

POLITECNICO DI TORINO

Master Degree in Biomedical Engineering

Master Thesis

Non-invasive methods for continuous blood pressure estimation from sensor data



Supervisor
Dr. Luigi Borzì

Candidate
Simona Di Biase

Academic Year 2024-2025

Summary

Hypertension represents a significant health risk, potentially leading to serious complications, making it a primary cause of death around the world. Continuous blood pressure (BP) monitoring is considered essential for ensuring reliable diagnosis. However, invasive methods, even though a gold standard for BP monitoring, can lead to serious complications. Non-invasive techniques are commonly used but cannot monitor BP continuously. In this context, it is crucial to develop a non-invasive, safe, and comfortable method for continuous BP monitoring during daily activities. Alternatively, photoplethysmography (PPG), a portable optical sensor that continuously detects volumetric changes in blood circulation, offers a promising solution for non-invasive and continuous BP monitoring. However, BP estimations derived exclusively from PPG signals frequently suffer from limited accuracy and fail to meet clinical standards. To overcome these limitations, recent approaches have focused on integrating PPG with electrocardiographic (ECG) signals, using Pulse Transit Time (PTT) based method and artificial intelligence (AI) based techniques. Despite their promising performance, PTT-based methods require subject-specific calibration, whereas AI-based models are often validated in an overoptimistic manner due to data leakage. In this thesis work, the estimation of continuous BP through PTT using regression techniques is investigated, evaluating models from literature and training them with a limited number of individual BP values per subject. Subsequently, various machine learning (ML) models are assessed using a large set of features extracted from ECG and PPG, with a subject-wise validation strategy to ensure generalizability and avoid data leakage. This work involves the analysis and processing of PPG, ECG and arterial blood pressure waveform extracted from the MIMIC II online database. These data were used as input for three relevant PTT-based methods, characterized by 2, 3, and 4 coefficients, and the models were calibrated using a limited number of BP values. As a further step, a set of features was extracted from the ECG and PPG to enable the application of different ML techniques. The dataset was partitioned by subject into 40% for training, 30% for validation, and 30% for testing to ensure subject-independent evaluation. Among the various model and training data combinations evaluated, 2 coefficient model trained with 5 BP values achieved the best performance, with a mean absolute error (MAE) of 5.309 ± 4.561 mmHg. This configuration balances model simplicity with minimal training. Furthermore, the distribution of MAE reveals that approximately 60% of subjects attain a MAE of 5 mmHg or less and about 30% fall between 5 and 10 mmHg. Gaussian Process Regression emerged as the best-performing method during validation and achieved

a MAE of 8.613 ± 8.315 mmHg on the test set. Moreover, MAE distribution reveals that approximately 50% of individuals attain a MAE of 5 mmHg or less and around 30% fall between 5 and 10 mmHg. In conclusion, regression model offers high predictive accuracy but its dependence on subject-specific calibration represents a significant limitation. In contrast, ML approach, although slightly less accurate, offers the advantage of being calibration-free and more generalizable across individuals. These promising results suggest that the algorithm could be further optimized and mark an initial step toward its future integration into an automatic wearable device.

Contents

| | |
|--|----|
| List of Tables | 5 |
| List of Figures | 6 |
| 1 Introduction | 7 |
| 1.1 Blood pressure and hypertension | 7 |
| 1.2 Conventional blood pressure measurement methods | 8 |
| 1.3 Electrocardiography | 12 |
| 1.3.1 Electrocardiographic recording | 12 |
| 1.3.2 Electrocardiographic waveform | 13 |
| 1.4 Photoplethysmography | 14 |
| 1.4.1 Photoplethysmographic recording | 14 |
| 1.4.2 Photoplethysmographic waveform | 15 |
| 1.5 Electrocardiography and photoplethysmography for blood pressure monitoring | 16 |
| 1.6 Contributions and thesis organization | 17 |
| 2 Related work | 19 |
| 2.1 Photoplethysmography-based methods | 19 |
| 2.2 Pulse transit time-based methods | 21 |
| 2.3 Artificial Intelligence-based methods | 22 |
| 3 Materials and Methods | 25 |
| 3.1 Data | 25 |
| 3.2 Data analysis | 25 |
| 3.2.1 Pre-processing | 25 |
| 3.2.2 Regression analysis | 27 |
| 3.2.3 Machine learning analysis | 30 |
| 4 Results | 38 |
| 4.1 Regression analysis | 38 |

| | | |
|----------|--|-----------|
| 4.2 | Machine learning analysis | 41 |
| 4.3 | Comparison with Existing Methods | 44 |
| 5 | Discussion | 46 |
| 6 | Conclusion and future work | 49 |

List of Tables

| | | |
|-----|--|----|
| 4.1 | Results in terms of MAE \pm SD for the three regression methods evaluated with different numbers of BP values. | 39 |
| 4.2 | Selected features grouped by category | 42 |
| 4.3 | MAE and Pearson correlation coefficient for each model evaluated on the validation set | 42 |

List of Figures

| | | |
|-----|-------|----|
| 1.1 | | 8 |
| 1.2 | | 11 |
| 1.3 | | 12 |
| 1.4 | | 13 |
| 1.5 | | 15 |
| 1.6 | | 16 |
| 1.7 | | 16 |
| 3.1 | | 26 |
| 3.2 | | 26 |
| 3.3 | | 28 |
| 3.4 | | 28 |
| 3.5 | | 32 |
| 3.6 | | 33 |
| 3.7 | | 34 |
| 3.8 | | 35 |
| 4.1 | | 40 |
| 4.2 | | 43 |

Chapter 1

Introduction

1.1 Blood pressure and hypertension


Blood pressure (BP) refers to the pressure exerted by blood against the arterial wall. It is influenced by cardiac output, blood volume and viscosity, peripheral vascular resistance, and vessel wall elasticity. BP is one of the most important vital signs of the cardiovascular system and the general health. Along with respiratory rate, heart rate (HR), oxygen saturation, and body temperature, it is assessed by healthcare professionals to evaluate a patient's physiological status [1].

BP is measured in millimetres of mercury (mmHg). It is typically expressed with two values: systolic blood pressure (SBP) which corresponds to the maximum pressure exerted on the arteries during left ventricular contraction, and diastolic blood pressure (DBP) which represents the lowest pressure on the arteries, occurring between heartbeats when the heart's chambers are filling with blood [1, 2].

According to the classification outlined in Figure 1.1, BP is recognized as healthy when the systolic value is under 120 mmHg and the diastolic is below 80 mmHg. High BP is indicated by consistent systolic readings of 130 mmHg or higher and diastolic readings of 80 mmHg or higher. High BP, also known as hypertension, represents a significant health risk, potentially leading to serious complications such as heart attack and stroke [3].

As reported by the World Health Organization (WHO) in 2023, approximately 1.28 billion people aged 30 to 79 worldwide are affected by hypertension. An estimated 46% of individuals with hypertension remain unaware of their condition, while less than half of adults with hypertension (42%) have received proper diagnosis and treatments. This makes hypertension a leading cause of premature deaths globally [4].

Blood Pressure Categories



| BLOOD PRESSURE CATEGORY | SYSTOLIC mm Hg (upper number) | | DIASTOLIC mm Hg (lower number) |
|--|----------------------------------|--------|-----------------------------------|
| NORMAL | LESS THAN 120 | and | LESS THAN 80 |
| ELEVATED | 120-129 | and | LESS THAN 80 |
| HIGH BLOOD PRESSURE (HYPERTENSION) STAGE 1 | 130-139 | or | 80-89 |
| HIGH BLOOD PRESSURE (HYPERTENSION) STAGE 2 | 140 OR HIGHER | or | 90 OR HIGHER |
| HYPERTENSIVE CRISIS (consult your doctor immediately) | HIGHER THAN 180 | and/or | HIGHER THAN 120 |

heart.org/bplevels

Figure 1.1. BP Categories [4]

High BP typically causes damage over time, and if not properly monitored and treated, it may lead to heart attack, stroke, heart failure, kidney disease or failure, vision loss, sexual dysfunction, heart disease, and atherosclerosis [3].

Individuals with hypertension often rely on occasional measurements to monitor their BP. However, since BP fluctuates throughout the day due to many factors such as food intake, stress levels, physical condition, and type of medication, continuous monitoring is needed [5, 6].

Therefore, continuous BP monitoring is considered essential not only for ensuring reliable diagnosis and effective treatment, but also for supporting clinicians in prescribing appropriate diet and medications according to each patient's needs.

In this context, the development of a non-invasive continuous BP monitoring technique that is cost-effective, safe, comfortable, and widely accessible becomes particularly relevant. The main purpose is to enable the continuous recording of a patient's BP during routine daily activities, without the individual's awareness of being monitored [6].

1.2 Conventional blood pressure measurement methods

Accurate and regular monitoring of BP is essential for diagnosis and management of hypertension, for assessing cardiovascular risk and predicting acute cardiovascular events. As illustrated in Figure 1.2, many invasive and non-invasive conventional BP measurement techniques have been developed [7].

Arterial catheterization represents the invasive gold standard for BP measurement. It involves the use of a catheter connected to a piezoelectric transducer equipped with a diaphragm. As blood flows through the cannula toward the measurement system, it applies pressure on the diaphragm resulting in its deformation and the stretching of the piezoresistive element. This deformation causes a change in the resistance of the piezoresistive material, leading to variations in the system's electrical output that are proportional to the arterial pressure. After amplification and filtering, the signal is displayed as a waveform representing the pressure profile. This technique is primarily employed in acute and critical care settings, such as intensive care units or during surgical procedures, where high measurement accuracy is essential [7, 8, 9].

Manual auscultation is a non-invasive technique that involves the use of a stethoscope placed over the brachial artery and a properly sized cuff that is inflated to temporarily occlude blood flow. As the cuff is gradually deflated, the operator listens for Korotkoff sounds while simultaneously monitoring cuff pressure via an external manometer. The appearance of the first Korotkoff sound indicates the onset of turbulent flow and corresponds to the SBP, whereas the disappearance of sounds marks the return to laminar flow and corresponds to the DBP. The main limitations of this method are its intermittent nature and the requirement for a trained operator [7, 10].

Oscillometry is a non-invasive and fully automated method, currently the most widely used technique in clinical practice. It involves placing a cuff around the upper arm which is gradually inflated and then deflated between supra-systolic and sub-diastolic pressures. During this process, the cuff pressure is continuously recorded, indicating the applied pressure and capturing small oscillations that reflect the pulsatile arterial blood volume. BP is estimated from the oscillogram amplitude using a specific algorithm. SBP and DBP are identified by the onset and disappearance of these oscillations, respectively. This method is also employed in ambulatory blood pressure monitoring (ABPM), which involves a wearable device programmed to take measurements at regular intervals, typically every 15 minutes, over a period of 24 to 72 hours. However, since the method still relies on arterial compression, it may cause discomfort for the patient, particularly during nighttime measurements, potentially leading to sleep disturbances. This is one of the primary reasons for non-adherence to the procedure [7, 8, 10].

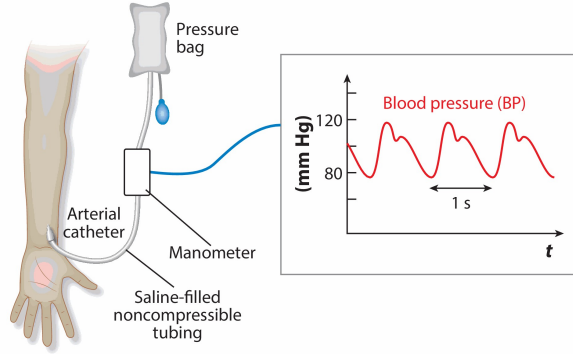
Volume clamping is a non-invasive and automatic method that use a cuff, placed around the finger, equipped with a PPG sensor and a manometer. Initially, the cuff pressure is gradually increased until it completely occludes blood flow then the cuff slowly deflates, allowing blood flow to resume. During this process, PPG sensor records blood volume oscillations while manometer records the cuff pressure. SBP is identified at the onset of oscillations, while DBP is determined when the oscillations stop. The system includes a rapid control mechanism that adjusts

the cuff pressure in response to blood volume variations, automatically inflating or deflating the cuff to maintain stable blood flow in the finger. This ensures that BP measurements remain accurate and reliable over time [7, 8].

