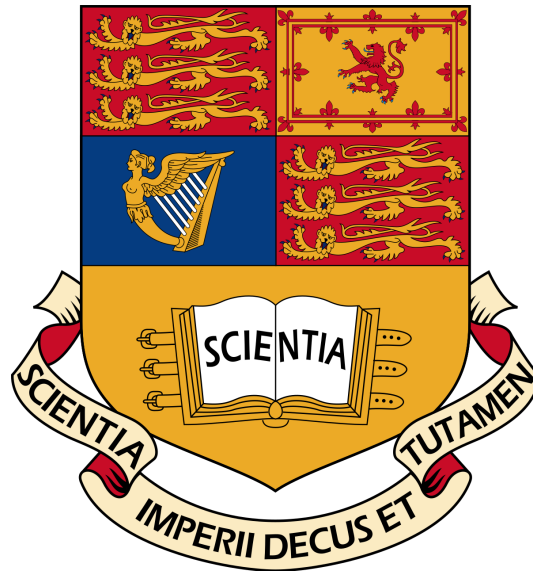


Imperial College London

Department of Electrical and Electronic Engineering

Final Year Project Interim Report 2022



Project Title: **Extraction of blood pressure from ECG and PPG signals**

Student: **Arijit Bhattacharyya**

CID: **01496199**

Course: **MEng Electrical and Electronic Engineering**

Project Supervisor: **Professor Esther Rodriguez Villegas**

Co Supervisor: **Dr Zaibaa Patel**

Second Marker: **Dr Christos Bouganis**

Contents

1 Introduction

1.1 Project Outline and deliverables

This is the Interim Report for my Final Year Project, which is titled "Extraction of blood pressure from ECG signals". The report aims to present the initial background research and early implementation results. The report will contain the following sections as suggested in the project guide ?:

- **Project Specification.** This chapter will establish the expected deliverables of this project. The chapter will also describe the context of the underlying problem that the project will try to solve.
- **Background.** This chapter will first introduce the medical background, signal processing theories and machine learning principles needed to understand the project. A literature survey will be taken both to justify the proposed BP estimation method and for the justification of the use of wearable technology in future implementations.
- **Implementation.** This chapter will describe what work has been done in regards to the chosen method for estimating blood pressure. In addition, a mathematical overview of the chosen implementation will be discussed in detail. Finally, the justification for using Python as the sole programming language for this project will be given.
- **Evaluation plan.** This chapter will detail how the project deliverables will be evaluated. In addition, the chapter will show the estimated timeline and possible extensions.
- **Ethical, Legal and Safety Considerations.** This chapter will look if there are any ethical legal or safety concerns both in regards to the project as is, but also briefly with potential extensions.

1.2 Context

Cardiovascular disease is one of the main causes of death around the world. High blood pressure (BP), which is also known as hypertension, is a common condition which can be a cause of cardiovascular disease ?. According to the World Health Organization (WHO), the mortality rate due to hypertension is 9.4 million per year and it causes 55.3% of total deaths in cardiovascular patients ?. If hypertension is detected early and prevented, this will greatly lower the number of deaths associated with cardiovascular diseases ?.

Recent developments in technology have made wearable sensors, such as Electrocardiogram (ECG) and Photoplethysmography (PPG) sensors significantly more popular in today's world. These sensors provide real-time 24 hour monitoring of the human bodily function. Hence there is great potential in using these sensors to diagnose medical conditions, such as hypertension, in real-time, thus helping to save lives ?. Ambulatory BP monitoring is seen as a promising method for detecting early symptoms of hypertension ?. There is a lot of existing research to predict ambulatory BP using methods which are cuff-less, continuous and non-invasive ?. Hence wearables are seen as a viable option for this. ECG and PPG sensors have been discovered to be a potential estimator of blood pressure that cause minimal harm to patients compared to existing cuff-based methods ? ?.

The aim of this project is to implement and evaluate the different techniques that can be used to measure Ambulatory BP from ECG and PPG signals. The aims are to seek out implementations that have minimal computational complexity, whilst maintaining accuracy.

2 Background

In this chapter, I will highlight all the necessary theory required to understand the basis of this project.

2.1 Hypertension

The heart can suffer from a variety of diseases and pathologies. Low blood pressure, or hypotension, has the potential to cause a lack of oxygen flowing to the brain and other organs, causing shock ?. Whilst hypotension is a serious issue, hypertension has been identified by the World Health Organization (WHO) as the most significant risk factor for cardiovascular diseases ?. According to the 2017 American Heart Association guidelines for hypertension, the risk of developing stage two hypertension, ≥ 140 mmHg systolic or ≥ 90 mmHg diastolic is almost 90% ? (see Table 1 below). Over 20% of adults have hypertension and its complications cause a major number of diseases, including heart attacks, strokes and heart failure. If hypertension is not diagnosed and properly treated it can even cause death ?.

Hypertension or high blood pressure (BP) is where blood continues to exert more and more pressure on the arterial walls. One particular disease linked to hypertension is hypertrophic cardiomyopathy, as indicated in Figure ??,

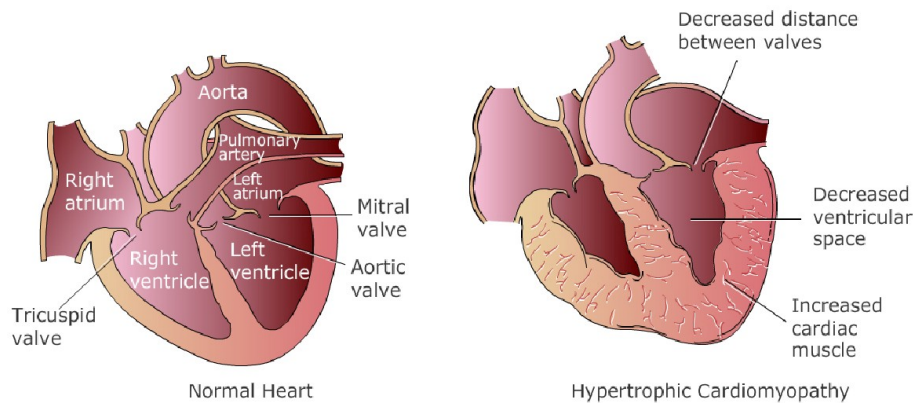


Figure 1: The effects of hypertension on the heart ?

Hence it is clear that hypertension is one of the largest motivating factors for this project.

2.2 Blood Pressure measurements

Blood pressure (BP) is the force of the blood pushing against the arterial walls as the heart pumps blood. It is measured in millimeters of mercury (mmHg) ?. BP is measured in terms of systolic blood pressure (SBP) and diastolic blood pressure (DBP). These values are the maximum and minimum pressure values of an arterial pressure tracing during a cardiac cycle respectively ? ?.

BP has oscillations or pulses that mirror the oscillatory nature of the heart. The blood is propelled during systole, also known as heart contraction, and the blood is rested during diastole, known as heart relaxation.

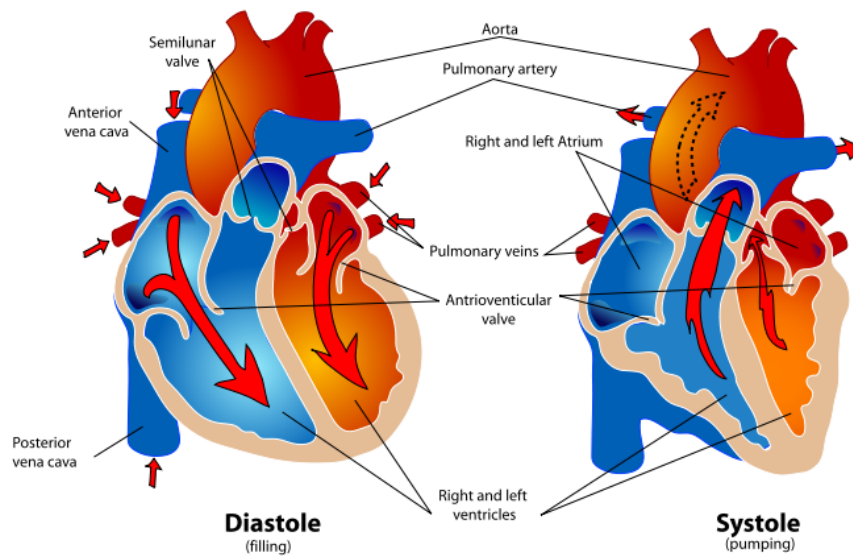


Figure 2: Visualisation for SBP and DBP ?

