DATA-DRIVEN METHODOLOGIES FOR CUFF-LESS BLOOD PRESSURE ESTIMATION

by

WEINAN WANG

A thesis submitted to the

School of Graduate Studies

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Master of Science

Graduate Program in Electrical and Computer Engineering

Written under the direction of

Laleh Najafizadeh

and approved by

_____

_____

_____

New Brunswick, New Jersey

January, 2021

ABSTRACT OF THE THESIS

Data-Driven Methodologies for Cuff-Less Blood Pressure Estimation

By Weinan Wang

Thesis Director:

Laleh Najafizadeh

Cuff-less blood pressure (BP) estimation methods are the highly-desired replacement for conventional cuff-based methods, as they enable long-term and continuous monitoring of the BP with limited disturbance or need for manual operation. Cuff-less BP estimation methods can be generally classified into model-driven and data-driven methods. Both methods often utilize photoplethysmogram (PPG) or electrocardiogram (ECG), which can be continuously and non-invasively acquired.

Model-driven methods are based on the pulse wave transition theory, which relates the pulse transit time (PTT) to the blood pressure. An advantage that model-driven methods offer is that very few parameters have to be learned from the training set, thereby, making them efficient and computationally inexpensive. However, these methods require individual calibration, their accuracy decreases over time, and they generally require recordings of both ECG and PPG, which is hard to maintain over long periods. On the other hand, in data-driven methods, the need for subject-specific calibration and the requirement of using two or more physiological recordings are released. However, these methods typically require high computational budget and massive training datasets for learning a much larger quantity of parameters.

In this thesis, we first provide a comprehensive review of seven models that have been

utilized in model-driven BP estimation methods, and discuss their advantages and limitations. We then present an overview of existing data-driven methods that have used ECG alone, PPG alone, or both signals for BP estimation.

Motivated by the notable performance of PPG-based data-driven BP estimation methods, we then present a novel transfer learning-based blood pressure estimation algorithm that utilizes visibility graph to form images from PPG recordings. The proposed method provides accurate BP estimation with only one PPG beat, while being computationally efficient requiring training of only one dense layer. Experimental results demonstrate that the proposed method offers comparable or better BP estimation accuracy compared to other data-driven methods with higher computational complexity, making it a suitable candidate for continuous BP monitoring applications.

Acknowledgements

I want to sincerely thank my advisor and professors who have been endeavoring to offer guidance and support along my graduate researches.

I would like to acknowledge my advisor, Dr. Laleh Najafizadeh, for guiding me through this research topic. She has been trusting and motivating me to follow my own ideas and interests, while kindly and timely offering directive suggestions when I am confused. It is impossible to form this comprehensive thesis without her guidance.

I would also like to acknowledge my M.Sc. thesis committee members, Dr. Yingying Chen and Dr. Pedram Mohseni, for the precious time they have taken from their busy schedule to examine and comment on this thesis. Their feedbacks and suggestions further perfected this research.

I would like to thank Dr. Sophocles Orfanidis and Dr. Waheed U. Bajwa for their instructions on digital signal processing and machine learning, which are the most essential components of this research. Dr. Orfanidis's thorough instructions as well as his textbook, *Introduction to Signal Processing*, have been very helpful references for me to understand the pre-processing and spectral analysis procedures of physiological signal. Dr. Bajwa's profound lectures, on the other hand, have deepen my comprehension of machine learning methodologies.

Also, a special acknowledgement to Dr. Shantenu Jha, Dr. John J. McGarvey in the Electrical and Computer Engineering department, and Prof. Nela Navarro from the School of Arts and Sciences, whose dedicated instructions have always been very important components of my domain knowledge.

Finally, I would like to express my appreciation to the unconditioned support and care from my beloved family.

This work was supported by Craig H. Neilsen Foundation.

Dedication

A very special dedication belongs to my college friends who have maintained unvaried connections with me after my departure from China.

Table of Contents

Chapter 1

Introduction

1.1 Background

As one of the leading causes of lethal diseases across the world, hypertension has long remained a wide-spread public health challenge that leads to cardiomyopathy, arterial fabrication, and other cardiovascular diseases accounting for more than 15% of health problems among mid-aged people, and 20% for those who are older [1]. What makes it exclusively detrimental besides its direct damage to the cardiovascular system is that it boosts other chronic diseases, such as brain and kidney impairment causing strokes [2] and diabetes [3].

Hypertension is noticeably pervasive yet has not been effectively controlled in the United States. A recent study [4] reported that the prevalence of hypertension among US adults is 29.0%, with a trend that increases with age and reaches as high as 63.1% among elders aged over 60. Even under relaxed standards of controlling the systolic and diastolic blood pressure (SBP, DBP) under the level of stage 2 hypertension, only 48.3% of the patients have their blood pressure (BP) under control, which is below the goal of 61.2% set by the U.S. Department of Health and Human Services for 2020 [5].

Generous number of home-measured blood pressures are informative and helpful for treating hypertension. Free from the bias caused by observers at the scene, the well-known Ohasama follow-up study on local residents [6] concluded that home-measured blood pressure has better reprehensibility than those readings measured for screening at the clinic, being more significantly related to patient's risk of cardiovascular mortality even when the number of measurements are the same. The study also suggested an increment in health status predictability with the growth of number of home-measured blood pressures, as averaging among vast records eliminates the bias and yields a more representative value. Additional studies conducted on

comprehensive medical records from U.S., Canada and European countries also supported similar conclusions [7]. Besides the rich information that these readings provide for therapy, they also enhance patients' long-term adherence and compliance to clinical advice [8], which is also significant for keeping the blood pressure under control. Therefore, organizations such as the European Society of Hypertension recommends taking an adequate number of blood pressure readings of 2 daily measurements, 2 readings for each measurement, for appropriate diagnosis or treatment [9].

However, patients' actual adherence to the guidelines is not optimistic. Recent study shows that among patients with known or suspected hypertension, who have a BP monitoring device at home, merely 29% of them measured their blood pressure more frequently than twice a day [10], let alone problems of not strictly following the guidelines such as properly waiting between readings or taking multiple readings in single measurement.

The abovementioned issues originate from the intrinsic inconvenience of conventional blood pressure monitoring devices such as sphygmomanometer [11] and oscillometry-based monitors [12]. For sphygmomanometers, SBP and DBP are measured by recording the instant pressure reading on the manometer when the first and fifth Kortokoff sound are detected [11]. For the oscillometry monitor, the mercury manometer is replaced by an electronic pressure sensor. The variation in the amplitude of obtained pressure signal is analyzed to estimate the mean arterial pressure (MAP), and SBP and DBP are calculated consequently [13]. For both types of devices, an inflatable cuff is required to occlude the main artery of the user's upper arm, which fully occupies user's time and attention and makes multiple daily measurements an unpleasant and cumbersome experience, as the user feels obvious bondage on the arm. It is thus impossible to acquire blood pressure readings continuously and automatically without distracting the user with these cuff-based devices. Figure 1.1 demonstrates an example of such a device.

Figure 1.1: A user taking blood pressure readings from a commonly used cuff-based oscillometry monitor [14]. User has to manually initiate the measurement by wearing the cuff on the upper arm and pressing the button on the device. After the measurement has started, the device inflates the cuff and its pressure to the whole upper arm fully occupies user's attention in the following $40 \sim 50$ seconds until one pair of SBP and DBP reading is obtained.

1.2 Existing Works

To address the limitations of cuff-based methods, several cuff-less approaches have been proposed over the past decades, which can be categorized as model-driven and data-driven methods.

Model-driven BP estimation techniques aim at deriving BP from its physiological relationship with pulse wave velocity (PWV). By manually analyzing and properly simplifying the physiological processes that relates PWV to BP, parametric models can be formed and calibrated via solving a simple regression problem with manually-specified model parameters. PWV can be measured from a pair of synced physiological signals, such as electrocardiogram (ECG) and photoplethysmogram (PPG). These signals can be automatically and continuously recorded from sensors and electrodes attached to human body without disturbing user's daily activity. Figure 1.2 demonstrates a device utilizing this method to estimate blood pressure.

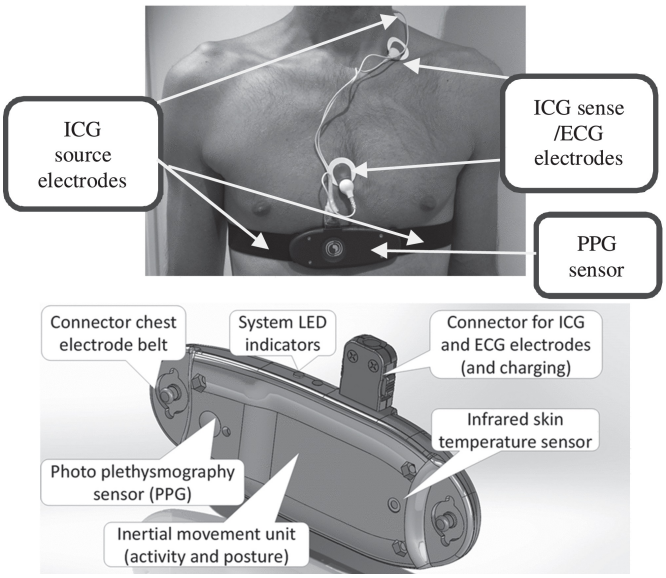

Figure 1.2: A cuff-less blood pressure monitor [15]. Blood pressure values are indirectly estimated from ECG, PPG and ICG signals, and no cuff is required. These signals can be automatically and continuously recorded, so blood pressure readings can be measured and monitored without any manual intervention or any disturbance to user's daliy activity.

Model-driven techniques have their advantages of simplicity and interpretability. Parameters in the model are manually specified to represent some physiological parameters or processes, which can be calibrated with a small training set. However, measuring PWV requires at least 2 synced signals as inputs, which could not be feasible over a long period of time. Also, the simplified model parameters often lacks generalizability, which have to be frequently calibrated individually for each subject.

Apart from the pressure-velocity relationship requiring both ECG and PPG as inputs, there is also a pressure-volume relationship that correlates PPG signal with BP, which potentially yields an algorithm requiring only PPG as the input. The pressure-volume relationship is highly dynamic and inconsistent to be manually modeled with a few parameters, which is thus studied with statistical and data-driven methods such as feature selection and machine learning (ML). Many morphological features with good BP predictability can be extracted from the PPG signal, whose relationships with BP are formed as a multiple regression problem, and is solved by ML algorithms without manually specifying the parameters in the model.

Data-driven methods addressed some drawbacks of the model driven methods. By including more features and more parameters in the model, the model out of ML methods has good generalizability to be applied on a group of subjects without requiring individual calibration. These methods also release the rigid requirement of having 2 signals, making them easier to implement. However, these models lack interpretability and have to be trained on massive datasets with high computational budget.

In the past two decades, the literature has proposed lots of studies using both methods. Most of these works share a similar processing pipeline from acquiring the datasets to the assessment of final performance of model, which is summarized below and in Figure 1.3.

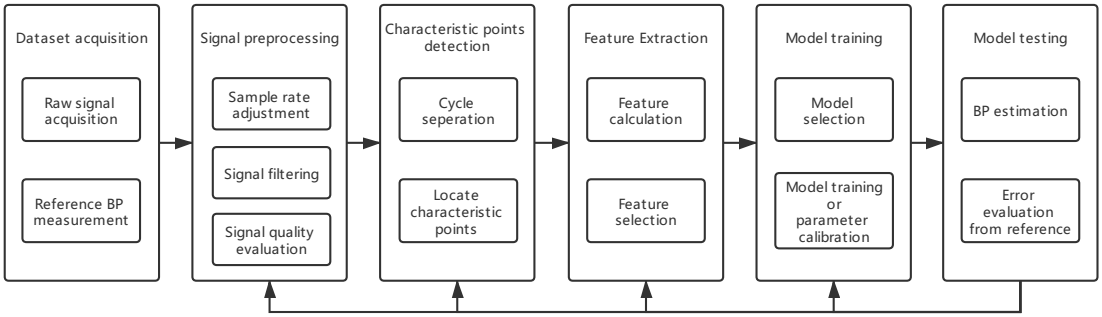

Figure 1.3: A typical flow chart for cuff-less BP estimation studies.

1. **Dataset acquisition**: A dataset including digital recordings of raw physiological signals and reference BPs is collected from customized experiments or taken from online datasets, such as the MIMIC waveform database [16]. For model-driven methods, recordings of ECG and PPG are required to be synced for accurately extracting PWV indicators. Usually, the dataset includes recordings for multiple subjects to show model generalizability and meet main stream requirements for blood pressure monitoring devices, such as the the AAMI SP-10 protocol from the American Association for the Advancement of Medical Instrumentation (AAMI), which requires having more than 15 subjects [12].

2. **Signal preprocessing**: Raw physiological signals from sensors or open datasets are prone to be affected by multiple sources of interference, such as motion artifacts, baseline wandering, and powerline noises. These interferences introduce additional ripples or spikes to signal morphology and obstruct the detection of characteristic points and proper extraction of features. Therefore, the signals are often processed via band pass

digital filters or wavelet decomposition to rejects the effect of interference. For random or time-variant interference being hard to remove from recorded signal, the signal quality within a segment in the recording is often evaluated in order to exclude segments with poor quality for later analysis.

3. **Characteristic point extraction**: PWV and PPG features are often defined on the temporal position or amplitude of characteristic points in each cycle, so the first step before feature extraction is to separate the signal into individual cycles and locate positions of characteristic points in each cycle. In most cases, the characteristic points are related to peaks of the original signal or its derivatives, so a robust peak detection algorithm is often essential.

4. **Feature extraction**: Features are often calculated from the relative position of characteristic points, such as the time delay between them or the ratio of their amplitudes. If a large feature pool is considered, a feature selection algorithm is often utilized to exclude redundant features having low correlation with BP or colinearity with other features to optimize the subsequent training process.

5. **Model training**: In model-driven cases, a parametric model is manually designed to depict the relationship between the PWV indicator and the BP. These manually defined parameters in the model are related to physiological properties of the human artery, which are subject specific. Being hard or impossible to measure directly, these parameters are determined using simple regression methods such as least square regression with a training set including part of the extracted indicators and their corresponding blood pressure values. In data-driven cases, instead of a parametric model, a ML model such as support vector machine, regression tree or neuron network is selected, which is expected to spontaneously forms the relationship between the features and the BP from the training set.

6. **Model testing**: After all parameters in the model have been determined, the algorithm can be used to estimate BP values from new inputs in the testing set. Estimated BP values are compared to the corresponding reference BP values to evaluate the performance of

obtained model. Regression metrics such as Pearson's correlation coefficients, mean error, mean absolute error and root mean square error are often calculated for performance comparison among multiple models if same or comparable dataset is used.

Existing works have made major improvements for both approaches. For the model-driven approach, studies have sought for new forms of parametric models that better describe the PWV-BP relationship, while for the data-driven approach, studies have focused on extracting new features and their combination with proper ML model to improve the regression accuracy. We describes these advances below.

- **Parametric models in model-driven studies**: Great efforts have been made by researchers to derive new PWV-BP models with better regression accuracy. Examples include using various linear or non-linear equations to describe the relationship between PWV and BP as well as adding more parameters to consider additional physical factors or effects that alter the relationship. Chapter 2 of this thesis will review derivation and performance of the following 7 commonly-used models:

$$\text{Linear Model: } BP = a_1 \times PAT + a_2, \tag{1.1}$$

$$\text{Inverse Model: } BP = a_1 \times \frac{1}{PAT} + a_2, \tag{1.2}$$

$$\text{Logarithmic Model: } BP = a_1 \times \ln(PAT) + a_2, \tag{1.3}$$

$$\text{Inverse Square Model: } BP = a_1 \times \frac{1}{PAT^2} + a_2, \tag{1.4}$$

$$\begin{aligned} \text{MAP Model: } MAP &= a_1 \times \ln(PAT) + a_2, \\ PP &= a_3 \times \frac{1}{PAT^2}, \end{aligned} \tag{1.5}$$

$$\text{Trimmed Inverse Square Model: } BP = a_1 \times \frac{1}{(PAT - a_2)^2} + a_3, \tag{1.6}$$

$$\text{Square Root Model: } BP = \sqrt{a_1 \times \frac{1}{PAT^2} + a_2} + a_3, \tag{1.7}$$

  where a_1, a_2 and a_3 are the model parameters to be determined from the training set.

- **Exploration of features in data-driven studies**: Data-driven studies have developed along the expansion of the feature pool, especially features extracted from PPG. Explicitly-defined PPG features calculated from the duration, amplitude, area under waveform, and

statistical indices of the signal have been added to the feature pool, with studies proving their BP predictability. Later, frequency domain features of the PPG were used as semi-defined features, and some of the recent works take a whole-based feature approach, in which samples of the PPG signal are given to the ML model without extracting any predefined feature. Data-driven approaches are reviewed in Chapter 3 of this thesis.

1.3 Motivation

Motivated by the promising accuracy of recent data-driven BP estimation algorithms, we endeavor to develop a new PPG-only method addressing the limitation of existed methods. Utilizing transfer learning, our proposed method is free of individual calibration while having much less parameters to be trained compared to existed works with similar performance, lowering the computational budget of data-driven method. The proposed approach will be presented in Chapter 4.

1.4 Thesis Layout

The rest of this thesis is organized as follows. Chapters 2 and 3 present comprehensive literature review of key methodologies in existing model-driven and data-driven studies. In Chapter 4, we present our study proposing a new data-driven BP estimation method using only PPG as input, which utilizes transfer learning by converting the time series of PPG signal to visibility graph. Finally, the thesis is concluded in Chapter 5.

Chapter 2

Model Driven Methodologies for Cuff-less Blood Pressure Estimation

2.1 Introduction

For decades, researchers have been looking for alternative ways other than cuff-based methods, such as sphygmomanometers or oscillometry devices, to measure the blood pressure (BP) noninvasively, not only for eliminating the inconvenience and discomfort caused by the pressure and the large size of the inflatable cuff, but also to enable the possibility of continuous BP monitoring that offers essential information for the diagnosis of cardiovascular diseases and personalized healthcare.

Among the studies, model-driven BP estimation techniques derive BP indirectly from pulse wave velocity (PWV)-related indicators. The velocity of wave propagation is related to the density and the bulk modulus of the media, while these stiffness parameters of human artery is related to BP. Therefore, the relationship between PWV and BP can be described using manually derived mathematical models.

Being inversely related to PWV, the pulse transit time (PTT) is an indicator of PWV defined as the time for the pulse wave to travel over a certain distance, which can be obtained from two synced physiological recordings taken from proximal and distal spots. However, the pulse arrival time (PAT), which is a biased measure of PTT, is often utilized as the PWV indicator due to the wide utilization of electrocardiogram (ECG) and photoplethysmogram (PPG) signals.

Previous works have been endeavoring on proposing new models with better BP prediction accuracy. Some models are derived from fundamental physical relationships between PWV and BP, while others are trimmed from existing models by adding extra parameters to compensate potential measurement bias, considering additional factors that alter arterial physiology, or

simplifying the model by neglecting some parameters.

In this chapter, we present an overview of some existing models that are used to estimate the blood pressure. After explaining some of the basic terminologies and background theories, this chapter reviews and summarizes these models in terms of their derivation, and makes a comparative discussion of their application and performance.

2.2 Background

2.2.1 Pulse Wave Velocity (PWV) and Pulse Transit Time (PTT)

Pulse wave is repeatedly generated by heart contraction as the heart beats. In each cardiac cycle when the blood is ejected from the heart chamber to the aorta, an acute increase of blood volume dilates the artery's wall, and creates a pressure wave in the vessel [17], which is referred to as the pulse wave. The pulse wave velocity (PWV) is therefore defined as the speed that this pressure wave created by the heart systole propagates in the vessel in each cardiac cycle.

The mechanism of pulse wave generation through heartbeats determines its transmission direction, originating from the ventricle and transitioning along the arterial tree, through the main aorta to the peripheral arteries. Consequently, the PWV is often measured as the average wave speed over a known length of artery by measuring the pulse transit time (PTT), which is defined as the time taken for the pulse wave to transit from a proximal spot to a distal spot [18]. Although PWV and PTT are related by the transit distance L by

$$PWV = \frac{L}{PTT}, \tag{2.1}$$

we don't have to actually measure the distance since L can be obtained from regression.

2.2.2 Electrocardiogram (ECG) and Photoplethysmogram (PPG)

Two synced physiological signal recordings are required to measure PTT from the time difference between their beat-to-beat characteristic points, indicating the transmission of the pulse wave. In order to accommodate the demand of measuring the blood pressure in a continuous and disturbance-free way, the signal must be available through a non-invasive method without requiring for manual interventions. ECG and PPG are signals that are most frequently used for

estimating BP from PWV due to their easy, safe and standardized sensor placement and low hardware budget.

- **ECG**: ECG is a voltage waveform evaluated from multiple electrodes attached to the body, which measures the electrical activity of the heart. It represents the change of electric potential of cardiac muscles as the heart intakes and ejects blood at each heartbeat cycle [19]. A simple ECG setup with 3 electrodes is enough to capture the procedure of ventricle depolarization, in which the muscles in the heart chamber, under the control of electrical signal generated by the pacemaker cells [20], contracts and pushes the blood from the heart chamber to the main artery, generating the pulse wave. The procedure appears as the QRS complex on the ECG waveform, which is morphologically the major spike of the ECG signal in a given cycle.

  The characteristic point of ECG indicating the generation of pulse wave is selected to be the peak of the QRS complex in each cycle, which is known as the R-peak. Figure 2.1 depicts the shape of an ECG waveform, the location of the QRS complex, and the position of the R-peak. The ECG R-peak is often used as the proximal indicator of the pulse wave.

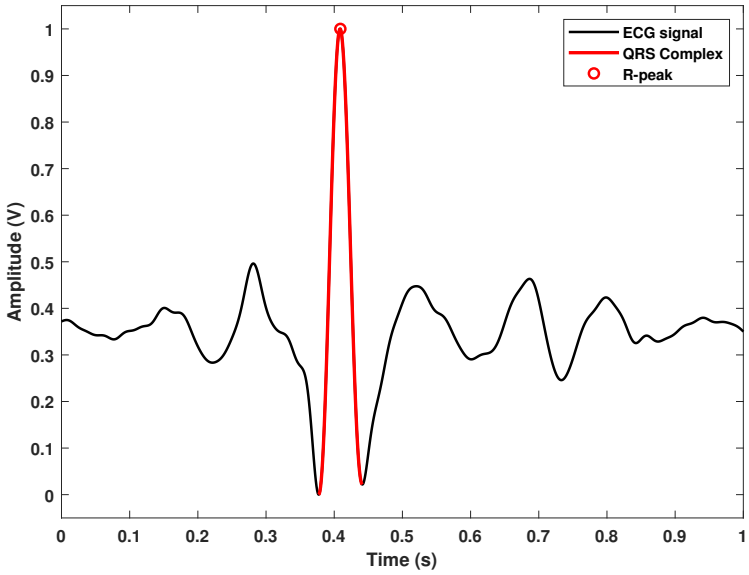

Figure 2.1: Plot of the ECG waveform, its QRS complex and its R-peak in a cycle.

- **PPG**: PPG is another voltage waveform measured from optical sensors placed on the

peripheral body tissues such as fingertip, toe or earlobe. The sensor records changes in the blood volume of peripheral artery in each cardiac cycle using a light emitter and a photodector placed at the same (reflection mode) or opposite (transmission mode) sides of the tissue. The arterial pulsation at the distal spots, caused by the transmission of pulse wave, changes the peripheral blood volume in a quasiperiodic way, which alters the path length of the light transmission between the emitter and the detector, changing the light energy absorbed across the path and received by the photodector [21]. Consequently, the time instance of the pulse wave arrival can be located in the waveform.

In each cycle, the PPG waveform consists of a rising systolic phase and a falling diastolic phase. Unlike ECG with its distinct R-peak, various characteristics can be extracted from the PPG waveform [22, 23, 24]. Some of the mostly-used characteristics are summarized below:

- **Foot**: End of the diastolic phase of the previous cycle and start of the systolic phase of the current cycle.
- **Maximum slope**: The time instance at which the maximum dPPG amplitude in the systolic phase of the current cycle is obtained, where dPPG represents the derivative of the PPG signal.
- **Peak**: The time instance at which the maximum PPG amplitude is obtained in each cycle.

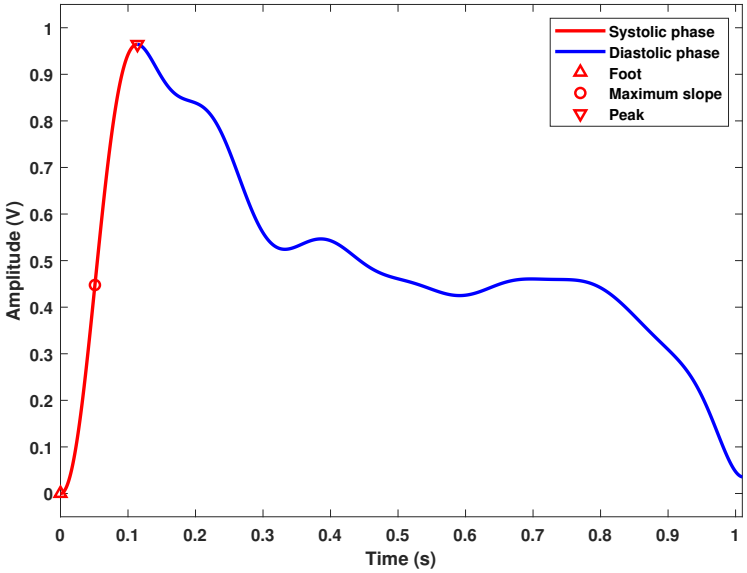

Figure 2.2: Plot of a PPG waveform, its systolic and diastolic phases, and position of various characteristic points in a cycle.

2.2.3 Pulse Arrival Time (PAT) and Pulse Transit Time (PTT)

Although many existing works do not distinguish between the pulse transit time (PTT) and the pulse arrival time (PAT), the concepts are different. PTT can be measured from various signal sources as long as the delay between their characteristic points reflects the time taken for the pulse wave to travel in the vessel. However, PAT is defined as the time interval between the occurrence of the ECG R-peak and the PPG characteristic point in each cardiac cycle. Sometimes depending on the definition of the used PPG characteristic point, the derived PAT gets an additional notation. For example, 'PAT_f' is used when PPG foot is used for the calculation of PAT, or 'PAT_d' is used when the maximum slope of PPG is used for PAT calculation. As an example, Figure 2.3 demonstrates the extraction of PAT_d from the ECG and PPG signals.

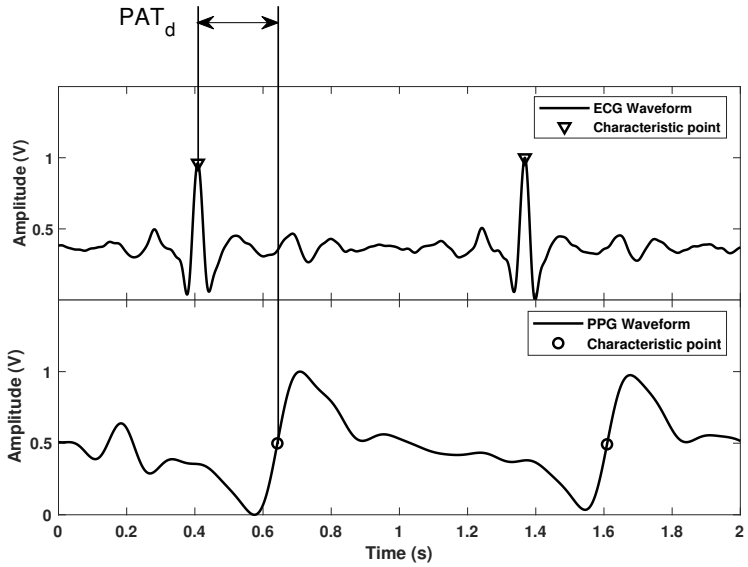

Figure 2.3: Plot of extracting PAT_d from the time difference of the characteristic points of ECG and PPG.

The subtle difference between PAT and PTT does not end at their definitions. An electrical-mechanical delay exists between ventricle depolarization, recorded as R-peak in the ECG waveform, and the actual contraction of the muscles in the heart chamber, which ejects the blood to the artery and generates the pulse wave [25]. This delay is referred to as the pre-ejection period (PEP). This additional delay will be included when using the ECG R-peak as proximal characteristic point to extract PAT, with their relationship given as

$$PAT = PEP + PTT. \tag{2.2}$$

Unlike PTT, which is physically related to PWV by (2.1), there is no direct theoretical relationship between PEP and PWV.

While PAT has been widely utilized due to convenience and wide signal availability in large datasets, some previous works have endeavored to exclude PEP and measure accurate PTT values by using additional signal sources. In the literature, two major methods are utilized to address the problem:

1. **Replace proximal signal**: Instead of using ECG reflecting electrical activities of heart, signals indicating volume changes of central artery after generation of pulse wave excludes the electrical-mechanical delay. Impedance cardiogram (ICG), for example, is an

ideal alternative. It measures the variation of thorax impedance, which is caused by two heartbeat-related factors: the blood volume changed by ventricle blood ejection, and the orientation of erythrocytes changed by blood flow [26].

2. **Use additional distal signal**: For example, we can measure PTT from the delay between characteristic points of PPG signals taken from the fingertip and the earlobe, as is shown in Figure 2.4. The theory behind this method is explained below.

   In the first method using one proximal and one distal signals, the waveforms are utilized as indicators to calculate the time taken for the pulse wave to travel along single artery. For example, as is shown in Figure 2.5, if we use ICG and fingertip PPG to measure PTT, then we are estimating the PWV in the axillary artery. However, in the second method where we use two distal signals, two different paths are involved. Taking our example depicted in Figure 2.4, and for better explanation, suppose that we calculate PTT by subtracting the $\text{PAT}_{\text{Fingertip}}$ and $\text{PAT}_{\text{Earlobe}}$ as

$$\begin{aligned} PTT &= PAT_{Fingertip} - PAT_{Earlobe} \\ &= (PEP + \frac{L_{AxillaryArtery}}{PWV_{AxillaryArtery}}) - (PEP + \frac{L_{CarotidArtery}}{PWV_{CarotidArtery}}) \\ &= \frac{L_{AxillaryArtery}}{PWV_{AxillaryArtery}} - \frac{L_{CarotidArtery}}{PWV_{CarotidArtery}}. \end{aligned} \tag{2.3}$$

   Consequently, PEP is canceled out. Furthermore, we assume that PWV is the same in both arteries, i.e.,

$$PWV = PWV_{AxillaryArtery} = PWV_{CarotidArtery}, \tag{2.4}$$

   and from (2.3) and (2.4), we have

$$\begin{aligned} PTT &= \frac{L_{AxillaryArtery} - L_{CarotidArtery}}{PWV} \\ &= \frac{L_{Difference}}{PWV}. \end{aligned} \tag{2.5}$$

   Therefore, the PTT estimated with this method does not reflect the speed of pulse wave in any physical section of an artery, but is measured from the length difference between two arteries. It is worth noticing that the equality in (2.4) needs to be discussed as a stand-alone topic.

