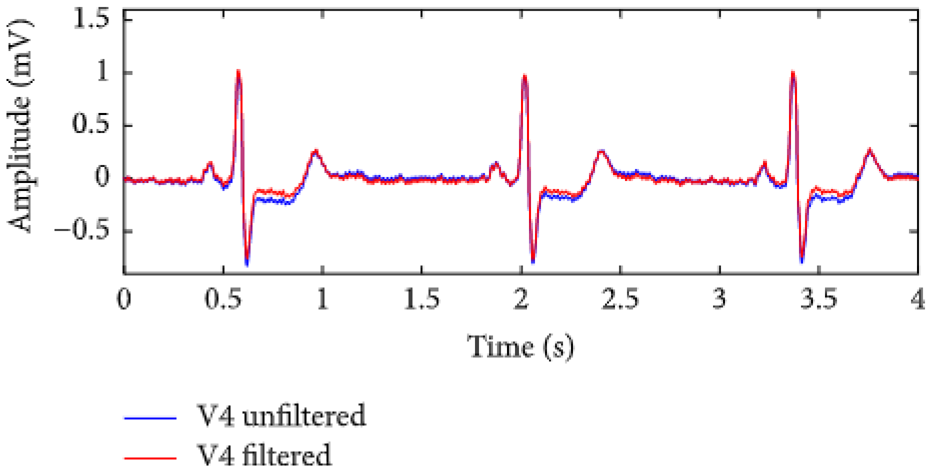
2.3.2 Baseline Wander and Noise Filter Removal Techniques

The baseline of an ECG can move unpredictably up and down during testing, when this happens this is known as Baseline Wandering (BW). BW is low frequency noise picked up by the ECG. The main cause for this noise is movement due to respiration, electrode impedance changes and patient movement (Froning et al. 1988).

The removal of BW is important, as feature extraction methods which rely on the amplitude of signals become flawed if the signal is not clean. For instance elevation or depression in the ST segment is an important ECG indicator of ischemia or myocardial infarction (Lenis et al. 2017). These small changes in the ST segment allow us to determine whether a patient can be diagnosed with these diseases. So removal of baseline wandering is essential. If you refer to figure 2.1.a you can see the baseline of the signal is not constant, making possible diagnosis extremely hard. In figure 2.1.b, there is more of a depression in the ST segment for the unfiltered signal, this is the effect noise has on ECG signals leading to misdiagnosis.



(a) Corrupt ECG signal



(b) Unfiltered and Filtered ECG Signal

Figure 2.1: Baseline Wandering ECG signals

Since baseline wandering refers to low frequency noise, there are several techniques which incorporate a high pass filter, to remove the low frequencies. One notable high pass filter technique is the Butterworth highpass filter (Chavan, Agarwala & Uplane 2008). The key thing about this filter technique, it is designed to have a frequency response as flat as possible. It is recommended by the American Heart Association to have a max cutoff frequency of 0.05Hz, because when a signal is filtered it undergoes a nonlinear phase shift causing distortion to the ECG signals. For Butterworth filtering we can overcome the effect of phase shift by filtering the signal in both directions (Lenis et al. 2017).

(Kannathal et al. 2003) used a band pass filtering algorithm to remove noise from the ECG signals, this method requires two cutoff frequencies, hence only allowing frequencies between these two values. Band pass is essential for dealing with baseline wandering noise and muscular contraction noise, which occur at two different frequency levels.

2.3.3 Wavelet and Fourier Transforms

Wavelet transforms and Fourier transforms are both techniques used to analyse signals and remove noise. The Fourier transform is a useful tool to analyse the frequency components of the signal, by converting the signal from the time domain to the frequency domain. Fourier analysis states that any periodic function can be represented or approximated by sums of sinusoidal waves. The problem with Fourier transform, we cannot locate the time at which a particular frequency occurs (Mallat 2009), if we are dealing with non-stationary signals for example ECG signals, hence making noise removal harder. The wavelet transform can decompose the signal into different coarseness levels which allows us to show the main parts of the signals and the high fluctuating parts of the signal reflecting higher detail, Noise is removed and the signal is then restructured (Mallat 2009).

(Naima & Timemy 2009) produced a comparative study showing us the effects of Discrete Fourier Transform (DFT) and Discrete Wavelet Transform (DWT) on the performance of neural network classifiers for ECG classification, across the metrics they used, accuracy, sensitivity and specificity the NN model using DWT outperformed the other NN model using DFT in terms of accuracy and sensitivity and achieved the same for specificity. Here we can conclude that DWT is a more refined technique for noise removal of non-stationary signals. They're study was one of the first to look into DWT for noise removal, and going forward multiple studies have used Wavelet Transform as common practise for removing noise from ECG signals.

