CARDIAC PARAMETERS ESTIMATION USING SEISMOCARDIOGRAPHIC AND REMOTE PHOTOPLETHYSMOGRAPHIC SIGNALS

Mousumi Das

CARDIAC PARAMETERS ESTIMATION USING SEISMOCARDIOGRAPHIC AND REMOTE PHOTOPLETHYSMOGRAPHIC SIGNALS

A

Thesis submitted

for the award of the degree of

DOCTOR OF PHILOSOPHY

By

Mousumi Das

DEPARTMENT OF ELECTRONICS AND ELECTRICAL ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI

GUWAHATI - 781 039, ASSAM, INDIA

Feb 2025

Certificate

This is to certify that the thesis entitled "**Cardiac Parameters Estimation Using Seismocardiographic and Remote Photoplethysmographic Signals**", submitted by **Mousumi Das** (186102009), a research scholar in the *Department of Electronics and Electrical Engineering, Indian Institute of Technology Guwahati*, for the award of the degree of **Doctor of Philosophy**, is a record of an original research work carried out by her under our supervision and guidance. The thesis has fulfilled all requirements as per the regulations of the institute and in our opinion has reached the standard needed for submission. The results embodied in this thesis have not been submitted to any other University or Institute for the award of any degree or diploma.

Prof. M. K. Bhuyan,
Professor
Dept. of Electronics and Electrical Engg.
Indian Institute of Technology Guwahati
Guwahati - 781039, India.

Dr. L. N. Sharma,
Senior Technical Officer
Dept. of Electronics and Electrical Engg.
Indian Institute of Technology Guwahati
Guwahati - 781039, India.

Dated:
Guwahati.

Dated:
Guwahati.

To

The Almighty Lord Krishna

for His blessing

My Parents and My Husband

Mr. Syamanta Kumar Das, Mrs. Lucky Das, & Mr. Bitu Handique

for their blessings, support and encouragement

&

My in-laws and other family members

for their love and support

Acknowledgements

I am profoundly grateful to God for his divine guidance, grace, and blessings, to whom I dedicate this thesis. The completion of this journey would not have been possible without the unwavering support of numerous individuals.

Foremost, I extend my deepest gratitude to my esteemed Ph.D. supervisors, Prof. M.K. Bhuyan and Dr. L.N. Sharma. Their guidance, assistance, and encouragement have been valuable throughout my research. I admire their dedication to the pursuit of knowledge and their enthusiasm for work. Words cannot adequately express my appreciation for their constant motivation, positive influence, and moral support. The insightful feedback they provided has significantly contributed to my scientific and personal growth. The EMST Laboratory's facilities and the freedom to work within it have been indispensable in shaping my research and this thesis.

I express sincere thanks to my doctoral committee members, Prof. S.R. Ahamed, Dr. R.D. Kulkarni, and Dr. Navin Gupta, for their support, encouragement, and valuable suggestions. I am also grateful to the anonymous reviewers for their constructive insights. The Department of Electronics and Electrical Engineering at IIT Guwahati, including the faculty members and office staff, played a crucial role in facilitating my research.

My heartfelt appreciation goes to my friends at IIT Guwahati for their care, love, and unwavering support during my doctoral studies. The members of the EMST Laboratory have been instrumental with their significant support whenever I needed it. Special thanks to my senior colleague, Dr. Tilendra Choudhary for his valuable support throughout my Ph.D. journey. Above all, I extend my deepest gratitude to my parents, my husband, my in-laws, and other family members. Their love, support, and sacrifices were indispensable to my Ph.D. completion. This achievement is as much theirs as it is mine, attributed to their constant blessings, silent prayers, and unwavering support, placing me in this position of accomplishment.

Mousumi Das

Abstract

Cardiovascular diseases (CVDs) are major risk factors contributing to the increasing death rate. Effective and regular monitoring of cardiac activities are useful for early detection and clinical management of the CVDs. Many vital parameters, such as heart rate (HR), heart rate variability (HRV), blood pressure (BP), oxygen saturation (SpO2), and respiratory rate (BR) provide insight to cardiac health and help in diagnosing and treating life-threatening diseases. In this study, two emerging cardiac modalities, such as seismocardiography (SCG) and remote photoplethysmography (rPPG) are considered for the estimation of cardiac vital parameters. The SCG is a non-invasive technique that captures the chest wall vibrations induced by cardiac mechanical activities. The acquired SCG signal needs precise delineation and feature extraction prior to the measurement of human vital parameters. The first part of the thesis involves the detection of the prominent peaks of SCG cycles and investigates their possible clinical applications. A data-adaptive modified variational mode decomposition (MVMD) method along with simple decision rules are incorporated to extract two fiducial points, AO and post-AC (pAC) peaks. Later, these points are utilized to derive systolic blood pressure (SBP), diastolic blood pressure (DBP) and HRV parameters. Another application is explored, which utilizes these feature points along with the demographical information of the volunteers to identify ventricular depolarization events using a deep feedforward neural network (DFN). The proposed methods are evaluated on publicly available CEBS database (at PhysioNet archive) and in-house recordings created using a small electronic circuit board consisting of a 3D MEMs-based accelerometer, pre-amplifier, and a filter.

In the second part of the thesis, camera-based rPPG technology is employed to generate the cardiac waveforms and further used them as an application in vital measurements. rPPG is a non-contact camera-based method that senses the variation in reflected light intensity associated with the change in blood flow volume. Accurate estimation of cardiac

parameters using this technology needs robust framework that facilitates the extraction of reliable rPPG signal embedded within the video frames. Signal extraction frameworks involve preprocessing steps, region of interest (ROI) selection, and proper filtering schemes. In this direction, both conventional and deep neural network (DNN) based frameworks are explored. The first method is a conventional technique that is based on 2D-variational mode decomposition (VMD) scheme. Performance analysis suggests that the VMD method performs well. However, this method is based on prior assumptions that may not hold well for real time data. Therefore, in the second approach, few DNN architectures are designed and developed without any prior assumptions to fit into the real dynamic settings. The designed architectures are based on 1D, 2D, and 3D convolution neural networks (CNNs). The input variable for the 1D-CNN is a raw signal generated by averaging the pixel values within the region of interest (ROI), whereas a scalogram-based feature image is chosen as an input for the 2D-CNN. Both the networks are proposed for the extraction of a reliable rPPG waveform that is successively used for HR estimation and HRV analysis. In contrast to 1D and 2D CNN frameworks, the input to the 3D-CNN is the raw video data. This architecture is designed for multitask where remote electrocardiography (rECG) along with rPPG is generated. The generated signals are quantified by estimating HR values and comparing them with the corresponding ground-truth values. The robustness of the proposed rPPG methods is evaluated using publicly available UBFC-rPPG, COHFACE, and ECG-fitness datasets.

Contents

List of Figures

List of Tables

List of Acronyms and Abbreviations

AO	Aortic Valve Opening
AC	Aortic Valve Closure
ACC	Accuracy related to detection of fiducial points
ADR	Accelerometer-derived respiration
ANN	Artifcal Neural Network
AV	Atrioventricular
ANS	Autonomic Nervous System
ADMM	Alternate Direction Method of Multipliers
AAPSD	Azimuthally Averaged Power Spectrum Density
BR	Breathing Rate
BP	Blood Pressure
BSS	Blind Source Separation
BB	Bounding box
BPF	Band pass filter
BVP	Blood Volume Pulse
BPM	Beats Per Minute
CEBS	Combined measurement of ECG, Breathing, and SCG
CCE	Cardiac Cycle Envelope
CTI	Cardiac Time Interval
CNN	Convolutional Neural Network
CCNN	Cascade Convolutional Neural Network
CAN	Convolutional Attention Network
CVDs	Cardiovascular Diseases

CP	Contrast-Phys
DMK	Dominant Multiscale Kurtosis
DMCF	Dominant Multiscale Central Frequency
DER	Detection Error Rate
DE	Energy of diastolic region
DBP	Diastolic Blood Pressure
DRM	Dichromatic Reflection Model
DRMF	discriminative response map fitting
DAN	Deep alignment network
DFT	Discrete Fourier Transform
DCT	Discrete Cosine Transform
DL	Deep Learning
DNN	Deep Neural Network
DFN	Deep Feedforward Network
DelNet	Delineating Network
ECG	Electrocardiogram
ECHO	Echocardiogram
EMAC	Electro-Mechano-Acoustic Cardiovascular
EMD	Empirical Mode Decomposition
EEMD	Ensemble Empirical Mode Decomposition
EVM	Eulerian Video Magnification
ER	Exercise recovery
EMST	Electro Medical and Speech Technology
FFT	Fast Fourier Transform
FH	Forehead
FLM	Facial landmark detection
FBG	Fiber Bragg Gratings
FC	Fully-Connected
GCG	Gyrocardiogram
GDFM	Gaussian Derivative Filtered Mode

GRD	Green-red difference
GAN	Generative Adversarial Network
HR	Heart Rate
HRV	Heart Rate Variability
HF	High Frequnecy
HVD	Hilbert Vibration Decomposition
IM	Isovolumic Moment
IC	Isotonic Contraction
IVCT	Isovolumetric Contraction Time
IVRT	Isovolumetric Relaxation Time
ISS	International Space Station
ICG	Impedance cardiography
ICA	Independent Component Analysis
ICU	Intensive Care Unit
IRB	Institutional Review Board
IMFs	Intrinsic Mode Functions
KLT	Kanade-Lucas-Tomasi
KCF	Kernelized correlation filters
LVET	Left Ventricular Ejection Time
LDV	Laser Doppler Vibrometry
LF	Low Frequency
LR	Logistic Regression
LBL	Lambert-Beer Law
LPF	Low pass filter
LMSAF	Least mean square adaptive filter
LOA	Limits of Agreement
LPF	Low Pass Filter
MEMS	Micro-Electro-Mechanical Systems
MRI	Magnetic Resonance Imaging
MO	Mitral Valve Opening

MC	Mitral Valve Closure
MDR	Microwave Doppler Radar
MWD	Multiscale Wavelet Decomposition
meanNN	Mean of all NN intervals
meanHR	Mean of Heart Rates
MTCNN	Multi-Task cascaded convolutional networks
MAF	Moving average filter
MVMD	Modified Variational Mode Decomposition
ME	Mean Error
MAE	Mean Absolute Error
MWPPG	Multi-Wavelength PPG
NCC	Normalized Cross Correlation
NLMS	Normalized Least Mean Square
NB	Naive Bayes
NIR	Near Infrared
OMIT	Orthogonal Matrix Image Transformation
PPG	Photoplethysmogram
PCG	Phonocardiogram
PEP	Pre-ejection Period
PTT	Pulse Transit Time
pNN50	Percentage of successive NN intervals that differ by more than 50 ms
PAM	Pulmonary Artery Mean Pressure
PCWP	Pulmonary Capillary Wedge Pressure
PCA	Principle Component Analysis
POS	plane orthogonal to the color space of the skin
PD	peak detection
PF	PhysFormer
RF	Rapid Filling
RE	Rapid Ejection
RHC	Right Heart Catheterization

RR	Interval between two consecutive R-peaks in ECG
RMSSD	Root Mean Square of the Successive Differences
RSDMK	Relative Squared Dominant Multiscale Kurtosis
RGE	Relative GDFM Energy
ROI	Region of Interest
RFE	Recursive Feature Elimination
RFT	Random Forest
RMSE	Root Mean Squared Error
rPPG	Remote Photoplethysmography
rECG	Remote Electrocardiography
SCG	Seismocardiogram
STI	Systolic Time Interval
SVM	Support Vector Machine
SDNN	Standard deviation of all NN intervals
SDSD	Standard Deviation of Successive Differences (in HRV measurement)
SE	Shannon energy
Se	Sensitivity
SBP	Systolic Blood Pressure
SNR	Signal to Noise Ratio
SK	Skin Detection
SBB	Subset of Bounding Box
STFT	Short Time Fourier Transform
SSA	singular spectrum analysis
SSD	Single shot multibox detection network
SpO2	Blood Oxygen Saturation
SHAP	SHapley Additive exPlanations
SB	Supine position with stopped breathing
SI	Sitting
ST	Standing
STD	Standard Deviation

SV	Stroke Volume
SDE	Standard Deviation Error
THD	Total Harmonic Distortion
TFL	Time Frequency Learning
VMD	Variational Mode Decomposition
VJ	Viola-Jones
VD	Ventricular Depolarization
XGBoost	eXtreme gradient boosting

1

Introduction

Contents

1. Introduction

Cardiac vital parameters play an essential role in assessing cardiac anomalies and help reduce the mortality rate. This thesis documents the estimation of major cardiac parameters using two emerging biosensor modalities. The first one is a low-cost MEMS-based accelerometric device that acquires the seismocardiogram (SCG) waveform. SCG is generated due to chest wall vibrations and it captures clinically relevant information associated with cardiac mechanical activities. In comparison to other standard measurement tools, such as electrocardiography (ECG), phonocardiography (PCG), echocardiography (ECHO), and magnetic resonance imaging (MRI), SCG provides more insight to the cardiovascular activities [1]. The fiducial points reflecting the cardiac events are useful in the estimation of cardiac vital parameters that are effective for cardiac analysis and disease diagnosis. SCG has the potential to be a part of cardiac measurement tools. It may gain popularity due to its low cost and ease of use. However, for data acquisition, the sensor needs direct contact with the subjects, which may cause discomfort and limit the moving flexibility. To overcome the limitation of SCG modality, the second approach explores a non-contact camera-based method for cardiac parameters estimation. This method is based on the optical principle of photoplethysmography (PPG) and named as remote PPG (rPPG) due to its contactless nature. In this approach a camera is used for the recording of facial video data that is further processed for the extraction of rPPG waveform. Extracting a reliable rPPG waveform for clinical applications involves robust frameworks consisting of preprocessing and filtering schemes for eliminating the environmental interferences.

The first part of the thesis manifest the delineation and extraction of fiducial points of SCG waveform. Whereas, the second part focuses in the design and development of robust frameworks for rPPG waveform extraction. Overall the dissertation centers around the estimation of significant cardiac parameters for cardiac health assessment. The rest of this chapter is organized as follows.

The first section 1.1 provides an overview of SCG including the recording procedure, morphological characteristics of the SCG waveform and its physiological aspects alongside with ECG and PPG. It also covers the literature surveys on SCGs potential applications. Section 1.2 presents the rPPG technology, its physiological aspects, and the challenges in rPPG waveform extraction. Also, it presents progressive state-of-the-art works on rPPG frameworks for rPPG waveform extraction followed by clinical applications. Section 1.3 presents the motivation and the work-plan of the thesis. Finally, the organization of the thesis is presented via graphical-abstract in Section 1.4.

1.1 Seismocardiography

This section delves into the evolution of seismocardiography (SCG) technology, starting from its origin and exploring its numerous clinical possibilities. It encompasses the genesis, recording process, and characteristics of SCG waveform. Additionally, it outlines the correlation between SCG's fiducial points and cardiac mechanical events. Subsequently, the challenging factors confronting SCG technology, alongside a comprehensive literature review highlighting its progress are addressed.

1.1.1 Seismocardiographic Waveform

Seismocardiography is a non-invasive method to record the chest-wall vibrational data caused due to change in volume, pressure, and shape of the heart during different stages of a cardiac cycle [14]. The SCG waveform generated from the chest-wall vibration can be acquired by placing a micro-electro-mechanical system (MEMS)-based tri-axial accelerometer sensor on the precordial area of the chest, generally at the lower end of the sternum on the xiphoid process [15]. A tri-axial accelerometer captures the SCG events in three orthogonal axes (x, y, and z), where x, y, and z represents the lateral [left-right], longitudinal [foot-head], and sagittal [dorso-ventral] components, respectively. Among all the axes, the z-component is mostly considered for the studies due to its significant variations with heart events as compared to the other two components. The sensor is able to capture the mechanical activities of both systolic and diastolic phases of heart with important fiducial points, such as opening and closing of valves, isovolumetric moment, isotonic contraction, rapid blood filling and its ejection. All these events are reflected as peaks and troughs on an SCG cycle, which are called fiducial or feature points. Fig 1.1 shows the placement of a SCG sensor and a cycle of SCG waveform with its corresponding fiducial points.

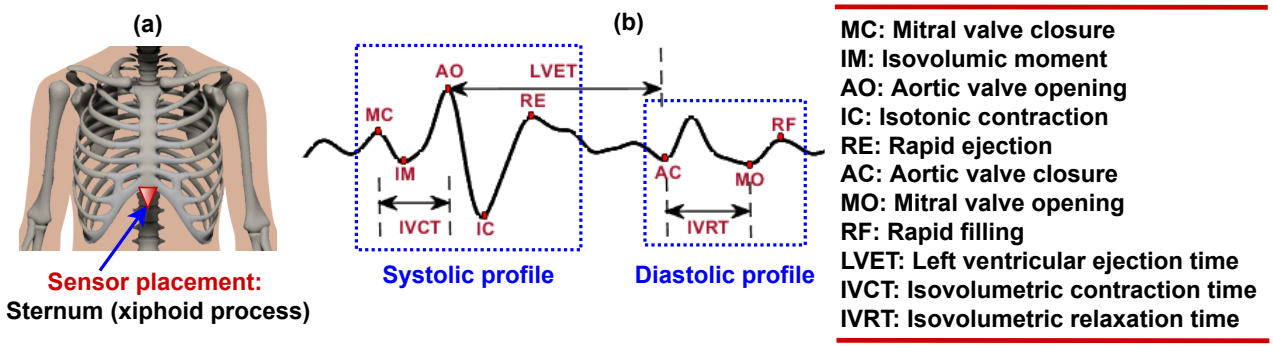


Figure 1.1: (a) Sensor placement for SCG acquisition and (b) SCG cycle with its fiducial points and time-intervals.

The SCG signal has a frequency range from 0.5 to 40 Hz, but its maximum energy lies in the infrasonic range (≤30 Hz) [16]. The typical peak to peak amplitude range of SCG ranges from 5–10 mg (milli-gravity, g = 9.81 m/s^2) [17], however in the presence of motion artifacts it extends up to 20–50 mg. The annotation of the SCG fiducial points is yet to be standardized, and it is being annotated through mapping with other cardiac signals, such as echocardiogram, ECG, heart sound, and arterial blood pressure [18]. The fiducial points and the time intervals between them can be used for clinical applications. It captures many mechanical events that can be used to derive important cardiac intervals, such as pre-ejection period (PEP), left ventricular ejection time (LVET), isovolumetric contraction time (IVCT), and isovolumetric relaxation time (IVRT). Thus, annotation of SCG is a primary task before deploying it for clinical usage.

1.1.2 Physiological Aspects of SCG

The heart is an electromechanical pump constituting two different pumps separated by a septum. Each of these pumps are further subdivided into lower and upper chambers, called atrium and ventricle. These chambers are partitioned by atrio-ventricular and semi-lunar valves, as shown in Figure 1.2. The heart involves four primary phases: 1) isovolumetric relaxation, 2) inflow, 3) isovolumetric con-

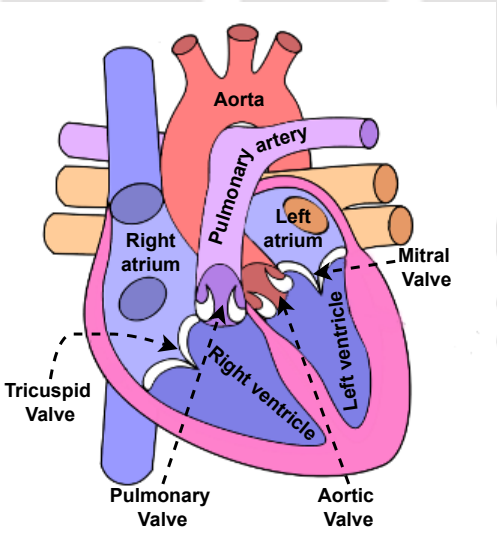


Figure 1.2: Human heart anatomy

traction, and 4) ejection. These stages involve mechanical events like valve opening and closing, rapid filling, and rapid ejection of blood from the ventricles. The SCG waveform effectively depicts these mechanical phenomena of the heart. Figure 1.3 shows the SCG waveform along with the simultane-

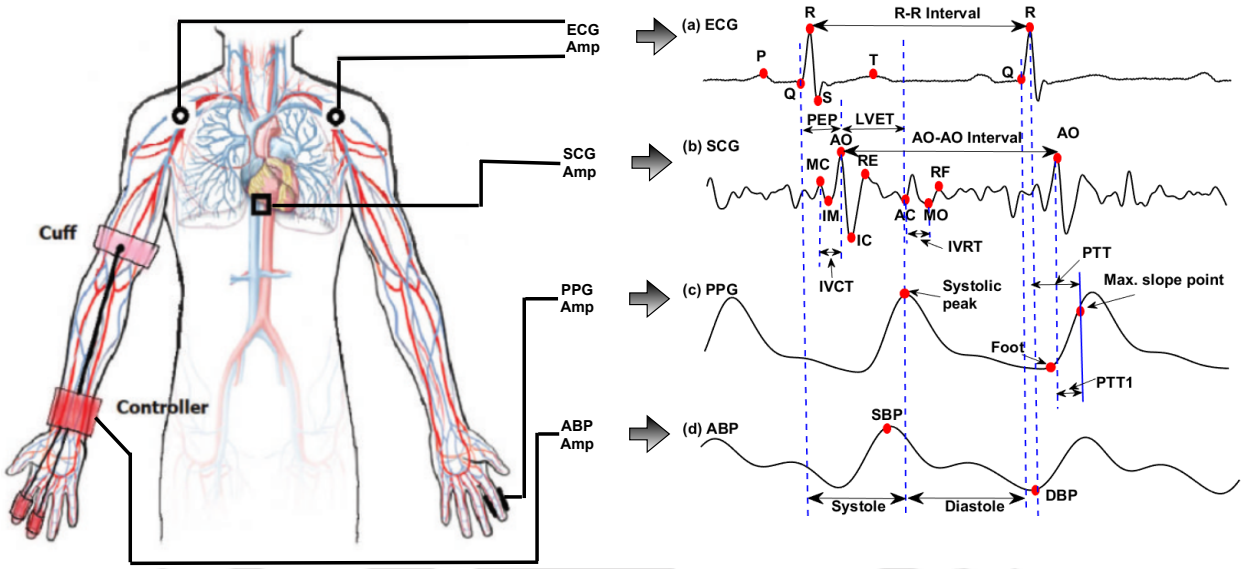


Figure 1.3: Simultaneously recorded ECG, SCG, PPG, and ABP waveforms. **MC**: Mitral valve closure, **IM**: Isovolumic moment, **AO**: Aortic valve opening, **IC**: Isotonic contraction, **RE**: Rapid ejection, **AC**: Aortic valve closure, **MO**: Mitral valve opening, **RF**: Rapid filling, **PEP**: Pre-ejection period, **LVET**: Left ventricular ejection time, **IVCT**: Isovolumetric contraction time, **IVRT**: Isovolumetric relaxation time, **PTT**: Pulse transit time.

ously recorded ECG, PPG, and ABP waveforms. A cardiac cycle in each cardiac waveforms consists of two phases, systole and diastole. The MC notch in the SCG is resulted due to mitral valve closure after complete ejection of blood from left atrium to left ventricle. At this point, the major left ventricular depolarization through Purkinje fibers takes place, and in a standard ECG, the R-wave is seen. The ventricular depolarization makes the ventricles contract leading sharp inward wall motion, and decrease acceleration waves. At this moment, aortic valve is not opened. The ventricle works as a closed chamber. The volume remains the same, the pressure increases rapidly due to rapid contraction. This phase is known as isovolumetric contraction and it cause the generation of the IM valley point. This process is also reflected as S-wave in ECG which indicates complete depolarization of ventricles. The period until the opening of aortic valve from the depolarization of the left ventricle is known as pre-ejection period (PEP). When the ventricular pressure reaches to the aortic pressure, the apex moves inward, and the walls are swelled, which increases the overall acceleration and opens the aortic valve. This is indicated by AO peak point. The opening of aortic valve decreases and stops the acceleration of inward wall motion, which is mapped with the IC valley point. Then, the ventricular pressure rises and blood is rapidly ejected out from the left ventricle to aorta, which gives the RE peak point. In ECG, this condition is traced as an isoelectric line and followed by onset of T-wave

caused due to ventricular repolarization. The opening of AO peak also reflects the beginning of blood flow as a onset point in PPG. Also, the similar behaviour is observed in the pressure waveform, where the pressure gradually increases with respect to aortic blood pressure. Till the opening of aortic valve, blood is circulated through out the body and it creates systolic peaks in PPG and pressure wave. Once the aortic valve is closed, the AC point is generated. Also, the ventricular repolarization stops that corresponds to the offset of T-wave in ECG. At this point, the volume and pressure of peripheral blood starts reducing as reflected in PPG and ABP waveform. The closing of AC point in SCG and offset of T-wave in ECG reflects the end of systolic phase. After this phase, the blood enters into the atrium and the atrial pressure increases leading to rise in the acceleration. Once the mitral valve opens, the atrium volume decreases and creates MO valley point. Consecutively, the blood flows into the ventricles, and the heart walls move outward with increased acceleration. This action generates the RF peak point. The time duration of rapid filling is traced by an isoelectric line. Within this time period, the whole blood is circulated to the peripheral sites and at the end the volume and pressure decreases, as depicted by the foot of PPG and ABP waveforms. This phenomenon ends the diastole phase and this way, an SCG cycle is generated [15].

1.1.3 Challenging factors in SCG domain

The high sensitivity of accelerometeric sensor towards unwanted noises degrades the quality of the recorded SCG signals, and creates difficulty in the interpretation of SCG fiducial points. The morphological structure of SCG is directly affected by large vibration, inhale and exhale phases of respiration, lung volume variations, and motion artifacts [14, 18, 20]. In addition, SCG gets easily distorted by many other factors, such as inter-subject variabilities (age, gender, body mass index), body postures (sitting, standing, supine, walking), sensor displacement from the measurement site, heart rhythm, cardiac contractility, valvular function, and blood vessels stiffness [21, 22].

In addition to unwanted effects causing morphological issues, there are some other difficulties associated with SCG signal. The first and foremost is the requirement of universal standardization for the annotated fiducial points and their corresponding clinical significances [14, 20]. Further, the frequency range of SCG needs standardization. Some studies have suggested that SCG lies in the infrasonic range i.e., ($\leq$25 Hz) [23], while others have specified that SCG contains some higher frequency components up to 1000 Hz [24]. Next, is the optimal selection of the measurement axis among all the three axes recorded using the accelerometer. Mostly, the axis corresponding to the dorso-ventral direction

is considered for the studies. However, further investigations are needed to study the usefulness of other two axes.

1.1.4 Advancement of SCG and its applications - A Review

The word seismocardiography was derived by combining the seismology technology with chest wall vibrations generated from cardiac activities. It was termed by Russian researchers R. Baevski and Bozhenko in the early 1960s [25], [26]. Bozhenko and his team had used a magnetic-spring oscillator as an accelerometer to study the chest vibration signals and analyzed them with the concurrent ECG. In contrast, Baevski's group measured SCG by using the procedure followed by seismologists to measure underground vibrations caused by earthquakes. Initially, the SCG signal was used in the Russian aerospace-program for astronauts health monitoring [18]. Later, SCG was considered as a research work and several experiments were performed in the International Space Station (ISS) [27]. After 2 decades i.e. around 1991, the clinical utility of SCG was first introduced by Salerno and Zanetti, and later it was commercialized by *SeisMed Instruments* company in USA [28]. Successively, during 1991 to 1994, various studies were conducted clinically by integrating SCG with other cardiac sensing modalities, such as ECG, ECHO, and MRI imaging to provide a more comprehensive assessment of cardiac function [18, 19, 29–31]. Also, the combined information from multiple sources was used to develop waveform nomenclature for cardiovascular dynamics and for improving the diagnostic accuracy. However, due to bulky and heavy sensing device, the SCG technology didn't get much popularity in the early 1990s. Later, with the advancement in technology, small-sized, cost effective, noise resistive, and highly sensitive MEMS accelerometers were developed and SCG signal was reinvestigated with better accuracy. Few portable, wearable, and wireless SCG sensors were also developed to facilitate continuous and ambulatory monitoring. In 2007, Castiglioni *et al.* introduced a wearable device by integrating an external triaxial MEMS accelerometer with their previously designed textile-based MagIC smart shirt for long-term monitoring. The MagIC was designed in 2005 to monitor ECG along with respiratory movements [32]. In their follow-up work they presented a modified version of this system, named *MagIC-SCG* in which a smart vest is designed by integrating the accelerometer into electronics with the textile electrodes and the piezoresistive plethysmograph [33], [34], [35]. In 2017, Andrew *et al.* developed a SeismoWatch that involved an accelerometer and an optical sensor to monitor cuffless blood pressure [36]. The sensors embedded in the watch acquires both SCG and PPG waveforms to obtain the proximal and distal points for pulse transit time (PTT) measurement. Successively, the

measured PTT is modelled to estimate BP. In contrast to wearable SCG sensors based on either rigid accelerometers or non-stretchable piezoelectric membranes, Taewoo *et al.* introduced an ultra-thin and stretchable electronic tattoo (e-tattoo) based sensor to simultaneously record ECG and SCG signals [37]. In their work, a soft SCG sensor and a pair of soft gold electrodes were integrated on a single e-tattoo platform that was named as soft electro-mechano-acoustic cardiovascular (EMAC) sensing tattoo. This tattoo was investigated to extract various cardiac time intervals including systolic time interval (STI). In 2020, a wearable *SeisMote* system was presented that consists of multiple wireless sensor nodes to simultaneously record ECG, PPG, acceleration, and rotational velocity from different body areas [38]. The developed system was used to study the propoagtion of precordial vibrations along the thorax and to investigate strategies of hemodynamic regulation in different vascular sites. In another research, Shandhi *et al.* made a small wearable patch that simultaneously records the ECG and triaxial SCG signals [39]. The data was recorded during right heart catheterization (RHC) procedure and it was used to estimate the changes in pulmonary artery mean pressure (PAM) and pulmonary capillary wedge pressure (PCWP). In 2023, a non-contact, wearable synchronous ECG and SCG measuring system was developed. The effectiveness of the system was shown by recording the data through multi-layer cloth [40]. Here signals were recorded from multiple sites and the bottom of the sternum was found to be the optimal measurement point. In the same year, another multi-node wireless wearable system named *SeismoNet* was designed to record concurrent ECG and SCG signals from multiple locations of the human body for enhanced physiological sensing [41]. This study showed that the combined recordings from multiple nodes improves the accuracy of heart rate (HR) and breathing rate (BR) estimation.

SCG sensing technologies

Over the time, there have been an extensive growth in the SCG sensing technologies that contributed in the quality improvement of recorded SCG signals. Different SCG sensing mechanisms used in the recent studies are as follows [22]:

- Uniaxial/triaxial piezoelectric accelerometers [42–50],
- Uniaxial/triaxial MEMS accelerometers [11, 21, 51–55],
- 3-D axial gyroscopes [56–59],
- Smartphone accelerometers and gyroscopes [60–62],
- Laser Doppler vibrometers [63],

- Microwave Doppler radars [64–66],
- Airborne ultrasound surface motion camera (AUSMC) [67]

Based on the type of sensors employed, SCG signals may contain axial and rotational components. For instance, employing a uniaxial accelerometer allows measurement of the dorso-ventral SCG component. Conversely, the utilization of both a triaxial accelerometer and a triaxial gyroscope enables the extraction of heart-induced motion information in axial and rotational directions. Usually, SCG sensors are affixed to the sternum or its left-side lower border. However, in many methods alternative chest locations have been employed for recording SCG signals, such as the left clavicle, over the heart apex, and the upper right sternal border [22].

The application of adhesive electrodes or sensors for recording ECG or SCG signals faces limitations in certain scenarios, such as in highly infectious patients, burn victims, and premature infants. Consequently, there is ongoing research into the development of contactless and efficient SCG acquisition schemes. The utilization of these techniques has the potential to reduce skin coupling artifacts commonly observed in SCG signals recorded using contact sensors on the body. Contactless methods include technologies like laser Doppler vibrometry (LDV), microwave Doppler radar (MDR), and airborne ultrasound imaging [22].

LDV method measures the chest vibration velocity by comparing the frequency shift between the outgoing and reflected laser beams [63]. To precisely measure SCG using LDV, the laser beam must be incident perpendicular to chest surface, the chest surface must be reasonably reflective, and chest movement resulting from respiration needs to be considered. These criteria can be addressed by creating an algorithm that can automatically guide the beam to track a specific measurement point on the surface of the chest. MDR is another non-contact method for SCG measurements [64–66]. When acquiring the SCG signal with microwave Doppler radars, the SCG will manifest as phase variations in the microwave signal. Subsequently, SCG signals can be extracted from this phase fluctuation. However, the quality of the extracted SCG signals may degrade due to background microwave signals reflection, commonly called as radar clutter. In addition to LDV and MDR, ultrasound imaging was also investigated to capture 2D SCG maps of the body surface [67]. In comparison to other non-contact methods, this technology is highly reliable as it enables the collection of SCG data from various locations through distinct channels. However, it necessitates a planar measurement surface that is parallel to the emission panel [67]. Also, it's worth noting that these non-contact approaches

have limitations over their sensor cost and size.

Piezoelectric and MEMS-based sensors are more compact and lightweighted compared to contactless sensors. Thus, these types of sensors can be applied in clinical settings for the ongoing and routine monitoring of subjects. All this advanced SCG sensing technologies had opened new possibilities and offered the advantage of capturing data during physical activities and sleep, providing deep insight into cardiac performance in a wider range of conditions.

Processing of SCG signal

Apart from SCG sensing techniques, there has been an extensive growth in signal processing tools for enhancing the quality of SCG waveforms. Usually, SCG is contaminated with noises from diverse sources, such as friction noise, motion artifacts, sensor mechano-electronics, and environmental vibrations [22, 68]. The presence of these noises can pose challenges in signal interpretation, resulting in inaccurate feature extraction. Hence, the preprocessing of SCG signals is regarded as a crucial research objective. The preprocessing of SCG involves noise removal, segmentation and fiducial points extraction. This section provides an overview of studies focused on the denoising and annotation of SCG signals.

Signal denoising methods

Researchers have investigated multiple sensing approaches and signal processing tools to eliminate the unwanted effects from SCG waveforms. In [69], a dual-sensor approach was employed to acquire the SCG signals from posterior and anterior chest-surface. Subsequently, the motion artifact from the signals was eliminated by utilizing a normalized least mean square (NLMS) filter. The results indicated that using multiple sensors offers better resistance against motion noise compared to a single sensor. However, employing multiple sensors raises the overall complexity of the system. In contrast to sensor-based denoising schemes, most of the research groups had applied signal processing methods to eliminate a wide-range of noises. Unwanted effects caused due baseline wandering, respiratory artifacts, body vibrations, and white Gaussian noise were compensated by using conventional band-pass filter [42, 48, 50, 58, 59, 70–72], Kalman filter [73], fuzzy neural network [74], NLMS based adaptive filter [75], empirical mode decomposition (EMD) [76], ensemble empirical mode decomposition (EEMD) [77], median filter, Savitzky-Golay filter [17], and autoencoders [78]. The information of these noise removal methods is summarized in Table 1.1.

Table 1.1: Noise removal techniques used for SCG denoising.

Reference	Method	Suppressed noises
[11, 42, 44, 45, 48–50, 58, 59, 70–72, 76, 79–87]	Low-pass, High-pass, Band-pass, Notch filtering	Baseline drift, respiration and body-movement artifact
[33, 88]	Adaptive filtering	Motion artifact
[89, 90]	Averaging theory	Motion artifact
[64]	Comb filtering	Breathing noise from radar signal
[76, 77, 88, 91]	Empirical mode decomposition, Hybridized variational mode decomposition	Baseline drift, respiration and body movement artifact
[92]	Ensemble empirical mode decomposition (EEMD)	Vehicle-vibration artifacts
[93]	Independent component analysis	Motion artifact
[90, 94]	Median filtering	–
[88]	Morphological filtering	–
[17]	Polynomial smoothing	Motion artifact
[17, 55, 90]	Savitzky-Golay filtering	Motion artifact
[88, 95]	Wavelet denoising	Segmentation of HSs and SCG
[77]	Wiener filtering	–
[73]	Two-stage Kalman filtering	SCG artifacts
[78]	Autoencoders	White Gaussian noise

Annotation methods of SCG fiducial points

The fiducial points of SCG corresponds to different mechanical activities of the heart. Their annotation is clinically significant for pathological interpretations and cardiac health assessment. The attributes of fiducial points, such as amplitude, frequency, and temporal instants are useful for clinical applications including hemodynamic parameter estimation, HRV analysis, detection and identification of cardiac anomalies [96]. However, detecting these fiducial points becomes challenging in the presence of noises. Hence, annotation is considered as an another step in the processing of SCG signals. The techniques used for SCG annotation are well reported in [14, 18, 20, 97]. Most of the existing annotation works are based on the time-domain or the frequency-domain approaches. There are several methods that

are being focused to detect only the aortic valve opening (AO) point, as it is prominent in an SCG cycle. The annotation of the AO peak is studied extensively in [16, 42, 71, 92, 98–105]. The majority of these studies identifies cardiac events in an SCG signal by utilizing the ECG signal as a reference [42, 71, 92, 99, 103]. In contrast to ECG assistance, [104] employed PPG signal as reference to estimate systole and diastole phases, such as isovolumic moment (IM), AO, isotonic contraction (IC), aortic valve closure (AC), post-AC (pAC), and mitral valve opening (MO) points. Also, few other studies proposed algorithms to detect multiple fiducial points including IM, AO, and IC [16, 98]. In few studies, SCG signals are annotated using gyroscopic recordings of the chest wall [58, 59].

Apart from relying on other cardiac signals for SCG annotation, a few researchers have developed standalone algorithms for delineating SCG [101,102,105,106]. In [106], multiple envelopes were derived to estimate IM, AO, and AC fiducial points. The envelopes: heart rate (HR), AO, and AC were extracted using high pass filtering, absolute signal construction, and triple integration with optimal parameters, such as window size and percentage of casualty index. Subsequently, peaks of each envelope is detected that are named as PHRE, PAOE, and PACE corresponding to the envelopes HR, AO, and AC, respectively. Finally, the locations of IM, AO, and AC peaks are determined within a 30 ms time-window from the locations of detected envelope peaks PHRE, PAOE, and PACE. This study has shown the feasibility of standalone framework for fiducial point localization. However, constructing an envelope for SCG signals with severe inter- and intra-beat variabilities is deemed inappropriate. Identifying the optimal peak of an envelope is also challenging. Moreover, the window size is chosen empirically for the integration task. In [101], Tilendra *et al.* proposed an automated method for AO instants detection using multiscale wavelet decomposition (MWD). The steps involved in the detection of AO peak instants are as follows:

- SCG signal is decomposed into nine different subbands using biorthogonal-3.9 mother wavelet,
- The probable AO peaks rich subbands are selected by a newly proposed dominant multiscale kurtosis (DMK) and dominant multiscale central frequency (DMCF)-based criterion,
- SCG is reconstructed from the selected subbands,
- AO peaks in the reconstructed SCG are enhanced with weights, which are based on relative squared dominant multiscale kurtosis (RSDMK),
- Systole envelope is constructed using Shannon energy (SE) and autocorrelation coefficients,

- AO peaks are detected by a Gaussian-derivative-filtering-based logic.

The method has been tested on combined measurement of ECG, breathing and seismocardiogram (CEBS) database. The method achieved overall detection rates as an average (sensitivity = 94.3%, prediction rate = 90.19%, accuracy = 85.16%). Although the method achieves good sensitivity and prediction rate, its accuracy was poor due to the presence of various artifacts that adversely affect the quality of the signal. In their follow-up work, a data-adaptive time-frequency distribution technique was presented to effectively resolve and analyze the SCG signal in the presence of noises [102]. They employed variational mode decomposition (VMD) technique to suppress baseline drift from the signal, and extracted significant modes to detect the AO peak instants. They modified the traditional VMD method to yield the optimal parameter combination, including mode bandwidth constraint (α) and the number of mode (K), which were manually selected in the original VMD. The systolic profiles of the extracted modes were enhanced in the derived Gaussian derivative filtered modes (GDFM). Subsequently, GDFMs with rich AO peaks were selected based on relative GDFM energy (RGE) to reconstruct the signal. The AO peaks were approximated by envelope construction technique and Hilbert transform-based peak detection logic. Finally, the assistance of a cardiac cycle envelope (CCE) facilitated the identification of valid AO peaks. In 2022, a new time-varying decomposition method based on the Hilbert Vibration Decomposition (HVD) was presented to estimate the HR envelope from SCG [105]. The method consists of the following steps: SCG components decomposition using the HVD algorithm, HR envelope extraction, peak detection on the extracted envelope, and interbeat intervals calculation from the detected envelope peaks. A block diagram, depicted in Figure 1.4, illustrates the algorithm of the method. The method has been tested and validated on CEBS database

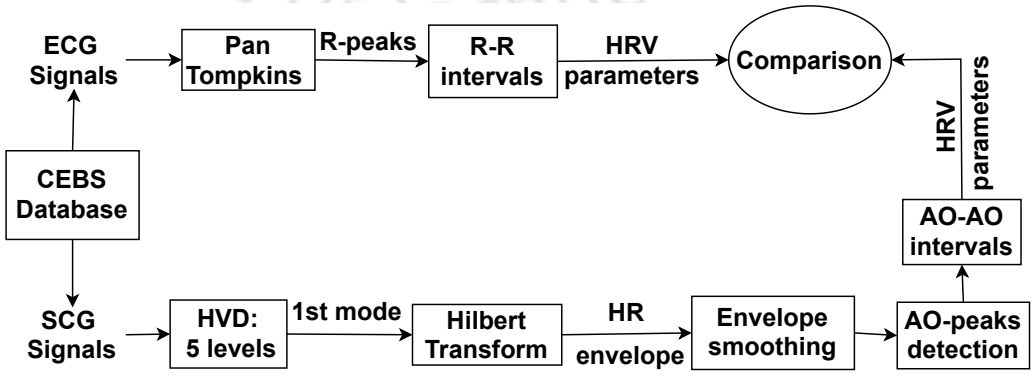


Figure 1.4: Block diagram representation of a previous annotation method, where AO peaks are detected using HVD and envelope construction scheme.

and its been compared with two other decomposition techniques, such as EMD and VMD. The cardiac intervals obtained from the SCG were compared with those derived from the concurrent ECG signals. The utility of the cardiac intervals was shown by estimating the HRV parameters in both time-domain (SDNN, SDSD, RMSSD, pNN50, meanHR) and frequency domain (LF/HF). Notably, the HVD method exhibited superior performance in estimating HRV parameters compared to the other two techniques.

In [42], an automatic delineation algorithm was proposed to detect IM, AO, and AC points for cardiac time intervals estimation. The algorithm was developed in a manner, where it can detect fiducial points of SCG with or without using the ECG as the reference signal. In the standalone approach, the points were detected from heart rate envelope that was calculated with the help of moving average filtering. Further, they employed machine learning-based approach to precisely estimate the fiducial points. The algorithm was designed to robustly eliminate the low-quality cardiac cycles using probabilistic measure. Thakkar *et al.* presented automatic and fast annotation scheme using binary classifiers, such as Naive Bayes (NB), Logistic Regression (LR), and Support Vector Machine (SVM) [107]. The annotation framework was divided into three phases (i) preprocessing, (ii) training, and (iii) testing. For training, three features were employed, namely amplitude, time of appearance, and count. Preprocessing was conducted before testing phase to identify the candidate peaks. The preprocessing was done to reduce the search area by dividing each SCG cycle into three distinct zones. The peaks were detected using the maxima and minima of zones. The classifiers were finally trained to select the desired peaks by filtering out the unwanted peaks. The method was evaluated using 9000 SCG cardiac cycles of 20 healthy subjects. Metrics such as, precision, recall, and F-measure are evaluated followed by 5-fold cross-validation to validate the model. The method was compared with state-of-the-art schemes that demonstrates its superior performance. Several machine learning methods have been used in SCG studies for purposes beyond annotation, including SCG classification of different phases of respiratory cycle [108–111], classification of low risk and normal patients [62, 112, 113], and identification of artifacts [114].

In brief, the above SCG algorithms can be categorized in three different categories: SCG delineation frameworks with reference to other cardiac modalities [16, 42, 58, 59, 71, 92, 98, 99, 103, 104], standalone temporal envelope construction [101, 102, 105, 106], and machine learning approaches [42, 62, 107–114].