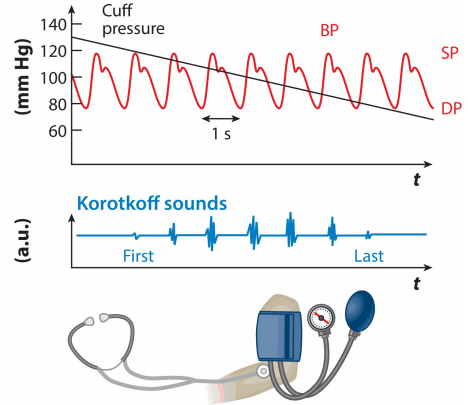
Tonometry is a non-invasive and continuous method for monitoring BP by applying a force sensor to the skin above large and superficial arteries. This technique requires that the artery be sufficiently applanated so that the transduced force is normal to the vessel wall and is uniformly distributed over the sensor's contact area. BP is then estimated as the detected force by the known sensor area. Proper sensor alignment is achieved using an array of small force sensors while the holding force is gradually adjusted automatically or manually. However, due to challenges in maintaining the contact force and accurately capturing the pressure wave, the device must be regularly calibrated [7, 8, 11].

The various conventional BP measurement methods present several issues. Although it is considered the gold standard due to its high accuracy, the invasive method carries significant risks for the patient such as hematoma formation, air embolism, blood loss, vasospasm, systemic infection, nerve damage and injury to nearby tissues. The auscultatory and oscillometric methods are the most commonly used technique for measuring BP. However, these methods require an inflatable cuff, which makes them unsuitable for continuous monitoring and may lead to discomfort due to physiological limitations. Tonometry and volume clamp can provide continuous measurements; however, their accuracy is highly sensitive to motion artifacts caused by sensor placement and patient movement [7].

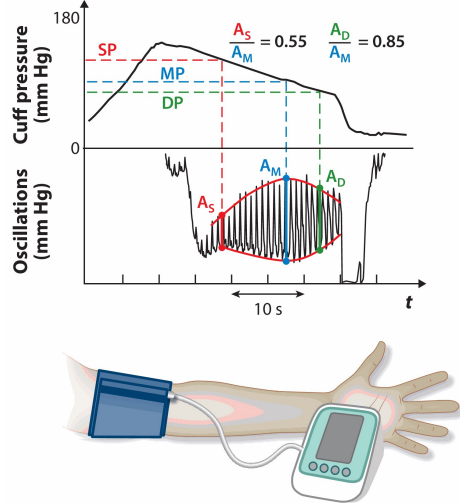
a Catheterization



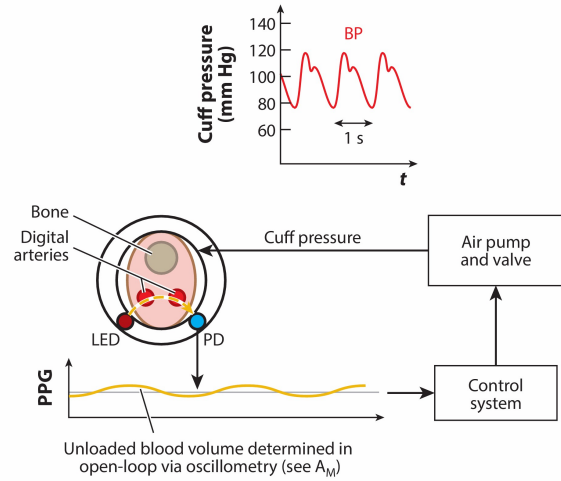
b Auscultation



c Oscillometry



d Volume clamping



e Tonometry

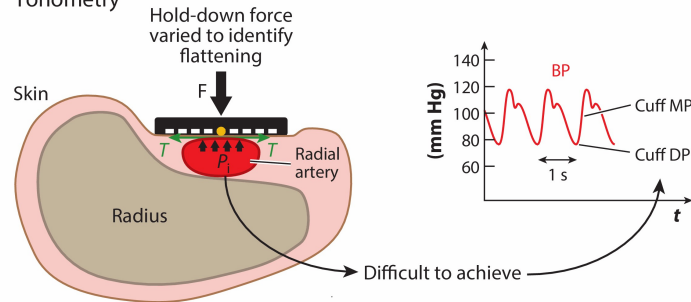


Figure 1.2. Conventional BP measurement methods: catheterization (a), auscultation (b), oscillometry (c), volume clamping (d) and tonometry (e) [7]

1.3 Electrocardiography

Electrocardiography (ECG) is a method for recording cardiac electrical activity and provides an essential tool for detecting heart disease. The ECG signal is recorded by placing electrodes on the surface of the body, which detect small potential differences between various recording sites that change during the cardiac cycle [12]. The ECG signal typically has an amplitude on the order of millivolts (mV), with overall variations ranging from 0.1 to 5 mV. For diagnostic purposes, clinically relevant information about cardiac activity is found within the frequency range of 0.05 to 120 Hz, while for monitoring applications, a narrower range of 0.5 to 40 Hz is typically used [13].

1.3.1 Electrocardiographic recording

Cardiac potentials are recorded using at least two electrodes applied to the body surface. The measurable potential differences depend on electrode placement, which should be standardized to ensure reliable comparison of ECG recorded from different individuals or over time in the same individual, especially for diagnostic purposes [12, 13]. ECG recordings are based on the standardization system proposed by Willem Einthoven, in which human body is considered as a large-volume conductor with the heart assumed to be the central source of electrical activity. Electrodes are positioned on the right arm (RA), left arm (LA) and left leg (LL), defining an imaginary equilateral triangle built around the heart. A potential difference, known as a lead, can be recorded from each pair of electrodes [12, 13]. The three leads defined by Einthoven are:

- Lead I: measures the potential difference between RA and LA.
- Lead II: measures the potential difference between RA and LL.
- Lead III: measures the potential difference between the LA and LL.

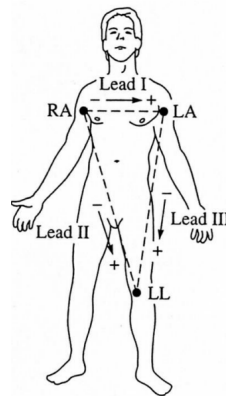


Figure 1.3. Einthoven triangle [13]

1.3.2 Electrocardiographic waveform

A physiological ECG waveform, as shown in Figure 1.4, exhibits a well-defined shape, characterized by distinct deflections that reflect the generation and propagation of the electrical impulse through the heart's conduction system. The reference morphology is generally based on the second lead [12]. A typical ECG trace is characterized by three main features:

- P wave: appears as the first deflection in the ECG signal, reflecting atrial depolarization that triggers the contraction of both atria, driven by the impulse generated at the sinoatrial node.
- QRS complex: represents ventricular depolarization and consists of a sequence of downward and upward deflections.
- T wave: represents ventricular repolarization and typically appears as an upward deflection.

The physiological state of the heart can be assessed by analyzing intervals and segments of the ECG signal:

- PR interval: represents the time interval between the onset of the P wave and the beginning of the QRS complex, reflecting the conduction time from atria to ventricles.
- QT interval: reflects the total duration of ventricular depolarization and repolarization.
- ST segment: represents the time interval between the end of the QRS complex and the beginning of the T wave, representing the period during which the ventricular cells are fully depolarized.
- PR segment: corresponds to the isoelectric line between the end of the P wave and the beginning of the QRS complex, representing the transition between atrial and ventricular activation.

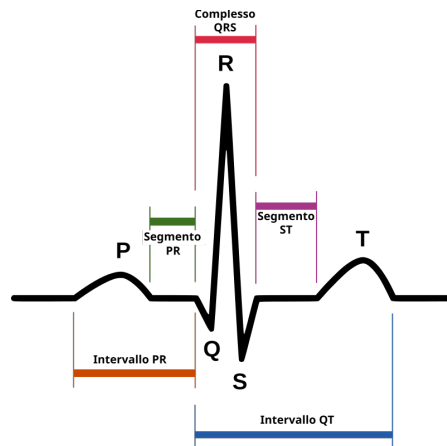


Figure 1.4. ECG waveform [14]

1.4 Photoplethysmography

Photoplethysmography (PPG) is an optical technique used to detect variations in blood volume at the skin surface. The basic form of PPG technology requires only two opto-electronic components: a light source and a photodetector. PPG is commonly employed non-invasively and operates at a red or a near infrared wavelength. PPG signal can be acquired from multiple anatomical locations, such as fingertip, wrist, earlobe, forehead, and ankle, each offering different levels of accuracy. In recent decades, the growing demand for small, reliable, low-cost, and simple-to-use non-invasive cardiovascular assessment techniques has been a key factor in the increased focus on PPG. This optical technique is now widely integrated into a wide range of applications, including clinical devices, activity trackers, and HR monitoring devices. One of the main challenges in employing PPG signals in wearable systems is their high sensitivity to motion artifacts caused by body movements [15, 16].

1.4.1 Photoplethysmographic recording

A PPG device consists of a light source that emits light towards the tissue and a photodetector that captures variations in light intensity caused by fluctuations in blood volume in the illuminated tissue. The detected light is converted into electrical signals, which are then recorded to obtain the PPG signal. Changes in blood volume can be estimated based on the detected light intensity [15, 16].

PPG devices generally use infrared or green light-emitting diodes (LED) as light sources. Wavelength selection is a critical factor influencing both the quality of the PPG signal and the device's susceptibility to motion artifacts. Green wavelengths have shown considerable benefits compared to traditional infrared-based PPG devices. As longer wavelengths are known to penetrate deeper into tissue, the LED can be adjusted to emit light ranging from 532 nm (green light) for superficial measurements to 1064 nm (infrared) when deeper tissue penetration is required [15, 16].

Depending on the relative positioning of the light source and the photodetector, PPG can be implemented in two main configurations:

- Transmission mode: the tissue is located between the emission source and the detection unit. The photodetector captures the portion of light that propagates through the tissue without being absorbed.
- Reflection mode: The light source and the photodetector are placed adjacent to each other on the same side of the tissue. A portion of light is reflected and captured by the photodetector.

The selection of the appropriate PPG configuration depends on both the anatomical measurement site and the specific application. Thanks to its functional mechanism, the transmission mode enables the acquisition of high-quality signals and is best suited for body sites that are easily accessible and light-permeable, such as the finger, ear, or foot, where the tissue is relatively thin. In contrast, the reflection mode offers greater flexibility in terms of application, as its two components are placed side by side, enabling measurements in different body locations. However, in this case the signal quality may be lower and more susceptible to external artifacts or ambient light interference [15].

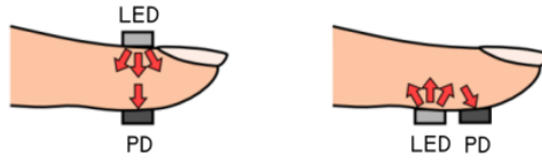


Figure 1.5. PPG configurations. Transmission mode (left) and reflection mode (right). [16]

1.4.2 Photoplethysmographic waveform

PPG waveform consists of a direct (DC) and an alternating (AC) component. The AC component, also referred to as the pulsatile component, reflects the pulsatile nature of blood flow. It varies synchronous with each heartbeat and is therefore used to assess changes in the blood volume. In contrast, the DC component remains relatively constant and represents the non-pulsatile part of the signal, which is influenced by respiration, vasomotor activity, autonomic nerve activation, and thermoregulation. During the cardiac cycle, the DC level does not significantly change, as it is not influenced by the rhythmic variations in blood flow [15, 16].

A physiological PPG waveform is typically divided into two phases: the anacrotic phase, corresponding to the rising edge of the pulse and primarily associated with systole, and the catacrotic phase, representing the falling edge and reflecting both diastole and the return of pressure waves from peripheral sites. In individuals with healthy compliant arteries, a dicrotic notch typically appears during the catacrotic phase [15].

As shown in Figure 1.6, a typical PPG waveform comes with four fiducial points: onset, systolic peak, dicrotic notch and diastolic peak. These points, along with those identified in the first derivative of the PPG signal (VPG) and in the second derivative of the PPG signal (APG), are essential for extracting a wide range of morphological features, such as areas, amplitudes, and widths, that provide valuable information for the diagnosis of cardiovascular diseases [17].