2.3.4 Normalisation

(Yu & Chou 2008) recommends subtracting the mean and dividing by the standard deviation from the points collected from the ECG signal to eliminate the DC offset effect, which is a common effect when recording analog signals to digital signals, this helps prevent misclassification from signal amplitude biases due to instrumental and human error. (Wang, Chiang, Yang & Hsu 2012) also use this method referring to it as the Z score method. Most studies in this research area use this method to account for inconsistencies in the signal amplitude.

2.4 Feature Extraction

Selecting and extracting features is an important task for ECG classification, due to the high dimensionality of the ECG signals, we cannot process all points as features. Sampling techniques and specialised algorithms are used to select points from the ECG signals.

2.4.1 Morphological and Temporal Features

Morphological and Temporal features of an ECG signal, refers to the maximum amplitudes of the PQRST waves, the duration of various segments such as the ST segment, RR interval and the QRS complex etc. These morphological and temporal features are used by trained physicians for ECG analysis as indicators for different CVD diagnoses, so including such features for our models should be beneficial in the training process.

(Kannathal et al. 2003) used the QRS detection algorithm developed by (Pan & Tompkins 1985). They defined a feature based on the duration of the QRS complex, by measuring the time between the QRS onset and offset. The onset is defined as the beginning of the Q wave or R wave if Q is not present. Offset defined as the end of the S wave. Using the slopes of the waves we can determine when they start using differentiation, when they change in amplitude with respect to time is equal to zero.

(Vijayavanan 2014) who focused his solution on only using morphological features of the ECG, found the R peaks values by looking for the maximum amplitude in the signal. He also employed a windowing function to find the other peaks on both sides of the R peak. The window is traversed on both sides of R to find the max negative values, these will correspond to the Q and S points of the QRS complex. Then using the window function to find the next max peak on the left of Q to find the P point and to the right of S to find the T point. Other various segments such as ST, QT,RR and PR, were extracted using the onsets and offsets of these points.

2.4.2 Sampling

(Wang et al. 2012) focused on collecting time domain features. They focused on the R point as the focal point of the signal, and collected 200 points before and after the R point at a sampling frequency of 360Hz. This ensures all the important points such as PQRST are one of the 200 sampling points.

2.4.3 Convolutional Neural Networks

(Mousavi & Afghah 2019) used a one dimensional CNN to extract features from the ECG signal. It has been shown that 1D CNNs are adept at extracting features from signals which show a lot of variation, requiring little to no feature engineering for the classification task at hand. Due to 1D nature they have fewer parameters and require less training than their 2D counterparts, hence proven to be more suitable for real time ECG classification (Kiranyaz et al. 2019).

2.5 Feature Reduction

Feature reduction are techniques used to reduce the amount of feature while keeping the integrity of the original data, meaning the reduction in features should still explain the variance exhibited within the dataset (García & Luengo 2014).

(Nasiri et al. 2009) recommended the use of genetic algorithms for feature reduction. This is a unique method of choosing the best subset of features, from the extracted morphological features. The technique models the features as a vector of length 22, with each element being 0 or 1, to represent whether a feature is omitted or present in the chromosome. The chromosomes undergo the normal genetic algorithm process using the SVM as the fitness function to evaluate the chromosomes. This process should eventually converge to optimum state, where chromosome produces the highest fitness where in this case would be a subset of features which produce the highest classification rate in the SVM classifier.

Principal component analysis (PCA) another technique which has been used for feature reduction in a couple of studies. PCA maps a set of feature variables which are in a high dimensional spake k to a lower dimensional space p. The idea behind the PCA is that the reduced number of variables should still be able to represent majority if the variance exhibited in the high dimensional space (Castells et al. 2007).

2.6 Models

A variety of models have been developed for the task of ECG classification, where most have performed highly on the MIT arrhythmia dataset.

(Kannathal et al. 2003) produced three types of neural networks. The first model, was a dense neural network, second model was a self organising map (SOM), which is an unsupervised model, and the last was a neural network implemented with a Radial Basis Function (RBF). The models were trained on 13 morphological features which were extracted from the ECG signal. They achieved accuracies of 94.0%, 95.5% and 97.0% respectively.

(Jadhav et al. 2010) developed a modular neural network (MNN) for this task. An MNN uses subnetworks to work on different parts of the input, essentially this neural network breaks tasks into subtasks, then combines output of these subnetworks in the output layer to produce a result. The MNN only achieved an accuracy of 82.22%, they concluded that this model would require more fine tuning, to be useful in practical situations.