Applications of SCG

The annotated fiducial points of SCG can be utilized to assess various cardiovascular parameters, including HRV analysis, estimation of hemodynamic parameters, and the diagnosis of cardiovascular diseases [97]. Many research work has demonstrated the application of SCG for monitoring HR and BR [21,22], left-ventricular functions [115,116], valvular activities, blood flow in coronary artery during angioplasty [29,115,116], and ventricular filling and ejection during ischemia [115]. The clinical utility of the SCG signal reported in the literature, are outlined as follows:

- **Heart rate and heart rate variability estimation:**

  Tadi *et al.* presented a Hilbert adaptive beat identification technique for heartbeat timings and inter-beat time intervals measurement from healthy volunteers in supine, left and right recumbent positions [21]. In [117], a hidden Markov model approach was employed for SCG signals processing that was followed by the expectation-maximization algorithm to learn the characteristic points. Subsequently, Viterbi sequence of the signal obtained using Viterbi algorithm was used to estimate instantaneous HR, HRV indices, and cardiac time intervals. Siecinski *et al.* studied the influence of the heartbeat detector on the SCG signals by comparing the HRV indices measured from SCG and ECG signals of CEBS database [118]. In [119], SCG signals recorded in supine, 45 degree heads-up, and sitting postures were used to estimate HR during low- and high-lung volume phases. Respiratory flow rate was used to extract lung volume information, which was then employed to measure HR during the high and low phases of lung volume.

  Recently, a automated procedure based on unsupervised learning was presented for HR and HRV monitoring [120]. Initially, the process started with self-calibration solely relying on the SCG signals to extract suitable heart template for each subject. Subsequently, beat detection and timing annotation were performed to estimate HR and HRV indices. In [121], a comparison was investigated between the time domain, frequency domain , and nonlinear HRV indices measured from ECG, SCG, and gyrocardiogram (GCG) signals. The signals were recorded from 29 healthy male volunteers and 30 patients with valvular heart diseases. The cardiac intervals for HR and HRV were measured in three steps: (i) R-peaks on ECG were detected using Pan–Tompkins algorithm, (ii) detection of local maxima in the SCG and GCG signals within 100 ms window from the occurrence of R-peak, (iii) computation of intervals between each heartbeat for each signal.

- **Extraction of cardiac time intervals:**

  In the existing literature, various methods demonstrated the utility of SCG for the estimation of cardiac time intervals (CTIs), including PEP, LVET, systolic time, and diastolic time [42, 50, 117, 122–125]. The timing information of cardiac activities and phases provides valuable insight into various cardiovascular functioning, and opens a broad range of applications for cardiac disease diagnosis and cardiac health assessment. Dehkordi *et al.* presented a comparative study to assess the accuracies of estimated CTIs based on SCG, PCG, and impedance cardiography (ICG) in relation to multimodal echocardiography technique [122]. The results showed that, compared to ECHO, the accuracies of estimated PEP were 86%, 43%, and 43%, for SCG, PCG, and ICG, respectively. The study concluded that SCG performs superior in comparison to the other techniques in the estimation of systolic time intervals. In [50], a regression model was employed to accurately measure the PEP using SCG. Multiple SCG sensors were placed at different measurement sites to extract the set of timing features. The extracted features were averaged and used along with their corresponding PEP to train the model. Then, the trained model was used to measure PEP. Another regression model for PEP estimation was introduced by Shandhi *et al.* [123]. Here, the model was trained using the data of accelerometer-based and gyroscope-based SCG signals. The combination of both accelerometer and gyroscope was observed to yield a more accurate estimation of PEP. In [124], patients suffering from myocardial infarction were considered for CTI. Several systolic, diastolic, and global parameters (myocardial performance index, blood pressure, HR) were estimated during an exercise tolerance test involving ECG. Tavakolian *et al.* presented a SCG-based method to predict the diastolic time [126]. They extracted AC and MC fiducial points and used ECG Q-wave as a reference to measure the timing information. The average duration of timing intervals spanning 15 cardiac cycles from Q to AC and Q to MC were employed to approximate the onset and end of the diastole phase, respectively.

- **Investigating cardiac-diseases using SCG:**

  The delineation of SCG has also been investigated to diagnose various cardiac diseases, such as valvular defects, atrial fibrillation, atrial flutter, hypertension, myocardial ischemia, and coronary artery diseases [22]. An overview of existing SCG applications associated with cardiac diseases diagnosis are presented in Table 1.2. The table clearly depicts that SCG's fiducial points along

Table 1.2: Details of SCG fiducial points used for cardiac disease diagnosis.

Reference	Extracted fiducial points	Cardiac disease diagnosis
[96]	MC, AO, AC, MO	Valvular disease
[52, 54, 127, 128]	AO-AO interval	Atrial fibrillation
[70]	MC, AO, AC, MO	Atrial flutter
[84]	PEP, LVET	Hypertension
[29, 30, 115]	AO and RF amplitudes	Myocardial ischemia
[49, 50]	PEP, Normalized-SCG	Heart failure
[129]	LVET interval	Haemorrhage
[129]	AC, MC	Myocardial infarction

with timing intervals covers a wide range of applications for assessing cardiac health.

- **Extracting respiratory patterns using SCG:**

  The morphology of an SCG waveform is influenced by various respiratory conditions. Therefore, SCG can serve not only to measure cardiac health but also for assessing lung condition [22]. In existing studies, few works have utilized SCG or ECG to detect respiratory inhale and exhale phases, breathing rate sequence, and sleep apnea [46, 108, 130–132].

  Pandia *et al.* devised an approach to investigate the cardiac vibrations for the recognition of cardiopulmonary health [130]. The frequency domain analysis of the SCG signal was used to study the impact of respiration, which prove more effective than time-domain analysis. In [108], SVM was used to identify the breathing phases by incorporating the average value of 512 data points from each systolic-profiles as a feature. In their follow-up work, they segmented SCG beats into blocks of equal size in both temporal and spectral domains, and used their average values as features [46]. Subsequently, optimal features group with higher identification rate was selected using SVM. Alamdari *et al.* computed lower and upper envelopes of the SCG signals to yield two accelerometer-derived respiration (ADR) signals [131]. For each volunteer, the piecewise total harmonic distortion (THD) metric was employed to determine the best ADR among the lower and upper envelopes for detecting respiratory phases. The results confirmed that the suggested ADR technique can accurately detect respiratory phases of heartbeats with an accuracy above 75%. In [132], a preliminary work was presented where the morphological structure of SCG was studied under the influence of respiratory phases, such as normal and stopped breathing. The interbeat variability of the SCG signal was estimated and employed to

demonstrate morphological differences at various breathing levels. The features such as kurtosis (K), energy of diastolic region (DE), amplitude of IM or IC instant (IA) were also explored in the work to classify the breathing levels. The combination of these three features showed robust performance in classification of the breathing levels.

Thus, the information captured by SCG has been explored for numerous clinical applications including cardiac health monitoring and diagnosis of heart anomalies. Despite its advantages and wide range of applications, it still has some limitations. SCG waveforms are usually corrupted from various environmental interferences. Among all the noises, SCG is mostly vulnerable to motion artifacts and its elimination is a challenging task. Since SCG gets easily affected by motion artifacts, its utility may get limited for the states other than the resting condition. Also, many factors affects the morphology of the waveforms, such as sensor positioning, heart displacements, body postures, gender, and age. The feature points are yet to be standardized and needs universal acceptance before deploying it into medical field. Furthermore, acquiring SCG needs direct contact of the sensor to the body. The signal recording process may cause discomfort and limit the movement of the user. Also, placement of the sensor for female subject may be a bit challenging.

Looking into the limitations of the SCG, the research direction is diverged to a more advance technology for cardiac parameters estimation. The objective of this thesis is to consider all the ergonomic aspects while estimating the vitals for health monitoring in clinical and non-clinical settings. To overcome the problems of SCG, a non-contact camera based method is studied and investigated for cardiac parameter measurements. The non-contact functionality of the camera will play a significant role by providing comfort during continuous health monitoring. In addition, the camera sensors are widely available and easily accessible by any individual. Thus, the second part of the thesis is documented on this technology which is named as remote photoplethysmography (rPPG). The rPPG technology is explained in details in the follow-up Section 1.2.

1.2 Remote Photoplethysmography

Remote photoplethysmography (rPPG) is an emerging technology in computer vision that facilitates non-contact monitoring of cardiac health using widely available cameras. Over the past two decades, this domain has grown rapidly, with an exponential increase in research undertaken along with technology. This technology has a vast potential in the future of digital healthcare due to its low cost, ease

of use and convenience. Hence, we explored this technology for the estimation of cardiac parameters for both clinical and non-clinical environments. This section provides a summary of the rPPG technology, covering its foundation, state-of-the-art computational methods, and subsequently challenges and applications associated with the method.

1.2.1 rPPG Signal

The characteristics of rPPG signals is identical to the conventional PPG signals. The sole difference lies in the recording process of the signals: one is obtained using contact sensors, while the other is recorded without any contact using imaging devices such as professional, mobile, or web cameras. Therefore, the signal recorded by the camera is termed remote PPG, signifying the generation of the PPG signal from a remote location. The PPG signal is one of the cardiac waveforms that captures the activities of cardiovascular system, reflecting variations in blood flow volume in the microvascular bed beneath the skin [133]. The waveform is generated based on optical properties, such as absorption, scattering, and transmission properties of measurement sites under a specific wavelength of light. The PPG signal is composed of a pulsatile (alternating current (AC)) component and a non-pulsatile (direct current (DC)) component. The pulsatile component reflects cardiac synchronous variations in the blood volume with each heartbeat. It can be used to detect ventricular fibrillation and ventricular tachycardia [134]. The non-pulsatile component is a much larger slowly varying quasi-static baseline that is affected by biological characteristics, including tissue composition, basic blood volume of the measurement site, and external factors such as measurement device specifications and ambient light. It carries valuable information about respiration, thermoregulation, sympathetic nervous system activities, and venous flow [134]. Fig 1.5 shows the principle of PPG signal generation and temporal characteristics of the waveform. Here, the PPG signal is derived by inverting the recorded light intensity from a photodetector after the transmission or reflection of light through human tissue.

The frequency range of the pulsatile DC and AC components of PPG ranges from [0.15-0.5 Hz] and [0.5-5 Hz], respectively [134]. Similar to other cardiac waveforms, PPG also reflects both systolic and diastolic phases of a cardiac cycle as shown in Figure 1.5. Various features of PPG are derived using the contour of the waveform, such as systolic amplitude, pulse width, pulse area, pulse-to-pulse interval and many more. All these features are applied in clinical applications for real-time monitoring of various physiological parameters, including HR, BP, cardiac output, breathing rate, and oxygen saturation. In addition, the PPG is utilized for the assessment of vascular (arterial diseases, stiffness, and ageing)

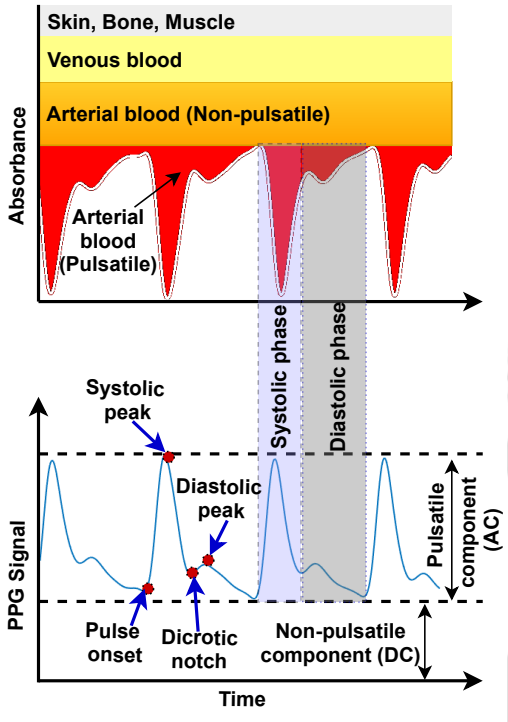


Figure 1.5: Principle of PPG generation and temporal characteristics.

and autonomic function (HRV/pulse rate variability (PRV)) [135]. Figure 1.6 displays the features extracted directly from PPG signal and its derivatives. The definition and clinical significance of these features are listed in Table 1.3. Since rPPG replicates the PPG waveform, all the characteristics of PPG can be mapped to rPPG. Consequently, the applications of PPG can be explored remotely using the rPPG waveform.

Acquisition of rPPG signals

The working of rPPG is based on the optical principle of the conventional PPG system that relies on the absorption spectrum of haemoglobin and other tissue components [135]. The system usually involves a light source (red, green, or IR light) and a photodetector to function. The measurement site is illuminated by a dedicated light source, and the photodetector senses the small variations in the reflected or transmitted light intensity due to the change in blood volume. The magnitude of the light intensity is directly proportional to the amount of blood present at the peripheral site. Thus, the PPG technology is able to acquire a pulsatile waveform that reflects synchronous blood volume variation during systolic and diastolic phases of each heartbeat. In contrast to PPG, rPPG does not require a dedicated light source, and it is operational under ambient lighting conditions using the readily avai-

Table 1.3: Summary of existing photoplethysmogram (PPG) features and their clinical applications

Reference	Feature type	Feature	Feature description	Clinical use
[136]	Basic	Systolic amplitude	Maximum amplitude of the systolic phase	Pulsatile component of blood volume
[137]				Stroke volume
[138]				Local vasodilatation
[139]		Pulse width	Time interval between the x% of the maximum systolic amplitude	Systemic vascular resistance
[140]		Pulse Area	The total area of the PPG in a pulsation	Surgical skin incision
[141]			The area of the systolic section, or the area of the diastolic section, divided based on the dicrotic notch	Total peripheral resistance
[142, 143]		Pulse-to-pulse interval	The time interval between the maximum systolic amplitudes of two adjacent pulsations	Cardiac cycle
[144]			The time interval between the pulse onsets of two adjacent pulsations	Heart (or Pulse) rate variability
[145, 146]	Combined	Perfusion index	The ratio of the amplitude of the pulsatile component to the non-pulsatile component	Peripheral perfusion
[147]		Large artery stiffness index	Ratio of subject's height and the time interval between the systolic and diastolic peak	Arterial stiffness
[148]		PPG augmentation index	The ratio of the systolic peak amplitude to the diastolic peak amplitude of a PPG	Arterial stiffness
			The ratio of the difference between the systolic peak amplitude and diastolic peak amplitude to the diastolic peak amplitude of a PPG	
[149]		Pulse transit time	Time difference between the specific features of PPGs measured at two different body sites	Cuffless blood pressure
[150]	1^{st} derivative	Crest time	Time interval between the pulse onset and the first zero-crossing of the derivative	Longer in vascular disease or hypertension patients
[151]		ΔT	Time difference between the first and the second zero-crossing points proceeding in the positive to negative value of PPG derivative	Time taken for the blood ejected from the heart to pass to the peripheral blood vessel
[152–154]	2^{nd} derivative	b/a	Ratio of the amplitude of the early systolic negative peak to the amplitude of the early systolic positive peak of SDPTG	Proportional to the stiffness of blood vessels, and increases with age; Inversely related to lead poisoning; Proportional to the Framingham risk score

Continued on next page

Table 1.3 – *Continued from previous page*

Reference	Feature type	Feature	Feature description	Clinical use
[152, 155]	2^{nd} derivative	c/a	Ratio of the amplitude of the late systolic re-increasing peak to the amplitude of the early systolic positive peak of SDPTG	Vascular stiffness, and decreases with age; Identifying hypertensive patients
[152]		d/a	Ratio of the amplitude of the late systolic re-decreasing peak to the amplitude of the early systolic positive peak of SDPTG	Inversely proportional to vascular stiffness, and decreases with age; Evaluation of vasoactive agents
[152, 156]		e/a	Ratio of the amplitude of the early diastolic positive peak to the amplitude of the early systolic positive peak of SDPTG	Inversely proportional to vascular stiffness, and decreases with age

*SDPTG: Second derivative of PPG

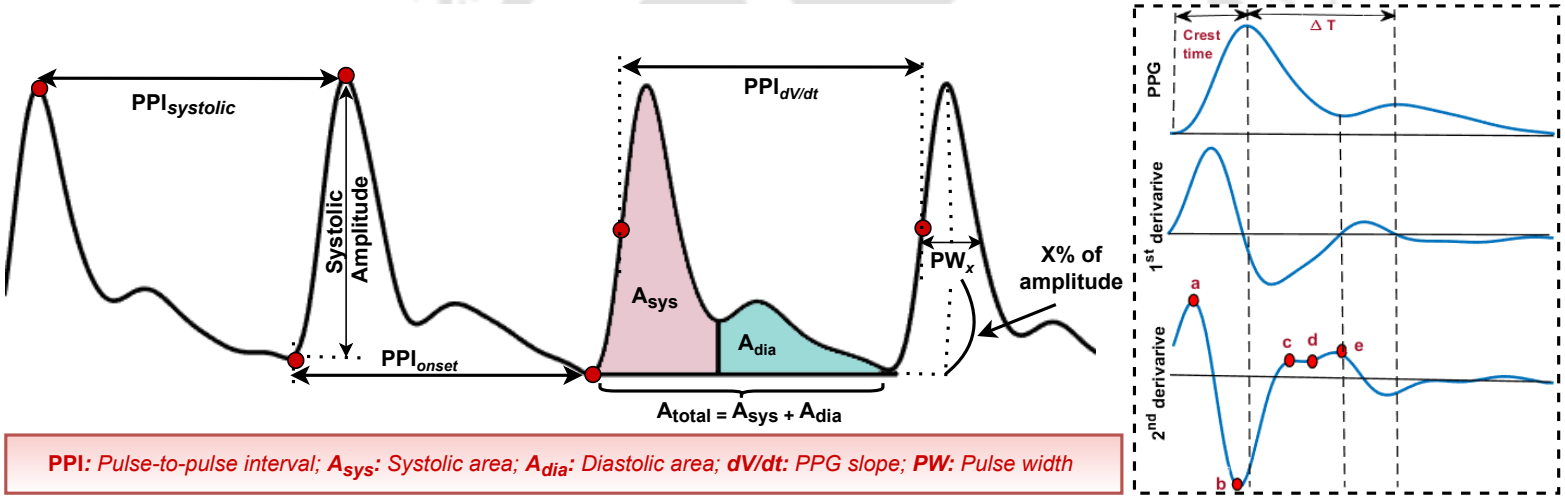


Figure 1.6: Basic features of PPG.

lable low-cost camera as a photodetector sensor. The camera sensor senses the reflected light intensity from the skin surface for the measurement of the rPPG signal. However, due to the involvement of a camera, the desired rPPG signal gets embedded in a series of video frames unlike in the case of PPG system where the signal is readily available. Therefore, this system involves various processing steps to extract a rPPG waveform. Figure 1.7 illustrates the schematic rPPG setup and the general pipeline for signal extraction. The acquisition setup involves a imaging device that records the facial video of a subject under natural or artificial light source. Subsequently, the recorded facial video is processed for the extraction of rPPG signal. The video can be processed using both conventional and deep learning techniques. The conventional algorithms consists of different steps, such as region of interest (ROI) selection and tracking, spatial pooling, color channel selection, and digital filtering schemes. In contrast, deep learning techniques involves 1D, 2D, or 3D neural networks for rPPG estimation task.

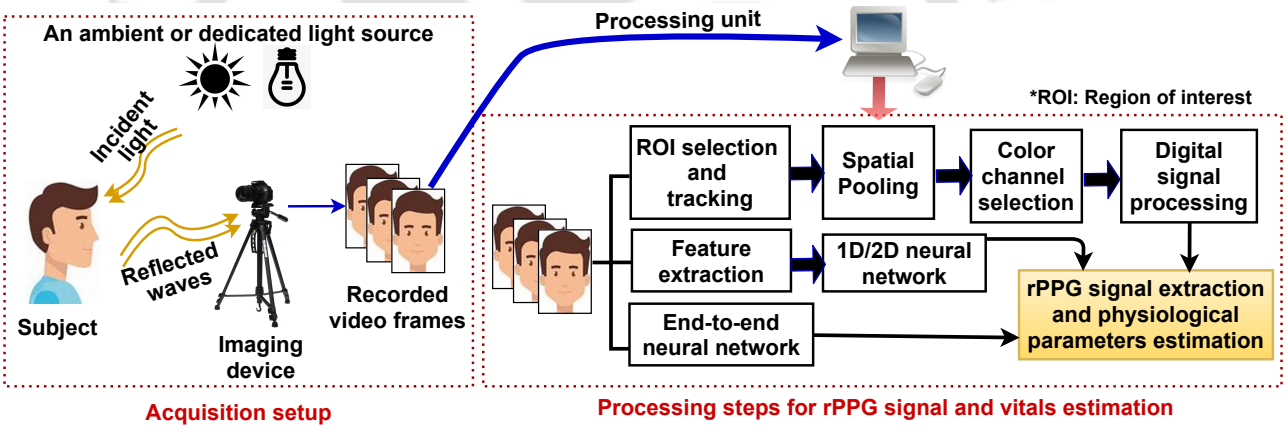


Figure 1.7: Schematic of rPPG system and basic building blocks of rPPG technology.

1.2.2 Physiology behind PPG waveform

The PPG waveform undergoes morphological changes in response to variations in subsurface blood volume pulses associated to cardiac activities. Additionally, changes in the waveform can also occur as a result of respiration, arterial activity, venous activity, and autonomous nervous system. Both systolic and diastolic phases of a cardiac cycle can be identified in a PPG waveform in correspondence to the amount of blood volume present at the peripheral sites. The PPG waveform and its characteristic points are depicted in Figure 1.5. It exhibits an ascending and a descending curve, indicating the rise and fall in capillary blood volume resulting from cardiac contraction and dilation, respectively. The rising and descending curve corresponds to the systole and diastole phase of the PPG waveform. The

initiation of pulsation, where the blood volume is minimum before the systolic phase, is represented by the onset point. The point at which the blood volume is at its maximum is indicated by the systolic peak. Following this, the blood volume starts to decrease, and the diastole phase takes place, marked by a transient rise and fall in PPG. The transient rise and fall occurs due to temporal elevation of blood volume in capillaries, caused due to a pressure gradient in the reverse direction to the blood flow, occurring just before the closure of aortic valve [157, 158]. At this moment, recessed instant is depicted as a dicrotic notch, and the point where the first derivative of the signal is nearest to zero following the systolic peak is identified as a diastolic peak [159]. The morphological structure of the AC component of the PPG waveform is primarily contributed by the blood flow volume associated with cardiac phases. In contrast, various physiological factors, such as respiration, vascular compliance, vascular tone, pain, and drug use contributes to the baseline or DC part of the PPG.

1.2.3 Parameters Affecting rPPG Morphology

In rPPG technology, the signal is derived from the recorded video, capturing the subtle skin colour changes imperceptible to the naked human eye. The subtle changes in pixel values within the video frames lead to relatively low amplitudes and SNR value in rPPG signal. These can be easily influenced by motion artifacts, illumination variation, facial expression, and inter-subject skin tone variability [160]. Consequently, these unwanted effects can easily distort the small pixels variations corresponding to cardiac and pulmonary events. In addition, there are significant ethical and privacy considerations related to rPPG measurements that needs to be taken into account before handling the signal.

Challenges from sensing point of view:

- Identifying the optimal qualities of an image sensor (sensor type, color filter array, bit depth, frequency bands specification, imager size, frame rate, and resolution) and assessing the sensitivity of measurements with respect to all these parameters is highly valuable. However, it is a difficult task to precisely characterize the impact of each property of the camera on rPPG measurements.

- Standardizing the distance between the camera and measurement sites is essential. Exploration of the recording setup and its impact on the signal quality needs further investigation.

- Regulatory approval is one of the biggest challenges in the rPPG measurements due to the

availability of wide range of cameras with varying specifications.

Challenges from environmental factors:

- rPPG measurements are highly susceptible to motion artifacts, rendering the system less practical compared to the conventional PPG approach. These motion usually includes rigid and non-rigid motions. Rigid motions are caused due to head rotation or translation, whereas eye blinking, mouth talking, and emotion expressing contributes to the non-rigid motions [161]. The existence of these motions has a significant impact on the desired rPPG signal, introducing undesirable periodical noises, displacing measurement sites in consecutive frames, and causing variation in reflected light. The removal of motion artifacts becomes more challenging when the frequency of unwanted noises overlaps with the signal of interest.

- The composition and dynamics of ambient lighting conditions alters the rPPG. Dynamic changes in ambient light intensity, position, and direction may yield relatively large pixel variations compared to those resulting from physiological events.

- Subject's body appearance, including skin-tone, facial structure, facial hair, and obstructions caused by apparel, also impact rPPG measurements.

  Additionally, the designed rPPG models for physiological measurements exhibit bias that are often challenging for detection. Most of the available datasets are recorded from fair skin type subjects, leading to methods being primarily evaluated on these data. This creates difficulties in generalizing the methods. Thus, significant validation of the camera-based technology is required before deploying it in real-world dynamic settings.

1.2.4 rPPG, its Extraction and Application - A Review

The rPPG-data is vulnerable to various environmental interferences, such as motion artifacts and illumination variation, as well as intrinsic quantization noise from sensors. Hence, the main objective of the rPPG system is to develop algorithms capable of isolating pulsatile information from environmental noises. Various methods have been proposed in the literature for extracting reliable rPPG signals, with some specifically designed for the direct estimation of cardiac vitals. The reported methods can be broadly classified into two categories: (i) Conventional signal processing approaches and (ii) Deep neural network (DNN)-based frameworks. In this section, we delve into the in-depth discussion of

state-of-the-art methods and potential applications of rPPG.

In the early 2000s, Blazek, Wu, and Hoelscher introduced the use of CCD-near infrared (NIR) imaging system for the first time to measure cardiopulmonary signals [162]. Their work established an evidence that an imaging system can be used to visualize skin vessels and analyze local changes in peripheral blood volume. After a short span of time, Verkruysse *et al.* studied this imaging approach, employing a digital consumer-level visual band RGB camera [163]. Since then, extensive research has been conducted, consolidating the rPPG concept and leading to the emergence of a new research domain in non-contact physiological measurement [160, 161].

All computational rPPG algorithms are designed based on the principle of optical models, such as the Beer-Lambert law (BLL) [164, 165] and Shafer's Dichromatic Reflection Model (DRM) [166]. These models serve as the foundation for modeling lighting, imagers, and physiology. The Figure 1.8

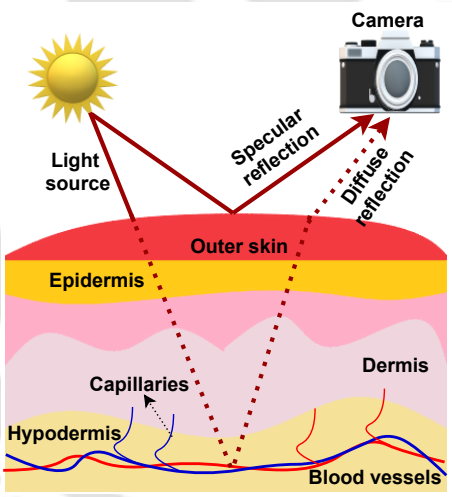


Figure 1.8: Principle of rPPG based on the dichromatic reflection model (DRM).

illustrates the DRM model which is the basis for comprehending the algorithmic principles behind rPPG. Assume the light source with a constant spectral composition but varying intensity. The RGB values of the k-th skin pixel in a video frame is defined by a time-varying function as [166]:

$$\mathbf{C}_k(t) = I(t) \cdot (\mathbf{v}_s(t) + \mathbf{v}_d(t)) + \mathbf{v}_n(t) \tag{1.1}$$

where $\mathbf{C}_k$ denotes a vector of the RGB values captured by the camera, $I(t)$ is the luminance intensity level that varies (i) with the light source and (ii) with the distance between the light source, camera, and skin-tissue. $I(t)$ is modulated by the specular reflection $\mathbf{v}_s(t)$, which is a mirror like reflection from the skin surface, and the diffuse reflection $\mathbf{v}_d(t)$, the absorption and scattering of light in skin

tissues. $\mathbf{v}_n(t)$ represents the quantization noise of the camera sensor, which can be effectively mitigated by averaging certain pixels. $\mathbf{v}_d(t)$, $\mathbf{v}_s(t)$, and $I(t)$ can be decomposed into a stationary and a time-dependent part using linear transformation.

Therefore, the diffuse reflection $\mathbf{v}_d(t)$ can be written as:

$$\mathbf{v}_d(t) = \mathbf{u}_d \cdot d_0 + \mathbf{u}_p \cdot p(t) \tag{1.2}$$

where $\mathbf{u}_d$ is the skin-tissue unit color vector, d_0 is the stationary reflection strength, $\mathbf{u}_p$ is the relative pulsatile strength caused by haemoglobin and melanin absorption, and $p(t)$ denotes the underlying physiological signals of interest.

The sepcular reflection $\mathbf{v}_s(t)$ can be expressed as:

$$\mathbf{v}_s(t) = \mathbf{u}_s \cdot (s_0 + \Phi(m(t), p(t))) \tag{1.3}$$

where $\mathbf{u}_s$ is the unit color vector of the light source spectrum; s_0 and $\Phi(m(t), p(t))$ depicts the stationary and varying parts of specular reflections; $m(t)$ represents all the non-physiological variations associated with flickering of the light source, facial expressions, and head rotation.

The luminance intensity $I(t)$ can be expressed as:

$$I(t) = I_0 \cdot (1 + \Psi(m(t), p(t))) \tag{1.4}$$

where I_0 is the stationary component of the luminance intensity, and $I_0 \cdot \Psi(m(t), p(t))$ is the intensity variation captured by the camera. The interaction between physiological and non-physiological motions are usually represented by non-linear functions, $\Phi(\cdot)$ and $\Psi(\cdot)$. The stationary skin reflection can be represented using a single component, formed by combining the stationary components from both specular and diffuse reflections:

$$\mathbf{u}_c \cdot c_0 = \mathbf{u}_s \cdot s_0 + \mathbf{u}_d \cdot d_0 \tag{1.5}$$

where $\mathbf{u}_c$ indicates the unit color vector of the skin reflection and c_0 is the reflection strength. Using Eqs. 1.2 to 1.5, $\mathbf{C}_k$ can be derived as:

$$\mathbf{C}_k(t) = I_0 \cdot (1 + \Psi(m(t), p(t))) \cdot (\mathbf{u}_c \cdot c_0 + \mathbf{u}_s \cdot \Phi(m(t), p(t)) + u_p \cdot p(t)) + \mathbf{v}_n(t) \tag{1.6}$$

The products of time-varying components are relatively small and they can be discarded from Eq. 1.6,

resulting the approximated $\mathbf{C}_k$ as:

$$\mathbf{C}_k(t) \approx \mathbf{u}_c \cdot I_0 \cdot c_0 + \mathbf{u}_c \cdot I_0 \cdot c_0 \cdot \Psi(m(t), p(t)) + \mathbf{u}_s \cdot I_0 \cdot \Phi(m(t), p(t)) + \mathbf{u}_p \cdot I_0 \cdot p(t) + \mathbf{v}_n(t) \quad (1.7)$$

The goal of the camera-based physiological measurement algorithms is to extract signal of interest, $p(t)$ from vector of RGB values, $\mathbf{C}_k(t)$. The traditional rPPG algorithms have ignored $p(t)$ inside $\Phi(\cdot)$ and $\Psi(\cdot)$ by assuming a linear relationship between $\mathbf{C}_k(t)$ and $p(t)$. This assumption holds true only when the measurement site remains stationary under constant lighting conditions i.e., the variable $m(t)$ is small. However, in practical dynamic conditions, $m(t)$ is not negligible, and the linear assumption could degrade algorithm's performance. To capture a broader and more complex relationship between $\mathbf{C}_k(t)$ and $p(t)$, various machine learning techniques have been investigated.

1.2.4.1 Signal processing Approaches

Classical signal processing methods offer several advantages for estimating vital signs from video data: (i) simplicity of implementation, (ii) computational efficiency, (iii) easy interpretability, and (iv) relatively transparent. The general framework of the signal processing-based approach for measuring physiological parameters using a camera is illustrated in Figure 1.9 and explained in detail below.

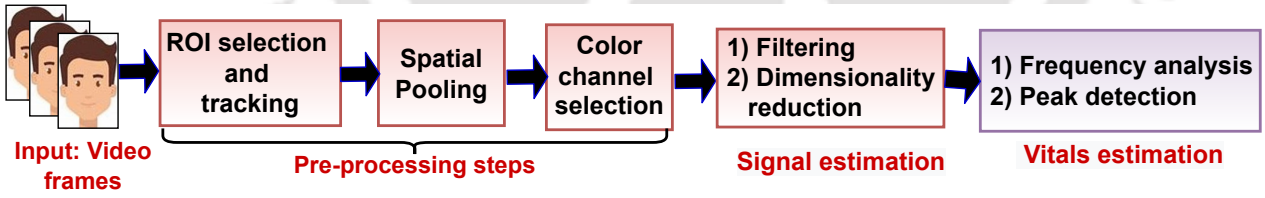


Figure 1.9: General framework of traditional rPPG algorithms.

Input Video data:

Researchers have investigated rPPG algorithms using video data recorded with cameras having different specifications and under dynamic environmental conditions. Most of conventional rPPG algorithms were assessed on privately owned databases. However, a wide range of public datasets are available for rPPG experimentations. Table 1.4 provides the summary of publicly available rPPG datasets. Each dataset consists of video recordings with simultaneously recorded physiological signals. The videos available within each dataset are of different resolutions and frame rates. Each dataset covers

Table 1.4: Summary of rPPG datasets

Dataset	Subjects	Video Specifications			Gold standard
		No.	Resolution	fps	
MANHOB	27	527	780*580	61	ECG, EEG, Breathing
VIPL-HR	107	3130	960*720/1920*1080	25/30	PPG, HR, SpO2
COHFACE	40	160	640*480	20	PPG
UBFC-rPPG	42	42	640*480	30	PPG, PR
UBFC-Phys	56	168	1024*1024	35	PPG, EDA
MR-NIRP	18	37	640*640	30	PPG
OBF	106	212	1920*1080	60	PPG, ECG, BR
PURE	10	59	640*480	60	PPG, SpO2
rPPG	8	52	640*480/1920*1080	15	PPG, SpO2
PFF	13	85	1280*720	50	PR
CMU	140	140	25*25	15	PR
MVPD	13	***	*****	30	ECG, PPG, BR, SpO2, BP, HR, temperature
AFRL	25	300	658*492	120	ECG, PPG, BR
MMSE-HR	140	102	1040*1392/ 640*480	25	HR, BR, BP
MR-NIRP-Car	19	190	640*640	30	PPG, SpO2
V4V	179	1358	1040*1392	25	HR, BP, BR
PFF	13	85	1280*720	50	PR
VicarPPG	10	20	720*1280	30	PPG
ECG-Fitness	17	207	1920*1080	30	ECG
MMVS	129	762	1920*1080/ 640*480	25	PPG, BR, BP
BSIPL-RPPG	37	***	640*480	30	PPG
RGB Video I	25	***	658*492	120	ECG, PPG
RGB Video II	18	***	1920*1080	24	PPG
Infrared Video	12	***	640*240	62	PPG
FaceForensics++	43	1000	*****	30	*****
SCAMPS	2800	2800	72*72	*****	PPG, PR, BR, Action Units

various dynamic conditions, including skin tone, lighting conditions, and physical states. These public datasets serves a transparent test sets for fair evaluation of methods and benchmarking.

Pre-processing steps

The pre-processing phase involves ROI selection and tracking, spatial pooling, and color channel selection, as explained below.

- **ROI selection and tracking:** Initially, the most appropriate region of interest (ROI) is identified and located in the first frame of the video and later the same ROI is tracked and detected in the consecutive frames using various tracking algorithms. As rPPG algorithms focus on the facial region as the ROI, it is essential to determine the bounding box of face. In most of the work,

Viola-Jones (VJ) algorithm [167] is employed to extract the bounding box of the facial region. Additionally, other popular face detection libraries, such as Dlib [168], discriminative response map fitting (DRMF) [169], Multi-Task cascaded convolutional networks (MTTCNN) [170], and MediaPipe face mesh [171] are also being utilized for the task. A fixed subset of the facial bounding box has been chosen by some researchers as the ROI, typically excluding regions with low pulsatile information, such as the eyes, lips, eyebrows, and background pixels.

The ROI identified in the initial frame is tracked across the video to ensure that the pixels within the ROI remain invariant to subject motion. The tracking algorithm updates the location of ROI in each frame, thereby reducing computational time by avoiding the re-detection of the ROI. Various tracking algorithms, including good-features-to-track [172], Kanade-Lucas-Tomasi (KLT) feature tracker [173], kernel-based object tracking [174], and tracking-by-detection with kernels (CSK) [175], have been employed to track the ROI.

- **Spatial pooling:** The conventional PPG provide blood volume variation directly in 1D signal form. However, due to utilization of a camera in rPPG technology, the required signal needs to be extracted from the sequence of 2D frames using the selected ROI. This involves averaging of pixels within the ROI in each frame, yielding the generation of raw RGB traces. This process is known as spatial pooling and it significantly improves the signal to noise ratio (SNR) in the presence of minor motion artifacts [135].

- **color channel selection:** The generated raw RGB traces show different strengths of the pulse wave and it is found that the green channel has better SNR due to its greater absorptivity for both oxyhemoglobin and deoxyhaemoglobin compared to other two channels [163]. Therefore, several works were done with the single green channel to study the remote pulse signal [161,176]. However, some researchers extracted the pulse waveform by a linear combination of RGB channels using blind source separation (BSS) techniques, including independent component analysis (ICA), and principle component analysis (PCA) [177–180]. They suggested that incorporating three channels and employing the BSS approach could further enhance the quality of the estimated pulse signals by reducing the motion artifacts.

Signal estimation

- **Filtering:** Despite ROI tracking and spatial pooling, the raw traces may include external noises caused due to subject's body movement, illumination change, intersubject skin tone variability, facial expression, and other factors. Many researchers have applied one or more digital filters on the raw traces by utilizing the feasible frequency range information of HR and expected noise sources. The purpose of filtering step is to increase the SNR of the desired physiological waveform. Typically, this filtering step is performed prior to dimensionality reduction. However, in some cases, it has been applied post dimensionality reduction or both before and after the reduction process. Most commonly bandpass filter has been utilized by researchers to eliminate both low- and high- frequency noises. A frequently selected frequency range is 0.7 Hz to 4 Hz, which corresponds to a HR ranging from 42 bpm to 240 bpm [177,178,181]. Apart from bandpass filter, researchers have experimented with several other digital filters and signal decomposition schemes, such as, moving average filter, adaptive filter, wavelet transformation, and fast Fourier transform (FFT)-based methods, to reduce the noises [182–184].

- **Dimensionality reduction:** It is assumed that a linear combination of raw RGB traces can result a more accurate and precise rPPG signal. Hence, dimensionality reduction methods are employed to reduce the three RGB traces into a single robust signal that encapsulates pulse information. The methods used to extract rPPG signal from RGB traces can be classified into two categories: (i) Blind source separation (BSS) approach [177, 178, 185–188] and (ii) Skin-reflection model based methods [166, 179, 189, 190].

  BSS methods, including ICA and PCA, determines the optimal combination of raw signals to recover the pulse signal, and their formulation can be expressed as follows:

$$\mathbf{S}(t) = \mathbf{W} \cdot \mathbf{C_k}(t) \tag{1.8}$$

  Where $\mathbf{S}(t)$ represents underlying source signals which includes the pulse and artifacts, $\mathbf{C_k}(t)$ is a vector of observed RGB traces, and $\mathbf{W}$ denotes the de-mixing matrix obtained from PCA or ICA. After BSS operation, the task is to determine the pulse signal by selecting the most periodic signal from the obtained source-signal $\mathbf{S}(t)$. The ICA approach was first introduced by Poh *et al.* where they used a laptop webcam to extract HR from the facial video recording of a subject [177]. In their work, they selected the second component produced by ICA as the

desired rPPG signal. However, the selection of a fixed ICA component cannot be generalized for all the cases. Later, the limitation was addressed in a revised version of their algorithm [178]. Cardiac parameters were estimated from an ICA component having highest peak response in the measured power spectrum. In other studies, the preferred pulse signal was determined by evaluating the power spectral percentage attributed to the first harmonic [185, 186]. PCA is the second popular BSS algorithm that has been utilized to separate raw traces into linearly uncorrelated components and arranged them based on variance [187, 188]. The criterion used in the ICA method for selecting desired component can also be applied to the components generated by PCA. In [187] and [188], the authors utilize both ICA and PCA schemes for pulse rate estimation and reveal that PCA is computationally efficient and outperforms the ICA method. The authors investigated these BSS techniques only under a constrained environment, demonstrating their effectiveness in successfully eliminating noise. However, later, the focus got shifted from lab conditions to a variety of realistic settings. In these scenarios, researchers introduced model-based approaches to tackle more challenging conditions, including dynamic light changes and large body motions. In this direction the classical examples are chrominance (CHROM) [189], blood volume pulse (PBV) [179], and plane-orthogonal-to-skin (POS) [166] based methods. De Haan *et al.* designed the CHROM method [189] to eliminate specular reflections by building two orthogonal chrominance signals from the original RGB traces. The performance assessment demonstrated that the CHROM method outperformed the BSS-based ICA and PCA techniques when dealing with exercising motions. Further, in their follow-up research, they proposed a PBV-based method to enhance the motion robustness [179]. Using the signature of blood volume variation, this model distinguished pulse-induced color changes from motion artifacts in temporal RGB signals. The approach exhibited significant improvement compared to their previously proposed CHROM method, particularly in handling data with prominent motion artifacts. In 2017, Wang et al. proposed the robust physically-grounded demixing approach for rPPG recovery that was based on defining a plane orthogonal to the color space of the skin (POS) [166]. The evaluation of this approach involved a comparison with BSS approaches, CHROM, PBV, and other state-of-the-art methods, using a private dataset that incorporates dynamic challenges. Among all the methods, POS achieved the best performance, making it well-suited for situations involving motion. In another work [190] an orthogonal

matrix image transformation (OMIT) method was proposed to remove compression artifacts that involves face stabilization, dynamic selection of region of interest (ROI), and RGB to rPPG transformation. All these model based techniques used the prior knowledge about the physical and optical properties of the skin and physiological waveform pattern that eventually helps to improve the estimated signal quality. These traditional rPPG algorithms are summarized in Table 1.5.

Vitals estimation

- **Frequency analysis:** After recovering the rPPG signal, physiological parameters like HR can be estimated through frequency analysis. For this purpose, the pulse-train undergoes transformation from the time domain to the frequency domain, and the frequency index with the highest peak response in the measured power spectrum is selected as the estimate for the HR frequency. For this various approaches have been employed, including discrete Fourier transform (DFT), FFT, discrete cosine transform (DCT), Welch's method, STFT [177, 180, 181, 186, 211].

- **Peak detection:** Apart from frequency analysis, parameters can be estimated in the time-domain through peak detection. Usually, the signal undergoes refinement through interpolation before the peak detection process [178, 195, 212]. Using the peaks information, HRV indices, along with HR values, can be determined.

  Traditional signal processing methods have made significant contributions towards the growth of rPPG technology. However, these methods have similar pitfalls. They made naive assumptions on the underlying waveform properties to simplify the noise reduction process that may deviate from the practical scenarios and result in inaccurate measurements. Additionally, the pixel averaging in the pre-processing step overlooks valuable spatial and color space information crucial for generating reliable physiological signals. This motivated researchers to explore the benefits of supervised learning and deep neural models.