There are two conventional methods for measuring blood pressure (BP). These are invasive and non-invasive methods.

Firstly, a description on invasive methods. The most popular form of invasive BP measurement is catheterization [1]. Invasive BP measurements are continuous and the most accurate from heartbeat to heartbeat. As a result these measurements are recognised as the gold standard internationally [2]. However, this method is usually restricted to hospitals, as medical supervision is required [3]. In addition, this method poses several health risks, including bleeding and infection. As a result, invasive measurements are only utilised for critically ill patients in intensive care units and for use during surgery [4].

Now a description on non-invasive methods. The gold standard for BP measurement is the use of a cuffed sphygmomanometer. Cuff-based methods provide BP measurements without any major side effects as opposed to BP measured invasively [1]. However, patients will feel uncomfortable with long term monitoring due to the painful cuff inflation which interrupts the regular blood flow [2]. In addition, these methods can only measure BP intermittently with intervals between measurements greater than at least two minutes. These devices are too cumbersome to wear during measurements. Also, it has been found that over three in ten home BP monitoring cuffs have produced inaccurate results [3].

As a result, the existing invasive and non-invasive BP measurement techniques are not feasible for an implementation involving continuous ambulatory BP monitoring [?]. Hence, after having assessed the viability of all aforementioned methods, it is clear that it is difficult for these methods to be integrated with wearable technologies, which continue to gain popularity in commercial sectors and clinical practice [?].

Blood pressure classification	BP (mmHg)	
	Systolic	Diastolic
Hypotension	≤ 90	And ≤ 60
Normal	90-119	And 60-79
Prehypertension	120-139	Or 80-89
Stage 1 hypertension	140-159	Or 90-99
Stage 2 hypertension	≥ 160	Or ≥ 100
Isolated Systolic hypertension	≥ 140	And < 90
Hypertensive crisis	≥ 180	Or ≥ 110

Table 1: Categories of blood pressure in adults ? ?

2.3 Ambulatory Blood Pressure (ABP)

ABPM is when BP is measured as the patient moves around, and it allows patients to still live their normal daily lives ?. It has been classed as the gold standard for detecting and diagnosing hypertension and also assessing BP values over a 24 hour period ?. ABPM provides data on several important and unique parameters ?. This data can explain how changes in your BP may correlate with your daily activities and sleep patterns ?. Conventionally, ABP is monitored by using a cuff attached to a portable device which is worn on the patient's waist ?. The 24-hour measurement duration is what makes this method a powerful tool to be used as a part of wearable technologies, which will be discussed later in Chapter 3. As mentioned previously, ECG and PPG signals have the potential to be used as an estimator for BP, and they will both be discussed now in more detail.

2.4 Electrocardiogram (ECG) signals

ECG signals provide an overview of the electrical impulses occurring in the heart ?. Electrical changes in the heart conduct through the body and are received at skin level. The record of these electrical fluctuations during the cardiac cycle is called the Electrocardiogram (ECG) ?. The signals are recorded by measuring the electric potential difference by placing electrodes across the heart of an individual ? ?. These electrodes are connected to the ECG machine with recordings from 12 different places on the body, which is known as the 12-lead ECG. The standard EKG leads are I, II, III, aVF, aVR, aVL, V1, V2, V3, V4, V5, V6. Leads I, II, III, aVR, aVL, aVF are classed as the limb leads and the others are precordial leads ?.

The QRS complex of an ECG signal is detailed in Figure ???. This complex is first created through the generation of the electrical impulses from the heart. These signals then move along the electrical highway and as a result cause the ventricles to contract and pump oxygenated blood into the arteries. Physically, this whole describes the QRS complex ?.

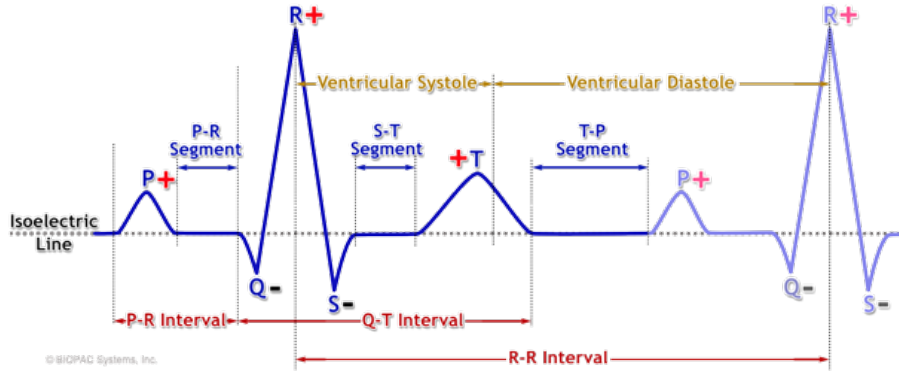


Figure 3: Structure of an ECG signal

2.5 Photoplethysmography (PPG) signals

Photoplethysmography (PPG) measures the blood volume changes per pulse. It is an optical and non-invasive technique that can determine a wide range of medical values, including an estimate for BP ?. Physically, the PPG signal is acquired by measuring the optical signal transmitted through or reflected from the subject's tissue ?. The PPG sensor consists of two components. The first component is an Light Emitting Diode (LED) to light up the surface of the skin. The second component is a photodetector, which is utilised for measuring the changes in light absorption over a period of time ?.

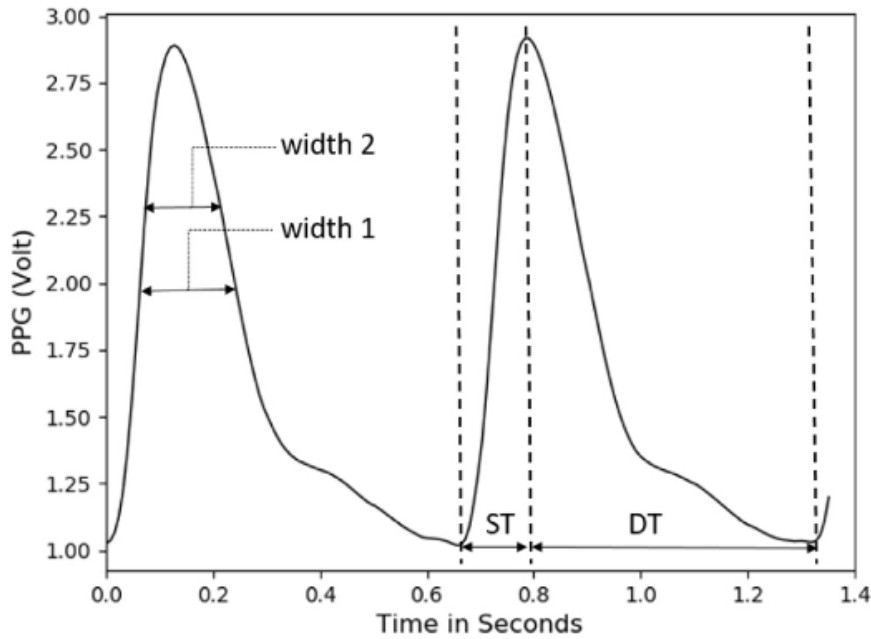


Figure 4: Structure of a PPG Signal ?

In Figure ??, the four features are the Systolic upstroke Time (ST), Diastolic Time (DT), width at $\frac{1}{2}$ amplitude (width 1) and width at $\frac{2}{3}$ amplitude (width 2) ?. PPG waveforms have a wide range of temporal features ?. These features have been utilised in several experimentations, creating models to estimate blood pressure ?.

2.6 Cuff-less methods for deriving BP

Cuff-less methods have great potential in being used to estimate BP. This is because they provide continuous measurements, they cause minimal harm to the patients and they produce BP values over a long period of time. There are three fundamental cuff-less methods which will now be discussed which can be used for deriving BP. These three methods rely on Pulse Transit Time (PTT), Pulse Arrival Time (PAT) and Pulse Wave Velocity (PWV) respectively. These will each now be discussed in more detail.