(Vijayavanan 2014) implemented a probabilistic neural network (PNN). This is a 3 layer network, which contains an input layer, hidden layer and output layer. The hidden layer contains K groups of hidden nodes, K would be equal to the number of classes in the dataset. All thel nodes in the hidden node group have a gaussian function, we sum the gaussian functional values from each group and choose the group, with the highest sum as the class for the input. They achieved an accuracy of 96.5% on the MIT-BIH.

(Mousavi & Afghah 2019) implemented a seq to seq model for ECG classification. They used two recurrent neural networks (RNN) to form an encoder-decoder architecture alongside an 1D CNN to extract features from the ECG data. At each time step a sequence of ECG beats is fed into the CNN which extracts the features which is then fed into the RNN which is the encoder part of the model. The decoder then takes this encoded input to produce the class as output. The advantages of this model, it is able to catch non-linear dependencies within the data. This model achieved an extremely high accuracy of 99.53% on the MIT-BIH dataset.

2.7 Summary

Many studies have shown that ECG classification is a complex task which requires a lot of processing and careful experimentation to produce robust models. Studies have shown that noise is a substantial problem in ECG signals, and has to be removed with relevant techniques, such as Butterworth Filtering or Discrete Wavelet Transform, also taking into account any phase shift which could attenuate the signal. Feature extraction is another important task, whether you decide to use morphological, temporal, or a variety of sampling points, the selection of the features will influence the performance of the model on the task. Data reduction has to be considered if there are too many features, as the complexity would be too high and cause slow training times. A variety of models had been developed, from the studies I have researched, the models which used the most relevant features, and combined with fine tuning experienced the best results. I can conclude that the seq to seq model (Mousavi & Afghah 2019) implemented, proved to be a model which showed real improvement in how we extract features, which also takes into account the time dependant nature of the features, I feel this is the direction ECG analysis is going in using CNNs to obtain high level abstract features.

Chapter 3

Methodology

3.1 Overview

The first half of this chapter will focus on the methodology behind the deep learning approach to classifying ECG signals using convolution neural networks, trained on 2D images of ECG signals which are based on temporal and morphological features of an ECG signal. The second half will focus on the application of transfer learning of the model which is trained on the PTB-XL dataset which is focused on cardiovascular diagnosis, to the MIT-BIH arrhythmia data, which is used for classifying different types of arrhythmia from the ECG signals.

3.2 Dataset Collection

The PTB-XL which I described in chapter 2 for the Literature Review is the main dataset, used to train the CNN models to meet the first aim of this thesis. The PTB-XL contains 21837 12-lead ECG signals of 10 second length. The dataset is annotated by two trained cardiologists, to cover diagnostic, form and rhythm statements. For this thesis, only the ECG signals which have been annotated with a diagnostic class such as Myocardial Infarction , have been selected. The MIT-BIH dataset will be used to evaluate the transfer learning method developed using the model from the first part. I will also train and test the model on the MIT-BIH dataset, for comparison purposes, to evaluate how it measures up against other state of the art models for ECG arrhythmia classification.

3.3 Data Preprocessing

3.3.1 Noise Removal With Discrete Wave Transform

When ECG signals are recorded they are more often than not, corrupted with noise such as baseline wandering, or powerline interference noise. The noise induced on the signals will affect any ECG practitioner or automated ECG classification systems, decision making when making a diagnosis. The key indicators within an ECG signal become blurred due to noise. This is why it is imperative the right noise removal algorithm is used to remove the largest amount of noise while maintaining the integrity of the signal.

The chosen noise removal method I have used is the discrete wavelet transform (DWT). The advantages of this technique over classical filter techniques is that the attenuation of the signal due to phase shift is minimised, whilst also maintaining a high signal to noise ratio. (Naima & Timemy 2009) showed the improvement in accuracy for ECG classification using DWT vs Discrete Fourier Transform (DFT), as the focus of the study was comparing the effects of these two signal analysis techniques. DFT converts a signal from the time domain to the frequency domain, providing information on all the constituent frequencies present within the signal alongside their amplitudes. The downfall of the DFT, is that it does not have the ability to tell the time at which these particular frequencies occur, hence it is only an ideal method for time varying signals that are considered stationary, this is when the mean and variance of the signal is constant. In short, it cannot provide frequency information for a specific moment in time (Kehtarnavaz 2008). This led to the formulation of the Short Time Fourier Transform (STFT), to account for the poor time resolution of DFT. The STFT slides a window over the signal which corresponds to a time period, and performs a fourier transform at each time interval, providing frequency information for that time period. The downfall of this method is that, the smaller the window is, you attain a higher time resolution but a lower frequency resolution and vice versa. So there will always be a tradeoff through the STFT method.