1.2.4.2 Deep neural network-based frameworks

In recent years, deep learning (DL) technique has been successfully utilized by researchers from different domains to solve practical problems. The DL technique can perform effectively if the designed network is trained with a large amount of data with various scenarios related to the desired task. This characteristic makes the DL method robust and flexible to solve rPPG problem under different pract-

Table 1.5: Summary of conventional rPPG algorithms

Reference/ Year	Signal Extraction			Signal estimation		Parameters	HR estimation	Dataset
	ROI detection	ROI definition	Raw signal	Filtering	Dim reduction			
[163]/ 2008	Manual	BB, FH	RGB	BPF	***	HR, BR	***	Private
[177]/ 2010	VJ [60% W]	Face	RGB	***	ICA	HR	FFT	Private
[178]/ 2011	VJ [60% W]	Face	RGB	MAF, BPF	ICA	HR, HRV, BR	Lomb periodogram	Private
[180]/ 2012	Manual	BB, FH	(RG/ GB/ RB)	BPF	ICA/ PCA	HR	FFT	Private
[191]/ 2012	VJ	Face	RGB	***	ICA	HR	FFT	Private
[185]/ 2013	Manual	Single point from head	RGB	Normalize, BPF	ICA	HR	FFT	Private
[192]/ 2013	***	Region below eyes	Monochrome	BPF	***	HR, SpO2, BR	FFT	Private
[193]/ 2013	Manual	SBB	G	***	***	HR	FFT	Private
[194]/ 2013	VJ [60% W : 80 % H]	Face	NIR	MAF	ICA	HR, BR	FFT	Private
[189]/ 2013	VJ+ SK	Skin regions	RGB	Normalize, BPF	Fixed linear	HR	FFT	Private

Continued on next page

Table 1.5 – *Continued from previous page*

Reference/ Year	Signal Extraction			Signal estimation		Parameters	HR estimation	Dataset
	ROI detection	**ROI definition**	**Raw signal**	**Filtering**	**Dim reduction**			
[195]/ 2014	FLM	FLM based	RGBCO	Detrending, BPF	ICA	HR, HRV, BR	FFT, PD	Private
[181]/ 2014	VJ + DRMF	FLM based	G	Centralize, detrending, LPF	***	HR	FFT	Private, MANHOB
[182]/ 2014	VJ	Cheeks	RG	BPF, Adaptive BPF	Adaptive GRD	HR	FFT	Private
[183]/ 2014	Manual + FLM	Multiple ROIs	RGB	Spatial pruning, adaptive BPF	PCA	HR	FFT, PD	Private
[186]/ 2014	VJ	SBB	y comp.	MAF, BPF	PCA	HR	DCT	Private
[179]/ 2014	VJ+SK	Skin regions	RGB	Normalize, BPF	Fixed linear	HR	FFT	Private

Continued on next page

Table 1.5 – *Continued from previous page*

Reference/ Year	Signal Extraction			Signal estimation		Parameters	HR estimation	Dataset
	ROI detection	**ROI definition**	**Raw signal**	**Filtering**	**Dim reduction**			
[164]/ 2015	Manual	Multiple facial patches	G	Detrending, MAF	Fixed linear	HR	Welch PSD	MANHOB
[196]/ 2015	Face tracking method	Face	G	****	****	HR	STFT	Private
[197]/ 2015	Deformable face model	Seven facial regions	G	BPF	****	HR, HRV	FFT, PD	Private
[198]/ 2016	****	Skin regions	RGB	2SR algorithm		HR	PD	Private
[199]/ 2016	GFTT	Mouth area	RG	CHROM algorithm		HR	FFT	Private
[200]/ 2017	****	Skin pixels	RGB	POS algorithm		HR	DFT	Private
[188]/ 2017	Manual	Forehead + Cheeks	RGB	Normalize, HPF	ICA/ PCA	HR	FFT	Private
[201]/ 2018	Grid-based approach	Multiple	LWIR	LPF, BPF, HPF	****	BR	FFT	Private
[202]/ 2018	DRMF	Cheeks	RGB	BPF	EEMD	HR	Welch PSD	MANHOB

Continued on next page

Table 1.5 – *Continued from previous page*

Reference/ Year	Signal Extraction			Signal estimation		Parameters	HR estimation	Dataset
	ROI detection	**ROI definition**	**Raw signal**	**Filtering**	**Dim reduction**			
[203]/ 2018	Cascade detector	Face	RGB	BPF	EMD	HR	FFT	Private
[184]/ 2019	VJ [60% W : 80% H]	Four Face regions	RGB	LMSAF	ICA/ PCA	HR, HRV,IBI, SpO2	FFT	Private
[204]/ 2019	VJ	Face	RGB	BPF	SSA	HR	STFT, PD	Private
[205]/ 2019	Face tracking method	Forehead + Right cheek	RGB	Laplacian pyramid, BPF, MAF	ICA	HR, PTT, BP	PD	Private
[206]/ 2019	****	Face	RGB	Detrend, Normalize	ICA	BP	PD	Private
[207]/ 2020	KCF	Skin regions	Monochrome	Gaussain smoothing, BPF	PBV + PCA	HR, SpO2, BR	FFT	Private
[208]/ 2020	CCNN	Nostril + Chest	Monochrome	Median filtering, BPF	****	BR, Apnea	Lomb-Scargle	Private

Continued on next page

Table 1.5 – *Continued from previous page*

Reference/ Year	Signal Extraction			Signal estimation		Parameters	HR estimation	Dataset
	ROI detection	**ROI definition**	**Raw signal**	**Filtering**	**Dim reduction**			
[209]/ 2020	MTCNN	Cheeks	RGB	BPF, LPF	CHROM	HR	STFT	Private, PURE
[210]/ 2021	VJ [70% W : 60 % H]	Multiple super pixels	RGB	BPF, LPF	***	HR, HRV, BR	FFT	Private
[190]/ 2023	SSD + DAN	Skin regions	RGB	Detrending, BPF, MAF	OMIT	HR	STFT	PURE, COHFACE, Face-Video-Database, UBFC, MANHOB

Note: VJ: Viola-Jones algorithm, BB: Bounding box, FH: Forehead, BPF: Band pass filter, MAF: Moving average filter, SBB: Subset of bounding box, SK: Skin detection, FLM: Facial landmark detection, DRMF: Discriminative response map fitting, LPF: Low pass filter, FFT: Fast Fourier transform, GRD: Green-red difference, STFT: Short time Fourier transform, LMSAF: Least mean square adaptive filter, SSA: singular spectrum analysis, PD: peak detection, CCNN: Cascade convolutional neural network, MTCNN: Multi-task cascaded convolutional networks, SSD: Single shot multibox detection network, DAN: Deep alignment network, KCF: Kernelized correlation filters.

-ical scenarios. Different deep neural networks (DNN) models involving both supervised and unsupervised approaches have been found in the current state-of-the-art. These studies include convolutional models [213–217], transformers [218–220], generative adversarial networks [221, 222], and contrastive learning [223, 224].

Chen and McDuff [213] were the first to propose a convolutional attention network (CAN)-based DeepPhys model for physiological measurement. They utilized both spatial information (appearance) and temporal dynamics (motion) of video data through which the network learns an appropriate ROI and recovers the blood volume pulse (BVP) and respiration waveforms. Niu *et al.* [214] also employed a CNN-based ImageNet network to estimate HR. They designed a novel spatiotemporal feature map using the raw RGB traces extracted from multiple ROIs. Further, they improved the network's performance by incorporating an attention mechanism in their succession work [215]. In 2019, Qiu *et al.* [216] were motivated by Eulerian video magnification (EVM) technique [225] and applied it to design a robust feature. The feature image was intended to enhance the blood flow information by amplifying the facial pixel color changes using EVM. Subsequently, HR was estimated by mapping the feature map to a CNN. Song *et al.* [217] constructed another type of feature map using the pulse waveform estimated from the conventional CHROM method and used it with a ResNet18 model to measure HR. Other than CNN models, transformer-based architectures are becoming a choice for many computer vision tasks. In 2021, Liu *et al.* [218] presented the preliminary work on remote physiological measurement using a transformer-based model. Aliicja *et al.* [219] have also applied a transformer model for breathing measurement. Later, Yu *et al.* [220] proposed an end-to-end PhysFormer network to enhance the rPPG representation using local and global spatiotemporal features. Although transformers have provided good accuracy in many works, they are computationally expensive, and training them for an extensive data sequence is challenging. Few researchers have investigated generative adversarial networks (GAN) to generate realistic rPPG waveforms. Hao *et al.* [221] presented a Dual-GAN that jointly estimates the BVP and noise distribution. BVP and noise GAN promote each other's capability and yield a noise-robust BVP waveform. PulseGAN [222] is another such example where authors used the chrominance signals as an input to the GAN model to estimate realistic rPPG pulse signals.

In contrast to all the above supervised learning methods where labeled data are required during the training process, few works have demonstrated rPPG estimation with unlabelled video data using un-

supervised learning scheme. The unsupervised learning methods are desirable in the research domain where synchronized input source and target datasets are difficult to obtain. Gideon *et al.* [223] presented a fully self-supervised approach that involves the resampling of videos using a saliency sampler module. The videos were resampled to create positive pairs with a matching HR and negative pairs with different HR. Further, the model was fine tuned in a supervised manner on a smaller dataset. Similarly, Sun and Li [224] proposed a Contrast-Phys unsupervised model that involved a 3D-CNN architecture for spatiotemporal rPPG (ST-rPPG) blocks extraction, and the training was performed with a contrastive loss to generate the rPPG waveform. In short, the above-discussed on DL techniques can be classified into two groups: (i) feature map based methods and (ii) end-to-end methods. The feature based methods [214–217, 221, 222] usually focus on deriving handcrafted features that serve as an input to a defined neural network. In this approach, the quality of the generated feature map affects the network's performance. In contrast, the end-to-end methods [213,218–220,223,224] directly learn the informative features from the raw input video data. However, this require a large dataset to fit the model and they are computationally expensive due to the involvement of 3D networks. The DNN-based rPPG methods are summarized in Table 1.6.

Table 1.6: Summary of DNN-based rPPG algorithms

Category	Network name	Dataset	Parameter	Reference
Feature-based	SynRhythm	MANHOB-HCI, MMSE	HR	[214]
	********	MMSE, VIPL-HR		[215]
	EVM-CNN	MANHOB-HCI, MMSE		[216]
	********	MANHOB-HCI, VIPL-HR, ECG-Fitness, UBFC-rPPG		[217]
	NAS-HR	VIPL-HR, PURE		[226]
	ViT-rPPG	COHFACE, UBFC-rPPG		[227]
	MSDN	COHFACE, VIPL-HR, PURE		[228]
	Dual-GAN	UBFC-rPPG, VIPL-HR, PURE	HR, HRV, BR	[221]
	PulseGAN	UBFC-rPPG, VIPL-HR, PURE, MANHOB-HCI, BSIPL		[222]
	********	MMVS, VIPL-HR, UBFC-rPPG		[229]

Continued on next page

Table 1.6 – *Continued from previous page*

Category	Network name	Dataset	Parameter	Reference
End-to-End	EfficientPhys	AFRL, UBFC-rPPG, MMSE, PURE	HR	[218]
	********	PURE, COHFACE, MR-NIRP-Car, UBFC-rPPG		[223]
	Contrast-phys	MR_NIRP, OBF, UBFC-rPPG, MMSE, PURE		[224]
	HR-CNN	COHFACE, MANHOB-HCI, PURE, ECG-Fitness		[230]
	Siamese-rPPG	COHFACE, UBFC-rPPG, PURE		[231]
	GraphPhys	VIPL-HR, MANHOB-HCI, MMSE, OBF		[232]
	MSSTNet	PURE, UBFC-rPPG		[233]
	ETA-rPPGNet	PURE, COHFACE, UBFC-rPPG, MMSE		[234]
	LSTC-rPPG	UBFC-rPPG		[235]
	TranPhys	UBFC-rPPG, COHFACE, VIPL-HR		[236]
	DeepPhys	RGB Video I, RGB Video II, MANHOB-HCI, Infrared Video	HR, BR	[213]
	PhysFormer	VIPL-HR, MANHOB-HCI, MMSE, OBF	HR, HRV, BR	[220]
	Shuffle-rPPGNet	UBFC-rPPG, PURE	HRV	[237]
	CL-SPO2Net	UBFC-rPPG, FaceForensics++, MVPD	SpO2	[238]

Potential Applications of rPPG

Physiological signals extracted from cameras can be employed for both clinical and non-clinical applications. Over the past few years, many research works have explored the rPPG technology, demonstrating its feasibility across various potential applications.

Clinical applications:

In clinical settings, rPPG can be utilized to continuously monitor various physiological parameters for health assessment. In the existing literature, various methods illustrated the use of rPPG for the measurement of physiological parameters, such as HR [161, 163–165, 177–184, 213–217, 220, 221], HRV indices [178, 184, 195, 220–222], BR [163, 178, 195, 201, 213, 219–221, 239–242], pulse transit time/ pulse wave velocity [199, 205, 243–245], blood perfusion [162, 225], BP [205, 206, 244, 246–250], atrial fibrillation [251, 252], blood oxygen saturation (SpO_2) [184, 207, 253, 254], glucose [255, 256], and systolic and diastolic peaks [212, 257, 258]. Contactless measurement of all these vital parameters is highly valuable for patients with fragile skin injury or infections. This approach can provide a convenient experience for infants and elderly patients, whether they are at home or in the ICU. Furthermore, rPPG is well-suited for telehealth, an emerging health assessment approach utilized by over half of the healthcare organizations in the U.S. [259]. The incorporation of telehealth with rPPG technology brings numerous advantages to society: (i) enables social distancing and ensures safety for frontline healthcare workers, particularly during pandemics; (ii) decreases the workload of hospitals; (iii) manages users' daily routines by minimizing the time required for clinic visits and reducing waiting hours for doctor consultations. In addition, telehealth facilitates vitals measurement from any location using a consumer-level camera, assisting in the early detection of symptoms related to various health problems. A list of articles that have investigated numerous clinical applications of rPPG is provided in Table 1.7.

Table 1.7: Reported clinical applications of rPPG

Applications	References
Neonatal monitoring	[193, 201, 239–242, 260–262]
Kidney dialysis	[253]
Sleep monitoring	[207, 208, 254, 263]
Health screening	[264]
Telehealth	[194, 251, 265]

Non-clinical applications:

Apart from clinical applications, rPPG finds extensive applicability in consumer-grade products. The potential non-clinical rPPG applications are as follows:

- **Affective computing:** As rPPG utilizes consumer level cameras, it has great potential for affective computing and human-computer interaction applications. It has been investigated for human affects interpretation, such as emotion recognition [252, 266], pain recognition [267–269], cognitive stress estimation [212, 270], and engagement detection [271]. Moreover, rPPG can also be used as polygraph during lie detection, as telling a lie involves the activation of autonomic nervous system (ANS), leading to variations in mental and physiological parameters [161].
- **Driver safety:** In the automotive sector, rPPG can be employed to reduce the number of road accidents. In this aspect, rPPG can monitor drivers and their physiological status, assisting in identifying factors such as drowsiness, illness and fatigue that may contribute to accidents. This monitoring can serve to alert drivers, enabling them to adjust their behavior to prevent accidents. Several research works based on rPPG schemes are conducted for monitoring drivers' health status [196, 202, 203, 209, 272–277].
- **Deepfake detection:** This technology has also grabbed attention for identifying fake videos and verifying the subject's liveliness. Few attempts have been made to identify deepfake videos by capturing anomalies in physiological signals [278, 278–280].
- **Face anti-spoofing:** Researchers have utilized rPPG for anti-spoofing systems [281, 282]. This advancement in the technology can be used as a biometric authentication tool. It helps in preventing intruders from bypassing core security conditions by using a fake ID similar to that of an actual user, all without physical contact. The potential of rPPG technology for security purposes is demonstrated in various articles [283–287].
- **Fitness monitoring:** Monitoring physiological conditions during fitness training can help individuals in avoiding over-exertion and adjusting the training process according to their needs and condition. rPPG technology seems more attractive for tracking significant vitals compared to contact based devices like smartwatches and digital bracelets, which may cause discomfort and pain during vigorous exercise. Few researchers have studied the performance of various rPPG methods under fitness training set-ups [179, 200, 204, 288, 289].

1.3 Motivation and objectives

This thesis examines two modalities for estimating cardiac vital parameters using signals, namely SCG and rPPG. The research on both the signals had made remarkable progress in the field of biomedical engineering. However, there are certain research gaps in both domains that prompted us to formulate the objectives for this thesis. The motivation for our research objectives arises from the benefits and existing research gaps in these technologies, which are outlined along with the thesis's work plan below.

SCG is immensely used in recent years due to technological advancements and miniaturization of accelerometers. In contrast to other cardiac signals, an SCG is able to capture more diagnostic information and has potential applications in clinical and non-clinical aspects. Also, SCG sensor is somewhat simple, lightweight, portable, user friendly, and cost-effective than the other cardiac sensing devices. The estimation of fiducial points from SCG are useful to derive essential parameters for disease diagnosis and cardiac health assessment. However, detecting and extracting the fiducial points is a crucial task due to presence of noises and inter- and intra-beat variabilities. In most of the studies the maxima point within a fixed window is considered to detect the prominent AO peaks. However, this windowing technique may produce inaccurate estimation for the case where the AO peak has smaller amplitude than the MC or RE peaks under a systole profile. In other state-of-the-art methods, the detection of SCG fiducial points rely on other cardiac modalities besides SCG. Acquiring more than one signal and estimating the location-specific feature points makes the system computationally complex and expensive. Also, many methods have utilized the fiducial points of other cardiac signals with the fiducial points of SCG to estimate and analysis the systolic time intervals. For example, R-peak of ECG along with AO peak of SCG is used to derive PEP. Also, some of the blood pressure (BP) estimation methods have introduced the use of concurrent SCG and PPG signals for the PTT estimation. For this purpose, they suggested the AO peak of the SCG as a proximal point and systolic peak or onset of the PPG as a distal point. The mapping of SCG events with other cardiac signals helps in the precise estimation of the critical points. However, involvement of different sensing modalities reduces the flexibility of the system and may cause discomfort to the subject. Motivated from these points, the first part of the thesis is oriented towards the estimation of clinically used cardiac parameters using only the feature points of SCG signal. The main objectives are given as follows:

- To utilize a standalone framework to detect various SCG fiducial points, and employing them in a mathematical model to estimate systolic blood pressure (SBP) and diastolic blood pressure (DBP).
- To study the prominent AO peaks for HRV analysis.
- To detect ventricular depolarization events from standalone SCG fiducial points using a deep feedforward network.

Contact-based cardiac modalities have played an important role in estimating vital signs. However, they rely on sensors that must be in direct contact with the body, limiting the flexibility and mobility of users. This approach may cause inconvenience and discomfort for subjects undergoing long term monitoring. Moreover, contact-based sensors are not appropriate for patients with burned skin and infants with fragile and sensitive skin.

As rPPG is the non-contact counterpart of PPG, all the cardiac parameters that are estimated using PPG can also be measured remotely using rPPG. After an extensive state-of-the-art literature survey, it is found that the research community has a significant interest in rPPG technology owing to its numerous advantages as listed below:

- Measurement comfortability due to noncontact functioning
- Ability to sense large area at a time
- No deformation of the blood vessels unlike in the case of the contact PPG finger probes, and
- Potential to be used ubiquitously with the increasing number of smart-phone cameras.

All these advantages makes rPPG as the most promising technology for remote health assessment. In recent years, the rPPG research community has significantly evolved and has developed several rPPG algorithms that demonstrate promising results, suggesting the potential to achieve accuracy comparable to traditional contact-based methods. However, the magnitude of the signal acquired remotely has low SNR compared to contact-PPG and the system is highly sensitive to environmental interferences. Most of the state-of-the-art methods are not efficient in the presence of dynamic noises. Majority of the methods were designed under certain assumptions for simplification of noise reduction. However, those assumptions may not fit for practical applications where noises are more diverse

and may lead to inaccurate results. Moreover, the rPPG system is only investigated to capture the blood volume variation that is associated with the cardiac mechanical phenomenon. However, till date, none of the studies have experimented to capture the electro-cardiac activity using the camera. Studying both electrical and mechanical events of heart can promote the rPPG technology beyond its fundamental contribution towards cardiac health assessment. Motivated by the benefits of rPPG technology, the goal of the second part of the thesis is to formulate robust algorithms for noise compensation and precise estimation of vitals. The objectives are outlined as follows:

- To develop a time-frequency based framework for reconstructing reliable rPPG signals, aiming to facilitate accurate HR values estimation.
- To devise deep neural networks (DNN) for generalizing the rPPG system in dynamic environments. The framework includes both 1D- and 2D- convolutional neural networks (CNNs) for HR and HRV analysis.
- To investigate electro-cardiac activity by generating rECG from video data alongside the pre-investigated rPPG signal, using a 3D-CNN based end-to-end bitask network.

1.4 Thesis Outline

The works contributed to the thesis are depicted in Figure 1.10. It is divided into three parts, which are signal recording using two different sensing modalities, fiducial points extraction from the signals, and their application for cardiac health assessment. The first block represents the data acquisition, where concurrent cardiac waveforms, such as SCG, ECG, and PPG are acquired from the corresponding sensors. Also, simultaneous facial video data is recorded along with ECG and PPG for rPPG investigation. In the second block, the approaches used to detect the prominent peaks of SCG and rPPG are illustrated. For extracting features from the signals both conventional and deep neural network-based methods are explored. Application of both the modalities for cardiac health-parameters estimation by analysing the extracted fiducial points are shown in the third block. The standalone delineated features of SCG are used to measure BP, HRV analysis, and systolic time intervals. The camera modality show its applications in HR measurement and HRV analysis. The investigated research works are organized in the thesis as follows: Chapter 2 presents the standalone approaches to delineate the fiducial points of SCG. It further shows the applications of the extracted feature points in the estimation of

cardiac parameters. Whereas, Chapter 3 presents a conventional method to extract rPPG waveforms from video data to estimate HR. In Chapter 4, 1D and 2D convolutional neural networks (CNNs) are explored for HR and HRV analysis from the estimated rPPG waveforms. Chapter 5 presents an end-to-end bitask network based on 3D-CNN to estimate rECG signals alongside with the rPPG waveforms. A summary of the contributed works and their future scopes are discussed in Chapter 6.

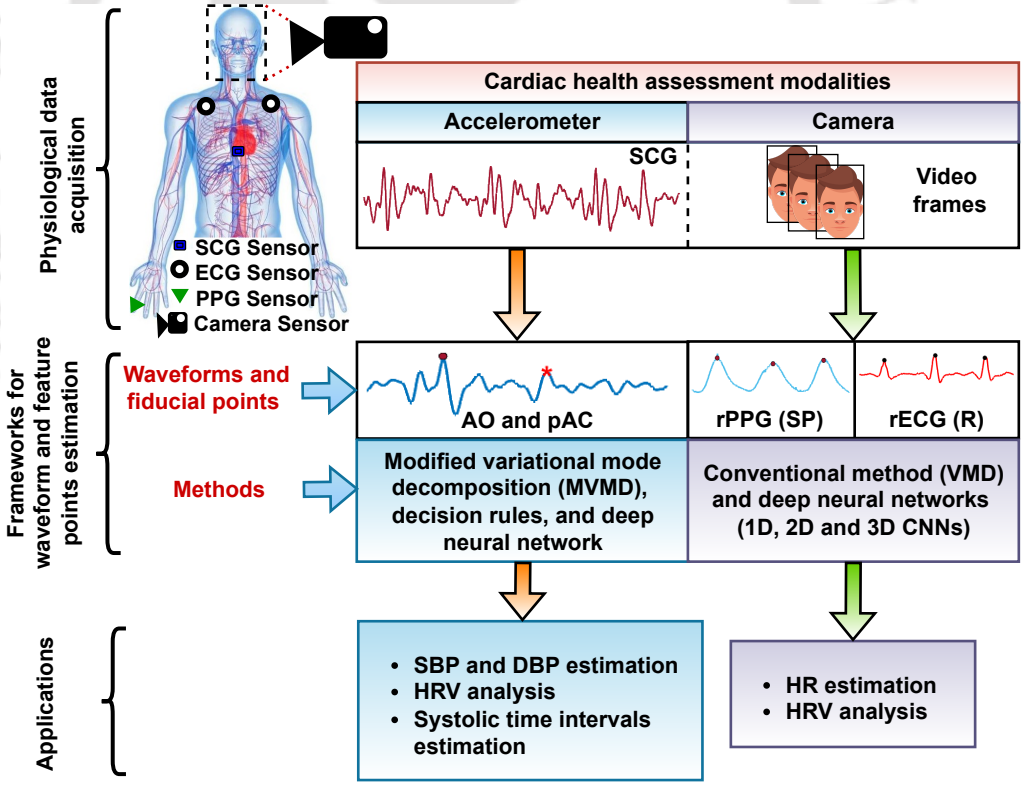


Figure 1.10: Graphical abstract representing thesis work-flow

2

Cardiac Parameters Estimation Using Seismocardiography

Contents

This chapter presents the estimation of cardiac parameters using the SCG signal. The parameters are derived from the fiducial points of SCG waveform, involving precise delineation and feature extraction of the signal. The delineation of the SCG signal is accomplished using a standalone approach that does not require any other reference cardiac signal, such as ECG. The method focused on estimating the aortic valve opening (AO peak) and post aortic valve closing (pAC peak) points. Usually, the AO peak is prominent within a cardiac cycle of the SCG waveform, and its detection is considered as a primary step that helps delineating other fiducial points. However, detecting the AO point is a critical task in an SCG signal. To address this issue, the proposed study employs a data-adaptive modified variational mode decomposition (MVMD) method and proposed simple decision rules to extract fiducial points. The extracted fiducial points are utilized for the application of BP measurement and HRV analysis. Another application is investigated, where fiducial points of SCG is combined with demographical information of subjects, to identify ventricular depolarization events using a deep feedforward neural network (DFN).

The chapter is organized as follows: The dataset used for the experimentation and evaluation of the methods is described in Section 2.1. The delineation algorithms for AO and pAC peaks are outlined in Section 2.2. Section 2.3 presents a mathematical model designed using SCG fiducial points to estimate BP. Estimation of HRV indices is illustrated in Section 2.4. Section 2.5 describes a DFN framework to generate R-peaks instants on SCG waveform. Finally, a summary is provided in Section 2.6.

2.1 Data Collection

A small electronic circuit board is employed for non-invasive acquisition of the SCG signals. The circuit board is designed using a low-cost and miniaturized MEMS-based tri-axial accelerometer sensor, a pre-amplifier, and a fourth-order Butterworth low pass filter (LPF). The utilized accelerometer is ADXL335 with a minimum full-scale range of ±3 g [290]. The SCG signals are acquired by mounting the sensor node near the lower end of the sternum with the help of an elastic chest-band. In addition, ECG, PPG, and ABP signals are also recorded for the experimentations.

Figure 2.1 shows the experimental setup for simultaneous recording of SCG, ECG, PPG and ABP signals. Data was recorded for 6 minutes from 10 healthy subjects (9 male and 1 female, age: 25-31 yrs, weight: 74±10 kg, height: 169±9 cm) in supine position. For all the subjects, recordings were done for ECG in Lead-II configuration using Ag-AgCl electrodes, z-component of the SCG signal at

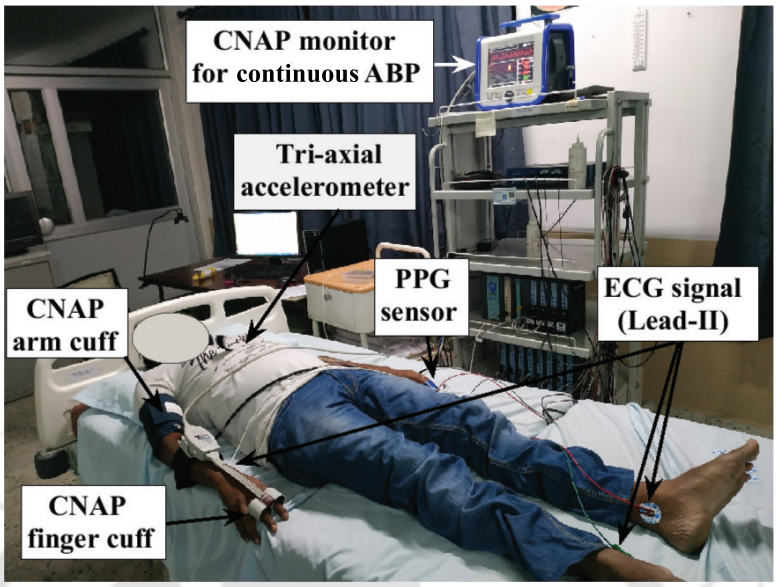


Figure 2.1: Experimental setup for signal acquisition.

the lower end of the sternum with self-built SCG acquisition circuit board [1], PPG at fingertip using Nonin medical's SenSmart® optoelectronic finger clip sensor, and ABP using finger-cuff provided by CNAP® Monitor 500. All signals, including reference ABP signal, were synchronously obtained using the MP150 DAQ system (BIOPAC Systems, Inc., Santa Barbara, CA, USA) at a sampling rate of 1 kHz. Data acquisition was performed at the Electro Medical and Speech Technology (EMST) Laboratory, IIT Guwahati with prior consent of the subjects and ethical approval from the institutional review board (IRB).

2.2 AO and pAC fiducial points detection

All the studies presented on SCG signal analysis involves the extraction of fiducial points, including AO and pAC peaks instants. These points are employed for various measurements of cardiac parameters that can be used for both clinical and non-clinical applications. The methodology for detecting AO and pAC is described in details in the subsequent sections.

2.2.1 AO peak detection

For the detection of AO peaks, a data-adaptive variational mode decomposition (VMD) is employed as a major tool on dorso-ventral SCG signals [1]. The VMD scheme can resolve a multicomponent signal $x(t)$ into its principal modes (m_k), nonrecursively [291]. Each of the decomposed modes is band-

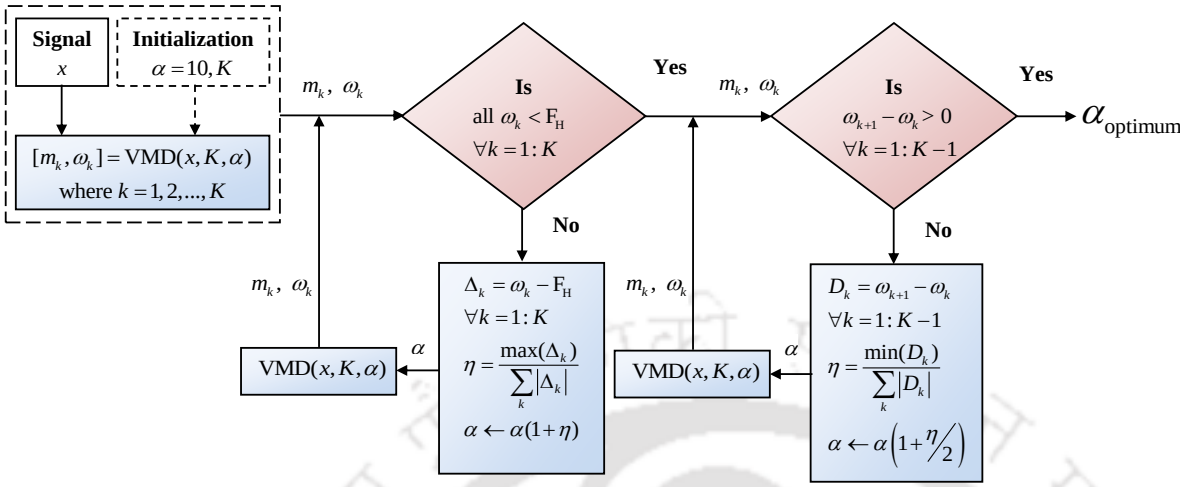


Figure 2.2: MVMD model for automatic adjustment of mode bandwidth constraint (α) for given number of modes (K) [1]. Note that F_H denotes the desired low pass cut off frequency for the VMD decomposition, and ω_k denotes center frequency of mode-k.

limited around a center frequency (ω_k). The constrained variational problem for the decomposition can be formulated as [291]:

$$\min_{m_k,\omega_k} \sum_{k=1}^{K} \left\| \partial_t \left([m_k(t) + j\widehat{m}_k(t)]\, e^{-j\omega_k t} \right) \right\|_2^2$$
$$\text{s.t.} \sum_{k=1}^{K} m_k(t) = x(t) \tag{2.1}$$

where $\widehat{m}_k = \frac{1}{\pi t} \otimes m_k(t)$ represents the Hilbert transform of mode m_k and k ($k = 1, 2, ..., K$) denotes the mode index. Here, '$\otimes$' is used as mathematical convolution operator. The unconstrained representation of above optimization problem [Eq. (2.1)] can be achieved using quadratic penalty term and Lagrange multipliers [291]. Now, the modified equation is expressed as follows:

$$\mathcal{L}(m_k, \omega_k, \lambda) = \alpha \sum_k \left\| \partial_t \left([m_k(t) + j\widehat{m}_k(t)]\, e^{-j\omega_k t} \right) \right\|_2^2 + \left\| x(t) - \sum_k m_k(t) \right\|_2^2 + \left\langle \lambda(t), x(t) - \sum_{k=1}^{K} m_k(t) \right\rangle \tag{2.2}$$

where α represents the balancing parameter of data-fidelity constraint. The VMD optimization problem is solved using alternate direction method of multipliers (ADMM) [291]. In the VMD algorithm, selecting an appropriate set of mode bandwidth constraint (α) and number of modes (K) is a challenging task. Thus, it needs further development to automatically control the allowable bandwidth (α) in each variational mode for a particular application. For this purpose, a modified VMD (MVMD) model was proposed to optimize the α parameter [1]. It involves two nested and recursive criteria, as shown in Figure 2.2. Through an example, the decomposition results obtained during the intermediate stages

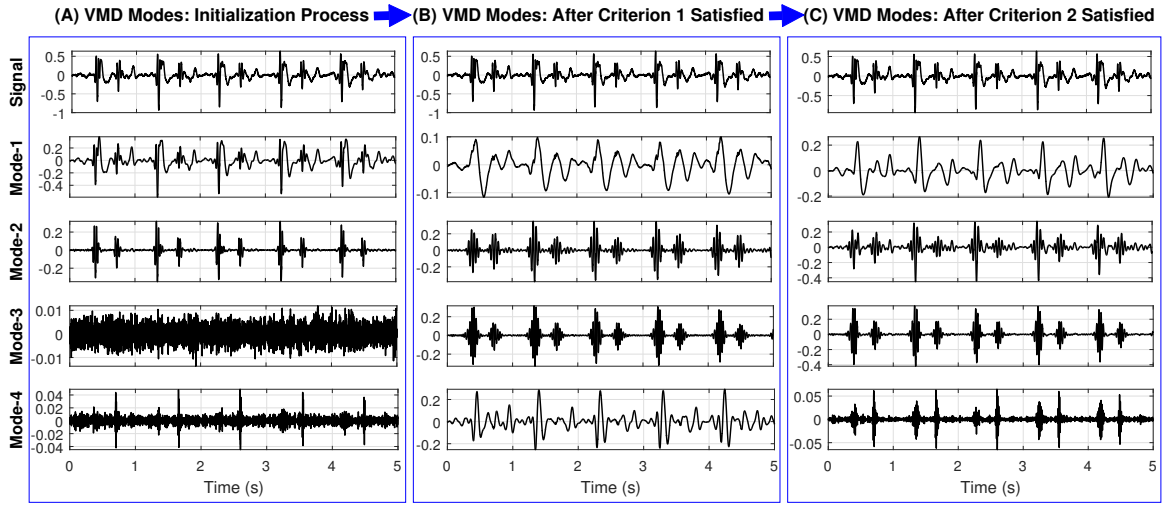


Figure 2.3: Decomposed modes during intermediate stages of the MVMD process (a) during initialization stage, (b) stage when criterion 1 is satisfied, and (c) stage when criteria 1 and 2 are satisfied and the mode bandwidth constraint is optimized.

of the model are shown in Figure 2.3. The pseudocode for AO peak detection is given in Algorithm 1. Although the method provides a robust detection of the AO peaks, it produces misdetections and false detections in some critical cases. To address this issue, a post-processing scheme is proposed.

2.2.2 Post-processing for AO peak detection

A post-processing technique is developed and employed in Algorithm 1 to improve the performance of the AO peak detection by reducing the false alarming conditions.

Let, $\boldsymbol{pks} = [p_1, p_2, \cdots, p_K]$ is a vector of detected AO peaks in a SCG segment. At first, the difference of each consecutive AO peak instants is calculated as:

$$d_k = p_{k+1} - p_k, \tag{2.3}$$

where, k ($k = 1, 2, \cdots, K-1$) denotes the peak index. Then, median-difference of these AO instants is computed as a reference, say M. The error between d_k and M is compared with a threshold (say TH). Usually, the following three false alarming situations are encountered in the AO peak detection process. These cases are shown via synthetically generated SCG amplitude envelopes for systolic and diastolic profiles in Figure 2.4.

(i) **Missing AO peaks:** This false negative condition is identified if $(d_k - M) > 3 \times TH$, and it can be avoided by determining the maxima as new AO peak in time-interval of $p_k + W$ to $p_{k+1} - W$ in the SCG, where $W = 60$ ms.

Algorithm 1 AO instant detection in an SCG signal

Input: SCG signal $x[n]$; $n = 1, 2, ..., N$

Low-Frequency Artifacts Removal

Decompose signal $x[n]$ using MVMD

$[m_i^I, \omega_i^I] = \text{VMD}(x, \alpha_1, K_1); \qquad i = 1, 2, ...K_1$

// where $m_i^I :=$ decomposed i^{th} mode of $x[n]$,

$\omega_i^I :=$ central frequency of m_i^I,

$\alpha_1 :=$ allowing mode-bandwidth,

$K_1 :=$ number of decomposed modes

Extract the detrended SCG signal

$x_d[n] = x[n] - \sum_i m_i^I[n]$

Systolic Profile Enhancement

Resolve useful signal components using MVMD

$[m_j^{II}, \omega_j^{II}] = \text{VMD}(x_d, \alpha_2, K_2); \qquad j = 1, 2, ...K_2$

Extract Gaussian derivavtive filtered modes (GDFMs)

$g_j[n] = d[m] \circledast m_j^{II}[n]; \qquad m = 1, 2, ...L - 1$

// where Gaussian derivative kernel $d[m] := g[m+1] - g[m]$,

Gaussian window $g[l] := \exp\left(\frac{-1}{\sigma^2}\left[\frac{l}{(L-1)/2}\right]^2\right)$; $0 \leq \mid l \mid \leq \frac{(L-1)}{2}$

Compute relative GDFM energy (RGE) for mode selection

RGE: $e_j = \frac{E_{g_j}}{\sum_j E_{g_j}}; \qquad j = 1, 2....K_2$

Reconstructed signal: $s[n] = e_{j^*}^2 g_{j^*} + \text{P}e_{j^*-1}^2 g_{j^*-1} + \text{Q}e_{j^*+1}^2 g_{j^*+1}$

// where e_{j^*} denotes maximum RGE and j^* is corresponding mode index

P=1, Q=0 $\leftarrow$ if $\mid e_{j^*} - e_{j^*-i} \mid < \rho; \qquad \rho \in [0, 1]$

P=0, Q=1 $\leftarrow$ if $\mid e_{j^*} - e_{j^*+i} \mid < \rho;$

P=0, Q=0 $\leftarrow$ otherwise

AO Peak Approximation

Construct envelope on systolic profiles

Difference envelope: $D_{env} = U_{env} - L_{env}$

// where U_{env}, L_{env}: upper and lower envelopes of s[n], respectively

Thresholding: $T_{env}[n] = (D_{env}^2[n] > \tau) \times D_{env}[n]$;

$\tau = \eta \cdot std(D_{env}^2[n])$ // where $\eta = 2$

Find zero-crossing locations

Approximated AO peaks: $a[v] \leftarrow$ Positive zero crossings of $\widehat{T}_{env}[n]$

// where $\widehat{T}_{env}$ corresponds Hilbert transform of T_{env}; and $v = 1, 2, ..., V$

Estimate True AO peaks using cardiac cycle envelope (CCE)

Construct CCE using absolute operator and triple integration on GDFMs

True AO peaks: $\widetilde{a}[k] \Longleftarrow a[v]$ elements lying within the range $\mathbb{Q}$

// where $\mathbb{Q} \in (\text{CCE\_peak}_k, \text{CCE\_peak}_k + 350ms)$; $k = 1, 2, ..., K$

Output: Detected AO peak locations $\widetilde{a}[k]$

(ii) **False detection on systole profile:** An AO peak detected in the systole profile is considered as false positive, if the condition $(d_k - M) < -0.7 \times M$ is satisfied. This false positive can be corrected by discarding p_{k+1} peak.

(iii) **False detection on diastole profile:** After taking care of the first two conditions, the ***pks*** and $\boldsymbol{d_k}$ vectors are updated. Subsequently, the method looks for the false positive in the diastole profile by checking the conditions, $(d_k - M) < -TH$ and $(d_{k+1} - M) < -TH$. If both the conditions are satisfied, p_{k+1} peak is discarded to correct this false positive case. In this way, an AO peak correction strategy is adopted in our proposed method.

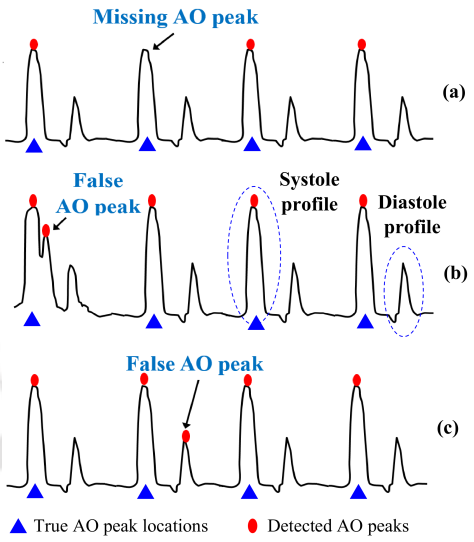


Figure 2.4: SCG envelopes showing false alarming conditions in AO peak detection. (a) Case 1: missing AO peak, (b) case 2: false detection on systole profile, and (c) case 3: false detection on diastole profile

2.2.3 pAC peak detection

The pAC peaks are detected using a technique given in [103]. At first, it divides each of the consecutive AO-AO intervals at their midpoint (M_1), which is expressed as:

$$M_1 = \frac{p_k + p_{k+1}}{2} \tag{2.4}$$

where, p_k and p_{k+1} are consecutive AO peaks from the updated $\boldsymbol{pks}$ vector. In a similar manner, the interval between p_k and M_1 is again divided to obtain the pAC as:

$$M_2 = \frac{p_k + M_1}{2} = \frac{3p_k + p_{k+1}}{4} \tag{2.5}$$

Now, the SCG signal is segmented from M_2 to M_1 for each of the cycles. Subsequently, the DC-offset is removed and amplitude normalization is performed. Subsequently, moving average filtering is applied to reduce irregularities. To detect all pAC points, the trends in the segments are removed by applying a Butterworth high pass IIR filter with an empirical cut-off frequency of 10 Hz. Finally, the pAC peaks are detected as maximas of the processed SCG segments.

2.3 Cuffless Continuous Blood Pressure Measurement

High blood pressure is a major risk factor that leads to serious cardiovascular diseases (CVDs), such as hypertensive heart disease, stroke, and coronary artery disease [292]. Effective and continuous

monitoring of BP gives important insights of cardiovascular regulation patterns. It provides clinically relevant information useful for early detection of hypertension and helps in the clinical maintenance of the CVDs. To monitor the continuous BP in real time in an intensive care unit (ICU), the most common method is to insert a cannula needle in a suitable artery. For arterial blood pressure measurement, it is a gold standard method. Because of its invasive nature, it may introduce some other clinical complications, such as bleeding, infection, and ischemia. Also, it requires a standard clinical setting and an expert. Among other modalities, mercury sphygmomanometer is considered as the most accurate non-invasive device. However, it requires an inflatable cuff that may cause discomfort and pain to the user. Moreover, it only provides the instantaneous blood pressure, and thus, cannot be used for long term ambulatory BP monitoring. Arterial tonometry and arterial volume clamp are other primary techniques to obtain continuous BP in non-invasive way, but they need occlusive cuff [2]. All these devices have limitations for continuous and long term BP measurement. Therefore, in such a perspective, techniques that can provide continuous BP monitoring without occluding cuff would be advantageous for day-to-day cardiac health assessment.