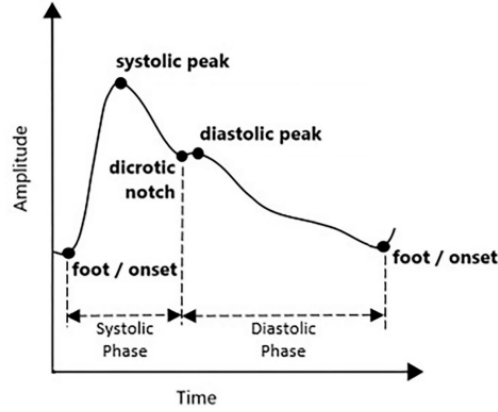


Figure 1.6. PPG waveform [17]

1.5 Electrocardiography and photoplethysmography for blood pressure monitoring

Recent studies have demonstrated that PPG signal, when used in combination with ECG can be employed to extract Pulse Transit Time (PTT). Since blood propagates from the heart to other parts of the body in the form of a pressure wave, PTT is defined as the time it takes for this pulse wave to travel from the heart to a peripheral site in the vascular system. It can be calculated as the time interval between the R-peak of the ECG, indicating the onset of ventricular contraction, and the systolic peak of the PPG, which reflects the arrival of the pulse wave at a distal site [18, 19]. Therefore, the PTT is influenced by multi-factors, such as cardiac output, BP, stiffness of the arterial wall, venous return, age, and temperature [20].

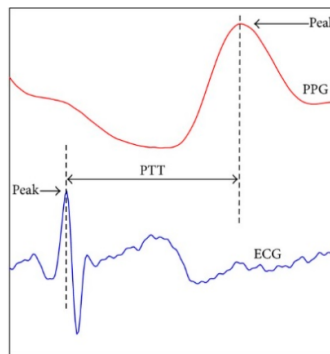


Figure 1.7. Measurement of PTT using PPG and ECG [18]

Previous studies have confirmed that variations in PTT reflect changes in BP and have also demonstrated a high coherence between PTT and BP fluctuations [20, 21]. It has been demonstrated that SBP can be estimated using Pulse Wave Velocity (PWV), which represents the speed at which the pulse wave propagates through the arteries between two specific points in the vascular system. PWV is calculated by dividing the distance between the two points (d) by PTT, according to Equation 1.1:

$$PWV = \frac{d}{PTT} \quad (1.1)$$

where d refers to the distance between heart and peripheral site which changes according to individual height. PWV is strongly correlated with arteries elastic properties and this relationship is formally described by the Moens-Korteweg:

$$PWV = \sqrt{\frac{Et}{2R\rho}} \quad (1.2)$$

where R is the internal radius of arteries, ρ is the blood density, t is the vessel wall thickness and E represents arterial elasticity and is exponentially related to BP through the following Equation 1.3:

$$E = E_0 e^{\alpha P} \quad (1.3)$$

where E_0 is the arterial elasticity at zero pressure, P is the arterial BP and α a constant.

Therefore, by combining Equations 1.1 and 1.2, the relationship between PTT and BP is obtained, allowing the determination of the BP through the following equation:

$$\frac{d}{PTT} = \sqrt{\frac{tE_0 e^{\alpha P}}{2R\rho}} \quad (1.4)$$

In this way, an inverse relationship between PTT and BP has been demonstrated; however, it depends on various physiological factors that may not be directly measurable and can change among individuals. Therefore, initial calibration, based on empirical data or regular subject-specific recalibrations, is essential to obtain pseudo-personalized parameters [22, 23].

1.6 Contributions and thesis organization

This thesis focuses on the challenging problem of estimating BP from non-invasive physiological sensor data. The motivation for this work derives from the increasing demand for solutions that are reliable, accurate, and suitable for real-world applications.

The aim of this work is to develop and evaluate models that capture the relationship between physiological signals and BP, comparing regression approaches and machine learning (ML) techniques under realistic validation scenarios relevant to potential clinical applications. This thesis makes several key contributions to the current state of research in cuffless BP estimation:

- For regression methods, the study conducts a systematic comparison of various models to identify the best-performing approach. These models differ in terms of the number of coefficients and the types of parameters involved. In contrast to previous studies that often rely on continuous segments of signal data for calibration, we propose a calibration strategy based on a small number of discrete BP reference values. This approach is designed to emulate realistic clinical scenarios in which only a few measurements are available, such as during periodic check-ups or initial device setup. The aim is to identify the best-performing combination of regression model and number of reference BP values needed to obtain accurate and reliable estimation.
- In ML approach, we design a feature extraction pipeline that includes an extended and more varied set of features compared to those commonly adopted in the literature. This richer representation is intended to better capture the dynamics of physiological signals and supply the model with more informative inputs, potentially improving its predictive performance. To ensure a realistic evaluation of model generalization, we adopt a subject-independent validation strategy, where data from each individual is assigned exclusively to one of the sets. This choice is intended to overcome a common limitation found in many existing studies, where subjects appear across different data splits. Such overlap often leads to overoptimistic results, as models may exploit subject-specific characteristics rather than learning generalizable physiological patterns. By adopting subject separation, we aim to obtain a more reliable and realistic evaluation of performance, better reflecting the model's ability to generalize to unseen individuals.

The structure of this thesis is organized as follows. Chapter 2 provides a review of the most relevant works in the field, focusing on PPG-based methods, PTT-based approaches, and Artificial Intelligence (AI) techniques for BP estimation. Chapter 3 describes the materials and methods used in the study. Section 3.1 presents the dataset, while Section 3.2 outlines the data analysis pipeline, including preprocessing procedures (Section 3.2.1), regression models (Section 3.2.2), and ML-based approaches (Section 3.2.3). Chapter 4 reports the experimental results obtained and concludes with a comparison with existing methods. Chapter 5 discusses the implications of the findings, highlighting the strengths and limitations of the proposed approaches. Finally, Chapter 6 summarizes the main conclusions and outlines possible future developments.

Chapter 2

Related work

Considering the significant limitations of conventional BP measurement methods, in combination with the need for non-invasive continuous monitoring, this topic has attracted growing attention in recent years, leading to the development of numerous innovative approaches.

2.1 Photoplethysmography-based methods

PPG provides a potential alternative for monitoring BP in a non-invasive and continuous manner. PPG devices have traditionally been employed for the assessment of HR and blood oxygen saturation, given their affordable price and portability. Thanks to these advantages, growing interest has recently been shown in the use of PPG for BP monitoring, particularly through the application of ML and Deep Learning (DL) techniques.

Khalid et al. proposed a single PPG-based algorithm for BP estimation and applied three ML techniques (regression tree, multiple linear regression, and support vector machine) to compare their accuracy in classifying BP categories. Raw PPG signals and its corresponding BP reference were extracted from the University of Queensland Vital Signs Dataset that contains a wide range of patient monitoring data and vital signs recorded during surgical cases.; only segments with acceptable quality were saved and subsequently processed through baseline removal and normalization. Three most significant pulse features, along with their corresponding BP references, were used to train and test ML algorithms. BP estimation was obtained by 10-fold cross-validation. The regression tree demonstrated superior overall estimation accuracy compared to the other techniques, resulting in a Mean Absolute Error (MAE) and standard deviation (SD) of -0.1 ± 6.5 mmHg for SBP and of -0.6 ± 5.2 mmHg for DBP. The other two techniques failed to achieve

acceptable results [24].

Mousavi et al. proposed an algorithm that uses whole-based features from PPG signals for estimating BP. A subset of the MIMIC II database is used as PPG and BP reference data source. To reduce noise, a Fast Fourier Transform (FFT) filter was applied, followed by the extraction of whole-based features. The dataset was divided using 10-fold cross-validation, and several ML techniques were evaluated: support vector regression, decision tree regression, adaptive boosting regression, and random forest. Among these, adaptive boosting regression obtained the best performance, achieving the lowest MAE and SD: 0.187 ± 4.173 mmHg for SBP and -0.050 ± 8.901 mmHg for DBP [25].

Yi et al. considered five algorithms to build a BP measurement model and assessed the performance of each. PPG and arterial blood pressure (ABP) were both derived from the MIMIC database. After baseline drift correction, nine morphological features are derived from each PPG signal. The dataset was split into training and testing sets, with 80% of the data used for training and the remaining 20% used for testing. ML algorithms employed in the study were: linear regression, lasso regression, classification and regression tree, elastic network, and K-nearest neighbors (KNN). Furthermore, the performance of these techniques was compared, with KNN achieving the highest accuracy in terms of MAE and SD: 1.66 ± 4.90 mmHg for DBP and 2.09 ± 5.61 mmHg for SBP [26].

Duan et al. proposed a framework that employs feature evaluation to identify the most relevant features for BP prediction model. PPG signals and BP measurements of the University of Queensland’s vital signs dataset are considered. As a first step, noise is reduced by applying average filters and wavelet transform, followed by the construction of a feature pool and the definition of three criteria to select relevant features for BP prediction. Support vector regression was employed during the learning phase, using 10-fold cross-validation to evaluate performance. This algorithm achieves prediction accuracy of 4.77 ± 7.68 mmHg for SBP and 3.67 ± 5.69 mmHg for DBP [27].

In most of the reviewed works, cross-validation is employed as the primary evaluation strategy. However, a crucial aspect that is often left unclear is the attention given to ensuring subject-independent data partitioning. In the absence of such clarification, it is possible that data from the same subject was distributed across both training and test sets. This would inadvertently result in a subject-specific evaluation scenario, leading to overestimated performance metrics. Consequently, the reported results, although impressive, may not reflect true generalization across

unseen subjects. It is possible to conclude that PPG-based methods frequently suffer from limited accuracy and still require substantial research to meet the requirements of medical validation standards. A potential solution could be the inclusion of additional physiological signals, such as ECG, to achieve higher accuracy. To overcome PPG-based methods limitations, recent approaches have focused on integrating PPG with ECG, using PTT-based methods and artificial intelligence based techniques [28].

2.2 Pulse transit time-based methods

The relationship between PTT and BP has been extensively studied in previous research and is currently recognized as a valid approach for BP estimation. Most recent methods are based on the Moens-Korteweg fluid dynamic law, previously presented in Equation 1.4.

It has been demonstrated that, assuming constant arterial diameter, vessel wall thickness, and arterial elasticity, a mathematical relationship can be established between PTT and BP [28]. As a result, several formulations have been proposed [22]:

- Logarithmic model: Moens-Korteweg’s equation gives a logarithmic relationship between BP and PTT and with the assumption done is possible to consider the following relationship:

$$BP = a \cdot \ln(PTT) + b \quad (2.1)$$

Where a and b are subject-specific and can be determined through regression analysis between the reference BP and the corresponding PTT. This model tends toward negative infinity as PTT approaches zero, making it difficult to use this relationship to represent small BP.

- Linear model: BP and PTT can be linearly related by differentiating the Moens-Korteweg equation, leading to the following expression:

$$BP = a \cdot PTT + b \quad (2.2)$$

Where the constants a and b are subject-specific and can be estimated by performing a regression analysis between the reference BP values and the corresponding PTT measurements. Several other studies have integrated the linear model with other influencing factors, such as HR and arterial stiffness index (ASI), that would affect BP.

One of the main limitations of PTT-based methods is the requirement for an initial calibration using a highly accurate reference method. Research shows that the

validity of parameters deteriorates within 24 hours after calibration, leading to increased estimation errors over time. Therefore, frequent recalibration is necessary to maintain accuracy.

Furthermore, several factors influence the relationship between PTT and BP, including posture, ambient temperature, physical activity, and individual characteristics such as arterial elasticity and vascular conditions. These factors can alter the correlation between PTT and BP, making challenging to achieve reliable and consistent estimations [22].

2.3 Artificial Intelligence-based methods

Despite the simplicity of PTT-based methods, due to their algebraic inverse relationship between PTT and BP, they have the significant drawback of requiring subject-specific calibration. For this reason, in recent years, the use of AI-based methods for BP estimation using PPG and ECG signals has gained increasing popularity. In this context, three categories of algorithms can be distinguished based on the type of input and output data they process [28].

- Feature-to-Label (*Feat2Lab*): refers to models that utilize handcrafted features extracted from signals as inputs to predict discrete BP values. The most significant and commonly utilized features include time-based, frequency-based, and statistical features.
- Signal-to-Label (*Sig2Lab*): includes models that directly take signals as input to generate discrete BP values as output.
- Signal-to-Signal (*Sig2Sig*): refers to models that generate continuous BP waveform starting from signals in input.

Chen et al. introduced a ML-based model to establish discrete BP values using PTT and characteristics of pulse waveform. ECG, PPG, and reference BP signals were extracted from the MIMIC-III Waveform Database. After removing corrupted segments and missing waveforms, denoising techniques were applied. Subsequently, relevant features were extracted and selected by applying the Mean Impact Value method, which reduces features redundancy. A genetic algorithm was introduced to implement parameter optimization. Support vector regression was employed to estimate BP from the selected features. The dataset was randomly split into a training set (80%) and a test set (20%), and a 5-fold cross-validation was performed. The model achieved promising results in terms of MAE and SD: 3.27 ± 5.52 mmHg for SBP and 1.16 ± 1.97 mmHg for DBP.[29].