The wavelet transform (WT) overcomes the shortcoming of the STFT. The WT is able to analyse a non-stationary signal and output different frequencies at different resolutions; this is known as multiresonal analysis (Singh & Tiwari 2006). With the wavelet transform we are able to achieve good time resolution and poor frequency resolution at high frequencies, while also achieving good frequency resolution and poor time resolution at low frequencies (Misiti et al. 2004). This proves pivotal for noise removal, as noise is a high fluctuating frequency superimposed on to the original signal. WT allows local analysis of the signal. A wavelet is a waveform with limited duration and has a mean of a zero, they are considered as asymmetrical waveforms (Misiti et al. 2004).

Compared to Fourier transform, wavelet transform results in splitting a signal into a shifted and scaled version of the mother wavelet. To understand the wavelet transform formula we must look at the fourier transform first.

$$F(\omega) = \int_{-\infty}^{\infty} f(t)e^{-i\omega t}dt \tag{3.1}$$

The formula 3.1 which is the Fourier transform represents the sum over time of the signal f(t) multiplied by a complex exponential.

$$F(\tau, \omega) = \int_{-\infty}^{\infty} f(t)w(t-\tau)e^{-i\omega t}dt \tag{3.2}$$

The formula 3.2 is the Short Term Fourier Transform has the addition of a window function and the addition of the parameter tau τ which represents the translation factor.

$$F(\tau, s) = \frac{1}{\sqrt{|s|}} \int_{-\infty}^{\infty} f(t)\Psi^{*}\left(\frac{t-\tau}{s}\right) \tag{3.3}$$

The formula 3.3 is the Wavelet Transform and is similar to the Short Term Fourier Transform function. Instead of the complex exponential function and window function this is replaced with the Psi ψ function known as the wavelet function, the wavelet now acts as the window. The S is now our scale parameter. Depending on the size of S, this can shrink or enlarge our wavelet, small values of S are better at resolving higher frequency components of a signal, where larger values of S are better at resolving the low frequency components of a signal.

The output of the wavelet transform is known as the wavelet coefficient. The wavelet coefficient represents how well a certain part of the signal correlates with the mother wavelet, so when there is a good match between signal and wavelet we are going to get high wavelet coefficient values.

(Singh & Tiwari 2006) stated, making the right choice for the mother wavelet is important as this affects how well defined, the wavelet coefficients are going to be. These coefficients then correspond to different parts and frequencies within the ECG signal. The decision for the mother wavelet can be done via a quantitative route, where we find a mother wavelet which is highly correlated with the ECG signal. (Ngui, Leong, Hee & Abdelrhman 2013) states that the optimal choice of wavelet can be done by a qualitative method. The qualitative method is simply choosing the wavelet which is most similar to the signal which is going to be decomposed. For a simplistic reason I went down this approach. The mother wavelet I chose, was the Daubechies wavelet 4.

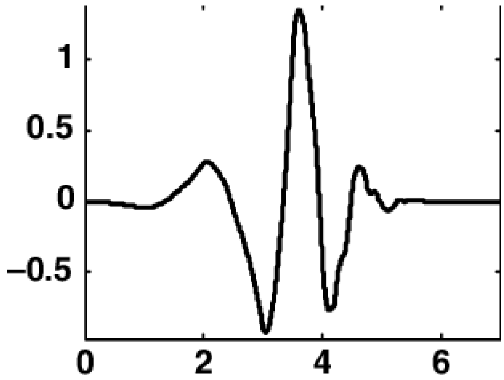


Figure 3.1: Daubechies Wavelet 4

The db4 shared a lot of similarities to the ECG waveform, where you can see key features such as the P wave, T wave and the QRS complex. Denoising an ECG signal using wavelet decomposition involves finding the right waveform, finding the wavelet coefficients, then using a thresholding technique to suppress those coefficients which were generated by noise frequencies and then finally reconstructing the waveform from the coefficients which are left.