In recent years, research on cuffless BP measurement methods for continuous BP monitoring has gained much attention. In literature, it is found that the pulse transit time (PTT) is employed to estimate blood pressure. The PTT is defined as the time taken for a cardiac pulse in artery to travel from a proximal point to a distal point. Usually, for the calculation of PTT interval, R-peak of electrocardiogram (ECG) and different characteristic points of photoplethysmogram (PPG), such as sytolic peak, onset, maximum slope, and mid-point, are used [2–9]. However, the existing literature has reported the maximum slope of PPG as a better distal point for accurate BP estimation. Recently, some of the BP estimation methods have introduced the use of concurrent SCG and PPG signals for the PTT estimation [10–13]. For this purpose, they suggested AO peak of the SCG as a proximal point, and systolic peak or onset of the PPG as a distal point. All of these approaches require two correlated cardiac pulse signals, either ECG-PPG or SCG-PPG. Acquiring more than one signal and estimating the location specific feature points is computationally complex. It also reduces the flexibility of the system due to the involvement of different sensing modalities. This type of complex system involving two or more signals to measure the BP parameters may also be expensive. However, a few attempts have also been made for BP estimation with a single sensor, such as Fiber Bragg Gratings (FBGs) [293]. But, due to its bulky physical configuration, the method is not yet matured and reliable

for clinical use. Non-contact SCG signals acquired from Microwave Doppler radar or Laser Doppler Vibrometry (LDV) were also used along with ECG and PPG to measure the PTT [22]. However, the acquired signal gets affected by radar clutter and other environmental noises, leading to a poor signal-to-noise ratio. Also, a single-channel impedance plethysmography (IPG) based method was proposed to devise a BP estimation model [294]. It uses Ag-AgCl based multiple wet electrodes integrated in a single patch. Although it involves a single IPG sensor, there are challenging factors of choosing the excitation source frequency and arranging the sensing electrodes. Additionally, the user comfortability is another issue with this method for the long-term measurement. Another method used a wearable multi-wavelength PPG (MWPPG) module that consists of light sources composed of a blue, green, red and IR LEDs [295]. It acquires the MWPPG signals to measure the characteristic features required for the BP estimation model. The model parameters are measured from two principle components that are generated using a PCA-based method. However, the system complexity increases due to the involvement of multiple sensors. Another research work was presented by [296], which introduces an epidermal electronics-based multiple polyaniline sensor array. The performance of the method relies on identifying the exact location of the radial and ulnar arteries. Additionally, the SCG components of two different orthogonal axes have also been investigated [297], where the components are acquired from a single sensor. The method relies on time-interval from systole-profile onset in x-axis to diastole-profile onset in z-axis of SCG for systolic blood pressure (SBP) estimation. However, determination of these onset points is a crucial task. Additionally, achieving a good accuracy is still a problem. The research on BP estimation using SCG signals is still far from maturity and has not yet advanced to a stage where it can be successfully deployed as a medical user interface in consumer-grade applications. The proposed method is the first ever one of its kind where BP is estimated using a single SCG component. Looking at the above facts, the main aim of this proposed work is to estimate blood pressure using a single SCG signal component acquired through an accelerometric device. The proposed method does not rely on SCG components of x- and y-axis. Only the z-axis component is used for processing. The detected AO and pAC points of the SCG signal are used to determine the LVET and HR. Successively, these features are used to compute the BP metrics using a calibration procedure.

2.3.1 Proposed Method for SBP and DBP estimation

During the systolic phase of the cardiac cycle, blood is pumped out from the heart. It exerts force on the wall of arteries in systematic circulation and it is measured as blood pressure (BP). In a normal

subject, the systolic and diastolic pressure is clinically written as:

$$\frac{\text{Systolic blood pressure}}{\text{Diastolic blood pressure}} = \frac{SBP}{DBP} = \frac{120\,\text{mmHg}}{80\,\text{mmHg}} \tag{2.6}$$

The computation of SBP and DBP is based on the principle of pulse propagation through major arteries. The systolic and diastolic blood pressure in a major artery are directly proportional to the total cardiac output (Q) and total peripheral resistance (Ω). The Ω includes resistances of the arteries, arterioles, capillaries, venules, and veins in systematic circulation. The BP is given as:

$$BP = Q \times \Omega \tag{2.7}$$

The cardiac output, Q, mainly depends on heart rate (HR) and stroke volume (SV). Therefore, it is expressed as:

$$Q = HR \times SV \tag{2.8}$$

The SV depends on end systolic volume (SV_l) and end diastolic volume (SV_m) of a ventricle with a relation,

$$SV = SV_l - SV_m \tag{2.9}$$

The LVET gives the systolic time intervals that corresponds the interval between ejection and termination of the aortic flow. Thus, the LVET (systolic time interval) is directly influenced by the SV. The LVET is expressed as [22]:

$$LVET = AC - AO \tag{2.10}$$

where, AO and AC represents aortic valve opening and closing instants, respectively. At the AO instant, the blood is ejected out of the heart, and therefore, can be considered as a proximal point for a pressure pulse. At the AC instant, the blood reaches the measurement sites and returns back from the whole body to the heart, and therefore can be considered as a termination point of the pulse. This justifies the use of LVET in place of PTT to calculate BP. Also, the LVET is inversely related to HR [22, 298]. It is difficult to identify AC due to large morphological variability of SCGs among the subjects, and the SCG is susceptible to the distortions caused due to body movements, respiration and other environmental noises [22, 103]. The proposed model uses the timing-information of pAC peak for the computation of LVET interval. With this modification, an approximation of LVET, say

LVET′, is used for modelling the BP in the proposed method. The LVET′ is computed as:

$$LVET' = pAC - AO \tag{2.11}$$

The proposed system design for the BP estimation can be described mathematically. Suppose $\mathbf{X} = \{x_1, x_2, \cdots, x_N\}$, $\forall x_i \subset \mathbb{R}^2$ be the input data sequence of the SCG signal. It can also be rewritten as $\mathbf{X} = [A, T]$, where $A = \{a_1, a_2, \cdots, a_N\}$ and $T = \{t_1, t_2, \cdots, t_N\}$ correspond sample amplitudes and time stamps, respectively. The AO peak vector will be a subset of $\mathbf{X}$ as $\{x_{i_1}, x_{i_2}, \cdots, x_{i_M}\}$, where $M << N$. Similarly, pAC peak vector is represented as $\{x_{k_1}, x_{k_2}, \cdots, x_{k_M}\}$. Subsequently, cardiac cycle difference, HR, and LVET′ are computed as: $CC_j = t_{i_j} - t_{i_{j-1}}$, $HR_j = 60/CC_j$, $LVET'_j = t_{k_j} - t_{i_j}$, where i_j, k_j denote indexes for AO and pAC peaks, respectively. Then, one can define a set function for BP as $BP : \mathbf{X} \rightarrow \mathbb{R}$ that maps a set of sample to a vector. By utilising the logarithmic LVET′ and HR information, the BP is modelled as:

$$BP(x_1, x_2, \cdots, x_N) = a \cdot ln(LVET') + b \cdot HR + c \tag{2.12}$$

where, a, b, and c are used to parametrize the proposed model. These parameters are subject-specific and are estimated via modelling in a least square sense. The HR is calculated from the time interval between consecutive AO peaks. The proposed model can be represented in a vector-matrix form as:

$$\begin{bmatrix} BP_1 \\ BP_2 \\ . \\ . \\ . \\ BP_{K^*} \end{bmatrix} = \begin{bmatrix} ln(LVET'_1) & HR_1 & 1 \\ ln(LVET'_2) & HR_2 & 1 \\ . & . & . \\ . & . & . \\ . & . & . \\ ln(LVET'_{K^*}) & HR_{K^*} & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} \tag{2.13}$$

where, K^* $(= K - 1)$ is the total number of beats, the BP vector with dimension $K^* \times 1$, represents the original arterial blood pressure and it is used for training the model for both SBP and the DBP estimation. The $LVET'$ and HR are the parameters that are estimated simultaneously for each beat along with the original BP and represented by a $K^* \times 3$ matrix. The unknown coefficients a, b and c are represented by a 3×1 vector. The unknown vector is estimated using linear regression by applying pseudo inverse as:

$$Z = (\mathbf{D}' \cdot \mathbf{D})^{-1} \cdot \mathbf{D}' \cdot BP \tag{2.14}$$

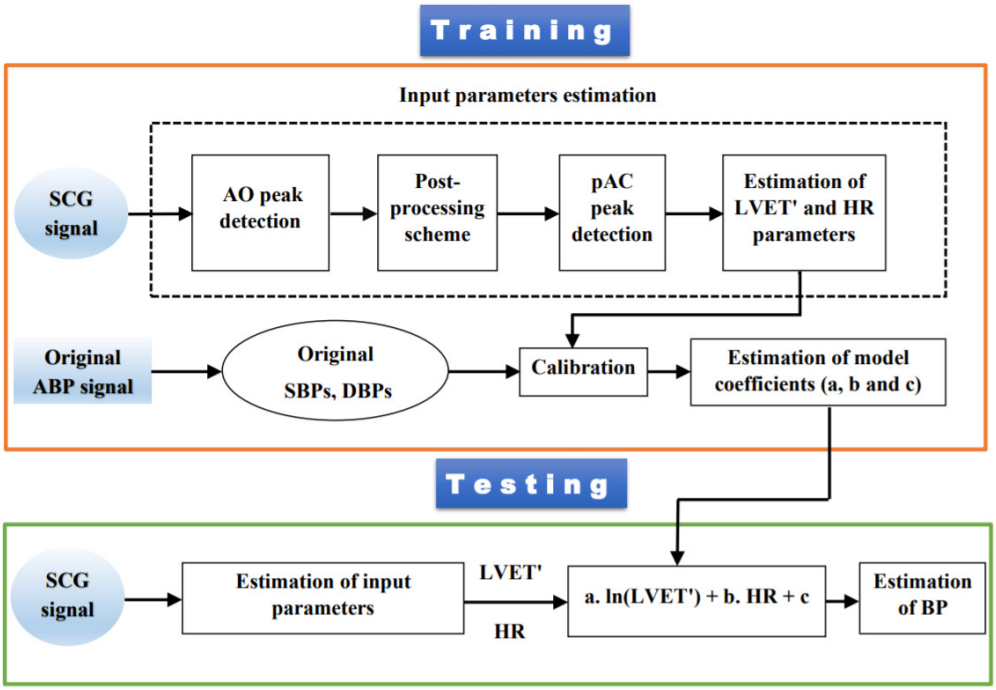


Figure 2.5: Block diagram of the proposed BP estimation method.

where, Z is the unknown coefficient vector and $\mathbf{D}$ is the matrix that carries the information related to the estimated $LVET'$ and HR values. The coefficients are estimated for each individual for both SBP and DBP, separately. Figure 2.5 shows a block diagram of the proposed cuffless BP estimation algorithm. It has training and testing phases. The training phase is required only during initialization where the coefficients a, b and c are computed using the regression method. Then, the testing phase becomes independent of BP parameters for a personalized measurement.

2.3.2 Performance Evaluation

The estimated SBP and DBP values are compared with the reference ABP measurements. The performance of the proposed approach is evaluated in terms of Mean error (ME), mean absolute error (MAE) and standard deviation (STD) measures, and the results are shown in Table 2.1. For the DBP measurement, the minimum ME and MAE values are -0.074 and 1.877, respectively for subject S005 with a STD value of 2.107. Similarly, for the SBP measurement, minimum ME is closed to zero for subject S004 with a STD value 2.649. The minimum MAE value is 1.036 for subject S005 with a STD value 1.405. For all the subjects, the average ME with STD of the estimated DBP and SBP are -1.299 ± 2.578 mmHg and -0.197 ± 3.332 mmHg, respectively, whereas the average MAE are, 2.598

Table 2.1: Comparison of estimated BPs using our proposed LVET′-based method with reference ABP measurements.

Subject	DBP (mmHg)			SBP (mmHg)		
	ME	MAE	STD	ME	MAE	STD
S001	-1.662	1.981	2.153	1.087	2.689	3.284
S002	-2.932	2.932	1.959	-2.807	3.074	2.248
S003	-1.503	2.985	3.335	0.567	3.031	3.857
S004	1.783	2.545	2.747	-0.004	2.142	2.649
S005	-0.074	1.877	2.107	0.437	1.036	1.405
S006	-0.737	1.943	2.044	-0.692	2.486	3.045
S007	-1.216	2.315	2.598	1.954	4.732	5.417
S008	-2.035	2.661	2.190	-4.991	4.991	2.907
S009	-2.182	3.516	3.728	-0.395	1.724	2.113
S010	-2.432	3.221	2.916	2.878	5.859	6.390
Average	-1.299	2.598	2.578	-0.197	3.176	3.332

and 3.176 mmHg, respectively. The obtained results suggest that the errors in BP estimations from the proposed method satisfy the requirements of the IEEE standard 5 ± 8 mmHg [3, 299].

The estimated BP values are also compared with the BP values computed through earlier PTT- [2–9] and PTT1- [10–13] based approaches. In PTT-based approach, the duration of PTT is calculated using the combination of 'ECG-PPG', whereas in the PTT1-based approach, 'SCG-PPG' signal pair is used. Table 2.2 shows the comparison of the proposed method with the existing methods. The average ME and MAE between the proposed LVET′-based and PTT-based methods are found as 0.109 ± 0.669 mmHg and 0.617 mmHg for DBP, and 0.052 ± 0.615 mmHg and 0.547 mmHg for SBP measurement, respectively. While, for the PTT1-based method, these metrics are observed as 0.011 ± 0.777 mmHg and 0.384 mmHg for DBP measurement. While for SBP, they are found as -0.012 ± 0.326 mmHg and 0.183 mmHg, respectively. Additionally, the average correlation of the BPs estimated from the proposed method with that of the earlier methods is also computed. For the PTT-based method, the average correlation coefficients, denoted by $r1$, are obtained as 0.510 and 0.419 for the DBP- and SBP-estimation, respectively. Similarly, for PTT1-based method, the correlation coefficient ($r2$) values are observed as 0.596 and 0.740 for the DBP and the SBP, respectively. The good correlation exhibited by the proposed method clearly shows that it is comparable with the existing methods. Also, for subject S003, the traces of beat-to-beat DBP and SBP values estimated by the proposed and the existing methods are shown in Figure 2.6. It is observed that the proposed method and the existing state-

Table 2.2: Statistical comparison of the proposed LVET′-based method with the PTT [2–9] and PTT1 [10–13] based methods for DBP and SBP measurements

Subject	DBP (mmHg)							
	ME_{PTT}	ME_{PTT1}	MAE_{PTT}	MAE_{PTT1}	STD_{PTT}	STD_{PTT1}	r1	r2
S001	-0.289	-0.016	0.406	0.096	0.397	0.115	0.134	0.778
S002	0.005	0.017	0.048	0.107	0.063	0.153	0.969	0.839
S003	-0.090	-0.061	0.290	0.305	0.383	0.461	0.863	0.836
S004	-0.439	-0.035	1.311	0.323	1.568	0.839	0.163	0.434
S005	1.372	0.081	1.569	0.361	1.084	0.507	0.136	0.305
S006	0.066	-0.002	0.072	0.019	0.059	0.026	0.921	0.985
S007	0.027	-0.006	0.102	0.108	0.119	0.134	0.645	0.646
S008	0.005	-0.004	0.088	0.123	0.145	0.211	0.433	0.238
S009	0.106	0.104	0.697	0.336	0.870	0.527	0.086	0.518
S010	0.326	0.042	1.587	2.065	2.005	4.802	0.749	0.381
Average	0.109	0.012	0.617	0.384	0.669	0.778	0.510	0.596
Subject	**SBP (mmHg)**							
	ME_{PTT}	ME_{PTT1}	MAE_{PTT}	MAE_{PTT1}	STD_{PTT}	STD_{PTT1}	r1	r2
S001	-0.619	-0.037	0.888	0.301	0.886	0.379	-0.048	0.750
S002	0.013	0.013	0.375	0.236	0.473	0.378	0.343	0.586
S003	-0.090	-0.028	0.340	0.484	0.576	0.779	0.886	0.822
S004	-0.269	-0.025	0.825	0.304	0.987	0.752	0.573	0.596
S005	0.336	0.016	0.383	0.044	0.265	0.068	0.403	0.795
S006	1.339	0.000	1.445	0.085	1.108	0.099	0.182	0.924
S007	0.036	0.033	0.081	0.083	0.096	0.096	0.685	0.683
S008	0.034	-0.006	0.082	0.051	0.166	0.067	0.135	0.260
S009	-0.079	-0.003	0.561	0.010	0.698	0.017	0.055	0.993
S010	-0.182	-0.084	0.489	0.232	0.897	0.621	0.972	0.989
Average	0.052	-0.012	0.547	0.183	0.615	0.326	0.419	0.740

of-the-art methods produce similar DBP and SBP measurements. Therefore, a single accelerometric sensor based proposed method may replace the use of multi-modal system for the continuous BP estimation.

2.3.3 Discussion

The key factor of the proposed model is to estimate BP measurements with a single cardio-accelerometric sensor. To show the performance and its comparison with the existing methods, the BP values estimated using the testing data of all the subjects are taken for Bland-Altman and regression plots. Figure 2.7 shows the Bland-Altman plots in terms of measurement-difference in BP estimation. The

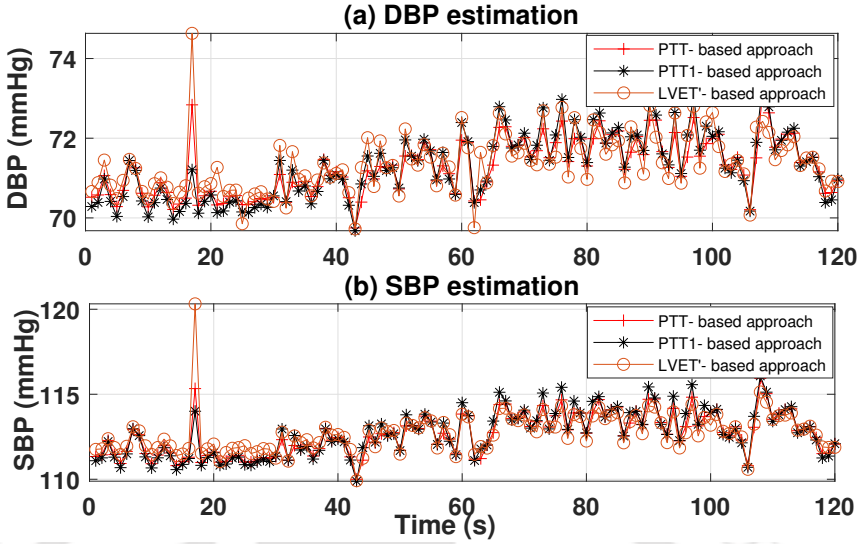


Figure 2.6: Comparison of the proposed LVET′ based method with PTT and PTT1 based methods for estimated DBP and SBP values.

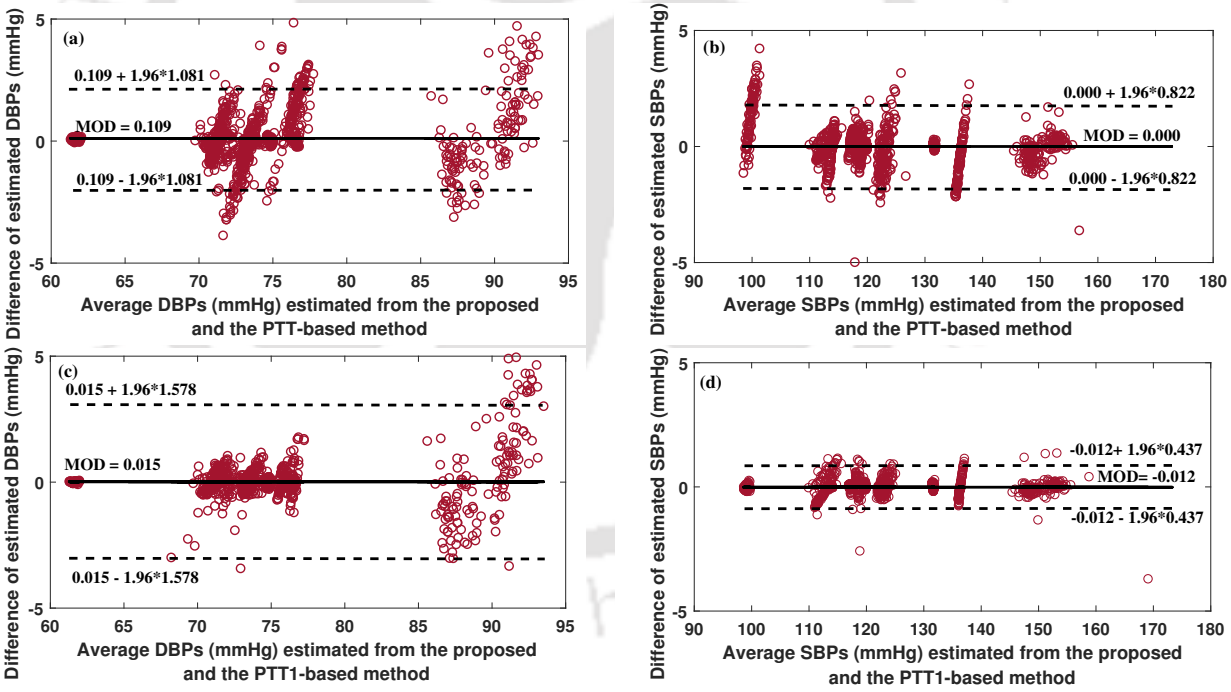


Figure 2.7: Bland-Altman plot to compare the performance of the proposed method with PTT- and PTT1-based methods. Panels (a) and (b) represent comparison with PTT-based method for DBPs and SBPs, respectively, and panels (c) and (d) represent comparison with PTT1-based method for DBPs and SBPs, respectively.

x-axis represents the average of the BP values obtained by the proposed and the existing method, whereas the difference between both the BP values is shown in the y-axis. It is desirable to have a low value of bias or mean difference. This result indicates that the proposed method can give similar measurement values as that of the method used for comparison of performance. As shown by the bold

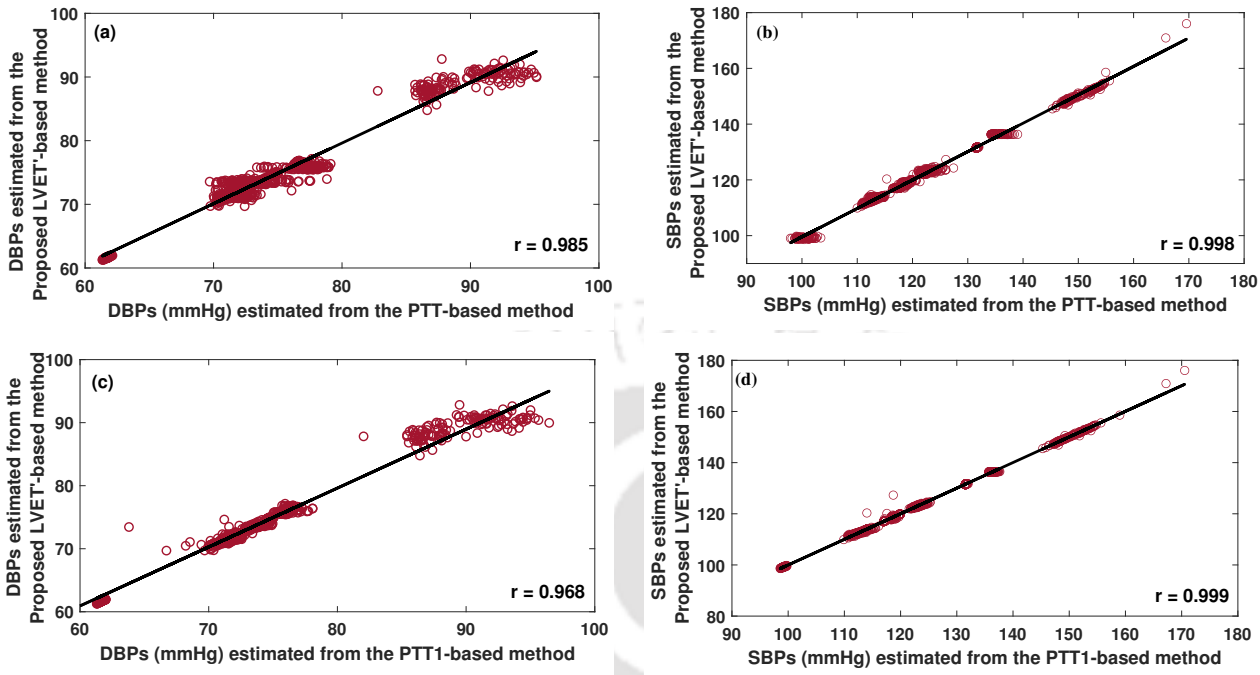


Figure 2.8: Regression plots of estimated BP measurements for a single subject. (a) and (c) show the regression of DBPs estimated by the proposed method with that of the PTT- and PTT1-based methods, respectively, (b) and (d) show the regression of SBPs estimated by the proposed method with that of the PTT- and PTT1-based methods, respectively.

line in Figure 2.7, the mean differences of the estimated DBPs and SBPs between the proposed and the PTT-based methods are 0.109 ± 1.081 mmHg and -0.000 ± 0.822 mmHg, respectively. Whereas, they are found to be -0.015 ± 1.578 mmHg and -0.012 ± 0.437 mmHg, respectively for the PTT1-based method. It can be observed that the majority of the points lie within the limits of agreement (LOA), *i.e.*, $m \pm 1.96s$, where m and s denote the mean and standard deviation of the BP difference, respectively. The LOA are shown by dashed lines in the figure. This shows the suitability of our method over the existing methods. The performance comparison is also shown with the help of regression plots in Figure 2.8. The corresponding correlation coefficients between the proposed LVET′-based and the PTT-based methods are approximately 0.985 and 0.998 for the DBP and SBP estimation, respectively, as shown in Figure 2.8(a) and (b). In case of the PTT1-based method, they are found to be 0.968 and 0.999 for the DBPs and SBPs as shown in Figure 2.8(c) and (d).

Although the current research has demonstrated the feasibility of an accelerometric device for continuous BP measurement, the proposed work still needs further improvement for its practical applicability. The limitation of the proposed work is that the method has been tested and validated

with limited number of subjects. Additionally, the recordings were done in a relatively stable condition, where large body movements were not allowed. For more generalization of the method, its performance should be evaluated for different physiological and pathological conditions of individuals with varying demographics and under different dynamic settings. The proposed work could be further analyzed and validated for BP measurements with respect to different body postures, like supine, left and right lateral, sitting and standing. Also, an investigation needs to be performed on a longer duration data to ensure the applicability of our method for long-term BP monitoring.

2.3.4 Conclusion

In this work, a method is proposed for noninvasive and cuffless continuous BP measurement using a single accelerometric based sensor. Initially, the AO and pAC peaks are detected for the estimation of LVET′ and HR parameters. With these parameters, the BP modeling is performed for the SBPs and DBPs using individual calibration. The existing methods require a PPG signal along with either ECG or SCG signal to perform this task. However, the use of bio-electrodes for the ECG and the photo-detector for the PPG acquisition has been avoided in the proposed method by using only an accelerometer for SCG data. This makes the system more flexible, and thus, offers mobility to the user. Also, the lightweight of the sensing device and its simple operation make it versatile to be operated independently without any expert help. With in-house recordings, quite satisfactory correlation results between the proposed method and the existing methods are obtained. The performance results clearly show that our method is able to achieve the performance up to the level obtained by the state-of-the-art methods only with a single cardiac sensor. Hence, the proposed method shows the potential to replace the existing methods for continuous BP monitoring. However, in the current pilot study, the proposed method was tested and validated only on a limited number of healthy individuals, who were instructed to be in a supine position without significant body movements during signal acquisition. Thus, as a future research direction, the performance could be evaluated for different physiological modulations and pathological conditions of individuals with varying demographics. Also, it is required to validate the efficacy of the proposed method on long duration signals for long-term BP monitoring. We believe that the present work would surely open opportunities to bring and build new ideas and research efforts in this field.

2.4 Heart Rate Variability (HRV) Measurement

HRV analysis is usually employed to assess the autonomic nervous system (ANS) functioning in cardiac studies. The HRV is a physiological phenomenon that describes the oscillation in consecutive heart cycles [99]. Both sympathetic and parasympathetic activities can be apparently characterized by this tool. Clinically, HRV has proven its abilities in major applications, such as diagnosis of sudden cardiac death due to acute myocardial infarction and early prediction of diabetic neuropathy [99,300]. Besides these, HRV has been investigated for several cardiac diseases, physical exercise, renal failure, stress conditions, sleep disorder, age, gender, drugs, alcohol consumption, and smoking [300, 301]. The HRV analysis gives many temporal, spectral, and nonlinear indices solely relying on beat-to-beat interval traces.

The prevalence of HRV estimation has led to increased development of several commercial and non-business software tools [300, 302]. SCG is immensely used for HRV analysis in recent years due to technological advancements and miniaturization of accelerometers. Usually, AO point creates a prominent peak in an SCG cycle, and hence, it has been studied more for heartbeat extraction and HRV applications [99,105,303]. In this study, we proposed an SCG-based algorithm for HRV analysis, which entirely depends on the AO peaks of an SCG signal. The method performs time-domain and frequency-domain HRV analysis, and the results derived from SCG are cross-validated with ECG as a reference. The performance results are validated using absolute error and normalized linear cross-correlation measures.

2.4.1 Cardiac interval estimation in SCG signal

The detected AO peaks are used to derive the tachogram from AO-AO intervals in an SCG signal. The AO-AO interval time series is investigated for HRV analysis that is obtained as the traces of time-difference between consecutive AO peaks. The AO-AO tachogram (AA) is created as:

$$AA_t = AO_{t+1} - AO_t \tag{2.15}$$

where, AO_t $(t = 1, 2, ..., T)$ represents timing information of the t^{th} AO peak. Similarly, RR intervals are also measured from the time instants of successive R-peaks of concurrent ECG for validation purposes. The well known Pan-Tompkins algorithm [304] is employed for this.

2.4.2 Measurement of HRV indices

Both time-domain and frequency-domain parameters are estimated for the HRV analysis [300]. In this work, the studied HRV parameters are average cardiac interval (meanNN), average heart rate (meanHR) and their corresponding standard deviations SDNN and SDHR, root mean square of successive differences (RMSSD), percentage of successive intervals differing more than 50 ms (NN50), triangular index (HRV TI), and power ratio of low to high frequency-band (LF/HF). All the parameters are individually determined using AO-AO and RR interval time series.

2.4.3 Materials and methods

In this study, a publically available CEBS database (at PhysioNet archive) [305] and a private database are used for performance analysis. The CEBS database consists of twenty multi-channel recordings acquired from twenty healthy subjects. Each recording possesses four channels: ECG (I and II), respiration signal, and SCG. Recordings were taken in still awake condition and supine position of subjects in three subsequent states: before listening to music at basal state for 5 min., music listening state for 50 min., and the music stopped state for 5 min. However, the authors used all the twenty basal stated recordings (ECG-II and SCG-z) for the study.

The private database contains fifteen multi-channel signals recorded from three healthy male subjects in five different sessions. These sessions involve various physiological modulations and postures: (i) supine position with normal breathing for 6 min., (ii) supine position with hold or stopped breathing for 40 s, (iii) sitting for 2 min., (iv) standing for 2 min., and (v) exercise recovery. The exercise recovery includes rope-skipping (1 min.) and plank exercise (30 s) followed by a recovery period of 20 s. The measurement protocols for these sessions (abbreviated as NB, SB, SI, ST, and ER) were suitably designed. All the signals are resampled at 1 kHz for the experimentation.

2.4.4 Results and discussion

In the proposed method, the AO peaks of an SCG are detected initially along with the identification of R-peaks in the concurrent ECG signal. An experimental result for the detection of these peaks is shown in Figure 2.9(a) and (b). The derived AO-AO and RR intervals are used to create traces of cardiac cycles. Before going deep further into HRV analysis using AO-AO interval time-series, it is interesting to perform the reliability-test with RR intervals. For instance, it is observed that the derived AO-AO intervals and RR intervals plotted together in Figure 2.9(c) show similar traces.

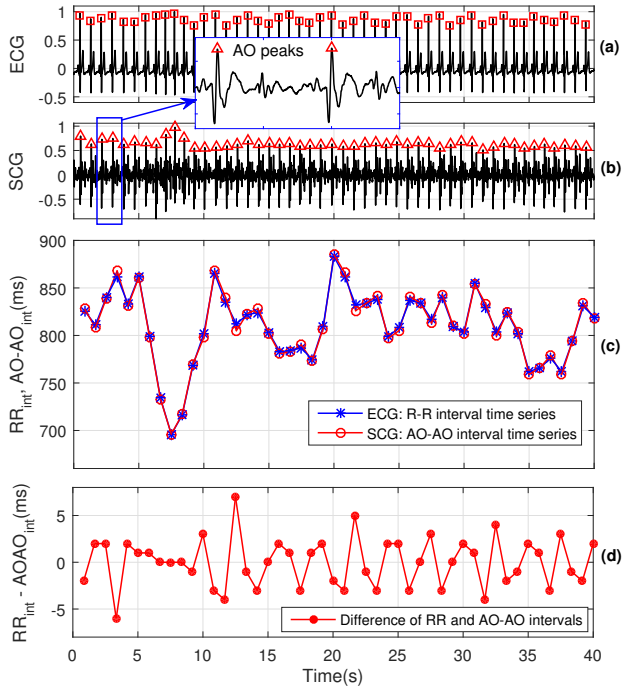


Figure 2.9: Experimental results of cardiac cycle extraction. (a) Detected R-peaks in ECG, (b) detected AO peaks in SCG, (c) extracted heartbeats from RR and AO-AO intervals, and (d) difference of RR and AO-AO intervals.

Thus, from the qualitative analysis, it can be said that insignificant differences are observed between these intervals, as depicted in Figure 2.9(d); however, a quantitative analysis will put more value on this. Performance results for the validation of heartbeat extraction and HRV analysis are discussed in the upcoming subsections for both the databases. For further experimentation, we used *HRVTool*, an open source software toolbox, available in MATLAB (MathWorks Inc.) to measure the HRV parameters [302].

2.4.4.1 Validation of Heartbeat Extraction

An experiment is performed to show the possible use of our method in the HRV analysis application. For this, the cardiac cycles extracted from the AO peaks of an SCG and RR intervals of concurrent ECG signal are compared, where each of the 20 signals are taken for 30 s duration. The comparison is shown via regression and Bland-Altman plots in Figure 2.10. It shows Pearson's R-squared correlation coefficient equal to 0.98, which reveals a strong correlation. Root mean squared error (RMSE) is achieved as 17 ms for RR and AO-AO intervals. Also, a good agreement can be observed in the Bland-Altman plot, as shown in Figure 2.10(b). It shows around 34 ms limits of agreement in 95%

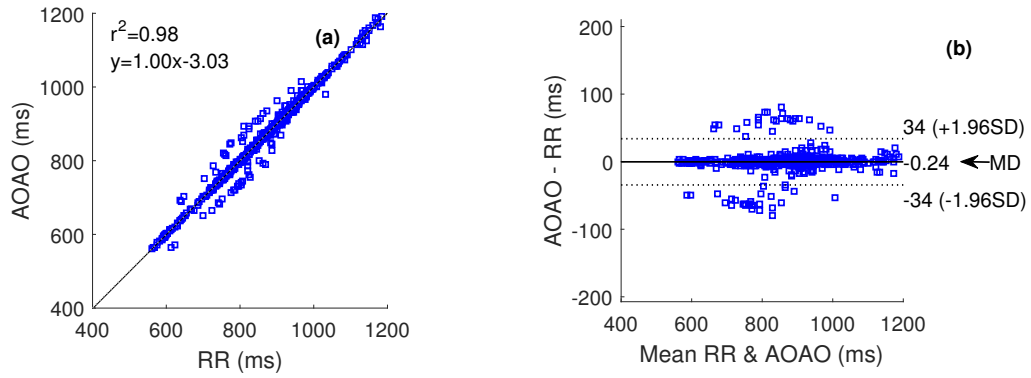


Figure 2.10: Comparison of heartbeats derived from RR and AO-AO intervals for CEBS data using (a) correlation plot, and (b) Bland-Altman plot. Limits of agreement (LOA) in 95% confidence intervals are shown via dotted lines. (Abbreviations used – r^2: Pearson's R-squared correlation coefficient, MD: mean difference, SD: standard deviation of difference).

confidence interval, and a minimum mean difference (MD) of -0.24 ms.

In a similar manner, correlation and Bland Altman plots are prepared for real-time scenarios of different postures and physiological modulations. Figure 2.11 shows the Bland-Altman plots along with regression plots for all the five sessions. The obtained parameters from these plots are also listed for individual cases in Table 2.3. It is observed that the Pearson's R-squared correlation coefficient is equal to 0.84, 0.99, 0.97, 0.94, and 0.97 for NB, SB, SI, ST, and ER cases, respectively. Whereas RMSE measures are found as approximately 27 ms, 13 ms, 15 ms, 26 ms, and 20 ms, respectively, for these cases. They reveal a quite strong correlation of AO-AO intervals with RR intervals. A very few percentage of data lying outside the limits of agreement for all the conditions represent good agreements with RR intervals, as shown by the LOA measure in the Bland-Altman plots. This experiment validates our AO peak detection method in heartbeat extraction applications for healthy volunteers at various physiological variations.

Table 2.3: Parameters obtained after the Bland-Altman analysis of RR and AO-AO intervals at different conditions.

Conditions	Regression Plot			Bland-Altman Plot	
	RMSE (ms)	r^2	y=ax+b	MD (ms)	LOA [U, L] (ms)
NB	27	0.84	$1.00x - 2.37$	0.02	-53, 54
SB	13	0.99	$1.01x - 6.23$	0.37	-26, 26
SI	15	0.97	$0.98x + 10.2$	-0.57	-31, 30
ST	26	0.94	$1.00x - 1.94$	-0.47	-52, 51
ER	20	0.97	$1.00x + 1.19$	0.05	-40, 40

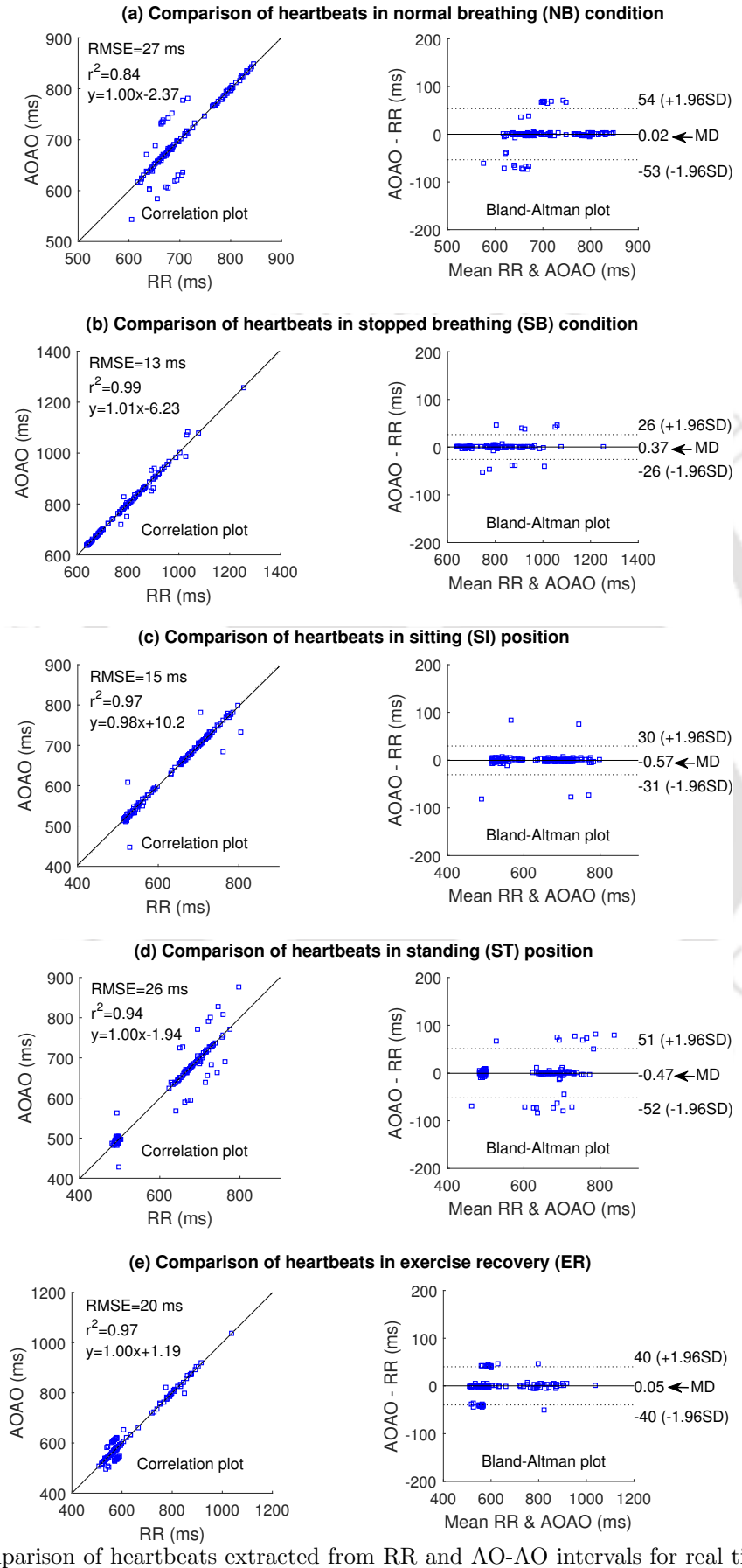


Figure 2.11: Comparison of heartbeats extracted from RR and AO-AO intervals for real time data recorded in different physiological conditions using correlation [left] and Bland-Altman plots [right]: (a) NB (b) SB (c) SI (d) ST and (e) ER.

2.4.4.2 Validation of HRV Parameter Estimation

After the validation of extracted cardiac cycles, the HRV estimation results obtained from AO-AO intervals are compared with the ones derived from RR intervals. The comparison of HRV indices is evaluated via absolute error and Pearson's correlation coefficient measures. Here, the absolute error is measured in Min-Max normalized HRV indices. Min-Max normalization, scales the data range from 0 to 1. It is defined as follows [306]:

$$H_i^n = \frac{H_i - min(H_i)}{max(H_i) - min(H_i)} \tag{2.16}$$

where, $H_i \ \ \forall i (i = 1, 2, \cdots, 8)$ represents a vector corresponding to i^{th} HRV parameter. Then, absolute error is computed for the normalized indices obtained from ECG and SCG signals as:

$$|e_i| = |H_i^n(ECG) - H_i^n(SCG)| \tag{2.17}$$

For the CEBS database, Figure 2.12(a) shows a boxplot representing the absolute errors for HRV indices derived from ECG and SCG signals. The evaluation results show median absolute errors of 5.8×10^{-4} for meanNN, 0.0177 for SDNN, 4.2×10^{-4} for meanHR, 0.0122 for SDHR, 0.0560 for RMSSD, 0.0911 for NN50, 0.1657 for HRV TI, and 0.0468 for LF/HF ratio. For all the HRV parameters, it is observed that the absolute error is very low, indicating the fact that the obtained parameters from the SCG signal using the proposed method are similar to the parameters derived from the reference signal.

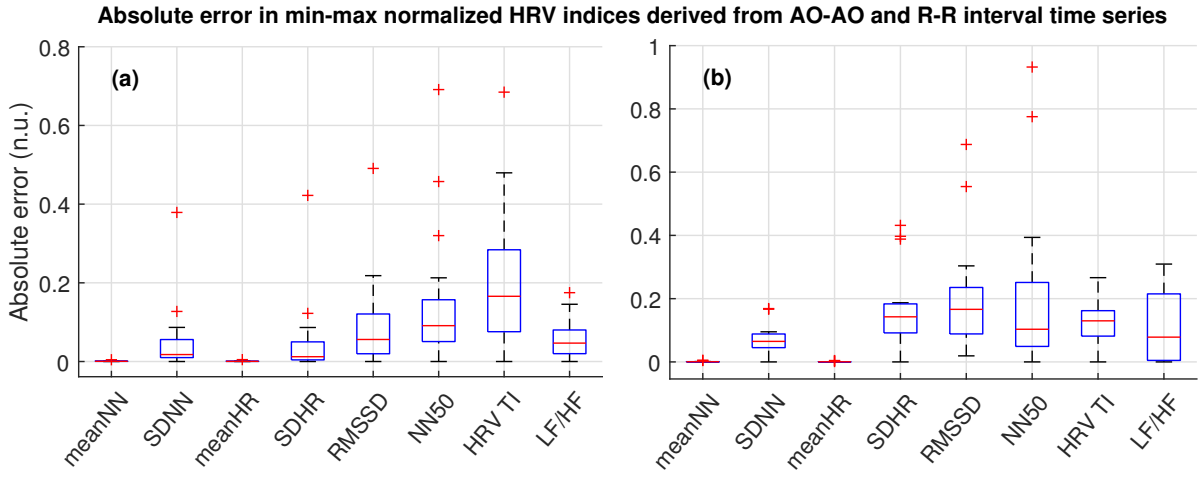


Figure 2.12: Boxplot representation of absolute errors between RR and AO-AO derived HRV parameters for (a) CEBS database and (b) real-time recorded data. (Note that all the HRV indices were processed in normalized units (n.u.)).