He et al. used a random forest algorithm to explore the relationship between ECG, PPG, and reference BP signals. In this study, features were extracted by PPG and ECG signals from the MIMIC II database. Feature importance is assessed through correlation analysis and ranking using the random forest algorithm, enabling the identification of the most significant variables. Subsequently, a random forest regression model is trained and evaluated using 10-fold cross-validation. Results show that this method performs better than the commonly used PTT method achieving error in terms of MAE and SD: 8.29 ± 5.84 mmHg for SBP and 4.44 ± 3.72 mmHg for DBP [30].

Miao et al. conducted a study to estimate continuous BP by integrating a mechanism-driven model with data mining techniques. Features were extracted from ECG and PPG signals and a genetic algorithm was employed to identify the most relevant indicators for each individual. Signals were acquired with the patients in resting and seated conditions. Multivariate linear regression and support vector regression models were trained under static, dynamic, and long-term follow-up conditions. Experimental results demonstrated an improved accuracy compared to traditional PTT-based methods. In particular, static BP estimation achieved excellent accuracy: -0.001 ± 3.102 mmHg for SBP, and -0.004 ± 2.199 mmHg for DBP [31].

Kachuee et al. proposed a framework for predicting BP values by extracting physiological parameters from ECG and PPG signals from MIMIC II in combination with ML algorithms. After pre-processing and denoising the raw signals, two types of features, physiological parameter-based and whole-signal based, were extracted. Dimensionality reduction is utilized to reduce the feature length. Four ML algorithms were employed. Model performance was evaluated using MAE and SD of the estimation errors: 11.11 ± 10.09 mmHg for SBP and 5.35 ± 6.14 mmHg. An optional subject-specific calibration process is proposed to improve estimation performance [32].

In the development of ML models, it is standard practice to divide the dataset into training, validation, and test sets. However, BP datasets often contain multiple records for the same individual. A commonly used approach is to split the data into folds based on a uniform probability distribution. This strategy can result in data from the same subject appearing in different sets, leading to information leakage. This issue is particularly relevant in typical cross-validation strategies, which shuffle data before partitioning, potentially placing segments from the same subject into multiple subsets at once. Moreover, it increases the risk of generating folds that contain few or no instances of underrepresented BP values.

A comparative study evaluated models from the three different categories under two conditions: with information leakage versus without. In the leakage condition, data partitions were formed using uniform random splitting, allowing segments from the same subject to appear in multiple folds. In all cases, models performed substantially better in the leakage condition. Specifically, the Mean Absolute Scaled Error (MASE) for both SBP and DBP showed a dramatic reduction, from around 92–98% in the leaked condition down to below 60% when leakage was prevented. Besides, the significant drop is also shown in the SD metric. These findings highlight how overlooking subject overlap during data partitioning can lead models to rely on subject-specific features rather than learning generalizable patterns, with the consequence of overestimated performance and potentially misleading conclusions [28].

Chapter 3

Materials and Methods

3.1 Data

The Cuff-Less Blood Pressure Estimation Dataset is used in this work as the reference database. The main aim of this dataset is to offer clean and reliable signals suitable for developing cuffless BP estimation algorithms. Derived from the MIMIC-II Waveform Database, this dataset includes ECG, PPG, and ABP signals originally acquired from intensive care unit patients using the Philips CareVue Clinical Information System (models M2331A and M1215A; Philips Health-care, Andover, MA). Data were divided into fixed-size blocks, and cleaned by applying smoothing and removing blocks with abnormal BP or HR values, severe discontinuities or with altered consistency. Waveform signals were sampled at a frequency of 125 Hz with at least 8-bit accuracy and a minimum duration of 1 minute. In this work, a subset of the dataset was used by extracting ECG, PPG, and ABP signals corresponding to the first 100 patients. The provided MATLAB file is organized as a cell array of matrices, with each cell representing a record. In each matrix, rows correspond respectively to PPG signal acquired from the fingertip, ABP signal measured invasively via catheterization, and ECG signal recorded from lead II [33].

3.2 Data analysis

3.2.1 Pre-processing

A pre-processing strategy was chosen to enhance signal quality while preserving relevant physiological information, and it is well-suited for both ML and regression-based analyses. In both cases, the quality of the input signals is a critical factor influencing the reliability and accuracy of the resulting BP estimation. Therefore,

the same pre-processing strategy was uniformly applied to support robust and comparable performance across different analytical methodologies.

Based on this strategy, specific filtering operations were applied to the signals according to their characteristics and acquisition method. Since ABP signals were recorded invasively, they were largely free of artifacts and did not require additional filtering. In contrast, PPG and ECG signals, being more susceptible to noise and motion artifacts, were processed using a 4th-order Butterworth band-pass filter. Cutoff frequencies were set to 0.5–10 Hz for PPG and 0.5–40 Hz for ECG, effectively reducing noise and interferences while preserving physiologically relevant information. Figures 3.1 and 3.2 show the effect of filtering in 10-second segments of raw and filtered ECG and PPG signals, extracted from Subject 1. In the ECG trace, baseline drift and high-frequency noise have been effectively removed, and the PPG signal appears smoother and more regular post-filtering.

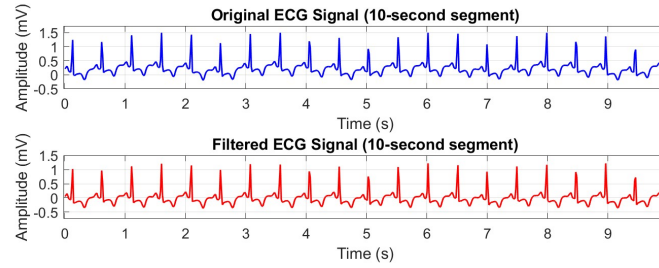


Figure 3.1. Comparison between raw and filtered ECG of 10-second signal from Subject 1. Top panel: original ECG signal. Bottom panel: ECG signal after filtering.

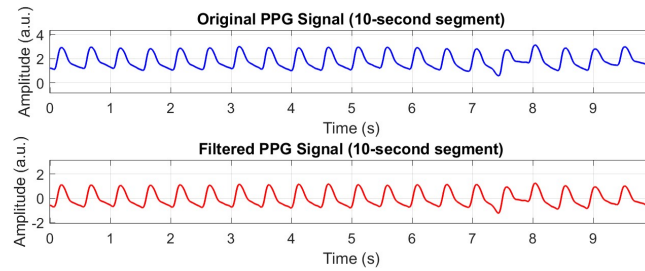


Figure 3.2. Comparison between raw and filtered PPG of 10-second signal from Subject 1. Top panel: original PPG signal. Bottom panel: PPG signal after filtering.

3.2.2 Regression analysis

As discussed in Section 2.2, the relationship between PTT and BP has been extensively investigated, and it is well established that this relationship can be described using either linear or logarithmic models.

In this study, we selected three representative models from the literature, including one that incorporates HR, which has been recognized as a relevant influencing factor in BP estimation. A key novelty of our approach lies in training these models using only a limited number of BP measurements per subject. This strategy aims to assess the feasibility and reliability of subject-specific calibration when only a small amount of data is available, a scenario that closely reflects real-world situations.

In order to obtain PTT, it is necessary to accurately identify the R-peaks in the ECG and the systolic peak in the PPG. For this purpose, two algorithms have been applied:

- Pan-Tompkins is a real-time algorithm for detecting QRS complexes in ECG signals. It performs digital analysis of the slope, amplitude, and width, combined with adaptive thresholds that automatically adjust to variations in HR and signal morphology.

The MATLAB implementation called *pan_tompkin* was used to detect R-peaks in the ECG signal. This function takes as input an ECG signal vector and the sampling frequency in Hz, and returns the amplitudes and indices of the detected R waves, along with the delay introduced by the process [34, 35].

- Multi-Scale Peak and Trough Detection (MSPTD) is an improved algorithm designed to detect peaks and troughs in physiological signals. It operates using a scalogram-based approach, which analyzes the signal at multiple scales. It calculates a local maxima scalogram that identifies potential peaks at various zoom levels, marking these points based on their prominence relative to neighboring values.

The MATLAB implementation called *msptd_beat_detector* was used to detect peaks and onsets in the PPG signal. This function takes as input a vector of PPG values and the sampling frequency in Hz, and returns the indices of the detected pulse peaks and onsets [36, 37]

Figures 3.3 and 3.4 illustrate representative 10-second segments of the ECG and PPG signals with the detected points superimposed. In the ECG, R peaks are highlighted, while in the PPG, both onsets and systolic peaks are identified.

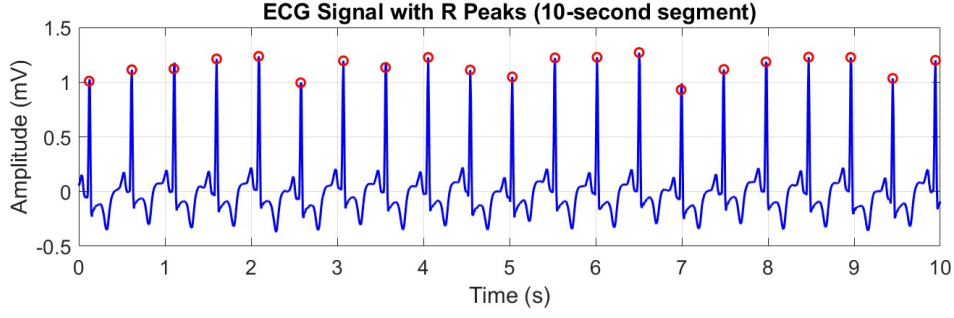


Figure 3.3. Representation of a 10-second segment of the filtered ECG signal, illustrating the detected R peaks (red markers).

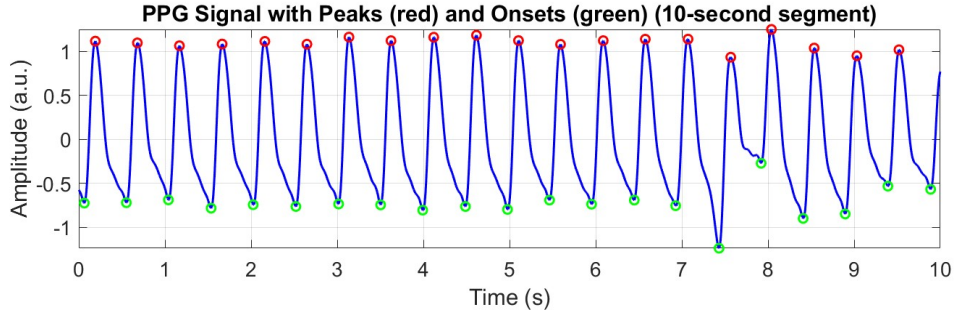


Figure 3.4. Representation of a 10-second segment of the filtered PPG signal. Green markers indicate pulse onsets, while red markers indicate the systolic peaks.

In this way, PTT values can be obtained by calculating the time interval between the R peaks of the ECG signal and the systolic peak of the PPG signal. HR plays a significant role in BP estimation. It is physiologically linked to vascular tone, cardiac output, and autonomic nervous system activity, and represents a valuable complementary parameter in BP modeling. In particular, when BP decrease, the autonomic nervous system is activated and responds by increasing HR and inducing vasoconstriction to restore pressure levels. This regulatory mechanism suggests that BP is influenced not only by HR, but also by its own previous values [38].

In this work, HR was derived from the ECG signal by measuring the time interval between two successive R peaks, known as RR interval. The HR was then calculated using the following expression:

$$HR = \frac{60}{RR_{interval}} \quad (3.1)$$

where the RR interval is measured in seconds, and the resulting HR is expressed in beats per minute (bpm).

In this work, three different regression models were developed to estimate BP based on physiological parameters. Each model differs in the number of coefficients and parameters involved:

- Two-coefficient linear model: represents the simplest approach, where BP is modeled as a linear function of PTT. It is defined by the following equation:

$$BP = a \cdot PTT + b \quad (3.2)$$

- Three-coefficient linear model: extends the previous one by incorporating HR as an additional variable. The model is expressed as:

$$BP = a \cdot PTT + b \cdot HR + c \quad (3.3)$$

- Four-coefficient logarithmic model: introduces a logarithmic transformation of PTT and includes a term based on the previous BP value. This model aims to capture more complex physiological dynamics and is defined as:

$$BP_n = a \cdot \ln PTT + b \cdot HR + c \cdot BP_{n-1} + d \quad (3.4)$$

In all three models, the coefficients were estimated using the least squares method. This method finds the values of the coefficients that make the model's predictions as close as possible to the actual BP values. Specifically, it works by minimizing the sum of the squared differences between the measured BP values and those predicted by the model, as shown in the following expression:

$$\sum_{i=1}^n (BP_i - \hat{BP}_i)^2 \quad (3.5)$$

In this formula, BP_i represents the actual BP for the i -th observation, and \hat{BP}_i is the corresponding predicted value from the model. The smaller this quantity, the better the model fits the data. The least squares method is widely used due to its simplicity, efficiency, and closed-form solution in regression settings [39].

