Chapter 1

Introduction

1.1 Background

Health is an important aspect of living a happy life, especially due to the pandemic, we are taking more care of our health. We check our vital signs and body signs to monitor our health status in different scenarios. For example, our blood pressure and blood oxygen level [114] are measured when we visit the clinics. We check our heart rate and/or breathing rate during exercise via contact sensing, such as fitness watch [101], or contactless sensing [25,27,87,171]. Those are all applications of digital health in our daily lives. One of the key technologies behind those applications is photoplethysmography.

1.1.1 Photoplethysmography (PPG) and Remote PPG

Photoplethysmography (PPG) is a low-cost, user-friendly, and noninvasive optical technique that measures the periodic change of blood volume in the microvascular bed of tissue in the pace of heartbeat, and can be obtained from an optoelectronic device clipped to a person's fingertip or by other wearable devices such as an Apple Watch, as shown in

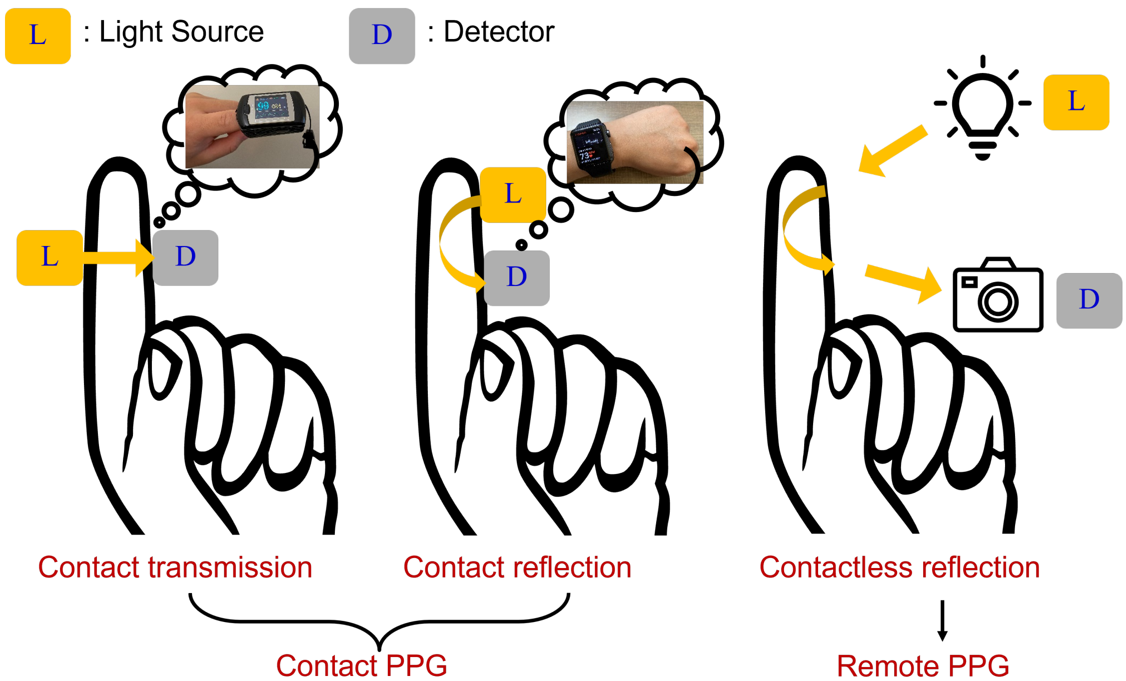

Figure 1.1: PPG measurement in both contact and contactless methods.