To validate the proposed method for HRV analysis, the comparison of HRV results is also evaluated for the real-time recorded data. In a similar way, absolute deviations of SCG derived HRV indices are evaluated with the ECG derived indices, as shown in Figure 2.12(b). For the estimated HRV parameters, the median absolute deviations are found to be around 4.8×10^{-4}, 0.0649, 3.5×10^{-4}, 0.1425, 0.1661, 0.1032, 0.1299, and 0.0783, respectively. Furthermore, we have measured the normalized cross-correlation (NCC) values between the HRV parameters estimated from AO-AO intervals and RR intervals of both CEBS and our real-time recorded data, as shown in Figure 2.13. The obtained insignificant absolute errors and good correlations for HRV indices signify that HRV parameters can be accurately achievable by the proposed method. To be more specific, the quantitative analysis of the experimental results indicates that four SCG derived indices, including meanNN, SDNN, meanHR and LF/HF, are the best surrogates of the ECG derived respective indices among all. Thus, the statistical comparison justifies our proposed method for HRV analysis.

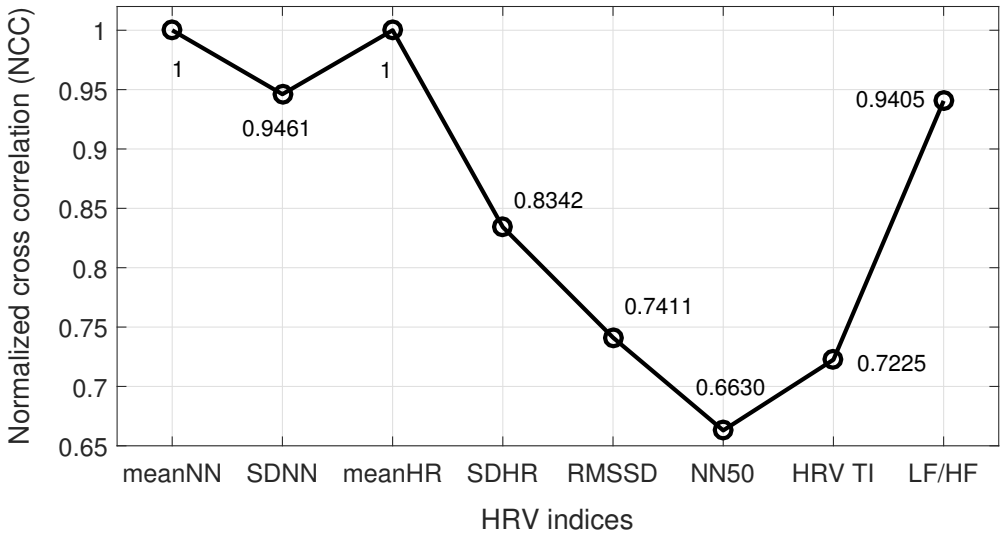


Figure 2.13: Correlation plot analysis for estimated HRV indices from ECG and SCG signals of both real-time and CEBS data.

During rigorous exercises and ambulatory conditions, the SCG signal quality is usually affected by motion artifacts caused due to body movements, friction noise, and respiratory efforts. In this work, HRV measurements from the SCG signal are analyzed under five different physiological conditions without considering large body movements. However, for more generalization, our HRV estimation method should be robust towards these artifacts. The introduction of multi-sensor data recording along with appropriate adaptive filtering techniques can make our method more robust toward body movements of a subject. Thereafter, it can be tested and validated on different physiological and

pathological conditions under dynamic set-up in future works. We believe that the current work would open opportunities to explore and implement new research ideas and efforts for the cardiovascular assessment using SCG.

2.4.5 Conclusion

In this work, we have proposed a reliable SCG-based HRV estimation method. For this purpose, our MVMD-based AO instant detection method integrated with the proposed post-processing scheme was investigated. All the experiments were performed on a publicly available CEBS database and our own registrations acquired at various physiological variations. The method has been validated for a robust heartbeat extraction by comparing the AO–AO intervals with RR intervals extracted from a reference ECG. It shows a high correlation and agreement between these intervals. The obtained HRV indices from the AO–AO intervals are also statistically compared with the parameters derived from the RR intervals in terms of absolute errors and NCC coeffcients. The performance results clearly show that our method is able to produce similar HRV results that are obtained from a standard ECG signal. However, a clinical validation is still needed for cardiac pathological conditions.

2.5 Identification of Ventricular Depolarization Events

Systolic time intervals (STIs) offer clinical assessments of risk factors associated with the cardiac muscle dysfunction [307]. Systolic intervals such as, pre-ejection period (PEP), left ventricular ejection time (LVET), and their ratio (PEP/LVET), also called Weissler index are useful for diagnosing cardiac systolic dysfunctions [308]. PEP is defined as the time interval between the start of left ventricular depolarization (onset of R-peak in ECG) and aortic valve opening (AO peak in the SCG), and LVET corresponds to the time interval between the opening and closure instants of the aortic valve, representing the time required for left ventricles to eject blood from the aorta. These STIs are considered as a marker for the left ventricular function that reflects changes in sympathetic activity and myocardial contractility, left ventricular end-diastolic volume, aortic diastolic pressure, and cardiac output [309]. Despite the individual application of PEP and LVET, the ratio between them is also regarded as a useful measurement of left ventricular dysfunction. The left ventricular performance is inversely proportional to the PEP/LVET ratio and the ratio value 0.423 is used to diagnose patients with heart failure caused due to systolic dysfunction [125]. Clinically, echocardiography is considered as a gold standard for the measurement of STIs [310, 311]. However, this method is not so convenient for long-

term continuous monitoring in personalized home healthcare due to having numerous reasons, such as use of bulky, and expensive sensor system and expert requirement for system setup and interpretation. In literature, few other sensing techniques are being proposed for STIs measurement, such as oxygen saturation, radial pulse pressure, impedance cardiography (ICG), simultaneous recording of ECG, phonocardiogram (PCG) and carotid arterial pulse [125]. However, these methods requires more than one cardiac sensing devices for the measurements. Involvement of multiple cardiac waveforms and estimating the precise time instants of the required fiducial points is computationally complex and may be expensive. Additionally, the flexibility and mobility of the user gets limited due to use of multiple sensing devices.

To address the limitation of the existing methods, our objective is to extract all the fiducial points required for measuring STIs from the SCG waveform. The time information of fiducial points of SCG is useful in the estimation of clinically significant systolic and diastolic time intervals. Among STIs, the fiducial points required to measure LVET, i.e. opening and closing instants of aortic valve are well captured by SCG. However, measurement of PEP relies on two essential timing information, (a) ventricular depolarization events, usually embedded in the ECG's R-peaks, and (b) SCG's AO peaks. Thus, estimation of PEP involves multiple cardiac sensing modalities. To avoid the use of multiple sensors, this study investigates to generate R-peaks instants on SCG waveform for PEP measurement using a single accelerometric sensor.

As an emerging signal, although SCG has been explored in various cardiac investigations since last decades, its utility to detect its generating cause would be quite interesting. The SCG waveform carries a most prominent AO peak corresponding to aortic valve opening, a cardiac action in response to the ventricular depolarization (VD). VD is a major cardiac event and its assessment provides diagnosis and prognosis of multiple life-threatening diseases, such as atrioventricular (AV) blocks and bundle branch blocks [312]. It is often characterized by R-peak of an ECG waveform. Thus, generation of AO peaks in SCG and R-peaks in ECG can be explicitly mapped as cause and effect. Due to this fact, it is extremely intriguing to investigate the association between R-peak events and SCG features. By estimating R-peaks on SCG, the SCG may put a step closer to become an integrated cardiac health solution exhibiting vital mechanical and electrical specifics of heart activities. This investigation is the first ever one of its kind where R-peak instants are estimated using cardio-mechanical events rendered by SCG fiducial points. The detected fiducial points of SCG signal are provided into the trained deep

feedforward neural network (DFN) to detect R-peaks. The estimated R-peaks along with the SCG feature points can be used to measure STIs for the assessment of ventricular dysfunction.

2.5.1 R-peak detection using SCG information

The detected AO and pAC peak instants are used to derive two more useful cardiac features, such as AO-AO tachogram and LVET′. Along with these four features estimated from SCG signal, static demographic features such as age and gender of each individual are also considered as input features to train the models for R-peak estimation. In this work, we employed a DFN to predict the R-peak instants. Here, the network is trained on R-peaks of reference concurrent ECG signals.

Deep Feedforward Neural Network (DFN):

The DFN is based on artifical neural network (ANN) architecture. It is composed of a node layers, containing one or more hidden layers between the input and output layers. With the advancement in hardware and technology, ANNs have progressed immensely from research to industrial applications. Also, ANN effectively works on unstructured data with comparable performance to human experts. Looking into the advantages of ANN, in this study, a five-layer DFN is designed and used. Here, the weights of the network kernels are initialized using normal initializer. Using an Adam optimizer, the network was trained for 500 epochs with batch size 32. The DFN architecture used in this study is illustrated in Table 2.4.

Table 2.4: The DFN architecture

Layer	Layer type	Activation function	Neurons count	Parameters
1	Input	ReLU	32	224
2	Hidden	ReLU	64	2112
3	Hidden	ReLU	64	4160
4	Hidden	ReLU	64	4160
5	Output	Linear	1	65

2.5.2 Results and Discussion

This study was conducted using self-recorded data, as described in Section 2.1. The detected AO and pAC feature points of an SCG were further used to derive AO-AO and LVET′ features. These four features along with the age and gender of all individuals, were used as input features for the proposed model to generate R-peak instances. Since the gender feature (male/female) is represented

as a string, it was one-hot encoded before model training. Male subjects were coded as 1, while female subjects were coded as 0. The final dataset can be represented in $(N \times M)$ matrix form, where M denotes the features and N represents the total number of cardiac cycles. While, the corresponding reference R-peak values, detected using the Pan–Tompkins algorithm [304] are placed in $N \times 1$ vector form. Finally, the dataset was split into 80% and 20% for training and testing purposes. The size of the input feature vector is $32 \times 1 \times 6$ (B × C × L), where B is the batch size, C is the number of channels, and L is the length of the feature vector. While, the dimension of the model's output is 32×1 (B × 1). The dataset contains a total of 6609 observations, corresponding to the number of cardiac cycles obtained from ten subjects. Figure 2.14 shows a SHapley Additive exPlanations (SHAP) plot representing the importance of features and their contributions in model for predicting VD events, where impact is shown for neural network model. Among all, AO seemed the most significant feature and was directly proportional to the model outcomes, i.e., R-peak instants. Whereas pAC was identified as the second most prominent feature and a similar relation was observed with the model output. The plot encouraged us to perform an ablation study on features for the prediction analysis. The utilized features are distinctive, and AO and pAC are highly correlated with the ground-truth R-peak values. Based on feature importance, an ablation study was carried out for the input features that was implemented using recursive feature elimination (RFE). In case-1, all features were used to build the models. Then, less relevant features were eliminated iteratively in successive cases, and the models were trained with the remaining features. Table 2.5 shows the experimental results of the ablation test performed on the proposed model. Also, the performance of the model is compared with the traditional random forest (RF) [313] and eXtreme gradient boosting (XGBoost) [314] models as presented in Table 2.5. The performance of all the three methods were evaluated and compared using statistical parameters, such as MAE, RMSE, and SDE. Additionally, a statistical ANOVA test was performed on each of the methods to reveal the significance of model

Table 2.5: Statistical comparison between the estimated and the ground-truth R-peak instants

Ablation Test	RF			XGBoost			DFN		
	MAE	RMSE	SDE	MAE	RMSE	SDE	MAE	RMSE	SDE
Case 1: AO, pAC, LVET, AO-AO, Gender, Age	0.123	0.200	0.200	0.171	0.238	0.238	0.080	0.167	0.164
Case 2: AO, pAC, LVET, AO-AO, Gender	0.123	0.201	0.200	0.171	0.238	0.238	0.077	0.166	0.165
Case 3: AO, pAC, LVET, AO-AO	0.123	0.200	0.200	0.171	0.238	0.238	0.069	0.167	0.162
Case 4: AO, pAC, LVET	0.123	0.200	0.200	0.174	0.248	0.248	0.073	0.175	0.168
Case 5: AO, pAC	0.120	0.198	0.197	0.174	0.248	0.248	0.077	0.184	0.174
Case 6: AO	0.107	0.184	0.184	0.167	0.241	0.241	0.061	0.162	0.160

Abstract-TH-3588_186102009

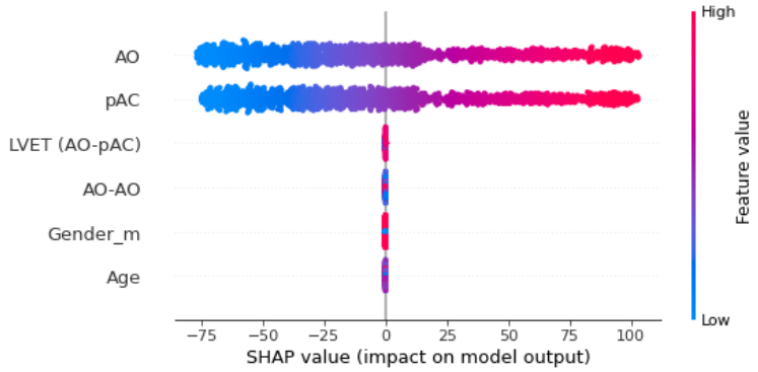


Figure 2.14: SHAP beeswarm plot exhibiting importance of features and their association with DFN-model output.

Table 2.6: Median absolute error values of ANOVA test for all the ablation cases

Case	RF	XGBoost	DFN
1	0.078	0.131	0.040
2	0.078	0.131	0.042
3	0.077	0.131	0.021
4	0.076	0.132	0.025
5	0.077	0.132	0.017
6	0.069	0.124	0.032

performances in all ablation cases. It is presented using boxplot of the absolute error between the ground-truth and the predicted R-peak values as shown in Fig 2.15. Also, the median absolute error values of the ANOVA test are presented in Table 2.6. From the tables and the graphical results, it was observed that all the methods were able to detect R-peaks using the provided SCG features with minimal error values. However, the proposed DFN model had achieved better results in comparison to the other two models. Also, from the ablation study, we found that the most significant feature i.e AO had produced the minimum error values for all the three methods. The significance of the AO feature is highlighted by the scatter plot of reference and estimated R-peak instants, as shown in Fig. 2.16(a). The plot clearly demonstrates that the estimated values align closely with their corresponding ground-truth values. Furthermore, the agreement between the R-peaks estimated from SCG and the ideal ECG signals is illustrated via Bland–Altman plot in Fig. 2.16(b). It shows around 0.3 limits of agreement (LOA) in 95% confidence interval, and a minimum mean difference (Mean) of -0.02. This experiment reveals good agreement of the estimated R-peaks with reference R-peaks. Thus, this shows the standalone ability of the AO peaks in the detection of R-peak instants. As a future direction, the

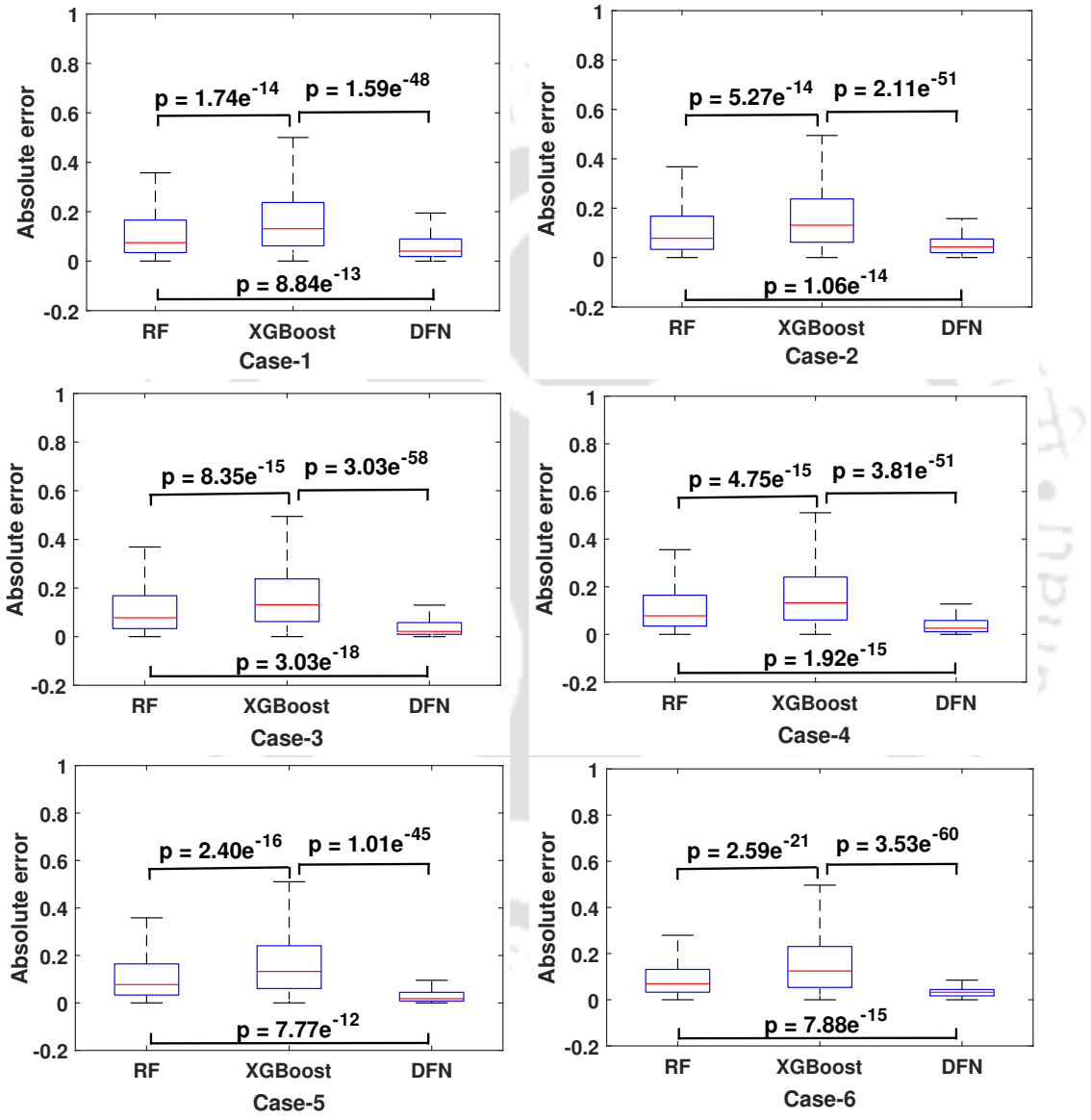


Figure 2.15: Boxplots representing performance of three ML models with an absolute prediction error metric for all six ablation cases ($p < 0.001$).

Abstract-TH-3588_186102009

detected R-peaks along with SCG's feature points can be utilized to measure PEP, LVET, and PEP/ LVET ratio for the assessment of ventricular dysfunction.

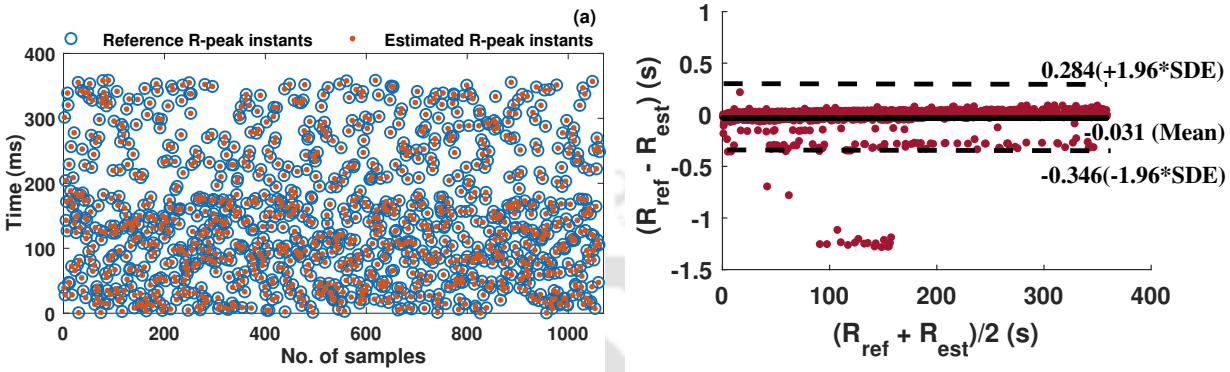


Figure 2.16: Experimental results for case 6. (a) Scatter plot of reference and estimated R-peak instants (b) Bland-Altman plot showing the agreement between reference and estimated R-peak timings with minimum mean error values.

2.5.3 Conclusion

In this study, R-peaks of ECG signal are detected on SCG waveform. The feature points of SCG along with the age and gender of each individual are considered as input features of the models to predict the R-peaks. For this, we utilized a deep feedforward neural network (DFN). The method is tested and validated using our in-house recorded data that are collected from ten healthy individuals in supine position. The estimated R-peaks from the model are statistically compared with the reference R-peaks of ECG signal in terms of MAE, RMSE and SDE. Also, the method is compared with traditional machine learning (ML) algorithms, such as random forest (RF) and extreme gradient boosting (XGBoost). The performance results shows the efficiency of each model in the detection of R-peaks. The evaluation metrics indicates that all the three models performed similarly with minimum error values. However, in comparison to RF and XGBoost, the proposed model achieved the best performance of approximately MAE = 0.080, RMSE = 0.167, SDE = 0.164. Thus, this study show the potential application of an accelerometric sensor to integrate both cardio-electrical and mechanical events for ventricular function assessment using systolic time intervals information.

2.6 Summary

In this chapter, three cardiac parameters are estimated using the feature points of the SCG signal. Initially, SCG signals are delineated by employing a data-adaptive MVMD model and a few simple

decision rules to estimate AO instants. The utilized method prioritize the extraction of prominent AO phase information as an initial step, which further helps in the estimation of other fiducial points of an SCG. The identified AO instants were used for estimating pAC instants, which collectively contribute to the derivation of cardiac parameters, including BP, HRV indices, and STIs. A mathematical model based on LVET′ and HR is designed to estimate both SBPs and DBPs using an individual calibration procedure. In the second study, the detected AO instants are used to derive the AO-AO interval time series, which are investigated for HRV analysis. In the third section, all the extracted mechanical features of SCG are combined with demographical information of subjects to identify R-peaks on SCG using a DFN. Subsequently, the identified R-peaks on SCG are used to demonstrate the potential standalone application of SCG in the measurement of STIs. The performance evaluation of each measurement study shows the potential of SCG to be a part of health assessment tools. However, acquiring SCG data requires direct contact of the sensor to the subject's body, which may cause discomfort and limit mobility. To address the drawbacks of SCG modality, in the next chapter, we will discuss a non-contact camera-based method for vital signs measurement.

3

Camera-Based Remote HR Estimation Using Conventional Method

Contents

The heart rhythm is a vital cardiac parameter that plays a key role in noninvasive assessment of various cardiovascular-phenomena, such as cardiac output and BP in hemodynamics, and HRV measures in assessing autonomous nervous system [315]. Estimating HR information is of a great importance in determining physiological health and mental state of an individual. The HR is influenced by many factors, such as air temperature, body posture, physical exercise, medication, emotions and numerous other factors. Therefore, estimation of precise HR is a challenging task. For medical practice, there exists a lot of HR measurement tools like ECG, PPG, and heart sound. However, these cardiac modalities use the sensors that need to be in contact with the body. This contact-based approach limits the flexibility and the mobility of users. Moreover, recording and analysis of the signals need a well-established clinical setup and an expert. In last few decades, researchers are more concerned about developing contactless cardiac monitoring systems to overcome the limitations of the conventional contact-based methods. Non-contact cardiac health assessment has been accomplished by laser and microwave doppler radar [316, 317]. However, these techniques require the use of highly specialized and expensive sensors. To overcome these limitations, the proposed method utilizes rPPG technology for remote HR estimation. The proposed model combines an adaptive and non-recursive 2D variational mode decomposition (2D-VMD) model with azimuthally averaged power spectrum density (AAPSD) and kurtosis measures for the extraction of a reliable rPPG signal. Subsequently, HR values are measured from the extracted rPPG signal. The organization of this chapter is as follows: the description of VMD-based framework is discussed in Section 3.1. Section 3.2, describes the utilized dataset and examines the experimental results of the framework to assess its efficacy.

3.1 Variational Mode Decomposition Based Framework

The 2D-VMD method is analogous to its 1D-predecessor, where it non-recursively decomposes a 2D analytical signal/image into a finite number of intrinsic mode functions (IMFs). Ensembling these IMFs reproduces the original image either exactly or in least-square sense [318]. The 2D-VMD has a wide ranging applications in the field of image processing like image fusion, feature extraction and classification, image dehazing, and image denoising [319–321]. However, the 2D-VMD algorithm has not been explored to denoise the video frames corrupted with noises, such as illumination variation and motion artifacts. Earlier, illumination variation was suppressed by an adaptive signal processing algorithm known as empirical mode decomposition (EMD) [322]. Bhagavatula *et al.,* used the EMD

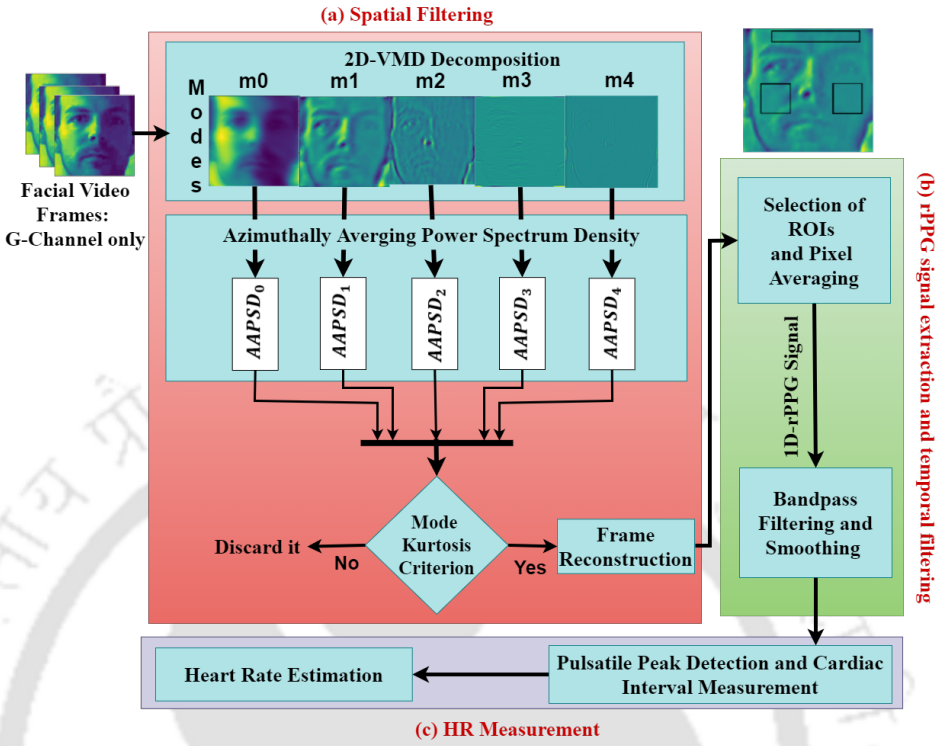


Figure 3.1: Block diagram of the proposed method

algorithm to decompose a facial image into their IMFs and subsequently, the facial image was reconstructed by selective ensembling of IMFs that do not contain illumination effects. The EMD technique has a good time-frequency resolution. However, it has some limitations, such as extremal point finding, interpolation of envelopes, stopping criterion, mode mixing, sensitivity to noise and the lack of mathematical theory [318]. These limitations are well addressed by the VMD scheme. The VMD algorithm is based on well-founded mathematical concepts, and it is less susceptible to noise because wiener filter is used to update the modes in the Fourier domain. Additionally, convergence criterion is well controlled to effectively eliminate mode mixing with a high noise immunity [318]. In our proposed method, the VMD model is used to compensate illumination variation and motion artifacts. Earlier studies show that only the low-frequency component is sensitive to illumination changes, whereas the corresponding high-frequency components are more robust to illumination changes [323, 324]. Also, most of the image information is concentrated in lower order IMFs as compared to higher order IMFs [321, 324]. Hence, a proper selection of modes is important to get the significant information of an image. However, it is a crucial task to select the informative modes from the decomposed modes. To overcome this problem, we have combined the 2D-VMD with the AAPSD and kurtosis measures to retain the useful information of the image. It may help to extract a reliable rPPG signal. The

block diagram of the proposed method is shown in Figure 3.1. It consists of three modules: (i) spatial filtering (ii) rPPG signal extraction and temporal filtering (iii) peak detection and HR estimation. Each of the modules is explained in detail in the following subsections.

3.1.1 Spatial filtering of facial video frames

The basic building blocks of the VMD algorithm are Wiener filter, Hilbert transform, and harmonic mixing. A real valued multicomponent 2D signal/image $f(x)$ is decomposed using 2D-VMD into a finite number of 2D IMFs $m_k(x)$, which are mostly centered around a central frequency ω_k. The bandwidth of each mode is assessed by the following steps [318]:

(i) Construction of an analytic signal for each mode $m_k(x)$ by means of 2D-Hilbert transform to obtain a single-sided frequency spectrum.

$$m_{AS,k}(x) = m_k(x)^* \left(\delta(\langle x, \omega_k \rangle) + \frac{j}{\pi \langle x, \omega_k \rangle} \right) \delta\left(\langle x, \omega_{k,\perp} \rangle\right) \tag{3.1}$$

where $j = \sqrt{-1}$ and k $(k = 1, 2, \cdots, K)$ denotes the number of modes. $m_{AS,k}(x)$ is the analytic signal of k^{th} mode, $\delta(\cdot)$ is the unit impulse function, * is the convolution operator, and ω_k, $\omega_{k,\perp}$ are the orthogonal pair.

(ii) Spectrum of each analytic signal is shifted to the baseband by mixing with an exponential $e^{-j\langle \omega_k, x \rangle}$ tuned to the corresponding estimated centre frequency ω_k.

$$m_{AS',k}(x) = \left[m_{AS,k}(x)\right] e^{-j\langle \omega_k, x \rangle} \tag{3.2}$$

(iii) Finally, bandwidth is estimated through the squared L2-norm of the gradient (denoted by ∇) in Eq. (3.2):

$$\| \nabla \left[m_{AS',k}(x) \right] \|_2^2 \tag{3.3}$$

The resultant constraint variational problem of the 2D analytic signal is obtained by the above Eqs. (4.1)–(3.3) as:

$$\min_{\{m_k, \omega_k\}} \sum_k \alpha_k \| \nabla \left[m_{AS',k}(x) \right] \|_2^2$$
$$\text{s.t.} \sum_{k=1}^{K} m_k(x) = f(x) \tag{3.4}$$

where α represents the balancing factor for bandwidth allowance, $m_k = \{m_1, m_2, \cdots, m_K\}$ and $\omega_k = \{\omega_1, \omega_2, \cdots, \omega_K\}$ represents the sets of decomposed modes and their corresponding center frequencies. The above constraint optimization problem (Eq. (3.4)) is addressed by using augmented Lagrange multipliers and a quadratic penalty term. Lagrange multipliers are generally used to enforce constraints strictly, whereas quadratic penalty term is used to stimulate reconstruction fidelity. The resultant unconstrained problem is expressed as:

$$LM\left(m_k, \omega_k, \lambda\right) = \sum_k \alpha_k \| \nabla \left[m_{AS',k}(x) \right] \|_2^2 + \ \| f - \sum_{k=1}^{K} m_k(x) \|_2^2 + \langle \lambda(x), f - \sum_{k=1}^{K} m_k(x) \rangle \quad (3.5)$$

where λ is the Lagrange multiplier. To solve Eq. (3.5), alternate direction method of multipliers (ADMM) is used [318]. The optimized solution of the 2D variational constraint model is obtained as:

- Minimization of the argument of unconstrained problem (Eq. (3.5)) w.r.t. m_k, shows the Wiener filter behavior and is expressed as:

$$\widehat{m}_k^{n+1}(\omega) = \frac{\widehat{f}(\omega) - \sum\limits_{k=1}^{K} \widehat{m}_k(\omega) + \frac{\lambda}{2}}{1 + 2\alpha_k(\omega - \omega_k)^2} \quad (3.6)$$

  where n is the iteration index, $\widehat{f}(\omega)$ and $\widehat{m}_k(\omega)$ are Fourier transform of $f(x)$ and $m_k(x)$, respectively.

- Minimization w.r.t. center frequency ω_k on the half plane Ω_k is expressed as:

$$\omega_k^{n+1} = \frac{\int_{\Omega_k} \omega |\widehat{m}_k(\omega)|^2 d\omega}{\int_{\Omega_k} |\widehat{m}_k(\omega)|^2 d\omega} \quad (3.7)$$

- Maximization w.r.t λ is achieved by gradient ascent and expressed as:

$$\widehat{\lambda}^{n+1}(\omega) \longleftarrow \widehat{\lambda}^{n}(\omega) + \tau \left(\widehat{f}(\omega) - \sum_k \widehat{m}_k^{n+1}(\omega) \right) \quad (3.8)$$

where the convergence criterion is given as:

$$\sum_{k=1}^{K} \frac{\| \widehat{m}_k^{n+1} - \widehat{m}_k^{n} \|_2^2}{\| \widehat{m}_k^{n} \|_2^2} < \epsilon \quad (3.9)$$

The above optimization solution is obtained in the frequency domain and therefore, the corresponding time domain modes are estimated as the real part of the inverse Fourier transform of the respective filtered analytic signal.

Due to having better pulsatile strength with low noise power, only green channel is used for the processing in the proposed method. The employed VMD scheme mainly depends on two major parameters: the balancing parameter (α) and the number of modes (K). Bandwidth of a mode is controlled by the α parameter, whereas K controls the distribution of energy. For a specific application, both the parameters are required to be selected optimally. The proposed algorithm uses a small alpha ($\alpha = 10$) for decomposition of each frame into five modes ($K = 5$). These parameters were determined to be optimal for separating the frequency components of a frame. The criterion of mode-separability was achieved by estimating the minimum spectral correlation between successive modes. Subsequently, the frame is reconstructed after discarding non-useful modes. The process is explained in the following subsections. For further computations on 2D image data, the notations for original image $f(x)$ and mode $m_k(x)$ are changed by $f(x, y)$ and $m_k(x, y)$, respectively, and their Fourier transform pairs are represented by $\widehat{f}(u, v)$ and $\widehat{m}_k(u, v)$ from 3.1.2.

3.1.2 Proposed reconstruction scheme

The original image contains both signal and noise information that are resolved into different IMFs. Let's suppose, $f(x, y)$ be the original frame that is expressed as:

$$f(x,y) = m_0(x,y) + \sum_{k=1}^{L(<K)} \phi_k(x,y) + \sum_{k=L+1}^{K} \psi_k(x,y) \tag{3.10}$$

where $m_0(x, y)$ is the lowest order mode that captures the illumination effect [322], $\phi(x, y)$ represents a set of useful modes (mode-1 to mode-L) containing the signal information, where L relies on AAPSD of all the modes and an applied criterion based on their kurtosis measurements. From the experimentation, it was found that in most of the cases, modes 1 and 2 satisfy the selection criterion for obtaining useful modes. The last component $\psi(x, y)$ carries the unwanted modes that may contain non-rigid motion artifacts caused due to eye blinking, emotion expressing, and mouth movement. The proposed reconstruction method relies on AAPSD of modes and their kurtosis measurements. For each of the decomposed modes, the 1D-PSD is measured using AAPSD, and subsequently, the corresponding kurtosis measurement is performed. Here, reconstruction of the frame is achieved by synchronized aggregation of the modes having super-Gaussian behaviour. The AAPSD and kurtosis measurements for each mode are explained as follows:

3.1.2.1 Azimuthally averaged power spectrum density (AAPSD)

The 2D-PSD is a Fourier analysis-based technique that is used to measure the frequency components of an image. For each mode, the 2D-PSD is measured as the squared amplitude per unit area of its corresponding spectrum and is expressed as [325]:

$$PSD_k(u,v) = \frac{1}{A}|\widehat{m}_k(u,v)|^2 \tag{3.11}$$

where u, v are the spatial frequencies, and A corresponds the total number of pixels in an image. The power spectrum contains 2D information, but it is often useful to summarize the spectrum with its 1D representation. The 1D-PSD can be approximated from 2D-PSD representation by gathering and averaging similar frequency components into a vector that shows the same profile as the original PSD function, without loosing relevant information. Hence, an azimuthal integration is applied to compute the 1D-PSD information from the 2D power spectrum. The azimuthal PSD either sums or averages 2D spectrum values radially from the origin [325]. The radial frequency variable, ρ, is defined as:

$$\rho = \sqrt{u^2 + v^2} \tag{3.12}$$

The expression to obtain AAPSD from the mode spectrum is given as:

$$\gamma^k(\rho) = \frac{1}{P} \sum_{\|u,v\|_2 < \rho} PSD_k(u,v) \tag{3.13}$$

where P is the total number of pixels within the radii ρ of the 2D spectrum, and $\rho \in [0, W_k/2]$, where W_k represents the width of the 2D spectrum of k^{th} mode. Figure 3.2 shows the decomposed IMFs of a video frame recorded under natural light. The IMFs are processed to estimate the 2D power spectrums, and further summarized by the AAPSD for better visualization and understanding of the spectral components. The AAPSD clearly shows that the frequency components are distributed in ascending order in the consecutive modes as depicted in Figure 3.2(d)-(h). It is observed that the illumination component of the frame is captured by the zeroth mode, whereas, the unwanted information falls in the higher order modes. In most of the cases, modes 1 and 2 satisfies the condition of frame reconstruction.

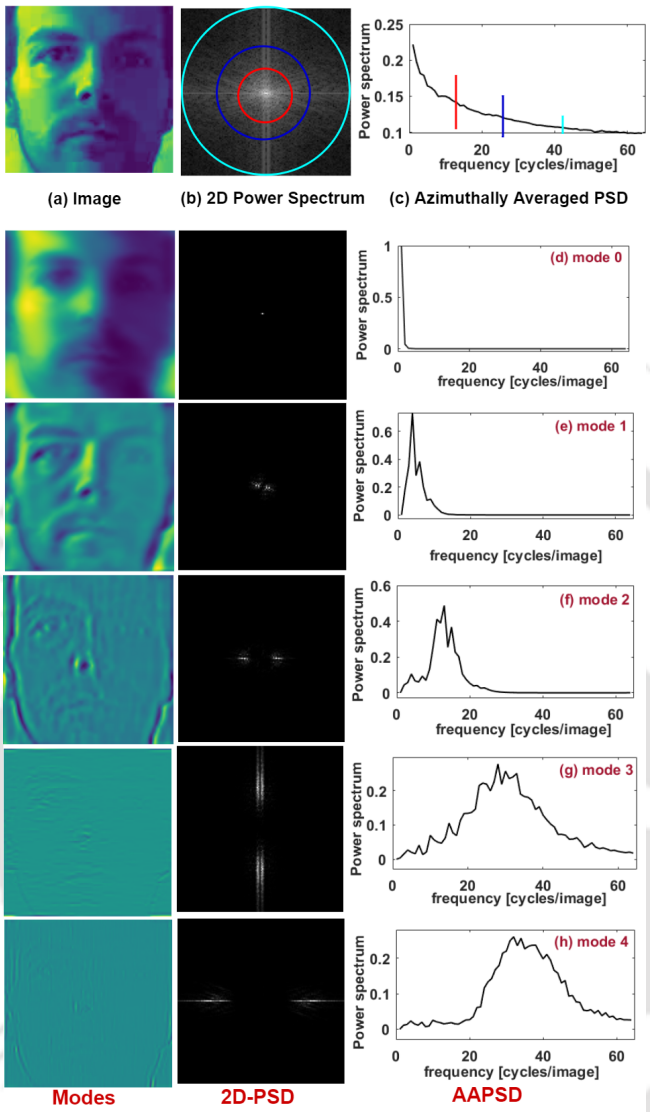


Figure 3.2: Power spectrum representation of a video frame. (a)–(c) Estimation of azimuthally averaged power spectrum density (AAPSD), where spectral power is measured radially over the 2D spectrum. An example is shown via red, blue and cyan rings as different frequency components. (d)–(h) modes [left] with their corresponding 2D-PSDs [middle] and AAPSDs [right].

3.1.2.2 Proposed Multimode Kurtosis

Kurtosis is the fourth standardize moment that indicates the sharpness of the peak of probability distribution [326]. Here, the kurtosis is used to measure Gaussianity of each AAPSD sequence, and hence, it is named as multimode kurtosis. Based on the kurtosis value, the Gaussanity of any probability distribution is classified into three categories, (i) Gaussian (kurtosis = 3), (ii) sub-Gaussian (kurtosis < 3), and (iii) super-Gaussian (kurtosis > 3) [326]. The multimode kurtosis for the k^{th} mode

AAPSD ($\gamma^k(\rho)$) is computed as:

$$\beta^k = \frac{\frac{1}{W_k/2}\sum_{\rho=1}^{W_k/2}\left(\gamma^k[\rho] - \overline{\gamma^k}[\rho]\right)^4}{\left(\frac{1}{W_k/2}\sum_{\rho=1}^{W_k/2}\left(\gamma^k[\rho] - \overline{\gamma^k}[\rho]\right)^2\right)^2} \tag{3.14}$$

where $\bar{(.)}$ corresponds to mean operator. Here, we assumed that the non-rigid motion artifacts in a video frame are mainly captured by higher order modes and they will show sub-Gaussian behavior (i.e., $\beta < 3$) for their AAPSDs. As an experiment, when the multimode kurtosis is estimated, it is found to have a value less than 3 for the higher order AAPSDs. This shows the sub-Gaussanity behaviour of the noisy modes and thus, enables us to classify the signal dominant modes from the noisy modes. Figure 3.3(a) and (b) show the measured kurtosis values of AAPSDs for video frames under studio and natural light, respectively. It is concluded that that the kurtosis value decreases for the higher order modes. Finally, the image is reconstructed by selecting the signal dominant modes

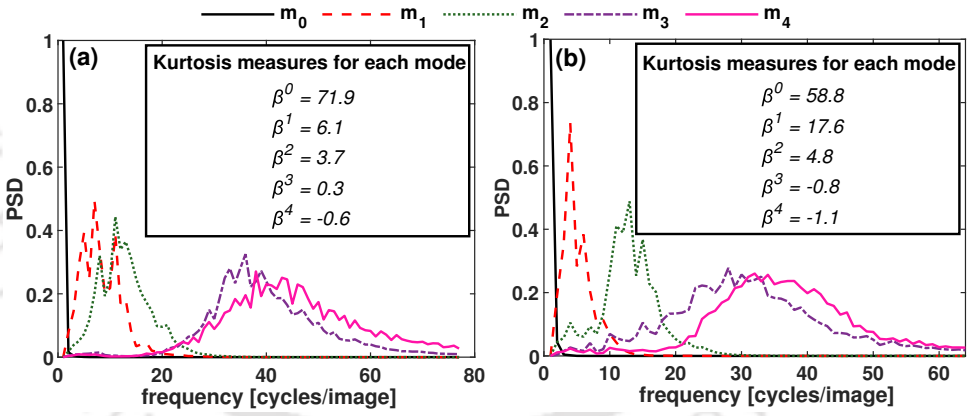


Figure 3.3: Multimode kurtosis measurement of the estimated AAPSDs under — (a) studio light, and (b) natural light.

having the super-Gaussian AAPSDs. The reconstructed frame is expressed as:

$$r(x,y) = \sum_{k\in\zeta} m_k(x,y) \quad \text{where } \zeta = \underset{k}{arg}(\beta^k > 3) \tag{3.15}$$

Figure 3.4 shows the decomposed modes and reconstructed image under both the lighting conditions. It shows the distribution of the original image components by different modes, and the reconstructed frame $r(x,y)$ that is obtained by summing useful subbands.