Multiple linear regression was used to estimate the model coefficients by applying the least squares method, which minimizes the sum of the squared differences between the observed and predicted values. The approach involves constructing an input matrix composed of relevant predictor variables, such as PTT, HR, and previous BP measurements, and solving for the coefficients that best fit the measured BP values. Different model structures were implemented by modifying the

composition of the input matrix. For example, in the two-coefficient model, the matrix includes PTT values and a column of ones to account for the intercept term. In the three- and four-coefficient models, additional columns were added for HR and previous BP values as needed.

The regression process was structured into two main phases: a training phase, in which model parameters were estimated, and a testing phase, in which the performance of the models was evaluated. A key characteristic of the proposed approach is that, during training each model is calibrated using only a limited number of BP values per subject. Unlike many existing methods, this strategy aims to replicate realistic conditions where only a few BP measurements may be available for calibration.

For each subject, a fixed number of BP values was used to train the models, specifically, 5, 10, 20, 50, or 100, along with the corresponding input features, such as PTT and HR. These data were used to estimate the regression coefficients for each of the three models under study.

Once the models were trained using the selected BP values, the remaining data for each subject were used for testing. In this phase, the trained models were applied to the unseen PTT and HR values to predict BP. In the case of the four-coefficient logarithmic model, predictions were made recursively, with each output depending also on the previously estimated BP value. This subject-specific training approach, based on minimal calibration data, allowed for a robust evaluation of the models' ability to generalize and provide accurate predictions under data-limited condition.

3.2.3 Machine learning analysis

As mentioned in Section 2.3, traditional PTT-based methods are limited by the need for subject-specific calibration. This limitation has led to a growing interest in AI-based techniques for estimating BP from ECG and PPG signals.

Among the different categories, this work focuses on the Feature-to-Label approach, which involves extracting a predefined set of features from ECG and PPG signals to predict discrete BP values. In developing and evaluating the proposed model, we employed a subject-independent validation strategy that prevents subject-wise information leakage across training, validation, and test sets. In other words, data from the same individual is never included in more than one subset. This is essential to ensure that the model learns patterns that can generalize across subjects, rather than relying on individual-specific characteristics. Avoiding such leakage is crucial, as it can lead to unrealistically high performance estimates that do not reflect the model's true ability to generalize to unseen individuals.

Fiducial points extraction. The first crucial step in this process is the extraction of fiducial points from the ECG and PPG waveforms, which are necessary for subsequent feature computation. The detection process relies on the algorithms previously described in Section 3.2.2: the Pan-Tompkins algorithm was used to detect R-peaks in the ECG signal, while the MSPTD algorithm was employed to identify onsets and systolic peaks in the PPG waveform. After initial detection, a physiological criterion was applied to retain only those onset–peak pairs whose temporal separation falls within a plausible physiological range, between 0.1 and 0.4 seconds. This filtering step helps reduce false positives and enhances the reliability of the features derived in the next stages. These detected points serve for further morphological analysis. In particular, for each valid beat, the following additional fiducial points were extracted from the PPG:

- Maximum slope: computed as the peak of the first derivative between the onset and the systolic peak.
- Diastolic peak: located between the systolic peak and the next onset, based on local minima in the second derivative of the PPG signal.
- Dicrotic notch: identified from local maxima in the second derivative between consecutive onsets, selecting the second most prominent peak in amplitude.

This process produced a consistent and physiologically coherent set of fiducial points for each heartbeat, forming the basis for the computation of relevant features. The accuracy of this detection is essential to ensure the robustness and validity of the model.

Features extraction. Once the fiducial points had been reliably extracted from the ECG and PPG signals, a comprehensive set of features was computed to characterize each cardiac cycle [6, 29, 40, 41, 42]. These features aim to capture morphological, temporal, frequency-based, and nonlinear aspects of cardiovascular dynamics, and serve as the input for the predictive model [43]. The extracted features are organized into the following main categories:

- Pulse Wave Analysis Features
- Pulse Transit and Hemodynamic Features
- Amplitude-Based and Area-Derived Features
- Level Crossing and Complexity-Based Features
- Statistical Features

The first group consists of Pulse Wave Analysis Features, which describe the shape, timing, and dynamics of the PPG waveform within a single cardiac cycle. These features are illustrated in Figure 3.5:

- Cardiac Period (CP): Time interval between two consecutive systolic peaks of the PPG signal, representing the duration of one complete heartbeat.
- Systolic Upstroke Time (SUT): Time from the onset of a PPG beat to its systolic peak, associated with the rapid rise in blood flow during systole.
- Diastolic Time (DT): Time from the systolic peak of a beat to the onset of the next, corresponding to the gradual decrease in blood flow during diastole.
- Downstroke Width at x% (DW_x): Time for the signal to drop from the systolic peak to x% of the amplitude. Evaluated at 10%, 25%, 33%, 50%, 66%, and 75%.
- Sum of Upstroke and Downstroke Times at x% ($SW_x + DW_x$): Total time for the signal to rise to and fall from the peak through the same x% amplitude. Evaluated at 10%, 25%, 33%, 50%, 66%, and 75%.
- Downstroke-to-Upstroke Ratio at x% (DW_x/SW_x): Ratio between falling and rising durations at different amplitude levels. Evaluated at 10%, 25%, 33%, 50%, 66%, and 75%.

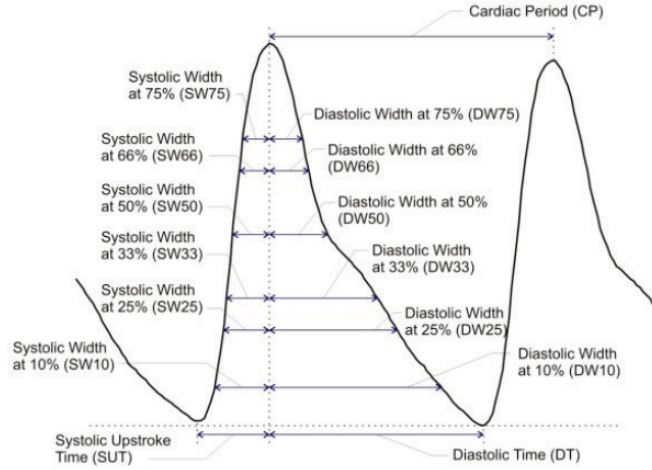


Figure 3.5. Schematic diagram of Pulse Wave Analysis Features. [40]

The second group includes Pulse Transit and Hemodynamic Features, which reflect arterial stiffness, wave reflection, and the temporal relationship between ECG and PPG events. These are illustrated in Figure 3.6.

- Augmentation Index (AI): Ratio between the diastolic and systolic peak amplitudes. It provides an estimation of arterial wave reflection.
- Large Artery Stiffness Index (LASI): Time interval between the systolic peak and the diastolic peak. It is associated with arterial compliance.
- S1, S2, S3, S4: Areas under the PPG curve between selected fiducial points; represent different phases of the pulse energy.
- PTT_p : Time delay between the ECG R-peak and the systolic peak of the PPG.
- PTT_f : Time delay between the ECG R-peak and the onset of the PPG waveform.
- PTT_d : Time delay between the ECG R-peak and the point of maximum slope in the PPG signal.
- HR: Derived from the interval between successive ECG R-peaks.

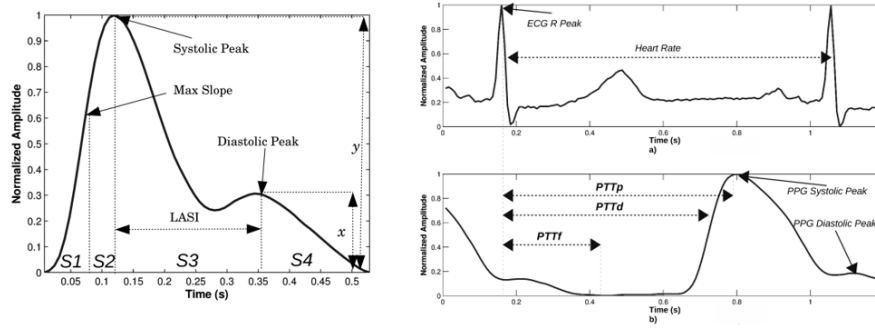


Figure 3.6. Schematic diagram of Pulse Transit and Hemodynamic Features. [6]

The third group comprises Amplitude-Based and Area-Derived Features, which provide a compact representation of energy distribution and relative waveform proportions. These are shown in Figure 3.7:

- K-value: Ratio between the mean PPG amplitude minus onset amplitude and the difference between systolic peak and onset amplitude. It is closely related to the peripheral resistance and the hardening degree of the arterial wall, and is one of the important physiological indicators for clinical research of cardiovascular diseases.
- T_{upr} : Ratio of the time from onset to systolic peak to total beat duration.

- T_{downr} : Ratio of the time from systolic peak to the next onset to total beat duration.
- C_{slope} : Systolic peak amplitude divided by upstroke time.
- H_{ar} : Ratio of amplitude at the maximum slope point to the systolic peak amplitude.
- H_{er} : Ratio of amplitude at the dicrotic notch to the systolic peak amplitude.
- H_{gr} : Ratio of amplitude at the diastolic peak to the systolic peak amplitude.
- $S1_{\text{norm}}$: Normalized area from onset to diastolic peak.
- $S2_{\text{norm}}$: Normalized area from diastolic peak to the end of the beat.
- $S1_{\text{norm}}/S2_{\text{norm}}$: Ratio of the systolic to diastolic normalized areas, indicating the balance of waveform energy.

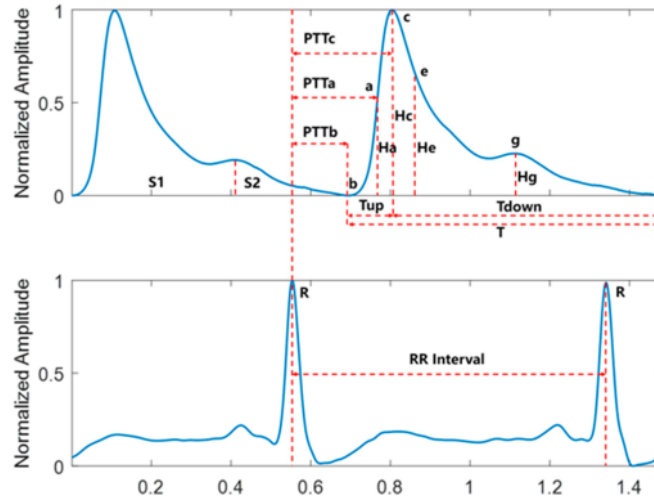


Figure 3.7. Schematic diagram of Amplitude-Based and Area-Derived Features. [29]

The fourth group includes Level Crossing and Complexity-Based Features, which are derived from the second derivative of the PPG waveform and aim to capture local signal irregularities. These are illustrated in Figure 3.8:

- Level Crossing Feature 1 (LCF1): Mean number of crossings per beat between the normalized APG and contour levels at 5% intervals from -100% to $+100\%$.

- Level Crossing Feature 2 (LCF2): Mean cumulative duration per beat during which the APG remains above the contour levels at 5% intervals from -100% to $+100\%$.

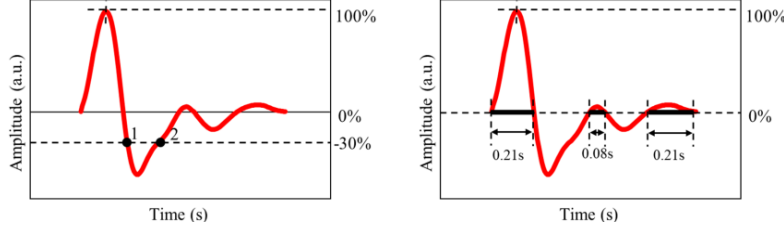


Figure 3.8. Schematic diagram of Level Crossing and Complexity-Based Features. [41]

The final group includes Statistical, Spectral, and Entropy-Based Features, which summarize global properties of the signals. These features were extracted from both ECG and PPG.