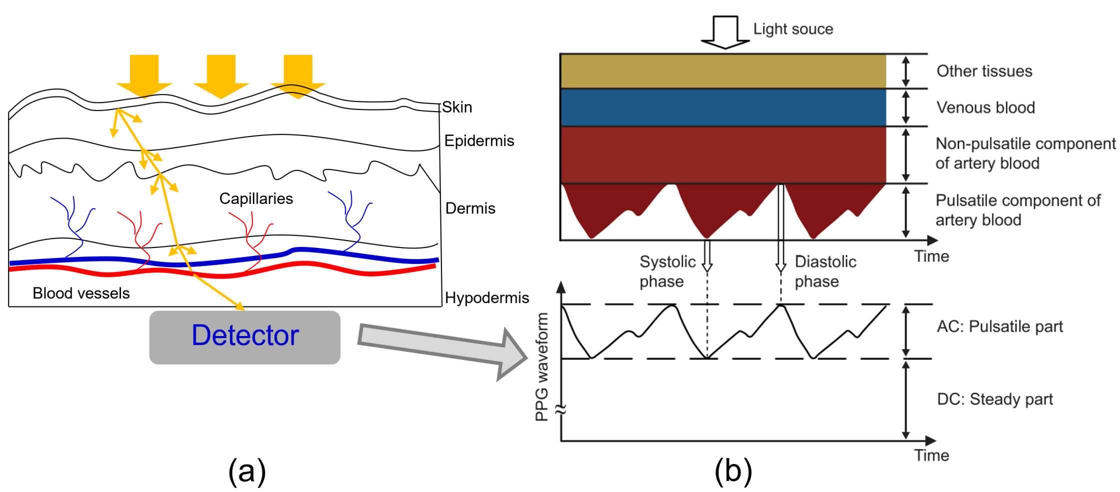

Figure 1.2: (a) Anatomical cross-section structure of human skin tissues and transmitted light captured by a detector when the skin is illuminated by a light source. (b) Variations in light attenuation by tissue, illustrating the rhythmic effect of arterial pulsation. Figures modified from [156] and [135].

the "contact PPG" part of Fig. 1.1. PPG has become a common modality for heart activity monitoring in clinics, hospitals, and homes for healthcare and fitness purposes [7]. As illustrated in Fig. 1.2(a), the measurement of PPG requires a light source to illuminate the tissue and a photodetector to receive the light transmitted or reflected by the tissue. During each cardiac cycle, the blood is first pumped into the body so that the blood volume increases in the capillaries in the skin, which causes increased light absorption. Then as the blood travels back to the heart via the venous network, the light absorption at the capillaries decreases. Therefore, as shown in Fig. 1.2(b) the PPG signal is composed of the 'AC' component which reflects the cardiac change with each heartbeat, and the 'DC' component which contains the information including respiration, venous flow, and thermoregulation [135].

To facilitate long-term and comfortable sensing, advances in video signal processing, computer vision, and artificial intelligence have opened up opportunities to use a camera captured video to monitor a person's health related vital signs remotely as illustrated in the "remote PPG" part of Fig. 1.1. This technology is commonly referred to as remote PPG (or rPPG), which is first proposed by Verkruysseet et al. [152]. The basic principle of rPPG is to illuminate tissue and to use the camera as the receiving sensor to capture the re-emitted light from the tissue. The captured video contains the periodically variational light absorption in the microvascular bed underneath the skin, thus can convey information about the cardiovascular system. The rPPG has been utilized to monitor important physiological parameters, including heart rate [37, 91, 148, 152, 157, 176], breathing rate [27, 28, 115], heart rate variability [47, 69, 100, 115], blood pressure [71], and blood oxygen saturation [80].

In this dissertation, we study the following two application directions of contact and contactless PPG, one is electrocardiogram waveform inference from contact PPG waveform and its application and contribution to the emerging digital twin technology (Ch. 2 and Ch. 3), and the other is noncontact blood oxygen saturation measurement using remote PPG captured by RGB cameras (Ch. 4 and Ch. 5).

1.1.2 Electrocardiogram (ECG) and Its Physiological and Signal Relation With PPG

Cardiovascular diseases (CVDs) have become a leading cause of death globally. From alarming reports of the World Health Organization, an estimated 17.9 million people died from CVDs in 2019, representing 32% of all global deaths [20]. However, some CVDs, such as heart muscle dysfunction, show no obvious symptoms in the early stage. The presence of symptoms usually indicates the onset of heart failure. A study conducted on the aged population shows that around one-third to one-half of heart attacks are clinically unrecognized [38]. The unawareness of diseases makes some patients miss the opportunities of receiving an early medical intervention.