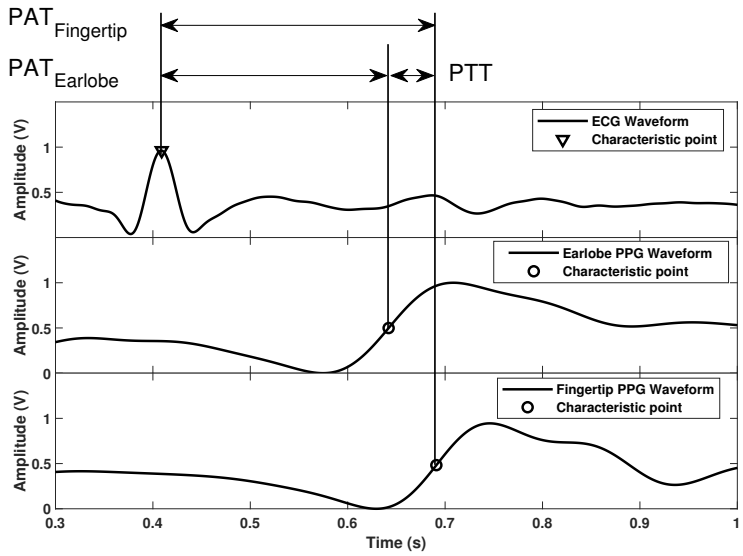

Figure 2.4: Plot of extracting accurate PTT from two distal signals, using the example of earlobe and fingertip-measured PPG signals. ECG signal is not required to obtain PTT, but is used as a reference for explaining how PEP is excluded.

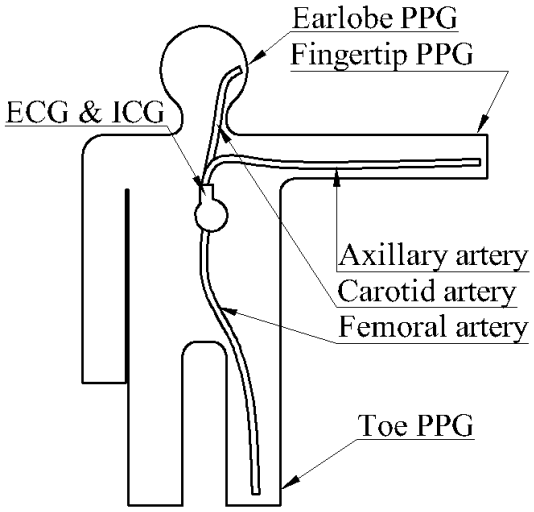

Figure 2.5: A simplified figure of human arteries, as well as the locations that ECG, ICG and PPG signals measure the pulse wave with their characteristics.

The expectation is that using same parametric model, by accurately measuring PTT that excludes PEP, a more accurate BP estimation results will be achieved compared to using PAT. As will be discussed in Section 2.4, this is due to the fact that the models are derived based on the physical relationships between PWV and BP, without expecting PEP to be included

in the measurement. However, results from [27, 28] show that a theoretically more accurate estimation of PWV from PTT does not always have better BP regression performance than using PAT. Their details are discussed below.

Using Method 1, the work from Wong et al. (2011) [27] utilized ECG, fingertip PPG, and ICG signals to extract PAT, PTT and PEP from 22 normotensive subjects to compare their accuracy of BP estimation using a linear model. They found significant variation in SBP, PAT and PEP after intervening subject's BP level through exercising, while PTT and DBP varied insignificantly. Similar conclusion indicating that PEP has higher influence to PAT than PTT, when BP level is intervened by exercising, is also reported in [29]. Pearson's correlation coefficients (R) between the indicators and the BP values were calculated, and are summarized in Table 2.1. Their results showed that PAT has better SBP predictability compared to PTT, in terms of more significant correlation and lower error deviation. For DBP, the comparison is less meaningful due to a generally insignificant correlation close to 0.

Table 2.1: Summary of correlation between PWV-related indicators and BP values proposed by the work from Wong et al. [27]. Since the linear model is used to estimate BP values, absolute value of presented correlation will be equivalent to the absolute value of correlation between estimated and reference BP values.

Average Pearson's Correlation Coefficient (R)	SBP	DBP
PAT	-0.81	-0.16
PTT	-0.61	-0.09
PEP	-0.25	-0.19

Additionally, Proença et al. (2010) [28] utilized ECG, ICG, fingertip PPG and earlobe PPG signals from 20 healthy subjects to make similar comparison that Wong's work did, but using a logarithmic model, with additional PTT extracted with Method 2 from fingertip and earlobe PPG. As is summarized in Table 2.2, the SBP values estimated from PTT extracted with both methods show much lower correlation to the reference values than the results from PAT. No results were reported for DBP because the DBP variation from exercise intervention is small, which is of the same situation observed in Wong's work.

While [27, 28] seem to advocate the advantage of PAT over PTT for SBP estimation, studies from Chen et al. (2009) [30] suggests different results for DBP. Their works used Method 2 in the same way as [28], but replaced fingertip PPG with toe PPG to yield the largest possible

Table 2.2: Summary of correlation results between BP values estimated using a logarithmic model and reference BP values by the work from Proença et al. [28].

Average Pearson's Correlation Coefficient (R)	SBP	DBP
BP= $a_1 \times \ln(PAT) + a_2$	0.85	N/A
BP= $a_1 \times \ln(PTT) + a_2$, Method 1	0.22	N/A
BP= $a_1 \times \ln(PTT) + a_2$, Method 2	0.22	N/A

distance difference of artery ($L_{Difference}$). Using an exponential model, their study involving 23 patients reported an optimal DBP correlation coefficient of 0.94 (converted from reported R-square value), which is better than most of PWV parametric model based works using PAT, in terms of DBP correlation. However, the work lacks results of SBP performance and comparison with BP estimations from PAT under same datasets.

Although the aforementioned works intentionally distinguished PAT and PTT and compared their performance of BP regression, the majority of PWV model-based works, which will be discussed below, neglect the effect of PEP and treat PAT as PWV indicator without differentiating it from PTT.

2.3 Protocol of Literature Review

7 PWV-based parametric models are discussed in the next section, along with selected studies that utilized them with various implementation. For comparison, studies related to each model are summarized as tuples in a table. With each publication reporting various methods and results based on different standards, it is necessary to clarify the protocol we used to decide what to include in our tables of comparison. The columns in each table are explained as follow:

Year, Citation: Studies in each table are sorted in chronological order.

Dataset: The following items are recorded:

- Number of subjects involved in the study as well as their health status are recorded, reflecting the size of dataset as well as the model's adaptability. As most related studies use subject-specific calibration, it is significant to verify if the model works for many individuals. Moreover, the model will be more promising for clinical applications if good results come from dataset consisting of heterogeneous subjects, i.e., including both

healthy individuals and those suffering from BP-related problems.

- Method for acquiring reference BP values in the dataset is recorded. This information indicates the location that reference BPs are measured and the frequency that the reference values have been acquired. It also gives a hint about the total number of estimations in the regression analysis, because each BP value estimated from the model needs to have a reference value for error and correlation calculations. Intervention methods are categorized as follow:

  - **Cuff BP reference**: Using a sphygmomanometer or an oscillometry device, the ordinary non-invasive cuff-based BP measurement method takes a $30 \sim 40$ s measuring procedure to acquire 1 pair of brachial SBP and DBP values from the subject's upper arm.
  - **Arterial BP reference**: Formally named as catheterization, this invasive method measures continuous and instantaneous brachial BP values by placing a gauge in direct contact with blood. 1 pair of SBP and DBP values can be extracted from each beat, which is much more frequent than the ordinary cuff BP method.
  - **Volume clamping BP reference**: This non-invasive method uses a cuff wrapped on subject's finger to acquire continuous and instantaneous finger BP values. For each beat, 1 pair of SBP and DBP values can be extracted.
  - **Finapres BP reference**: Finapres is a medical system that noninvasively measures continuous brachial BP values with both brachial and finger cuff. The brachial BP values are converted from the finger BP values acquired from the finger cuff with volume clamping method. The conversion is based on the reference values from the brachial cuff.

Intervention: BP is relatively steady in short time intervals when the subject is at rest. Since most models include a constant term, least square model parameter calibration with non-varying BP references will assign most model parameters (other than the constant term) to values close to 0, which conceals the potential advantage of the model in tracking BP variations. Therefore, in the data collection procedure, obtaining sufficient reference BP variation

from the subject is of significance for both model calibration (training) and validation (testing). Intervention methods are categorized as follow:

- **None**: This term means that no manual operation is conducted to elevate or lower the BP level. This is applicable when subjects are undergoing surgeries. Absence of intervention will also be reasonable if the BP values are constantly recorded along very long intervals, e.g. for hours, in which BP values can naturally fluctuates.

- **Rest**: This term means that subjects are required to sit still and / or minimize their activity (e.g. keep quiet) to maintain a stable BP value.

- **Exercise**: This term means that physical exercises such as running, stair climbing and cycling are utilized to elevate subject's blood pressure values.

If the employed method does not belong to above categories, it will be specified in the table.

PWV Feature Type: The inputs to PWV models are recorded. Terms are explained below:

- **PAT, PAT_f, PAT_d, PAT_p**: If no additional specification is made, then this term implies that the PAT extracted from the ECG R-peak and the characteristic point of fingertip PPG is utilized as the PWV feature.

- **PTT**: This term implies that PTT, with PEP being excluded using any method, is used as the PWV feature. The particular signals used for extracting PTT are specified after the term.

Recent studies tend to enhance model's performance by adding new indicators to the original model. These additional features, as well as other methods that do not belong to ordinary PAT or PTT, will be specified in tables. If applicable, the total number of extracted features or feature vectors (when more than 1 feature is used), each corresponding to 1 set of BP estimation (i.e., 1 pair of SBP and DBP estimations), is recorded as N.

Calibration: The methods that the studies have employed to determine the unknown parameters are recorded. These methods are categorized as follow:

- **Subject-specific**: This term implies that the model parameters are determined for each

subject using the extracted PWV features and the reference BP values, coming exclusively from that subject. Consequently, each subject will have different model parameters.

- **Grouped**: This term implies that the model parameters are determined for a group of subjects (or all subjects) using the extracted PWV features and the reference BP values chosen from more than one subjects. Consequently, subjects in the group will have same model parameters.

- **Training set**: This term implies that there exists non-overlapping training and testing sets. Model parameters are determined using the training set, while the metrics of error performance are calculated using the model parameters obtained from the training set, and the data in the testing set.

- **Regression**: This term implies that there is no separation of training and testing sets. Both model parameters and the BP estimations are acquired using the whole dataset. Metrics of error performance are evaluated between the estimated model and the reference values over the whole dataset.

Proposed Model: Studies are classified in each table with respect to the parametric models they proposed for the estimation of BP. If a study is categorized in a table corresponding to a model, it means that either the work advocates for the usage of the model or it presents a new method based on the modification of the model, or it reports the best results with the model compared to other methods.

Performance: Recorded in the tables are the optimal **overall** or **typical** performance that reflects the model's performance on most or all the subjects involved in the study. The values are presented as a stand alone $Value$, or in the form of $Value1 \pm Value2$, or N/A indicating that corresponding value is not reported in the publication.

Metrics are evaluated independently for SBP, DBP and mean arterial BP (MAP). These metrics are:

- **Pearson's correlation coefficients (R)**: This metric represents the correlation between the estimated BP and the reference BP.

  - If the study reports R-squared instead of R, it has been converted to R in the tables, assuming that R is positive.
  - Some studies use linear model and report the negative correlation between PAT/PTT and BP instead of the positive correlation between estimated BP and reference BP. In such case if no other results are available, the absolute value of PAT/PTT-BP correlation is regarded as the positive correlation between estimated and reference BP.

- **Error (E)**: Error is defined as the difference between the estimated BP and the reference BP, in mmHg.
  - *Value* indicates the mean error (ME) evaluated on all estimations from all subjects in the group.
  - $Value1 \pm Value2$ indicates ME$\pm$ standard deviation of errors evaluated on all estimations from all subjects in the group.
    - Some studies report the confidence interval instead of the standard deviation. In such situations, the standard deviation is converted from the confidence interval using the z-score table.

- **Absolute error (AE)**: Absolute error is defined as the absolute value of difference between the estimated BP and the reference BP, which is in mmHg.
  - *Value* indicates the mean absolute error (MAE) evaluated on all estimations from all subjects in the group.
  - $Value1 \pm Value2$ indicates MAE $\pm$ standard deviation of absolute errors evaluated on all estimations from all subjects in the group.

- **Root mean square error (RMSE)**: Root mean square error is defined as

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N} Error_i^2}, \tag{2.6}$$

  which is also in mmHg.
  - *Value* indicates the RMSE evaluated on all estimations from all subjects in the group.

– $Value1 \pm Value2$ indicates that the RMSE is individually estimated for each subject's estimation, and is reported as an overall metric in the form of mean value $\pm$ standard deviation.

Key Conclusion: The important points taken from the study are summarized.

2.4 Review of Parametric Models

2.4.1 The Logarithmic Model

Model Derivation

PWV, the speed that the pressure wave created from heart systole travels in the main artery, is related to the stiffness of the blood vessel wall. This relationship is often described by the well-known Moens-Korteweg equation proposed in Moens's study [31]. By modeling the human artery as a cylinder tube with negligible wall thickness filled up with incompressible fluid, PWV can be related to the elasticity of the wall of the tube, and is given as

$$PWV = \sqrt{\frac{Eh}{2\rho R}}, \tag{2.7}$$

in which E is the Young's modulus of artery wall elasticity, h denotes the thickness of artery wall, ρ is the density of blood, and R represents the radius of the artery.

On the other hand, the variation of Young's modulus with the pressure inside the tube is explained by the Hughes' empirical equation proposed in his experiments of dogs [32] as

$$E = E_0 e^{\alpha P}, \tag{2.8}$$

where E is the Young's modulus under loaded pressure, P is the intrinsic pressure, E_0 denotes the Young's modulus under zero pressure, and α is the constant parameter obtained by regression. Using (2.7) and (2.8), the logarithmic model is derived as

$$\begin{aligned} BP &= \frac{2}{\alpha}\ln(PWV) + \frac{1}{\alpha}\ln(\frac{2\rho R}{E_0 h}) \\ &= a_1 \times \ln(PTT) + a_2, \end{aligned} \tag{2.9}$$

where the unknown parameters a_1 and a_2 sum up the effects of aforementioned physical and mechanical constants of the artery, such as stiffness, wall thickness and blood density, to the

relationship between BP and PWV. These constants have been proven to be subject-specific in Hughes' study, and are determined for each particular subject by using regression with some reference BP values measured with conventional methods from that subject. The process of finding these parameters for each subject is known as 'model calibration', which is widely applied in model-based methods reviewed in this chapter.

A problem of this model is that as a_1 will be calibrated to a negative value, large PTT values will correspond to negative BP values, making the model hard to use as the calibrated model has a risk of yielding unrealistic BP estimation with given PTT.

Performance of Related Works

Table 2.3: Summary of studies related to the logarithmic model.

Year	Ref.	Dataset	Intervention	PWV Feature Type	Calibration	Proposed Model	Performance			Key Conclusion
							SBP	DBP	MAP	
2008	[33]	4 hospital subjects during 8 h neurosurgical operation; arterial BP reference	None	PAT_d	Subject specific; regression, 15 beats (max correlation, min deviation of error)	Logarithmic	0.78(R) $N/A \pm 6.8$(ME)	N/A	N/A	Correlation and accuracy degeneration between PAT estimated BP and reference BP is fast
					Subject specific; regression, 360 beats (min correlation, max deviation of error)	Logarithmic	0.63(R) $N/A \pm 12.4$(ME)	N/A	N/A	
2010	[28]	20 healthy subjects; volume clamping BP reference	Exercise	PAT_f	Subject specific; regression	Logarithmic	0.85(R)	N/A	N/A	Compared to ln(PAT) having high correlation, ln(PTT) (either way) and SBP has very low correlation
				PTT (ICG and PPG)			0.22(R)	N/A	N/A	
				PTT (finger and earlobe)			−0.22(R)	N/A	N/A	
2015	[34]	15 healthy subjects; volume clamping BP reference	Exercise	PTT (BCG and finger cuff BP)	Subject specific, regression	Logarithmic	0.67(R)	0.70(R)	0.71(R)	BCG is adequate as a proximal signal; BCG-based PTT may be superior to PAT in estimating DBP
				PAT (ECG and finger cuff BP)			0.59(R)	0.51(R)	0.47(R)	

Derived from the Moens-Korteweg equation [31] and the Hugh's equation [32] proposed in 1963 and 1979, the logarithmic model is one of the earliest models with limited regression performance, and is nowadays used more as a performance baseline to be compared with newly-proposed models (see Table 2.4). However, the simplicity and robustness of the logarithmic model and the linear model, especially compared to models including the $\frac{1}{PTT}$ term being very sensitive to PTT changes and outliers, makes them ideal for preliminary studies when new methods for extracting PAT or PTT are tested out.

Poon et al. (2008) [33] proposed one of the earliest works discussing the problem of fast accuracy degeneration and the need of frequent calibration when applying the parametric model for BP estimation. While [33] referred to this problem on the logarithmic model, the same

problem would later be discussed for the MAP model and the linear model, revealing its universality. [33] utilized linear regression between reference SBP, continuously acquired from invasive arterial BP, and the value of ln(PAT) calculated from beat-to-beat PAT values, on different numbers of consecutively acquired beats. The correlation and accuracy degeneration of the logarithmic model is demonstrated, as the R value between the estimated and reference SBP drops from 0.78 to 0.63, and the standard deviation of the error rises from 6.8mmHg to 12.4mmHg, as the number of beats increases from 15 to 360. The results can be interpreted as follows. If beat-to-beat SBP is estimating using the logarithmic model through PAT after an initial parameter calibration with reference values, then the maximum possible estimation error could double in about 5 minutes if no repetitive calibration is done.

We have already discussed the details of the work from Proença et al. (2010) [28] in Section 2.2.3 distinguishing PTT from PAT. Although the reported SBP correlation (0.85) is among the highest for the logarithmic model, as will be shown in the next sections, it is still not very impressive as other models generally have SBP correlation between 0.85 and 0.95, which again states the inferior performance of the model.

Kim et al. (2015) [34] utilized the logarithmic model in their study proposing a new way for extracting PTT, using ballistocardiogram (BCG) as a substitution for ECG. Their result suggests that the logarithmic model provides better estimates for DBP or MAP, compared to SBP. One limitation of this work is that the distal signal is finger BP (instead of finger PPG).

2.4.2 The Inverse Square Model

Model Derivation

Another model that has been proposed for estimating BP from PTT is the inverse square model [30, 35]. The key difference in deriving this model compared to the logarithmic model is to replace the empirical exponential relationship in (2.8) with a linear version, which is based on the Laplace law in cylindrical tube, when the wall is considered to be thin [36] as follow

$$\sigma = \frac{\Delta P R}{h} = E \times \frac{\Delta R}{R}, \tag{2.10}$$

in which E is the Young's modulus, σ is the circumferential wall stress, ΔP denotes the change of blood pressure, h represents the wall thickness, R is the inner radius of circular cylinder, and

ΔR is the change of radius caused by change of blood pressure, as is displayed in Figure 2.6:

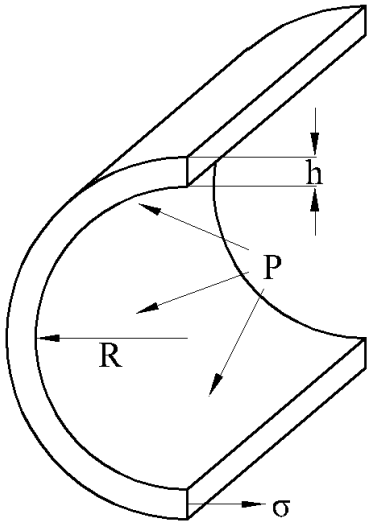

Figure 2.6: Sectional drawing of blood vessel modeled as cylinder tube with thin wall.

Consequently, the E-P relationship will be given as

$$\begin{aligned} E &= \frac{\Delta P R^2}{h \Delta R} \\ &= c_0 \times \frac{P - P_0}{\Delta R}, \end{aligned} \tag{2.11}$$

with $c_0 = \frac{R^2}{h}$. Usually, ΔR is considered to be a small constant that is independent of the blood pressure [35], thus yielding an approximately linear relationship between E and P. Finally, (2.7) and (2.11) yields

$$\begin{aligned} BP &= \frac{2\rho R \Delta R}{c_0 h} \times PWV^2 + BP_0 \\ &= a_1 \times \frac{1}{PTT^2} + a_2. \end{aligned} \tag{2.12}$$

In another approach for deriving the inverse square model [37], the BP-PWV relationship is re-considered from the perspective of the kinetic energy theorem. By considering the work done by the pulse wave on the blood and the change in potential and kinetic energy of the blood as the pulse wave pushes the blood from proximal to distal spot, the kinetic energy theorem yields

the following equivalences

$$\begin{aligned} F \times d &= \frac{1}{2} m \times PWV^2 + mgh, \\ F &= \Delta BP \times s, \\ \rho &= \frac{m}{ad}, \end{aligned} \tag{2.13}$$

in which F is the force created by blood pressure difference loaded on blood, d is the distance that the pulse wave travels, m is the mass of blood, mgh denotes the change in gravitational potential energy of the blood, ΔBP represents the pressure difference, s is the area of cross section of vessel, and ρ is the density of the blood. Rewriting (2.13) as the relationship between ΔBP and PWV yields

$$\Delta BP = \frac{1}{2} \rho \times PWV^2 + \rho gh. \tag{2.14}$$

Finally, the pressure difference is considered to be proportional to arterial blood pressure [38]:

$$\begin{aligned} BP &= \frac{\Delta BP}{0.7} \\ &= a_1 \times \frac{1}{PTT^2} + a_2. \end{aligned} \tag{2.15}$$

Performance of Related Works

Table 2.4: Summary of studies related to the inverse square model.

Year	Ref.	Dataset	Intervention	PWV Feature Type	Calibration	Proposed Model	Performance			Key Conclusion
							SBP	DBP	MAP	
2004	[37]	22 hospital subjects undergoing cesarean section spinal anesthesia; cuff BP reference	Drug	PAT_d ($N = 4660$)	a_1: subject specific, calculated from subject height; a_2: subject specific, recursively calibrated for new-coming reference BP by total least square	Inverse Square	-0.0790 ± 11.32(ME)	N/A	N/A	Derived the inverse square model from theorems of kinetic and gravitational energy
2006	[29]	18 healthy subjects; cuff BP reference	Exercise	PAT_f	Subject specific, regression	Inverse Square	3.6(RMSE)	N/A	N/A	PEP dominants the PAT variation in short-term exercising; proposed a subject-specific calibration method requiring only 1 reference value for new subjects
						Comparison: Inverse	3.9(RMSE)	N/A	N/A	
						Comparison: Logarithmic	4.4(RMSE)	N/A	N/A	
					Averaged sensitivity factor (a_1), subject specific offset factor (a_2)	Inverse Square	7.3(RMSE)	N/A	N/A	
						Comparison: Inverse	6.9(RMSE)	N/A	N/A	
						Comparison: Logarithmic	7.5(RMSE)	N/A	N/A	
2013	[39]	2 anesthesized female sheep; arterial BP reference	Dopamine injection	PAT_d ($N = 12000$)	Subject specific, training set	Inverse Square	0.94(R) $N/A \pm 6.7$(ME) $N/A \pm 5.5$(MAE)	0.85(R) $N/A \pm 6.5$(ME) $N/A \pm 5.3$(MAE)	0.89(R) $N/A \pm 6.4$(ME) $N/A \pm 5.3$(MAE)	PAT is promising for BP estimation
2018	[40]	15 healthy subjects; cuff BP reference	Exercise	PTT (IPG and PPG), IPG	Subject specific, training set	Inverse Square, with a_1 being an additional variable related to IPG	0.88(R) 0.31 ± 8.55(ME) 8.47 ± 0.91(RMSE)	0.88(R) -0.5 ± 5.07(ME) 5.02 ± 0.73(RMSE)	N/A	Added Impedance cardiography to model variation of artery's cross sectional area; results are better than MAP model and linear model
				PTT		Comparison: MAP	0.59(R) 26.43 ± 6.85(RMSE)	0.64(R) 14.53 ± 3.9(RMSE)	N/A	
				PTT		Comparison: Linear	0.59(R) 17.96 ± 3.54(RMSE)	N/A	N/A	

The inverse square model is a decent combination of accuracy and simplicity. With studies

[29, 39, 40] advocating accuracy advantages of the inverse square model compared to logarithmic and linear model, it is also admirable to see that parameter calibration of the model is easily done with a least square linear regression between $\frac{1}{PTT^2}$ and BP values.

The study from Fung et al. (2004) [37] proposed the abovementioned derivation of the inverse square model. In [37], the parameter a_1 is directly calculated for each subject with subject's height and the blood density constant ρ using (2.13), with the pulse transition distance L approximated from subject's height.

The accuracy advantage of the inverse square model can be found in the study from Muehlsteff et al. (2006) [29], in which the regression performance of the inverse square model, the inverse model, and the logarithmic model were compared against each other on the same dataset. Although results for DBP were not presented, it was reported that the inverse square model offers the best accuracy in terms of RMSE under subject-specific calibration condition. When physical exercise is utilized to create large BP elevation (especially for SBP compared to DBP [27]), it seems that the $\frac{1}{PTT^2}$ term, being the most sensitive to PTT variations compared to $\frac{1}{PTT}$ or PTT, offers the advantage of tracking large BP variations. However, such sensitivity could also lead to larger errors if the parameters are not optimally tailored for each subject, which could explain its inferior accuracy when averaged sensitivity factor is used in the comparative results.

Furthermore, the work from Theodor et al. (2013) [39] reported more promising results with the inverse square model using animal subjects. Very high SBP correlation (0.94), decent DBP and MAP correlation (≥ 0.85) and a relatively small error deviation was presented. Demonstrating the highest correlation among other works in Table 2.4, we highlight specified points in this work that may have contributed to enhancing the BP estimation performance besides selection of the model:

- **PPG placement**: In this animal study on two sheep, the reflective PPG sensor is placed directly on the neck muscle under the skin. This setup greatly reduces the power of the AC component of the PPG signal, compared to the PPG signals acquired from human finger, because the density of capillary under the neck muscle is much lower than that under fingertip skin. This setup could be helpful to alleviate the negative effects of pulse wave reflection and change of PWV caused by shrinking of the vessel diameter

and branching from the main artery to peripheral capillaries, which is the case when PPG is acquired from the fingertip. With PPG measured from spot closer to the main artery (carotid artery), the estimated PAT could better represent PWV in the main artery. However, realizing such ideal situations could be impractical when migrating the method to human subjects.

- **Respiration filtering of PAT and BP**: [39] presented two sets of results from raw beat-to-beat PAT and BP values, and from the values filtered with 20-point moving average. The work assumes that PAT and BP values are affected by respiration, and such filtering is helpful for removing the interference, with results confirming better performance after filtering (recorded in Table 2.4), compared to using raw data (R=0.89, ME=$N/A \pm 9.8$ for SBP). This could be essential for enhancing the performance of the inverse square model, since the sensitivity of this model to small PAT variations also means being more prone to producing outlier estimations (i.e. unreasonably large or small BP values), and moving average filtering to PAT values (and reference BP values) is a good practice of smoothing the PAT values. However, this approach could also result in degenerated accuracy when BP changes abruptly instead of smoothly.

Also reported in [39] is an outstanding pulse pressure (PP) estimation performance with inverse square model (R=0.95, ME=$N/A \pm 2.4$). This strengthened the idea presented in MAP model of using inverse square model to estimate PP.

In (2.12), ΔR is considered as a constant to simplify the variables in the model to only PTT and BP. The most recent study from Huynh et al. (2018) [40], however, considered its variation with a new indicator: the arterial impedance measured from the impedance cardiogram (ICG) signal. The inverse square model is consequently modified as

$$BP = a_1 \times \frac{1}{PTT^2} \times \ln(1 + a_2(Z(t) - Z_0)) + a_3, \tag{2.16}$$

in which $Z(t)$is the time-varying body impedance acquired from the ICG signal.

Study [40] concluded that model (2.16) has better performance compared to MAP or linear models, from its comparative study under same dataset. Having $Z(t)$ acquired from the ICG signal as an additional indicator in model (2.16) compared to the original inverse square model,

model (2.16) still maintains its simplicity of requiring only 2 signal inputs, since ICG is also used for PTT extraction. Moreover, with ICG measured at wrist, and PPG measured at fingertip, the method can be integrated compactly on one small device that the user wears on wrist, which is more compact compared to methods requiring ECG nodes to be adhered on body surface with long wires.

Having its advantages, extracting PTT from wrist ICG and fingertip PPG could yields drawbacks in PTT accuracy. The SBP performance from the MAP model ($RMSE = 26.43$ mmHg) and the linear model ($RMSE = 17.96$ mmHg) reported in [40] appear to be much inferior to the SBP performance from the inverse square model ($RMSE = 3.6$ mmHg), the inverse model ($RMSE = 3.9$ mmHg) and the logarithmic model ($RMSE = 4.4$ mmHg) reported in [29], under subject-specific calibration. Besides the difference in dataset and model, it is also expected that while in [40] PTT is acquired from wrist ICG signal and fingertip PPG signal, the short wrist-fingertip distance yields smaller PTT values and larger relative deviation when the signal is affected by noise or artifact, compared to the classical PAT method applied in [29] having a much larger proximal-distal distance. In addition, as is discussed for study [39], the physiological complexity of peripheral vessels when extracting PTT in [40] could also negatively affect the performance reported in [40].