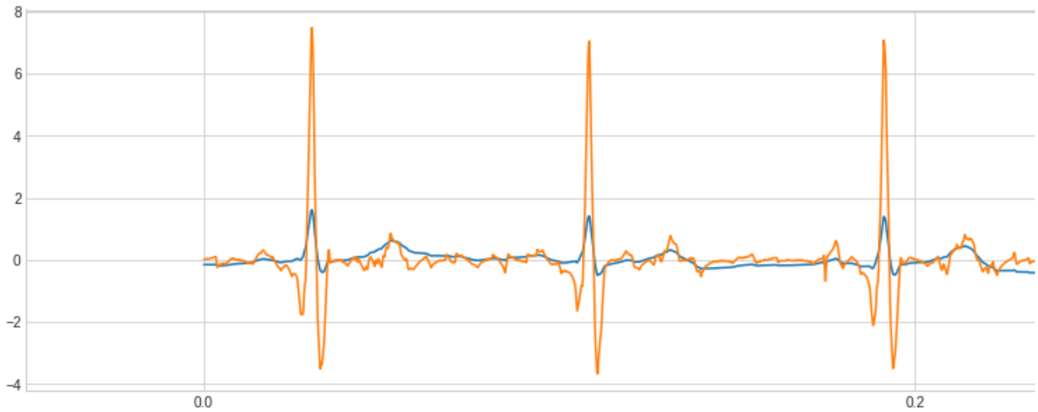


Figure 3.2: Denoised ECG waveform

Figure 3.2 is the denoised ECG waveform alongside its previous self. As you can see, the ECG signal is more aligned with baseline than before, this is important as if it wasn't, it would affect feature extractions based on amplitude. The whole signal became more pronounced, as we have removed all those signals which were superimposed on to the ECG signal.

3.3.2 ECG Normalisation

After cleaning the ECG signals, they were normalised between 0 and 1. The method used for normalisation was the Z-score method. This is when you subtract the mean and divide by the standard deviation. This causes the mean to centered around zero with a standard deviation of 1(Wang et al. 2012). This ensure there are no amplitude biases when the model is being trained, as most models have activation functions which are capped between 0 and 1.

3.3.3 ECG segmentation

The ECG signals were segmented into 1 second segments, as these 1 second segments contained a whole cardiac cycle within them. I had to consider the reduction in size as the ECG signals were sampled at 500hz, so in 10 seconds, that is a recorded amount of 5000 sampling points.

3.4 Feature extraction

3.4.1 Temporal Features

Feature extraction for ECG signals was focused on extracting the morphological and temporal features of the signal. The morphological features were extracted from each 1 second segment by locating the R peak within the signal which typically had the largest amplitude.

To find the RR interval I went back and the split signals into 2 second intervals. I sliced into the segment and found the first peak. Then found the other peak with other half. using the difference between two peaks, I could find the RR interval. As the signal were sampled at 500Hz for 10 seconds, the time between each point is 0.002 hence the difference between the two indices is multiplied by this value to find the RR interval. All RR intervals were normalised within the signals they came from, as heart rates can vary from patient to patient, and the RR interval needs to be accurate if it is going to be used as a feature. The code below was the function I made to do this. The RR interval was appended on to the onto the corresponding 1 second

ECG segment, to make the full feature set.

```
def get_RR(segment):
    segmentList = segment.tolist()
    max1Index = segmentList[:400].index(max(segmentList[:400]))
    max2Index = segmentList[max1Index+250:].index(max(segmentList[max1Index+250:]))
    trueMax2Index = max1Index + 250 + max2Index
    RR_interval = (trueMax2Index - max1Index)*sample_interval
    return RR_interval
```

3.4.2 Morphological Features

These features encompass all the points surrounding the QRS complex. It is a common technique to sample points around the R peaks, as this will gather the PQRST points which are key focal points of an ECG signal. As the segment which I extracted where focused on the R peak, I selected all values within 1 segment, which came to 500 sampling points.

3.4.3 Feature Reduction

I used Principal component analysis (PCA) as I wanted two datasets, so I could compare the effect of data reduction on the models learning. Principal Components maps a set of features in a higher dimension to a set of feature in lower dimension. The set of features in the lower dimension are known as principal components. Principal Components main aim is to to explain the variance exhibited within data, using smallest amount of component it is allowed to use. I used PCA to convert the 500 point ecg signal into 256 components, so I could the signal as 16x16 greyscale image.

3.5 Creating ECG Feature Images

This is the main part of the thesis, as this goes back to one of the aims, which is representing the ECG signal, as an image of its features. So I created two datasets, one where PCA has been applied to ECG segments, and one where it hasn't. The PCA segments have a length of 256 whereas the non PCA segments have a length of 501, I reshaped all the signal as 2D images with the shape (16x16) and (23x23). This was done using the numpy.reshape() function. You can visually see the R peak in figure 3.3 (a) as it has the largest amplitude therefore appears white. Figure 3.3 B shows how the features have bee spread across the 256 components. All these images where made for every ECG segement. The values are already between 0 and 1 so normalisation is needed.

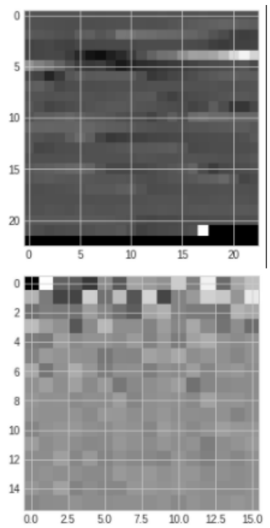


Figure 3.3: (a) ECG feature image no PCA. (b) ECG feature image with PCA

3.6 Cleaned Dataset Summary and Finetunig

From the PTB-XL 10,000 ECG signals were extracted, this came down to 9868 as some signal were removed as they did not have any cardiac diagnosis labels. Splitting the signal into 1 second segments produced 98,680 signals. The class distribution of the signals is down below.