Electrocardiogram (ECG) is a widely-used clinical gold-standard for the detection of irregular heart rhythms, cardiovascular diseases, and determination of how certain heart disease treatments are working in a painless and noninvasive manner. By measuring the electrical activity of the heartbeat and conveying information regarding heart functionality, timely and continuous ECG monitoring is proven to be beneficial for the early detection of CVDs [120, 134]. The clinical standard ECG measurement device is the 12-lead

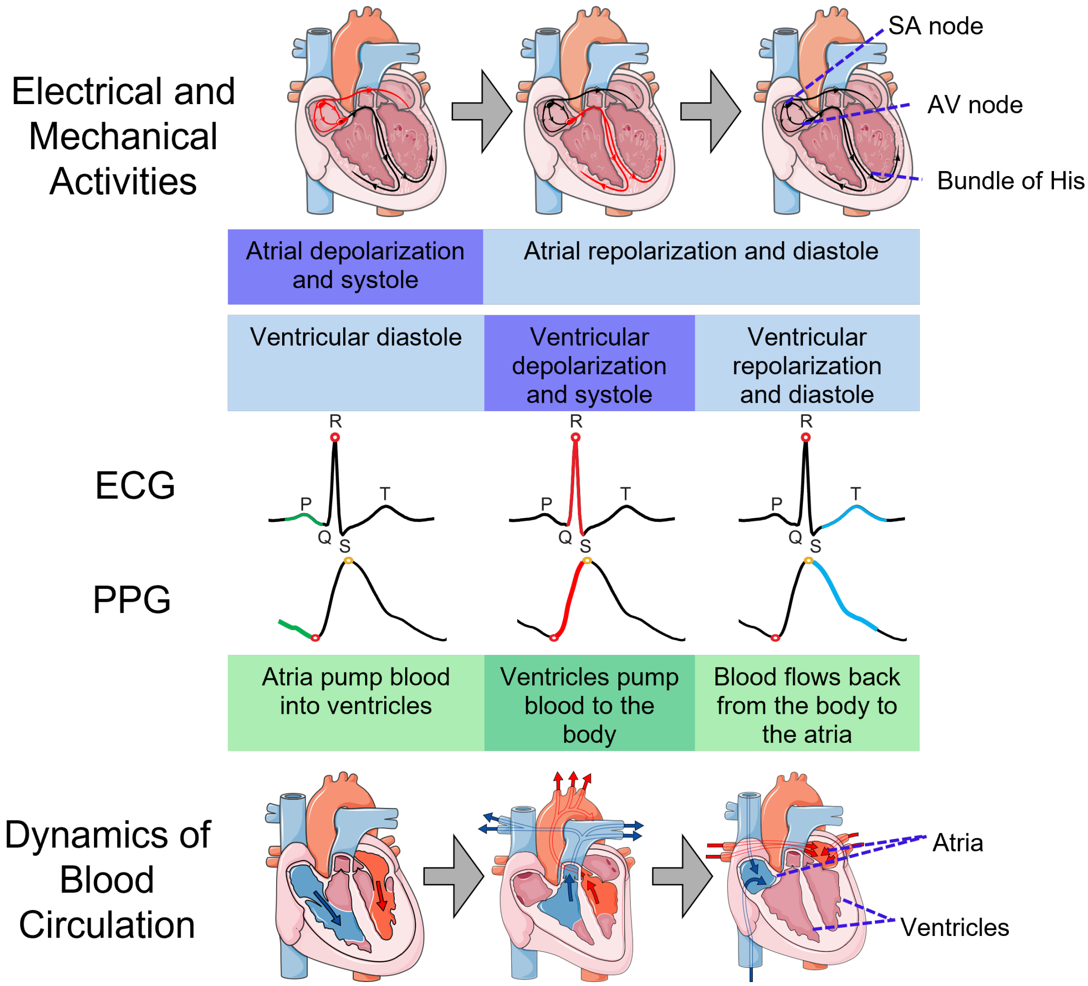

Figure 1.3: Association between the electrical and mechanical activities of the heart and the simultaneous blood flow dynamics represented by ECG and PPG, respectively. The heart images are adopted from Servier Medical Art [126].