2.6.1 Pulse Transit Time (PTT)

PTT is the time required for the arterial pressure wave to travel from the left ventricle to a distal arterial site. PTT holds an inverse relationship to blood pressure and as a result it is dependent on arterial compliance, arterial wall thickness, arterial radius, and blood density. PTT is conventionally found with the use of two PPG sensors, as indicated in Figure 5.

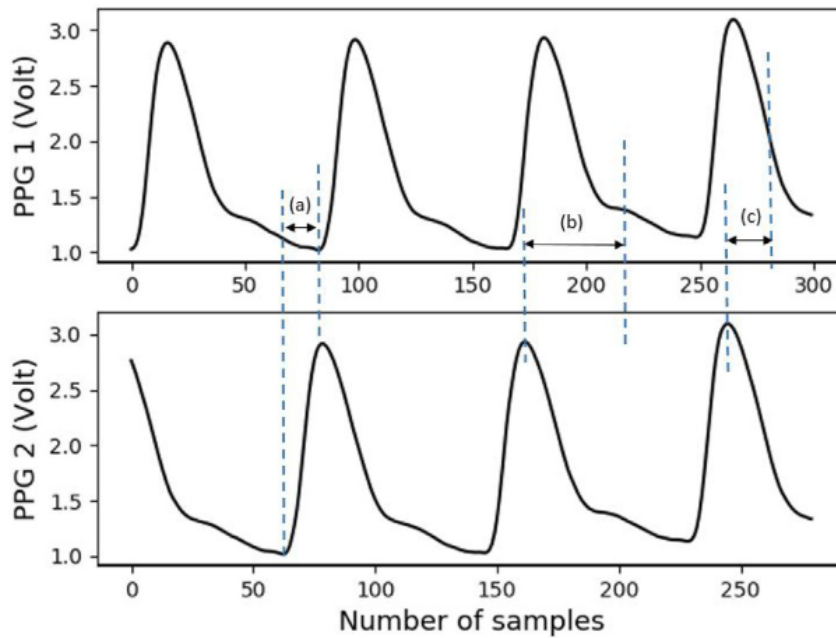


Figure 5: Pulse Transit Time (PTT) visualisation

It is important to note for Figure 5 that the PTT can be measured at different points along the PPG waveforms. (a) represents a foot-to-foot time delay, (b) is a peak-to-dicrotic notch time delay and (c) is a peak to mid-point of the falling edge time delay. As a proof of concept, increasing BP leads to an increase in the tension along the arterial wall tension, which therefore reduces the PTT. Hence, the opposite also applies.

2.6.2 Pulse Arrival Time (PAT)

The PAT is the difference in time between the R-peak of the ECG signal and the systolic peak of the PPG signal when measured during the same cardiac cycle, as indicated in Figure 6. Physically, PAT is the interval in time between the activation of electrical impulses at the heart and the arrival of the pulse wave at a location on the body, such as the finger. PAT is measured using two sensors, an ECG and a PPG sensor.

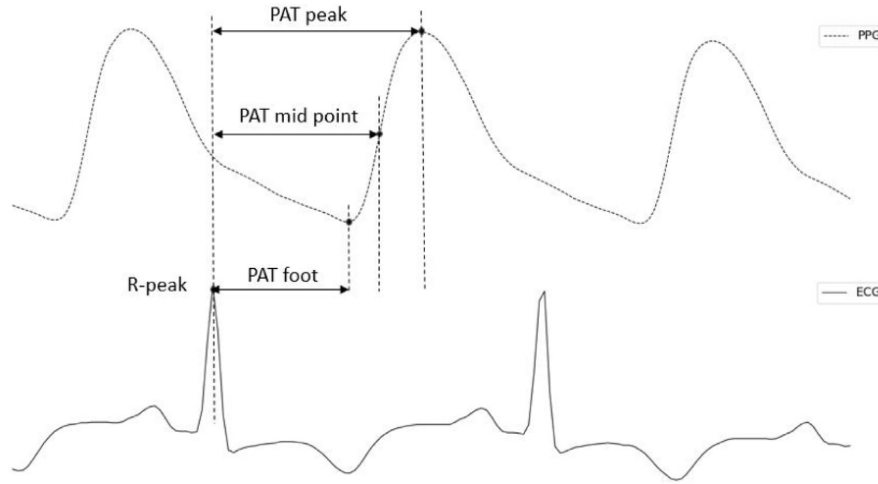


Figure 6: Pulse Arrival Time (PAT) visualisation ?

The Pre-ejection Period (PEP) delay can also be briefly discussed. PEP is the time needed to convert the electrical signal into a mechanical pumping force and isovolumetric contraction to open the aortic valves,

$$PAT = PTT + PEP \quad (1)$$

2.6.3 Pulse Wave Velocity (PWV)

The PWV calculates the velocity of the pulse wave using two PPG sensors located on the same arterial branch at a known distance apart ? ?. The relation between PTT and PWV can be expressed as

$$PWV = \frac{d}{PTT} \quad (2)$$

where d is the arterial distance travelled by the pressure wave. PWV is related to the Young's modulus of the vessel wall by the Moens-Kortweg equation,

$$PWV = \sqrt{\frac{Eh}{\rho d}} \quad (3)$$

where

PWV = Velocity of the pulse wave (m/s)

E = Young's modulus of vessel wall (Pa)

h = vessel thickness (m)

ρ = blood density (kg/m^3)

d = arterial diameter (m)

The Young's modulus of the vessel wall is then related to the arterial pressure by the Bramwell-Hills equation,

$$E = E_0 e^{\lambda P} \quad (4)$$

where E_0 and λ depend on the thoracic and abdominal aortas and P is the vessel blood pressure (mmHg) ? ? ?.

By solving Equations (2) and (3), we have the final equation for estimated blood pressure as,

$$P = \frac{1}{\lambda} \ln \left(2r\rho \frac{\Delta X^2}{E_0 h} \right) - \frac{2}{\lambda} \ln (PTT) \quad (5)$$

where

$r = \frac{d}{2}$ = arterial radius

ΔX = distance from heart to vessel

2.6.4 Limitations

The blood pressure (BP) can be derived through mathematical models as soon as estimates have been calculated for PTT, PAT and PWV. Although these models are common approaches for BP monitoring in an environment that is non-invasive and cuff-less, there are many challenges to these implementations. As a result, none of these techniques by themselves have been established as a reliable indicator for the estimation of BP.

Firstly, all three of the aforementioned methods require two separate measurements from two synchronised sensor devices. This can be a very inconvenient process for patients who are uncomfortable with this method ?.

In addition, there is a very likely possibility that these sensor devices will have different real-time sampling rates. Their operability depends on rather complicated arterial wave propagation models ?.

In order to be able to continuously measure BP, constant calibration of the methods is required. This is due to individual patients having different physiological parameters ?.

Finally, even with per-person calibration, these models can only provide BP estimation for a short period of time. As a result, this makes the models unreliable for the estimation of BP with every heartbeat ?.

To conclude this chapter, there is a lot of potential in the use of the three above parameters in the estimation of Ambulatory BP. However, there are still overarching limitations which currently hinder the progress of these parameters. As it will be discussed in Chapter 4, it will be indicated that these parameters still play a role in the estimation process but in methods which are more data driven.

3 Overview on Cuff-Less BP device options

Wearable technology provides an opportunity for real-time monitoring of human vital signs, thus enabling the possibility for preventive, timely notification and real-time diagnosis. Unlike commonly-used BP sensors, which demand a specific measurement procedure, modern wearable bio-sensors monitor vital signals online and all day long, presenting no additional burden other than wearing the device. A wide range of wearable devices achieve very positive results, even those wearables which are cheaper in price ?. Some of the systems developed for the purpose of non-invasive BP monitoring will be discussed in more detail now. In addition, they will be critically analysed against two other viable alternatives.

The motivation behind this project is to replace the current cuff-based BP devices. Cuff-based devices often require the supervision of an expert to work correctly and do not provide continuous measurements for BP. In addition they can cause irritation and inconvenience for patients due to cuff inflation and deflation. As a result, current clinical cuff-based BP devices are not suitable for providing continuous BP monitoring which could play a significant role in the early detection of diseases which affect the heart ?.