- Normal ECG (NORM) : 47450
- Conduction Disturbance (CD): 21610
- Myocardial Infarction (MI) : 22270
- ST/T change (STTC): 23540
- Hypertrophy (HYP): 11930

From this class distribution Normal ECG signals are clearly dominant, this is a problem when training models as this can lead to huge biases within the model towards that top heavy class. So balancing the dataset was required. As this is a multi label classification problem, I couldn't just drop labels. I made sure to count the number of time a label has appeared in the dataset once it reaches the amount Hypertrophy has since it has the least, no labels of that type could be added to the dataset.

```
def downsample(images,labels):
  dic = {'HYP' :0, 'NORM' : 0, 'MI' : 0, 'STTC' : 0, 'CD' : 0}

  new_images = []
  new_labels = []

  for i in range(len(labels)):
    temp = dic
    temp1 = new_images
    temp2= new_labels
    if len(labels[i])>1:
      num = len(labels[i])
      for label in labels[i]:
        if dic[label] == 11930:
          dic = temp
          new_images = temp1
          new_labels = temp2
          continue
        else:
          dic[label]+=1
          if len(labels[i]) == num:
            new_images.append(images[i])
            new_labels.append(labels[i])
            num -= 1
    else:
      for label in labels[i]:
```

```
        if dic[label] == 11930:
          continue
        else:
          dic[label]+=1
          new_images.append(images[i])
          new_labels.append(labels[i])

  return new_images, new_labels, dic
```

The function above is quite lengthy but it successfully down samples the dataset so all labels appear 11930 times with a corresponding image as well. The labels of the dataset were all one hot encoded to represent a different class. This is important as the model would output a vector of length 5 of different probabilities for different classes, so we can apply a threshold which pushes the prediction to 1 or 0 determining the diagnosis for ECG feature image.

3.7 Models

I created 2 CNN models. One is shallow and the other is Deep. I wanted to compare the performance and see if there is any comparative difference.

3.7.1 Software

Both models were trained in Google Colab on their Tesla K80 GPUs. The neural network architectures was composed in tensorflow Keras using Python.

3.7.2 Neural Network Topology

Both models were CNN models connected to a fully dense neural network.

- The deep model had 11 convolutional layers including 4 MaxPool2D layers. Followed by 2 layer network, with an output of 5, as there is 5 classes. A sigmoid activation function was used in the outer layer instead of a softmax function because in multi-label classification, we are predicting probabilities for different classes independant of the classes.

- The shallow model was similar, but instead had 7 convolutional layers with 3 MaxPool2D layers.

3.7.3 Neural Network Training

Both models were trained for 10 epochs each. They were both trained on the PCA images and the NON PCA images. The Train Test Val split was 70/15/15.

Chapter 4

Results and Discussion

4.1 Results

4.1.1 Shallow VS Deep Models

	PCA Images				Non PCA Images			
Model	Train acc	Val acc	Train loss	Val loss	Train acc	Val acc	Train loss	Val loss
Shallow	0.3387	0.3387	0.6063	0.6062	0.3387	0.3387	0.6063	0.6061
Deep	0.3387	0.3387	0.6063	0.6062	0.3387	0.3387	0.6063	0.6063

Initial training of both models, showed poor performance, across the metrics used. Neither managed to climb above 40 percent for accuracy on a balanced dataset for both PCA image and Non PCA Images and maintained a very high loss. Using the unbalanced dataset they both manged to attain an accuracy 47 percent but the result is not significant due to how many normal ECG segment there are in the dataset. It is very clear from the results that both models struggled to extract any meaningful features from the ECG feature images. I decided to proceed with the deeper model for fine tuning, as it can attain higher abstraction levels in terms of features.

4.1.2 Adjusting The Learning Rate - Deep Model

Investigating whether learning rate would have positive influence on the model deep is where my went investigation next. decreasing learning rate (LR) from the default LR of the Adam optimiser which is 0.01, made no change only very minuscule changes occurred. Increasing the LR made the performance worse for the model, and did not administer any improvements.