ECG monitor that is commonly seen in the health care provider's office, a clinic, or a hospital room. It sensitively picks up electrical potential changes spread in the heart from the skin during a cardiac cycle from 12 different perspectives by attaching ten electrodes with sticky patches to each of the limbs and six positions across the chest of the patient.

A normal ECG waveform and the corresponding electrical and mechanical activities in the heart are shown in Fig. 1.3. Specifically, different phases in one cardiac cycle progress as follows [8,60]: A cardiac cycle begins with the atria depolarization and systole triggered by the heart's pacemaker at the sinoatrial (SA) node, represented by the P-wave of ECG. The electrical impulse then spreads across the atria to the atrioventricular (AV) node and proceeds to the ventricular walls through the bundle of His to initiate ventricular contraction that is recorded by the QRS complex of ECG. After the ventricles are completely activated, they start to repolarize (return to the resting electrical state) and relax, and the T-wave in ECG depicts this phase. Finally, both the atria and ventricles complete repolarization and a new cycle is about to start.

As a dynamically involved system, the electrical stimulus that spreads in the heart drives the orderly contraction and relaxation of the heart muscles, leading to blood pumped into the vessels and peripheral ends that can be captured by PPG. As a result, the dynamics of blood flow are coupled with the transmission of electrical stimulus throughout the heart, indicating that PPG and ECG represent the same cardiac process measured in different signal sensing modalities. As shown in Fig. 1.3, the ascending slope of PPG is caused by ventricle contraction represented by the QRS wave complex in an ECG cycle that pumps blood to the vessels and microvasculature and increases the blood volume there correspondingly [7]. And the descending slope of PPG forms when the blood flows back from the body towards the heart during the ventricle repolarization and relaxation represented by the T-wave and the P-wave of the next cycle.

From the signal processing perspective, we view the ECG as the source signal and PPG as the signal on the receiver side through our cardiovascular system. The electrical cardiac activity caused our heart to beat, followed by the variations of the aortic blood pressure and the blood circulation all over our body, which includes the peripheral site used to measure the PPG signal. The system from ECG to PPG can be treated as an equivalent lowpass filter given the time-frequency representation of the two signals shown in Fig. 1.4, which also inspires a way to model the relationship from PPG to ECG as an inverse engineering. In particular, the low frequency component of ECG can be recovered via some inverse filtering process from the low frequency of PPG. And the high frequency part of the ECG signal can be reconstructed by examining the correlation between the high and low frequency parts of ECG. In our previous work, the transition from PPG to ECG (route (a) + (b) in Fig. 1.4) in the frequency domain can be characterized by a linear transform relating PPG and ECG in the DCT domain [174, 175] that embodies the underlying electrical, biomechanical, and optophysiological principles.