2.4.3 The Square Root Model

Model Derivation

Apart from focusing on the change of the Young's modulus in the logarithmic and the inverse square model, the relationship between the artery stiffness and PWV can be described in other forms. The square root model discusses PWV as the the wave speed in a transmission line model of artery as

$$PWV = \frac{1}{\sqrt{LC}}, \tag{2.17}$$

where C is the compliance, and L is the inertance measuring compressibility (stiffness) and inertia (pressure required to accelerate blood) of the artery. L is a constant defined as ρ/A, where ρ denotes the density of the blood, and A represents the cross-section area. C is related

to the blood pressure via an empirical quadratic equation following [41], and is given by

$$C = \frac{C_m}{1 + (\frac{BP - BP_0}{BP_1})^2}, \tag{2.18}$$

where C_m indicates the maximum compliance, and BP_0, BP_1 denote the essential BP levels acquired from regression.

Finally, using (2.17) and (2.18), a square root relationship between BP and PWV is obtained as [22]

$$\begin{aligned} BP &= BP_1 \times \sqrt{\frac{\rho C_m}{A} \times PWV^2 - 1} + BP_0 \\ &= \sqrt{\frac{\rho C_m BP_1^2}{A} \times PWV^2 - BP_1^2} + BP_0 \\ &= \sqrt{a_1 \times \frac{1}{PTT^2} + a_2} + a_3. \end{aligned} \tag{2.19}$$

Performance of Related Work

Table 2.5: Summary of study related to the square root model.

Year	Ref.	Dataset	Intervention	PWV Feature Type	Calibration	Proposed Model	Performance			Key Conclusion
							SBP	DBP	MAP	
2017	[22]	32 healthy subjects; cuff BP reference	Exercise	PAT_p ($N = 173$)	Subject specific; training set	Square Root	0.95(R) 0.12 ± 6.15(ME) 4.71(MAE)	0.84(R) 1.31 ± 5.36(ME) 4.44(MAE)	N/A	The model is reliable for both PAT and PTT
				PTT_p (PCG and PPG)			0.89(R) -0.28 ± 9.44(ME) 6.22(MAE)	0.84(R) 1.03 ± 5.15(ME) 3.97(MAE)	N/A	

One problem of the square root model in (2.19) is the existence of parameter a_2, which complicates the calibration procedure. A recent study from Esmaili et al. (2017) [22] presents BP estimation results among the top-of-the-line performances of model-based works. In this work, a novel method is proposed to address the parameter calibration problem, in which the model is first calibrated as the inverse model assuming a_2 is equal to 0. After that, gradient descent optimization is applied to acquire the final values of a_1, a_2 and a_3. Realization of this method on other datasets, however, could be more complicated than described. (2.19) implies a_2 to be a negative value, which, during the gradient descent procedure, could drive the model to produce negative values in the $a_1 \times \frac{1}{PTT^2} + a_2$ term under the square root. Restrictions to training epochs, learning rates or gradient directions are therefore required.

2.4.4 The Linear Model

Model Derivation

Another model used for estimating BP from PTT is the linear model. It is natural to assume linear relationship between PTT and BP, if high correlation is observed, which is often the case for SBP. Alternatively, the linear model can be considered as a small signal linearization of non-linear model at a given nominal point, around which the PTT does not change much and all other related parameters can be assumed to be constant within a short interval around the considered point. This situation is true in many cases, which can explain the reason that changing from non-linear models to linear model sometimes does not significantly alter the correlation between the estimated and reference BPs. To derive the linear model, one can start with the logarithmic model in (2.9). Applying linearization yields [42]

$$\begin{aligned} BP &= BP_0 + (PTT - PTT_0) \times \left. \frac{d(c_1 \times \ln(PTT) + c_2)}{dPTT} \right|_{PTT=PTT_0} \\ &= BP_0 + (PTT - PTT_0) \times \frac{c_1}{PTT_0} \\ &= a_1 \times PTT + a_2, \end{aligned} \tag{2.20}$$

where BP_0 and PTT_0 denote BP and PTT at the nominal point, c_1 and c_2 represent the model parameters determined for the logarithmic model, and a_1 and a_2 are the model parameters of the linear model, which is given by

$$\begin{aligned} a_1 &= \frac{c_1}{PTT_0}, \\ a_2 &= BP_0 - c_1. \end{aligned} \tag{2.21}$$

In practice, a_1 and a_2 are often acquired by simple linear regression.

Performance of Related Works

With linearization of the logarithmic model proposed by Chen et al. (2000) [42], the linear model is undoubtedly the most cited and used model. In [42], BP estimation is conducted with consecutive beat-to-beat PAT and reference BP (SBP, particularly) values. The time-varying BP values of the subject is considered as the sum of two time series, given by

$$BP(t) = BP_{HF}(t) + BP_{LF}(t), \tag{2.22}$$

Table 2.6: Summary of studies related to the linear model.

Year	Ref.	Dataset	Intervention	PWV Feature Type	Calibration	Proposed Model	Performance			Key Conclusion
							SBP	DBP	MAP	
2000	[42]	20 cardiovascular surgery subjects; arterial BP reference	None	PAT_f, beat-to-beat, interpolated to PAT signal @1Hz	Subject specific; $BP_{HF} = PAT \times a_1$; BP_{LF} calibrated from interpolation of intermittent calibrated BP values (every 5 minute)	Linear ($BP_{LF} + BP_{HF}$)	0.97(R) 0.06(ME) 3.7(RMSE)	N/A	N/A	High frequency components of PAT can track high frequency variation of SBP (small signal model method)
2006	[43]	20 healthy male subject; arterial BP reference	Drug	PAT	Subject specific; regression	Linear	0.62(R) $N/A \pm 8.67$(ME)	0.14(R)	0.28(R)	PEP significantly contribute to PAT; PAT is not usable for predicting DBP or MAP
				PTT (ICG and PPG)			0.57(R)	0.64(R) $N/A \pm 8.83$(ME)	0.67(R)	
2009	[44]	14 healthy subjects, repeated after half-year; cuff BP reference	Exercise	PAT_d	Subject specific; regression, initial dataset	Linear	0.92(R) 0 ± 4(ME)	0.38(R) 0 ± 3.5(ME)	N/A	Regression coefficients obtained half year ago could not predict BP well in all subjects
					Subject specific; regression, repeated dataset		0.87(R) 0 ± 5.3(ME)	0.3(R) 0 ± 2.9(ME)	N/A	
					Subject specific; repeated dataset, using parameters obtained from initial dataset		1.4 ± 10.2(ME)	2.1 ± 7.3(ME)	N/A	
2011	[27]	22 healthy subjects; cuff BP reference	Exercise	PAT_d	Subject specific; regression	Linear	0.81(R) 3.1 ± 5.0(2, rest) -1.4 ± 5.4(2, exercised)	0.16(R) 2.2 ± 3.5(2, rest) -1 ± 3(2, exercised)	N/A	PAT with inclusion of PEP is significant for achieving good SBP estimation accuracy
				Comparison: PTT (ICG and PPG)			0.25(R) 4.8 ± 5.2(2, rest) -2.2 ± 9.8(2, exercised)	0.09(R) 2.2 ± 3.7(2, rest) -1 ± 3.1(2, exercised)	N/A	
				Comparison: PEP			0.61(R) 4.4 ± 4.8(2, rest) -2 ± 7.2(2, exercised)	0.19(R) 2.4 ± 3.1(2, rest) -1.1 ± 3.0(2, exercised)	N/A	
2016	[45]	22 subjects; volume clamping BP reference	Mental arithmetic, cold pressor, exercise (3 elevated BP and 3 rest BP for each subject)	PTT_f (BCG and foot PPG)	Subject specific; regression	Linear	0.80(R) 8.50 ± 0.70(RMSE)	0.80(R) 5.7 ± 0.4(RMSE)	N/A	Whether PEP tracks BP variation is dependent of BP intervention method
				Comparison: PAT_f (ECG and finger PPG)			0.66(R) 10.1 ± 0.7(RMSE)	0.60(R) 7.1 ± 0.6(RMSE)	N/A	

where BP_{HF} is the high frequency variation, and BP_{LF} is the baseline (DC component) and low frequency variation. The work aims at estimating BP_{HF} by linearly scaling the high frequency component of extracted beat-to-beat PAT values, while directly acquiring BP_{LF} from interpolation of intermittently-acquired reference BP values. Consequently, BP is estimated using

$$BP(t) = a_1 \times PTT(t) + BP_{LF}(t). \tag{2.23}$$

While a_2 in (2.20) is a constant determined by the DC component of BP and PTT, it is replaced by $BP_{LF}(t)$ in (2.23), which is not a constant as it includes the low frequency variation of BP. This highlighted difference between (2.20) and (2.23) adds some restrictions to the application of (2.23) that produced excellent SBP estimation performances in [42]. We illustrate this via an example. Assume we would like to estimate BP at 2.5 min from PTT measured at 2.5 min. This is doable using (2.20). However, to estimate BP from (2.23), we first need to find BP_{LF} at 2.5 min, which according to [42] is obtained from linear interpolation between BP measured at 0 min and at 5 min. This dependency on future values limits the application of (2.23) in realistic scenarios.

Similar to [27] discussed in Section 2.2.3, Payne et al. (2006) [43] compares the linear

regression performance of PAT and PTT, while changing the method of BP intervention from exercise to drug. The work drew a similar conclusion that PEP dominates PAT variation after BP intervention, while reporting a better DBP correlation when using PTT ($R^2 = 0.41$ between PTT and DBP, $|R| = 0.64$) compared to [27] ($R = -0.09$ between PTT and DBP), which could be the result of intervention difference, as in [27] insignificant DBP increment of 11.8% after exercise was reported, while in [43] DBP decreases as much as 44.6% after the usage of salbutamol, i.e., a much larger variance of reference DBP value to be correlated with PTT. [43] intuitively demonstrated how the difference in intervention method could affect the outcomes.

In another work that employed linear model from Wong et al. (2009) [44], the problem of accuracy reduction with data obtained half a year later than the initial data was collected. The results clearly demonstrated how variations in estimation error doubled when the model parameters were obtained with the data from half a year earlier ($ME = 1.4 \pm 10.2$ mmHg for SBP, $ME = 2.1 \pm 7.3$ mmHg for DBP), compared to results from re-calibrated parameters ($ME = 0 \pm 5.3$ mmHg for SBP, $ME = 0 \pm 2.9$ mmHg for DBP).

In another work, Martin et al. (2016) [45] proposes a weight-scale like device that extracts PTT from BCG and foot PPG, demonstrating a better linear regression performance from PTT compared to conventional PAT. With BCG signal marking time instances of ventricle blood ejection, it seems that the PTT measured with this method takes the advantages of smaller relative error from the long transmission distance compared to conventional combination of ECG and fingertip PPG, and less PWV variation as the pulse wave travels along femoal artery without much interference from vessel branching. Another novelty of this study is its rich selection of BP intervention methods, in which mental arithmetic, cold pressor and physical exercise (stair climbing) are experimented on each subject in sequence. The study chose these methods to demonstrate how they affect BP and PEP differently: mental arithmetic and physical exercise increase BP and decrease PEP, while cold pressor increases both BP and PEP, revealing a potential bias in the datasets from works advocating PAT over PTT, since most of them used physical exercise for BP intervention in which PEP varies concertedly with PTT.

2.4.5 The Inverse Model

Model Derivation

Similar to the linear model, a direct linear correlation can be established between PWV and BP [46, 47]. Under such an assumption, the inverse PTT model is just the linear model between PWV and BP.

Derivation of the inverse model [48] is a linear approximation of the square root model depicted in (2.19)

$$\begin{aligned} BP &= BP_1 \times \sqrt{\frac{\rho C_m}{A} \times PWV^2 - 1} + BP_0 \\ &\approx \sqrt{\frac{\rho C_m BP_1^2}{A}} \times PWV + BP_0 \\ &= a_1 \times \frac{1}{PTT} + a_2. \end{aligned} \tag{2.24}$$

Performance of Related Works

Table 2.7: Summary of studies related to the inverse model.

Year	Ref.	Dataset	Intervention	PWV Feature Type	Calibration	Proposed Model	Performance			Key Conclusion
							SBP	DBP	MAP	
2011	[49]	33 healthy subjects; cuff BP reference	Exercise	PAT_d	Subject specific; training set	Inverse	0.94(R) -0.058 ± 6.61(ME)	0.88(R) -0.25 ± 5.64(ME)	N/A	The mid-term (30 days) BP estimation accuracy is only slightly deteriorated
		10 recalled subjects after 30 days			Using earlier calibrated parameters		0.90(R) -1.74 ± 7.72(ME)	0.86(R) -0.36 ± 5.80(ME)	N/A	
2012	[50]	63 subjects free of cardiovascular diseases; cuff BP reference	Exercise	PAT_d, height	$a_1 \sim a_4$: training set; b: subject specific one-point calibration; train on 13 subjects ($N = 58$)	Exponential + polynomial of PWV (from PAT and height)	0.89(R)	N/A	N/A	Group calibration on most parameters, combined with individual calibration on the rest, is feasible for SBP estimation
					Test on 50 subjects ($N = 267$)		0.83(R) $N/A \pm 10.1$(ME)	N/A	N/A	
2014	[51]	31 healthy subjects; cuff BP reference	Posture change	PAT_p	Subject specific; train on sit and stand dataset, test on another sit dataset	Inverse +Linear +std(PAT)	0.980(R) -0.2 ± 2.4(ME)	0.841(R) -0.5 ± 3.9(ME)	N/A	Proposed model is better than the inverse model, especially for DBP
						Comparison: Inverse	0.978(R) 0.1 ± 2.5(ME)	0.517(R) 1.3 ± 7.4(ME)	N/A	

Although the literature review [48] demonstrated derivation of the inverse model by simplifying the square root model, other works [50, 51] have developed their own models starting from a linear relationship between PWV and BP, which results in the inclusion of the $\frac{1}{PTT}$ term. Assuming a linear relationship between PWV and BP seems to be more promising than that of between PTT and BP, since no restriction of small signal approximation is required: a large PTT value will neatly predict a small, positive BP value, instead of a negative BP value that the linear model may do.

Masè et al. (2011) [49] used the inverse model to test the validity of model parameter with subjects' data obtained 30 days after their initial measurement, producing a high SBP and DBP accuracy with acceptably small accuracy degeneration after 30 days. It is worth noticing that results reported from this work are very close to the results reported in [22] using the square root model, which could imply the efficiency of simplification from square root model to inverse model, in terms of much easier parameter calibration with no compromise to the accuracy.

Gesche et al. (2012) [50] proposed an empirical model between PWV and BP as

$$\begin{aligned} BP &= a_1 \times PWV \times e^{a_3 \times PWV} + a_2 \times PWV^{a_4} + b, \\ PWV &= \frac{c \times h}{PTT}, \end{aligned} \tag{2.25}$$

where h is the height of the subject and c is a known constant related to sensor placement. This study is a very rare example in model-driven works in which most model parameters are trained with grouped data from multiple subjects. In the proposed method taking PAT_d and subject's height as model input, parameters a_1, a_2, a_3 and a_4 are obtained from a training set consisting of reference values from 13 subjects. When the model is applied to the testing set of 50 subjects, only b is determined with one-point calibration for each subject. Although no corresponding metric is reported to compare this work with [29], which also endeavored to yield a mostly-universal model with constant parameters for all subjects, reported SBP correlation and error performance of this work can still be rated as decent.

In addition, [50] provides a great idea for many models to relieve the requirement of individually calibrating a_1. In most models, the sensitivity factor a_1 summarizes the effects of multiple properties, including the subject specific pulse wave transition distance. If metrics such as subject's height can be used as an additional parameter in the model to explain the difference of pulse wave transition distance among subjects, then the difference of transition distance among subjects will no longer varies a_1, which makes it possible to obtain a unified model with decent performance.

Ma et al. (2014) [51] proposed another form of the inverse model, which generates 1 pair

of BP estimations from 5 successively-acquired beat-to-beat PAT values as

$$\begin{aligned} BP &= a_1 \times \frac{1}{\text{mean}(PTT)} + a_2 \times \text{mean}(PTT) + a_3 \times \text{std}(\Delta PTT) + a_4, \\ PTT &= \{PTT_i\}, i = 1,2,3,4,5\ , \\ \Delta PTT &= \{\Delta PTT_i\} = \{PTT_i - PTT_{i-1}\}, i = 1,2,3,4,5\ , \end{aligned} \tag{2.26}$$

where mean(x) and std(x) stand for the average value and the standard deviation of elements in set x, respectively. The study reported considerable improvement in DBP estimation accuracy compared to the inverse model, and an excellent SBP accuracy with smallest error deviation and highest correlation coefficient among all works reviewed in this chapter. However, considering the fact that the testing set of this work only included data measured from seated subjects, the high accuracy could be due to smaller SBP variation in the dataset, compared to other works having both rest and elevated BP levels in their testing set. An example can be find in [27], in which the linear model showed larger error variation when it is tested on the exercised dataset, compared to on the resting dataset.

2.4.6 The MAP Model

Model Derivation

The MAP model combines the logarithmic model and the inverse square model as a trial to make utilization of these equations as close as possible to their original definition. Note that while in (2.8) P is the absolute pressure, in (2.10) the linear relationship is between E and ΔP. Therefore, the MAP model estimates PP, the increment between SBP and DBP, with inverse square model, while estimating the MAP with the logarithmic model [52].

The relationship between SBP, DBP, MAP and PP is shown as

$$\begin{aligned} MAP &= \frac{1}{3} SBP + \frac{2}{3} DBP, \\ PP &= SBP - DBP, \end{aligned} \tag{2.27}$$

while MAP and PP are estimated as:

$$\begin{aligned} MAP &= a_1 \times \ln(PTT) + a_2, \\ PP &= a_3 \times \frac{1}{PTT^2}, \end{aligned} \tag{2.28}$$

and SBP and DBP can be derived as:

$$\begin{aligned} SBP &= MAP + \frac{2}{3}PP \\ &= a_1 \times \ln(PTT) + \frac{2a_3}{3} \times \frac{1}{PTT^2} + a_2, \\ DBP &= MAP - \frac{1}{3}PP \\ &= a_1 \times \ln(PTT) - \frac{a_3}{3} \times \frac{1}{PTT^2} + a_2. \end{aligned} \tag{2.29}$$

Performance of Related Works

Table 2.8: Summary of studies related to the MAP model.

Year	Ref.	Dataset	Intervention	PWV Feature Type	Calibration	Proposed Model	Performance			Key Conclusion
							SBP	DBP	MAP	
2006	[53]	85 subjects (39 hypertensives); cuff BP reference	None	PAT, averaged over 45 s ($N = 999$)	Subject specific; training set	MAP	0.6 ± 9.8(ME)	0.9 ± 5.6(ME)	N/A	Proposed the MAP model and its potential for cuff-less BP estimation
2013	[54]	15 healthy subjects; cuff BP (every 2 min, for 30 min)	None	PAT_p	Subject specific; intermittent calibration every 8 min (min deviation of error)	MAP	1.08 ± 9.74(ME)	N/A	N/A	Large error, while smaller change of error indicates that MAP model better capability of explaining BP variation than linear model
					Subject specific; intermittent calibration every 6 min (max deviation of error)		1.4 ± 10.59(ME)	N/A	N/A	
					Subject specific; intermittent calibration every 4 min (min deviation of error)	Comparison: Linear [42]	-0.03 ± 0.07(ME)	N/A	N/A	Rapid accuracy degeneration after the interval between two calibrations exceeds 8 min
					Subject specific; intermittent calibration every 12 min (max deviation of error)		-3.42 ± 29.22(ME)	N/A	N/A	
2014	[52]	10 healthy subjects; cuff BP reference (every 30 min, for 24 h)	None	PAT_d, acquired at nighttime, PAT and SBP interpolated and smoothed to align the time stamp ($N = 70$)	Subject specific; training set	MAP	-0.8(R, between SBP and PAT) 2.4 ± 5.7(ME) 6.2(RMSE)	N/A	N/A	Revealed problem of synchronization and energy consumption in long time signal and reference acquisition
2015	[55]	27 healthy subjects; Finapres reference BP	None	PAT_d, PPG ($N = 1713$)	Subject specific; training set	MAP, replaced MAP estimation with DBP estimation from PPG intensity ratio	0.91(R) -0.37 ± 5.21(ME) 4.09(MAE)	0.88(R) -0.08 ± 4.06(ME) 3.18(MAE)	0.89(R) -0.18 ± 4.13(ME) 3.18(MAE)	Beat-to-beat BP is more effectively estimated with PPG intensity ratio estimating low frequency variations (DBP) and PAT estimating high frequency variations (PP)
						Comparison: MAP	-0.11 ± 7.31(ME) 5.76(MAE)	0.19 ± 6.03(ME) 4.80(MAE)	0.09 ± 6.25(ME) 4.96(MAE)	
						Comparison: Linear	0.19 ± 6.21(ME) 4.94(MAE)	N/A	N/A	

Earliest usage of the MAP model can be traced back to the work from Poon et al. (2006) [53], reporting an inferior SBP accuracy ($ME = 0.6 \pm 9.8$ mmHg) and a similar level of DBP accuracy ($ME = 0.9 \pm 5.6$ mmHg), compared with works using only the inverse square model [39] ($ME = N/A \pm 6.7$ mmHg for SBP, $ME = N/A \pm 6.5$ mmHg for DBP). It seems that more evidence is required to prove the advantage of the MAP model over the logarithmic model and the inverse square model that it combined.

After the MAP model and the original linear model from [42] were proposed, McCarthy et al. (2013) [54] tried to find their performance with variable intervals between intermittent calibrations, concluding that none of these two models are reliable. The authors validated the

performance of these two models by gradually increasing the interval between two intermittent calibrations from 2 min to 12 min. For the linear model, large accuracy degeneration was clearly observed, as the standard deviation of error rises from a minimum of 0.07 mmHg (4 min) to a maximum of 29.22 mmHg (12 min). For the MAP model, the difference between the best case and worst case accuracy was relatively smaller, showing the advantage of the MAP model over the linear model on estimating BP. However, the accuracy of the MAP model was still mediocre compared to other models.

Zheng et al. (2014) [52] proposed a unique study of cuff-less BP estimation considering a situation closer to real-life continuous BP monitoring. In most studies, data from each subject, including physiological signals and reference BP values, are acquired consecutively during short sessions that last from minutes to hours, in which fast BP variation is manually created via intervention. The accuracy of estimation results consequently only reflects model's predictability over a short interval, while in real-life cases of continuous BP monitoring, subject's BP value could vary slowly and naturally in longer intervals, while the monitor is expected to provide accurate BP readings 24 hours a day. To address this limitation, subjects in this study wore ECG, PPG and reference BP monitor for 24 hours, and signals were automatically recorded every 30 minutes. Challenges were encountered in keeping the sensor signals and reference BP readings synced under limited power budget and a long running time of 24 hours, but were addressed with an approximate method including outlier removal, signal interpolation and data smoothing. The results show low correlation between daytime reference BP values and PAT values, and a relatively higher correlation of -0.8 with data acquired during nighttime. The MAP model was tested on nighttime data, and a biased error performance of 2.4 ± 5.7 mmHg was obtained under the best case settings. This study is a reminder of the awaiting challenges between the decade-long model-based BP estimation studies and their real-life application outside laboratory environments.

Ding et al. (2015) [55] improved the MAP model by replacing the MAP estimation obtained from the logarithmic model with DBP estimation from a new indicator, the PPG intensity ratio (PIR), similar to the model described in (2.16). Furthermore, spectral analysis of time-varying BP, PAT and PIR values demonstrated that both low frequency and high frequency variations are present in SBP, while DBP only has low frequency components. The study states that PAT

has only high frequency variations, while the complementary low frequency components are found in PIR. Consequently, the study advocates for the usage of additional indicators in the model as

$$DBP = a_1 \times \frac{1}{PIR}, \tag{2.30}$$

$$PP = a_2 \times \frac{1}{PTT^2}, \tag{2.31}$$

$$SBP = DBP + PP, \tag{2.32}$$

$$PIR = \frac{\max(VPPG)}{\min(VPPG)}, \text{ within each PPG cycle,} \tag{2.33}$$

where $VPPG$ is the amplitude of PPG signal. The proposed method shows improvement in SBP, DBP and MAP accuracy in comparison to the MAP model and linear model without requiring additional inputs besides ECG and PPG.

2.4.7 The Trimmed Inverse Square Model

Model Derivation

The trimmed inverse square model aims at addressing problems in the linear model described in (2.20). Note that in linear model, parameter a_1 will be calibrated to a negative value, which is inappropriate for dealing with large BP variation because:

- The linearization is under the assumption of small PTT variation around the nominal point.
- The linear model indicates that large BP values correspond to negative PTT values, which is unrealistic.

To address these problems, Wibmer et al. (2014) [56] empirically derived the trimmed inverse square model based on the inverse square model to create an asymptotic behavior for larger BP values ($PTT \to PTT_0$ as $BP \to \infty$), by adding an additional degree of freedom to the inverse square model:

$$BP = a_1 \times \frac{1}{(PTT - a_2)^2} + a_3. \tag{2.34}$$

In the inverse square model, parameter a_1 changes the curvature of the quadratic curve that relates $\frac{1}{\text{PTT}}$ and BP, while parameter a_2 shifts the PTT-BP regression curve on Y (BP) axis. The

trimmed inverse square model, in addition, allows for the regression curve to shift on X (PTT) axis, which is expected to help further reduce the BP estimation error.

Performance of Related Work

Table 2.9: Summary of study related to the trimmed inverse square model.

Year	Ref.	Dataset	Intervention	PWV Feature Type	Calibration	Proposed Model	Performance			Key Conclusion
							SBP	DBP	MAP	
2014	[56]	20 patients; cuff BP reference	Exercise	PAT_d, smoothed ($6 \sim 9$ estimations per subject)	Subject specific; regression	Trimmed inverse square model	0.975(R) 0 ± 5.56(ME)	0.653(R) 0 ± 4.54(ME)	N/A	Proposed the trimmed inverse square model and its advantage over linear model
						Comparison: Linear	0.967(R) -0.05 ± 6.71(ME)	0.566(R) 0 ± 4.85(ME)	N/A	

Few discussion have been made to this model besides its original publication from Wibmer et al. (2014) [56]. The work only reported subject specific correlation performance, therefore the recorded values are the mean correlation coefficient among all subjects. Achieving better SBP and DBP accuracy over the linear model, the SBP estimation performance is excellent among all subjects with very high correlation and top-level accuracy, while the DBP correlation is not so novel.

2.5 Conclusions, Limitations and Suggested Future Works

In this chapter we provided an overview of most commonly used models for estimating blood pressure, and discussed their advantages and disadvantages. In summary, the following points can be made for the model-based BP estimation methods:

- **Nonlinear vs linear models**: As seen in Table 2.6 and in studies that used the same dataset to compare models, we can conclude that nonlinear models perform better than linear models. Theoretically, the linear model (as seen in (2.20)) is derived from small signal linearization of the logarithmic model and as such, its validity is only true under very restrictive conditions [42], and it won't be usable under large BP and PTT variations. Practically, the nonlinear term in nonlinear models enables the model to have large BP sensitivity under small PTT changes, which is more feasible considering the fact that the large pulse wave speed and short transition distance makes both the PTT value and its changes small.

- **Performance variability of the same model**: Differences in dataset in terms of subject's health status and BP intervention methods, as well as other factors including calibration, data smoothing and signal filtering could result in different regression results for the same model, which leads to variable performance. Therefore, optimizing the entire process from raw signal to BP estimations appears to be more important than finding the best performing model.

- **Subject-specific calibration**: The vast majority of model-based methodologies in the literature require subject-specific parameter calibration, which is incompatible with some health standards [23] and could potentially limit the application of all model-based methods. However, the problem could hopefully be addressed by adding more indicators to the model, such as subject's height, as shown in studies from Fung et al. [37] and Gesche et al. [50]. Moreover, study like [57] advocates for developing a new standard (IEEE 1708 project [58]) that is specifically designed for cuff-less BP monitors, which accepts calibration as part of the device validation procedures.

- **Accuracy degradation over time**: Decreases in model accuracy have been reported over short-term ([54, 33], from beats to minutes), mid-term ([49], 30 days) and long-term ([44], half a year). These results suggest that physiology parameters in the model can only be assumed to be constant over a short time period, suggesting the need for frequent intermittent calibration. Additional indicators extracted from sensor signals could help reflecting the real-time physiological status of the subject, addressing this problem. However, few studies have validated the performance of modified models with additional indicators over time, and future studies are needed.

In addition, limitations of reviewed studies in the chapter are summarized as follow, with recommended future works given accordingly:

- **Limitation of short sessions**: Most studies were limited to lab environments with manually created fast BP variations. Such settings are not consistent with real-life environment for continuous BP monitoring. The low daytime correlation between PAT and BP reported in [52] could suggest that some PWV model-based methods may lack the ability to track slow varying BP values during long sessions. While the literature has already

devoted much attention to study the accuracy degradation with data acquired in short sessions, further studies that bring the experiments closer to real-life application are needed.

- **Limitation of BP intervention**: It can be observed that most studies advocating the advantage of PAT over PTT were using physical exercises for BP intervention, while [45] has provided evidence that the performance of PAT and PTT is dependent on the method used for BP intervention. In addition, while most intervention methods can effectively trigger an variation in SBP, some of them are not so effective at changing the DBP [27]. It is thus suggested that multiple intervention methods should be used in the same study to select the optimal model with the presence of heterogeneous BP variation.

- **Limitation of comparability across studies**: Different choice of performance metrics reported in different studies makes the comparison among different works difficult, because the 4 widely applied metrics (R, ME, MAE, RMSE) are not interconvertible. Although differences in datasets might imply that such comparisons are meaningless, reporting multiple metrics, the distribution of reference SBP and DBP levels, as well as the total amount of estimations used when calculating these metrics can enable fair comparison across works.

- **Larger transition distance and alternative sensor placement**: Advantages of larger pulse wave transition distance and alternative sensor placement setup being closer to the main artery were discussed in [39] and [45]. Longer distance between proximal and distal spots scales up the PTT value and reduces the relative change caused by the interference, while alternative sensor placement could help reducing the negative effect of artery branching and wave reflection that deviate the experiment situation from the theoretical situations. Future work can further explore the feasibility of placing the sensors at different locations as alternatives to ECG and fingertip PPG.