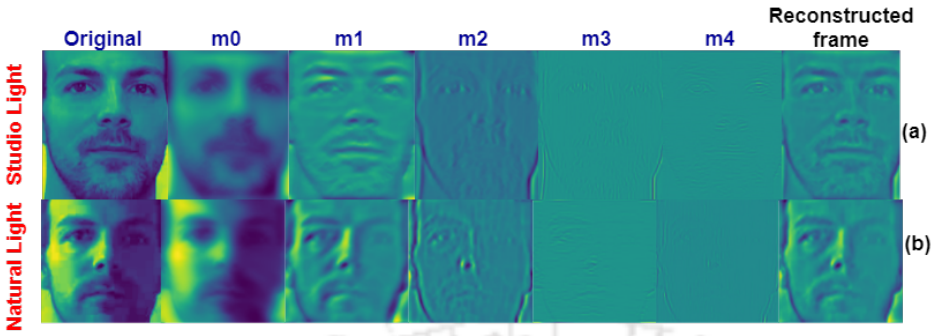


Figure 3.4: Reconstruction of frames after denosing from the 2D-VMD approach under — (a) studio light, and (b) natural light.

3.1.3 rPPG signal extraction and HR estimation

After the frame reconstruction, few facial regions, including forehead, right and left cheeks are selected as regions of interest (ROIs) for the extraction of rPPG signal. The signal is computed by spatial averaging of the pixels in the selected ROIs as:

$$rPPG = \frac{1}{TP}\sum_{i=1}^{3}\left(\sum_{x=1}^{M-1}\sum_{y=1}^{N-1} r_i(x,y)\right) \tag{3.16}$$

where i acts as an ROI index representing forehead, right cheek and left cheek regions, and TP is the total number of pixels within all the three ROIs. An interpolation operation is performed to increase the temporal resolution of the obtained trace from 20 Hz to 256 Hz. The unwanted frequencies other than cardiovascular pulse is filtered out by applying a fifth-order Butterworth bandpass filter with cut-off frequencies 0.8 and 4 Hz (that corresponds to 48-240 beats per minute (bpm)). The order of the filter is selected empirically starting from 2, and an optimal smallest order is found to be 5. Then peaks are detected in the filtered signal by employing simple amplitude-temporal thresholding-based technique. Figure 3.5 shows the detected peaks in the ground truth PPG signal and the filtered rPPG signal. The time intervals between the consecutive peaks are used to measure the HR in bpm that is calculated as:

$$HR = \frac{60}{\frac{1}{J}\sum_{j=1}^{J} CI_j} \tag{3.17}$$

where $j(j = 1, 2, ...J)$ corresponds to the heartbeat index. The cardiac interval, also denoted by CI, is measured as: $CI = t_{i+1} - t_i$, where t_i is the time instant of the i^{th} peak of the filtered rPPG signal.

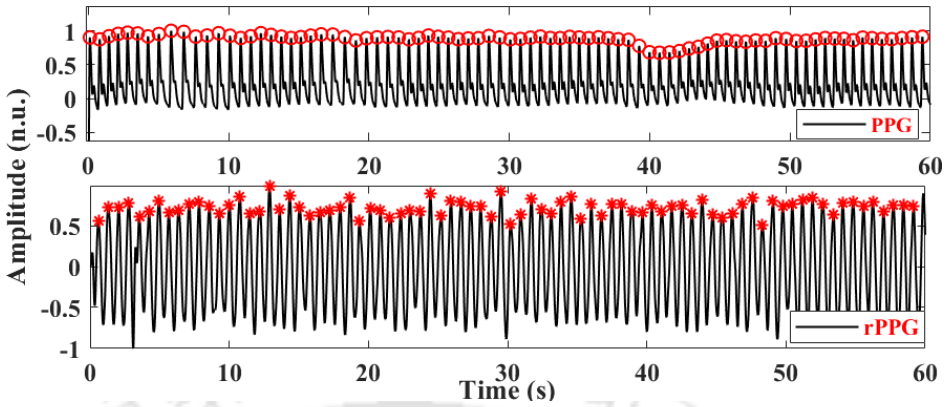


Figure 3.5: Peak detection results for the reference PPG and the filtered rPPG signals.

3.2 Materials and evaluation

3.2.1 Datasets

Experiments are performed on two datasets, a private one and the publicly available COHFACE database. For both the datasets, videos were recorded in sitting position without any large body movements. The datasets consist of subjects with varying demographics, such as age, gender and skin tone. The COHFACE dataset contains 160 videos of 40 healthy individuals (12 females and 28 males, age: 35±11 yrs), each one minute long. Videos were recorded at a rate of 20 fps with a resolution of 640 × 480. Simultaneously, blood volume pulse (BVP) signal was acquired at a sampling frequency of 256 Hz. Each participant contributes four sets of data, two under studio light using a 400W halogen light, followed by other two under ambient light by keeping the blinds open.

The private dataset was recorded for five minutes from 25 healthy subjects (4 females and 21 males, age: 28±2 yrs) under both studio and natural lighting conditions. The videos were recorded under a 36W tubelight for studio lighting condition, whereas blinds were kept open for the natural lighting condition. Facial videos were captured using a webcam from QHMPL at a 6 fps rate with a 640 × 480 resolution and saved in the AVI format. As a reference, the PPG was recorded at fingertip using Nonin medical's SenSmart® optoelectronic finger clip sensor. The video and reference signal were synchronously obtained using MP150 DAQ system (BIOPAC Systems, Inc., CA, USA). The data was recorded with the prior consent of the subjects after getting approval from institutional ethics review board, IIT Guwahati.

3.2.2 Results

The performance of the proposed method is statistically and graphically evaluated by comparing the estimated HR values with the reference HR measurements. Additionally, a comparison is made with the well established ICA-based method [178]. Figure 3.6 shows the block diagram of the ICA-based method. In this method, the ICA approach was applied on the facial RGB traces followed by temporal filtering for HR estimation. Table 3.1 presents the statistical comparison between the estimated HRs in terms of mean error (ME, Eq. (3.18)), root mean square error (RMSE, Eq. (3.19)) standard deviation of error (SDE, Eq. (3.20) and correlation coefficient (R, Eq. (3.21)) for both COHFACE and private datasets.

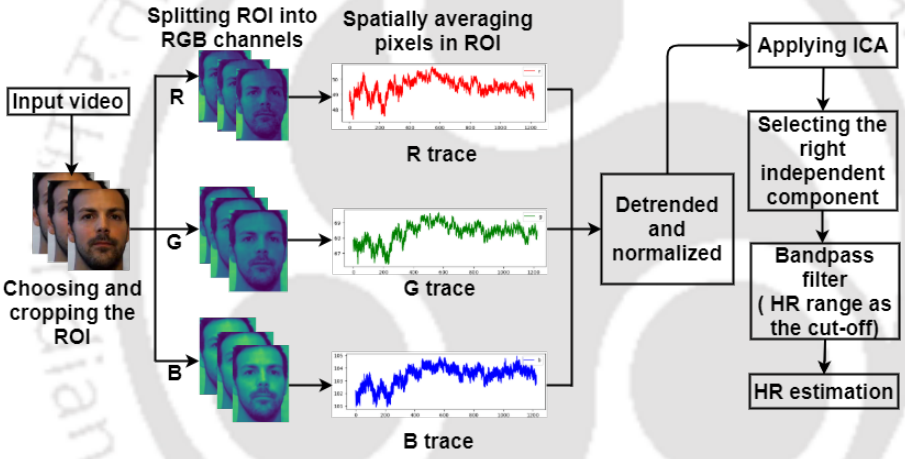


Figure 3.6: Block diagram of the earlier ICA-based method

$$ME = \frac{1}{N}\sum_{i=1}^{N}(HR_i - HR_i^{\dagger}) \tag{3.18}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(HR_i - HR_i^{\dagger})^2} \tag{3.19}$$

$$SDE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}[(HR_i - HR_i^{\dagger}) - ME]^2} \tag{3.20}$$

$$R = \frac{\sum_i (HR_i - \overline{HR})(HR_i^{\dagger} - \overline{HR^{\dagger}})}{\sqrt{\sum_i (HR_i - \overline{HR})^2 (HR_i^{\dagger} - \overline{HR^{\dagger}})^2}} \tag{3.21}$$

In the above equations, the reference heart rate is denoted as HR, the heart rate estimated from the video is denoted as $HR^{\dagger}$, and N is the total number of heartbeats.

For the COHFACE dataset under studio lighting scenario, the ME, RMSE and SDE measures are

Table 3.1: Statistical comparison between HRs (bpm) estimated from the camera-sensor-based methods and conventional PPG sensor technique.

Dataset →		COHFACE		Private Dataset	
Methods	**Statistical parameters**	**Studio Light**	**Natural Light**	**Studio Light**	**Natural Light**
ICA-based	ME	5.95	8.53	25.72	25.20
	RMSE	16.89	16.93	28.49	27.25
	SDE	15.91	14.72	12.51	10.59
	R	-0.04	0.60	0.06	0.32
	$y = ax + b$	-0.04x+66.67	0.52x+23.95	0.05x+53.35	0.22x+39.15
	ME+1.96(SDE)	37.14	37.38	50.26	45.97
	ME−1.96(SDE)	-25.24	-20.33	1.18	4.43
Proposed	ME	-0.94	-0.51	-0.28	-0.32
	RMSE	2.41	2.51	0.82	0.80
	SDE	2.23	2.47	0.79	0.75
	R	0.98	0.99	0.99	0.99
	$y = ax + b$	0.92x+6.05	0.98x+2.03	0.98x+2.07	0.97x+3.02
	ME+1.96(SDE)	3.43	4.33	1.27	1.15
	ME−1.96(SDE)	-5.31	-5.35	-1.83	-1.79

found to be 5.95 bpm, 16.89 bpm, and 15.91 bpm, respectively for the HR estimated using the ICA-based method. Whereas, they are found to be -0.94 bpm, 2.41 bpm, and 2.23 bpm, respectively using the proposed method. For natural lighting condition, the results are obtained as 8.53 bpm, 16.93 bpm, and 14.72 bpm, respectively using the ICA-based method, and they are found to be -0.51 bpm, 2.51 bpm and 2.47 bpm for our proposed method. Similarly, on private dataset the results under studio lighting condition are found to be 25.72 bpm, 28.49 bpm, and 12.51 bpm for ME, MAE and SDE, respectively using the ICA-based method, and -0.28 bpm, 0.82 bpm and 0.79 bpm using the proposed method. Under natural light, the parameters are found to be 25.20 bpm, 27.25 bpm and 10.59 bpm using the ICA method, and -0.32 bpm, 0.80 bpm and 0.75 bpm using our method. It is observed from Table 3.1 that the errors obtained from the proposed method is considerably low as compared to the ICA-based method, thus showing the robustness of our method in accurate detection of HR parameter. However, due to the involvement of both spatial and temporal filtering, the proposed method requires a bit more computational time than the ICA-based method, with a large extent of improvement in the performance. The overall HR traces obtained using rPPG methods and ground truth sensor are shown in Figure 3.7(a) and (b) under studio and natural lights, respectively. Here, x-axis represents the video sample number and the y-axis represents the estimated average HR of the corresponding video data. It clearly show that our proposed method tracks well to the ground truth values. For more quantitative results, the Bland-Altman analysis is performed on these reference and estimated

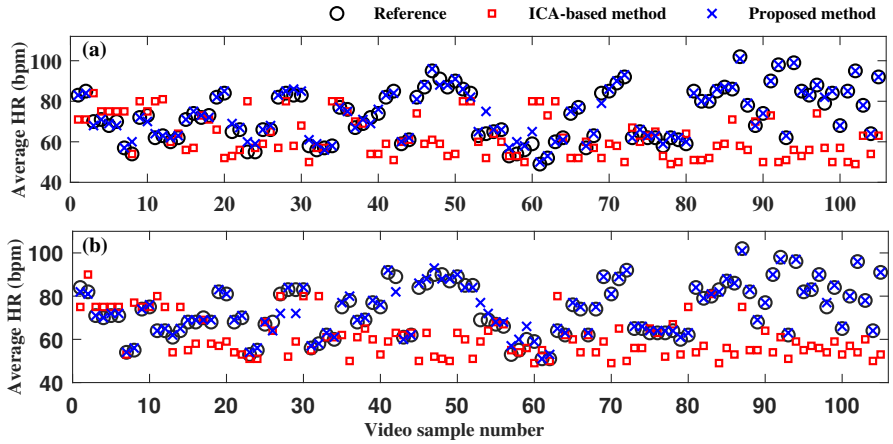


Figure 3.7: Comparison of HR information derived from the proposed method with the estimated HR from the ICA-based method and the reference HR from the PPG under — (a) studio light, and (b) natural light.

HRs. Figs. 3.8 and 3.9 show the Bland-Altman plots to reveal the agreement between estimated and reference HRs for both COHFACE and private datasets, respectively. Here, the estimated HR values using both the proposed method and the ICA-based method are plotted against the HR measured from the reference PPG signal. The x-axis represents the mean HR obtained from the reference PPG and camera-based rPPG signal, whereas the difference between both the HR values is represented by the y-axis. Successively, the bias or mean of the differences is measured along with 95% limits of agreement (LOA) that are defined as mean$\pm$1.96$\times$SDE. Both mean and LOA values are also listed in Table 3.1 for both the methods. The private dataset under both the lighting conditions is showing the minimum value of bias and LOA for our proposed method. For studio lighting condition, the mean bias is -0.28 bpm with the LOA of -1.83 to 1.27 bpm. While, they are found as -0.32 bpm and -1.79 to 1.15 bpm, respectively under natural lighting condition. Also, for the COHFACE dataset, the obtained bias and LOA using our proposed method are significantly small as compared to the ICA-based method. This remarkable results show that the proposed method improves the estimation of HR under various dynamic condition. Further, the performance comparison is also shown in terms of correlation coefficients (R). The correlation coefficients between the HR values obtained from the PPG sensor and the camera-based system using the two methods are presented in Table 3.1. For the COHFACE dataset, the ICA-based method produces correlation as -0.04 and 0.60 under studio and natural light, respectively. Whereas, they are found to be 0.98 and 0.99 using our method. For the private dataset, both the methods achieve the correlation coefficient pairs, (0.06, 0.32) and (0.99, 0.99), respectively for the studio and natural lighting conditions.

Abstract-TH-3588_186102009

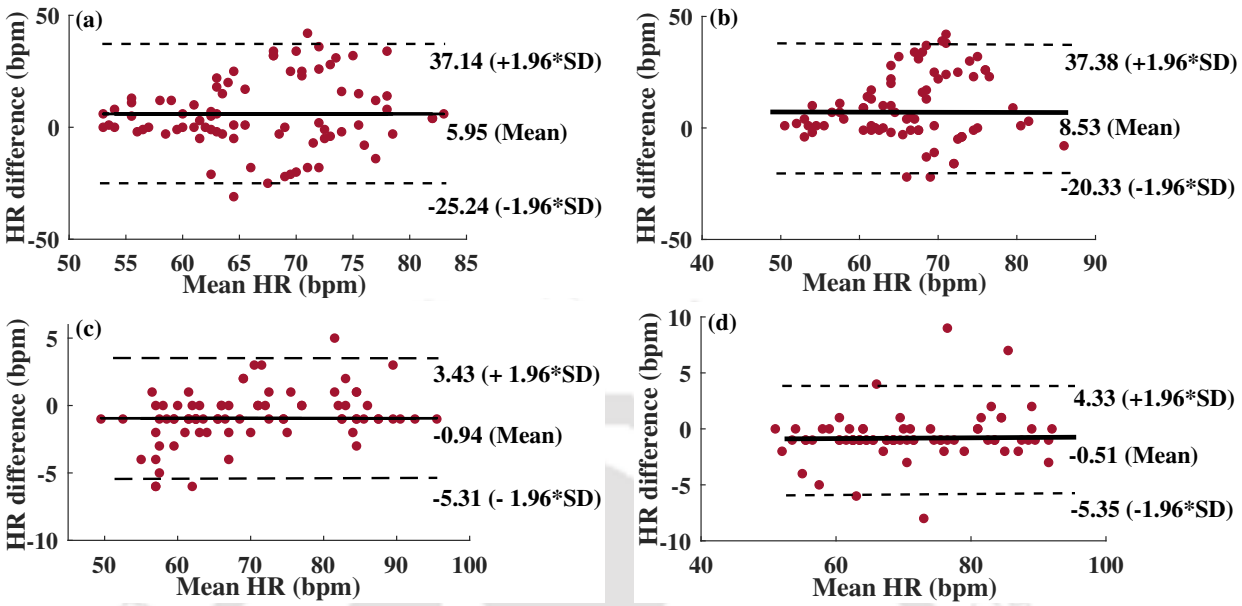


Figure 3.8: Bland-Altman plots for evaluating our HR estimation method under studio [left] and natural [right] lights for the COHFACE dataset. (a)-(b) ICA-based method, and (c)-(d) proposed method. The PPG derived HR is used as a reference for the Bland-Altman analysis.

As an initial step, the proposed method utilizes the well-known Dlib library for the face detection [168]. However, it is unable to detect the facial regions when facial features are absent due to head rotations. Therefore, the proposed work was validated only on the videos containing no head rotations. Although the current study has achieved remarkable results on the datasets with varying lighting conditions, inter-subject skin tone variabilities and facial expressions, the proposed work still needs further validation on different real time conditions. As a future direction, the performance of the proposed work can be evaluated on a large dataset consisting of various body movements and postures. It will aid more impact to the proposed work and will ensure its applicability for real-time dynamic states.

3.3 Conclusion

In this work, a 2D-VMD based method is proposed for non-contact HR measurement using a low-cost web camera. It works well for both non-rigid motions and varying lighting conditions. The main contribution of our work is the spatial filtering of video frames using the 2D-VMD approach. The proposed method decomposes the original frame into different modes and later, the frame is reconstructed using the modes that are selected on the basis of AAPSD and multimode kurtosis

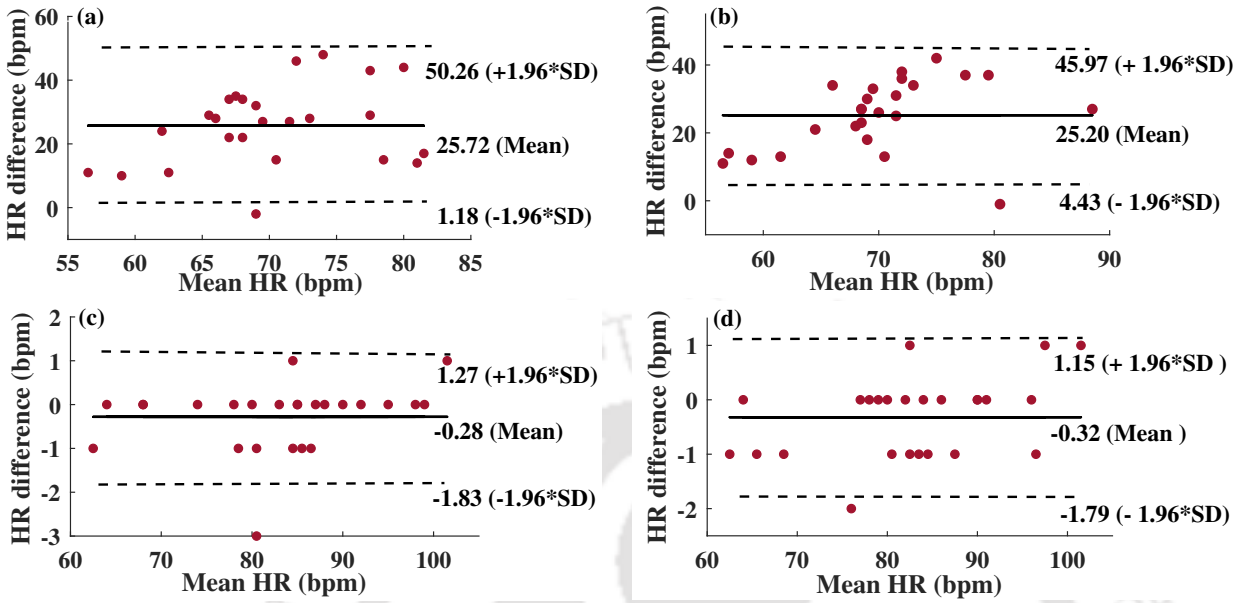


Figure 3.9: Bland-Altman plots for evaluating our HR estimation method under studio [left] and natural [right] lights for the private dataset. (a)-(b) ICA-based method, and (c)-(d) proposed method.

values. From the reconstructed frame, three best ROIs, namely, forehead, right and left cheeks, are selected to detect the rPPG signal. Further, filtering and smoothing is applied to clean the rPPG signal. Subsequently, estimated HRs are statistically compared with the HRs obtained from the PPG signal. The obtained results show an insignificant error and a strong correlation. Also, the Bland-Altman analysis reveals a good agreement between the proposed system and the reference system. Additionally, the proposed method proves to be more accurate when compared with the earlier ICA-based method. Thus, the proposed spatio-temporal based method is effective for a reliable HR estimation under dynamic conditions. Hence, as a low-cost, easy-to-use, and contact-free system, it could be a better solution for robust and ubiquitous HR monitoring in both clinical and non-clinical environments.

3.4 Summary

In this paper, a novel spatial-temporal filtering based method using a 2D-VMD is proposed. It extracts a reliable rPPG signal under various realistic scenarios. The proposed method spatially decomposes the recorded facial video frames into different modes and further reconstructs the original frames by selecting the modes, free from lighting and motion variation effects. Here, the rPPG signal is extracted from the ROIs of the reconstructed frames and subsequently, the HR is measured. Although the proposed method seems to be promising, it still have some limitations. The selection

of modes are based on certain assumptions for the simplification of noise reduction. However, the involvement of noise in real-time applications are more complicated and diverse. This may deviate from the assumptions of proposed method and may produce incorrect results. Additionally, the method involves both spatial and temporal filtering which makes it computationally expensive and thus require further optimization to implement it for real time application. Recently deep learning (DL) techniques have attracted many researchers for addressing various practical problems. The DL technique demonstrated excellent performance when trained with extensive datasets covering scenarios relevant to the targeted task. This characteristic endows DL methods with robustness and flexibility to effectively solve rPPG problems across different practical scenarios. Hence, the subsequent chapters describes DL frameworks that are being explored for remote cardiac measurements.

4

Camera-Based Remote HR and HRV Estimation Using Deep Neural Network Frameworks

Contents

The conventional rPPG methods have played a significant role in the growth of low-cost camera based cardiac health monitoring system. However, they rely on certain assumptions and their performance may degrade with real-time dynamic interferences. Recently, deep neural network (DNN) based methods have emerged in the rPPG domain, and they prove to perform well in noisy environments. This chapter presents proposed rPPG-based cardiac parameter measurements with the use of DNN frameworks. The chapter includes two different approaches. The first approach utilizes a 1D-CNN based DelNet network to accurately identify systolic peaks in the rPPG signal. The input to the network is the raw green trace that is obtained by spatially pooling the facial pixels of the green channel frame. In the second approach, we introduce a novel 2D-CNN based time-frequency learning (TFL) framework for estimating the rPPG signal using a scalogram-based feature map extracted from facial videos. The feature map is obtained using the raw RGB traces and it involved the enhancement of pulsatile profile within the RGB traces with the help of wavelet transformation. This enhancement enables the proposed framework to characterize a wide range of frequencies associated with cardiac activity. Both networks are designed to estimate reliable rPPG signals for robust measurement of HR and HRV indices. The current chapter is organized as follows: the systolic peak delineation in rPPG signal using 1D-CNN based DelNet framework is presented in Section 4.1. Section 4.2 provides details of our proposed TFL framework. Finally, the chapter is summarized in Section 4.3.

4.1 DelNet based Systolic Peak Delineation in rPPG Signal

Figure 4.1 shows the block diagram of the proposed method. It consists of two module: (i) Feature signal extraction and (ii) Delineating network (DelNet). Both the modules are explained in detail in the following subsection.

4.1.1 Feature signal extraction

Initially, the process starts with the extraction of face from the sequence of video frames that are provided to the module. Here, the face is detected by utilizing the well-know Dlib library [168]. Few facial regions, such as eyes, eyebrows and lips are not considered as reliable elements for beat detection. Therefore, the module segments only the facial skin as ROI that consists most of the blood flow information for cardiac health assessment. The skin segmentation is achieved with simple thresholding scheme on HSV color space [327]. Subsequently, the raw RGB traces are obtained by spatial averaging of the facial skin pixels. In most of the studies, it is observed that green channel

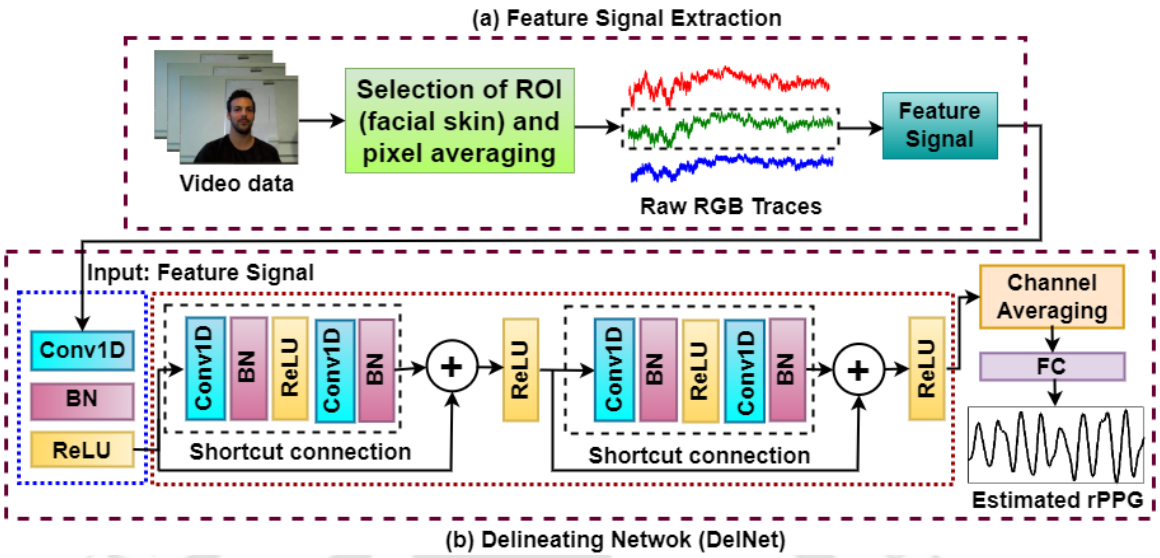


Figure 4.1: Block diagram of the proposed DelNet framework

has the maximum strength of the pulse wave in comparison to red and blue channels [133]. Hence, we have selected the raw green trace as the feature signal for the DelNet module.

4.1.2 Delineating network (DelNet)

DelNet is a deep neural framework that is trained to model a reliable rPPG waveform. The complete network architecture is presented in Figure 4.1 (b). The input to this network is a feature signal that is extracted from the prior module. The shape of the input is $16 \times 1 \times 512$ ($B \times C \times L$, where B is the batch size, C is the number of channels and L is the length of the signal), and the shape of output is 16×512 ($B \times L$). Inspired by the advantages and success of residual networks (ResNets) [328], the proposed DelNet has also introduced the skip connections within the framework. The network starts with a convolution block and followed by two layers of similar behaviour. Both the layers consists of skip connection after every two convolution blocks and perform the operation with fixed feature map dimensions 2 and 4, respectively. Subsequently, the output feature maps are pooled together to obtain a single channel at the output. Finally, the averaged output is flattened and passed to a fully connected (FC) layer having linear activation units. The convolution operations are performed with a kernel size of 3 followed by a batch normalization and ReLU nonlinearity. An additional dropout operation of 20% is introduced before the final layer to regularize the network against overfitting. The details of the DelNet are tabulated in Table 4.1.

Since, the framework is expected to estimate a reliable rPPG waveform with systolic peak instants

Table 4.1: DelNet architecture

Layer name	Kernel size, channels / Stride	Output shape
Conv_1	3, 2 / 1	16×2×512
Conv_2	[3, 2 / 1]×2	16×2×512
Conv_3	[3, 4 / 1]×2	16×4×512
Pooling	-	16×512
Dropout	ratio 0.2	16×512
FC	-	16×512

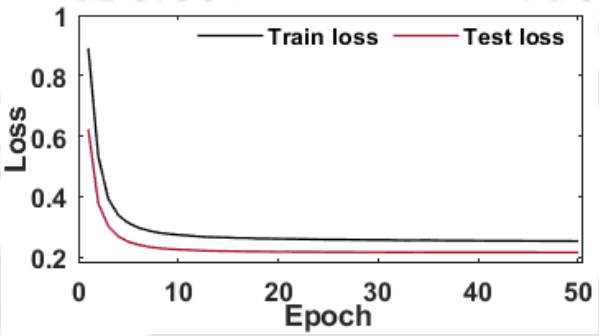


Figure 4.2: Training and testing loss curve

corresponding to the ground truth PPG signal, a negative Pearson correlation coefficient is used as a loss function [231]. The loss function is defined as:

$$Loss = 1 - r(x, y)$$
$$R(x, y) = \frac{\sum_i (x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_i (x_i - \overline{x})^2 (y_i - \overline{y})^2}} \tag{4.1}$$

where, $R(x, y)$ is the Pearson correlation coefficient, x and y are the ground truth PPG and estimated rPPG waveform, respectively. The network was trained using standard backpropagation algorithm and Adam optimizer. The training was done for 50 epochs with a learning rate of 0.0001 and batch size 16. Figure 4.2 shows the loss curve for both the training and testing phases. From the figure it is clear that the models hyperparameters and the loss function are effective in minimizing the error between the estimated and the ground-truth signals. The network was implemented and trained in Python using Pytorch library.

4.1.3 Materials and Evaluation

4.1.3.1 Dataset

Experiments were performed on the publicly available UBFC-rPPG dataset. The dataset contains 42 videos along with the simultaneously recorded PPG signals. The dataset was collected from 42 healthy subjects (11 females and 31 males) under varying amount of ambient light. Also, the subjects were instructed to play a time sensitive mathematical game to create a wide range of heart rate values. The videos were recorded at 30 fps with a resolution of 640×480 in the uncompressed 8-bit RGB format using a low-cost Logitech C920 HD pro webcamera. While, the ground-truth PPG signals were acquired using a CMS50E transmission pulse oximeter at a sampling rate of 30 Hz. Since, the number of videos in the dataset is relatively small for neural network training, we have sliced each video and its corresponding PPG signal by sliding a 2 sec window with 50% overlapping. This way, a large number of data segments were generated for both training and testing of the proposed framework. Moreover, each segment was interpolated to increase the temporal resolution from 30 Hz to 256 Hz. The training and testing set contains 1840 and 782 segmented data of 2 sec duration ($L = 2 \times 256$), respectively.

4.1.3.2 Results

The effectiveness of the proposed method is measured on the average HR value and HRV parameters that are derived from the estimated rPPG waveform. Initially, the systolic peak instants on the signals are detected using a simple amplitude-temporal thresholding scheme. Then the peak instants are provided to an open source Heartpy toolkit that measures the average HR value and computes the time domain and the frequency domain HRV parameters [329]. Figure 4.3 shows the detected systolic peaks in the reference PPG signal and the estimated rPPG signal. Further, the method is compared with four state-of-the-art methods: ICA, CHROM, PBV and OMIT [178,179,189,190]. The performance of the framework is evaluated using a few benchmarked statistical metrics, such as MAE, RMSE, SDE, and R.

Averaged heart rate (HR)

The statistical comparison of the averaged HR values estimated from the rPPG methods and the reference PPG signal are presented in Table 4.2. From the table it is clear that the proposed DelNet has the minimum error values compared to the state-of-the-art methods. The MAE, RMSE, and SDE

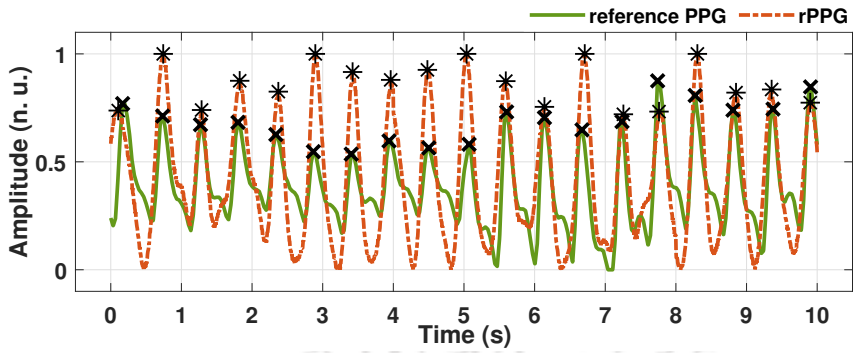


Figure 4.3: Peak detection in the ground-truth PPG and estimated rPPG signals (Green: PPG signal, Red: rPPG signal)

Table 4.2: Statistical comparison between averaged HR values estimated from the rPPG methods and the reference PPG sensor

Methods	MAE	RMSE	SDE	R	Mean+1.96*SDE	Mean	Mean-1.96*SDE
ICA	6.382	7.478	4.005	0.811	12.858	-1.772	-16.402
CHROM	1.883	2.68	1.959	0.986	3.477	-1.273	-6.022
PBV	5.843	6.549	3.04	0.883	11.330	-1.505	-14.341
OMIT	3.312	4.345	2.89	0.953	5.214	-2.259	-9.732
DelNet	1.132	1.448	0.927	0.994	2.479	-0.350	-3.178

for the proposed framework are found to be 1.132, 1.448, and 0.927, respectively. For each subject, the averaged HR values estimated using the rPPG methods and the reference sensor are shown in Figure 4.4. It clearly show the HR values estimated from the proposed framework twin significantly with the reference values. Further, the performance is evaluated graphically in terms of Bland-Altman (BA) and regression plots as shown in Figure 4.5. From the BA plots it is observed that the proposed framework has minimum mean of -0.350 along with LOA as 2.479 and -3.178. This result shows the highest agreement of the proposed framework with the reference measurement device. The mean and LOA values are listed in Table 4.2. Additionally, R values that are computed between referenced HR values and estimated HR values are found to be 0.811, 0.986, 0.883, 0.953 and 0.994 for ICA, CHROM, PBV, OMIT, and DelNet, respectively. Thus, the proposed work estimates the average HR values with insignificant error in comparison to the state-of-the-art methods.

HRV analysis

In this study, both time domain and frequency domain HRV parameters are measured, such as average cardiac interval (meanNN) and its corresponding standard deviation SDNN, standard deviation

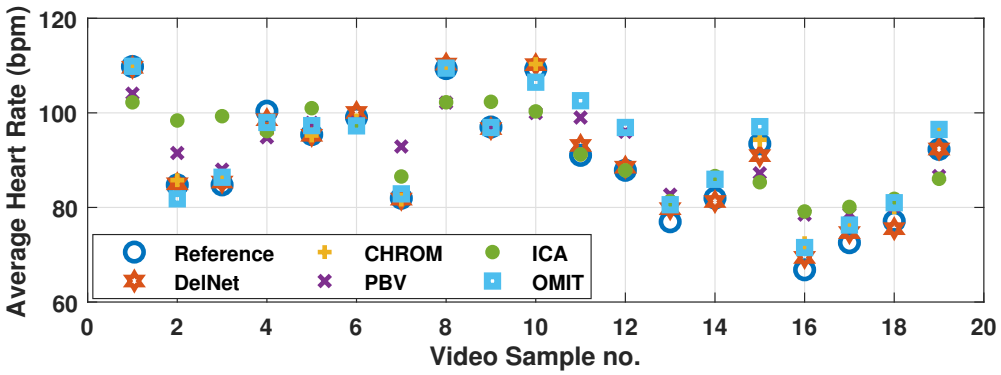


Figure 4.4: Comparison of HR information derived from the DelNet with the estimated HR from the state-of-the-art methods and the reference HR

Table 4.3: Time and frequency domain HRV parameters estimated from the rPPG methods and the reference PPG sensor

	Parameters	Reference	Methods				
			ICA	CHROM	PBV	OMIT	DelNet
Time domain parameters (ms)	meanNN	632.089	620.033	632.849	621.531	623.244	632.839
	SDNN	14.439	32.940	23.641	34.417	25.630	18.816
	SDSD	12.896	27.926	23.394	30.071	22.899	16.590
	RMSSD	18.613	46.560	35.441	48.738	34.378	23.547
	PNN50	0.022	0.306	0.160	0.317	0.153	0.049
Frequency domain parameters (ms^2)	LF	0.302	1.584	0.722	1.943	1.161	0.507
	HF	0.463	3.065	1.236	4.424	2.148	0.956
	LF/HF	0.827	0.494	0.462	0.438	0.571	0.849

of the successive differences (SDSD), root mean square of successive differences (RMSSD), percentage of successive intervals differing more than 50 ms (PNN50), power of low-frequency (LF), power of high frequency (HF), and power ratio of low to high frequency-band (LF/HF). The time domain features are calculated in milliseconds (ms), whereas frequency domain parameters are computed in milliseconds squared (ms^2). Table 4.3 presents the comparison between the HRV indices estimated from the rPPG methods with the corresponding values measured from the reference PPG sensor. Further, the performance is evaluated in terms of absolute error, as shown via the box plot in Figure 4.6. The figure clearly shows the HRV indices estimated from the proposed network have the minimum absolute error deviation compared to the state-of-the-art methods. The obtained remarkable results for HRV parameters show the significance of the proposed framework for HRV analysis.

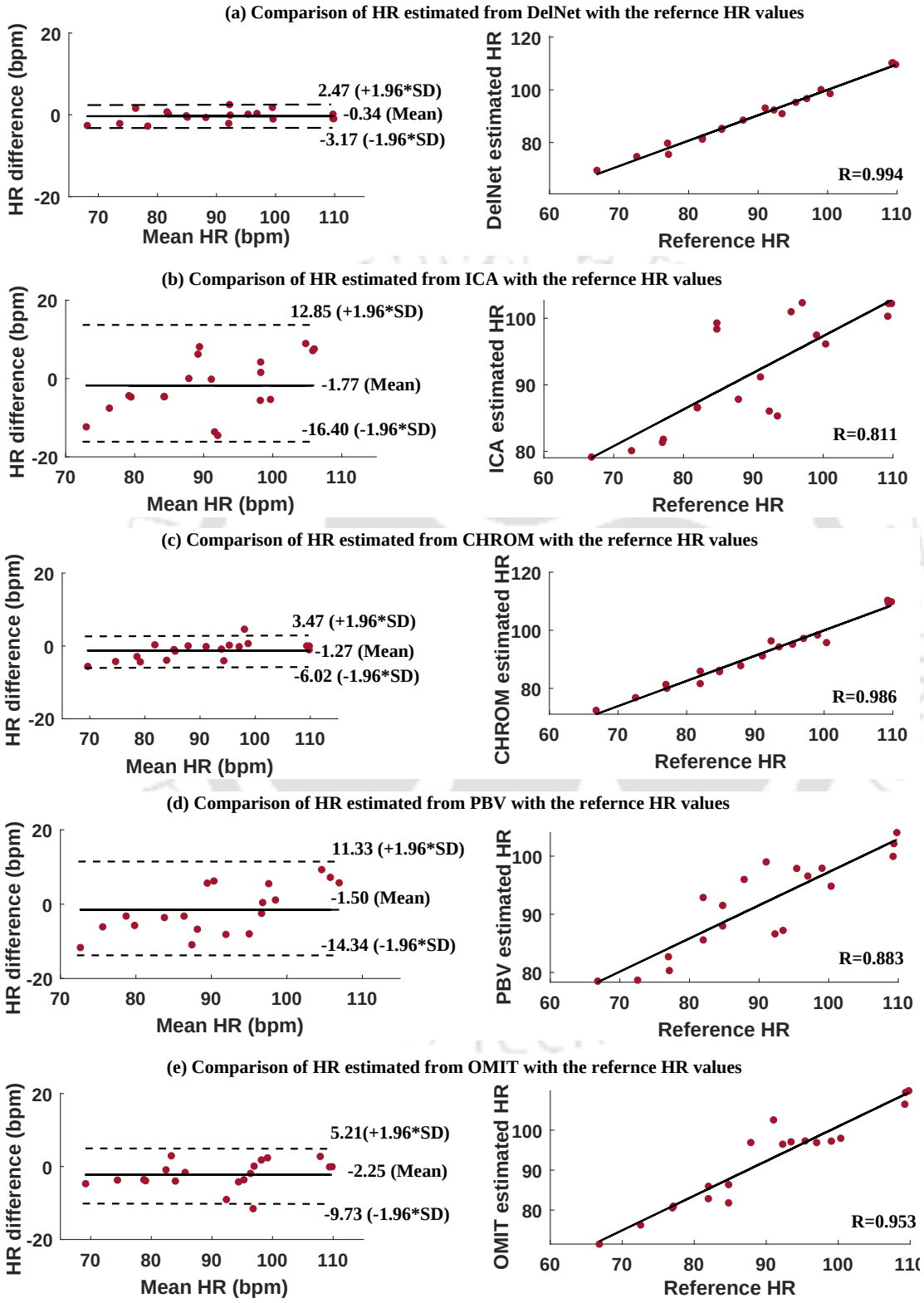


Figure 4.5: Comparison of HR values extracted from rPPG methods with the reference values (RV) using Bland-Altman plots [right] and correlation [right]: (a) Proposed-RV (b) ICA-RV (c) CHROM-RV (d) PBV-RV (e) OMIT-RV

Abstract-TH-3588_186102009

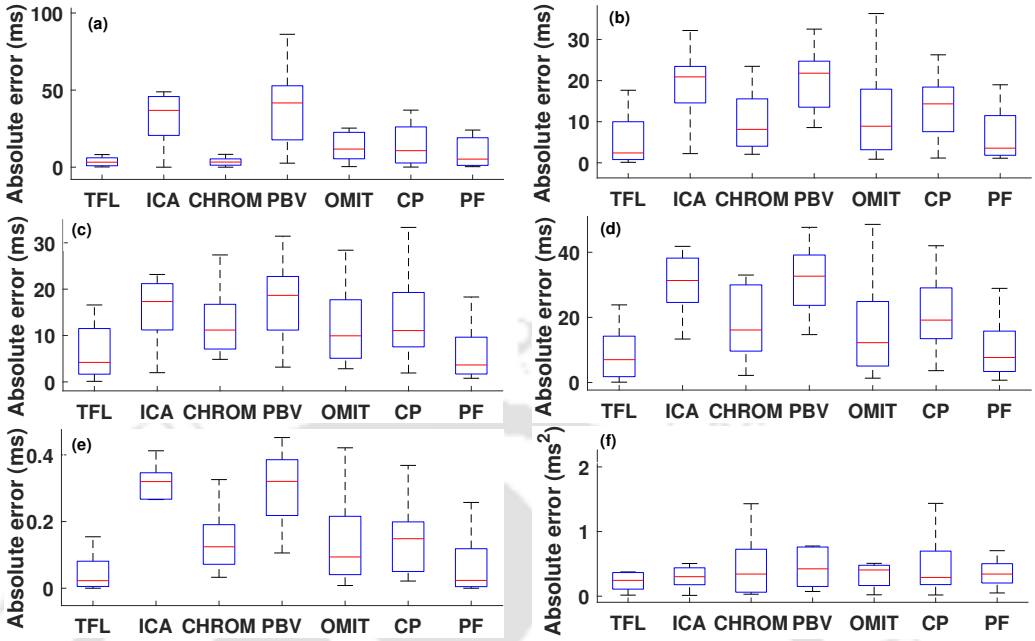


Figure 4.6: Boxplot representation of absolute errors between rPPG methods and reference PPG derived HRV parameters. (a) meanNN (b) SDNN (c) SDSD (d) RMSSD (e) PNN50 (f) LF/HF

4.1.4 Conclusion

This work presents a DNN-based DelNet architecture to estimate a reliable rPPG waveform with accurate systolic peak localization from a sequence of video frames. The raw trace obtained by averaging the skin pixels of the green channel is fed as an input to the network. The green trace constituting the maximum pulsatile information helps the proposed network in learning and extracting a reliable rPPG waveform. The proposed framework is made computationally efficient by designing it with simple 1D convolution filters. The network was trained and tested on the UBFC-rPPG dataset, and it was validated by statistically comparing the estimated average HR and HRV parameters derived from the rPPG waveform with the corresponding parameters measured from the reference PPG signal. Further, the network is compared with four state-of-the-art methods: ICA, CHROM, PBV, and OMIT. The performance evaluation reveals that the proposed framework is robust compared to the other methods. Thus, the proposed method is applicable for both HR and HRV measurements and may be used as a cardiac health assessment tool.