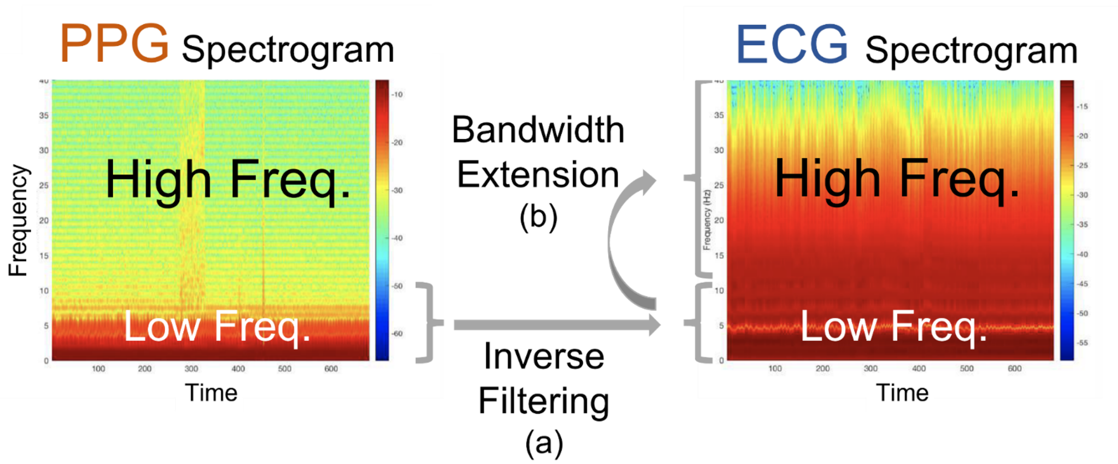

Figure 1.4: Spectrograms of PPG and ECG manifesting the route to discover and model their relationship from PPG to ECG. (a) The low frequency component of ECG can be recovered via some inverse filtering process from the low frequency of PPG. (b) The high frequency part of the ECG signal can be reconstructed by examining the correlation between the high and low frequency parts of ECG.

1.1.3 Digital Twins

Accompanying the industrial revolutions over the past several centuries have been four eras of healthcare revolutions [86, 88, 137], the most recent of which is just beginning. Technological advances have enabled improvements in patient care and monitoring by introducing portable and affordable devices such as pulse oximeters [88], as well as communication and computing infrastructures that make telehealth and remote care possible, enabling people to obtain service from the comfort of their homes [88, 137]. In addition, precision health takes various types of patient-specific information into account, enabling personalized monitoring for preventative care, early detection of diseases, and individualized treatment(s) with potentially improved outcomes.

The digital twin is a promising paradigm toward realizing precision health. As a digital representation of a physical artifact to facilitate the monitoring of the artifact's status [16], the notion of a digital twin was introduced by Michael Grieves in his 2002 presentation for product life cycle management [54, 55] and adopted by the U.S. National Aeronautic Space Administration in its aerospace missions [48, 51]. Broadening the scope of digital twins to healthcare has the potential of producing fine-grained tailor-built models of the biological phenomena relating to an individual's health [16].

Given the high anatomical complexity of human bodies, it is nearly impossible to build one digital twin model to account for all aspects of health needs. Thus, we apply a common engineering strategy of divide-and-conquer to cluster digital twins for healthcare into the following categories:

1. Organ and structural level digital twins [3, 16]: a major application of these digital twins is to build models to help clinicians understand an individual patient's organ structure, support surgery planning, and reduce treatment risks and uncertainties;

2. Omics level digital twins [16]: these digital twins relate genome and other omics in-

formation with patients' health risks, reactions to drugs, and other high-level health conditions;

3. Physiology level digital twins [16, 95]: these digital twins leverage sensing technologies and data science to facilitate the monitoring, analysis, and management of a patient's health conditions on-demand and/or at a chosen time scale.

We focus on the physiological digital twin in this work. One representative prior work in this area [16] describes a digital twin as a system for providing fast, accurate, and efficient medical simulation, which consists of a physical object, a virtual object/model, and healthcare data [95]. A physical object can be represented as a medical or wearable device for monitoring a person's health; and the healthcare data can include data from wearable devices, real-time monitoring data, and simulation data from digital models. While increasingly large amounts of data are becoming available to potentially support data-driven approaches, we believe that healthcare digital twins research should strive for constructing explainable digital twin models when considering both the complex ethics and social aspects of healthcare.

1.1.4 Blood Oxygen Saturation

Peripheral blood oxygen saturation (SpO_2) shows the ratio of oxygenated hemoglobin to total hemoglobin in the blood, which serves as a vital health signal for the operational functions of organs and tissues [131]. The normal range of SpO_2 is 95% to 100% [106]. Abnormality in the SpO_2 level can serve as an early warning sign of respiratory diseases [106]. The estimation and monitoring of SpO_2 are essential for the assessment