	PCA Images				Non PCA Images			
LR	Train acc	Val acc	Train loss	Val loss	Train acc	Val acc	Train loss	Val loss
0.1	0.2685	0.3387	0.6205	0.6292	0.2637	0.2784	0.6275	0.6110
0.01	0.3387	0.3387	0.6063	0.6062	0.3387	0.3387	0.6063	0.6063
0.01	0.3387	0.3887	0.6062	0.6063	0.3387	0.3387	0.6062	0.6061
0.001	0.3387	0.3387	0.6062	0.6063	0.3387	0.3387	0.6062	0.6061

4.1.3 Test Results - Deep Model

	PCA Images			Non PCA Images		
LR	Precision	Recall	Accuracy	Precision	Recall	Accuracy
NORM	0.0	0.0	69.8	0.0	0.0	72.1
STTC	34.3	1.0	52.1	34.3	1.0	34.3
MI	29.9	29.0	29.9	0.0	0.0	70.1
HYP	0.0	0.0	73.4	0.0	0.0	73.4
CD	30.0	1.0	30.0	30.0	1.0	30.0

The metrics used to assess the performance of the mode are:

$$Precision = \frac{TP}{TP + FP} \tag{4.1}$$

Precision is the percentage of the results which are relevant.

$$Recall = \frac{TP}{TP + FN} \tag{4.2}$$

Recall is the percentage of results which were accurately classified.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{4.3}$$

Accuracy is the percentage of correctly classified data points for TP and TN.

The Recall and Precision for most classes is consistently, not high performing on the test set. But one surprising insight I can extract from these results is that although the model may not be that good at predicting the correct class, it has achieved some high accuracy scores when it come to predicting True Negative, but this is also biased in the sense that is easier to say no than yes and achieve a high accuracy score so it is not truly a reflection of the models ability to discern between classes.

4.2 Discussion

4.2.1 Overview

The CNN image classification models, have not performed as well as I hoped on the ECG Feature Image Dataset I made. There are a few main areas which need to be discussed and evaluated towards aim of this thesis such as the Features and so on.

4.2.2 Features

Extracting sequential features from an ECG signal, may not have been the best way to represent the features of an ECG. A 2D CNN will not follow the signal, one amplitude value at a time. The CNN summarises the information into a new smaller information space based on the size of the kernel. The CNN will extract pattern if they are there, temporal sequential information which is important for the ECG is lost, as pattern in ECG are 1 dimensional, hence it would have been more appropriate to have used a 1D CNN. Convoluting different parts of the signal would cause the model lose track of the temporal information within the ECG signal. I feel this is where I went wrong in the feature extraction part of this pipeline. The feature extraction should of been more focused on collecting scalar and statistical features from the signal. This would have caused the ECG feature images to be based less on the temporal sequence of the features, hence find be able to see pattern vertically and horizontally. More features such as the length of time for the different segment should have been incorporated with the feature image.

4.2.3 Difficulties

Difficulties such as finding the perfect point to segment the ECG signal for feature extraction. This was one of the problems which was present in RR interval extraction. There would sometimes be 3 R peaks within one segment leading to confusion or false results for the RR interval. In those cases I had to result to the national average for the RR interval for that specific cardiac diagnostic class.

Another difficulty is having to balance the data in such a way that there is enough representation for signals which have multiple diagnoses (labels). The labels would prove as an extra learning point for the models to produce multiple diagnoses. Single label diagnoses may not always be case for all patients.

Producing the images in such a way which would be beneficial, and highlight the peculiarities within the ECG signal. Sequentially producing an image has proven to be very ineffective in the task of ECG cardiovascular disease diagnosis classification.

Chapter 5

Conclusion

5.0.1 Meeting The Aims

The aims was to produce a model which could detect and classify feature images, of ECG signals. This has not been successful, due to the nature of the data being fed into the CNN model, but has opened opened up new avenues in how this problem should be approached. The models produced failed to recognise accurately the cardiac diagnosis labels which had been placed on the ECG signals. The classification rate, would not make this model a suitable, if it were to be used in primary care, hence much more fine-tuning of ideas are needed. The other aim was to investigate transfer learning, if this proposed model was successful, given its ability mow it would not be a suitable model, in which we would be able to leverage and use the weights learned for fine-tuning on another task. Temporal sequences would be much better suited for 1D CNN's than 2D. So I can now conclude knowing that fact in regards to ECG sequences

5.0.2 Future Work

As mentioned in the discussion, future work would involve looking into more scalar and statistic based features for ECG signals, to produce the feature image of the ECG signal, as this would have less emphasis on the temporal aspects of the ECG. To follow on from that point, I would like into include other key indicator physicians use, other the ECG to determine your risk for cardiovascular diseases such as, alcohol consumption, cholesterol and BMI. If images are generated which are like info cards, with different health statistics about a patient we could leverage, a models ability to discern patterns and determine a patients risk a for particular type of disease. Summarising patient data in way which is not text, which machine learning can leverage, would be an area of great interest

Bibliography

ACLS (2020), 'The basics of ecg'.
URL: `https://www.aclsmedicaltraining.com/basics-of-ecg/`

AlGhatrif, M. & Lindsay, J. (2012), 'A brief review: history to understand fundamentals of electrocardiography', *Journal of community hospital internal medicine perspectives* **2**(1).