4.2 TFL Framework for rPPG Estimation Using Scalogram Based Feature Map

In Section 4.1, DelNet framework is proposed, where systolic peaks of rPPG are accurately delineated and utilized for HR and HRV estimation. The method effectively works on the dataset that were recorded under varying environmental conditions, such as illumination variation, rigid and no-rigid motions, inter-subject skintone variability, and facial expression. In addition, utilization of 1D-CNN in the work has made the model simple and robust in terms of computational power and memory. However, the performance of the network significantly dropped when experimented on the ECG-fitness dataset that was recorded from the subjects while performing various physical activities on fitness machines. Thus, the network is not effective for large body motions and therefore, it further requires redesigning and development that is essential to monitor cardiac health status during exercises. Also, the input to the network is only the green trace which may have effect the performance of the network. Thus, the network can be modelled to utilized the significant information of all the color channels of each frames.

To address this limitation, we have proposed a 2D-CNN based time-frequency learning (TFL) framework for the estimation of rPPG waveform. The network is fed with a novel scalogram image that is obtained from raw RGB traces using the wavelet transform. The generation of scalogram involves the enhancement of pulsatile profile by selecting a mother wavelet that is well-aligned with the desired PPG waveform. Therefore, generated feature maps have specific and well-oriented structure to be learned by the TFL network. The TFL network is designed in an effective way by combining convolution layers with powerful DL modules, such as encoder-decoder and attention blocks that helps the network to learn and characterize a wide-range of rPPG signal. Subsequently, the estimated rPPG waveform is used for HR and HRV measurements. Our approach has considered incorporating advanced signal processing tools and DL techniques. The proposed method is elaborately discussed in the subsequent sections.

4.2.1 Proposed Methodology

A scalogram based feature map is provided as an input to the TFL network that in return estimates a pulsatile waveform. A schematic block diagram of the proposed feature image extraction and TFL network architecture is illustrated in Figure 4.7.

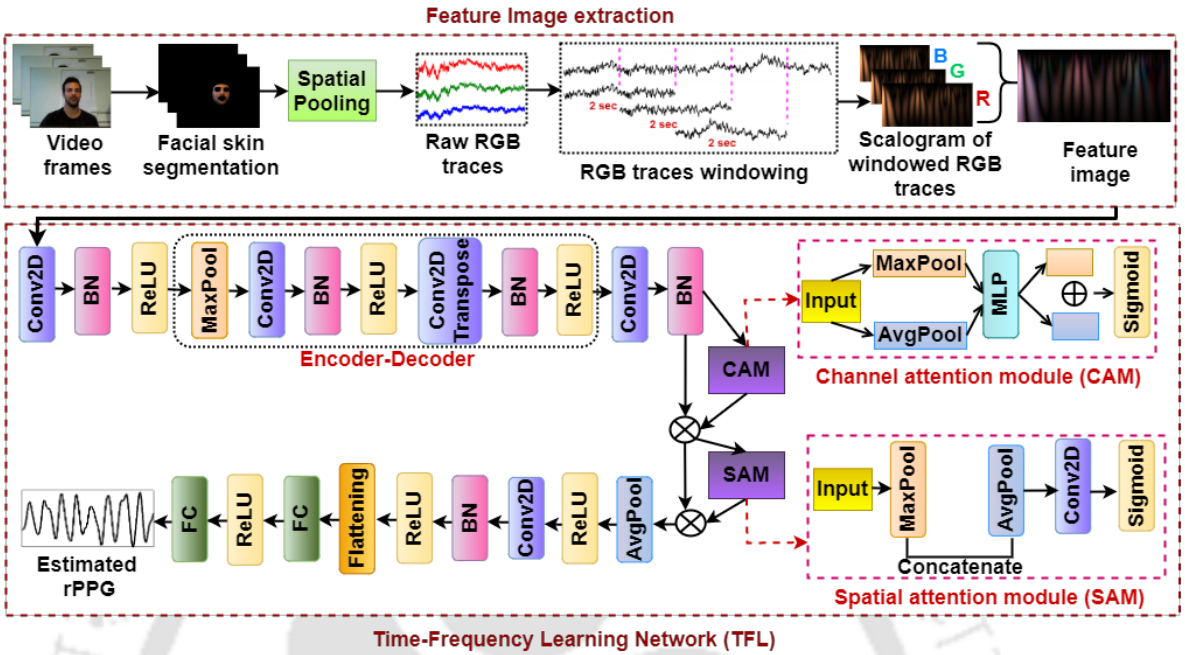


Figure 4.7: Block diagram of the proposed TFL network

4.2.1.1 Feature image extraction

The scalogram based feature map is the time-frequency representation of the raw RGB traces. The feature map is obtained by sequentially processing the input video data. Initially, 468 facial landmarks are detected and tracked throughout the video frames using the MediaPipe face mesh [171]. Subsequently, the facial skin is segmented using the convex-hull extractor. The skin region is segmented by subtracting the pixels associated with the eyes and mouth landmarks. Then, the raw RGB traces are obtained by spatially pooling the extracted skin pixels. Since, a large set of data is required by DNNs to perform effectively, a subset of each RGB-traces is created by segmenting the sequence of raw RGB traces using a 2 sec sliding window with 50% overlapping. Further, the segmented traces are interpolated at 256 Hz to increase the temporal resolution and to make the sampling rate identical for all the datasets. Usually, the time-domain traces are vulnerable to environmental interferences, thereby making their analysis complex and difficult in distinguishing different rPPG signals due to inadequate features. However, time-frequency representation assist to know how the traces are distributed with frequency and phase, thereby making the analysis simple and robust for signal characterization. Therefore, the proposed work has mapped the raw traces to their corresponding scalograms that represents the time-frequency distribution of the traces. The algorithm for scalogram processing is explained in

the following section.

Processing of Scalogram

A scalogram transforms a 1D-signal into a 2D-matrix with the help of wavelet transformation [330]. Wavelet transform (WT) is a powerful signal processing tool that decomposes a signal into time-scale plane, where each scale represents a certain frequency range [330]. The decomposition is achieved by means of scaling (a) and shifting (b) parameters of a mother wavelet function $\psi(t)$. In the proposed feature map extraction scheme, WT of the raw RGB traces is computed as [330]:

$$w_{a,b}(m_k) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} m_k(t) \psi^* \left(\frac{t-b}{a} \right) dt, \quad k \in [1,2,3] \tag{4.2}$$

where m_k represents the raw RGB traces, $w_{a,b}$ is the b^{th} coefficient at the a^{th} scale, and the complex conjugate operator is denoted by $(*)$. One of the advantageous features of this transformation is that it can eliminate a specific frequency and amplitude from the decomposed signal using an appropriate mother wavelet. Thus, selecting a mother wavelet is crucial and important for a specific task. Since the objective of the proposed work is to estimate a reliable rPPG signal from the input feature map, the feature map generation scheme aims to generate a feature map that captures the pulsatile profile of a PPG signal. Therefore, the proposed scheme has considered second derivative of a Gaussian function *(Gauss-2)* as the mother wavelet due to its morphological similarity with the pulsatile profile of a PPG signal [104]. The *(Gauss-2)* can be defined as [104]:

$$\psi(t) = -\frac{d^2 g(t)}{dt^2} = \lambda(1-t^2)e^{-t^2/2} \tag{4.3}$$

where $g(t)$ and λ denote a Gaussain function and normalization factor, respectively. Subsequently, scalogram of each RGB traces is measured separately as the absolute value of the wavelet coefficients. Finally, the normalized scalogram using each RGB scalogram is obtained as:

$$S_{a,b} = \frac{stack[|w_{a,b}(m_k)|]}{max(|w_{a,b}(m_k)|)} \tag{4.4}$$

The size of the obtained scalogram is 30×512 $(H \times W)$, where $H = a$ and $W = L$ $(t \times fs)$, where L is the length of raw trace, t (2 sec) is the trace duration, and fs (256 Hz) is the sampling rate of the trace.

4.2.1.2 TFL network architecture

Our proposed TFL network architecture is shown in Figure 4.7. Since the network learns the time-frequency representation of the video data, it is named as time-frequency learning (TFL) network. The size of the input scalogram is $128 \times 3 \times 30 \times 512$ ($B \times C \times H \times W$), where B is the batch size and C is the number of channels. The TFL architecture includes different blocks comprising convolution layers, encoder-decoder block that is used to reduce the noise in the input data, channel and spatial attention modules to enhance the significant pulsatile profile embedded within the feature map, and two fully-connected (FC) layers to regress the rPPG signal. In the framework, each convolution block is followed by batch normalization and ReLU activation function. Also, an additional dropout operation of 50% is introduced before the final layer to reduce the overfitting effects. The final output of the network is a 1D rPPG signal with a size of 128×512 ($B \times W$). The details of the TFL network architecture is illustrated in Table 4.4. In addition to the design of the TFL network architecture, an appropriate loss function is also required to guide the network for the estimation of a reliable rPPG waveform. Here, the network is trained with the ground-truth PPG signals that were recorded simultaneously with the video data. In order to recover the rPPG signal with accurate systolic peak instants, the negative Pearson correlation coefficient (L_ρ) is utilized as a loss function. Further, (L_ρ) is combined with L1 loss function (L_1) to reduce the difference between the ground-truth PPG and the estimated rPPG signal values. The (L_ρ), (L_1), and the combined loss function are expressed by Eq. 5.1.2.2, 5.1.2.2, and 5.2, respectively.

$$L_\rho(x,y) = 1 - \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}} \tag{4.5}$$

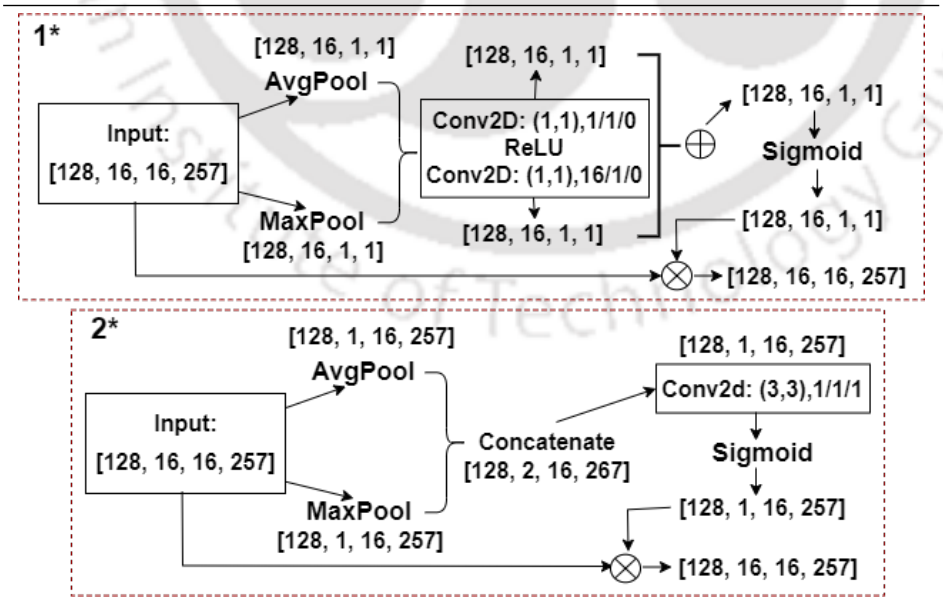
$$L_1(x,y) = \sum_{i=1}^{n} |x_i - y_i| \tag{4.6}$$

$$Loss = \alpha * L_\rho + (1 - \alpha) * L_1 \tag{4.7}$$

where x represents the ground-truth PPG signal values, y denotes the estimated rPPG signal values, n is the length of signal, and the optimal value of α is found to be 0.9.

Table 4.4: TFL network architecture

	Layer	K, C / S / P	Output shape
1.	Conv2d	(5×5), 6 / 1 / 2	[128, 6, 30, 512]
2.	BatchNorm2d	-	[128, 6, 30, 512]
3.	ReLU	-	[128, 6, 30, 512]
4.	MaxPool2d	(2×2) / 2 / 0	[128, 6, 15, 256]
5.	Conv2d	(3×3), 6 / 1 / 1	[128, 6, 15, 256]
6.	BatchNorm2d	-	[128, 6, 15, 256]
7.	ReLU	-	[128, 6, 15, 256]
8.	ConvTranspose2d	(2×2), 6 / 1 / -	[128, 6, 16, 257]
9.	BatchNorm2d	-	[128, 6, 16, 257]
10.	ReLU	-	[128, 6, 16, 257]
11.	Conv2d	(3×3), 16 / 1 / 1	[128, 16, 16, 257]
12.	BatchNorm2d	-	[128, 16, 16, 257]
13.	ChannelAttention	1*	[128, 16, 16, 257]
14.	SpatialAttention	2*	[128, 16, 16, 257]
15.	ReLU	-	[128, 16, 16, 257]
16.	MaxPool2d	(2×2) / 2 / 0	[128, 16, 8, 128]
17.	Conv2d	(3×3), 16 / 1 / 1	[128, 1, 8, 128]
18.	BatchNorm2d	-	[128, 1, 8, 128]
19.	ReLU	-	[128, 1, 8, 128]
20.	Flatten	-	[128,1024]
21.	FC1	-	[128, 256]
22.	ReLU	-	[128, 256]
23.	Dropout	0.5	[128, 256]
24.	FC2	-	[128, 512]



Note: K, S and P represents kernel size, stride and padding, respectively

4.2.2 Materials and evaluation

4.2.2.1 Datasets

The experiments were conducted on three publicly available datasets: UBFC_rPPG, COHFACE, and ECG-Fitness. All the datasets were recorded under realistic environmental conditions and consist of subjects with varying demographics, such as age, gender and skin tone. For the experimentation, the training data consists of 70% of the total subjects, and the remaining subjects are used for testing.

UBFC_rPPG

Details of this dataset can be found in Section 4.1.3.1. Here, the recorded videos and their corresponding PPG signals were sliced using a 2 sec sliding window with 50% overlapping. Further, the segments were interpolated from 30 Hz to 256 Hz. The resultant training and testing dataset contains 1840 and 782 segmented data, respectively.

COHFACE

The COHFACE dataset is described in Section 3.2.1. The dataset contains 160 videos of 40 subjects, each one minute long. The training set consists a total of 3248 data segments created from 28 subjects. Meanwhile, the testing set includes 1392 data segments that were created from the remaining 12 subjects.

ECG-Fitness

The ECG-fitness dataset was recorded from 17 subjects (14 male, 3 female) while performing different physical activities, such as speaking, rowing, exercising on a stationary bike and on an elliptical trainer. Twelve videos were collected from each subject during six different experimental states, each one min long. For each experimental states two videos were recorded at 30 fps with 1920×1080 pixels using two Logitech camera, where one camera was attached to the fitness machine and the other was positioned on a tripod. The six experimental states are (i) rowing and (ii) speaking under ambient lighting condition, (iii) rowing and (iv) speaking under halogen lighting condition, (v) exercising on elliptical trainer and (vi) exercising on a stationary bike under ambient light. Also, ECG signals were recorded simultaneously with the videos using two-lead Viatom CheckMeTMPro device with CC5 lead. Since the proposed network needs to be trained with ground-truth PPG signal, synthetic PPG signals are generated from the available ECG data using Algorithm 2. Figure 4.8

Algorithm 2 Synthetic PPG generation

Input: ECG signal $x[n]$; $n = 1, 2, ..., N$

1. **ECG signal normalization:** $x^{'}[n] = \dfrac{x[n] - min(x[n])}{max(x[n]) - min(x[n])}$
2. **R- peak detection using amplitude-temporal thresholding scheme:**
   $R_t = [R_1, R_2, \cdots, R_T] \quad (t = 1, 2, ...T)$
   //where R_t represents the timing information of t^{th} R-peak
3. **Derive RR tachogram ($\boldsymbol{RR_t}$) from detected R-peak instants:**
   $RR_t = [RR_1, RR_2, \cdots, RR_{T-1}]$
   //where $RR_t = R_{t+1} - R_t$
4. **PPG signal generation from RR tachogram:**
   Construct a vector $\boldsymbol{SP}$ using RR tachogram:
   $SP = [RR_1/2, RR_1, \cdots, RR_{T-1}/2, RR_{T-1}]$
   Final PPG signal is obtained by interpolating SP at a sampling rate of 256 Hz

Output: Synthetic PPG signal $y[n]$; $n = 1, 2, ..., N$

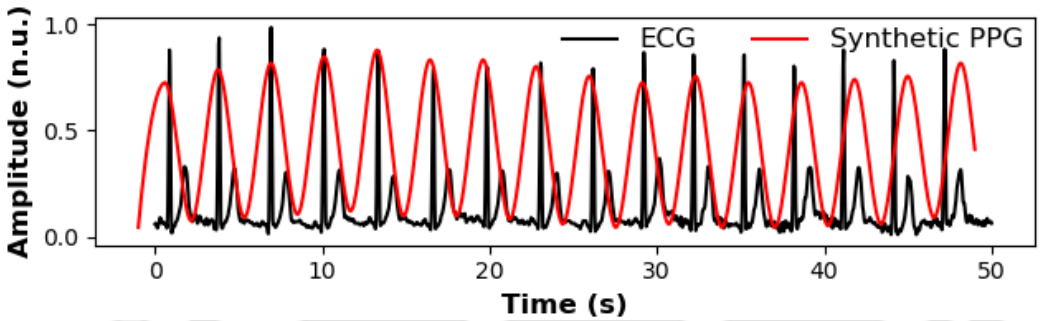


Figure 4.8: ECG signal along with the generated synthetic PPG waveform

shows a small segment of the synthetic PPG signal that is generated from its corresponding ECG signal. Finally, the dataset is available with videos and their corresponding ground-truth synthetic PPG signals for the experimentation purpose. A total of 3960 and 1800 data segments were created for the training and testing sets, respectively.

4.2.2.2 Experimental Settings

The weights of the network parameters were optimized using the Adam optimizer with a batch size of 128. The network was trained separately for each dataset with a learning rate of 0.025. The number of epochs was set to 100 for UBFC_rPPG and COHFACE dataset, whereas ECG fitness dataset was trained for 150 epochs. Figure 4.9 shows the loss curve of each dataset for both training and testing phases. The decrease in the loss value shows the effectiveness of the selected hyperparameter values for the network operation. The network components were implemented using PyTorch and trained on the NVIDIA Tesla P100-PCIE-16GB GPU platform.

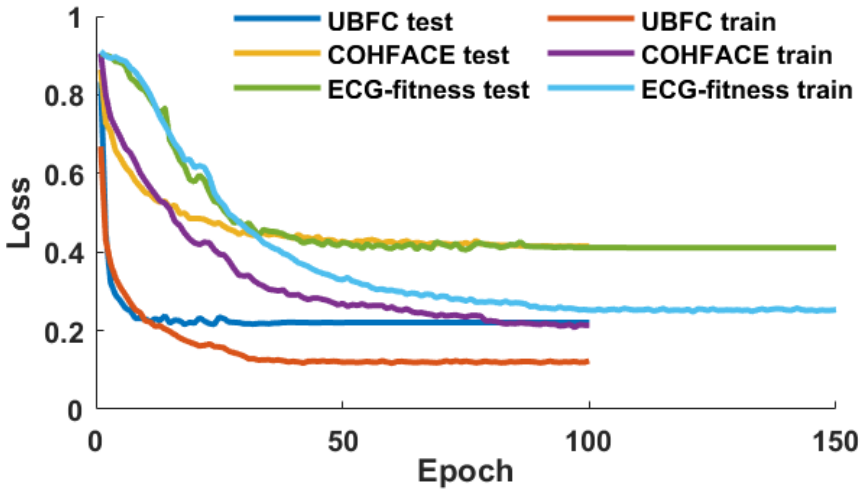


Figure 4.9: Training and testing loss curve for each dataset

4.2.2.3 Evaluation metrics

The effectiveness of the proposed network is shown with HR and HRV measurements using the estimated rPPG waveform. Since the network is designed to estimate an rPPG waveform of 2 sec duration, the output segments corresponding to the test data of a particular individual are concatenated before cardiac parameters measurement. The concatenated signals' duration is equivalent to the length of the recorded videos, which are about 1 min long. Here, the systolic peak instants of the estimated rPPG waveform are detected and provided to the HeartPy toolbox for HR and HRV parameter estimation. The estimated cardiac parameters are evaluated using statistical and graphical methods. The statistical metrics used for HR evaluation are MAE, RMSE, SDE, and R. Additionally, Bland-Altman (BA) and box-plot are used to graphically quantify the agreement between the ground-truth and the estimated measurements. Further the cardiac measurements are compared with state-of-the-art methods: ICA, CHROM, PBV, OMIT, Contrast-Phys (CP), and PhysFormer (PF) [178, 179, 189, 190, 220, 224]. The state-of-the-art methods [178, 179, 189, 190] are implemented using the publicly available open toolbox iPhys [331]. Whereas, [220] and [224] are trained from scratch and tested on each dataset separately.

4.2.2.4 Network comparison using UBFC_rPPG dataset

Averaged heart rate (HR): We first conducted our experiments on the UBFC_rPPG dataset. Figure 4.10 shows the rPPG waveform estimated from the proposed network along with the ground-

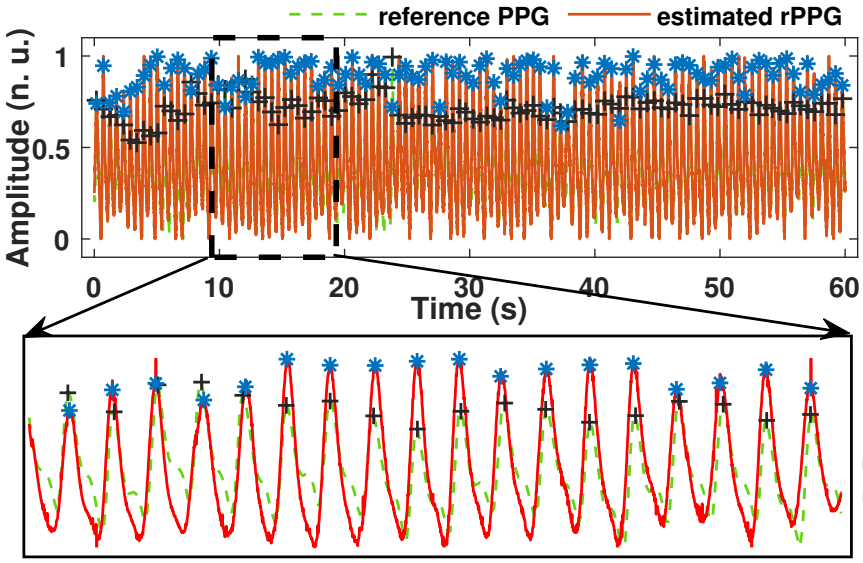


Figure 4.10: Ground-truth PPG signal (green) and the rPPG waveform (red) estimated from the proposed network

Table 4.5: Statistical comparison between averaged HR values estimated from the rPPG methods and the reference PPG sensor using UBFC_rPPG dataset

Methods	MAE	RMSE	SDE	R	Mean	LOA [L, U]
ICA	6.192	7.560	4.531	0.642	-0.994	[-16.336, 14.348]
CHROM	0.898	1.692	1.498	0.985	0.361	[-3.023, 3.746]
PBV	5.820	6.679	3.421	0.819	-0.645	[-14.254, 12.964]
OMIT	3.003	4.583	3.616	0.889	-1.062	[-10.189, 8.066]
CP	2.012	2.598	1.716	0.963	0.078	[-5.238, 5.394]
PF	1.514	2.243	1.728	0.987	-1.342	[-5.021, 2.337]
Proposed	0.569	0.774	0.549	0.997	0.019	[-1.565, 1.604]

truth PPG signal. Table 4.5 presents the comparison results between the averaged HR values measured from the estimated rPPG waveforms and the ground-truth PPG. The table clearly shows the effectiveness of the proposed network with the reduced error values. The MAE, RMSE and SDE are found to be 0.569, 0.774 and 0.549, respectively. Further, the comparison is shown graphically using BA and regression plots. Figure 4.11 shows the BA analysis to quantify the agreement between the estimated and the reference HR values. The mean and LOA values for each method are listed in Table 4.5. The BA plots reveal the highest agreement of the proposed network with the ground-truth PPG sensor with a minimum mean of 0.019 and LOA around 1.6. Additionally, the performance comparison is shown in terms of regression plot as shown in Figure 4.12. The R values for ICA, CHROM, PBV,

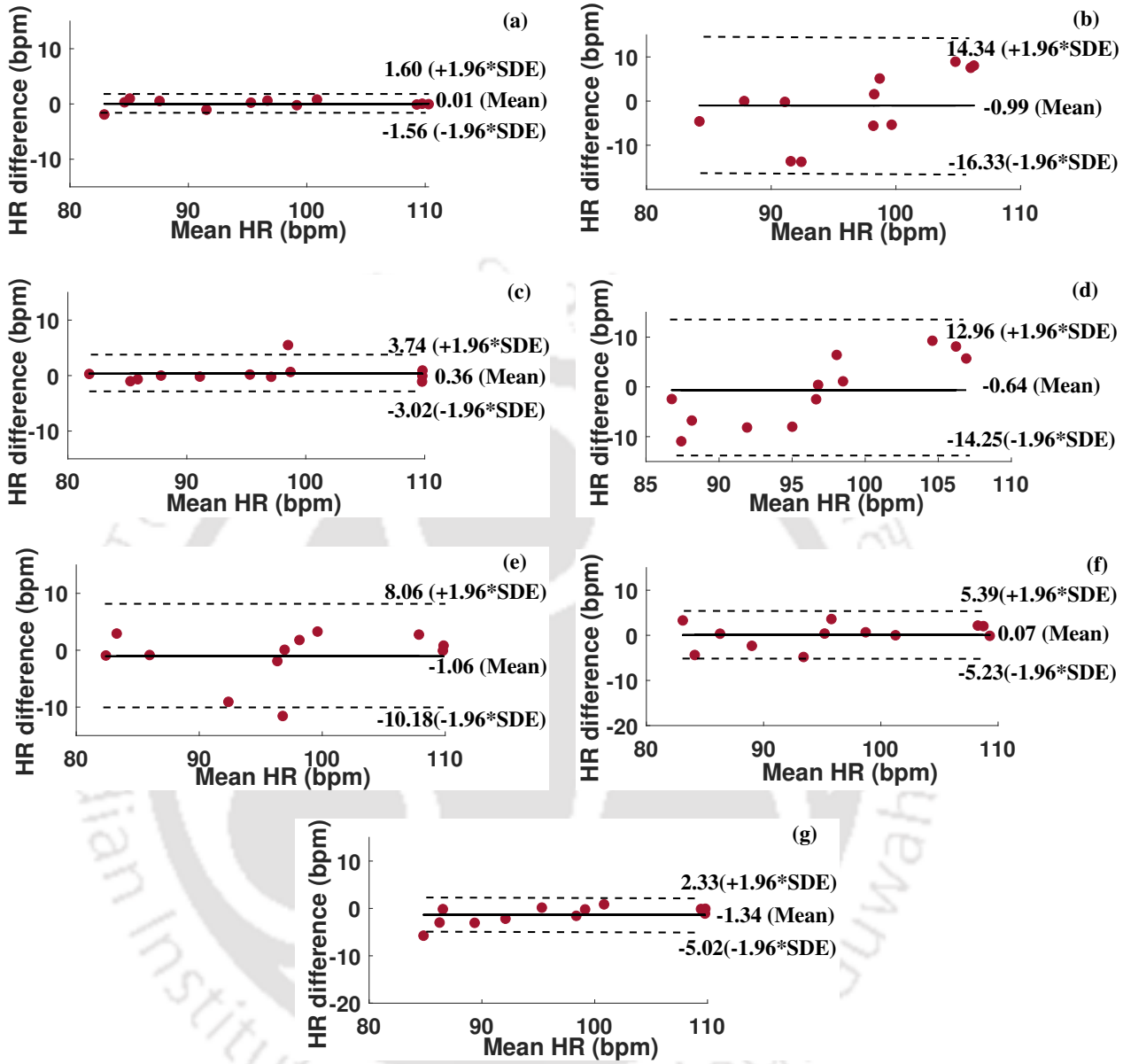


Figure 4.11: Bland-Altman plot to compare the performance of the camera-based methods with ground-truth (GT) using UBFC_rPPG dataset. (a) Proposed-GT (b) ICA-GT (c) CHROM-GT (d) PBV-GT (e) OMIT-GT (f) CP-GT (g) PF-GT

OMIT, CP, PF and the proposed network are found to be 0.642, 0.985, 0.819, 0.889, 0.963, 0.987, and 0.997, respectively. Thus, these results show the robustness of the proposed network in determining the HR values accurately.

Heart rate variability analysis (HRV): After the validation of HR, the HRV indices estimated from the rPPG signal are compared with the one derived from the ground-truth PPG signal. The performance of the network is validated by considering both time-domain and frequency domain HRV

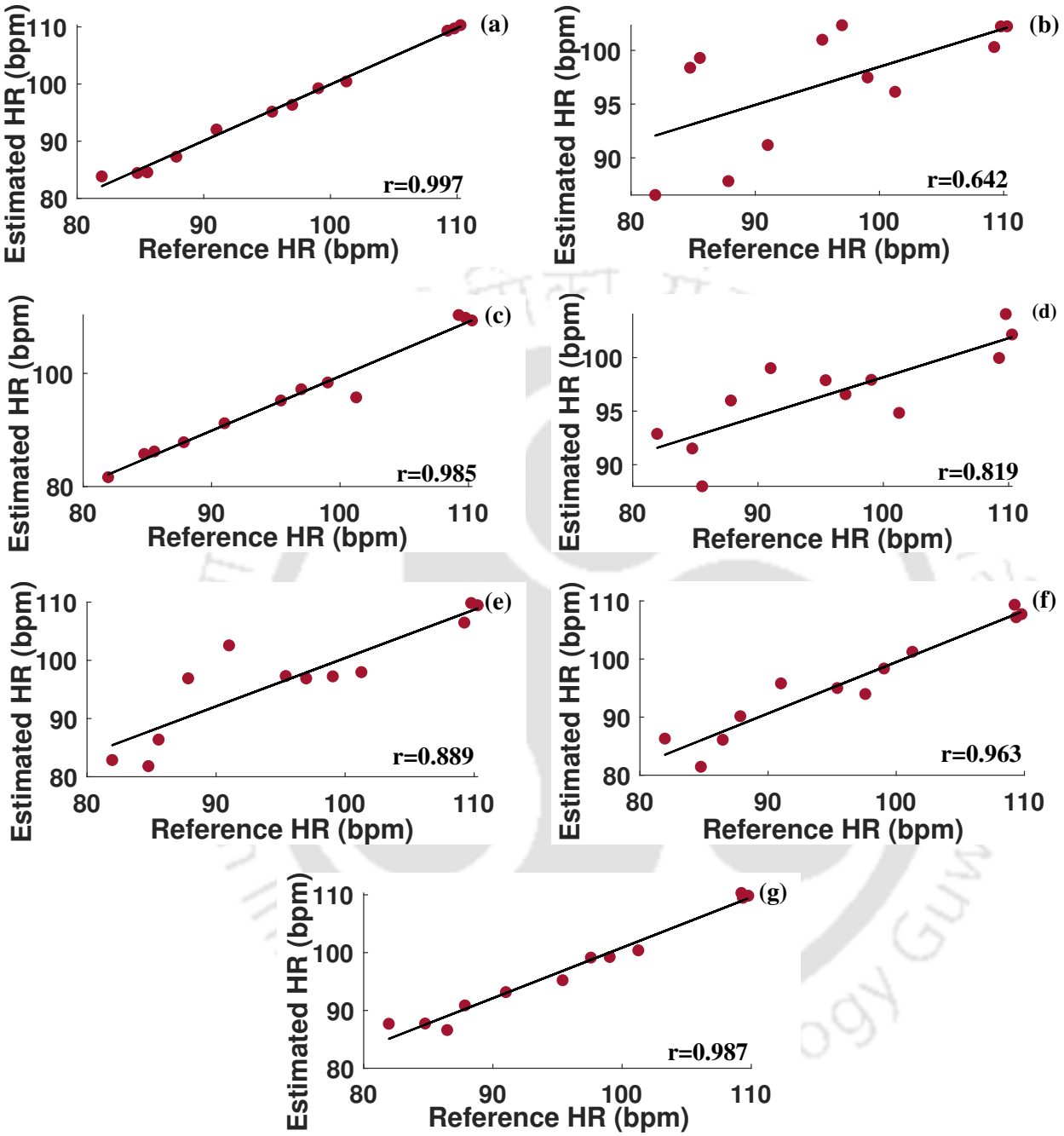


Figure 4.12: Regression plot to compare the performance of the camera-based methods with ground-truth (GT) using UBFC_rPPG dataset. (a) Proposed-GT (b) ICA-GT (c) CHROM-GT (d) PBV-GT (e) OMIT-GT (f) CP-GT (g) PF-GT

parameters. The studied time-domain parameters are – (i) meanNN, (ii) SDNN, (iii) SDSD, (iv) root RMSSD, and (v) PNN50, all are measured in millisecond (ms). The frequency domain parameters are – (i)LF, (ii) HF, and (iii) LF/HF, each calculated in milliseconds squared (ms^2). The studied HRV parameters measured from the reference PPG sensor, state-of-the-art methods, and the proposed

network are presented in Table 4.6. Further, the comparison is evaluated via box plot that represents the absolute error of the measured HRV indices. Figure 4.13 shows the box plot for meanNN, SDNN, SDSD, RMSSD, PNN50, and LF/HF. It is observed that the absolute error is very low for all the parameters that are estimated from the proposed network, indicating the significance of the network for HRV analysis.

Table 4.6: Time and frequency domain HRV parameters estimated from the rPPG methods and the reference PPG sensor for UBFC_rPPG dataset

	Parameters	Reference	Methods						
			ICA	CHROM	PBV	OMIT	CP	PF	Proposed
Time domain parameters (ms)	**meanNN**	630.819	620.033	632.848	621.531	623.244	629.850	620.042	630.763
	SDNN	14.023	32.940	23.640	34.417	25.630	27.559	20.089	18.693
	SDSD	12.540	27.926	23.394	30.071	22.899	26.730	17.388	17.265
	RMSSD	17.432	46.560	35.440	48.738	34.378	39.073	27.385	23.569
	PNN50	0.017	0.306	0.160	0.316	0.153	0.161	0.087	0.059
Frequency domain parameters (ms^2)	**LF**	0.419	1.584	0.722	1.943	1.161	1.252	0.502	0.739
	HF	0.701	3.065	1.236	4.424	2.148	2.734	0.975	1.168
	LF/HF	0.868	0.494	0.461	0.437	0.571	0.577	0.691	0.807

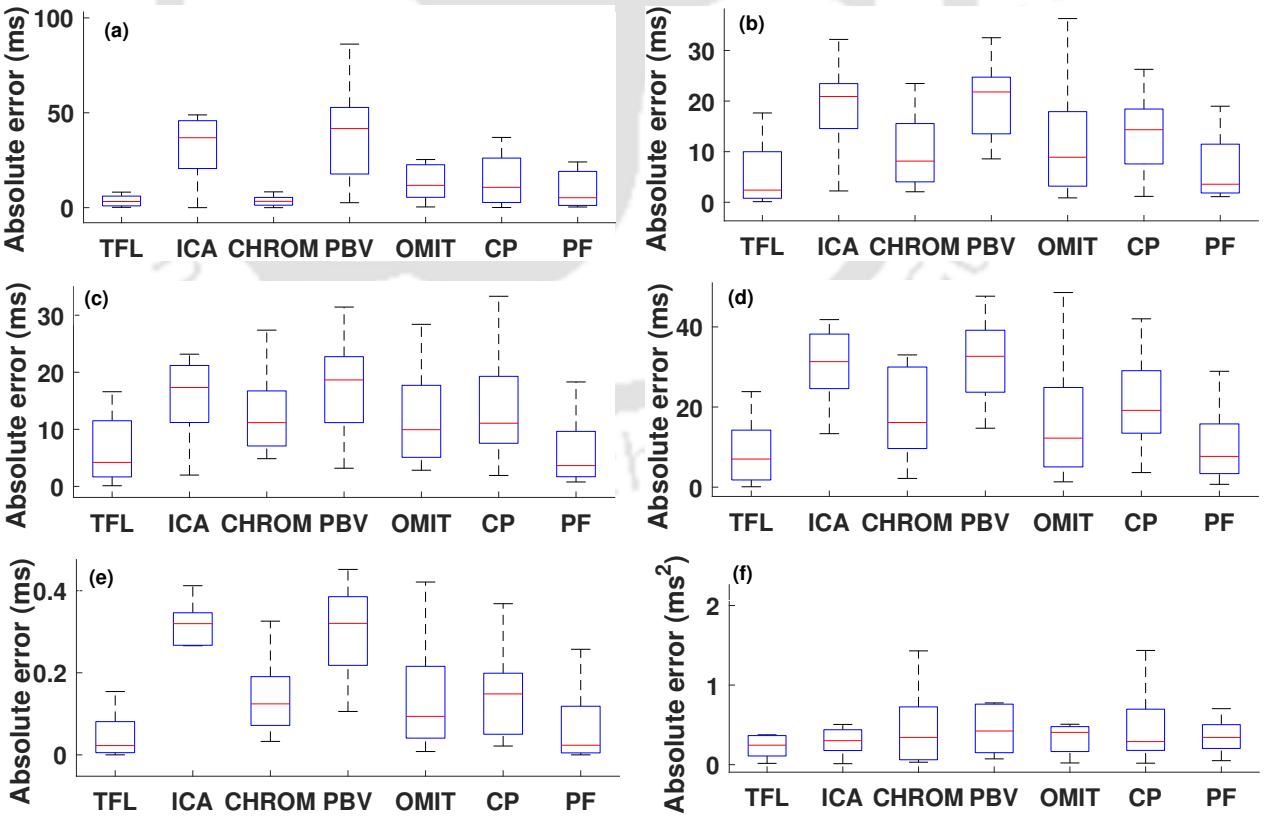


Figure 4.13: BOX plot of absolute error for UBFC_rPPG dataset. (a) meanNN (b) SDNN (c) SDSD (d) RMSSD (e) PNN50 (f) LF/HF

4.2.2.5 Network comparison using COHFACE and ECG-ftiness dataset

Similarly, we carried out the experiments on the widely used COHFACE dataset and a more challenging ECG-fitness dataset. Table 4.7 and 4.8 presents the experimental results of the HR values for COHFACE and ECG-fitness data, respectively. Further, Figure 4.14 and 4.15 presents the BA plots to assess the performance of the network on both the datasets. In addition to HR validation, HRV analysis is also shown in Table 4.9 for both COHFACE and ECG-fitness datasets. The tables show that the error values are high for COHFACE and ECG-fitness data compared to the UBFC_rPPG dataset. This increase in error may be caused due to high compression of videos and large body motion in COHFACE and ECG-fitness data, respectively. However, the computed experimental results show the feasibility and significance of the proposed network in estimating a reliable rPPG signal.

Although the current study has achieved insignificant error on the datasets with dynamic environmental conditions, the proposed work still needs investigation in other aspects to deploy it for practical usage. The proposed model needs training on a large dataset of random body movements to make it robust against large motion artifacts. It could be further tested and validated on a dataset capturing various body postures, such as standing, left and right lateral sitting, and supine, to enhance its performance. Additionally, the recorded videos in each dataset are about 1 min long, limiting the evaluation of the model's performance on long-term data for both HR and HRV analysis. Thus, an investigation is required on longer-duration data to make it applicable for long-term cardiac health monitoring.