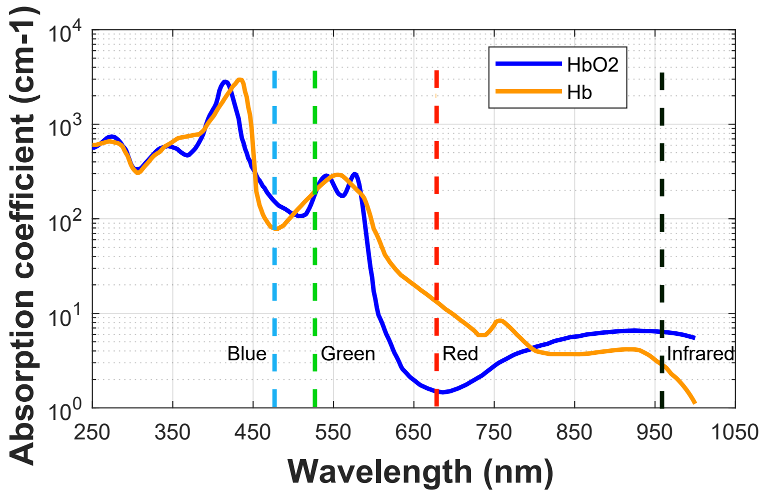

Figure 1.5: Extinction coefficient curves of hemoglobin, figure reproduced based on [40, 108]. The difference between oxygenated and deoxygenated hemoglobin at the red and blue wave lengths means these color channels contain most of the useful information for SpO_2 prediction.

of lung function and the treatment of chronic pulmonary diseases. It has become increasingly important in the COVID-19 pandemic, where many patients have experienced "silent hypoxia," a low level of SpO_2 even before obvious breathing difficulty is observed [35, 130, 145]. The vulnerable population with a high possibility of infection is recommended to monitor their oxygen status continuously for early COVID-19 detection [130, 141].

Pulse oximeters have been widely used for SpO_2 measurement at home and in hospitals in the form of a finger clip [127, 159], which adopts the *principle of ratio of ratios* (RoR) that was first proposed by Aoyagi in the early 1970s [127]. The RoR principle is based on the different optical absorption rates of the oxygenated hemoglobin (HbO_2) and deoxygenated hemoglobin (Hb) at 660 nm (red) and 940 nm (infrared) wavelengths as indicated in Fig. 1.5. By illuminating red and infrared lights on the fingertip, the more oxygenated hemoglobin in the blood, the less infrared light and the more red light are received by the detector after transmission. In other words, the relative AC and DC amplitudes between red and infrared PPG contains pulsatile information to derive SpO_2.

The gold standard for measuring blood oxygen saturation is blood gas analysis, which is invasive and painful and requires well-trained healthcare providers to perform the test. In contrast, the pulse oximeter is noninvasive and provides readings in nearly real time, and is therefore more tolerated and convenient for daily use. The pulse oximeter is known to have a deviation of $\pm 2\%$ from the gold standard when the blood oxygen saturation is in the range of 70% to 99% [114], which is well-known and accepted in clinical use.

1.2 Main Contributions

In this dissertation, we study the modeling of contact and contactless PPG signals to facilitate its promising applications in cardiovascular signal and vital sign sensing and learning for digital health. First, we explore the potential of user-friendly and continuous electrocardiogram (ECG) monitoring with the help of fingertip PPG sensors in Ch. 2. Next, we develop a physiological digital twin for personalized continuous cardiac monitoring in Ch. 3. Last, we study the noncontact methods of blood oxygen saturation (SpO_2) monitoring from remote PPG signals captured by smartphone cameras. Both principled signal processing method and neural network based method for explicit (handcrafted) and implicit (data-driven) feature engineering from multi-channel color signals are proposed in Ch. 4 and Ch. 5, respectively. Below are the detailed key contributions of this dissertation research.