Chapter 3

Data-driven Methodologies for Cuff-less Blood Pressure Estimation

3.1 Introduction

In Chapter 2, we presented an overview of model-driven methods for estimating BP. Another class of cuff-less BP estimation approaches is based on data-driven methods, which are going to be reviewed in this chapter. While model-driven methods are theoretically supported by the PWV-BP relationship, data-driven methods mostly originate from the PPG-BP relationship, with no intersection with PWV.

In this chapter, we focus on reviewing existing data-driven studies of cuff-less BP estimation. As will be shown, most data-driven methods use either only PPG features, or PPG features plus PAT. We first provide explanations about basic machine learning methods, and then review existing studies by categorizing them with respect to the input signal(s), which generally falls into three categories: studies using only PPG signal, studies using both PPG and ECG signals, and studies using only ECG signals. After discussing the selection of features, the machine learning methods and the BP estimation accuracy of these studies, we summarize the novelty and limitations of existing data-driven studies.

3.2 Review of Machine Learning (ML) Methods

While model-driven studies tend to use standardized features (PTT or PAT) and customized models, data-driven studies often utilize standardized regression models with customized selection of features. Therefore, a brief summary of standardized machine learning methods is presented before discussing features that lead to BP estimation. Specifically, two types of methods are addressed: the multiple regression methods for spontaneously forming the mathematical relationship between multiple features and the BP to be estimated, and the feature

selection methods for reducing and controlling the size of feature vector.

3.2.1 Multiple Regression Methods

Suppose that N training entries of the feature vector $\mathbf{x} \in \mathbb{R}^P$ and corresponding regression target $y \in \mathbb{R}$ (reference blood pressures, in our case) are given in pairs as a training set $\{(\mathbf{x}_i, y_i)\}_{i=1}^N$. A multiple regression method aims at finding a function that gives estimation of y: $\hat{y} = F(\mathbf{x})$ from input $\mathbf{x}$, with F determined by minimizing the loss function $L(\{(y,\hat{y})\})$ evaluated on the training set as

$$F = \underset{F:\mathbb{R}^P \to \mathbb{R}}{\operatorname{argmin}} L(\{(y_i, \hat{y}_i = F(\mathbf{x}_i))\}_{i=1}^N). \tag{3.1}$$

Specifically in our application of BP estimation, the loss function is most commonly selected to be the L2 loss, i.e. the (mean) squared error given by

$$L(\{(y_i, \hat{y}_i)\}_{i=1}^N) = \sum_{i=1}^{N} (\hat{y}_i - y_i)^2. \tag{3.2}$$

The major difference between regression methods is the assumptions they make to form function F.

Multiple Linear Regression (MLR)

The MLR method assumes F to be a linear model

$$\begin{aligned} F_{MLR}(\mathbf{x}) &= \mathbf{x}^\top \mathbf{a} + b \\ &= \begin{bmatrix} \mathbf{x}^\top & 1 \end{bmatrix} \times \begin{bmatrix} \mathbf{a} \\ b \end{bmatrix} \\ &= \tilde{\mathbf{x}}^\top \mathbf{w},\ \mathbf{a} \in \mathbb{R}^P,\ b \in \mathbb{R},\ \tilde{\mathbf{x}}, \mathbf{w} \in \mathbb{R}^{P+1}, \end{aligned} \tag{3.3}$$

where $\mathbf{a}$ is the normal vector, and b is the intercept term. To simplify the model into a matrix multiplication form, $\tilde{\mathbf{x}}$ and $\mathbf{w}$ are formed as the augmented input vector and augmented regression coefficients vector.

Advantages of this method are its simplicity and efficiency, as F (i.e. $\mathbf{w}$) has closed-form

solution subjecting to (3.1) and (3.2), which is given as

$$\begin{aligned}
\mathbf{w} &= (\tilde{\mathbf{X}}^\top \tilde{\mathbf{X}})^{-1} \tilde{\mathbf{X}}^\top \mathbf{y}, \\
\tilde{\mathbf{X}} &= \begin{bmatrix} \mathbf{x}_1^\top & 1 \\ \mathbf{x}_2^\top & 1 \\ \vdots & \vdots \\ \mathbf{x}_N^\top & 1 \end{bmatrix}, \\
\mathbf{y} &= \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix},
\end{aligned} \tag{3.4}$$

where $\tilde{\mathbf{X}}$ is the augmented matrix of independent variables and $\mathbf{y}$ is the vector of dependent variables, which are formed by concatenating the feature vectors $\mathbf{x}_i$ and the regression targets y_i observed in the training set $\{(\mathbf{x}_i, y_i)\}_{i=1}^N$.

The MLR method cannot form nonlinear relationships between the features in $\mathbf{x}$ and the regression target y, which is a major limitation in our application considering the nonlinearity of the physiological mechanisms and the empirical inferiority of simple linear model, as demonstrated in Chapter 2. Therefore, the key to improve the regression method for BP estimation is in involving nonlinear relationships in F.

Kernel-based Methods: Support Vector Regression (SVR) and Kernel Ridge Regression (KRR)

One way to introduce nonlinearity in the model is to manually add new values calculated with nonlinear functions from the original features to the feature vector before utilizing linear regression, which is similar to the formation of nonlinear PWV models described in Chapter 2. The kernel-based approaches can be interpreted as systematically-optimized methods of this idea, which enables the model to be nonlinear while using the kernel method to address the increase in computational complexity. In order to demonstrate the idea of kernel methods, we use KRR as example.

Suppose we have a mapping function $\phi(\mathbf{x}) : \mathbb{R}^P \to \mathbb{R}^Q$ that extends the original feature

vector $\mathbf{x}$ of dimension P to Q by adding nonlinear terms calculated from original features in $\mathbf{x}$. The problem is that as we solve for $\mathbf{w}$ using (3.4), the computational complexity approximately grows from $O(P^3)+O(P^2N)$ to $O(Q^3)+O(Q^2N)$ due to the high cost of calculating $\tilde{\mathbf{X}}^\top\tilde{\mathbf{X}}$ and the matrix inversion, which is an unacceptable growth. Two modifications are therefore made to this method:

Dual solution of w: First, an additional regularization term besides the L2 loss is added into the loss function

$$L(\{(y_i,\hat{y}_i)\}_{i=1}^N,\mathbf{w})=\sum_{i=1}^N(\hat{y}_i-y_i)^2+\lambda||\mathbf{w}||_2^2, \tag{3.5}$$

where λ is the shrinkage parameter that helps avoid over-fitting and keep $\mathbf{w}$ solvable. Then, under the same model assumption in (3.3), solution of $\mathbf{w}$ subjecting to (3.1) and (3.5) is given as

$$\mathbf{w}=(\lambda\mathbf{I}+\tilde{\mathbf{X}}^\top\tilde{\mathbf{X}})^{-1}\tilde{\mathbf{X}}^\top\mathbf{y}, \tag{3.6}$$

where $\mathbf{I}$ is the identity matrix with same dimension as $\tilde{\mathbf{X}}^\top\tilde{\mathbf{X}}$. Comparing (3.6) to (3.4), it is obvious that the modification to the L2 loss in (3.5) does not change the order of growth in computational complexity when solving for $\mathbf{w}$.

Next, note the following equality of matrix inverse

$$(\mathbf{D}^{-1}+\mathbf{B}^\top\mathbf{R}^{-1}\mathbf{B})^{-1}\mathbf{B}^\top\mathbf{R}^{-1}=\mathbf{D}\mathbf{B}^\top(\mathbf{B}\mathbf{D}\mathbf{B}^\top+\mathbf{R})^{-1}, \tag{3.7}$$

where $\mathbf{D}$, $\mathbf{R}$ and $\mathbf{B}$ are any matrices with proper dimension and invertibility. The equality can be verified by multiplying $(\mathbf{D}^{-1}+\mathbf{B}^\top\mathbf{R}^{-1}\mathbf{B})$ to the left side and multiplying $(\mathbf{B}\mathbf{D}\mathbf{B}^\top+\mathbf{R})$ to the right side as

$$\begin{aligned}
&(\mathbf{D}^{-1}+\mathbf{B}^\top\mathbf{R}^{-1}\mathbf{B})\times(\mathbf{D}^{-1}+\mathbf{B}^\top\mathbf{R}^{-1}\mathbf{B})^{-1}\mathbf{B}^\top\mathbf{R}^{-1}\times(\mathbf{B}\mathbf{D}\mathbf{B}^\top+\mathbf{R})\\
&=(\mathbf{D}^{-1}+\mathbf{B}^\top\mathbf{R}^{-1}\mathbf{B})\times\mathbf{D}\mathbf{B}^\top(\mathbf{B}\mathbf{D}\mathbf{B}^\top+\mathbf{R})^{-1}\times(\mathbf{B}\mathbf{D}\mathbf{B}^\top+\mathbf{R}),\\
&\Rightarrow\mathbf{B}^\top\mathbf{R}^{-1}\times(\mathbf{B}\mathbf{D}\mathbf{B}^\top+\mathbf{R})=(\mathbf{D}^{-1}+\mathbf{B}^\top\mathbf{R}^{-1}\mathbf{B})\times\mathbf{D}\mathbf{B}^\top,\\
&\Rightarrow\mathbf{B}^\top\mathbf{R}^{-1}\mathbf{B}\mathbf{D}\mathbf{B}^\top+\mathbf{B}^\top=\mathbf{B}^\top+\mathbf{B}^\top\mathbf{R}^{-1}\mathbf{B}\mathbf{D}\mathbf{B}^\top.
\end{aligned} \tag{3.8}$$

By letting $\mathbf{D}=\mathbf{I}$, $\mathbf{R}=\lambda\mathbf{I}$, and $\mathbf{B}=\tilde{\mathbf{X}}$, we have

$$(\lambda\mathbf{I}+\tilde{\mathbf{X}}^\top\tilde{\mathbf{X}})^{-1}\tilde{\mathbf{X}}^\top=\tilde{\mathbf{X}}^\top(\lambda\mathbf{I}+\tilde{\mathbf{X}}\tilde{\mathbf{X}}^\top)^{-1}, \tag{3.9}$$

and the solution of $\mathbf{w}$ in (3.6) can thus be rewritten as

$$\mathbf{w} = \tilde{\mathbf{X}}^\top(\lambda\mathbf{I} + \tilde{\mathbf{X}}\tilde{\mathbf{X}}^\top)^{-1}\mathbf{y}, \tag{3.10}$$

which is referred to as the dual form solution of $\mathbf{w}$. The key idea of these operations is to change the $\tilde{\mathbf{X}}^\top\tilde{\mathbf{X}}$ term inside the inverse operator to $\tilde{\mathbf{X}}\tilde{\mathbf{X}}^\top$. While the time complexity of calculating $(\tilde{\mathbf{X}}^\top\tilde{\mathbf{X}})^{-1}$ is approximately $O(P^3) + O(P^2N)$, calculating $(\tilde{\mathbf{X}}\tilde{\mathbf{X}}^\top)^{-1}$ takes $O(N^3) + O(N^2P)$, which grows much less with increment of P. However, the complexity can be further optimized.

The kernel method: Now assume that we map $\mathbf{x} \in \mathbb{R}^P$ to $\phi(\mathbf{x}) \in \mathbb{R}^Q$, and solve for $\mathbf{w}$ using (3.10). When calculating $\tilde{\mathbf{X}}\tilde{\mathbf{X}}^\top$ one gets

$$\begin{aligned}
\tilde{\mathbf{X}}\tilde{\mathbf{X}}^\top &= \underset{N\times(Q+1)}{\begin{bmatrix} \phi(\mathbf{x}_1)^\top & 1 \\ \phi(\mathbf{x}_2)^\top & 1 \\ \vdots & \vdots \\ \phi(\mathbf{x}_N)^\top & 1 \end{bmatrix}} \times \underset{(Q+1)\times N}{\begin{bmatrix} \phi(\mathbf{x}_1) & \phi(\mathbf{x}_2) & \dots & \phi(\mathbf{x}_N) \\ 1 & 1 & \dots & 1 \end{bmatrix}} \\
&= \underset{N\times N}{\begin{bmatrix} \phi(\mathbf{x}_1)^\top\phi(\mathbf{x}_1)+1 & \phi(\mathbf{x}_1)^\top\phi(\mathbf{x}_2)+1 & \dots & \phi(\mathbf{x}_1)^\top\phi(\mathbf{x}_N)+1 \\ \phi(\mathbf{x}_2)^\top\phi(\mathbf{x}_1)+1 & \phi(\mathbf{x}_2)^\top\phi(\mathbf{x}_2)+1 & \dots & \phi(\mathbf{x}_2)^\top\phi(\mathbf{x}_N)+1 \\ \vdots & \vdots & \ddots & \vdots \\ \phi(\mathbf{x}_N)^\top\phi(\mathbf{x}_1)+1 & \phi(\mathbf{x}_N)^\top\phi(\mathbf{x}_2)+1 & & \phi(\mathbf{x}_N)^\top\phi(\mathbf{x}_N)+1 \end{bmatrix}},
\end{aligned} \tag{3.11}$$

for which the computational complexity is $O(N^2Q)$ due to the dimension increase.

The growth of complexity is addressed with the kernel method. Calculation of inner product after the mapping: $< \phi(\mathbf{x}_i), \phi(\mathbf{x}_j) >= \phi(\mathbf{x}_i)^\top\phi(\mathbf{x}_j)$, takes a complexity of $O(Q)$. Instead of defining the mapping function $\phi(\mathbf{x})$, the kernel method directly defines the result of this inner product as a function given as

$$k(\mathbf{x}_i, \mathbf{x}_j) =< \phi(\mathbf{x}_i), \phi(\mathbf{x}_j) >, \tag{3.12}$$

which gives the results after mapping from the original feature vectors $\mathbf{x}_i$ without actually mapping it to $\phi(\mathbf{x}_i)$, and obtains a close-to-constant complexity. As many kernel-based ML methods use $Q >> P$, the kernel approach can significantly reduce the computational complexity.

SVR is the most commonly used kernel-based ML method for BP estimation [59, 60, 61].

Using the abovementioned kernel approach, SVR, KRR as well as other kernel-based regression methods introduce various nonlinear relationships to the models by selecting different kernel functions, while maintaining the computational efficiency.

Decision-based Methods: Regression Tree, Random Forest Regression (RFR) and Adaptive Boosting (AdaBoost)

Regression tree: Decision-based regression methods define F as a series of decision rules f on the input feature vector $\mathbf{x}$. Regression tree is the basic implementation of this idea, which is generated recursively by learning a series of tree-structured binary decision rules from the training set. Let x_j^i be the value on the j-st dimension of the P dimensional feature vector $\mathbf{x}_i$, and let $\{(\mathbf{x}_m, y_m)\}_{C_k}$ be a subset of the training set $\{(\mathbf{x}_i, y_i)\}_{i=1}^N$ that belongs to the class $C_k (k \geq 0)$. A binary decision rule $f_{(j_k,t_k)}(\mathbf{x})$ inspects the value on the j_k-st dimension of vector $\mathbf{x}$ with threshold t_k, and splits $\{(\mathbf{x}_m, y_m)\}_{C_k}$ into 2 sub-classes, $\{(\mathbf{x}_p, y_p)\}_{C_{2k+1}}$ and $\{(\mathbf{x}_q, y_q)\}_{C_{2k+2}}$, as

$$\begin{aligned} f_{(j_k,t_k)}(\{(\mathbf{x}_m, y_m)\}_{C_k}) &= [\{(\mathbf{x}_p, y_p)\}_{C_{2k+1}}, \{(\mathbf{x}_q, y_q)\}_{C_{2k+1}}], \\ \{(\mathbf{x}_p, y_p)\}_{C_{2k+1}} &= \{(\mathbf{x}_p, y_p) \in \{(\mathbf{x}_m, y_m)\}_{C_k} |\, x_{j_k}^p \leq t_k\}, \\ \{(\mathbf{x}_q, y_q)\}_{C_{2k+2}} &= \{(\mathbf{x}_q, y_q) \in \{(\mathbf{x}_m, y_m)\}_{C_k} |\, x_{j_k}^q > t_k\}. \end{aligned} \tag{3.13}$$

The decision rule (i.e. the selection of j_k and t_k) is optimal when the loss between the reference values y_m and their estimations $\hat{y}_m$ in class C_k is minimized as (3.1) describes, i.e.

$$(j_k, t_k) = \underset{j_k, t_k}{\operatorname{argmin}}\, L(\{(y_m, \hat{y}_m)\}_{C_k}). \tag{3.14}$$

The selection of j_k and t_k will affect the loss L, since $\hat{y}_m$ is given according to the sub-class that y_m belongs after the split as

$$\hat{y}_m = \begin{cases} \text{mean}(\{y_p | (\mathbf{x}_p, y_p) \in \{(\mathbf{x}_p, y_p)\}_{C_{2k+1}}\}) & \text{, if } (\mathbf{x}_m, y_m) \in \{(\mathbf{x}_p, y_p)\}_{C_{2k+1}} \\ \text{mean}(\{y_q | (\mathbf{x}_q, y_q) \in \{(\mathbf{x}_q, y_q)\}_{C_{2k+2}}\}) & \text{, if } (\mathbf{x}_m, y_m) \in \{(\mathbf{x}_q, y_q)\}_{C_{2k+2}} \end{cases}. \tag{3.15}$$

Without a closed-from solution, j_k and t_k are found by trying all possible values [62].

To form a regression tree from a training set, the recursion start from a root class C_0 : $\{(\mathbf{x}_m, y_m)\}_{C_0} = \{(\mathbf{x}_i, y_i)\}_{i=1}^N$, which is the training set. Following the aforementioned rules of splitting an existing class into 2 smaller sub-classes, the loss between the estimations $\hat{y}$ and the

references y on the entire training set is reduced in a greedy way. By repeating the split on the classes that have not been splitted previously, a binary tree with each node corresponding to a class can be formed, which becomes the regression tree. When a new input $\mathbf{x}$ is presented to the regression tree $F_{RT}(\mathbf{x})$, it recursively classifies $\mathbf{x}$ to one of its non-overlapping classes C_{Leave} at the leave nodes of the tree, and the output $\hat{y}$ is given as

$$\hat{y} = F_{RT}(\mathbf{x}) = \text{mean}(\{y_L | (\mathbf{x}_L, y_L) \in \{(\mathbf{x}_L, y_L)\}_{C_{Leave}}\}). \tag{3.16}$$

The working principle of the regression tree is very similar to the way that human makes a final decision $\hat{y}$ by inspecting and judging the circumstances of each factors x_j given in the new task $\mathbf{x}$, classifying the new task into a type C_{Leave}, and then come up with a solution $\hat{y}$ by referencing and averaging his previous experiences y_L for previous tasks $\mathbf{x}_L$ that belongs to the same type. As such, regression tree provides remarkable interpretability compared to other nonlinear methods.

Bagging and boosting: The accuracy of the regression tree method is often unremarkable in terms of large bias, because its output is restricted to the finite number of outputs that C_{Leave} in the regression tree can produce. Moreover, the method suffers from undesirable high variance being sensitive to the presented training set, which means that models trained from different subsets of the same training set will have very different decision rules [62], and consequently, very different outputs when the same input is presented. Therefore, bagging and boosting methods are often applied to improve the performance of the regression tree, with their specific implementations becoming the widely-applied RFR and AdaBoost methods.

- **Bagging and RFR**: Bagging is a method that trains multiple models from multiple training subsets in parallel, with all subsets generated from the original training set by bootstrapping. The target of this idea is to lower the estimation variance by averaging over multiple estimations from models with variety.

  Bootstrapping is an approach that generates training subsets by randomly sampling indices $\{(\mathbf{x}_k, y_k)\}_q$ from the original training set $\{(\mathbf{x}_i, y_i)\}_{i=1}^N$ with replacement, and adding it to the subset. The process is repeated for $K \times Q$ times to generate Q training subsets with K indices in each subset, allowing for duplication. Application of bagging on regression tree uses these training subsets to obtain an ensemble of Q regression trees,

$\{F_q(\mathbf{x})\}_{q=1}^{Q}$, which is referred to as bagged regression trees or a forest. When a new input is presented to the forest, each regression tree in the forest individually produces an output from the same input, and the output of the forest is given as

$$F_{Bagged,Q}(\mathbf{x}) = \frac{1}{Q}\sum_{q=1}^{Q} F_q(\mathbf{x}). \tag{3.17}$$

The RFR method is a further optimized implementation of bagged regression trees. It restricts the decision dimension j_k selected at each split to a randomly selected subset of all P available features when generating each tree. Such restriction forces most trees in the ensemble $\{F_q(\mathbf{x})\}_{q=1}^{Q}$ to be uncorrelated, which further reduces the variance of estimations that the forest produces [63].

- **Boosting and AdaBoost**: Boosting is a method that sequentially generates multiple models, with the training set renewed after each iteration. The target is to lower the estimation bias after every iteration by using the model trained in the previous iteration to guide the training of model in the next iteration.

  One basic method to achieve this target is to train the new model with the error between y_i and $\hat{y}_i$, where $\hat{y}_i$ are estimated with models formed in previous iterations, which is referred to as the boosted regression trees [62]. When the boosted regression trees are trained, an ensemble of regression trees is generated iteratively with renewed regression targets. After B iterations, the ensemble will have B trained regression trees $\{F_b(\mathbf{x})\}_{b=1}^{B}$, with the output from the ensemble given as

$$F_{Boosted,B}(\mathbf{x}) = \begin{cases} \sum_{b=1}^{B} \lambda F_b(\mathbf{x}) & \text{, if } B > 0 \\ 0 & \text{, if } B = 0 \end{cases}, \tag{3.18}$$

  with λ being the shrinking parameter affecting the rate that the ensemble improves. In the next iteration, the training target of $F_{B+1}(\mathbf{x})$ is the residual of current ensemble given as

$$F_{B+1}(\mathbf{x}) = \text{argmin}\, L(\{(F_{B+1}(\mathbf{x}_i), y_i - F_{Boosted,B}(\mathbf{x}_i))\}_{i=1}^{N}). \tag{3.19}$$

  The iteration stops if certain criteria is met.

  Another method is known as the AdaBoost [64] algorithm, which trains the new model in the successive iteration with training indices that have poor estimation accuracy when

estimated using the model formed in previous iterations. Bootstrap sampling with non-uniform probability distribution is applied in this method to generate different training subset at each iteration. In the b-st iteration, the training subset $\{(\mathbf{x}_k, y_k)\}_b$ is sampled from $\{(\mathbf{x}_i, y_i)\}_{i=1}^N$, with probability of sampling each index given as

$$\mathbf{p}_b = \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_N \end{bmatrix}. \tag{3.20}$$

After training, the obtained model $F_b(\mathbf{x})$ is validated on the whole training set $\{(\mathbf{x}_i, y_i)\}_{i=1}^N$, and the probability of selecting each index is updated to $\mathbf{p}_{b+1}$ with respect to the loss between y_i and $\hat{y}_i$, for $i = 1, 2...N$. If the loss between y_n and $\hat{y}_n$ is larger than others, then p_n will be raised so that $(\mathbf{x}_n, y_n)$ appears more frequently in the training subset in successive iterations. The loss also affects the confidence of model obtained in each iteration, with models achieving less loss assigned to higher confidence (i.e. weight). Finally, the output of the ensemble $\{F_b(\mathbf{x})\}_{b=1}^B$ after B iterations is considered over all models with respect to their confidences $\mathbf{c}$, with different forms in specific implementations, such as

$$F_{AdaBoost,B}(\mathbf{x}) = \begin{cases} \text{median}(\{c_b F_b(\mathbf{x})\}_{b=1}^B) & \text{, for AdaBoost.R2 [65]} \\ \sum_{b=1}^B c_b F_b(\mathbf{x}) & \text{, for AdaBoost.RT [66]} \end{cases}. \tag{3.21}$$

Neuron-based Method: Artificial Neuron Network (ANN)

Neuron-based methods imitate the functions of the biological neuron network, which forms the nervous system of creatures. The nonlinear behavior of biological neuron networks that creates human intelligence is formed with an essential activation mechanism: Each neuron aggregates the total amount of neurotransmitter it received from other neurons connected to it, and is either activated and delivers neurotransmitter to the neurons it connected to when the amount of received neurotransmitter is over certain threshold, or has no output at all. With its activation controlled by a threshold rule (i.e. a step function), this type of neuron is referred to as preceptron. With the development of other activation functions, this term now also indistinctively represent artificial neuron with any activation function.

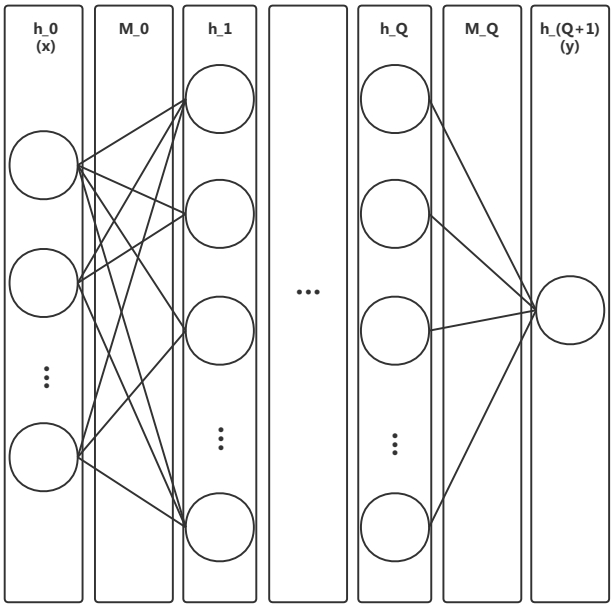

Figure 3.1: Structure of a MLP with Q hidden layers. $\mathbf{h}_q$ is the vector representing neurons in layer q, while $\mathbf{M}_q$ (weight matrices) and $\mathbf{b}_q$ (bias vectors, not shown in the figure) define the connections between consecutive layers.

The most basic implementation of the ANN is the multi-layer perceptron (MLP). The MLP setup of the regression problem described in (3.1) for this problem is $F_{MLP}(\mathbf{x})$, whose structure is shown in Figure 3.1. In MLP, each layer with k neurons corresponds to a k dimensional vector, including the input layer $\mathbf{h}_0$ with P neurons corresponding to the input vector $\mathbf{x}$, the output layer $\mathbf{h}_{Q+1}$ with 1 neuron corresponding to the output y, and Q hidden layers $\mathbf{h}_1, \mathbf{h}_2 ... \mathbf{h}_Q$ with number of neurons $k_1, k_2 ... k_Q$ determined by design. Connections between every 2 adjacent layers are determined with matrices $\mathbf{M}_0, \mathbf{M}_1 ... \mathbf{M}_Q$ as

$$\underset{k_{q+1}\times 1}{\mathbf{h}_{q+1}} = \Sigma_q\big(\underset{k_{q+1}\times k_q}{\mathbf{M}_q} \times \underset{k_q\times 1}{\mathbf{h}_q} + \underset{k_{q+1}\times 1}{\mathbf{b}_q}\big), \tag{3.22}$$

where Σ_q is the activation function of each layer defined as

$$\Sigma_q\big(\underset{k\times 1}{\mathbf{v}}\big) = \begin{bmatrix} \sigma_q(v_1) \\ \sigma_q(v_2) \\ \vdots \\ \sigma_q(v_k) \end{bmatrix}. \tag{3.23}$$

(3.22) and (3.23) demonstrate the following characteristics of artificial neuron networks, which imitate the behavior of biological neuron networks:

- **Aggregation**: The connection between any 2 neurons is weighted. The input of each neuron is the weighted sum of the outputs from the neurons in the previous layer that are connected to it, plus a bias. All weights and bias connecting neurons in layer q and $q+1$ are given as $\mathbf{M}_q$ and $\mathbf{b}_q$.

- **Activation**: The output of each neuron is a scalar function $\sigma_q(v)$ of its input, which is named as the activation function. Neurons in the same layer share the same activation function, as is shown in (3.23). The activation function of each layer is chosen from design, which is nonlinear in most cases.

- **Forward propagation**: From the first layer formed by the input vector, the outputs from neurons in a layer propagate to the next layer as inputs of its neurons. Such propagation repeats until the final layer is reached, whose output is the output of the network. The process of calculating output from input of ANN is referred to as forward propagation.

With the structure of $F_{MLP}(\mathbf{x})$ determined from design, minimizing the loss described in (3.2) is equivalent to optimizing $\{\mathbf{M}_q, \mathbf{b}_q\}_{q=0}^{Q}$. Without an analytical solution [67], the values are randomly initialized, and are determined by optimization algorithms with gradient information obtained from backward propagation.

The effectiveness of the MLP as a ML algorithm has been shown in [68], which proved that MLP is capable of approximating any continuous function with a finite size. MLP's approximation to the function gets better as the number of hidden neurons increases [69]. To increase the number of hidden neurons, an option is to include more hidden layers. ANNs containing MLP with $Q \geq 2$ are often referred to as deep neuron networks (DNNs). For cuff-less BP estimation, most studies use MLP with $Q = 1$ [70, 71, 72, 73], while applications of DNN are still preliminary [74, 75, 76].

3.2.2 Feature Selection Methods

While many model-driven methods use only the PTT, data-driven methods use much more features. A very recent study from Lin et al. (2021) [77] utilized 65 PPG features, 35 newly proposed and 30 summarized from previous studies, for BP estimation using MLR method. Their results showed that usage of all 65 features is achieving better accuracy than comparisons

using a subset of 14 or 6 features, which raises one question: does utilizing more features in the algorithm always imply better BP estimation accuracy?

Previous studies suggest a negative answer to this question, which is summarized as 'the curse of dimensionality'. It has been shown theoretically that the increment of feature vector dimension first improves the estimation accuracy of a Bayesian learning algorithm whose parameters are determined from a training set of finite size, then at certain point the accuracy reaches the maxima and start to consistently decrease [78, 79]. The degradation of accuracy is caused by data sparsity, which is explained as follow. Assume that a training set $\{(\mathbf{x}, y)\}$ of size N is sampled from the population, and assume that each element x_p in the feature vector $\mathbf{x}$ has M possible values in the population. If additional features are added to extend the dimension of feature vector from P to Q, then the quantity of possible $\mathbf{x}$ in the population grows exponentially from M^P to M^Q, and consequently the training set with fixed size N lose its generality to represent the majority of $\mathbf{x}$ in the population, which in the end manifests as under-fitting or over-fitting when the ML model is tested on new samples from the population. Although the data sparsity problem can be addressed by exponentially increasing N along the growth of P [80], the computation power required for keeping the running time of the model training process acceptable grows dramatically, which is often impracticable in real-life scenarios.