3.1 Smart Watches

Smart watches are becoming an increasingly popular form of wearable technology ?. Most of the existing smart watch produces currently measure BP through Pulse Transit Time (PTT) from a pulse wave measurement at the wrist. The Heartisans BP smartwatch uses ECG and PPG signals to measure PTT and estimate BP. The watch requires a motionless 20 second scan with the device held at heart level for measurements and provides systolic and diastolic BP readings. Calibration with a validated cuff-type BP device is required prior to standalone use. Despite its availability on the market, the Heartisans Watch has not undergone a formal validation study ?.

Another wearable method is through arterial tonometry. The BPro device, developed by the London company HealthSTATS Technologies, is one such device.

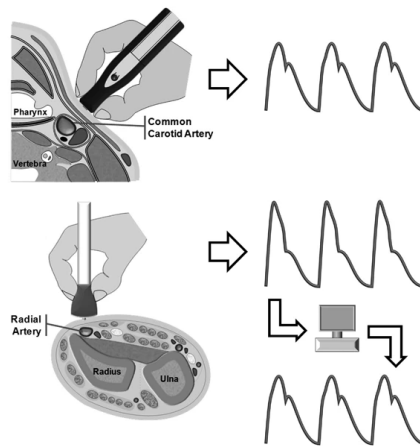


Figure 7: Arterial Tonometry

Figure ?? shows two different methods for using arterial tonometry to measure blood pressure. The first image shows that the recording is taking at the carotid artery level in order to estimate the central blood pressure waveform. In the second image, the recording is taken at the radial artery level to estimate the arterial pulse wave. The central waveform is then reconstructed from this pulse wave

using software.

The BPro allows for continuous 24 hour waveform analysis of BP. Calibration to the brachial artery BP is required using a validated upper arm cuff device prior to use. The BPro has been validated in a non-ambulatory study following the Association for the Advancement of Medical Instrumentation (AAMI) standards and the European Society of Hypertension (ESH) protocol.

3.2 Smartphone Applications

Smartphones offer a great potential to expand the continuous recording of BP if their sensors can be correctly utilised ?. A formal validation study on one particular iOS application, the Instant Blood Pressure app, revealed poor accuracy of BP measurements. Mean absolute differences of 12.4 mmHg for systolic BP and 10.1 mmHg for diastolic BP were found between the Instant Blood Pressure application and a reference device. This resulted in approximately four out of five hypertensive individuals being falsely classified as normotensive ?. The My BP Lab application measures BP through measurements of Pulse Transit Time (PTT). However, there is currently a lack of experimental data to justify it's dominance in the market.

3.3 Medical Tricorders

A medical tricorder is a handheld portable device used by consumers to self-diagnose medical conditions and take basic vital signs measurements. The BodiMetrics Performance Monitor uses ECG and PPG signal data to estimate BP through PTT. The Bodimetrics tricorder obtains measurements with a 20 second scan of the user's fingertip at heart level but only provides systolic BP data. This tricorder has a large spread in absolute bias against an automated sphygmomanometer, hence it is unlikely to meet formal accuracy and precision standards ?.

The FreeScan Personal Cardiovascular Monitor, which is developed by the Taiwanese company Maisense Inc., also estimates BP from PTT. However the device uses a force sensor to capture the systolic arterial waveform rather than PPG. This requires the user to physically apply the force sensor directly to the radial artery for around 10 seconds for measurements. The Freescan device has been verified according to the AAMI protocol ?.

The SOMNOtouch NIBP is a non-traditional medical tricorder that utilizes PTT data collected in a similar fashion to the Bodimetrics Performance Monitor. The device has met the ESH standards, however it begins to lose accuracy for higher SBP and DBP values ?.

3.4 Conclusion

After having discussed three viable devices for cuff-less measurements of BP, it has been finalised that smart watches are the most viable option. Whilst existing smart watch devices do suffer from inaccuracies in BP estimation due to motion, it is clear that they produce the most acceptable results in line with the AAMI and ESH standards. This chapter can be seen as a forward looking overview of how the estimation methods discussed in this report can be used to benefit future products. With regards to this project, this chapter can be treated as a supplementary overview.

4 Non-Invasive Cuff-Less Machine Learning methods for estimating BP

Due to advancements in technology related to machine learning, there has been a lot of research into neural networks algorithms that can offer continuous BP measurements that are non-invasive and also cuff-less ?. These ML models have been seen to be based on the parameters mentioned in Chapter 2.6 (namely Pulse Transit Time, Pulse Arrival Time and Pulse Wave Velocity). However, in this case BP estimation is motivated by how much data is available to the algorithm ?.

Artificial neural networks (ANNs) are a machine learning method that can be used to estimate blood pressure ?. ANNs are based on the neural networks found in the human body and aim to replicate their behaviour ?. The structure of ANN consists of multiple individual units called neurons. Each neuron has five critical components. These are inputs, weights, transfer function, activation function and bias ?.

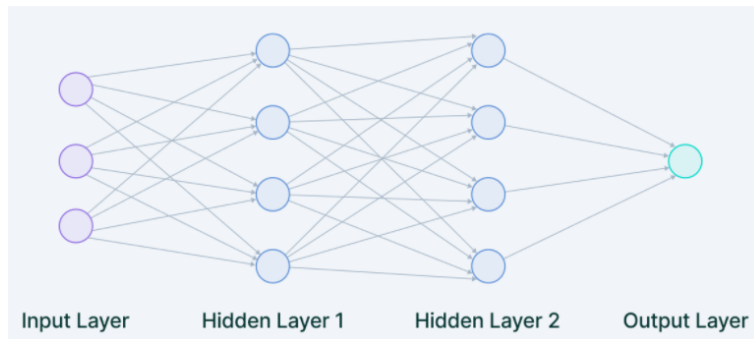


Figure 8: Structure of a multi-layer ANN

4.1 Convolutional Neural Networks (CNNs)

Convolutional Neural Networks (CNNs) are a class of feed-forward artificial neural architecture. The basic pipeline of common CNNs consists of an image as input and a stack of convolutional layers that extract a feature representation from the input image. The final shape of the image representation is conditioned on the type of problem/task that the architecture is facing. For instance, the output of the last layer in a classification problem is a probability vector. Each dimension of the probability vector represents how likely is that the input image belongs to a specific class. However, the architecture design is customisable, and therefore it is possible to code a network that outputs a single value for regression problems, or for example that generates a new image map for semantic segmentation.

One popular CNN architecture is the Residual Neural Network (ResNet) architecture, as depicted in Figure ???. ResNet is built by micro-architecture modules, called residual blocks. Those blocks introduced skip connections, which allowed to train huge architectures (152 layers) while still maintaining a lower complexity than other popular CNN architectures, such as VGGNet ?. Thus, residual connections allowed authors to design deeper architectures since the gradient could backpropagate easier through the skip connections.

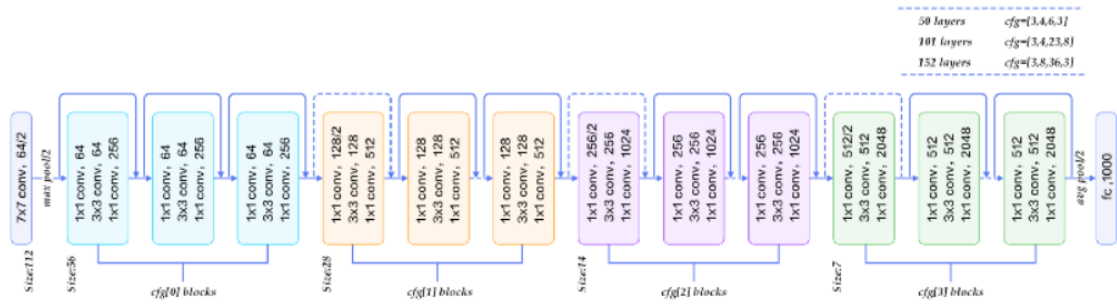


Figure 9: Architecture of the ResNet CNN with skip connections

4.2 Recurrent Neural Networks

A Recurrent Neural Network (RNN) is a specific type of architecture that is widely used to deal with sequential information ?.