1.2.1 Cross-domain Joint Dictionary Learning for ECG Inference from PPG

The inverse problem of inferring clinical gold-standard electrocardiogram (ECG) from photoplethysmogram (PPG) that can be measured by affordable wearable internet-of-healthcare-things (IoHT) devices is a research direction receiving growing attention. It combines the easy measurability of PPG and the rich clinical knowledge of ECG for long-term continuous cardiac monitoring. The prior art for reconstruction using a universal basis, such as discrete cosine transform (DCT), has limited fidelity for uncommon ECG waveform shapes due to the lack of representative power. To better utilize the data and improve data representation, in Ch. 2, we design two dictionary learning frameworks, the cross-domain joint dictionary learning (XDJDL) and the label-consistent XDJDL (LC-XDJDL), to further improve the ECG inference quality and enrich the PPG-based diagnosis knowledge. Building on the K-SVD technique, our proposed joint dictionary learning frameworks largely extend the expressive power by optimizing simultaneously a pair of signal dictionaries for PPG and ECG with the transforms to relate their sparse codes and disease information. The proposed models are evaluated with a variety of PPG and ECG morphologies from benchmark datasets that cover various age groups and disease types. The results show that the proposed frameworks achieve better inference performance than previous methods, suggesting an encouraging potential for ECG screening using PPG based on the proactively learned PPG-ECG relationship. By enabling the dynamic monitoring and analysis of the health status of an individual, the proposed frameworks contribute to the emerging digital twins paradigm for personalized healthcare.

1.2.2 Never-Miss-A-Beat: A Physiological Digital Twins Framework for Cardiovascular Health

Digital twins are emerging as a promising framework for realizing precision health for their ability to represent an individual's health status. Ch. 3 of the dissertation work introduces a physiological digital twin for personalized and precision continuous cardiac monitoring in the form of modeling the PPG-ECG relationship. Using the dictionary learning algorithm proposed in the previous chapter as the backbone model, the work in this chapter focuses on the problem of inferring ECG signals from PPG signals for continuous precision cardiac monitoring under realistic conditions in which available ECG data is scarce. By performing transfer learning, a generic digital twin model learned from a large portion of paired ECG and PPG data is fine-tuned to precisely infer the ECG from the PPG of a target participant whose available ECG data are scarce. Experimental results for interpolation and extrapolation testing scenarios show that the proposed transfer learning method yields better ECG reconstruction accuracy compared to other baseline comparison models. This suggests that it can be used as a reliable digital twin for precision continuous cardiac monitoring. In parallel, neural network and causality based backbone model designs are also proposed based on the underlying physiological process of ECG generation for better explainability.

1.2.3 Noncontact Hand Video Based SpO_2 Monitoring Using Smartphone Cameras

SpO_2 is an important indicator of pulmonary and respiratory functionalities. It is recommended, especially for the vulnerable population, to regularly monitor the blood oxygen level for precaution. Recent works have investigated how ubiquitous smartphone cameras can be used to infer SpO_2. Most of these works are contact-based, requiring users to cover a phone's camera and its nearby light source with a finger to capture reemitted light from the illuminated tissue. Contact-based methods may lead to skin irritation and sanitary concerns, especially during a pandemic. In this dissertation, we propose a noncontact method for SpO_2 monitoring using remote PPG signals in the hand videos acquired by smartphones. The whole algorithm pipeline includes 1) receiving video of the hand of a subject captured by a regular RGB camera of a smartphone; 2) extracting a region of interest of the hand video; 3) performing feature extraction of the region of interest based on spatial and temporal data analysis of more than two color channels; and 4) estimating a blood oxygen saturation level of the subject from the features.

The contributions of this dissertation mainly focus on the feature engineering and estimation parts of the pipeline. Considering the optical broadband nature of the red (R), green (G), and blue (B) color channels of the smartphone cameras, we exploit all three channels of RGB sensing to distill the SpO_2 information beyond the traditional ratio-of-ratios (RoR) method that uses only two wavelengths. In the principled signal processing method (Ch. 4), the features are explicitly extracted based on the multi-channel RoR after adaptively narrow bandpass filtering based on accurately estimated heart rate to obtain the most cardiac-related AC component for each color channel. Experimental results show that our proposed blood oxygen estimation method can reach a mean absolute error of 1.26% when a pulse oximeter is used as a reference, outperforming the traditional RoR method by 25%. With the understanding of the multi-channel based principled signal processing method, convolutional neural network based schemes (Ch. 5) are further proposed for implicit data-driven feature extraction by skin color channel mixing and temporal analysis. The neural network architectures are designed inspiring by the optophysiological models for SpO_2 measurement. Through the visualization of the weights for the RGB channel combinations, we demonstrate the explainability of our model and that the choice of the color band learned by the neural network is consistent with the suggested color bands used in the optophysiological methods.