- Time-Domain Features: Describe the statistical and morphological characteristics of the signal in the time domain. These include: mean, variance, skewness, kurtosis, peak-to-peak interval, heart rate variability (HRV), crest factor, maximum, and root mean square (RMS).
- Frequency-Domain Features: Describe the distribution of signal power across frequencies using Welch's method. The extracted parameters are mean frequency, median frequency, and total band power.
- Chaotic and Fractal Features: Describe the complexity and irregularity of the signal, helping to reveal hidden dynamics in physiological data. The extracted measures include: approximate entropy, sample entropy, Shannon entropy, fuzzy entropy, permutation entropy, Higuchi fractal dimension, and Katz fractal dimension.

Internal consistency check. To ensure data integrity, an internal consistency check was conducted to verify that all feature vectors were aligned in length with the number of valid cardiac cycles detected for each subject. This process led to the exclusion of a subset of features whose values could not be reliably extracted for every beat or subject, often due to variability in signal quality or segmentation errors. In total, 24 features were excluded from further analysis, ensuring that only consistent and reliable variables were retained for the subsequent steps. Following the feature extraction phase, a series of operations were applied to improve data quality and prepare it for model training.

Outlier detection. The process began with the identification and correction of outliers. For each subject, outlier detection was performed separately on each feature using MATLAB’s `isoutlier` function, which defines an outlier as a value that lies more than three scaled median absolute deviations (MAD) from the median. Rather than removing these data points, a replacement strategy was adopted: each outlier was replaced with the average of its neighboring values, preceding and succeeding. This approach helps to preserve the continuity of the data while mitigating the effect of anomalous values, which could otherwise introduce bias or instability during model training. In contrast, records associated with abnormal reference BP values were directly excluded from the dataset. This decision was made to avoid exposing the model to physiologically implausible targets that could distort the learning process or compromise the evaluation.

Dataset splitting strategy. Once the data was refined, the complete dataset was divided into three distinct sets for training, validation, and testing. This division was performed at the subject level: all samples belonging to a given subject were assigned entirely to a single subset. This strategy guarantees the absence of subject-wise overlap between training, validation, and test data. As previously discussed, avoiding such information leakage is fundamental to ensure that the model learns general patterns rather than memorizing subject-specific characteristics. Subjects were randomly assigned to the three subsets with a 40%-30%-30% split for training, validation, and testing, respectively. After assigning subjects to each group, their corresponding feature were aggregated into the respective sets. This resulted in three distinct datasets, each containing the full set of features and labels, but with no overlap in subject identity.

Feature normalization. At this stage, all features were standardized using Z-score normalization. The mean and SD were computed exclusively from the training set, and these values were then applied to normalize the validation and test sets. This approach ensures that no information from the validation or test sets influences the training process in any way. This avoids subtle forms of data leakage that, although not involving target values directly, can still result in unrealistically optimistic performance if distributional characteristics of unseen data are implicitly incorporated during training, in full compliance with the subject-independent design.

Feature selection. Following the data preparation steps, feature selection was performed using Boruta, an algorithm that builds upon random forest models to evaluate the contribution of each variable in a robust and statistically manner. In this implementation, a random forest regressor was used with 100 trees and a maximum depth of 5, meaning that each tree can make up to five consecutive binary decisions from root to leaf. The Boruta algorithm was run with automatic estimation of the number of trees and a maximum of 50 iterations. Unlike conventional selection methods that aim to find a minimal optimal subset, Boruta is

designed to identify all the features that carry useful information for prediction. It achieves this by comparing the importance score of each actual feature with that of a set of artificially constructed shadow features, randomly permuted copies of the original variables, which serve as a baseline representing noise. Through multiple iterations, the algorithm assigns importance scores using a random forest estimator and evaluates whether each feature performs consistently better than the best of its shadow counterparts. Features that repeatedly outperform their shadows are marked as selected; those that do not are progressively rejected. Following this process, Boruta provided a classification for each feature, indicating whether it was selected or rejected based on its relative importance.

Model selection and evaluation. To identify the most suitable modeling approach, a comparative analysis was carried out by applying multiple regression algorithms to the prepared dataset. Each model was trained on the training set and evaluated on the validation set, with performance quantified through MAE and Pearson correlation coefficient. The mathematical expressions for both metrics are reported below:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (3.6)$$

$$\text{Pearson correlation coefficient} = \frac{\sum_{i=1}^n (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2} \sqrt{\sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})^2}} \quad (3.7)$$

where y_i is the true value for the i -th sample, \hat{y}_i is the corresponding predicted value, \bar{y} is the mean of the true values, $\bar{\hat{y}}$ is the mean of the predicted values, and n is the total number of samples. MAE provides a measure of the average absolute difference between predictions and actual values, while the Pearson correlation coefficient quantifies the strength and direction of the linear relationship between them. Based on the results obtained on the validation set, the model that demonstrated the best performance was selected and subsequently applied to the test set for final evaluation. The models tested included gaussian process regression (GPR), ensemble methods based on decision trees, linear regression, support vector regression, and regression trees.

Chapter 4

Results

4.1 Regression analysis

The results of the three previously described regression methods were evaluated on a dataset of 100 subjects. For each method, performance was evaluated in terms of MAE and SD, calculated across different training set sizes ranging from 5 to 100 reference BP values. The analysis is structured to first compare the overall performance of the models, then investigate the impact of the number of calibration points, and finally identify the best trade-off between accuracy, robustness, and simplicity.

Comparison of Estimation Methods. The three BP estimation models considered in this study are distinguished by both formulation and complexity, as shown in Equations (3.2), (3.3), and (3.4). Method 1 relies on a simple linear model with two coefficients. Method 2 adds HR, extending the basic model to three-coefficients. Method 3 further increases complexity by incorporating a logarithmic transformation of PTT and a recursive term based on the previous BP estimate, resulting in a four-coefficient model. The performance of the different models is summarized in Table 4.1.

Method 3, being the most complex, might be expected to offer the best performance. However, its results with only 5 BP values shows a remarkably high error of 13.642 ± 12.342 mmHg, indicating instability due to the insufficient data for parameter estimation. As the number of training points increases, the error decreases substantially, down to 5.312 ± 4.467 mmHg with 100 BP, suggesting that this model demands significant calibration.

Method 2, while simpler than Method 3, still requires estimation of three parameters. Its performance is relatively stable across training sizes, but its MAE remains consistently slightly higher than that of Method 1. For instance, with 5 BP values, the MAE is 6.293 ± 5.028 mmHg, compared to 5.309 ± 4.561 mmHg for Method

1.

Method 1, based on a basic linear regression, shows competitive or superior performance across all training set sizes. With only 5 calibration points, it achieves an MAE of 5.309 ± 4.561 mmHg, which is already better than Method 2 and far better than Method 3 under the same condition. It is remarkable that its performance does not improve significantly as the number of calibration points increases, suggesting a degree of robustness and stability even with minimal calibration.

Influence of Calibration Size. An important aspect to consider is how the number of BP reference values used during training impacts the accuracy of each method. For Method 1, the MAE remains relatively consistent across different training sizes. This indicates that even a small number of calibration points is sufficient to fit the model parameters. In Method 2, performance appears slightly more sensitive to training size, although the improvements beyond 20 BP values are marginal. Method 3 exhibits a pronounced dependency on the number of training points. Its error reduces from 13.642 mmHg (5 BP) to 5.312 mmHg (100 BP), showing clear benefits from a larger calibration set. However, the large variance at low training sizes raises concerns about stability in real-world scenarios with limited calibration.

Calibration–Method Compromise. Based on the above analysis, Method 1 with 5 calibration points emerges as the best compromise between estimation accuracy and model simplicity. It provides low MAE values comparable to more complex models trained on far more data. These results allow to reduce the need for extensive calibration and complex models, a significant advantage in real-world scenarios.

| BP values | Method 1 | Method 2 | Method 3 |
|-----------|-------------------|-------------------|---------------------|
| 5 | 5.309 ± 4.561 | 6.293 ± 5.028 | 13.642 ± 12.342 |
| 10 | 5.092 ± 4.508 | 6.036 ± 4.981 | 8.869 ± 9.221 |
| 20 | 5.131 ± 3.974 | 5.166 ± 3.861 | 6.014 ± 7.569 |
| 50 | 5.545 ± 5.114 | 5.511 ± 4.988 | 5.953 ± 5.280 |
| 100 | 5.381 ± 4.431 | 5.400 ± 4.570 | 5.312 ± 4.467 |

Table 4.1. Results in terms of MAE \pm SD for the three regression methods evaluated with different numbers of BP values.

To further illustrate the performance consistency of Method 1 with minimal calibration, Figure 4.1 shows the distribution of the MAE across all subjects when using only 5 BP samples for training. The box plot highlights the presence of a compact interquartile range and a relatively small number of outliers, suggesting that the majority of predictions fall within a narrow error band. The cumulative distribution function confirms that approximately 90% of the subjects exhibit an

MAE below 10 mmHg. Furthermore, the distribution of MAE reveals that approximately 60% of subjects attain a MAE of 5 mmHg or less, and about 30% fall between 5 and 10 mmHg. This visual evidence supports the quantitative findings previously reported, reinforcing the notion that Method 1 achieves both stable and accurate results with minimal calibration effort. These characteristics make it particularly suitable for real-world deployment, where collecting a large number of reference BP values may not be feasible.

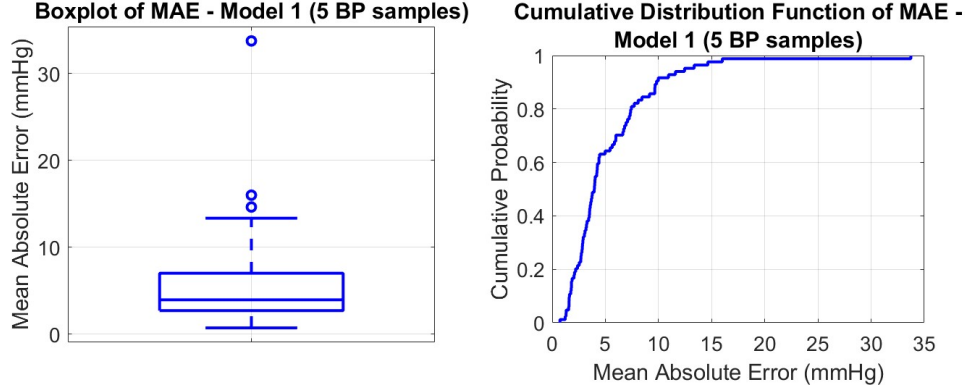


Figure 4.1. Box plot (on the left) and cumulative distribution function (on the right) of MAE for Model 1 using 5 BP samples for training.

Two additional experiments were conducted exclusively on Method 1 using 5 BP reference values, with the aim to test the model under less favorable or more generalized scenarios. In the first test, we examined whether the coefficients estimated for a single subject could be generalized to other subjects. Specifically, we computed the regression parameters using the 5 BP values from subject 1 and applied the resulting linear model to estimate BP for all other subjects in the dataset. Performance was again measured using MAE and SD across all test subjects. The results were significantly worse than those obtained when using subject-specific calibration: 23.706 ± 35.055 mmHg. This drastic degradation in performance highlights a critical limitation of the model: the high inter-subject variability in the physiological relationship between PTT and BP. Although Method 1 demonstrates stability and robustness when trained individually for each subject, its parameters do not transfer well across individuals. This is expected, given that physiological signals like PTT are influenced by various personal factors including vascular compliance, arterial stiffness, and individual cardiovascular dynamics. Therefore, applying a universal set of parameters leads to poor estimation accuracy and high dispersion in errors, indicating that at least a minimal degree of subject-specific calibration remains essential.

In the second test, we explored the performance under a more conventional validation scheme. Rather than selecting a limited number of BP reference points per subject, we performed a subject-level split of the dataset: 70% of the subjects were used for training, to estimate coefficients, and the remaining 30% were used for testing. This setup obtained the following result: 25.920 ± 23.980 mmHg. Again, the estimation error increased considerably compared to the individualized calibration strategy. The poor performance can be attributed to the inability of a global linear model to capture the subject-specific nature of the PTT–BP relationship, resulting in parameters that do not generalize well to unseen individuals.

4.2 Machine learning analysis

The evaluation of the proposed ML pipeline was carried out following the methodology previously described. We begin by examining the output of the Boruta algorithm, analyzing the set of features it identified as relevant. Next, we present the results obtained on the validation set, focusing on the MAE and Pearson correlation coefficient as performance metrics. Based on this comparison, the model that achieved the best validation performance was selected and subsequently applied to the test set for final evaluation.