Therefore, with restricted computational budget and dataset size, a challenge in data-driven studies is to select a limited amount of features in the feature pool that gives optimal BP estimation with best accuracy. Feature selection methods address this problem by providing references to judge if a feature is valid for improving the estimation, or is redundant, irrelevant or ineffective. These methods can be categorized into the following classes:

Empirical Feature Selection

A very straightforward method for feature selection is to empirically select a subset of features from all available features, extract the features from the waveform provided in the dataset, train the model with selected features, and compare the testing performances with other selections. The empirical selections of features can come from prior experiences, literature review, or trial-and-error. If no feature selection method is mentioned in the study, then it is very likely that the features are empirically selected.

Linear Correlation-based Feature Selection

Similar to the cases of linear models between PTT/PWV and BP discussed in Chapter 2, linear relationship is often a good practice for considering a regression problem. For feature selection, linear relationships are considered as the following principles:

- A feature is less likely to be inefficient or irrelevant if it can accurately estimates BP via a linear regression (LR) model.
- A feature is redundant if it can be accurately estimated from other features that have been added to the feature vector with a MLR model.

As such, the following linear correlation matrices are considered:

- **Pearson's correlation coefficient (R)**: R can be used to evaluate the feasibility of estimating BP with a feature by LR model. For explanation, we derive R from the coefficient of determination (R-squared).

  Let $\{(x_i, y_i)\}_{i=1}^{N}$ represent a training set, with x being the feature and y indicating the reference BP value, and let $\hat{y}_i$ indicate the BP value estimated from x_i. The R-squared metric compares the squared error between the reference values $\{y_i\}_{i=1}^{N}$ and the estimated values $\{\hat{y}_i\}_{i=1}^{N}$ from a given model $F(x)$, to the squared error between the reference values and the optimal estimation of y under the assumption that x and y are independent, which is the sample mean given as $F_{NULL}(x) = \bar{y} = \text{mean}(\{y_i\}_{i=1}^{N})$. As such, R-squared is defined as

  $$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2} \leq 1. \tag{3.24}$$

  The R-squared metric gives an insight of the performance of $F(x)$ in the training set, where a high value close to 1 indicates good accuracy.

  Then, let $F(x)$ be the LR model given as

  $$\hat{y} = F(x) = a_1 x + a_2, \tag{3.25}$$

  where a_1 and a_2 are the regression parameters. To evaluate the best-case feasibility of the LR model between the feature and the BP, we maximize the R-squared metric of

$F(x)$ over the training set by solving for $F(x)$ that yields minimal squared error, i.e., minimizing the L2 loss. As such, the solutions to a_1 and a_2 are given as

$$\begin{aligned} a_1 &= \frac{\sum_{i=1}^{N}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{N}(x_i-\bar{x})^2}, \\ a_2 &= \bar{y}-a_1\bar{x}, \end{aligned} \tag{3.26}$$

where $\bar{x} = \text{mean}(\{x_i\}_{i=1}^{N})$ and $\bar{y} = \text{mean}(\{y_i\}_{i=1}^{N})$. (3.25) and (3.26) indicate the following equalities

$$\begin{aligned} \hat{y}_i - \bar{y} &= (a_1 x_i + a_2) - (a_1\bar{x} + a_2) \\ &= a_1(x_i - \bar{x}), \end{aligned} \tag{3.27}$$

$$\begin{aligned} \sum_{i=1}^{N}(y_i-\hat{y}_i)(\hat{y}_i-\bar{y}) &= a_1\sum_{i=1}^{N}(y_i-\hat{y}_i)(x_i-\bar{x}) \\ &= a_1\sum_{i=1}^{N}[(y_i-\bar{y})-(\hat{y}_i-\bar{y})](x_i-\bar{x}) \\ &= a_1\sum_{i=1}^{N}[(y_i-\bar{y})-a_1(x_i-\bar{x})](x_i-\bar{x}) \\ &= a_1\sum_{i=1}^{N}(y_i-\bar{y})(x_i-\bar{x}) - a_1^2\sum_{i=1}^{N}(x_i-\bar{x})^2 \\ &= 0, \end{aligned} \tag{3.28}$$

(3.27), (3.28) and (3.24) yields

$$\begin{aligned} R^2 &= 1 - \frac{\sum_{i=1}^{N}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{N}(y_i-\bar{y})^2} \\ &= \frac{\sum_{i=1}^{N}(y_i-\bar{y})^2 - \sum_{i=1}^{N}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{N}(y_i-\bar{y})^2} \\ &= \frac{\sum_{i=1}^{N}[(y_i-\hat{y}_i)+(\hat{y}_i-\bar{y})]^2 - \sum_{i=1}^{N}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{N}(y_i-\bar{y})^2} \\ &= \frac{\sum_{i=1}^{N}(\hat{y}_i-\bar{y})^2 + 2\sum_{i=1}^{N}(y_i-\hat{y}_i)(\hat{y}_i-\bar{y})}{\sum_{i=1}^{N}(y_i-\bar{y})^2} \\ &= \frac{\sum_{i=1}^{N}(\hat{y}_i-\bar{y})^2}{\sum_{i=1}^{N}(y_i-\bar{y})^2} \\ &= \frac{a_1^2\sum_{i=1}^{N}(x_i-\bar{x})^2}{\sum_{i=1}^{N}(y_i-\bar{y})^2} \\ &= \frac{[\sum_{i=1}^{N}(x_i-\bar{x})(y_i-\bar{y})]^2}{\sum_{i=1}^{N}(x_i-\bar{x})^2\sum_{i=1}^{N}(y_i-\bar{y})^2} \\ &= \frac{\text{Cov}(x,y)^2}{\text{Var}(x,x)^2\text{Var}(y,y)^2}, \end{aligned} \tag{3.29}$$

where Cov and Var denotes covariance and variance. Finally, let $R = \pm\sqrt{R^2}$, with $+$ indicating positive correlation and $-$ indicating negative correlation, and the Pearson's correlation coefficient is given as

$$R = \frac{\text{Cov(x,y)}}{\text{Var}(x,x)\text{Var}(y,y)} \in [-1,1]. \tag{3.30}$$

The above derivation shows that the Pearson's correlation coefficient between x and y is a metric that instantly gives the feasibility of fitting x to y with a LR model, with its accuracy described by the coefficient of determination. If $|R|$ is close to 1, it means that the LR fitting is highly desirable as the model gives estimations that is almost identical to the references in the training set, which suggests that x and y are very likely to have linear correlation. A feature with high $|R|$ value has significant linear correlation with reference BP, and should be included in the feature vector.

- **Multiple correlation coefficient (R)**: The multiple correlation coefficient is just the Pearson's correlation coefficient evaluated between the estimations obtained from MLR model and the reference values, therefore having the same R notation. In feature selection, it is mostly used to judge if a feature can be accurately estimated with linear combination of other features. Let y be a feature, and let $\mathbf{x} \in \mathbb{R}^P$ be the P-dimensional feature vector formed by other features. In order to see if feature y can be estimated from other features, MLR is used to obtain a model $F(\mathbf{x})$ from the training set, with its solution obtained from (3.4). Afterward, Pearson's correlation coefficient between $\{\hat{y}_i = F(\mathbf{x}_i)\}_{i=1}^N$ and $\{y_i\}_{i=1}^N$ is calculated using (3.30). A feature having high R value with existed feature vector is said to have colinearity, and should be considered as redundant.

- **Variance inflation factor (VIF)**: VIF is another form of multiple correlation coefficient, which is given as:

$$VIF = \frac{1}{1-R^2}. \tag{3.31}$$

It is suggested that features with $VIF > 10$ should be discarded from the feature vector [81].

Mutual Information (MI)-based Feature Selection

MI is another method to determine if one random variable is correlated with another, which has its root in the information theory and is given as:

$$MI(x,y) = \begin{cases} \sum_{x,y} p(x,y) \ln \frac{p(x,y)}{p(x)p(y)}, \text{ discrete random varaibles} \\ \iint f(x,y) \ln \frac{f(x,y)}{f(x)f(y)} dxdy, \text{ continuous random varaibles} \end{cases}, \tag{3.32}$$

in which p and f are the probability mass functions and probability distribution functions. In actual applications where the probability distributions are not available, p or f are often estimated from the samples in the training set with various methods such as k^{th} nearest neighborhood [82]. Although further discussion is beyond the scope of this review, the method basically implies that if knowing the distribution of one variable is reducing the uncertainty of another variable [83], then it is more likely that those variables are correlated.

In the reviewed studies, MI based methods are often used to determine the strength of correlation between the feature and the BP. A basic application of MI is similar to the linear correlation-based feature selection, which ranks the available features with respect to their MI with BP and select features with high MI. However, in both cases if more than 1 features are showing strong correlation with BP, then it is likely that these features are also strongly correlated with each other, which indicates redundancy. While linear correlation-based methods check colinearity between selected features to avoid redundancy, MI-based methods have their own advanced implementations to deal with the redundancy problem, such as the maximum information coefficient (MIC) method [84] or the minimal-redundancy-maximal-relevance (mRMR) method [85].

3.3 Protocol of Literature Review

Some representative data-driven studies are discussed in the following sections. Summaries of studies presented in the tables in this chapter generally follows the same protocol as Chapter 2, except that the 'Feature Type' and the 'ML Method' columns replace the 'Proposed Model' column in Chapter 2 due to the difference in methodologies between model-driven and data-driven studies. In addition, we highlight another point that is exclusive in data-driven studies:

Estimation Type: PWV features, including PTT and PAT, can be extracted in every cardiac cycle. Since most model-driven methods use only PWV features to estimated BP, these methods are capable of producing BP estimations from signal recordings with duration as short as containing only 1 beat. Methods with such capability are named as beat-to-beat methods. However, when utilizing other features in data-driven methods, signal recordings with longer duration are sometimes required for extracting the features. Therefore, methods reviewed in this chapter are classified according to the shortest duration of raw signal required for making 1 set of BP estimation (i.e. simultaneously acquired SBP/DBP/MAP reading(s)), which belongs to one of the following categories:

- **Beat-to-beat**: This term implies that the method is capable of producing 1 set of BP estimation as long as the input signal contains 1 complete beat (cycle).
- **Window**: This term implies that signal recording with longer duration is required, either including multiple beats or achieving certain length. Specific requirement will be noted.

3.4 Methods Using Only PPG: A Review of Feature Studies

The pressure-volume relationship of human finger arteries has determined the correlation between PPG and BP in a theoretical way [86]. Elasticity of human artery leads to quasi-periodic changes in vessel diameter and blood volume as the arterial BP oscillates between SBP and DBP in each cardiac cycle, which is recorded by the PPG signal. An intuitive prove of relevance between the BP and the vessel volume is by volume clamping, which shows that if an external pressure being continuously equivalent to the internal BP is applied to the vessel via a finger cuff, then the blood volume monitored by the PPG sensor remains unchanged. The morphological and periodical similarity of PPG and arterial BP (ABP) waveforms depicted in Figure 3.2 is another concrete embodiment of their time and frequency domain correlations [87, 70]. Therefore, it is very desirable and promising to interpret the close relationship between BP and blood volume as mathematical models between characteristics of the PPG signals and the BP values.

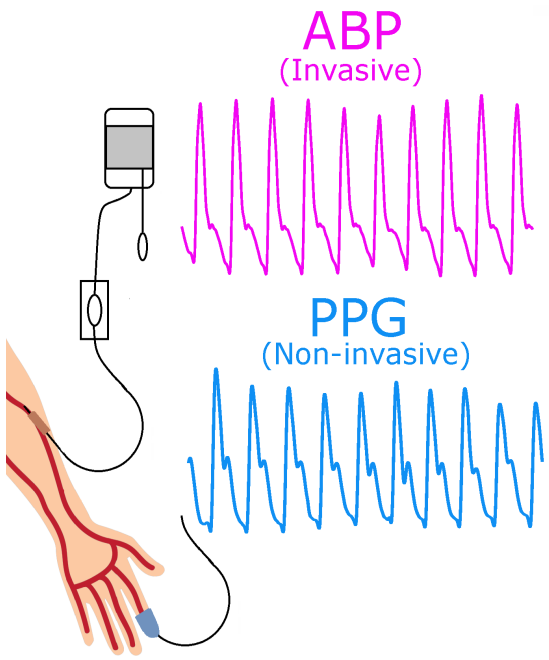

Figure 3.2: Similarity in the waveform morphology and cycle duration of the PPG and the ABP signal [87].

3.4.1 Temporal Domain Features Based on Waveform Morphology

Temporal domain features are the most commonly used PPG features in existing studies, which are extracted from the morphological characteristics of the PPG waveform. Temporal domain features are often defined as the relative positions of characteristics extracted from the PPG waveform, such as the ratio of amplitudes or the time difference between two points. Temporal domain features are explicitly defined, with each feature distinctively expressed as one numerical value, thereby having great interpretability. However, with some temporal domain PPG features defined only on the fingertip PPG waveforms with normal shape and quality, their application could be limited when migrating the features to other PPG signals, such as those measured from earlobe, toe, or animals.

The characteristics of temporal domain PPG features make one-by-one verification an ideal method for determining the BP predictability of each feature. Therefore, we divide the review into two parts. In **The Correlation Studies** part, we look into studies using only one or a few types of temporal domain features. These works could be limited to inferior accuracy, but are very helpful for validating the BP correlation of specific feature. In **The Comprehensive Studies** part, we review studies using inclusive and heterogeneous temporal domain features

with advantages in performance.

The Correlation Studies

Temporal domain PPG features can be classified into 4 categories: durations, amplitudes, areas, and statistical indices.

Durations: PPG duration features are often defined as the width of a PPG cycle at different percentage of normalized amplitude, with some commonly-used examples shown in Figure 3.3. To extract these features, the PPG cycle is firstly divided into the systolic phase and the diastolic phase with respect to the position of the systolic peak. Afterward, the systolic upstroke time (SUT), the diastolic time (DT), as well as the systolic widths (SW) and the diastolic widths (DW) at different percentage of amplitude are acquired. The total width (W) can be calculated from $SW + DW$. From Figure 3.3, one can seize the key idea of designing PPG duration features, that is, to normalize the amplitude of the PPG cycle and to represent its morphology by sampling its width at different height.

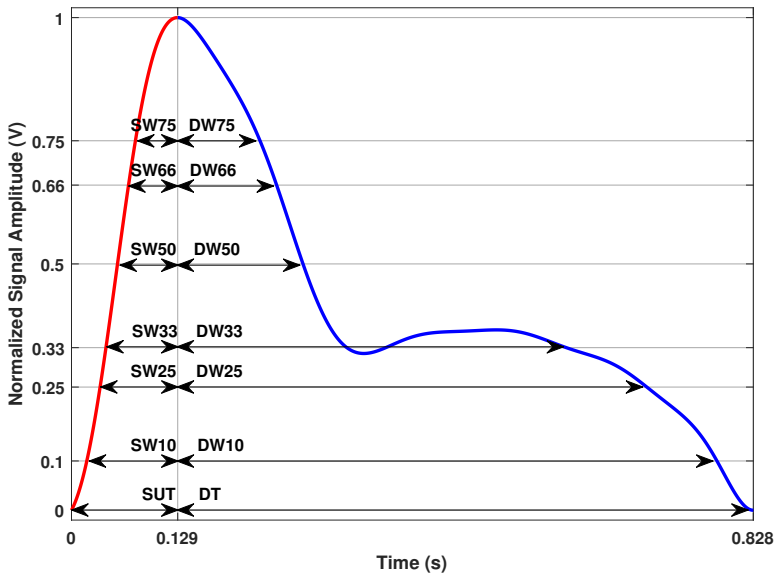

Figure 3.3: Definitions of some commonly-used PPG duration features.

The study from Awad et al. (2001) [88] is one of the earliest works that proposed the idea of estimating BP with features extracted from only PPG. In [88], only the width of PPG cycle at 50% of the amplitude (W50) was explored as the duration feature, which was compared against

Table 3.1: Summary of studies related to PPG duration features.

Year	Ref.	Dataset	Intervention	Feature Type	Estimation Type	ML Method	Training	Performance			Key Conclusion
								SBP	DBP	MAP	
2001	[88]	20 patients; arterial BP reference	None	PPG width	Beat-to-beat ($N = 41293$)	Linear regression	Grouped; 10 subjects training set	0.8 ± 0.1(R)	0.76 ± 0.1(R)	0.8 ± 0.1(R)	Width of earlobe PPG has high correlation with BP
					($N = 62077$)		10 subjects testing set	-8.96 ± 47.94(ME)	N/A	-5.04 ± 29.47(ME)	
2003	[89]	15 healthy subjects; cuff BP reference	Exercise	PPG Diastolic time	Beat-to-beat	Linear regression	Subject specific	0.811(R) 0.21 ± 7.32(ME)	0.690(R) 0.02 ± 4.39(ME)	N/A	Diastolic time has high correlation with BP
2013	[75]	MIMIC database; arterial BP reference	None	21 PPG duration features	Beat-to-beat ($N > 15000$)	MLP with 2 hidden layers	Grouped; 70% training, 15% validation, 15% testing	3.80 ± 3.46(MAE)	2.21 ± 2.09(MAE)	N/A	Diastolic time and BP has low correlation outside healthy subjects; improved accuracy with ANN
2018	[90]	32 subjects from University of Queensland database; non-invasive arterial BP reference	None	Area, SUT, W25	5 seconds window ($N = 8133$)	Multicollinearity-based feature selection; Regression tree	Grouped; 10 fold cross validation	-0.1 ± 6.5(ME)	-0.6 ± 5.2(ME)	N/A	Regression tree yields best performance; there exists colinearity within width features; inferior accuracy in hypotensive and hypertensive group
						Comparison: Support vector regression		-0.5 ± 15.3(ME)	0.05 ± 9.0(ME)	N/A	
						Comparison: Multiple linear regression		-3.9 ± 15.9(ME)	1.2 ± 8.5(ME)	N/A	

other features including the amplitude and the pulse area extracted from both fingertip and earlobe PPG. The results confirmed that W50 extracted from earlobe PPG has high correlation with SBP, DBP and MBP ($R \geq 0.76$), while all other features showed very low correlation ($R < 0.31$) in the grouped training set. However, trials of estimating BP from W50 with a linear model resulted in unacceptable large error in the grouped testing set.

Another study from Teng et al. (2003) [89] demonstrated the individual difference of model parameters similar to the cases in model-driven studies using PWV features, which utilized DT with same linear regression method as [88], but obtained much lower error from subject-specific parameter calibration. Early studies like [88, 89] has provided insights of the BP predictability of PPG features, while pointing out the limitations of simple regression methods with only a few features involved. It seems that including more features and increasing the model complexity is the key to develop universal models.

The idea of including more features and using advanced ML method was carried out in the study from Kurylyak et al. (2013) [75], which extended the PPG duration features to more widths in different vertical positions. Using all 14 features depicted in Figure 3.3 as well as their combinations and ratios, the work formed a 21-dimensional feature vector to train a MLP with 2 hidden layers, providing improved accuracy compared to [88]. The feature set with 21 PPG duration features proposed in [75] has been frequently cited and applied as part of the feature vector in the subsequent comprehensive studies.

While there is no feature selection process included in [75], it is natural to imagine that all the widths features depicted in Figure 3.3 increase synchronously if the duration of the PPG cycle increases, which implies potential colinearity in the PPG duration feature set. A work

from Khalid et al. (2018) [90] studied the colinearity of a 5-element feature set, including pulse area, SUT, W25, W50 and W75, by evaluating the VIF of each feature against the rest. Their results showed that W50 and W75 are redundant in this dataset, with $VIF > 10$, which implies that the feature set proposed in [75] can be further reduced and optimized. After removing W50 and W75, the 3-dimensional feature vector is applied to regression tree, SVR and MLR models for comparison of BP estimation accuracy. The optimal results from regression tree is close to the results reported in [89], while being free of subject-specific calibration.

Amplitudes and areas: Features extracted from amplitudes of the PPG signal or the areas under the waveform are often selected due to their indication of certain physiological parameters or processes, whose relationship with BP could be complicated and nonlinear. Therefore, unlike the duration features whose BP predictability can be validated via LR or MLR, the motivation of applying amplitude and area features could be more empirical than statistical.

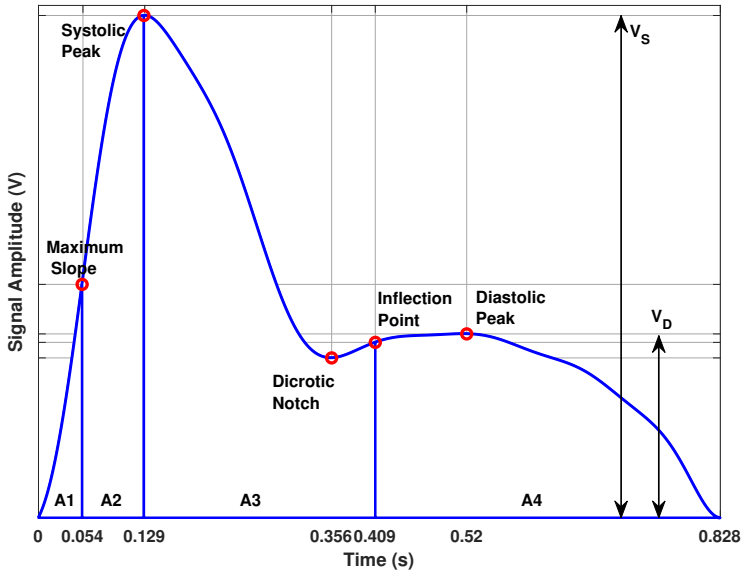

Figure 3.4: Definitions of some commonly-used PPG amplitude and area features.

As is shown in Figure 3.4, locating the position of characteristic points is the first step of extracting amplitude and area features. The characteristic points include the maximum slope, the systolic peak, the dicrotic notch, the inflection point, and the diastolic peak. The amplitudes of these points and their ratios form the basic amplitude features. Afterward, the area under the waveform is divided into 4 sections with respect to the positions of the maximum slope, the

systolic peak and the inflection point, which forms 4 area features A1∼A4.

Elgendi et al. (2012) [91] proposed a thorough review of PPG duration, amplitude and area features based on their physiological, vasculature and hemogynamic backgrounds. Some features reviewed in [91] have strong theoretical support to be correlated with specific biological properties, which could be particularly useful for BP estimation. These features are manually-designed indices calculated from the basic amplitudes and areas shown in Figure 3.4, which are summarized below:

- **Augmentation index (AI)** [92]: $AI = \frac{V_D}{V_S}$ is a measure of pulse wave reflection from peripheral location [93].
- **Inflection point area ratio (IPA)** [94]: $IPA = \frac{A4}{(A1+A2+A3)}$ is a measure of total peripheral resistance.
- **PPG intensity ratio (PIR)** [55]: $PIR = \frac{\max(VPPG)}{\min(VPPG)}$ has high correlation with DBP, which is verified not only by its original PWV-model based work, but also by later works comparing PIR with other features [95].

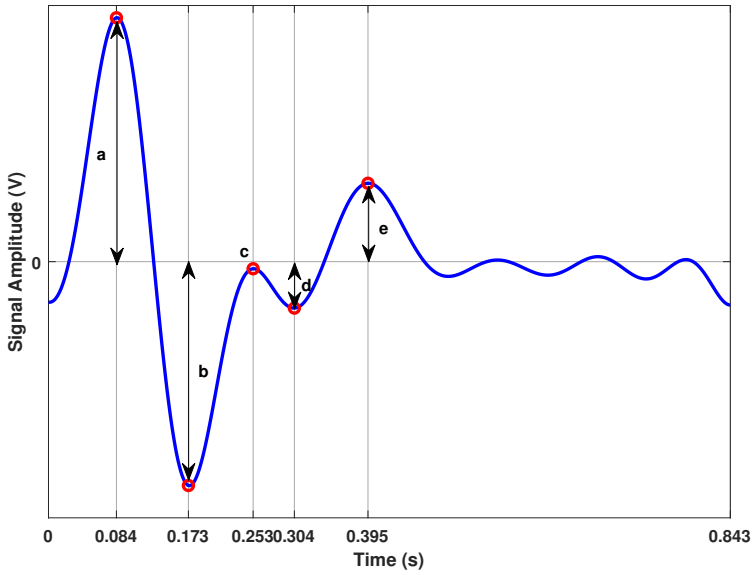

Figure 3.5: Definitions of amplitude features extracted from the second derivative of PPG signal.

Besides the original PPG signal, features extracted from the second derivative of PPG (SDPTG), which is also referred to as the acceleration PPG (APG), are also widely applied.

5 APG amplitude features are depicted in Figure 3.5 as a, b, c, d and e. Their ratios are related to arterial stiffness [92], which has close relationship with BP, as is discussed in Chapter 2.

Table 3.2: Summary of studies related to PPG amplitude and area features.

Year	Ref.	Dataset	Intervention	Feature Type	Estimation Type	ML Method	Training	Performance			Key Conclusion
								SBP	DBP	MAP	
2006	[96]	4 healthy subjects; Finapres BP reference	Posture change	Amplitude of PPG systolic peak	Beat-to-beat	Non-linear model	Subject specific; regression for 20 seconds signal	N/A	N/A	$N/A \pm 8.37$(ME)	Correlation between PPG and MBP is not consistent for larger time scale
2008	[97]	34 healthy, aged subjects; cuff BP reference	Rest	5 APG amplitude features	Beat-to-beat ($N = 155$)	Manual classification +regression	Grouped; regression	0.89(R) $N/A \pm 8.2$(ME)	N/A	N/A	APG features has great differences among age groups, which is better addressed through decision-based method
						Comparison: Regression only		0.67(R) $N/A \pm 13.7$(ME)	N/A	N/A	
2017	[98]	25 subjects; cuff BP reference	Rest	25 features: 15 APG (5 duration +10 amplitude), 4 subject information, heartrate	Beat-to-beat ($N = 122$)	MLR	Grouped; training with stand-alone dataset, testing on 25 subjects	0.8(R) 1.58 ± 8.54(ME)	N/A	N/A	APG features are feasiable for BP estimation

For correlation studies, one early work from Shaltis et al. (2006) [96] tried to estimate MAP from the amplitude of PPG systolic peak (V_s), with a non-linear model:

$$MAP = a_1 e^{a_2 V_s} + a_3 e^{a_4 V_s}. \tag{3.33}$$

While most duration features and some amplitude or area features expressed as ratios can be extracted from preprocessed PPG signals with normalized amplitude in each cycle, as is the case in Figure 3.3, direct application of PPG amplitudes in the model requires the relative signal magnitude among beats to be preserved, which could introduce additional difficulties in peak detection if there exists large PPG amplitude variation over time. Additionally, due to the differences in measurement hardware, tissue thickness and sensor placement, the amplitude of PPG signal is inevitably inconsistent across trials or subjects, which restricted the method to subject-specific calibration. Results from [96] confirmed the correlation between PPG amplitude and MAP in a short, 20 seconds time scale, while pointing out that the correlation is not consistence with prolonged timescales. The results thereby demonstrated the necessity of feature engineering when the model is relatively simple. However, as the model complexity increases in later works [99, 100], the raw amplitudes of the PPG signal is again considered as a reasonable input of ML algorithms.

The APG amplitude features have drawn special attention from the researchers. Suzuki et al. (2008) [97] and Atomi et al. (2017) [98] proposed their methods using only or mostly APG amplitude features. [97] directly utilized the 5 amplitudes (a$\sim$e), while [98] added their positions on the timeline, their ratios, as well as additional subject information including height,

weight, age and sex. [97] and [98] reported similar BP estimation accuracy ($R \geq 0.8$), which sufficiently validated the predictability of these features. [97] noted that the a∼e has large variation among the elder age groups, and suggested a method that combines classification and regression by first classifying the grouped dataset into 4 subgroups with manually designed decision rules on the value of a∼e, and then conduct subgroup-specific regression (unspecified regression method in publication). A better accuracy obtained from this method may imply advantages in decision-based methods when dealing with discontinuously changing feature vectors.

Statistical Indices: Statistical indices are evaluated over multiple samples of the PPG signal to reflect the properties of their value distribution. A review from Elgendi et al. (2016) [101] suggested that statistical indices evaluated over the amplitudes of PPG samples is indicative of the quality of PPG signal. With N samples of the PPG signal given as $\{x_i\}_{i=1}^{N}$, the following statistical indices can be evaluated according to [101]:

- **Mean (μ), Variance (σ)**
- **Skewness (S)**: $S = \frac{1}{N}\sum_{i=1}^{N}(\frac{x_i-\mu_x}{\sigma_x})^3$, usage on PPG suggested by [102].
- **Kurtosis (K)**: $K = \frac{1}{N}\sum_{i=1}^{N}(\frac{x_i-\mu_x}{\sigma_x})^4$, usage on PPG suggested by [103].
- **Entropy (E)**: $E = -\sum_{i=1}^{N} x_i^2 \ln(x_i^2)$, usage on PPG suggested by [103].
- **Zero crossing rate (Z)**: $Z = \frac{1}{N-1}\sum_{i=1}^{N-1} |\text{sgn}(x_{i+1}) - \text{sgn}(x_i)|$, usage on PPG suggested by [104].

Table 3.3: Summary of study related to PPG statistical features.

Year	Ref.	Dataset	Intervention	Feature Type	Estimation Type	ML Method	Training	Performance			Key Conclusion
								SBP	DBP	MAP	
2011	[59]	410 subjects; cuff BP reference	None	33 features: 30 PPG statistical indices + 3 subject information	1 min window ($N = 410$)	RFR	Grouped; 80% training, 20% testing	0.95(R)	0.93(R)	N/A	Proposed an effective BP estimation system using statistical features
						Comparison: SVR		0.85(R)	0.82(R)	N/A	
						Comparison: MLP		0.81(R)	0.79(R)	N/A	
						Comparison: MLR		0.77(R)	0.73(R)	N/A	

Besides the amplitudes of PPG samples, the abovementioned statistical indices can also be evaluated on many other parameters calculated from the PPG signal. A method using mostly statistical indices was proposed by Monte et al. (2011) [59]. Statistical indices are evaluated

over various parameters to form a feature vector for each 1-min PPG window. Utilization of some statistical features in this work are summarized as follows:

- First, 24 overlapping frames were separated from the 1-min PPG window.
- The Kaiser-Teager Energy (KTE) [105] was calculated for each of 24 frames in the window, and μ, σ, S and interquartile range (IQR) were evaluated over the KTE values in the window.
- Heart rate (HR) was calculated for each frame, and μ, σ, S and IQR were evaluated over all values from the frames in the window.
- Fast Fourier transform (FFT) of each frame wad calculated. E was evaluated over the FFT of each frame, and μ, σ, S and IQR were evaluated over the E values from all frames in the window.
- Log energy profile was evaluated for each frame, and σ and IQR were evaluated over all values from the frames in the window.