RNNs are called recurrent since they apply the same operation to each of the input sequences, with the output of an individual element being dependent on the previous one. Theoretically, RNNs establish a connection between the actual input and ALL the previous ones ?. Although this is assumed, in the practice, RNNs have proven to only remember a limited number of inputs. In other words, RNNs have a memory that allows them to remember previous elements and use their information to deal with the current input ?.

RNNs can be split into multiple types depending on their applications. For instance, if we want to predict one word given only the previous one, the topology of our network is a One to One. Another example is image captioning, where we can design a One to Many architecture to obtain a description from a single input image ?. The following diagram shows the different types of problems we can face:

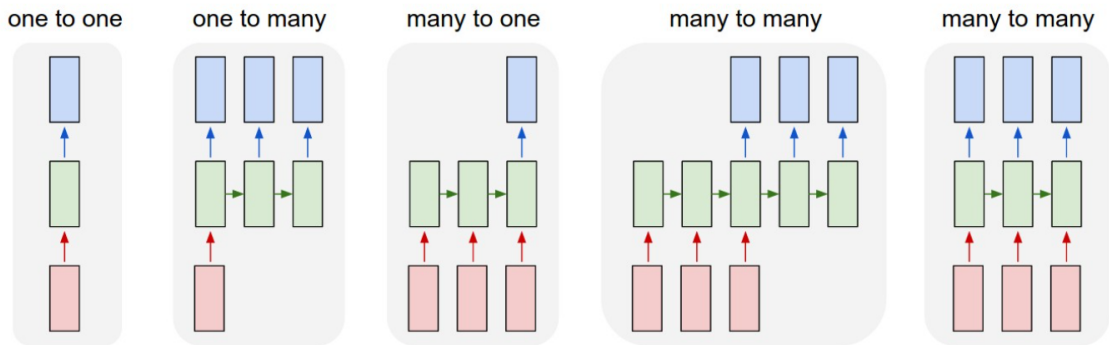


Figure 10: RNN problems

Figure ?? shows the simplest version of an RNN, which can be easily derived from a simple feedforward architecture by adding a single loop:

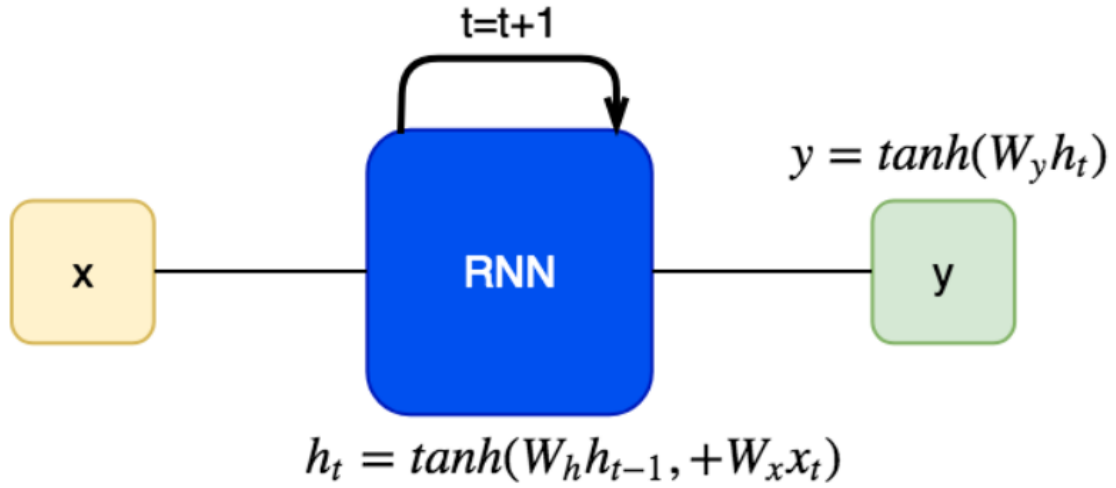


Figure 11: Simplification of a RNN

During training, the hidden state h is iteratively updated based on the input value x and the learned weights W_h and W_x . The final output y is estimated from the current state h_t and the matrix W_y . Although RNN can assure short-term dependencies within the network, simple RNNs become unable to learn to connect information as the gap between past and present information grows. To overcome this limitation, in practical applications LSTM unit is adopted, that is a special RNNs architecture composed of multiple interacting layers.

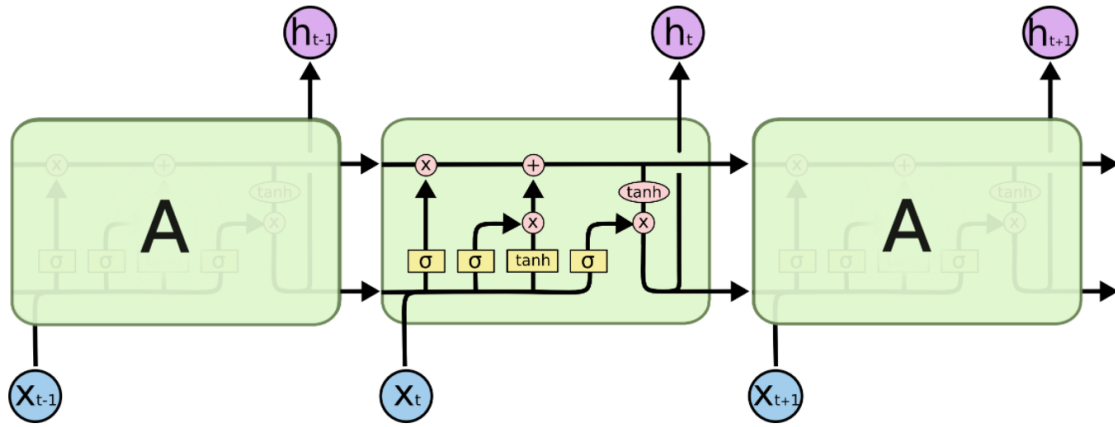


Figure 12: LSTM network

4.3 Transformer Neural Networks (TNNs)

RNNs are still quite used in several architectures and Natural Language Processing (NLP) tasks. However the Transformer Neural Network (TNN) has now started to outperform RNNs in several text/NLP benchmarks. A Transformer does not use any recurrence, it instead uses attention to focus on specific parts of a sentence.

Both text classification and text generation tasks have been successfully tackled by transformers. Transformers seem to scale better than recurrent neural networks. They can have great results when using large architectures and large datasets, whereas RNNs seem not to benefit as much from using more parameters and training examples. State-of-the-art transformer models can have even billions of parameters and can benefit from training in terabytes of text data.

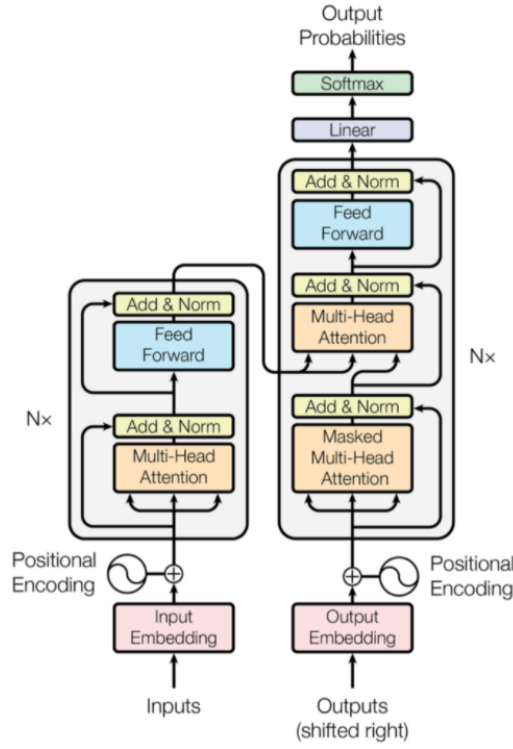


Figure 13: Model architecture for a transformer

4.4 Error metrics

The two considered error calculations used in this experimentation are the Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). They are defined by the following equations,

$$MAE = \frac{1}{N} \sum_{i=1}^N |a_{i_M} - b_{i_M}| \quad (6)$$

$$RMSE = \frac{1}{N} \sqrt{\sum_{i=1}^N |a_{i_M} - b_{i_M}|^2} \quad (7)$$

In the context of BP estimation, b_{i_M} and a_{i_M} represent the true value and BP estimate respectively for the M th element of the time sequence.