Feature selection output. Out of the full set of extracted features, the Boruta algorithm identified 40 variables as informative and retained them for further modeling, while 15 were rejected as irrelevant for the target prediction task. The selected features are summarized in Table 4.2. The reduced feature set improves model interpretability and computational efficiency while helping to mitigate the risk of overfitting.

Validation Results and Model Selection. To compare the performance of the different models trained on the selected features, we evaluated each on the validation set using MAE and Pearson correlation coefficient. Table 4.3 summarizes the results obtained for each model. Among the tested approaches, the GPR model achieved the lowest MAE of 9.966 mmHg, indicating the best prediction accuracy in terms of absolute error. Although the Pearson correlation coefficient for GPR, equal to 0.423, was not the highest, it still reflected a moderate linear relationship between the predicted and actual values. Interestingly, the Ensemble of Trees model showed a higher Pearson correlation coefficient (0.511), suggesting stronger linear agreement with the target, but it came with a higher MAE (11.026 mmHg), implying greater average deviation in predictions. This trade-off highlights that a higher correlation does not necessarily correspond to a more accurate model in terms of absolute error. Other models, such as linear regression (MAE = 11.883 mmHg, Pearson correlation coefficient = 0.212), support vector regression (MAE = 13.357 mmHg, Pearson correlation coefficient = 0.131), and

| Category | Selected Features |
|----------------------------------|---|
| Pulse Wave Analysis | Cardiac Period, Systolic Upstroke Time, Diastolic Time |
| Pulse Transit and Hemodynamic | Augmentation Index, Large Artery Stiffness Index, Areas under the PPG curve (S1,S3) |
| Amplitude-Based and Area-Derived | K-value, Ratio of amplitude PPG (maximum slope, diastolic notch, diastolic peak), Normalized areas PPG |
| Statistical Features | Time domain ECG (Mean value, Maximum value, Variance, Root mean square, Peak-to-peak amplitude, Kurtosis, Skewness) Time domain PPG (Mean value, Maximum value, Variance, Root mean square, Peak-to-peak amplitude, Crest factor, Kurtosis, Skewness) Frequency domain ECG (Mean frequency, Median frequency, Band power) Frequency domain PPG (Band power) Chaotic domain ECG (Approximate entropy, Shannon entropy, Sample entropy, Fuzzy entropy, Katz's fractal dimension) Chaotic domain PPG (Fuzzy entropy, Permutation entropy, Katz's fractal dimension) |

Table 4.2. Selected features grouped by category

regression tree (MAE = 12.432 mmHg, Pearson correlation coefficient = 0.443), demonstrated comparatively lower performance both in accuracy and correlation. Given the goal of minimizing prediction error, the GPR was selected as the most suitable approach and was subsequently employed for the final testing phase.

| Model | MAE (mmHg) | Pearson correlation coefficient |
|-----------------------------|------------|---------------------------------|
| Gaussian Process Regression | 9.966 | 0.423 |
| Ensembles of Trees | 11.026 | 0.511 |
| Linear Regression | 11.883 | 0.212 |
| Support Vector Regression | 13.357 | 0.131 |
| Regression Tree | 12.432 | 0.443 |

Table 4.3. MAE and Pearson correlation coefficient for each model evaluated on the validation set

Test Set Evaluation. Following the model selection process, the GPR model was evaluated on an independent test set to assess its generalization capability.

The model was trained using a squared exponential kernel function with automatic relevance determination. Hyperparameters were optimized by maximizing the marginal likelihood. On the test set, the GPR model achieved a MAE of 8.613 mmHg, with a SD of 8.315 mmHg. This result demonstrates a clear improvement over the validation performance (MAE = 9.966 mmHg), indicating that the model not only generalized well but also exhibited improved performance as a result of the richer information captured during training. The relatively low SD further suggests that the model maintains a stable performance across subjects, without significant degradation for particular individuals. This robustness is particularly important in physiological prediction tasks, where inter-subject variability can be considerable. Overall, these findings confirm the suitability of GPR for the task and reinforce the decision to select it as the final model. To further evaluate the generalization capability of the GPR model, we examined the distribution of MAE calculated across all test subjects. Specifically, a cumulative distribution function and a box plot, shown in Figure 4.2, were used to provide a more detailed view of the prediction variability across subjects.

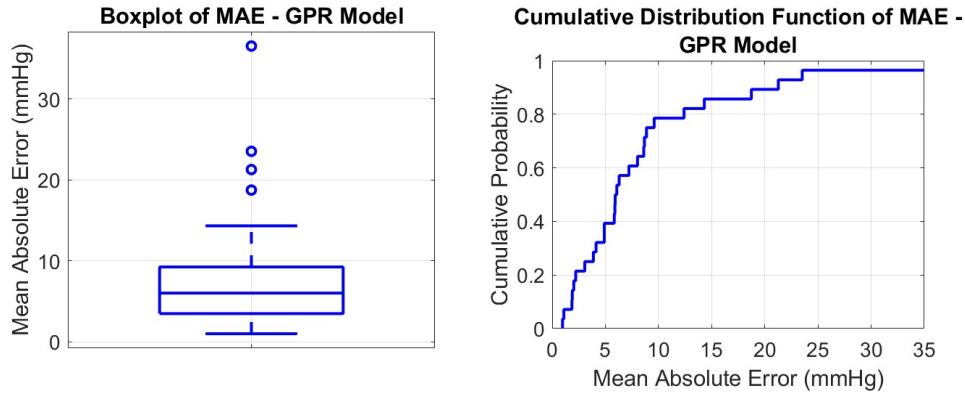


Figure 4.2. Box plot (on the left) and cumulative distribution function (on the right) of the MAE for the GPR model evaluated on the test set.

The box plot highlights a relatively narrow interquartile range, with the majority of subjects achieving errors below 10 mmHg, and only a limited number of outliers exceeding 20 mmHg. This indicates a consistent performance across individuals. In other words, the GPR model maintains a stable level of accuracy across different individuals, without large deviations in performance. This consistency is a desirable property, as it reduces the risk of subject-dependent degradation and enhances the model's reliability in diverse populations. The cumulative distribution function provides further insight into the distribution of MAE values. Approximately 50% of the subjects exhibit an MAE of 5 mmHg or less, around

30% fall in the 5–10 mmHg range, 10% lie between 10 and 15 mmHg, and fewer than 10% exceed 15 mmHg. This distribution highlights the model’s ability to deliver accurate predictions for the vast majority of subjects, with 80% remaining within a 10 mmHg error margin. The combined evidence from the box plot and the cumulative distribution function demonstrate that the GPR model achieves not only low average errors but also consistent performance across subjects, making it a reliable and robust solution.

4.3 Comparison with Existing Methods

Several studies have explored PTT-based BP estimation methods obtaining promising results. For example, Wong et al. conducted a study in which a linear model with only two coefficients was used to estimate BP in normotensive subjects. Remarkably, their approach yielded 0.0 ± 5.3 mmHg during the initial calibration phase. Even after a six-month interval without recalibration, the model maintained reasonable performances, with errors of 1.4 ± 10.2 mmHg [44]. Similarly, Choi et al. proposed a linear regression model based on PTT, applying it to data from the MIMIC database. Their method, which included advanced signal processing techniques such as the Hilbert-Huang Transform, achieved -0.44 ± 3.85 mmHg. Their study highlights that even with minimal model complexity, a two-parameter linear regression, accurate BP estimation is feasible, provided that a large number of subject-specific calibration points are available [45]. Despite these strong performances, a significant limitation shared by both models lies in their reliance on extensive individual calibration. Both studies required a substantial amount of reference BP data from each subject, which constrains their scalability and practicality in real-world applications.

In contrast, the approach proposed in the present study introduces a novel regression-based model designed to minimize the calibration need. Specifically, our method achieves a competitive estimation accuracy of 5.309 ± 4.561 mmHg, while requiring only five reference BP values per subject. This represents a significant reduction in calibration requirements compared to the large number of reference values needed in the previous studies. In summary, while earlier PTT-based models showed that simple linear methods can work well when trained with a lot of data from each person, our model focuses more on practical use. It reduces the amount of calibration needed, while still keeping good accuracy, making it more suitable for everyday and scalable cuffless BP monitoring.

Beyond regression-based approaches, a number of studies have explored ML-based approaches, reporting promising results under various experimental conditions, as described in Section 2.3. Miao et al. proposed a hybrid framework that combines a mechanism-driven model with data mining techniques, using features of

ECG and PPG signals selected through a genetic algorithm. Their model achieved very high accuracy in static conditions, with errors of -0.001 ± 3.102 mmHg [31]. Similarly, Chen et al. developed a ML pipeline based on PTT and morphological features, applying support vector regression optimized by a genetic algorithm. Their method yielded a MAE of 3.27 ± 5.52 mmHg [29]. Although these studies demonstrate promising performances in terms of accuracy, it is important to note that the reported results are derived from subject-dependent validation protocols. In such settings, data from the same individuals may be present in both the training and testing sets. This overlap enables the models to learn subject-specific patterns, which amplifies performance metrics. As a consequence, the evaluation does not faithfully represent the model’s ability to generalize to entirely unseen individuals, a critical requirement for clinical or wearable applications.

In contrast, our proposed model was developed and evaluated under a strictly subject-independent validation framework, ensuring a clear separation between training and testing individuals. This design choice prioritizes generalizability and robustness. While the overall MAE of 8.613 ± 8.315 mmHg may appear higher than that of subject-dependent models, our model maintained high consistency: 80% of test subjects showed MAEs below 10 mmHg, and the error distribution was tightly concentrated. These results highlight the importance of validation strategy. Although our model does not outperform all others in absolute terms, it offers a more realistic and reliable assessment of performance in unseen populations, an essential criterion for real-world applicability.

Chapter 5

Discussion

This thesis focused on developing and evaluating methods for non-invasive BP estimation based on physiological signals. Two distinct approaches were explored: regression-based models and a ML pipeline. The goal was to assess their performance under two realistic validation scenarios: a subject-specific calibration using a small number of reference BP points, and a subject-independent validation to avoid data overlap between sets. This dual approach provides a clear view of each method’s practical strengths and limitations.

Both approaches were developed and tested using a common dataset comprising ECG, PPG, and ABP signals from 100 patients who stayed in critical care units. To ensure consistency and reliability, all signals underwent a uniform pre-processing procedure, including tailored band-pass filtering for each signal type.

The regression-based analysis introduced three models of increasing complexity, each based on physiological relationships involving PTT and HR, and differing in the number of coefficients to estimate. Among them, Method 1 demonstrated notable performance when calibrated with only five reference BP values per subject. This minimal calibration strategy achieved a MAE of 5.309 ± 4.561 mmHg, which is not only competitive with more complex models but also exceeds them in stability and practicality. The novelty of this approach lies in its focus on reducing calibration effort, an aspect rarely addressed in existing literature. By showing that reliable predictions can be achieved with only five individualized data points, this work introduces a new perspective: balancing accuracy with usability, a critical requirement for scalable and accessible healthcare technologies. While more sophisticated models like Method 3 showed slightly better performance under large calibration datasets, they suffered significant degradation under minimal calibration conditions, highlighting their limited applicability in real-world low-data settings. In contrast, Method 1 consistently maintained narrow error margins across varying calibration sizes. This robustness was further confirmed through distribution analysis, showing that approximately 90% of subjects achieved an MAE

below 10 mmHg using only five reference BP values. To evaluate generalizability, two additional experiments were conducted with Method 1. The first trained the model on one subject and tested it on others, while the second used the dataset without individual calibration. Both approaches resulted in significantly degraded performance (MAEs above 23 mmHg), emphasizing that subject-specific calibration is essential. This demonstrates the effectiveness of the model when calibrated with minimal data for each individual, highlighting that coefficients estimated on one subject do not generalize well to others. These findings underscore the poor generalization of coefficients across individuals and support the superiority of personalized models over universal ones that ignore physiological variability.

The second part of the study proposed a ML pipeline incorporating feature extraction, feature selection, and model comparison. The ML analysis provided a complementary perspective, focusing on feature-based modeling using physiological signals. Following an extensive feature extraction process, the Boruta algorithm selected 40 relevant features encompassing time-domain, frequency-domain, morphological, and statistical-based characteristics from ECG and PPG signals. Among various models tested, GPR achieved the best performance in subject-independent validation and was therefore selected for the final evaluation on test set, where it reached a MAE of 8.613 ± 8.315 mmHg. Although this result does not surpass the performance of the regression model, it still demonstrates strong effectiveness, especially considering that the ML model was trained at the population level without subject-specific calibration. Furthermore, the GPR model demonstrated good consistency, with 80% of test subjects showing MAEs under 10 mmHg, and a relatively narrow interquartile error range. This highlights the model’s robustness and potential for applications where individual calibration is not possible.