Results from [59] showed notable high correlation using the RFR model, which demonstrated the possibility to interpret PPG-extracted parameters in a pure statistical way. However, the method proposed in [59] requires 1-min PPG window to make 1 set of BP estimation, which is a major drawback. Moreover, the statistical features often lack interpretability compared to the morphological features, which can be observed directly from the waveform.

The Comprehensive Studies

In the study from Gaurav et al. (2016) [106], 46 PPG features were utilized, including 4 PPG duration and amplitude features, 19 APG duration and amplitude features, 4 statistical indices evaluated on PPG amplitude, 11 features based on statistical analysis of successive heart rate in the 10-beats window, and 8 non-linear combinations of the aforementioned features. The ML model in [106] is an ensemble of 6 deep MLPs, 3 estimating SBP and 3 for DBP, with each MLP having 4 hidden layers. The reference DBP values are used as an additional 47^{th} feature to train the 3 MLPs estimating SBP. Their results are state-of-the-art among studies [70, 76] using dataset with similar scale.

Table 3.4: Summary of comprehensive studies using only PPG to estimate BP.

Year	Ref.	Dataset	Intervention	Feature Type	Estimation Type	ML Method	Training	Performance			Key Conclusion
								SBP	DBP	MAP	
2016	[106]	3000 subjects from UCI machine learning repository; arterial BP reference	None	46 PPG temporal features (8 PPG, 19 APG, 11 heartrate, 8 ratios of features)	10 beats window ($N = 151487$)	Combinatorial ANN with 6 MLPs	Grouped; 80%training, 20% testing	0.16 ± 6.85(ME) 4.47(MAE)	0.03 ± 4.72(ME) 3.21(MAE)	N/A	Proposed a novel method that can potentially be integrated on smartphone
2016	[60]	32 subjects from University of Queensland database; non-invasive arterial BP reference	None	57 PPG temporal features	10 seconds window ($N = 7678$)	MI and colinearity-based feature selection; SVR	Grouped; 6143 training, 1535 testing	4.76 ± 7.52(MAE)	3.42 ± 5.19(MAE)	3.33 ± 5.41(MAE)	Proposed an effective feature selection methodology
				18 selected features				4.63 ± 7.43(MAE)	3.29 ± 5.09(MAE)	3.30 ± 5.25(MAE)	
				11 selected features for SBP; 11 selected features for DBP and MBP				4.77 ± 7.68(MAE)	3.67 ± 5.69(MAE)	3.85 ± 5.87(MAE)	
2017	[61]	MIMIC-II database; arterial BP reference	None	35 PPG temporal features: 21 PPG, 14 APG	Beat-to-beat ($N = 910$)	SVR	Grouped; 648 training, 262 testing	8.54(MAE) 10.9(RMSE)	4.34(MAE) 5.8(RMSE)	N/A	Improved accuracy with APG features
				Comparison: 21 PPG duration features [75]				13.6(MAE) 13.6(RMSE)	7.7(MAE) 7.9(RMSE)		
2019	[107]	942 subjects from UCI machine learning repository; arterial BP reference	None	6 types of PPG temporal features and their non-linear combinations	Beat-to-beat	AdaBoost	Grouped; 10-fold cross validation	0.78(R) 0.09 ± 10.38(ME) 8.22(MAE)	0.72(R) 0.23 ± 4.22(ME) 4.17(MAE)	0.75(R) -0.02 ± 5.53(ME) 4.58(MAE)	Improved accuracy with Adaboost; proposed ranking of feature importance
						Comparison: RFR		0.75(R) -0.17 ± 10.35(ME) 10.29(MAE)	0.69(R) -0.14 ± 5.43(ME) 5.77(MAE)	0.72(R) 0.07 ± 6.62(ME) 6.38(MAE)	
						Comparison: Regression tree		0.54(R) -0.71 ± 15.29(ME) 13.87(MAE)	0.44(R) 0.17 ± 8.91(ME) 6.82(MAE)	0.49(R) 0.02 ± 9.02(ME) 8.18(MAE)	
						Comparison: MLR		0.37(R) 0.17 ± 10.03(ME) 16.12(MAE)	0.35(R) -0.11 ± 5.81(ME) 7.04(MAE)	0.34(R) 0.13 ± 6.25(ME) 8.89(MAE)	
2021	[77]	109 subjects from MIMIC-II database; arterial BP reference	None	69 PPG temporal features	Beat-to-beat ($N > 729000$)	MI-based feature selection; MLR	Subject specific; 75% training, 25% testing	0.73(R) -0.00 ± 6.00(ME) 4.59(MAE)	0.73(R) 0.00 ± 3.30(ME) 2.47(MAE)	N/A	Proposed a comprehensive ranking of feature importance
				10 PPG features from feature selection				0.64(R) $N/A \pm 6.90$(ME) 5.39(MAE)	0.65(R) $N/A \pm 3.71$(ME) 2.83(MAE)	N/A	
				Comparison: 14 PPG+ECG features: 11 PPG features, 3 PAT features[108]				0.62(R) $N/A \pm 7.02$(ME) 5.50(MAE)	0.61(R) $N/A \pm 3.80$(ME) 2.90(MAE)	N/A	
				Comparison: 10 PPG+ECG features: 7 PPG features, 3 PAT features [23]				0.65(R) $N/A \pm 6.74$(ME) 5.25(MAE)	0.64(R) $N/A \pm 3.70$(ME) 2.82(MAE)	N/A	

Another study from Duan et al. (2016) [60] demonstrated how proper feature selection can improve the efficiency of model formation. Starting with a 57-element feature pool, [60] utilized a feature selection method considering the computational complexity, the maximum information coefficient (MIC) with BP, and the colinearity of selected features to reduce the amount of feature to 11, individually selected for SBP and DBP. The comparison between 3 feature sets with 57, 18 and 11 elements shows that it is possible to significantly cut down the dimension of feature vector while only negligibly affect the estimation accuracy.

Liu et al. (2017) [61] combined the PPG duration feature set proposed in [75] and the features extracted from APG to form a larger feature vector. Results showed major accuracy improvement ($MAE = 8.54$ mmHg for SBP, $MAE = 4.34$ mmHg for DBP) compared to using only the duration features ($MAE = 13.6$ mmHg for SBP, $MAE = 7.7$ mmHg for DBP), which demonstrated the novel BP predictability of APG features.

Study from Hasanzadeh et al. (2019) [107] demonstrated how bagging and boosting methods improve the performance of regression tree. [107] compared the results out of a RFR model with 100 regression trees ($R = 0.78$ for SBP) and an AdaBoost model with 200 trees ($R = 0.75$

for SBP) against the results from basic regression tree ($R = 0.54$ for SBP) and MLR ($R = 0.37$ for SBP), which shows clear improvements. [107] also proposed a ranking of feature importance for MAP estimation out of the AdaBoost model, the top 10 of which is listed as follow: (from most significant to less significant)

For MAP estimation:

1. $\ln(HR \times mNPV)$, HR: heart rate; $mNPV = \frac{\max(VPPG) - \min(VPPG)}{\max(VPPG)}$.
2. $\ln(DRI)$, DRI: Dicrotic reflection index, $DRI = \frac{V_{Dicrotic\ notch}}{V_S}$.
3. $\frac{\mathrm{LF}(HRV)}{\mathrm{HF}(HRV)}$, the ratio of low frequency components and high frequency components of heart rate variation.
4. $\ln(IRI)$, IRI: Inflection point reflection index, $IRI = \frac{V_{Inflection\ point}}{V_S}$.
5. IPA: Inflection point area ratio.
6. $LASI$: Large artery stiffness index, the ratio of subject's height and the time duration between systolic peak and inflection point.
7. $A1 + A2 + A3 + A4$, the total area under the PPG pulse.
8. $\ln(HR)$.
9. μ_{HRV}.
10. SUT

Lin et al. (2021) [77] proposed a very inclusive study of 69 temporal domain PPG features. However, probably due to the utilization of a simple MLR method, [77] requires subject-specific model calibration. The feature selection method proposed in [77] considered the mutual information between the features and the BP for each subject, and the stability of the feature-BP correlation among all subjects. The feature ranking proposed in [77] is summarized as follows: (from most significant to less significant)

For SBP estimation:

1. Width at $\frac{10}{11}$ of amplitude.

2. *dPIR*, PIR of the 1th derivative of PPG.
3. *sdPIR*, PIR of APG.
4. *sdRIPV*,RIPV of APG, $RIPV = \frac{\max(VPPG)}{\min(VPPG_{Valy})}$.
5. Width at $\frac{9}{11}$ of amplitude.
6. Slope from onset to maximum peak of PPG.
7. $A4$.
8. Width at $\frac{8}{11}$ of amplitude.
9. $A3 + A4$.
10. $A3$.

For DBP estimation:

1. Width at $\frac{10}{11}$ of amplitude.
2. *dPIR*.
3. *sdPIR*.
4. *LASI*.
5. Width at $\frac{2}{11}$ of amplitude.
6. Width between systolic peak and end of cycle.
7. *sdRIPV*.
8. $A4$.
9. Width at $\frac{3}{11}$ of amplitude.
10. Width between maximum peak and end of cycle of the 1th derivative of PPG.

By comparing the results using only PPG features and the results using both PPG and PAT features, [77] concluded that PPG features are sufficient to substitute the usage of PWV features. However, the conclusion was not validated using a universal model in [77].

3.4.2 Spectral Domain Features Based on Waveform Spectrum

Temporal domain features are explicitly defined, with each feature distinctively expressed as a numerical value. Spectral domain features, on the other hand, can be described as semi-defined: while the feature is defined as the frequency domain information of the PPG signal, the feature is fed into the ML algorithm as a series of values. For example, spectral domain features are often presented to the ML models by using all FFT coefficients that belong to a frequency band as a feature vector [70, 71].

The similarity in cycle duration between PPG and ABP waveform depicted in Figure 3.2 intuitively suggests that both signals share similar harmonic components. Therefore, it could be feasible to design a filter whose input is PPG and output is estimated ABP. Millasseau et al. (2000) [109] tried to find the frequency response of such filter $H(f)$ by dividing the FFT coefficients of these two signals for each subject:

$$H(f) = \frac{FFT_{BP}(f)}{FFT_{PPG}(f)}, \tag{3.34}$$

Then, a universal filter is acquired by averaging $H(f)$ across all subjects. [109] involved 60 normaltensive and hypertensive subjects and measured the PPG and fingertip ABP waveform under rest condition, then estimated the ABP waveform with the PPG signal using $H(f)$, and reported an overall RMSE of 4.4 ± 2.0 mmHg. The novel results out of [109] as well as other studies sharing similar ideas [110, 111] imply a universal linear correlation between frequency components of PPG and ABP, suggesting that spectral domain information of PPG could be feasible features in ML algorithms.

Table 3.5: Summary of studies using spectral PPG features to estimate BP.

Year	Ref.	Dataset	Intervention	Feature Type	Estimation Type	ML Method	Training	Performance			Key Conclusion
								SBP	DBP	MAP	
2016	[70]	69 subjects from MIMIC-II database; arterial BP reference	None	FFT frequency bins under 10.8 Hz	1 window =10% previous cycle + complete cycle + 5% following cycle ($N = 175477$)	MLP with single hidden layer	Grouped; 70% training, 15% validation, 15% testing	0.06 ± 7.08(ME)	0.01 ± 4.66(ME)	N/A	PPG spectrum is feasible for BP estimation
					Averaged over 5 beats			0.06 ± 5.57(ME)	0.01 ± 3.69(ME)	N/A	
		23 additional subjects			Averaged over 8 seconds		Additional testing	-1.67 ± 2.46(ME)	-1.29 ± 1.71(ME)	N/A	
2018	[71]	72 subjects from MIMIC database; arterial BP reference	None	22 PPG features: 20 frequency bins below 10 Hz, SUT, DT	Beat-to-beat ($N = 58795$)	MLP with single hidden layer	Grouped; 70% training, 15% validation, 15% testing	-0.0217 ± 4.8950(ME) 4.02 ± 2.79(MAE)	0.0975 ± 2.9160(ME) 2.27 ± 1.82(MAE)	N/A	Advocate multitaper method for estimating PPG spectrum

Recent studies have tried to utilize spectral domain PPG features in ML models. Study from Xing et al. (2016) [70] formed a MLP with single 35-neuron hidden layer, while using the FFT coefficients (magnitudes and phases) calculated from the PPG window directly as the

feature vector. [70] reported comparable or better accuracy compared to methods using $40 \sim 60$ temporal domain features [106, 77]. Since temporal domain PPG features have to be extracted by manually-designed algorithms that accord with their definitions, the FFT-MLP method in [70] is much easier to implement than [106, 77]. The accuracy of the FFT-MLP can be further improved by averaging over multiple windows, as is reported in [70].

Another study from Wang et al. (2018) [71] utilized a method similar to [70], while adding 2 additional temporal domain features, SUT and DT, into the feature vector. [71] also advocated for using the multi-taper method [112] to estimate the spectrum of the PPG signal instead of using FFT. [71] reported better result than [70] under the same beat-to-beat condition. However, dataset used in [71] is only about half the size of dataset used in [70], which indicates a potentially unfair comparison.

Besides FFT, trials have also been made to extract the spectral domain PPG features for BP estimation with discrete wavelet transform (DWT) [113] or discrete cosine transform (DCT) [114]. However, biased BP estimation errors are reported in [113, 114].

3.4.3 Whole-based Features

While the temporal domain features are explicitly defined and the spectral domain features are semi-defined, both of them have to be extracted with predefined algorithms, such as the peak detection algorithms or the FFT. The whole-based PPG feature, on the other hand, is a feature vector formed by the PPG signal itself (sometimes with necessary processing for dimension reduction, e.g. principle components analysis (PCA)). The ML algorithm is therefore expected to spontaneously find and extract appropriate features from the PPG signal. Utilization of whole-based features often requires advanced ML models with enough complexity.

Whole-based PPG feature is by now the latest concept in cuff-less BP estimation studies. Mousavi et al. (2019) [99] proposed a method in which the feature vector is formed by preprocessed PPG samples within a cycle between 2 consecutive systolic peaks. The PPG samples are zero-mapped to a fixed length of 625 samples, and is subsequently reduced to 43 samples by PCA, which directly forms the 43-dimensional feature vector as the input of 500 AdaBoosted regression trees. SBP and MAP results ($ME = 0.187 \pm 4.173$ mmHg for SBP, $ME = 0.067 \pm 4.911$ mmHg for MAP) reported in [99] are top-of-the-line, while DBP results

Table 3.6: Summary of studies using whole-based PPG features to estimate BP.

Year	Ref.	Dataset	Intervention	Feature Type	Estimation Type	ML Method	Training	Performance			Key Conclusion
								SBP	DBP	MAP	
2019	[99]	> 441 subjects from MIMIC-II database; arterial BP reference	None	PPG samples between 2 consecutive systolic peaks	Beat-to-beat ($N = 1323$)	PCA for dimension reduction; AdaBoost on regression tree	Grouped; 10 fold cross validation	0.91(R) 0.187 ± 4.173(ME)	0.90(R) -0.050 ± 8.901(ME)	0.91(R) 0.067 ± 4.911(ME)	Advocate usage of whole-based features
						Comparison: RFR		0.196 ± 4.731(ME)	0.155 ± 10.683(ME)	0.196 ± 5.714(ME)	
						Comparison: SVR		-0.655 ± 7.506(ME)	-0.903 ± 16.717(ME)	-0.597 ± 9.055(ME)	
						Comparison: Regression tree		-0.247 ± 6.736(ME)	0.021 ± 18.543(ME)	-0.050 ± 9.594(ME)	
2019	[100]	478 young subjects, normaltensive and hypertensive; cuff BP reference	None	19 features: 4+8 whole-based PPG and APG features from PCA, 4 APG amplitude features, LASI, heart rate, BMI	Beat-to-beat ($N = 739$)	PCA for dimension reduction; RFR	Grouped; leave 1 subject out each time	0.86(R) 0.45 ± 11.3(ME)	0.83(R) 0.31 ± 8.55(ME)	N/A	Algorithm has better performance in the young group and in the normaltensive group
		754 aged subjects, normaltensive and hypertensive; cuff BP reference			Beat-to-beat ($N = 1340$)			0.79(R) -0.68 ± 14.1(ME)	0.81(R) -0.20 ± 9.0(ME)	N/A	

are inferior in terms of large error variance ($ME = -0.050 \pm 8.901$ mmHg). One drawback of study [99] is its relatively-small dataset size ($N = 1323$) among data-driven studies, which could imply unfair comparison.

Another study from Xing et al. (2019) [100] extracted whole-based features from both PPG and APG with a similar method to [99], but utilized a feature vector with lower dimension. Using PCA, a PPG cycle is reduced to 4 and 8 points whole-based features for PPG and APG, which are combined with other 6 temporal domain features and subject's BMI to form the final feature vector with 19 elements. Probably due to the reduced number of elements in whole-based features and the absence of inclusive temporal features, results from [100] is inferior to [99].

3.5 Methods Using Both ECG and PPG: The Pay and Gain From An Additional Sensor

From existing studies [72, 23, 73, 95], it seems that the only purpose of utilizing both ECG and PPG signals in a data-driven method is to include PAT in the feature vector, while most of other involved features are still PPG features. An algorithm requiring both ECG and PPG signals as input is less desirable for wearable device implementation compared to those requiring only PPG, as more power has to be consumed for measuring an additional signal, and more data has to be stored. In order to extract PAT, the ECG and PPG signals need to be continuously synced, which is hard to maintain over long period of time. Moreover, a device measuring both ECG and PPG will take up more space (see Figure 1.2 in Chapter 1, in which extra cables are required to connect ECG nodes to the device), and is less comfortable to wear

compared to devices that only measures PPG. With all these price paid for including PAT in the BP estimation algorithm, a question ask is whether PAT is so effective for estimating BP that the gains outweighs the pays.

Table 3.7: Summary of studies using PPG features and PAT to estimate BP.

Year	Ref.	Dataset	Intervention	Feature Type	Estimation Type	ML Method	Training	Performance			Key Conclusion
								SBP	DBP	MAP	
2006	[72]	45 healthy male subjects; SBP references	None	3 features: PAT$_p$, weight, arm length	Beat-to-beat ($N = 180$)	PCA and correlation-based feature selector; MLP with single hidden layer	Grouped; 160 training, 20 testing	4.53 ± 2.68(ME)	N/A	N/A	Proposed PAT method without individual calibration
2016	[23]	942 subjects from MIMIC-II database; arterial BP reference	None	10 features: PAT$_f$, PAT$_d$, PAT$_p$, heart rate, AI, LASI, A1$\sim$A4	2-beat window ($N = 3663$)	AdaBoost	Grouped; 10-fold cross validation	0.59(R) $N/A \pm 10.09$(ME) 11.17(MAE)	0.48(R) $N/A \pm 6.14$(ME) 5.35(MAE)	0.56(R) $N/A \pm 5.38$(ME) 5.92(MAE)	An additional calibration process could help enhancing the accuracy; AdaBoost yields better performance than SVR, RFR, MLR and decision tree
				Comparison: 15 whole-based PPG features from PCA				$N/A \pm 10.30$(ME) 11.87(MAE)	$N/A \pm 6.61$(ME) 5.78(MAE)	N/A	
				10 features			Additional one-point calibration for groups in testing set	0.54(R) $N/A \pm 5.45$(ME) 8.21(MAE)	0.57(R) $N/A \pm 3.52$(ME) 4.31(MAE)	N/A	
2017	[73]	7 healthy subjects; Finapres BP reference	Rest	15 features: PAT, 14 PPG temporal features	Beat-to-beat	MLP with single hidden layer	Subject-specific; first 5 min data for training, last 10 min data for testing	1.08 ± 4.87(ME)	-0.52 ± 3.84(ME)	N/A	Additional calibration could further improve performance
							Additional calibration by forming a regression curve in the training set	0.994(R) 0.41 ± 2.02(ME)	0.990(R) 0.46 ± 2.21(ME)	N/A	
2017	[95]	73 healthy subjects; Finapres BP reference	Rest	14 features: PAT$_f$, PAT$_d$, PAT$_p$, 4 PPG temporal features, 2 PPG first derivative temporal features, 4 APG temporal features, heart rate	Beat-to-beat	Conlinearity-based feature selector; MLR	Subject specific	0.824(R) 0.0016 ± 3.449(ME)	0.754(R) 0.0017 ± 2.468(ME)	N/A	Significance of feature is subject specific; feature importance is similar between Rest and Exercise subsets; proposed feature importance ranking shows that PAT is the most significant feature for SBP, while for DBP it is PIR
		35 healthy subjects	Exercise					0.941(R) -0.046 ± 4.705(ME)	0.923(R) -0.071 ± 2.839(ME)	N/A	
		10 healthy subjects; follow up after 6 months	Rest				Subject specific; use parameters trained from Rest dataset	0.619(R) -1.267 ± 5.98(ME)	0.549(R) -1.38 ± 5.49(ME)	N/A	
		Comparison Subset	N/A				Subject specific	-0.54 ± 3.39(ME)	-0.002 ± 2.60(ME)	N/A	
				Comparison: PAT+PIR		Comparison: Model-based [55]		-0.23 ± 8.57(ME)	-0.61 ± 5.51(ME)	N/A	

From Chapter 2, we have discussed that the relationship between PTT and BP is dependent of the transit distance that the pulse wave travels, which is subject-specific. A study from Kim et al. (2006) [72] tried to get rid of subject-specific model calibration by including addition subject information related to the transit distance as feature. [72] utilized MLP to form the relationship between PAT and BP in a data-driven way. Besides PAT, [72] selected 2 additional features from a pool of 6 individual information features including weight, BMI, body fat, height, arm length and arm circle. Feature selection is conducted using a PCA-based method, where features having highest correlation with the principle component and the BP are selected. The selected features are weight and arm length, which is reasonable since arm length is approximately the transit distance between heart and fingertip. However, SBP estimation performance reported by [72] has biased error.

A more recent study from Kachuee et al. (2016) [23] compared the BP estimation performance between an algorithm using a 10-element feature set including 3 types of PATs, and another algorithm using a 15-element whole-based PPG feature set, whose dimension is reduced from 190 PPG samples using PCA. Both algorithms used 1000 AdaBoosted regression trees for ML, and reported very close performance. [23] also proposed an addition calibration method after model training, in which the model trained on the training set is calibrated on 1

data in each subgroup in the testing set. The addition calibration procedure proposed in [23] lowers the variance of estimation error, but is not capable of improving the correlation between estimated and reference BPs.

Xu et al. (2017) [73] reported better results than [23] by turning back to subject-specific calibration and adding more PPG features to the feature set. Very small error ($ME = 0.41 \pm 2.02$ mmHg for SBP, $ME = 0.46 \pm 2.21$ mmHg for DBP) and extremely high correlation ($R = 0.994$ for SBP, $R = 0.990$ for DBP) has been reported by [73], with an additional calibration procedure after model training on the training set that forms a regression curve for error correction, which is more reasonable than [23] whose calibration procedure is on the testing set and could indicate data leakage. However, dataset used in [73] contains only 7 subjects.

Study from Miao et al. (2017) [95] discussed the accuracy degradation problem in data-driven method. In [95], results tested on follow-up data ($ME = -1.267 \pm 5.98$ mmHg for SBP) acquired 6 months after the subject-specific model calibration ($ME = 0.0016 \pm 3.449$ mmHg for SBP) showed that ML model can better preserve accuracy over time compared to model-driven methods based on PWV [44] (SBP: error increased from $ME = 0 \pm 4$ mmHg to $ME = 1.4 \pm 10.2$ mmHg after 6 months). The MLR model proposed in [95] also shows advantage in accuracy when compared to the PWV model proposed in [55] under same dataset. [95] also proposed a general feature ranking by averaging the feature importance evaluated over each subject among all subjects, which is summarized as follows: (from most significant to less significant)

For SBP estimation:

1. PAT_d.
2. PPG characteristic value.
3. PIR.
4. b.
5. Width of APG.
6. PAT_f.

7. Amplitude of 1st derivative of PPG.
8. Width of 1st derivative of PPG.
9. $a+b$.
10. a.

For DBP estimation:

1. PIR.
2. Width of APG.
3. b.
4. PPG characteristic value.
5. PAT_d.
6. Width of 1st derivative of PPG.
7. a.
8. Amplitude of 1st derivative of PPG.
9. DT.
10. HR.

The feature ranking proposed by [95] has its novelty of including PAT in the feature pool, compared to [107, 77] which only compared the importance of PPG features. The ranking shows that PAT could be a desirable indicator (especially for SBP), but probably only in circumstances where the model is subject-specifically calibrated.

3.6 Methods Using Only ECG: Thoughts Beside Ordinary Waveform Morphology

Unlike PPG signal having a large variety of interpretable morphological features with high correlation with BP, their is little correlation, both theoretically and empirically, between ECG

morphology and BP. Therefore, very few studies have considered ECG alone as a usable BP indicator. However, some very recent works are offering heuristic perspectives outside the signal morphology, which could be revolutionary.

Table 3.8: Summary of studies using only ECG to estimate BP.

Year	Ref.	Dataset	Intervention	Feature Type	Window / Estimation Type	Machine Learning Method	Training	Performance			Key Conclusion
								SBP	DBP	MAP	
2018	[115]	51 subjects from mixed sources; mix of cuff and arterial BP reference	None	7 features: mobility, complexity, fractal dimension, entropy, autocorrelation of ECG, age; hypertension level from stacked classifier	30 s window ($N = 3129$)	Stacked classifier to classify hypertension level; RFR	Grouped; 60% subjects training, 10% subjects validation, 30% subjects testing	8.64 ± 10.74(MAE) 10.97(RMSE)	18.20 ± 8.45(MAE) 19.34(RMSE)	13.52 ± 8.06(MAE) 15.07(RMSE)	Statistical index of signal complexity can be used for BP estimation
2020	[76]	1711 subjects from MIMIC-III database; arterial BP reference	None	ECG samples	3-beat window ($N = 897743$)	Deep learning combining 50-layer ResNet and 2-layer LSTM (404520 parameters)	Grouped; 65% subjects training, 10% subjects validation, 25% subjects testing	0.88(R) -0.11 ± 9.99(ME) 7.10(MAE)	0.71(R) 0.01 ± 6.29(ME) 4.61(MAE)	0.85(R) -0.03 ± 6.66(ME) 4.66(MAE)	Proposed deep learning-based method to estimate BP from ECG with novel accuracy
		30 arrhythmia patients from ARR database			($N = 1342$)		Additional testing	0.96(R) -0.22 ± 5.82(ME) 4.41(MAE)	0.74(R) -0.75 ± 5.62(ME) 4.37(MAE)	0.91(R) -0.57 ± 4.39(ME) 3.56(MAE)	

Simjanoska et al. (2018) [115] proposed a method in which features extracted from ECG are purely statistical indices, including mobility, complexity, fractal dimension, entropy and autocorrelation. Moreover, similar to the combination of classification and regression presented in [97], [115] first trains a stacked classifier that takes the feature vector as input and output a class corresponding to BP level (normaltension, pre-hypertension, hypertension), and then uses this class as an additional feature for training the RFR model. However, the reported accuracy is not novel.

Another recent study from Miao et al. (2020) [76] is the first ECG-based method that achieved BP estimation accuracy comparable to PPG-based or PWV-based methods. In [76], the BP estimation problem is interpreted under an intact whole-based, end-to-end deep learning framework. While previous studies utilizing whole-based features [99, 100] used PCA to reduce the dimension of feature vector after the feature vector is formed by samples of the PPG signal, [76] used a 50-layer ResNet [116] to learn the features from the raw ECG signal in a completely spontaneous way, and a 2-layer LSTM network [117] to form the relationship between the learned features and the BP. The model proposed in [76] involves 404520 parameters trained from a dataset with 897743 indices, which is a scale that has never been reached in previous studies.

3.7 Conclusions, Limitations and Suggested Future works

In this chapter we provided a review of data-driven studies for estimating BP, with respect to the features and the ML methods, the BP estimation performance and the novelty as well as limitations of these works. In summary, the following points can be highlighted:

- **Model generalization**: While model-driven works reviewed in Chapter 2 generally require subject-specific model calibration, most data-driven ML models discussed in this chapter can be trained and tested on grouped datasets involving multiple subjects, while obtaining comparable or ever better BP estimation accuracy compared to model-driven methods, such as [106] utilizing 46 PPG temporal domain features and [70] using PPG spectral domain features. As such, data-driven methods are very promising for developing universal models and devices that accurately measure cuff-less BP for the population.

- **PPG signal and PPG features**: Many studies reviewed in this chapter [59, 106, 60, 70, 71, 99] provide novel BP estimation accuracy from only the PPG signal. With evidences from theoretical studies revealing the biological connections between PPG and ABP [86, 87] and correlation studies validating the statistical correlations between PPG features and BP [88, 89, 75], we conclude that PPG is very feasible for BP estimation.

- **Model complexity and feature vector dimension**: Results from reviewed studies such as [107, 59, 99, 76] showed that advanced machine learning methods with more complexity such as bagging, boosting, SVR or MLP are more likely to produce better results compared simple methods such as MLR or regression tree. Moreover, it seems that studies including more features generally achieve better accuracy [75, 106, 61], as long as the features are properly selected from necessary redundancy removal processes [60, 77]. An increase of dataset size can also be observed in [106, 60, 77] for addressing the data sparsity problem discussed in Section 3.2.2.

- **Efficiency of utilizing both ECG and PPG**: Although correlation between PAT and BP has been validated when the dataset is limited to one subject [73, 95], it seems that adding ECG signal to include PAT as an additional feature of a universal model that mostly relies on PPG features is not an efficient idea [23]. While [72] has shown the importance of

including transition distance indicators in the model for getting rid of subject-specific calibration, it is hard to find such indicator in PPG features, which could explain the inferior results reported from [23] when the dataset included multiple subjects. However, with new methods of analyzing the ECG signal proposed in recent studies [115, 76], it is possible for future algorithms combining ECG and PPG signals to produce novel accuracy.