4.5 Literature Survey

As a reference, the original literature survey matrix can be found on the Github repository [?](#). Firstly, in Table [??](#), a simplified literature survey has been detailed out for the best performing methods which do not employ machine learning methods.

Study	Source	No. Subjects	Age	Implementation	MAE SBP
?	ECG, PTT-CP	10	24-63	Numerical solution	± 5.93
?	ECG	5	N/A	Analytical solution	9 ± 5.6
?	PPG	16	18-48	Frequency analysis	0.8 ± 7
?	ECG, PPG, PTT	N/A	N/A	Analytical solution	7.49 ± 8.8
?	PTT	127	N/A	Wavelet transforms	± 7.63
?	PTT, PPG	27	21-29	Analytical solution	-0.37 ± 5.21

Table 2: Overview of performance of the best non-invasive non-ML cuff-less methods for measuring BP

Study	Source	No. Subjects	Age	Method	MAE SBP
?	ECG, PPG	14 males	17-43	ANN	7.99 ± 10.34
?	PPG	65	22-65	Wavelet, SVM	5.1 ± 4.3
?	PPG	MIMIC II	Adults	Linear Reg., ANN, SVM	13.84 ± 17.56
?	ECG	51	16-83	Complexity analysis + ML	7.72 ± 10.22
?	PPG	72	N/A	ANN (MLP)	4.02 ± 2.79
?	ECG, PPG	MIMIC II	Adults, neonatal	ANN (150 neurons)	5.76 ± 6.39
?	ECG, PPG	39	20-100	ANN-LSTM	1.10
?	PTT, ECG, PPG	MIMIC I	N/A	SVM, Lin Reg.	3.27 ± 5.52
?	PTT	250	MIMIC I	ANN-RBM	3.70

Table 3: Overview of performance of the best non-invasive ML cuff-less methods for measuring BP

The Mean Absolute Error (MAE) of the Systolic BP (SBP) was used as the uniting accuracy measure in this paper, as it was the most readily available parameter in all of the aforementioned papers.

It is important to note that other factors were considered in this literature survey. These factors were,

- Range of SBP and DBP values
- Sampling frequency
- Denoising and detection techniques used
- Computational complexity
- Feasibility in a wearable context

However, due to a lack of regular occurrences of these factors over all the papers, they were not included in the above two tables.

The results show that the established non-ML methods do produce MAE values noticeably lower than the majority of the ML methods. However, a main factor to consider about these results is that machine learning based methods and neural networks are data driven. As shown in Table ??, there is a very limited number of subjects available for each study. If these studies had been extended to include more test patients, it is possible that these MAE SBP values were lower. An additional point is that another study ? did a study into using a 4-layer LSTM architecture to estimate BP from ECG and PPG signals with 84 patients. This study resulted in an RMSE SBP of 3.9 mmHg but no available MAE. Hence there is a lot of potential in ML methods when there is sufficient data available.

5 Implementation

The scope of the project can be split into 4 sections, as detailed in the subsections below. The outcome of each task from each section forms a milestone. The full display of sections and tasks is provided in a Gantt chart ? in the Appendix. The next 2 chapters will break down the respective sections of the chart.

5.1 Project Research and Understanding

This section includes the following, as depicted in Figure ??:

1. Exploration of literature on the heart and the underlying theory for blood pressure
2. Investigation into existing experimentations conducted with wearable technologies for estimating blood pressure
3. Research into signal denoising techniques and machine learning based methods for estimating BP
4. Experimentation with the **Mind the Gap: The PhysioNet/Computing in Cardiology Challenge 2010** (later referred to as the PhysioNet 2010 database) ?
5. Deciding on which methods to implement as a part of the project

TASK	START	END
Research and Understanding		
Heart and blood pressure research	25/10/21	9/1/22
Wearable technologies survey	29/11/21	9/1/22
Denoising + ML research	25/10/21	9/1/22
Familiarise with Mind the Gap database	6/12/21	9/1/22
Higher Level design implementation decision	17/12/21	10/1/22

Figure 14: Research and understanding matrix

For the first three tasks, the milestone is to write Interim Report sections on that particular topic. These three sections have been explained in Chapters 2-4 of this report. The fourth task was more experimental, and the sole aim was for myself to become more comfortable with working with this database. The final task is the basis of the conclusion in Chapter 4. The work completed to date has been on schedule. Some of these tasks have had a larger time-span allocated to account for clashes with exams and coursework occurring at the beginning of the Spring Term.

5.2 Implementation of chosen methods

This section includes the following:

1. The implementation of a Convolutional Neural Network (CNN) architecture
2. The implementation of a Recurrent Neural Network - Long Short Term Memory (RNN - LSTM) architecture
3. The implementation of a Transformer Neural Network (TNN) architecture

4. End-to-end testing of the three architectures

TASK	START	END
Implementation of chosen methods		
Implementation of CNN	10/1/22	1/2/22
Implementation of RNN-LSTM	1/2/22	15/2/22
Implementation of TNN	15/2/22	1/3/22
End-to-end testing	22/2/22	8/3/22

Figure 15: Implementation of chosen methods matrix

This section is the most crucial part of the project in terms of acquiring results. The BP estimation problem has been approached from 3 different directions: the CNN, the RNN-LSTM and the TNN architectures. The report section will include results of these unit tests, as well as a description of how the method functions and the process of programming the algorithm in Python. The main milestones for the first three points will be the comparison between the methods and unit testing. The last task in this section will be testing all methods from start to finish, with the aim to find any errors or bugs in the code. These tasks are currently set to take place during Spring term. However this may go beyond the set dates due to commitments from Spring Term coursework. This will be accounted for in the remaining 2 tasks to ensure the project progress is running on time.

5.3 Performance Testing of methods

This section includes the following:

1. Testing the three methods with the PhysioNet 2010 database ?
2. Testing the three methods with the MIMIC I database
3. Testing the three methods with the MIMIC II database
4. Applying any enhancements to the three methods based on results or new literature found

TASK	START	END
Performance Testing		
Testing methods with 2010 challenge data	15/2/22	15/3/22
Testing methods with MIMIC I database	15/3/22	29/3/22
Testing methods with MIMIC II database	29/3/22	12/4/22
Improvements to chosen methods	5/4/22	30/4/22

Figure 16: Performance Testing matrix

This section primarily consists of comparing the three methods on a variety of datasets. This section will also allow me to gather any information or data which may help to further optimise the methods used. The main milestones will be to gather this data and document the findings in the Final Report, along with potential reasons for the difference in performance. Extra time has been allocated to this stage due to exam commitments.

5.4 Assessment Preparation

This section summarises the main deliverables which embody the Final Year Project, as well as my own modules which run at the same time. The tasks are as indicated in Figure ??,

1. Literature survey presentation to supervisors
2. Interim Report submission
3. Accounting for commitments due to Spring Term coursework and Summer Term exams
4. Writing up the Final Report
5. Preparing for the final presentation

TASK	START	END
Assessment Preparation		
Presentation on literature review to ZP and ERV	13/12/21	20/12/21
Interim Report prep and submission	20/12/21	31/1/22
Spring coursework and Summer Exams	7/2/22	14/4/22
Final Report writeup and submission	14/4/22	15/6/22
Final Presentation prep and due date	13/6/22	27/6/22

Figure 17: Assessment Preparation matrix

The first two tasks are now effectively completed. It was noted that the first task was beneficial in the writing of Chapters 2 to 4 in the Interim Report. Regarding the third task, in order to stay focused on completion of the FYP, I will ensure to stay in regular weekly communication with my supervisor, and to ask any insightful questions where necessary. The final two tasks stem from successful completion of the FYP.

6 Evaluation Metrics

This section details how the success of each section will be measured, including backups in case a task for a particular section is not able to be completed.

6.1 Project Research and Understanding

The success of this section will be measured by the quality of the chapters of the Interim Report, in particular on the literature surveys. The report is the best factor to see how much has been understood during the Autumn and Spring Term. A primitive report draft was sent to my supervisor during this period to ensure information was concise, accurate and relevant. There is no definite backup plan for this section, except for implementing any changes based on the feedback for this report.