Ashley, E. (2004), *Conquering ECG*, Cardiology Explained, Remedica.

BHF (2018), 'British heart foundation: Early detection drive could prevent 115000 cases of heart and circulatory disease over next decade', British Heart Foundation Article.

Castells, F. et al. (2007), 'Principal component analysis in ECG signal processing', *EURASIP Journal on Advances in Signal Processing* **2007**(1).

Chavan, M. S., Agarwala, R. & Uplane, M. (2008), 'Suppression of baseline wander and power line interference in ecg using digital iir filter', *International journal of circuits, systems and signal processing* **2**(2), 356–365.

Froning, J. N. et al. (1988), 'Problems and limitations of ecg baseline estimation and removal using a cubic spline technique during exercise ecg testing: Recommendations for proper implementation', *Journal of Electrocardiology* **21**, 149–157.

García, S. & Luengo, J. (2014), 'Introduction. in: Data preprocessing in data mining', *Intelligent Systems Reference Library* **72**.

Hall, J. (2010), *Guyton and Hall textbook of medical physiology*, 12 edn, Cambridge Universty Press.

Hopkins, J. et al. (2010), 'Quality indicators for the prevention of cardiovascular disease in primary cares', *Canadian family physician Medecin de famille canadien* **56**(7), e255–e262.

Jadhav, S. M. et al. (2010), 'ECG arrhythmia classification using modular neural network model'.

Kannathal, N. et al. (2003), 'Classification of cardiac patient states using artificial neural networks', *Exp Clin Cardiol* **8**(4), 206–211.

Kehtarnavaz, N. (2008), Frequency domain processing, *in* 'Digital Signal Processing System Design', Elsevier, pp. 175–196.

Kiranyaz, S. et al. (2019), '1d convolutional neural networks and applications: A survey'.

Lenis, G. et al. (2017), 'Comparison of baseline wander removal techniques considering the preservation of ST changes in the ischemic ECG: A simulation study', *Computational and Mathematical Methods in Medicine* **2017**, 1–13.

Mallat, S. (2009), *A Wavelet Tour of Signal Processing*, Elsevier.

Misiti, M. et al. (2004), *Wavelet Toolbox*, 3 edn, MATLAB.

Moody, G. & Mark, R. (2001), 'The impact of the mit-bih arrhythmia database', *IEEE Engineering in Medicine and Biology* **20**(3), 45–50.

Mousavi, S. & Afghah, F. (2019), Inter- and intra- patient ecg heartbeat classification for arrhythmia detection: A sequence to sequence deep learning approach, *in* 'ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)', pp. 1308–1312.

Naima, F. & Timemy, A. (2009), *Neural network based classification of myocardial infarction: a comparative study of Wavelet and Fourier transforms*, Pattern Recognition.

Nasiri, J. A. et al. (2009), 'Ecg arrhythmia classification with support vector machines and genetic algorithm', pp. 187–192.

Neubauer, S. (2007), 'The failing heart–an engine out of fuel.', *The New England journal of medicine* **356 11**, 1140–51.

Ngui, W. K., Leong, M. S., Hee, L. M. & Abdelrhman, A. M. (2013), 'Wavelet analysis: Mother wavelet selection methods', *Applied Mechanics and Materials* **393**, 953–958.

NIH (2020), 'Seer training modules, anatomy and physiology'.
URL: `https://training.seer.cancer.gov/anatomy/cardiovascular/heart/structure.html`

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

Phoenix, C. (2020), 'The importance of the early detection of cardiovascular disease', Cardio Phoenix Article.

Roberts, W. et al. (2019), 'Across the centuries: Piecing together the anatomy of the heart', *Elsevier* .

Singh, B. N. & Tiwari, A. K. (2006), 'Optimal selection of wavelet basis function applied to ECG signal denoising', *Digital Signal Processing* **16**(3), 275–287.

Vijayavanan, M. (2014), 'Automatic classification of ecg signal for heart disease diagnosis using morphological features'.

Wagner, P. et al. (2020), 'Ptb-xl: A large publicly available ecg dataset', *Scientific Data* .

Wang, J.-S., Chiang, W.-C., Yang, Y.-T. C. & Hsu, Y.-L. (2012), An effective ECG arrhythmia classification algorithm, *in* 'Bio-Inspired Computing and Applications', Springer Berlin Heidelberg, pp. 545–550.

WHO (2020), 'World health organisation: Cardiovascular diseases (cvds)', `https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)`.

Yu, S. & Chou, K. (2008), 'Integration of independent component analysis and neural networks for ECG beat classification', *Expert Systems with Applications* **34**(4), 2841–2846.