In addition, limitations of reviewed studies and suggested topics for future works are summarized below:

- **Limitation of model interpretability**: Generally, data-driven cuff-less BP estimation methods are still at a preliminary stage. Studies with novel performance [59, 106, 60, 70, 71, 99] are very different from each other in terms of the selections of features and ML methods, such that it is hard to tell which factor is contributing most to a good model. Such limitation in model interpretability also casts doubts on the reliability of data-driven models when migrating to other groups of populations. With recent work [76] visualizing the features extracted by the ResNet with a deconvolution process, it is suggested that future studies can deepen our interpretations of data-driven methods.

- **Limitation of datasets**: Data-driven methods require much larger datasets ($N = 10^2 \sim 10^6$) compared to model-driven methods ($N = 10^1 \sim 10^3$). In model-driven studies, various BP intervention methods, such as exercise, drug, mental arithmetic and cold press, can be utilized because the required datasets are small enough to be acquired from customized experiments. However, most data-driven studies rely on online open datasets to train the ML models, because the ML models have much more parameters than the PWV models and require larger datasets to avoid over-fitting. With most of these online datasets acquired from hospitalized patients, there could potentially exist a bias in the training set. We therefore advocate for a combined training process in which the model is pre-trained on large, open datasets, and then tested or further trained on other customized datasets to address the potential training bias and further improve the generalizability of the model in different scenarios.

- **High computational budget for model training**: Implementing of some best-performing

methods [106, 70, 76] from scratch could be very hard due to their high computation cost to train the models on massive datasets ($N = 10^6$). It is thus suggested that the training efficiency of data-driven methods should be considered as another important factor besides their BP estimation accuracy. Future works can also explore the feasibility of utilizing pre-training methods for BP estimation, such as transfer learning.

- **Semi-defined and whole-based features**: Extraction of explicitly defined PPG features could be difficult or computational expensive in some situations, such as when the PPG signal is interfered by noise or artifacts, when the PPG signal is measured from animals or tissues other than fingertip, or when massive amount of features are required. Semi-defined or whole-based features such as spectrum or whole-based features have their advantage of easier implementation and better generality over PPG signals measured from different sources. Considering the promising accuracy of newly proposed works [99, 70, 76], it is very desirable if such methods can be further developed.

- **Interdisciplinary methods**: The data-driven machine learning problem can be considered in many other ways besides features and their correlations with BP: [109] has demonstrated how this problem can be interpreted as finding a filter whose input is the PPG signal and output is the the ABP signal; [76] exhibited another perspective in which the problem is addressed in a way closer to image recognition and computer vision problems. Ideas and concepts from other fields could potentially become breakthroughs. As such, we should be encouraged to seek for improvements outside the current scope.

Chapter 4

PPG-Based Blood Pressure Estimation Method Using Visibility Graph Features and Deep Learning

4.1 Introduction

From all studies reviewed in Chapter 2 and Chapter 3, we noticed that the state-of-the-art studies are among data-driven works in Chapter 3, which utilize explicitly-defined temporal domain features [59, 106, 60], semi-defined spectral domain features [70, 71], or whole-based features [99]. Advantages of these methods are top-of-the-line BP estimation performances, good model generalizability being free of subject-specific calibrations, and requiring only PPG signal as input. However, most of these methods are computationally expensive for implementation: in [59, 106, 60], more than 30 features have to be extracted from the PPG signal with different manually-designed algorithms that accord with their definitions, which is not only costly, but also prone to error; in [70, 71], the demanded size of training set is massive ($N = 10^4 \sim 10^5$), which again requires high computational budget to train the model from scratch.

Motivated by the novelty of the abovementioned PPG-based data-driven BP estimation methods, in this work, we present a new BP estimation method that preserves their advantages, while addressing their limitation of being computationally expensive. Our proposed framework utilizes a method, called visibility graph (VG) [118], for converting segments of PPG signals into images, followed by transfer learning with Inception v3 [119], a convolutional neuron network (CNN) with top-of-the-line image classification performance pre-trained on the ImageNet database [120]. Our method achieved the following targets:

- It requires only one type of physiological signal, which facilitates the data acquisition process.

- It eliminates the need for subject-specific calibration by having one universal model applicable to all subjects.
- It enables the BP estimation problem to be addressed with transfer learning method by converting time series into images, which greatly reduces the computational budget required for training a model from scratch.

The proposed method provides novel accuracy in terms of correlation and estimation error compared to the reference values. Under optimal setting, the mean and the standard deviation (SD) of the error between the estimated and reference BP with proposed method are -0.209 ± 9.476 mmHg for SBP, and -0.067 ± 4.491 mmHg for DBP, respectively. The SBP accuracy ranks grade C, while the DBP accuracy ranks grade A, under the British Hypertension Society (BHS) protocol.

4.2 Proposed Methods

4.2.1 Dataset Information

The dataset used in this study was selected from the UCI Machine Learning Repository of cuff-less blood pressure estimation [121], which is a subset of the Multi-parameter Intelligent Monitoring in Intensive Care (MIMIC) II waveform database [122]. The UCI dataset contains 12000 segments, with each segment containing continuous PPG, ECG, and invasively measured arterial BP (ABP) signal, with the sampling rate of 125 Hz. Our proposed method uses the PPG signal as input, and the ABP signal as reference values to evaluate the error of estimated SBP and DBP.

Manual segment selection was conducted to select a subset from the UCI dataset for this study with reduced size but sufficient SBP and DBP variation for training and validation. For all segments, the beat-to-beat SBP and DBP values were extracted from the ABP signal and 32 segments were selected. Figure 4.1 shows the histogram of extracted SBP and DBP in the original dataset, and in the selected dataset. As can be seen in the UCI dataset, the mean and the SD for SBP and DBP are 128.39 ± 22.05 mmHg and 66.44 ± 11.29 mmHg, respectively. In the selected subset dataset, these are 132.15 ± 18.50 mmHg and 68.18 ± 9.40 mmHg for SBP

and DBP, respectively, indicating that our selected subset sufficiently covers the variations in BP that exist in the original dataset.

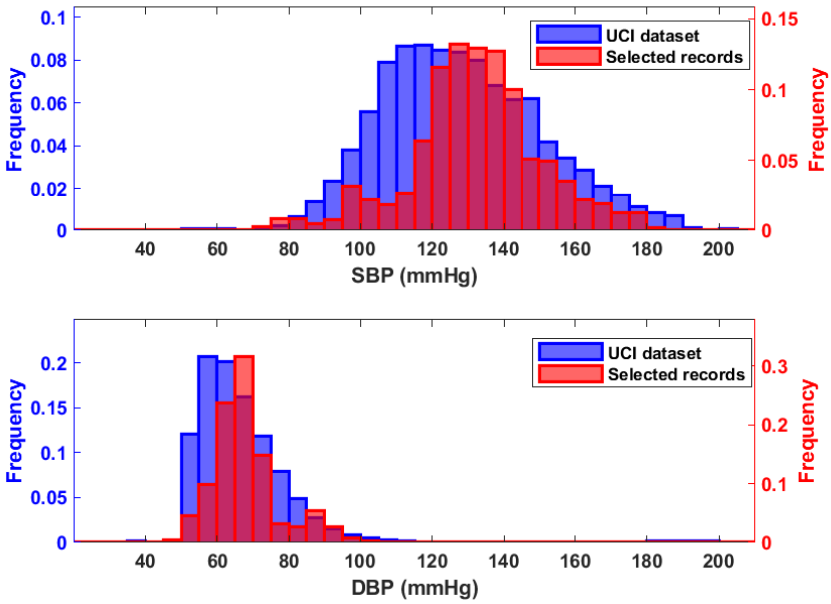

Figure 4.1: Histogram of all SBP (top) and DBP (bottom) values in the whole UCI dataset (blue bars), and in the subset (red bars) selected for this study.

4.2.2 Preprocessing

The signals in each segment were filtered with a forward-backward Butterworth filter. The forward-backward filtering process has zero phase response, which preserves signal morphology with no phase distortion and keeps the signals synced. As the effective frequency components of the PPG signal remain lower than 11 Hz [123], to remove noise and artifacts the PPG signal was band-pass filtered at $0.5 \sim 10$ Hz. The ABP signal was low-pass filtered at $0 \sim 10$ Hz to remove high frequency interference.

4.2.3 Peaks Detection and Windowing

After preprocessing, a windowing method was applied to the PPG signal, so the beat-to-beat BP values can be estimated from the corresponding non-overlapping PPG windows. Non-overlapping PPG windows were selected to prevent potential data leakage between the training and testing sets.

The windowing method starts with the detection of the systolic peaks in the PPG signal. First, each segment is divided into 20 seconds sub-segments. For each sub-segment the amplitude of the 1st derivative of the signal is derived and remapped between 0 and 1. Next, all possible peaks within the sub-segment are detected. A small interval of the 1st derivative of the signal before each peak is inspected, and its maximum amplitude is found. A threshold level can then be selected to separate the systolic peaks from other peaks, as the fast rising systolic phase before the systolic peak shows as a high spike in the 1st derivative, making the systolic peak distinguishable from other peaks, as is shown in Figure 4.2. After applying the threshold, the systolic peaks can be identified, while other peaks are excluded.

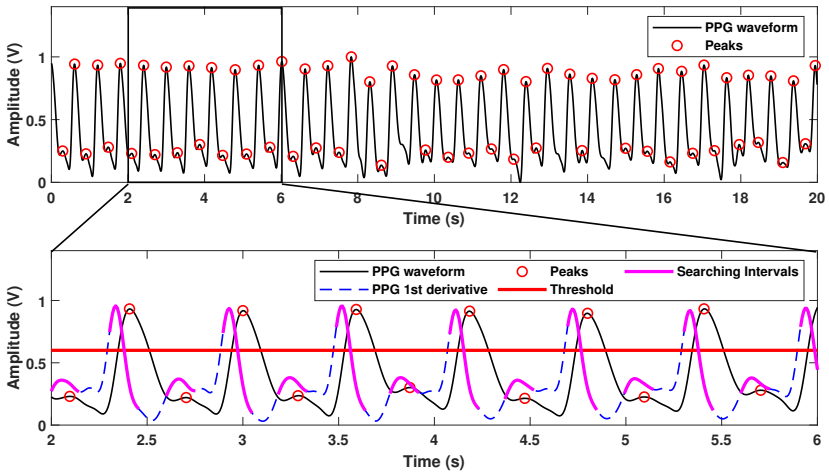

Figure 4.2: The proposed systolic peak detection method. Top: a 20 seconds sub-segment of PPG signal and all detected peaks for selection (red circles). Bottom: zoomed-in specification of the $2 \sim 6$ second of the signal and its first derivative. The threshold (red solid line) applied to the searching intervals (magenta solid line) extracted from the first derivative of signal (blue dashed line) selects systolic peaks from other peaks.

Once the systolic peaks in a segment are located, the PPG signal from the corresponding segment was divided into windows, with one pair of SBP and DBP estimation expected from each window. 3 settings of window duration, which is referred to as 1-beat, 2-beat and 3-beat settings, were experimented to find the optimal window duration that provides best BP estimation accuracy. For 1-beat, 2-beat or 3-beat setting, the window contains 1, 2 or 3 complete PPG cycle(s), which is located by the positions of systolic peaks. Figure 4.3 demonstrates the positions of 3 consecutive windows under 1-beat setting, with the position of each window located by every 3 consecutive systolic peaks. Under the 2-beat or 3-beat setting, the position

of each window is located by every 4 or 5 consecutive systolic peaks, respectively.

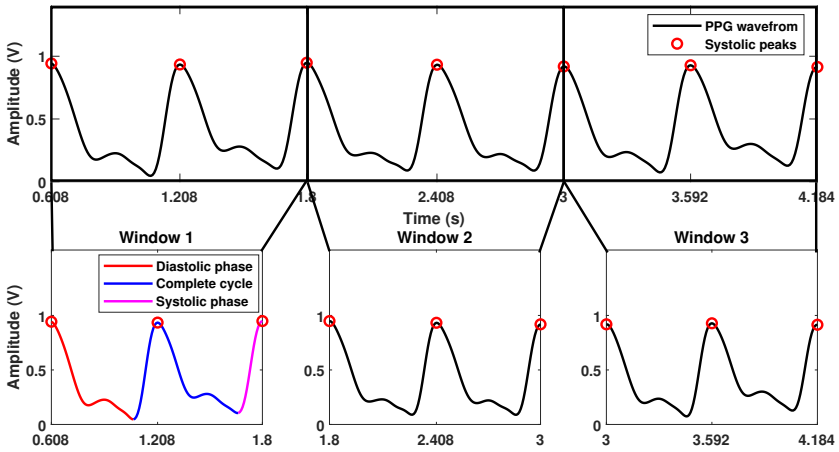

Figure 4.3: Result of the proposed windowing method under 1-beat setting. Top: consecutive PPG waveform. Bottom: 3 non-overlapping windows extracted from the PPG waveform.

4.2.4 VG Extraction

The challenge for using transfer learning in this application is to select a proper method for domain transformation, such that the key information is efficiently described in the alternative domain. To address this issue, we chose to convert the time-domain PPG signals into images because the similarity between characteristics of image recognition problem and the time-domain signal interpretation problem is significant to make our method promising.

To create images where temporal information of PPG signals is preserved, we used VG. VG is a method that maps a time series into an undirected graph [118, 124], and has shown great promise in extracting temporal information from physiological signals [125, 126, 127]. Let $\mathbf{x} = [x_1, \cdots, x_N]$ represent a time series of N points, where x_i $(i = 1, \cdots, N)$ denotes the i^{th} sample in the time series. Let t_i represent the time corresponding to occurrence of sample x_i. To construct the visibility graph for this time series, each sample is considered as a node in the graph. An undirected edge is formed between any two nodes if the nodes are considered to be ***naturally visible***, i.e., for two nodes h and l $(t_h < t_l)$, there will be an undirected and unweighted edge if

$$x_p < x_l + (x_h - x_l)\frac{t_l - t_p}{t_l - t_h}, \forall p \in \{p | t_h < t_p < t_l\}. \tag{4.1}$$

Here, 2 VGs are formed for every PPG window, as is shown in Figure 4.4 where the window

duration is under 1-beat setting. First, the number of samples in all PPG windows was adjusted to a fixed length by zero-mapping at the end. For the windows under 1-beat, 2-beat or 3-beat settings, the length is set to 250, 375 and 500 samples, respectively. Next, the inverted version of the waveform for each window was obtained. Finally, in all windows, the time series of both the original waveform and the inverted waveform were converted to VG, with each VG described with adjacency matrix filled with 0 and 1. The matrix here was considered as an image to form the input of CNN. We refer to the image corresponding to original PPG waveform as VG_POS, and the image corresponding to inverted PPG waveform as VG_INV.

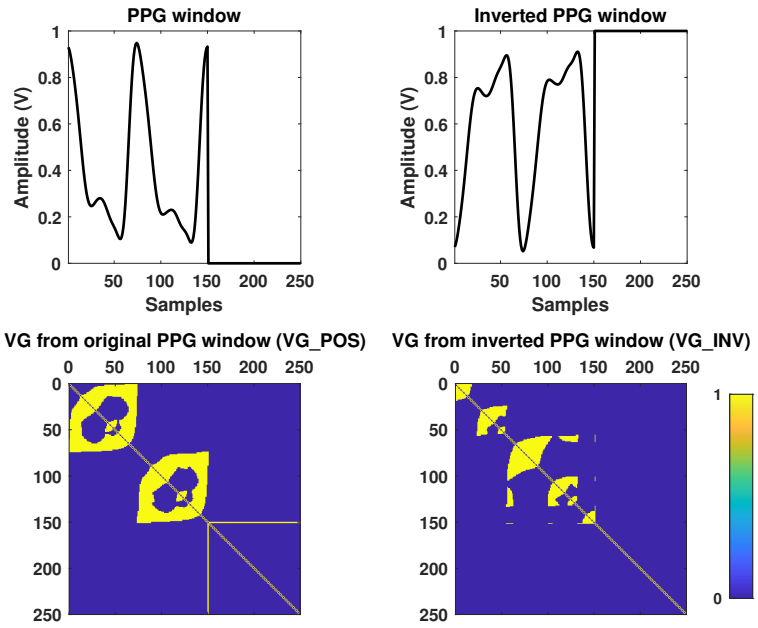

Figure 4.4: Plot of a zero-mapped window, the inverted window, and 2 VGs (VG_POS and VG_INV) formed with the PPG samples in that window. Here, the window duration is under 1-beat setting.

4.2.5 Transfer Learning with CNN

Since the Inception v3 model takes 3-channel RGB images as input, the VG matrix was firstly extended to 3 channels by replicating itself. Next, the VG matrix was directly passed to the bilinear rescaling layer of the model. By forward propagation, a 2048-dimension feature vector was obtained from the bottleneck layer before the softmax layer of the CNN, as is shown in Figure 4.5. Finally, we trained a dense layer between the feature vector and SBP and DBP

estimation outputs by minimizing the mean squared error between reference and estimated SBP and DBP values in the training set, under 3 settings: If only VG_POS or VG_INV is used, then the size of dense layer is 2048×2; If both VG_POS and VG_INV are used, then the feature vector is the concatenation of 2 feature vectors from each VG, and the size of dense layer is 4096×2, as is shown in Figure 4.6.

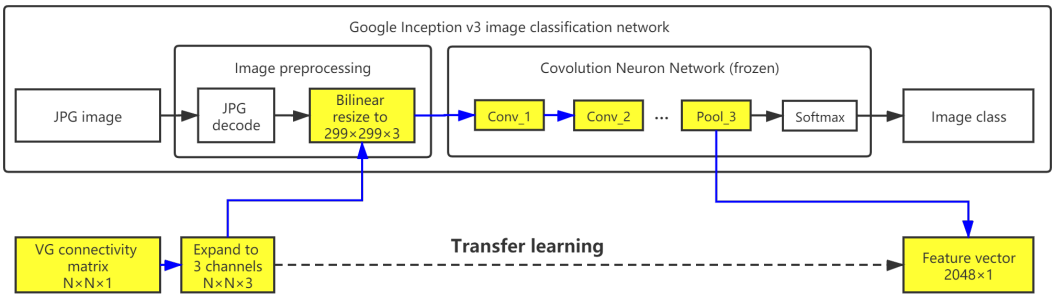

Figure 4.5: Flow chart of the transfer learning process. Yellow blocks indicate activated layers in the forward propagation process.

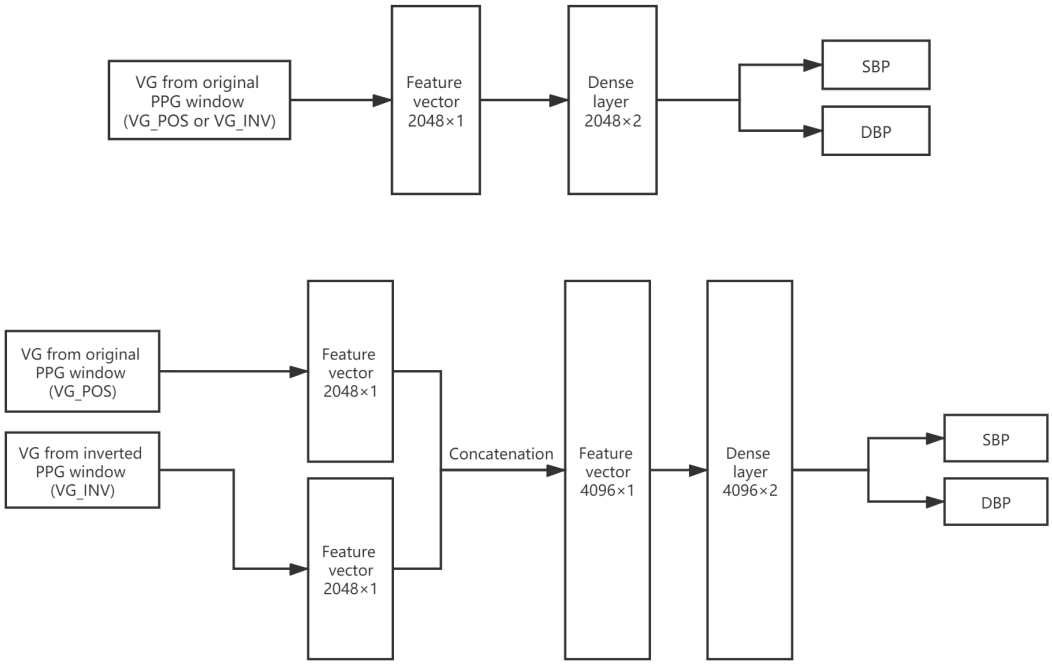

Figure 4.6: Formation of the feature vector and the dense layer using only VG_POS or VG_INV (top) or both VG_POS and VG_INV (bottom).

Table 4.1: Summary of SBP and DBP estimation performance under different combinations of window duration settings and VG usage settings. Bold font indicates optimal performances.

VG Features	Window	N	SBP R	SBP ME±SD (mmHg)	SBP MAE±SD (mmHg)	SBP RMSE (mmHg)	DBP R	DBP ME±SD (mmHg)	DBP MAE±SD (mmHg)	DBP RMSE (mmHg)
	1-beat	**8676**	**0.859**	**-0.209±9.476**	**6.890±6.509**	**9.478**	**0.879**	**-0.067±4.491**	**3.146±3.205**	**4.491**
POS+INV	2-beat	5907	0.832	0.121±10.175	7.418±6.964	10.175	0.850	0.075±4.929	3.418±3.552	4.929
	3-beat	4368	0.812	0.156±10.744	7.820±7.368	10.744	0.841	0.071±5.050	3.582±3.559	5.050
	1-beat	8676	0.838	-0.080±10.097	7.419±6.849	10.097	0.859	0.057±4.814	3.395±3.414	4.815
POS	2-beat	5907	0.804	0.028±10.931	8.136±7.299	10.930	0.834	0.031±5.153	3.619±3.668	5.152
	3-beat	4368	0.772	-0.144±11.706	8.679±7.854	11.705	0.806	-0.029±5.507	3.903±3.885	5.507
	1-beat	8676	0.802	-0.040±11.065	8.310±7.306	11.064	0.841	0.167±5.085	3.631±3.565	5.088
INV	2-beat	5907	0.778	-0.246±11.548	8.684±7.614	11.549	0.805	0.009±5.543	3.919±3.919	5.542
	3-beat	4368	0.756	-0.156±12.063	9.034±7.995	12.063	0.791	-0.027±5.698	4.061±3.996	5.697

4.3 Results

Using 10-fold cross validation, Table 4.1 summarizes the error performance under 9 different combinations of VG usage settings and window duration settings. The optimal result is achieved by utilizing both VG_POS and VG_INV, under 1-beat setting. For SBP and DBP, the correlation coefficient is 0.859 and 0.879, and the mean error and SD of error is -0.209 ± 9.476 mmHg and -0.067 ± 4.491 mmHg, respectively.

Under this optimal setting, the regression plot of reference and estimated BPs is shown in Figure 4.7, and the Bland-Altman plot of SBP and DBP are shown in Figure 4.8.

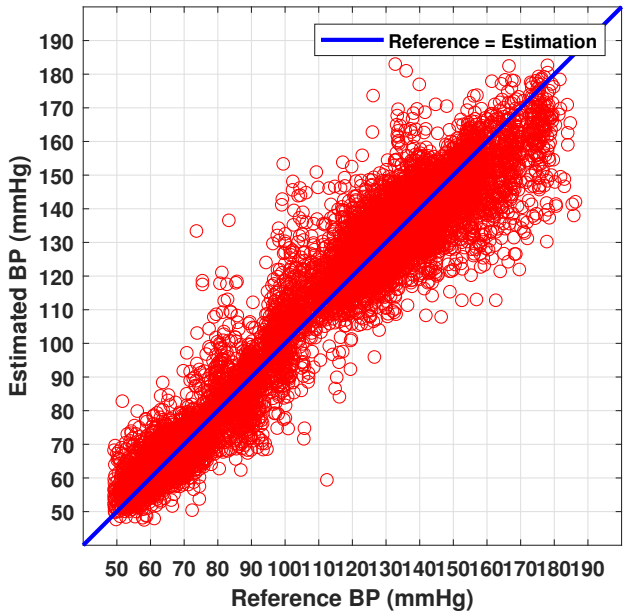

Figure 4.7: Regression plot between reference and estimated BP values.

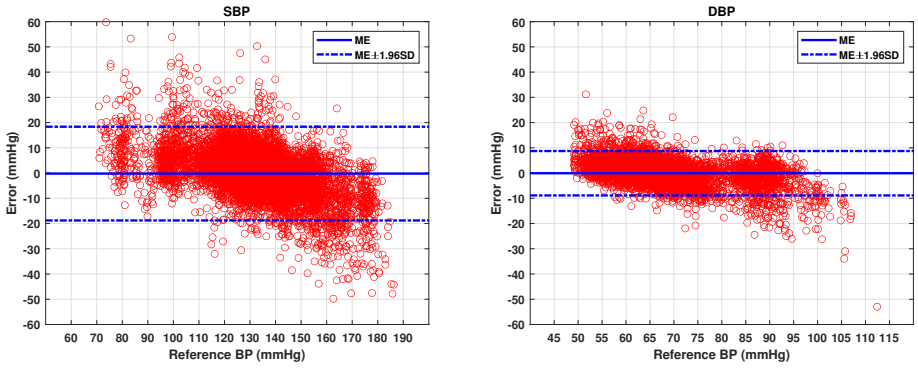

Figure 4.8: Bland-Altman plot of estimated SBP and DBP.

4.4 Discussions

The results listed in Table 4.1 demonstrated that using the proposed method, the DBP has a generally better estimation accuracy than the SBP under all settings in terms of smaller mean absolute error (MAE), root mean square error (RMSE), and standard deviation of error. The DBP performance under all settings are within the limits of the American National Standards of the Association for the Advancement of Medical Instrumentation (AAMI) [12], where the maximum acceptable error is 5 ± 8 mmHg. Under the optimal setting, the SBP estimation performance ranked grade C under the British Hypertension Society (BHS) protocol, while the DBP ranked grade A, as is shown in Table 4.2.

Table 4.2: Estimation accuracy of proposed work compared to the BHS protocol under the optimal setting using both VG_POS and VG_INV. SBP accuracy achieves grade C while DBP achieves grade A.

Item	**Precentage**		
	$\leq 5mmHg$	$\leq 10mmHg$	$\leq 15mmHg$
SBP	48.77%	78.31%	90.32%
DBP	81.49%	96.04%	98.78%
Grade A	60%	85%	95%
Grade B	50%	75%	90%
Grade C	40%	65%	85%

Under any given setting of VG usage, the results in Table 4.1 lead to the conclusion that

both SBP and DBP estimation accuracy drop as the window duration becomes longer. The decrease in accuracy could be relevant to the transfer learning procedure, as the Inception v3 model use bilinear resizing algorithm to shrink the input image to a 299×299 matrix. With the size of input VG matrix becomes larger along the increase of window duration, shrinking a large VG matrix to a fixed size could lead to losing more details compared to shrinking a smaller matrix, where here the details are relevant to the morphological features of the PPG signal. Additionally, previous studies utilizing temporal domain PPG features [59, 106, 60] and whole-based PPG features [99] have been focusing on the beat-wise morphology of the PPG instead of characteristics among multiple beats, which implies advantage of 1-beat setting over others.

On the other hand, under same window duration setting, comparisons among different VG usage settings show that using both VG_POS and VG_INV yields better results than using any one of them alone, which implies that VG_POS and VG_INV are complementary to some extent. However, the performances of proposed method under all 3 different VG usage settings are close to each other, which indicates redundancy between VG_POS and VG_INV.

Table 4.3: Summary of dataset and methods used in works for comparison. All listed works use only the PPG waveform for BP estimation. Bode font indicates this work.

Citation	**Dataset**	**Dataset size**	**Model**	**Validation**	**R**		**Error**	
					SBP	**DBP**	**SBP**	**DBP**
This Work	**UCI dataset**	**8676 beats**	**CNN**	**10 fold cross validation**	**0.859**	**0.879**	**6.890±6.509 (MAE±SD) -0.209±9.476 (ME±SD)**	**3.146±3.205 (MAE±SD) -0.067±4.491 (ME±SD)**
[107] Hasanzadeh et al. (2019)	UCI dataset	942 subjects	AdaBoost	10 fold cross validation	0.78	0.72	8.22 ± 10.38 (MAE±SD)	4.17 ± 4.22 (MAE±SD)
[100] Xing et al. (2019)	Self prepared	739 beats	Random forest	Leave one out	0.86	0.83	0.45 ± 11.3 (ME±SD)	0.31 ± 8.55 (ME±SD)
[99] Mousavi et al. (2019)	MIMIC database	1323 beats	Random forest (SBP) AdaBoost (DBP)	10 fold cross validation	0.91	0.90	0.19 ± 4.17 (ME±SD)	-0.05 ± 8.90 (ME±SD)
[71] Wang et al. (2018)	MIMIC database	58795 beats	ANN	Training and testing	N/A	N/A	4.02 ± 2.79 (MAE±SD)	2.27 ± 1.82 (MAE±SD)
[70] Xing et al. (2016)	MIMIC database	175477 windows	ANN	Training and testing	N/A	N/A	0.06 ± 7.08 (ME±SD)	0.01 ± 4.66 (ME±SD)
[90] Khalid et al. (2018)	University of Queensland database	8133 windows	Regression tree	10 fold cross validation	N/A	N/A	-0.1 ± 6.5 (ME±SD)	-0.6 ± 5.2 (ME±SD)
[60] Duan et al. (2016)	University of Queensland database	7678 windows	SVR	10 fold cross validation	N/A	N/A	4.77 ± 7.68 (MAE±SD)	3.67 ± 5.69 (MAE±SD)

For performance comparison, in Table 4.3 we summarized general information and best

recorded estimation accuracy of 7 related works published within the last four years that estimate BP from PPG, similar to this work. These studies have been discussed in Chapter 3 with more details. Some of the listed studies have specific comparability to us. For example, [107] shares the same dataset with this work but utilized the entire dataset; [90] and [60] use dataset with similar size to us; [99] used a way of beat-to-beat window segmentation method similar to this work. It can be seen that our work outperformed [107] in SBP and DBP correlation, [107, 100] in SBP error, and [107, 100, 90, 60, 99] in DBP error.