6.2 Implementation of chosen methods

Success is based on the output of the unit and end to end tests. If the methods all pass these tests by delivering similar results to those seen in the reviewed papers, then this is a total success. The main risk of this section will be the feasibility of implementing the solution in Python, i.e. whether it would take a deeper understanding of the mathematics or if certain capabilities are not possible in Python. If such a situation arises, the backup plan will be to switch to another programming language, such as MATLAB, or to implement another more suitable method, such as a wavelet-based method.

6.3 Performance Testing of methods

Success is based on the data acquired during the testing phase of the three potential methods. The Final Report is also an important indicator as it will detail the project findings. If balanced and suitable data, with regards to Ambulatory BP monitoring, is gathered for all of the potential methods under different use cases then this is seen as a total success. In addition, success is guaranteed if the report provides valid reasoning on the results. The main risk of this section is misunderstanding how the complexity can be correlated for the three different methods, leading to a misinterpretation of the results. The backup plan is to research more into the associated computational complexities when using these particular methods, to ensure a sound understanding of the principles.

With regards to improving the existing methods further, success will be measured based on the figures from the updated methods, e.g. reduced complexities, increased accuracies. The main risk of this section is the possibility of not finding any improvements to be made. As these are already well researched methods internationally, the chances of finding improvements in the methods used are minimal. If a potential improvement isn't found at all, then the backup plan is to organise a discussion meeting with the project supervisor.

6.4 Assessment Preparation

The success of this section will be based on several personal factors. Firstly, with regards to the project-related assessments, success will be measured based on how frequently I communicate with my FYP supervisors and how frequently I pose insightful questions that are beneficial to the development of my project. In addition, regular updates and additions to the Final Report as the Spring Term progresses will ensure that it is completed comfortably before the deadline of the 31st of May 2022. Going forward, emails will be sent detailing major progress to my main supervisor contact (Dr Zaibaa Patel) and other supervisor (Professor Esther Rodriguez Villegas) every 2-3 weeks in the Spring term, with more frequent updates (1 week) targeted during the Summer term. Meetings will be held as and when required.

7 Databases for the project

After having conducted a literature survey into the existing methods for estimating BP, I will now give an overview of what data will be used for this project.

7.1 Literature survey data

Referring back to the literature survey in Chapter 4, all of the studies used one of either Physionet's Multiparameter Intelligent Monitoring in Intensive Care (MIMIC) I or MIMIC - II databases.

The MIMIC Database includes data recorded from over 90 ICU patients. The data in each case include signals and periodic measurements obtained from a bedside monitor as well as clinical data obtained from the patient's medical record. The recordings vary in length; almost all of them are at least 20 hours, and many are 40 hours or more. In all, the database contains nearly 200 patient-days of real-time signals and accompanying data ?. The MIMIC-II database consists of 25,328 intensive care unit stays ?.

7.2 Data for this project

The database which is used for current experimentation is the Physionet 2010 database ?. The samples were taken from ICU patient monitor recordings collected from the MIMIC II database. After this, the Physionet team made ten-minute records containing at least six simultaneous and continuous signals (digitized at 125 Hz each). The signals vary across each of the records. The signals primarily include ECG and PPG signals, but some do occasionally contain other signals, such as respiration. 300 monitor recordings were randomly chosen and a random ten-minute interval was chosen within each recording. In each ten-minute recording, the final thirty seconds of one of the signals, also chosen at random, was designated as the target signal and it was replaced by a flat line signal. The aim was to recreate this missing 30-second signal in each database recording. The modified records were assigned to one of the available data sets of 100 records each. It is noted that this assignment was also done at random:

- Set A is a training dataset of 100 records. The target signals were provided for the records in this set.
- Set B was also a dataset of 100 records. However the target signals were not made available until the end of the competition.
- Set C is a set of 100 records, as was identical to Set B, in that the target signals were not made available.

Once a greater understanding of this database has been developed, there will be future plans to implement data found in the MIMIC-I and MIMIC-II databases, in order to produce results more in line with those found in the previous literature survey. The most important consideration is that the dataset has minimal bias with regards to age, gender, etc.

7.3 Software Implementation

Python will be used as the sole programming language for this project. Python has a wide variety of easy to use and powerful libraries. These include scientific libraries such as `numpy` and `pandas` and machine learning libraries, including `tensorflow`, `keras` and `scikit-learn`. In addition the `wfdb-python` package will be installed, which is library of tools for reading, writing, and processing Waveform-Database (WFDB) signals and annotations. These are commonly found on the Physionet

website.

MATLAB was also considered as a potential programming language to use, due to it having a wide range of signal processing and machine learning add-on toolboxes. However Python has been shown to offer a wider set of choices in graphics packages and toolsets, such as through `matplotlib`, and it also produces more compact and readable code.

Finally, a Github repository has been created for this project ?. It will be constantly updated with Python code and supporting documentation throughout the Spring and Summer Terms and will function as an organised directory for all files used in this project.

8 Ethical, Legal and Safety Plan

8.1 Ethical considerations

This section discusses the ethical implications of this project through the perspective of the Four Pillars of Medical Ethics: beneficence, non-maleficence, autonomy, and justice ?.

8.1.1 Beneficence

Beneficence is the duty to do good. During the process of this project, it must be ensured that actions must be taken with the intent of helping patients. The purpose of future wearable technologies is to help people to detect early risk factors of heart-related pathologies.

The issues that may arise would be due to the occurrence of false positives and negatives in the data, such as a normotensive result for a person who is actually hypertensive. Ideally, a blood pressure monitor which is cuffless should operate accurately over a range of blood pressure values after having been calibrated. However this has not been tested according to the established medical standards, such as AAMI and ESH. Measurements that are inaccurate will naturally cause a false sense of reassurance with the patients, unnecessary health care utilization, or improper administration of hypertension medications. An incorrect diagnosis also causes unnecessary psychological stress and the unnecessary reactions to the negative side effects of prescribed medication. For example, if there is a range of ± 5 mmHg in error, this could lead to approximately twenty-seven million hypertensive and twenty-one million normotensive patients being classified incorrectly ?. This can only be minimised through robust tuning of the parameters in the neural networks used.

8.1.2 Non-maleficence

Non-maleficence is the duty to not harm. Through the recording of ECG and PPG data, it must be ensured that any actions must not be harmful to the patients. An instance where the product could harm patients would be if the models falsely classify a patient to be normotensive, whereas the patient indeed is hypertensive in real life. If test accuracy could be proven, it would be possible to see if a model is provably 100% accurate, or otherwise justify to an Ethics Board that the risks would be worth the reward. However, with neural networks of this complexity, it is practically impossible to prove any estimate of test accuracy.

8.1.3 Autonomy

Autonomy is the respect of patients' freedom of choice and consent. During the data acquisition process, it must be ensured that there are no actions taken that are non-consensual to the patients. Those concerned should also be informed about the benefits and risks of using a particular wearable technology product for BP measurement, such as the validation accuracy and inference time, allowing them to make an informed decision.

Another general issue of using deep neural networks is that it is very difficult to explain how models obtain their results. It is impossible to explain what features an individual or end-to-end model detects and what any of their weights correspond to. This can be an issue if the user wants an explanation as to why they have been classified wrongly as hypertensive or hypotensive, as they should have the right to know.

8.1.4 Justice

Justice is the principle of fairness towards all patients and the idea of treating all people equally and equitably. The product must ensure it does not discriminate between patients and medical staff on any basis. An obvious concern is whether the dataset used to train the models are balanced in representing both gender and ethnicity. Although the dataset was created taking such biases into account, it is difficult to overcome subconscious biases and availability of a balanced dataset. This is important as it ensures that future products in wearable technologies have both a sustainable customer base but is also sustainable in today's social ecosystem, as it is actively standing against racism.

8.2 Legal considerations

Legally this project poses no risk. The Python programming language and its required libraries are Open Source, including the version which I intend to use (3.10). Python has a license agreement under the Python Software Foundation (PSF). Hence it is a tool that can be legally used freely for the purposes of this project. In addition, the ECG and PPG data is provided by a website called Physionet. PhysioNet is a data repository of medical signals which are free to use. The repository is managed by the MIT Laboratory for Computational Physiology. As this code will be published online, the produced code could be plagiarised in the future. However, I myself will bear no responsibility in the distribution of my code to others.