This work contributes to the field of cuffless BP monitoring by proposing and evaluating both a minimal-calibration regression strategy and a subject-independent ML pipeline. The comparative analysis reveals a clear trade-off, showing that the regression-based method is both highly accurate and stable with minimal calibration and low computational cost, making it ideal for practical deployment in scenarios with limited data, such as sensor initialization or periodic updates. On the other hand, the ML pipeline exhibits superior generalization capabilities and enables fully automated operation without requiring personalized calibration. Although associated with marginally higher error rates, this makes it well-suited for situations where calibration is not feasible.

Despite the encouraging results, several limitations remain that suggest directions for future research. First, the dataset used in this study was derived exclusively from intensive care unit patients, who are typically in a resting or immobile state. For this reason, it remains uncertain how well the models will perform

when applied to real-world scenarios. Second, the study was conducted on a relatively limited sample of 100 subjects. While the results are promising, a larger and more diverse dataset would help confirm these findings. Lastly, this work focused on regression-based and classical ML models, which were chosen for their interpretability and computational efficiency. While deep learning approaches represent a promising direction, their exploration falls outside the scope of this study and remains a valuable area for future investigation.

Despite these limitations, the work provides a substantial contribution to the field by introducing a minimally calibrated regression strategy alongside a subject-independent ML pipeline. Together, these approaches lay a strong foundation for developing scalable and practical cuffless BP monitoring systems.

Chapter 6

Conclusion and future work

This work contributes to the field of cuffless BP estimation by presenting and validating both a novel minimal-calibration regression strategy and a subject-independent ML pipeline. The demonstration that accurate BP estimation can be achieved with only 5 reference values using a simple linear model represents a practical innovation with immediate applicability in real-world systems. Simultaneously, the ML approach confirms that data-driven models can provide consistent results across subjects even when validated with subject-independent strategy, offering an alternative path for non-invasive BP monitoring.

Future work should evaluate the robustness of both methods on data collected from patients in motion, ideally using wearable sensors, to assess their performance under daily-life conditions, where signal quality and noise present additional challenges. Moreover, expanding the dataset to include a more diverse and larger population would further enhance the generalizability of the models.

In conclusion, this thesis demonstrates that it is possible to build scalable and practical cuffless BP monitoring systems by combining personalized calibration strategies with subject-independent learning models, paving the way for more accessible, non-invasive healthcare technologies.

Bibliography

- [1] Malcolm Elliott and Alysia Coventry. Critical care: The eight vital signs of patient monitoring. *British Journal of Nursing*, 21:621–625, 5 2012.
- [2] What is Blood Pressure? Vital Sign Measurement Across the Lifespan; 1st Canadian edition. <https://pressbooks.library.torontomu.ca/vitalsign/chapter/what-is-blood-pressure/>.
- [3] High Blood Pressure; heart.org. <https://www.heart.org/en/health-topics/high-blood-pressure>.
- [4] Hypertension; who.int. <https://www.who.int/news-room/fact-sheets/detail/hypertension>.
- [5] Mahtab Mehrabbeik and Saeid Rashidi. Estimation of cuffless systolic and diastolic blood pressure using pulse transient time. *5th Iranian Conference on Signal Processing and Intelligent Systems, ICSPIS 2019*, 12 2019.
- [6] Mohamad Kachuee, Mohammad Mahdi Kiani, Hoda Mohammadzade, and Mahdi Shabany. Cuff-less high-accuracy calibration-free blood pressure estimation using pulse transit time. *Proceedings - IEEE International Symposium on Circuits and Systems*, 2015-July:1006–1009, 7 2015.
- [7] Ramakrishna Mukkamala, George S. Stergiou, and Alberto P. Avolio. Cuffless blood pressure measurement. *Annual Review of Biomedical Engineering*, 24:203–230, 2022.
- [8] Agnes S. Meidert and Bernd Saugel. Techniques for non-invasive monitoring of arterial blood pressure. *Frontiers in Medicine*, 4:322154, 1 2017.
- [9] Heather H. Hager and Bracken Burns. Artery cannulation(archived). *StatPearls*, 7 2023.
- [10] Ahmadreza Argha, Branko G. Celler, and Nigel H. Lovell. Artificial intelligence based blood pressure estimation from auscultatory and oscillometric

- waveforms: A methodological review. *IEEE Reviews in Biomedical Engineering*, 15:152–168, 2022.
- [11] Matthew R. Nelson, Jan Stepanek, Michael Cevette, Michael Covalciuc, R. Todd Hurst, and A. Jamil Tajik. Noninvasive measurement of central vascular pressures with arterial tonometry: Clinical revival of the pulse pressure waveform? *Mayo Clinic proceedings*, 85:460–472, 2010.
- [12] J. R.. Levick. *An introduction to cardiovascular physiology*. Butterworths, 1991.
- [13] Texas Instruments. Patient monitoring 101: Part-2 understanding ecg basics and lead derivation prepared by: Ryan andrews. <https://www.ti.com>, 2021.
- [14] Electrocardiography Wikipedia en.wikipedia.org. <https://en.wikipedia.org/wiki/Electrocardiography>.
- [15] John Allen. Photoplethysmography and its application in clinical physiological measurement, 3 2007.
- [16] N. De Pinho Ferreira, C. Gehin, and B. Massot. A review of methods for non-invasive heart rate measurement on wrist, 2 2021.
- [17] Mohd Zubir Suboh, Rosmina Jaafar, Nazrul Anuar Nayan, Noor Hasmiza Harun, and Mohd Shawal Faizal Mohamad. Analysis on four derivative waveforms of photoplethysmogram (ppg) for fiducial point detection. *Frontiers in Public Health*, 10, 6 2022.
- [18] Heather Ting Ma. A blood pressure monitoring method for stroke management. *BioMed Research International*, 2014, 2014.
- [19] Piyawat Samartkit and Saroj Pullteap. Non-invasive continuous blood pressure sensors in biomedical engineering research: A review, 3 2024.
- [20] T. Ma and Y. T. Zhang. A correlation study on the variabilities in pulse transit time, blood pressure, and heart rate recorded simultaneously from healthy subjects. *Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings*, 7 VOLS:996–999, 2005.
- [21] Milton Landowne and Gerontology Branch. A method using induced waves to study pressure propagation in human arteries. *Circulation research*, 5:594–601, 1957.
- [22] Manuja Sharma, Karinne Barbosa, Victor Ho, Devon Griggs, Tadesse Ghirmai, Sandeep K. Krishnan, Tzung K. Hsiai, Jung Chih Chiao, and Hung Cao.

- Cuff-less and continuous blood pressure monitoring: A methodological review. *Technologies 2017*, Vol. 5, Page 21, 5:21, 5 2017.
- [23] Xiao Rong Ding, Yuan Ting Zhang, Jing Liu, Wen Xuan Dai, and Hon Ki Tsang. Continuous cuffless blood pressure estimation using pulse transit time and photoplethysmogram intensity ratio. *IEEE Transactions on Biomedical Engineering*, 63:964–972, 5 2016.
- [24] Syed Ghufuran Khalid, Jufen Zhang, Fei Chen, and Dingchang Zheng. Blood pressure estimation using photoplethysmography only: Comparison between different machine learning approaches. *Journal of Healthcare Engineering*, 2018, 2018.
- [25] Seyedeh Somayyeh Mousavi, Mohammad Firouzmand, Mostafa Charmi, Mohammad Hemmati, Maryam Moghadam, and Yadollah Ghorbani. Blood pressure estimation from appropriate and inappropriate ppg signals using a whole-based method. *Biomedical Signal Processing and Control*, 47:196–206, 1 2019.
- [26] Chen Yi, Cheng Jian, and Ji Wenqiang. Continuous blood pressure measurement based on photoplethysmography. *2019 14th IEEE International Conference on Electronic Measurement and Instruments, ICEMI 2019*, pages 1656–1663, 11 2019.
- [27] Kefeng Duan, Zhiliang Qian, Mohamed Atef, and Guoxing Wang. A feature exploration methodology for learning based cuffless blood pressure measurement using photoplethysmography. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 2016-October:6385–6388, 10 2016.
- [28] Sergio González, Wan Ting Hsieh, and Trista Pei Chun Chen. A benchmark for machine-learning based non-invasive blood pressure estimation using photoplethysmogram. *Scientific Data*, 10, 12 2023.
- [29] Shuo Chen, Zhong Ji, Haiyan Wu, and Yingchao Xu. A non-invasive continuous blood pressure estimation approach based on machine learning. *Sensors (Switzerland)*, 19, 6 2019.
- [30] Rui He, Zhi Pei Huang, Lian Ying Ji, Jian Kang Wu, Huihui Li, and Zhi Qiang Zhang. Beat-to-beat ambulatory blood pressure estimation based on random forest. In *BSN 2016 - 13th Annual Body Sensor Networks Conference*, pages 194–198. Institute of Electrical and Electronics Engineers Inc., 7 2016.
- [31] Fen Miao, Nan Fu, Yuan Ting Zhang, Xiao Rong Ding, Xi Hong, Qingyun He, and Ye Li. A novel continuous blood pressure estimation approach based on

- data mining techniques. *IEEE Journal of Biomedical and Health Informatics*, 21:1730–1740, 11 2017.
- [32] Mohammad Kachuee, Mohammad Mahdi Kiani, Hoda Mohammadzade, and Mahdi Shabany. Cuffless blood pressure estimation algorithms for continuous health-care monitoring. *IEEE Transactions on Biomedical Engineering*, 64:859–869, 4 2017.
- [33] UCI Machine Learning Repository archive.ics.uci.edu. <https://archive.ics.uci.edu/dataset/340/cuff+less+blood+pressure+estimation>.
- [34] Jiapu Pan and Willis J. Tompkins. A real-time qrs detection algorithm. *IEEE Transactions on Biomedical Engineering*, BME-32:230–236, 1985.
- [35] Hooman Sedghamiz. Matlab implementation of pan tompkins ecg qrs detector. Technical report, 2014.
- [36] Steven M. Bishop and Ari Ercole. Multi-scale peak and trough detection optimised for periodic and quasi-periodic neuroscience data. *Acta Neurochirurgica, Supplementum*, 126:189–195, 2018.
- [37] Peter H. Charlton, Kevin Kotzen, Elisa Mejía-Mejía, Philip J. Aston, Karthik Budidha, Jonathan Mant, Callum Pettit, Joachim A. Behar, and Panicos A. Kyriacou. Detecting beats in the photoplethysmogram: Benchmarking open-source algorithms. *Physiological Measurement*, 43, 8 2022.
- [38] A. C. Guyton. The relationship of cardiac output and arterial pressure control. *Circulation*, 64:1079–1088, 1981.
- [39] Least squares - Wikipedia — en.wikipedia.org. https://en.wikipedia.org/wiki/Least_squares.
- [40] Yuriy Kurylyak, Francesco Lamonaca, and Domenico Grimaldi. A neural network-based method for continuous blood pressure estimation from a ppg signal. *Conference Record - IEEE Instrumentation and Measurement Technology Conference*, pages 280–283, 2013.
- [41] Daisuke Fujita, Arata Suzuki, and Kazuteru Ryu. Ppg-based systolic blood pressure estimation method using pls and level-crossing feature. *Applied Sciences 2019, Vol. 9, Page 304*, 9:304, 1 2019.
- [42] Umit Senturk, Kemal Polat, and Ibrahim Yucedag. A non-invasive continuous cuffless blood pressure estimation using dynamic recurrent neural networks. *Applied Acoustics*, 170:107534, 12 2020.

- [43] Sumbal Maqsood, Shuxiang Xu, Son Tran, Saurabh Garg, Matthew Springer, Mohan Karunanithi, and Rami Mohawesh. A survey: From shallow to deep machine learning approaches for blood pressure estimation using biosensors, 7 2022.
- [44] Mico Yee Man Wong, Carmen Chung Yan Poon, and Yuan Ting Zhang. An evaluation of the cuffless blood pressure estimation based on pulse transit time technique: A half year study on normotensive subjects. *Cardiovascular Engineering*, 9:32–38, 2009.
- [45] Younhee Choi, Qiao Zhang, and Seokbum Ko. Noninvasive cuffless blood pressure estimation using pulse transit time and hilbert-huang transform. *Computers and Electrical Engineering*, 39:103–111, 1 2013.