4.5 Conclusions

In this chapter, we presented a new approach for effectively transforming the temporal domain information into image domain, which leads to development of a new data-driven framework for cuff-less BP monitoring with deep learning models. Our results demonstrate that the proposed method achieves our desire of using only PPG signal, eliminating the need for individual calibration, and reducing the number of parameters to be trained by utilizing transfer learning with pre-trained deep network. The method provides accurate BP estimation with only one PPG beat, while keeping its simplicity of implementation and requiring training only one dense layer. Our proposed work is thereby a suitable candidate for cuff-less and continuous BP monitoring.

Chapter 5

Conclusions

Cuff-less blood pressure (BP) estimation methods are the next generation replacement of traditional cuff-based BP methods, which can enable the possibility of 24-hour continuous and disturbance-free personalized healthcare. Most of such methods estimate BP values from physiological signals based on pulse wave velocity-BP relationship or pulse volume-BP relationship. There are two classes of methods for cuff-less BP estimation: model-driven methods which use parametric models, and data-driven methods which rely on extracted features from physiological recordings and machine learning. In this thesis, a thorough review of both classes of methods was presented to offer a summary of existing works, with respect to the derivation and application of parametric models and the feature pool of data-driven methods. It is worth mentioning that existing studies are still in preliminary stages, where model accuracy and reliability over time and population are only validated within limited range.

While model-driven methods offer advantages of simplicity and good interpretability, it seems that data-driven methods are more promising in the future due to their accuracy and good generalizability over the population. To reduce the computational budget of existing data-driven methods, we proposed a novel method that effectively transforms the temporal domain information in the PPG signal into image domain using visibility graph, thereby, enabling the machine learning model to be a pretrained deep convolutional image classification network instead of a model to be trained from scratch. The proposed method offers comparable or better results when compared to other methods with much more complicated and costly training process.

Future work should be focused on further improving and extending data-driven methods based on semi-defined or whole-based features combined with deep learning. Interpretability problem of machine learning models is also a valuable topic in future studies for improving our

understanding of feature predictability and model reliability.

References

[1] S. S. Lim, T. Vos, A. D. Flaxman, G. Danaei, K. Shibuya, H. Adair-Rohani, M. A. AlMazroa, M. Amann, H. R. Anderson, K. G. Andrews *et al.*, "A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the global burden of disease study 2010," *The lancet*, vol. 380, no. 9859, pp. 2224–2260, 2012.

[2] H. Girouard and C. Iadecola, "Neurovascular coupling in the normal brain and in hypertension, stroke, and alzheimer disease," *Journal of applied physiology*, vol. 100, no. 1, pp. 328–335, 2006.

[3] T. W. Gress, F. J. Nieto, E. Shahar, M. R. Wofford, and F. L. Brancati, "Hypertension and antihypertensive therapy as risk factors for type 2 diabetes mellitus," *New England Journal of Medicine*, vol. 342, no. 13, pp. 905–912, 2000.

[4] C. D. Fryar, Y. Ostchega, C. M. Hales, G. Zhang, and D. Kruszon-Moran, "Hypertension prevalence and control among adults: United states, 2015-2016," 2017.

[5] "Healthy people 2020." [Online]. Available: https://www.healthypeople.gov/2020/default

[6] T. Ohkubo, Y. Imai, I. Tsuji, K. Nagai, J. Kato, N. Kikuchi, A. Nishiyama, A. Aihara, M. Sekino, M. Kikuya *et al.*, "Home blood pressure measurement has a stronger predictive power for mortality than does screening blood pressure measurement: a population-based observation in ohasama, japan," *Journal of hypertension*, vol. 16, no. 7, pp. 971–975, 1998.

[7] F. P. Cappuccio, S. M. Kerry, L. Forbes, and A. Donald, "Blood pressure control by home monitoring: meta-analysis of randomised trials," *Bmj*, vol. 329, no. 7458, p. 145, 2004.

[8] D. Edmonds, E. Foerster, H. Groth, P. Greminger, W. Siegenthaler, and W. Vetter, "Does self-measurement of blood pressure improve patient compliance in hypertension?" *Journal of hypertension. Supplement : official journal of the International Society of Hypertension*, vol. 3, no. 1, p. S31—4, April 1985. [Online]. Available: http://europepmc.org/abstract/MED/3870467

[9] E. O'brien, R. Asmar, L. Beilin, Y. Imai, J.-M. Mallion, G. Mancia, T. Mengden, M. Myers, P. Padfield, P. Palatini *et al.*, "European society of hypertension recommendations for conventional, ambulatory and home blood pressure measurement," *Journal of hypertension*, vol. 21, no. 5, pp. 821–848, 2003.

[10] J. Levy, L. M. Gerber, X. Wu, and S. J. Mann, "Nonadherence to recommended guidelines for blood pressure measurement," *The Journal of Clinical Hypertension*, vol. 18, no. 11, pp. 1157–1161, 2016.

[11] D. Perloff, C. Grim, J. Flack, E. D. Frohlich, M. Hill, M. McDonald, and B. Z. Morgenstern, "Human blood pressure determination by sphygmomanometry." *Circulation*, vol. 88, no. 5, pp. 2460–2470, 1993.

[12] E. O'brien, B. Waeber, G. Parati, J. Staessen, and M. G. Myers, "Blood pressure measuring devices: recommendations of the european society of hypertension," *BMJ: British Medical Journal*, vol. 322, no. 7285, p. 531, 2001.

[13] E. Walma, C. Van Dooren, E. Van der Does, A. Prins, P. Mulder, and A. Hoes, "Accuracy of an oscillometric automatic blood pressure device: the omron hem403c." *Journal of human hypertension*, vol. 9, no. 3, pp. 169–174, 1995.

[14] "Omron 10 series® wireless upper arm blood pressure monitor." [Online]. Available: https://omronhealthcare.com/products/10-series-wireless-upper-arm-blood-pressure-monitor-bp7450/

[15] T. M. Seeberg, J. G. Orr, H. Opsahl, H. Austad, M. H. Røed, S. H. Dalgard, D. Houghton, D. E. Jones, and F. Strisland, "A novel method for continuous, noninvasive, cuff-less measurement of blood pressure: evaluation in patients with nonalcoholic fatty liver disease," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 7, pp. 1469–1478, 2016.

[16] M. Saeed, M. Villarroel, A. T. Reisner, G. Clifford, L.-W. Lehman, G. Moody, T. Heldt, T. H. Kyaw, B. Moody, and R. G. Mark, "Multiparameter intelligent monitoring in intensive care ii (mimic-ii): a public-access intensive care unit database," *Critical care medicine*, vol. 39, no. 5, p. 952, 2011.

[17] "Pulse wave velocity." [Online]. Available: https://www.datasci.com/solutions/cardiovascular/pulse-wave-velocity-(pwv)

[18] J. C. Bramwell and A. Hill, "Velocity of transmission of the pulse-wave: and elasticity of arteries," *The Lancet*, vol. 199, no. 5149, pp. 891–892, 1922.

[19] "Electrocardiography," Sep 2020. [Online]. Available: https://en.wikipedia.org/wiki/Electrocardiography

[20] L. S. Lilly, *Pathophysiology of heart disease: a collaborative project of medical students and faculty*. Lippincott Williams & Wilkins, 2012.

[21] K. H. Shelley, "Photoplethysmography: beyond the calculation of arterial oxygen saturation and heart rate," *Anesthesia & Analgesia*, vol. 105, no. 6, pp. S31–S36, 2007.

[22] A. Esmaili, M. Kachuee, and M. Shabany, "Nonlinear cuffless blood pressure estimation of healthy subjects using pulse transit time and arrival time," *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 12, pp. 3299–3308, 2017.

[23] M. Kachuee, M. M. Kiani, H. Mohammadzade, and M. Shabany, "Cuffless blood pressure estimation algorithms for continuous health-care monitoring," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 4, pp. 859–869, 2016.

[24] 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," in *2005*

IEEE Engineering in Medicine and Biology 27th Annual Conference. IEEE, 2006, pp. 996–999.

[25] D. B. Newlin and R. W. Levenson, "Pre-ejection period: Measuring beta-adrenergic influences upon the heart," *Psychophysiology*, vol. 16, no. 6, pp. 546–552, 1979.

[26] K. Visser, G. Mook, E. Van der Wall, and W. Zijlstra, "Theory of the determination of systolic time intervals by impedance cardiography," *Biological psychology*, vol. 36, no. 1-2, pp. 43–50, 1993.

[27] M. Y. M. Wong, E. Pickwell-MacPherson, Y. T. Zhang, and J. C. Cheng, "The effects of pre-ejection period on post-exercise systolic blood pressure estimation using the pulse arrival time technique," *European journal of applied physiology*, vol. 111, no. 1, pp. 135–144, 2011.

[28] J. Proença, J. Muehlsteff, X. Aubert, and P. Carvalho, "Is pulse transit time a good indicator of blood pressure changes during short physical exercise in a young population?" in *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*. IEEE, 2010, pp. 598–601.

[29] J. Muehlsteff, X. Aubert, and M. Schuett, "Cuffless estimation of systolic blood pressure for short effort bicycle tests: the prominent role of the pre-ejection period," in *2006 International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 2006, pp. 5088–5092.

[30] Y. Chen, C. Wen, G. Tao, M. Bi, and G. Li, "Continuous and noninvasive blood pressure measurement: a novel modeling methodology of the relationship between blood pressure and pulse wave velocity," *Annals of biomedical engineering*, vol. 37, no. 11, pp. 2222–2233, 2009.

[31] J. W. Remington, "The physiology of the aorta and major arteries," *Handbook of physiology*, vol. 2, p. 799, 1963.

[32] D. Hughes, C. F. Babbs, L. Geddes, and J. Bourland, "Measurements of young's modulus of elasticity of the canine aorta with ultrasound," *Ultrasonic Imaging*, vol. 1, no. 4, pp. 356–367, 1979.

[33] C. C. Poon, Y.-T. Zhang, G. Wong, and W. S. Poon, "The beat-to-beat relationship between pulse transit time and systolic blood pressure," in *2008 International Conference on Information Technology and Applications in Biomedicine*. IEEE, 2008, pp. 342–343.

[34] C.-S. Kim, A. M. Carek, R. Mukkamala, O. T. Inan, and J.-O. Hahn, "Ballistocardiogram as proximal timing reference for pulse transit time measurement: Potential for cuffless blood pressure monitoring," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 11, pp. 2657–2664, 2015.

[35] E. Kaniusas, H. Pfutzner, L. Mehnen, J. Kosel, C. Tellez-Blanco, G. Varoneckas, A. Alonderis, T. Meydan, M. Vázquez, M. Rohn *et al.*, "Method for continuous nondisturbing monitoring of blood pressure by magnetoelastic skin curvature sensor and ecg," *IEEE Sensors Journal*, vol. 6, no. 3, pp. 819–828, 2006.

[36] N. Westerhof, N. Stergiopulos, M. I. Noble, and B. E. Westerhof, *Snapshots of hemodynamics: an aid for clinical research and graduate education*. Springer, 2018.

[37] P. Fung, G. Dumont, C. Ries, C. Mott, and M. Ansermino, "Continuous noninvasive blood pressure measurement by pulse transit time," in *The 26th annual international conference of the IEEE engineering in medicine and biology society*, vol. 1. IEEE, 2004, pp. 738–741.

[38] G. Elert, "Density: The physics hyper text book," 1998.

[39] M. Theodor, D. Ruh, J. Fiala, K. Förster, C. Heilmann, Y. Manoli, F. Beyersdorf, H. Zappe, and A. Seifert, "Subcutaneous blood pressure monitoring with an implantable optical sensor," *Biomedical microdevices*, vol. 15, no. 5, pp. 811–820, 2013.

[40] T. H. Huynh, R. Jafari, and W.-Y. Chung, "Noninvasive cuffless blood pressure estimation using pulse transit time and impedance plethysmography," *IEEE Transactions on Biomedical Engineering*, vol. 66, no. 4, pp. 967–976, 2018.

[41] G. Langewouters, K. Wesseling, and W. Goedhard, "The static elastic properties of 45 human thoracic and 20 abdominal aortas in vitro and the parameters of a new model," *Journal of biomechanics*, vol. 17, no. 6, pp. 425–435, 1984.

[42] W. Chen, T. Kobayashi, S. Ichikawa, Y. Takeuchi, and T. Togawa, "Continuous estimation of systolic blood pressure using the pulse arrival time and intermittent calibration," *Medical and Biological Engineering and Computing*, vol. 38, no. 5, pp. 569–574, 2000.

[43] R. Payne, C. Symeonides, D. Webb, and S. Maxwell, "Pulse transit time measured from the ecg: an unreliable marker of beat-to-beat blood pressure," *Journal of Applied Physiology*, vol. 100, no. 1, pp. 136–141, 2006.

[44] M. Y.-M. Wong, C. C.-Y. Poon, and Y.-T. Zhang, "An evaluation of the cuffless blood pressure estimation based on pulse transit time technique: a half year study on normotensive subjects," *Cardiovascular Engineering*, vol. 9, no. 1, pp. 32–38, 2009.

[45] S. L.-O. Martin, A. M. Carek, C.-S. Kim, H. Ashouri, O. T. Inan, J.-O. Hahn, and R. Mukkamala, "Weighing scale-based pulse transit time is a superior marker of blood pressure than conventional pulse arrival time," *Scientific reports*, vol. 6, p. 39273, 2016.

[46] L. Geddes, M. Voelz, C. Babbs, J. Bourland, and W. Tacker, "Pulse transit time as an indicator of arterial blood pressure," *psychophysiology*, vol. 18, no. 1, pp. 71–74, 1981.

[47] B. Gribbin, A. Steptoe, and P. Sleight, "Pulse wave velocity as a measure of blood pressure change," *Psychophysiology*, vol. 13, no. 1, pp. 86–90, 1976.

[48] R. Mukkamala, J.-O. Hahn, O. T. Inan, L. K. Mestha, C.-S. Kim, H. Töreyin, and S. Kyal, "Toward ubiquitous blood pressure monitoring via pulse transit time: theory and practice," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 8, pp. 1879–1901, 2015.

[49] M. Mase, W. Mattei, R. Cucino, L. Faes, and G. Nollo, "Feasibility of cuff-free measurement of systolic and diastolic arterial blood pressure," *Journal of electrocardiology*, vol. 44, no. 2, pp. 201–207, 2011.

[50] H. Gesche, D. Grosskurth, G. Küchler, and A. Patzak, "Continuous blood pressure measurement by using the pulse transit time: comparison to a cuff-based method," *European journal of applied physiology*, vol. 112, no. 1, pp. 309–315, 2012.

[51] H. T. Ma, "A blood pressure monitoring method for stroke management," *BioMed research international*, vol. 2014, 2014.

[52] Y. Zheng, B. P. Yan, Y. Zhang, and C. C. Y. Poon, "An armband wearable device for overnight and cuff-less blood pressure measurement," *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 7, pp. 2179–2186, July 2014.

[53] C. Poon and Y. Zhang, "Cuff-less and noninvasive measurements of arterial blood pressure by pulse transit time," in *2005 IEEE engineering in medicine and biology 27th annual conference*. IEEE, 2006, pp. 5877–5880.

[54] B. McCarthy, C. Vaughan, B. O'flynn, A. Mathewson, and C. Ó. Mathúna, "An examination of calibration intervals required for accurately tracking blood pressure using pulse transit time algorithms," *Journal of human hypertension*, vol. 27, no. 12, pp. 744–750, 2013.

[55] X.-R. Ding, Y.-T. Zhang, J. Liu, W.-X. Dai, and H. K. Tsang, "Continuous cuffless blood pressure estimation using pulse transit time and photoplethysmogram intensity ratio," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 5, pp. 964–972, 2015.

[56] T. Wibmer, K. Doering, C. Kropf-Sanchen, S. Rüdiger, I. Blanta, K. Stoiber, W. Rottbauer, and C. Schumann, "Pulse transit time and blood pressure during cardiopulmonary exercise tests." *Physiological research*, vol. 63, no. 3, 2014.

[57] E. Balestrieri and S. Rapuano, "Calibration of automated non invasive blood pressure measurement devices," in *Advances in Biomedical Sensing, Measurements, Instrumentation and Systems*. Springer, 2010, pp. 281–304.

[58] I. S. Association *et al.*, "Ieee standard for wearable cuffless blood pressure measuring devices," *IEEE Std*, pp. 1708–2014, 2014.

[59] E. Monte-Moreno, "Non-invasive estimate of blood glucose and blood pressure from a photoplethysmograph by means of machine learning techniques," *Artificial intelligence in medicine*, vol. 53, no. 2, pp. 127–138, 2011.

[60] K. Duan, Z. Qian, M. Atef, and G. Wang, "A feature exploration methodology for learning based cuffless blood pressure measurement using photoplethysmography," in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2016, pp. 6385–6388.

[61] M. Liu, L.-M. Po, and H. Fu, "Cuffless blood pressure estimation based on photoplethysmography signal and its second derivative," *International Journal of Computer Theory and Engineering*, vol. 9, no. 3, p. 202, 2017.

[62] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An introduction to statistical learning*. Springer, 2013, vol. 112.

[63] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.

[64] Y. Freund, R. E. Schapire *et al.*, "Experiments with a new boosting algorithm," in *icml*, vol. 96. Citeseer, 1996, pp. 148–156.

[65] H. Drucker, "Improving regressors using boosting techniques," in *ICML*, vol. 97, 1997, pp. 107–115.

[66] D. P. Solomatine and D. L. Shrestha, "Adaboost. rt: a boosting algorithm for regression problems," in *2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No. 04CH37541)*, vol. 2. IEEE, 2004, pp. 1163–1168.

[67] C. M. Bishop, *Pattern recognition and machine learning*. springer, 2006.

[68] G. Cybenko, "Approximation by superpositions of a sigmoidal function," *Mathematics of control, signals and systems*, vol. 2, no. 4, pp. 303–314, 1989.

[69] M. A. Nielsen, *Neural networks and deep learning*. Determination press San Francisco, CA, 2015, vol. 2018.

[70] X. Xing and M. Sun, "Optical blood pressure estimation with photoplethysmography and fft-based neural networks," *Biomedical optics express*, vol. 7, no. 8, pp. 3007–3020, 2016.

[71] L. Wang, W. Zhou, Y. Xing, and X. Zhou, "A novel neural network model for blood pressure estimation using photoplethesmography without electrocardiogram," *Journal of healthcare engineering*, vol. 2018, 2018.

[72] J. Y. Kim, B. H. Cho, S. M. Im, M. J. Jeon, I. Y. Kim, and S. I. Kim, "Comparative study on artificial neural network with multiple regressions for continuous estimation of blood pressure," in *2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*. IEEE, 2006, pp. 6942–6945.

[73] Z. Xu, J. Liu, X. Chen, Y. Wang, and Z. Zhao, "Continuous blood pressure estimation based on multiple parameters from eletrocardiogram and photoplethysmogram by back-propagation neural network," *Computers in Industry*, vol. 89, pp. 50–59, 2017.

[74] J. Solà and R. Delgado-Gonzalo, *The Handbook of Cuffless Blood Pressure Monitoring: A Practical Guide for Clinicians, Researchers, and Engineers*. Springer Nature, 2019.

[75] Y. Kurylyak, F. Lamonaca, and D. Grimaldi, "A neural network-based method for continuous blood pressure estimation from a ppg signal," in *2013 IEEE International instrumentation and measurement technology conference (I2MTC)*. IEEE, 2013, pp. 280–283.

[76] F. Miao, B. Wen, Z. Hu, G. Fortino, X.-P. Wang, Z.-D. Liu, M. Tang, and Y. Li, "Continuous blood pressure measurement from one-channel electrocardiogram signal using deep-learning techniques," *Artificial Intelligence in Medicine*, vol. 108, p. 101919, 2020.

[77] W.-H. Lin, F. Chen, Y. Geng, N. Ji, P. Fang, and G. Li, "Towards accurate estimation of cuffless and continuous blood pressure using multi-order derivative and multivariate photoplethysmogram features," *Biomedical Signal Processing and Control*, vol. 63, p. 102198, 2021.

[78] G. V. Trunk, "A problem of dimensionality: A simple example," *IEEE Transactions on pattern analysis and machine intelligence*, no. 3, pp. 306–307, 1979.

[79] G. Hughes, "On the mean accuracy of statistical pattern recognizers," *IEEE transactions on information theory*, vol. 14, no. 1, pp. 55–63, 1968.

[80] R. E. Bellman, *Adaptive control processes: a guided tour*. Princeton university press, 2015, vol. 2045.

[81] M. A. Schroeder, J. Lander, and S. Levine-Silverman, "Diagnosing and dealing with multicollinearity," *Western journal of nursing research*, vol. 12, no. 2, pp. 175–187, 1990.

[82] B. C. Ross, "Mutual information between discrete and continuous data sets," *PloS one*, vol. 9, no. 2, p. e87357, 2014.

[83] T. M. Cover, *Elements of information theory*. John Wiley & Sons, 1999.

[84] D. N. Reshef, Y. A. Reshef, H. K. Finucane, S. R. Grossman, G. McVean, P. J. Turnbaugh, E. S. Lander, M. Mitzenmacher, and P. C. Sabeti, "Detecting novel associations in large data sets," *science*, vol. 334, no. 6062, pp. 1518–1524, 2011.

[85] H. Peng, F. Long, and C. Ding, "Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 27, no. 8, pp. 1226–1238, 2005.

[86] G. Langewouters, A. Zwart, R. Busse, and K. Wesseling, "Pressure-diameter relationships of segments of human finger arteries," *Clinical physics and physiological measurement*, vol. 7, no. 1, p. 43, 1986.

[87] G. Martínez, N. Howard, D. Abbott, K. Lim, R. Ward, and M. Elgendi, "Can photoplethysmography replace arterial blood pressure in the assessment of blood pressure?" *Journal of clinical medicine*, vol. 7, no. 10, p. 316, 2018.

[88] A. A. Awad, M. A. M. Ghobashy, R. G. Stout, D. G. Silverman, and K. H. Shelley, "How does the plethysmogram derived from the pulse oximeter relate to arterial blood pressure in coronary artery bypass graft patients?" *Anesthesia & Analgesia*, vol. 93, no. 6, pp. 1466–1471, 2001.

[89] X. Teng and Y. Zhang, "Continuous and noninvasive estimation of arterial blood pressure using a photoplethysmographic approach," in *Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE Cat. No. 03CH37439)*, vol. 4. IEEE, 2003, pp. 3153–3156.

[90] S. G. Khalid, J. Zhang, F. Chen, and D. Zheng, "Blood pressure estimation using photoplethysmography only: comparison between different machine learning approaches," *Journal of healthcare engineering*, vol. 2018, 2018.

[91] M. Elgendi, "On the analysis of fingertip photoplethysmogram signals," *Current cardiology reviews*, vol. 8, no. 1, pp. 14–25, 2012.

[92] K. Takazawa, N. Tanaka, M. Fujita, O. Matsuoka, T. Saiki, M. Aikawa, S. Tamura, and C. Ibukiyama, "Assessment of vasoactive agents and vascular aging by the second derivative of photoplethysmogram waveform," *Hypertension*, vol. 32, no. 2, pp. 365–370, 1998.

[93] N. Westerhof, P. Sipkema, G. V. D. Bos, and G. Elzinga, "Forward and backward waves in the arterial system," *Cardiovascular research*, vol. 6, no. 6, pp. 648–656, 1972.

[94] L. Wang, E. Pickwell-MacPherson, Y. Liang, and Y. T. Zhang, "Noninvasive cardiac output estimation using a novel photoplethysmogram index," in *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 2009, pp. 1746–1749.

[95] F. Miao, N. Fu, Y.-T. Zhang, X.-R. Ding, X. Hong, Q. He, and Y. Li, "A novel continuous blood pressure estimation approach based on data mining techniques," *IEEE journal of biomedical and health informatics*, vol. 21, no. 6, pp. 1730–1740, 2017.

[96] P. Shaltis, A. Reisner, and H. Asada, "Calibration of the photoplethysmogram to arterial blood pressure: capabilities and limitations for continuous pressure monitoring," in *2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*. IEEE, 2006, pp. 3970–3973.

[97] S. Suzuki and K. Oguri, "Cuffless and non-invasive systolic blood pressure estimation for aged class by using a photoplethysmograph," in *2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 2008, pp. 1327–1330.

[98] K. Atomi, H. Kawanaka, M. Bhuiyan, K. Oguri *et al.*, "Cuffless blood pressure estimation based on data-oriented continuous health monitoring system," *Computational and mathematical methods in medicine*, vol. 2017, 2017.

[99] S. S. Mousavi, M. Firouzmand, M. Charmi, M. Hemmati, M. Moghadam, and Y. Ghorbani, "Blood pressure estimation from appropriate and inappropriate ppg signals using a whole-based method," *Biomedical Signal Processing and Control*, vol. 47, pp. 196–206, 2019.

[100] X. Xing, Z. Ma, M. Zhang, Y. Zhou, W. Dong, and M. Song, "An unobtrusive and calibration-free blood pressure estimation method using photoplethysmography and biometrics," *Scientific reports*, vol. 9, no. 1, p. 8611, 2019.

[101] M. Elgendi, "Optimal signal quality index for photoplethysmogram signals," *Bioengineering*, vol. 3, no. 4, p. 21, 2016.

[102] R. Krishnan, B. Natarajan, and S. Warren, "Two-stage approach for detection and reduction of motion artifacts in photoplethysmographic data," *IEEE transactions on biomedical engineering*, vol. 57, no. 8, pp. 1867–1876, 2010.

[103] N. Selvaraj, Y. Mendelson, K. H. Shelley, D. G. Silverman, and K. H. Chon, "Statistical approach for the detection of motion/noise artifacts in photoplethysmogram," in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 2011, pp. 4972–4975.

[104] C.-h. Chen, *Signal processing handbook*. CRC Press, 1988, vol. 51.

[105] R. B. Dunn, T. F. Quatieri, and J. F. Kaiser, "Detection of transient signals using the energy operator," in *1993 IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 3. IEEE, 1993, pp. 145–148.

[106] A. Gaurav, M. Maheedhar, V. N. Tiwari, and R. Narayanan, "Cuff-less ppg based continuous blood pressure monitoring—a smartphone based approach," in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2016, pp. 607–610.

[107] N. Hasanzadeh, M. M. Ahmadi, and H. Mohammadzade, "Blood pressure estimation using photoplethysmogram signal and its morphological features," *IEEE Sensors Journal*, 2019.

[108] Y. Li, Z. Wang, L. Zhang, X. Yang, and J. Song, "Characters available in photoplethysmogram for blood pressure estimation: beyond the pulse transit time," *Australasian physical & engineering sciences in medicine*, vol. 37, no. 2, pp. 367–376, 2014.

[109] S. C. Millasseau, F. G. Guigui, R. P. Kelly, K. Prasad, J. R. Cockcroft, J. M. Ritter, and P. J. Chowienczyk, "Noninvasive assessment of the digital volume pulse: comparison with the peripheral pressure pulse," *Hypertension*, vol. 36, no. 6, pp. 952–956, 2000.

[110] H. Hsiu, C. Hsu, and T. Wu, "Effects of different contacting pressure on the transfer function between finger photoplethysmographic and radial blood pressure waveforms," *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, vol. 225, no. 6, pp. 575–583, 2011.

[111] J. Allen and A. Murray, "Modelling the relationship between peripheral blood pressure and blood volume pulses using linear and neural network system identification techniques," *Physiological Measurement*, vol. 20, no. 3, p. 287, 1999.

[112] D. J. Thomson, "Spectrum estimation and harmonic analysis," *Proceedings of the IEEE*, vol. 70, no. 9, pp. 1055–1096, 1982.

[113] S. C. Gao, P. Wittek, L. Zhao, and W. J. Jiang, "Data-driven estimation of blood pressure using photoplethysmographic signals," in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2016, pp. 766–769.

[114] Z. Wang and Y. Zhang, "A novel frequency domain method for estimating blood pressure from photoplethysmogram," in *Proceedings of the 9th International Conference on Signal Processing Systems*, 2017, pp. 201–206.

[115] M. Simjanoska, M. Gjoreski, M. Gams, and A. Madevska Bogdanova, "Non-invasive blood pressure estimation from ecg using machine learning techniques," *Sensors*, vol. 18, no. 4, p. 1160, 2018.

[116] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[117] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[118] L. Lacasa, B. Luque, F. Ballesteros, J. Luque, and J. C. Nuno, "From time series to complex networks: The visibility graph," *Proceedings of the National Academy of Sciences*, vol. 105, no. 13, pp. 4972–4975, 2008.

[119] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 2818–2826.

[120] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A Large-Scale Hierarchical Image Database," in *CVPR09*, 2009.

[121] M. Kachuee, M. M. Kiani, H. Mohammadzade, and M. Shabany, "Cuff-less high-accuracy calibration-free blood pressure estimation using pulse transit time," in *2015 IEEE international symposium on circuits and systems (ISCAS)*. IEEE, 2015, pp. 1006–1009.

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

[123] C.-H. Chen, E. Nevo, B. Fetics, P. H. Pak, F. C. Yin, W. L. Maughan, and D. A. Kass, "Estimation of central aortic pressure waveform by mathematical transformation of radial tonometry pressure," *Circulation*, vol. 95, no. 7, pp. 1827–1836, 1997.

[124] S. Sannino, S. Stramaglia, L. Lacasa, and D. Marinazzo, "Visibility graphs for fmri data: multiplex temporal graphs and their modulations across resting state networks," *Network Neuroscience*, 2017.

[125] L. Zhu, C. R. Lee, D. J. Margolis, and L. Najafizadeh, "Decoding cortical brain states from widefield calcium imaging data using visibility graph," *Biomedical optics express*, vol. 9, no. 7, pp. 3017–3036, 2018.

[126] L. Zhu, S. Haghani, and L. Najafizadeh, "On fractality of functional near-infrared spectroscopy signals: analysis and applications," *Neurophotonics*, vol. 7, no. 2, p. 025001, 2020.

[127] ——, "Spatiotemporal characterization of brain function via multiplex visibility graph," in *Microscopy Histopathology and Analytics*. Optical Society of America, 2018, pp. JTh3A–54.

Vita

Weinan Wang

2019 B. Sc. in Mechanical Design and Automation, University of Electronic Science and Technology of China

2019-2020 Research assistant in Department of Electrical and Computer Engineering, Dr. Laleh Najafizadeh, Rutgers University