8.3 Safety considerations

With regards to this project, there are no safety issues to consider, as all experimentation will be conducted online at home. It is critical that necessary measures are taken such that strains or injuries from desk work are prevented. This can easily be done by following NHS guidelines.

As highlighted in Chapter 3, there are a few points to consider if this work becomes implemented in a future wearable technology product. Firstly, it is important to remember these are only simulations. Training a machine learning model on Python will not accurately reflect the behavior in real life. Secondly, any results discussed in this report will not be officially recognised by the AAMI or ESH protocols. This is something that future active projects will need to consider, in order to have substantial results.

9 Current experimentation

The aim of this experimentation is to evaluate six implementations which were used in the Physionet 2010 database. These are detailed in Table ??,

Name	University	Method used
András Hartmann	Semmelweis University, Hungary	Adaptive Filtering
Mohamed Mneimneh	Sahar Elturk Marquette University, USA	Kalman Filtering
Ikaro Silva	MIT, USA	Kalman and Adaptive Filtering
Adam Sullivan et al.	University of Tennessee, Knoxville, USA	Artificial Neural Networks
Rui Rodrigues	Universidade Nova de Lisboa, Portugal	Neural Networks
Luiz Silva et al.	University of São Paulo, Brazil	Recurrent Neural Networks

Table 4: Overview of most popular methods used in the 2010 Challenge ?

It is noted that the first three methods adopt signal processing and filtering techniques ? ? ? and the other three methods implement machine learning based methods ? ? ?.

For the benefit of this project, I wanted to recreate their implementations to justify the choice of using machine learning based methods to estimate BP from ECG and PPG signals. The metric to compare all six implementations was the aggregated scores of the models on Set C of the previously mentioned PhysioNet 2010 database ? C_1 is a parameter between 0 and 100 which scores the overall accuracy of the reconstructed signal by the methods. C_2 is a similar parameter which instead measures the correlation between the original and reconstructed signals. Hence the final performance results are shown in Figure ??,

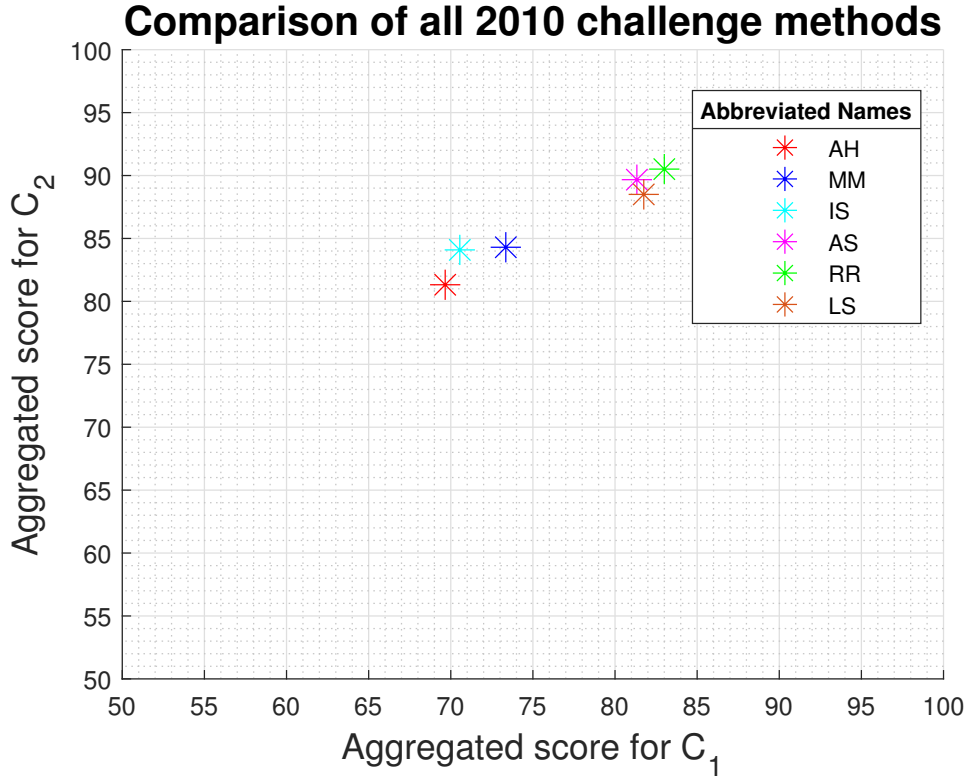


Figure 18: Performance of all implementations used in the 2010 CIC challenge

As the results in Figure ?? show, the machine learning methods appear to reconstruct the ECG signals much more effectively than the signal processing based methods. This is a chapter that will require more depth but for now, it shows promising signs that a machine learning method will be more effective in estimating blood pressure from at least ECG signals. An important point to note is that there were no Convolutional Neural Network (CNN) or Transformer Neural Network (TNN) methods available to implement on this dataset. In order to be in line with the arguments made in this report, a viable step would be to see the performance of a CNN on the Physionet 2010 database.

10 Conclusion

In Chapter 1, the project deliverables were highlighted and a motivation was provided to complement the background information. In Chapter 2, detailed background information was given on a variety of topics. Firstly, on hypertension, a significant disease associated with the heart. After this, the principles of blood pressure measurements were discussed, as well as the fundamentals of ECG and PPG signals. Finally, the underlying theory for PTT, PAT and PWV were discussed. In Chapter 3, an overview of the existing wearable implementations for estimating blood pressure were discussed, with smart watches being viewed as the most viable option due to the prevalence of accurate data. In Chapter 4, a literature survey was taken to assess the potential methods to estimate blood pressure from ECG and PPG signals. As a result, machine learning methods were favoured over adaptive filtering, wavelet-based and numerical methods. The implementation and evaluation plans were given in Chapters 5 and 6 respectively. After this, more information on the data used and the Ethical, Legal and Safety considerations were discussed.

The immediate future tasks are to implement each of the viable methods in Chapter 4 in Python and to begin to record results for the necessary chapters in the Final Report. In addition, a study will be taken into the existing denoising techniques used on the ECG and PPG waveforms, in order to assess what is suitable for this project. Finally, more empirical analysis is needed to assess the effects of complexity on the implementations. One potential factor which can be used to measure the effects of complexity on neural networks is the inference time.

11 Appendix

11.1 Miscellaneous Points

- The Advancement of Medical Instrumentation (AAMI) standard requires a mean BP difference of ≤ 5 mmHg with a standard deviation of ≤ 8 mmHg against auscultatory reference measurement
- Significant variation in BP measurements(12 mmHg systolic or 8 mmHg diastolic) from the validated reference device is an exclusion criterion in the AAMI protocol ?
- The European Society of Hypertension (ESH) protocol requires that the majority of subjects have investigational BP readings within ≤ 5 mmHg of the reference measurement.

11.2 Gantt Chart

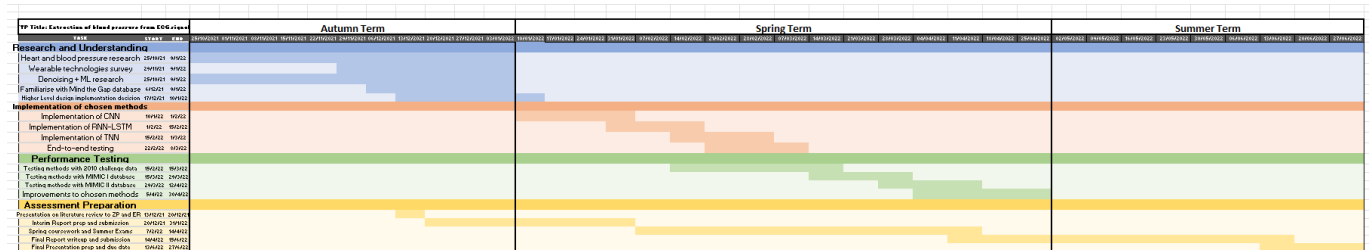


Figure 19: Gantt Chart