Table 4.7: Statistical comparison between averaged HR values estimated from the rPPG methods and the reference PPG sensor using COHFACE dataset

Methods	MAE	RMSE	SDE	ρ	Mean	LOA [L, U]
ICA	7.612	9.647	6.002	-0.147	-4.931	[-21.390, 11.526]
CHROM	9.701	10.798	4.802	0.669	-9.701	[-19.114, -0.289]
PBV	8.218	10.846	7.167	-0.034	-7.193	[-23.305, 8.918]
OMIT	6.942	8.967	5.747	0.015	-4.879	[-19.812, 10.052]
CP	19.739	22.334	10.587	0.006	-5.569	[-48.530, 37.392]
PF	6.334	8.852	6.267	0.089	-3.120	[-19.576, 13.334]
Proposed	0.987	1.982	1.740	0.958	-0.667	[-4.371, 3.037]

Table 4.8: Statistical comparison between averaged HR values estimated from the rPPG methods and the reference PPG sensor using ECG fitness dataset

Methods	MAE	RMSE	SDE	ρ	Mean	LOA [L, U]
ICA	22.734	26.820	14.356	-0.277	19.522	[-16.843, 55.887]
CHROM	19.037	23.840	14.478	0.102	17.043	[-15.918, 50.005]
PBV	24.530	28.352	14.342	-0.218	21.375	[-15.458, 58.207]
OMIT	19.214	23.646	13.904	-0.101	15.801	[-18.984, 50.586]
CP	23.720	28.750	16.392	0.183	13.557	[-36.584, 63.699]
PF	13.508	16.921	10.283	0.227	3.788	[-28.828, 36.405]
Proposed	4.267	5.439	4.267	0.937	1.243	[-9.227, 11.713]

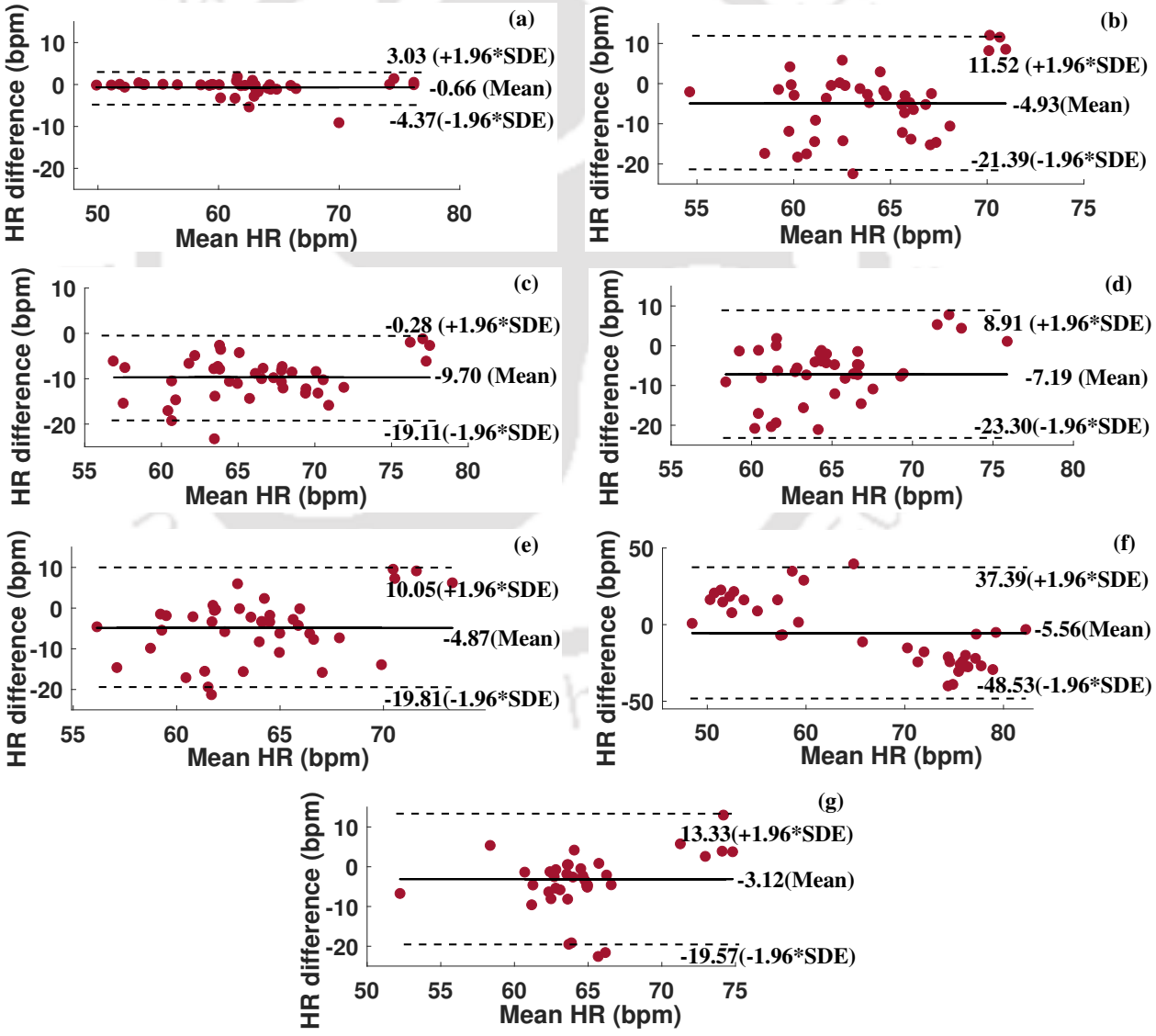


Figure 4.14: Bland-Altman plot to compare the performance of the camera-based methods with ground-truth (GT) using COHFACE dataset. (a) Proposed-GT (b) ICA-GT (c) CHROM-GT (d) PBV-GT (e) OMIT-GT (f) CP-GT (g) PF-GT

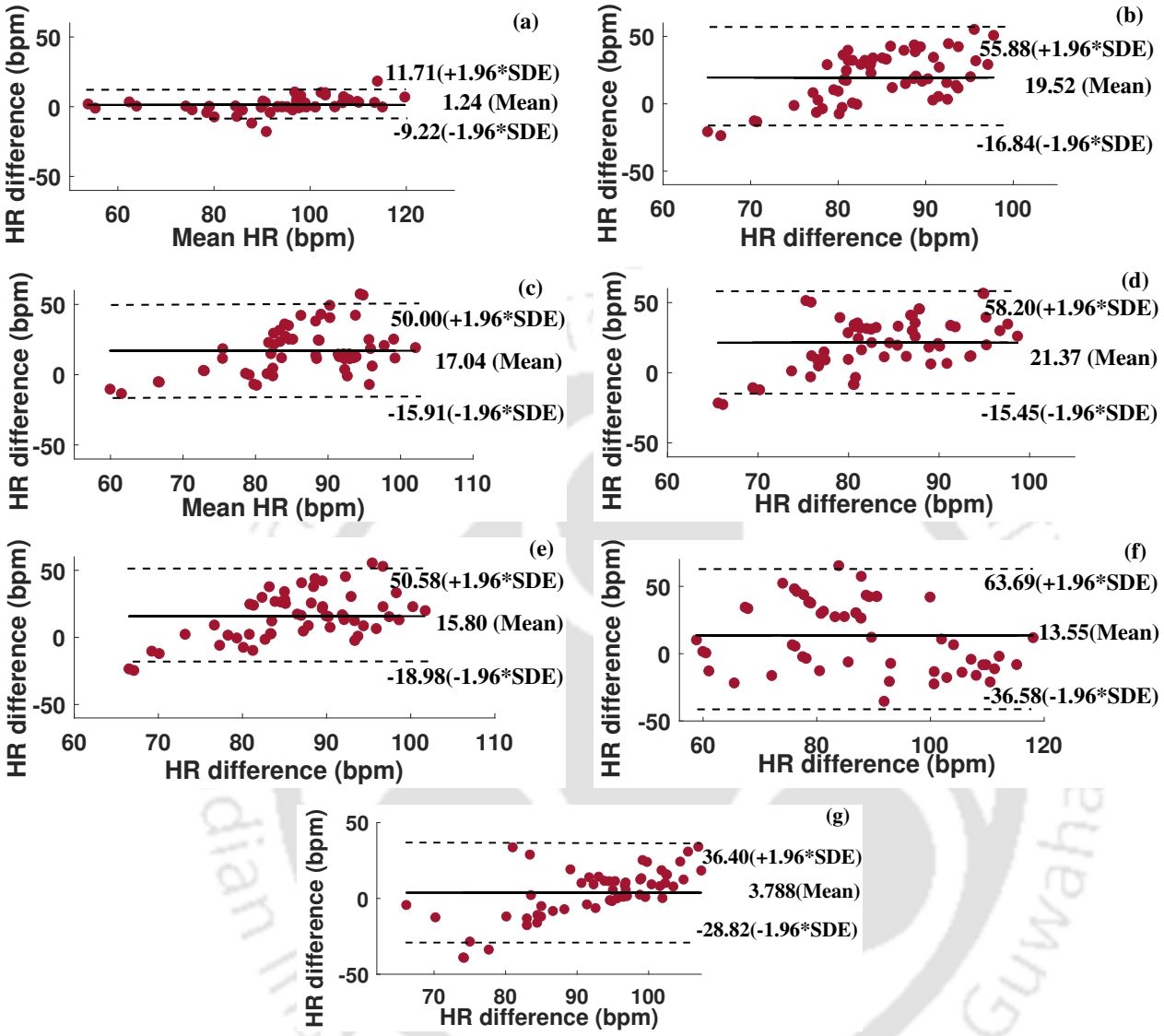


Figure 4.15: Bland-Altman plot to compare the performance of the camera-based methods with ground-truth (GT) using ECG fitness dataset. (a) Proposed-GT (b) ICA-GT (c) CHROM-GT (d) PBV-GT (e) OMIT-GT (f) CP-GT (g) PF-GT

4.2.2.6 Computational results

We conducted a further study about the computational complexity of the proposed network and compared it with state-of-the-art methods. The proposed TFL framework is measured for its computational requirements in FLOPs, number of parameters, memory consumption, and inference time. To calculate the run time and the memory usage, the proposed method is divided into two parts. The first part includes the processing and extraction of feature map. The second part is the estimation of rPPG waveform from the TFL network. The memory usage and run-time for part 1 and part 2

Table 4.9: Time and frequency domain HRV parameters estimated from the rPPG methods and the reference PPG sensor for COHFACE and ECG fitness dataset

Dataset ↓		Parameters	Reference	Methods						
				ICA	CHROM	PBV	OMIT	CP	PF	Proposed
COHFACE	Time domain parameters (ms)	**meanNN**	986.804	908.428	847.640	878.532	908.417	961.999	906.431	976.977
		SDNN	18.833	60.145	56.625	56.850	53.312	66.171	46.834	35.512
		SDSD	11.468	51.548	45.940	47.735	44.605	61.669	42.026	31.095
		RMSSD	18.139	84.737	76.219	78.584	71.914	90.121	67.278	47.873
		PNN50	0.019	0.569	0.530	0.536	0.474	0.450	0.423	0.266
	Frequency domain parameters (ms^2)	**LF**	3.196	12.363	8.531	6.777	15.234	23.343	7.278	6.728
		HF	5.724	20.458	37.560	16.389	15.253	28.343	16.515	18.682
		LF/HF	3.032	0.415	0.419	0.465	0.451	0.764	0.408	0.677
ECG fitness	Time domain parameters (ms)	**meanNN**	653.799	807.202	784.443	831.450	769.505	802.185	668.369	659.244
		SDNN	8.127	55.333	52.734	57.463	52.358	51.523	45.311	22.506
		SDSD	8.493	46.160	45.373	46.132	44.817	47.333	38.028	22.237
		RMSSD	12.895	68.251	72.313	74.932	69.730	64.452	61.197	26.611
		PNN50	0.028	0.460	0.453	0.477	0.415	0.354	0.384	0.193
	Frequency domain parameters (ms^2)	**LF**	0.141	8.200	4.945	20.233	18.646	3.117	3.274	0.892
		HF	0.610	20.519	10.525	11.022	45.014	3.637	7.305	2.115
		LF/HF	0.128	0.486	0.369	0.578	0.566	0.396	0.445	0.338

are (1361 MB, 824 s) and (0.0055 MB, 0.06 s), respectively. The detailed computational comparison of the TFL framework with the other models are listed in Table 4.10. Since the proposed method initially involves the extraction of a feature map, it requires an additional memory and time. However, if we only compare the TFL framework with the other DL techniques [220] and [224], it is found that the proposed framework has lower computational complexity than the others. Thus, the proposed approach is fast and efficient for real-time cardiac parameter estimation.

Table 4.10: Computational comparison of TFL and state-of-the-art methods

Method	Parameter	Flops	Size in memory	Inference
ICA	-	-	1371 MB	844 s
CHROM	-	-	1377 MB	943 s
PBV	-	-	1374 MB	918 s
OMIT	-	-	1375 MB	996 s
CP	858497	256.69G	0.142 MB	125.37 s
PF	7412097	47.01M	4.56 MB	33.02 s
Proposed	396065	13.80M	1361 MB	824.06 s

4.2.3 Conclusion

In this work, we present a time-frequency learning network that learns the time-frequency representation of the video data to estimate a reliable rPPG signal. The time-frequency representation is obtained from the raw RGB traces with the help of wavelet transformation. The scalogram captures

the pulsatile profile of the raw traces that helps the network to recover a reliable rPPG signal. The performance of the network was validated on three standard datasets. Additionally, the proposed network was compared with state-of-the-art methods. The experimental results clearly show that the proposed method is able to produce similar HR and HRV values that are measured from the ground-truth PPG signals. Thus, the proposed network is effective for both HR and HRV analysis using a low-cost camera that may be utilized for remote cardiac health assessment in both clinical and non-clinical setups.

4.3 Summary

In this chapter, two deep neural network (DNN) based frameworks are presented for estimating reliable rPPG waveforms. The first approach employs a DelNet network, which utilizes 1D-CNN to process the raw green trace of video data and produce an rPPG signal. Whereas second approach uses a 2D-CNN based TFL network to obtain the rPPG waveform. The input to this network is a scalogram-based feature map that is derived using the raw RGB traces with the help of wavelet transformation. The ability of the rPPG waveforms obtained from the frameworks is shown to estimate HR and HRV values. The qualitative and quantitative analysis of the results show robust performances for our proposed DNN frameworks. However, both frameworks rely on robust handcrafted features for generating reliable rPPG waveforms, and their performance is dependent upon the quality of these input features. Additionally, the process of feature generation involves extensive preprocessing steps, leading to increased computational costs. To overcome these limitations, the next chapter presents an end-to-end network based on 3D-CNN, which directly learns informative features from raw input video data. This network is designed for multitasking, generating both rECG waveform and rPPG signal.

5

Camera-Based Remote ECG Signal Generation: A Pilot Study

Contents

So far, the research on rPPG has reached a milestone in capturing the blood volume variation that is associated with the cardiac mechanical phenomenon. The mechanical activity constitutes the basic fundamentals of rPPG technology and have significantly contributed in vitals estimation, such as HR, HRV, BP, SpO2, cardiac output, and RR. Apart from the cardiac mechanical phenomenon, the electro-cardio information provides more insight of the events, such as atrial depolarization, ventricular depolarization and repolarization, and papillary muscle repolarization. Electrocardiography (ECG) is the gold standard that is clinically used to record the heart's electric activity and helps clinical experts to diagnose multiple life-threatening diseases, such as atrial fibrillation, ventricular tachycardia, myocardial infarction, atrioventricular (AV) blocks, and bundle branch blocks [312]. Thus, ECG monitoring facilitates the early detection of cardiovascular diseases and has proven to increase the survival rate in patients with cardiac abnormalities. However, the existing ECG-based wearables require direct contact with the user that may cause discomfort and restrict the user's mobility, making them infeasible for continuous monitoring. Therefore, this chapter presents a pilot study on the reconstruction of remote ECG (rECG) along with the rPPG waveform using a DNN-based end-to-end bitask network. The network involves a pre-trained transformer-based module that is shared between both rPPG and rECG reconstruction layers. The transformer module is adapted from a prior work which is designed to estimate rPPG waveform [220]. The output of the transformer module accomplishes the rPPG reconstruction task, and it is further connected to a simple CNN-based ECG header to estimate the rECG waveform. Subsequently, the estimated waveforms are used to measure HR and meanNN from the tachogram of their corresponding peak intervals. The measurements derived from rPPG and rECG are cross-validated with the parameters derived from the reference PPG and ECG, respectively.

The objective of this study is to promote the rPPG technology beyond its fundamental contribution towards cardiac health assessment. Blood flow information captured by PPG is associated with the heart's electrical conduction, establishing an intrinsic correlation between PPG and ECG. Systolic peak intervals derived from PPG are highly correlated with the corresponding R-R intervals of ECG, showing the possibility of deriving ECG from PPG. Building upon this insight, the pilot study explores the generation of rECG from corresponding rPPG waveforms. As rPPG has already been explored to capture the blood flow information in various research, detecting its primary electric source will be captivating, and it will surely expand the technology beyond its limited accomplishments. The

rest of this chapter is arranged as follows: Section 5.1 presents the description of the database used and proposed methodology. Experimental results and conclusion are drawn in Section 5.2 and 5.3, respectively. Finally, it is summarized in Section 5.4.

5.1 Materials and methods

5.1.1 Database

The rPPG and rECG reconstruction experiments using the proposed bitask network are conducted on the publicly available UBFC-rPPG dataset that is described in Section 4.1.3.1. In this dataset, videos were recorded concurrently with reference BVP signals. However, the ground-truth ECG signals that are required to train the ECG-header are not available in the dataset. Therefore, we employed a DNN-based method [332] to generate synthetic ECG waveforms from the available PPG signals. The PPG to ECG generation network is trained from scratch for 300 epochs with 256 batch size and a learning rate of 0.0001. Figure 5.1 shows a 10 s segment of the synthetic ECG signal along with its corresponding PPG signal. Since neural networks rely on extensive training data for effective performance, the videos and the corresponding PPG and ECG signals were segmented using a 2-s sliding window to create large data samples. A total of 1324 subsamples were created from the original data. Further, the segmented samples are split into 935 and 389 for training and testing, respectively.

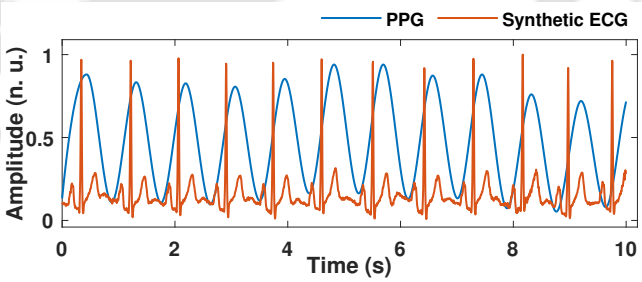


Figure 5.1: PPG signal along with the generated synthetic ECG waveform

5.1.2 Methodology

This section presents our end-to-end method to reconstruct rPPG and rECG using input video data ($V \in \mathbb{R}^{B\times C\times T\times H\times W}$, where B is the batch size, C is the number of channel, T represents temporal length, H and W is the height and width of the video frame). The overview of the proposed method is depicted in Figure 5.2, which involves two separate modules for rPPG and rECG estimation. A

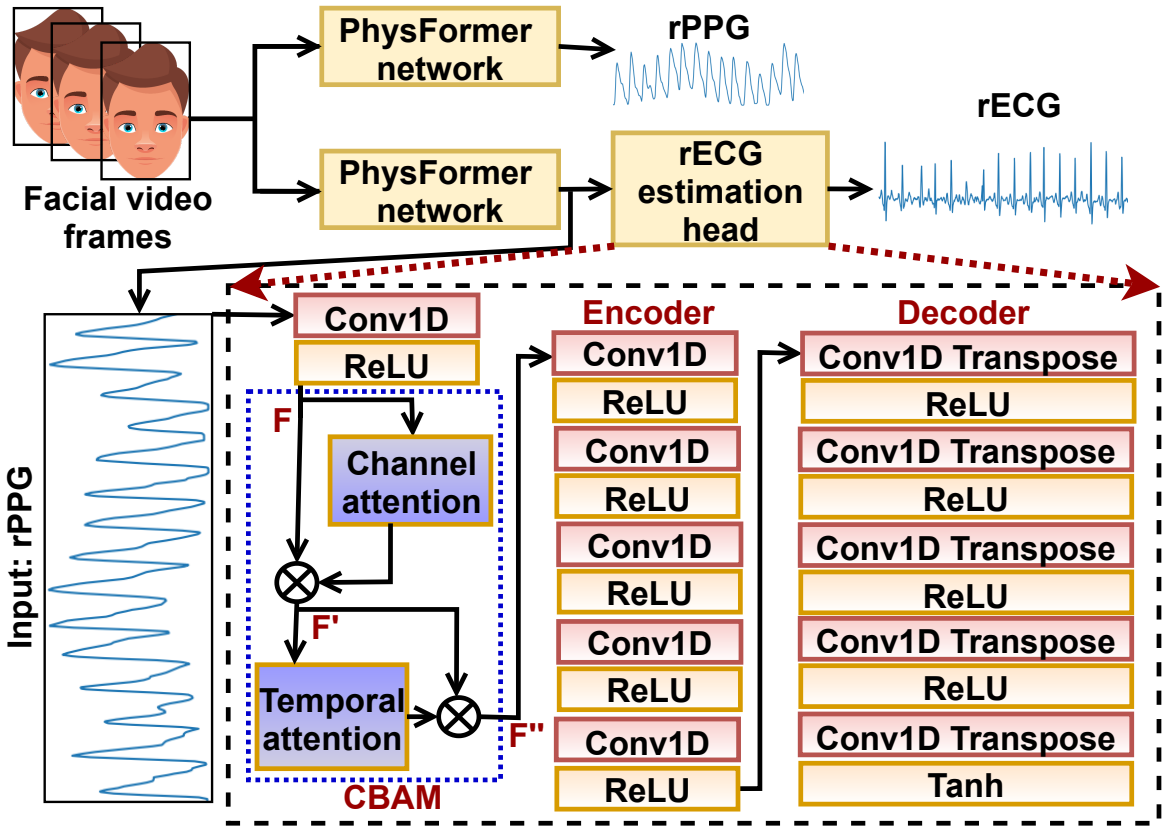


Figure 5.2: Block diagram of the proposed method

PhysFormer network is employed and trained from scratch to estimate the rPPG waveforms. The module is trained on the UBFC-rPPG dataset with a learning rate of 0.001 and batch size 8 for 250 epochs. After training, the PhysFormer module is leveraged into the bitask network, where it is concatenated with an rECG estimation header to generate the rECG waveforms. All the networks are implemented and trained on the NVIDIA Tesla P100-PCIE-16GB GPU platform using the PyTorch library.

5.1.2.1 PhysFormer Network

To accomplish the objective of the proposed work for estimating rECG along with rPPG, two separate networks would be required that will eventually increase the computational cost. As ECG and PPG are highly correlated, the information extracted from the rPPG network could be utilized for rECG generation. Therefore, in the proposed work, rPPG and rECG estimating tasks share an intermediate PhysFormer network [220] that helps reduce the computational budget. Because of

Table 5.1: PhysFormer network architecture

Name	Layer	K, C / S / P	Output shape
Stem0	Conv3d	(1×5×5), 24 / 1 / [0,2,2]	8×24×60×128×128
	BatchNorm3d	-	8×24×60×128×128
	ReLU	-	8×24×60×128×128
	MaxPool3d	(1×2×2), - / (1×2×2) / 0	8×24×60×64×64
Stem1	Conv3d	(3×3×3), 48 / 1 / 1	8×48×60×64×64
	BatchNorm3d	-	8×48×60×64×64
	ReLU	-	8×48×60×64×64
	MaxPool3d	(1×2×2), - / (1×2×2) / 0	8×48×60×32×32
Stem2	Conv3d	(3×3×3), 96 / 1 / 1	8×96×60×32×32
	BatchNorm3d	-	8×96×60×32×32
	ReLU	-	8×96×60×32×32
	MaxPool3d	(1×2×2), - / (1×2×2) / 0	8×96×60×16×16
Embedding	Conv3d	(4×4×4), 96 / [4,4,4] / 0	8×96×15×4×4
Transformer	TD-MHSA ST-FF	-	8×240×96
rPPG header	Upsample	scale_factor = 2×1×1	8×96×30×4×4
	Conv3d	(3×1×1), 96 / 1 / [1,0,0]	8×96×30×4×4
	BatchNorm3d	-	8×96×30×4×4
	ELU	-	8×96×30×4×4
	Upsample	scale_factor = 2×1×1	8×96×60×4×4
	Conv3d	(3×1×1), 48 / 1 / [1,0,0]	8×48×60×4×4
	BatchNorm3d	-	8×48×60×4×4
	ELU	-	8×48×60×4×4
	Conv1d	1, 1 / 1 / 0	8×1×60
	FC	-	8×512

Note: K, S and P represents kernel size, stride and padding, respectively.
TD-MHSA: temporal difference multi-head self attention; ST-FF: spatio-temporal feed forward

the robust performance of the PhysFomer network, we leveraged it into our architecture for rPPG estimation. Table 5.1 presents the details of PhysFormer network. It is a transformer-based network that involves a shallow stem, a tube tokenizer, transformer blocks, and an rPPG predictor head. The stems (0 to 1) are formed using three convolutional blocks to extract coarse local spatio-temporal features from raw input video data. This block helps in fast convergence and enhances the subsequent global self-attention. Subsequently, the embedding layer partitions the feature frames into spatio-temporal tube tokens that are forwarded with 12 transformer blocks. Each transformer block consists of a temporal difference multi-head self attention (TD-MHSA) module and a spatio-temporal feed forward (ST-FF) network that together captures the global-local refined rPPG features. Finally, the rPPG features are upsampled, spatially averaged and projected into a 1D rPPG signal by the rPPG

predictor head. Further, the estimated rPPG is fed to the CNN-based head to recover the rECG waveform.

5.1.2.2 rECG Header

The rECG header is designed using simple convolution blocks, including attention and an encoder-decoder module. The attention block is employed to learn the critical part of the input rPPG waveform for generating the rECG signal. At the same time, the encoder-decoder block can be considered a dimension-reduction process, making the network robust to noises. Additionally, the encoder maps the input rPPG to a significant latent space, and subsequently, the decoder decodes the latent space to generate the rECG waveform. For the attention block a convolution block attention module (CBAM) [333] is used that involves both channel and temporal attention module. Initially, the input rPPG waveform is convolved using a one-dimensional kernel and its generated feature maps ($F \in \mathbb{R}^{C \times T}$) are forwarded towards the channel attention module, where the temporal information is aggregated by using average pooling and maximum pooling. Subsequently, both average and maximum pooling features shares a common multilayer perceptron network (MLP) and their corresponding outputs are fused together to generate the channel attention map ($F' \in \mathbb{R}^{C \times 1}$). In contrast, the temporal attention module uses the time-series relationship of the rPPG waveform to form the temporal attention map ($F'' \in \mathbb{R}^{1 \times T}$). In temporal attention module, the feature maps are aggregated along the channel axis using average and maximum pooling operation. The generated average and maximum pooled features are concatenated and forwarded through a convolution layer to generate F''. The entire CBAM process can be expressed as:

$$\begin{aligned} F' &= M_c(F) \times F \\ F'' &= M_t(F') \times F' \end{aligned} \tag{5.1}$$

Where M_c and M_t are channel and temporal attention modules, respectively. Table 5.2 presents the details of rECG header. The rECG header is trained for 50 epochs with a batch size of 16 and is optimized using the Adam optimizer with 0.0001 learning rate.

In order to reconstruct the proper morphological structure of the rECG waveform, the rECG header needs to be guided by a suitable cost function, where ground-truth synthetic ECG signals are used to supervise the network. We have utilized the cost function from our previous work that is derived by combining a negative Pearson correlation coefficient (L_ρ) function and the L1 loss (L_1)

Table 5.2: rECG header architecture

Layer	K, C / S / P	Output shape
Conv1d	(3, 16 / 1 / 1)	16×16×512
ReLU	-	16×16×512
CBAM	-	16×16×512
Encoder		
Conv1d	31, 32 / 2 / 15	16×32×256
ReLU	-	16×32×256
Conv1d	31, 64 / 1 / 15	16×64×256
ReLU	-	16×64×256
Conv1d	31, 128 / 2 / 15	16×128×128
ReLU	-	16×128×128
Conv1d	31, 256 / 1 / 15	16×256×128
ReLU	-	16×256×128
Conv1d	31, 512 / 2 / 15	16×512×64
ReLU	-	16×512×64
Decoder		
ConvTranspose1d	31, 256 / 2 / 15	16×256×128
ReLU	-	16×256×128
ConvTranspose1d	31, 128 / 1 / 15	16×128×128
ReLU	-	16×128×128
ConvTranspose1d	31, 64 / 2 / 15	16×64×256
ReLU	-	16×64×256
ConvTranspose1d	31, 32 / 1 / 15	16×32×256
ReLU	-	16×32×256
ConvTranspose1d	31, 1 / 2 / 15	16×1×512
Tanh	-	16×1×512
Flattening	-	16×512

function. The L_ρ parts maintain the periodicity of the waveform, whereas the L_1 part tries to reduce the difference between the ground-truth ECG and the estimated rECG signal values. Further, to enhance the QRS complex of the generated rECG waveform, a QRS complex-enhanced loss (L_{QRS}) function [332] is incorporated with the prior parts. The combined cost function is expressed as:

$$Loss = (L_\rho + L_1 + L_{QRS})/3 \tag{5.2}$$

where,

$$L_\rho(x,y) = 1 - \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2 \sum_{i=1}^{n}(y_i-\bar{y})^2}},$$

$$L_1(x,y) = \sum_{i=1}^{n}|x_i - y_i|, \qquad \text{and}$$

$$L_{QRS}(x,y) = \sum_{i=1}^{n}|x_i - y_i| \left(1 + \beta \sum_{k=1}^{K} e^{\frac{-(i-c_k)^2}{2\sigma^2}}\right)$$

where x represents the ground-truth synthetic ECG signal values, y denotes the estimated rECG signal values, n is the length of signal, c_k denotes the index of k^{th} R-peak, σ and β are hyperparameters.

5.1.3 Evaluation metrics

The method's performance is evaluated by comparing the cardiac parameters measured from the estimated rPPG and rECG waveforms with their corresponding values obtained using the ground-truth PPG and synthetic ECG signals. Initially, the systolic peak instants of the waveforms are detected, and subsequently, the tachogram of their corresponding peak intervals is used to measure heart rate (HR) and average cardiac interval (meanNN) values. The measured values are reported using statistical metrics, MAE, RMSE, MER, SDE, and R. Additionally, the performance is evaluated graphically via Bland-Altman (BA) and regression plots.

5.2 Results

An experiment is performed to show the possible use of our proposed bitask network in the vitals measurement application using the estimated rPPG and rECG waveforms. Since we present a pilot study on the generation of rECG waveform using a camera sensor, we limit the evaluation to only two cardiac parameters, such as HR and meanNN. Figure 5.3 shows the rPPG and rECG waveforms estimated from the bitask network along with their corresponding reference signals. Initially, the systolic peaks of PPG (SP^1) and rPPG (SP^2) waveforms are detected using a simple amplitude-temporal thresholding-based technique as shown in Figure 5.3(a). Subsequently, the time difference between the adjacent peak instants are used to create the SS^i time interval series as:

$$SS_t^i = SP_{t+1}^i - SP_t^i \tag{5.3}$$

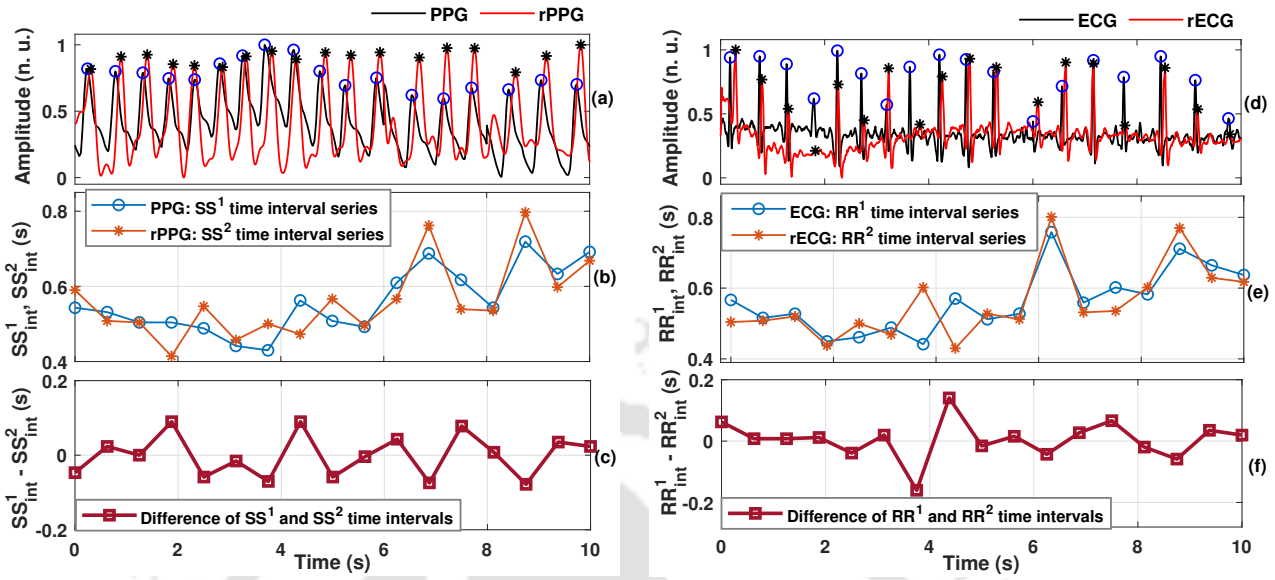


Figure 5.3: Cardiac cycle extraction from ground-truth and estimated waveforms. (a) Detected systolic peaks SP^1 and SP^2 in PPG and rPPG, respectively, (b) extracted cardiac interval time series SS^1 and SS^2, (c) difference of SS^1 and SS^2 intervals, (d) detected R-peaks R^1 and R^2 in ECG and rECG, respectively, (e) extracted cardiac interval time series RR^1 and RR^2, (f) difference of RR^1 and RR^2 intervals.

Table 5.3: Statistical comparison between HR and meanNN estimated from rPPG and rECG with the corresponding values measured from reference PPG and ECG

Statistical parameter ↓	PPG-rPPG		ECG-rECG	
	HR	meanNN	HR	meanNN
MAE	2.531	0.017	4.559	0.028
RMSE	3.763	0.026	6.440	0.041
MER	2.651	2.623	4.739	4.441
SDE	2.802	0.020	4.578	0.030

Where SP_t^i represents timing information of the tth SP^i peak and $i(=1,2)$ corresponds to systolic peaks detected from PPG and rPPG, respectively. Figure 5.3(b) represents the time interval series created from both PPG and rPPG waveforms and subsequently the error between SS^1 and SS^2 is presented in Figure 5.3(c). Similarly, Figure 5.3(d)-(f) represents the cardiac cycle extraction process for ECG and rECG waveforms. Finally, the derived cardiac intervals are used to measure both meanNN and HR values. Table 5.3 presents the comparison between the estimated values and their corresponding reference measurements. The evaluation results show MAE 2.531, RMSE 3.763, MER 2.651, and SDE 2.802 for HR values estimated from PPG and rPPG waveforms. Whereas, the metrics are found to be 0.017, 0.026, 2.623, and 0.020 for meanNN measurement. For HR and meanNN

estimated from ECG and rECG, the MAE, RMSE, MER, and SDE pairs are (4.559, 0.028), (6.440, 0.041), (4.739, 4.441), and (4.578, 0.030), respectively. It is observed that the error values between PPG and rPPG are relatively lower than the ECG and rECG error measures. However, the obtained ECG-rECG error values are insignificant that indicates the potentiality of the proposed network to generate rECG waveform. To validate the proposed network for HR and meanNN estimation, the comparison is also illustrated graphically using regression and BA plots in Figure 5.4. It shows Pearson's correlation coefficient pairs for HR and meanNN to be (0.919, 0.911) and (0.770, 0.736) using PPG-rPPG and ECG-rECG waveforms, respectively. Thus, these results reveals a strong correlation between the estimated and ground-truth measurement values. Also, a satisfactory agreement can be observed in the BA plots, where only few percentage of data are lying outside the limits of agreement (LOA). The obtained parameters from regression and BA plots are listed in Table 5.4. These experimental results validates our proposed bitask network in vitals measurement application using the robust reconstructed rPPG and rECG waveforms.

Table 5.4: Parameters obtained for regression and Bland–Altman (BA) plots of HR and meanNN parameters obtained from estimated and reference waveforms

Waveforms	Vitals	R	BA plot	
			Mean	LOA [L, U]
PPG-rPPG	HR	0.919	-0.254	[-7.660, 7.151]
	meanNN	0.911	0.002	[-0.047, 0.052]
ECG-rECG	HR	0.770	-1.978	[-14.068, 10.112]
	meanNN	0.736	0.012	[-0.064, 0.088]

This work depicts the possibility of a camera sensor to generate an rECG waveform along with the pre-investigated rPPG waveform using a DNN-based bitask network. Although the experimental analysis on the UBFC-rPPG dataset shows very promising results, the proposed method still needs to be investigated on a more dynamic dataset, including large motion artifacts. Also, the current study focuses only on the proper delineation of the QRS complex of the rECG waveform. However, other fiducial points of ECG waveforms are also clinically pivotal for diagnosing cardiac anomalies. Thus, as a future research direction, more exploration and investigation can be carried out to recover all the clinically relevant information from the rECG waveform.

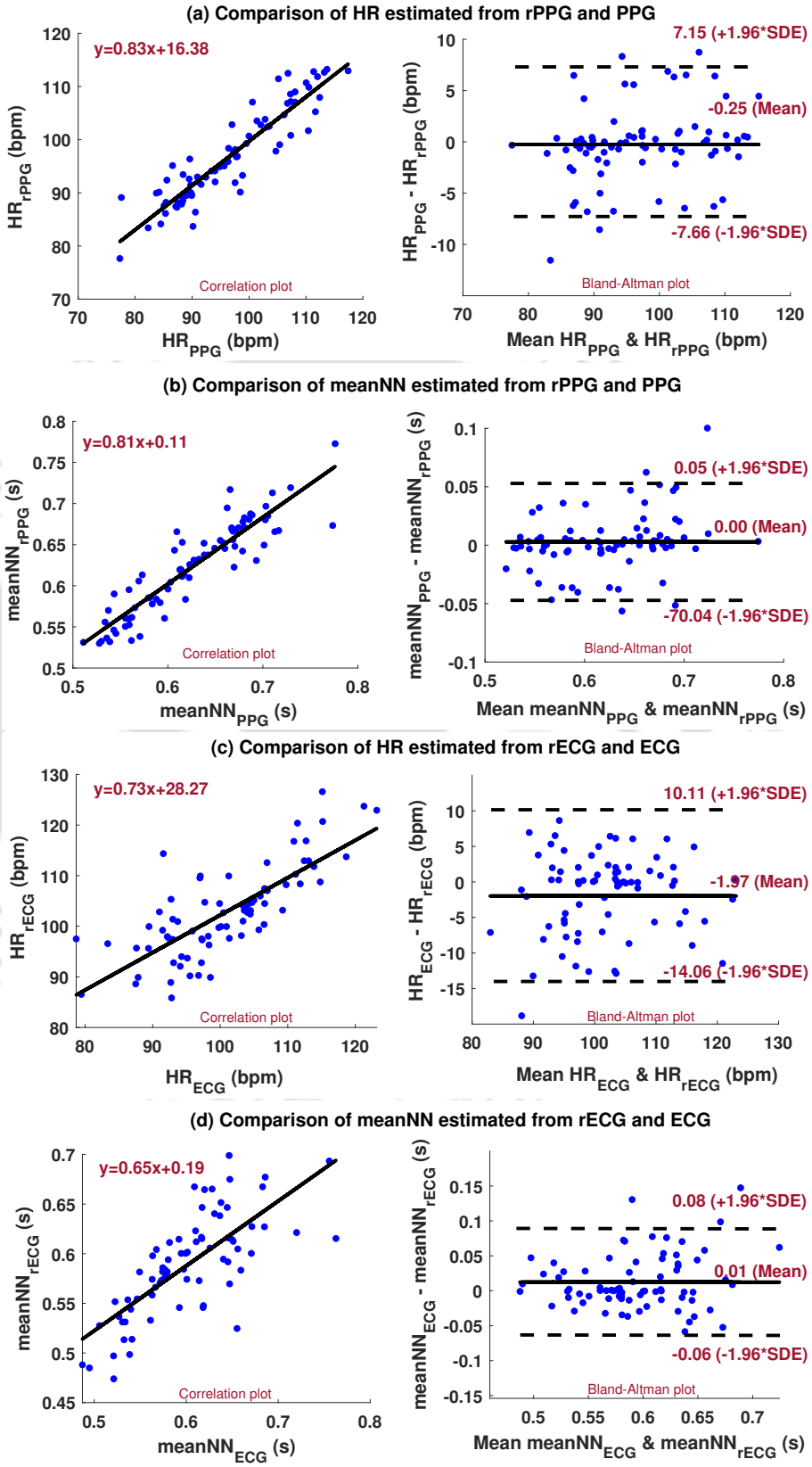


Figure 5.4: Comparison of HR and meanNN parameters using correlation [left] and Bland–Altman plots [right]: (a)-(b) Parameters estimated from PPG and rPPG, and (c)-(d) Parameters estimated from ECG and rECG

5.3 Summary

In this chapter, the established rPPG concept using video data is investigated a step ahead to extract the cardiac electrical events by reconstructing the rECG waveform using a DNN-based network. This chapter mainly presents a pilot study to look into the possibility of rECG reconstruction and its clinical application along with the generation of rPPG waveform. For this purpose, a bitask network is proposed that involves a state-of-the-art PhysFormer network and a CNN-based module to estimate rPPG and rECG waveforms, respectively. The network is trained and tested on a publicly available UBFC-rPPG dataset. Subsequently, the method is validated by comparing the estimated HR and meanNN parameters with their corresponding reference values. The experimental results show the promising outcomes of the proposed network, and thus, the estimated rPPG and rECG signals may be employed for cardiac health monitoring. However, the framework needs to be tested on a dynamic dataset before being deployed for clinical use. We believe the proposed work would open opportunities for rPPG domain researchers to explore and investigate all the hidden possibilities of rPPG technology to make it a robust cardiac multimodal system.

6

Conclusions

Contents

Living a healthy and quality lifestyle is the topmost priority of every individual. In order to reduce healthcare expenses and enhance convenience and accessibility for users, it is essential to transfer the health assessment process from hospitals to home settings. Therefore, the current healthcare systems are focused on developing efficient physiological measurement devices in terms of quality, cost, and ergonomic needs that can be useful for the general public outside the clinical environment. Personalized health assessment biosensors are designed to continuously monitor vital parameters essential for diagnosing and treating various life-threatening diseases.

In this thesis, two noninvasive and nonintrusive cardiac sensing modalities, including SCG and rPPG are examined and their potential applications are explored. The methods and algorithms developed for both modalities have improved the estimation of the targeted vital measurements, thereby making the frameworks applicable for personalized health monitoring. Chapter 1 provides an introduction to the fundamentals of both SCG and rPPG signals, covering their acquisition methods, physiological characteristics, and challenges related to their morphological structure. Along with these, it also includes the state-of-the-art methods for both SCG and rPPG that have been previously reported in the literature. Chapter 2 presents a standalone SCG delineation method to detect AO and pAC fiducial points. These fiducial points are leveraged to calculate SBP, DBP, and HRV parameters. Additionally, systolic time intervals are determined by utilizing these fiducial points along with the demographic data of volunteers through a neural network framework. Chapter 3 introduces a conventional rPPG technique based on 2D-VMD. This approach is employed to enhance the robustness of HR measurement in rPPG methods. In Chapter 4, two deep neural network architectures are designed to generate reliable rPPG waveforms, which are subsequently utilized for HR and HRV measurement. Chapter 5 demonstrates another application of rPPG technology, wherein an end-to-end multitask network is employed to generate rECG signals along with rPPG signals. The proposed network effectively captures both the electrical and mechanical phenomenon of the heart, bringing the framework one step closer to an integrated remote healthcare system for diagnosing cardiovascular risks. The current chapter is arranged as follows: the major contributions of this thesis are presented in Section 6.1. Possible future works are provided in Section 6.2.

6.1 Summary of Contributions

The major contributions of our work reported in this thesis are as follows:

- The cuffless BP estimation is first ever one of its kind where a single SCG component of an accelometric sensor is utilized. This approach proves effective for estimating both systolic and diastolic blood pressure. The method's efficacy suggests it could potentially replace current cuffless BP monitoring approaches, which typically rely on multiple cardiac sensing modalities.
- In critical scenarios, simple decision based rules are formulated to reduce false detections and misdetection rates of AO peaks identified by the MVMD model.
- A novel spatial-temporal filtering technique utilizing 2D-VMD, azimuthally averaged power spectrum density (AAPSD), and multimode kurtosis is devised to improve HR estimation.
- Two DNN based on 1D (DelNet) and 2D (TFL) CNN architecture are investigated for the extraction of reliable rPPG signals.
- A novel scalogram based feature map is generated from the raw RGB traces of video frames using the wavelet transformation. This feature map assists the TFL network to learn and characterize a variety of rPPG signals.
- The reconstruction of rECG from video frames using a 3D-CNN based end-to-end bitask network is first-ever investigated for the rPPG technology.

6.2 Scope for the Future Work

The first part of the thesis presents the study on SCG signal that is acquired non-invasively using an accelerometric sensor. This sensing modality is highly suitable for cardiac analysis and disease diagnosis in wearable healthcare due to its ease-of-use and low-cost. However, the conducted studies on SCG are validated only on a limited number of subjects. Also, the recordings were done in a relatively stable condition, prohibiting significant body movements. To enhance the applicability of these methods, their performance should be assessed across various physiological and pathological conditions in individuals under dynamic environments in future works.

While SCG technology offers numerous applications, its requirement for contact-based data acquisition may pose challenges in meeting the ergonomic needs of healthcare systems, including subject

comfort and mobility during long-term monitoring. Hence, the second part of the dissertation introduces non-contact camera based rPPG technology. It is one of the emerging approach in the healthcare system that addresses the limitations of the contact-based methods. The presented studies on rPPG can be employed in various applications, including telehealth, consumer fitness, neonatal monitoring, security and emotional computing. Despite considerable research in the field, certain limitations and unresolved issues still exist. Some of the research gaps that could serve as future directions in this field are highlighted as follows:

- The HR values estimated from the video data vary with many factors such as light intensity, skin-tone, facial expression, and motion artifacts. Several methods have been developed to compensate these challenges. However, it is essential to provide a comprehensive understanding of how these challenges are taken care from both biophysical and technical perspectives, rather than evaluating the performance on datasets containing noise factors.

- The accessibility of rPPG is facilitated by the widespread availability of cameras. However, the diverse specifications of available cameras can easily impact the performance of developed algorithms. It requires more attention from researchers to determine the optimal camera qualities and characterizing the sensitivity of measurements to camera parameters.

- A wide-range of public datasets are available to the rPPG research community, primarily addressing two key challenges: illumination variation and motion artifacts. However, future datasets should include additional challenges such as diverse skin tones, varying age groups of subjects, multiple persons in a frame, and variations in camera distance. This broader coverage will enhance the robustness and applicability of rPPG in real-world scenarios. Moreover, existing datasets have been recorded exclusively from healthy subjects. Consequently, there is a need to create a dataset comprising data from subjects with cardiac-disorders to investigate rPPG performance under pathological conditions.

- The rPPG methods and algorithms presented by this thesis enables the measurement of HR and HRV parameters across various dynamic conditions. Undoubtedly, both HR and HRV are essential physiological parameters to indicate health condition of a person. However, further research is required to explore more clinical applications, including the measurement of other vital signs such as RR, BP, and SpO2.

- The multitask network proposed for the generation of rECG signal focuses only on the delineation of the QRS complex of the wavfeorm. However, other fiducial points of ECG waveforms are also clinically useful for diagnosing cardiac anomalies. Thus, as a future research direction, more exploration and investigation can be carried out to recover all the clinically relevant information from the ECG waveform.

A

Developed Graphical User Interfaces

Contents

Graphical user interfaces (GUIs) are valuable tools for showing the output-behavior and practical applications of designed algorithms. They allow a visual means of interacting with graphical elements, making the usability of the system convenient for the end-user. In this section, we showcase our developed GUIs based on computer-interaction.

A.1 MATLAB-based GUI for BP estimation

For a continuous screening of BP values, a MATLAB (MathWorks, Inc.)-based graphical user interface (GUI) is developed. It facilitates the continuous monitoring of SCG signals. The system computes and displays LVET′ and HR traces. These parameters are further employed by the mathematical model (Chapter 2.3.1) for BP estimation. The GUI is shown in Figure A.1, which illustrates the estimation of SBP, DBP, and mean arterial pressure (MAP) values from the SCG signal of an individual.

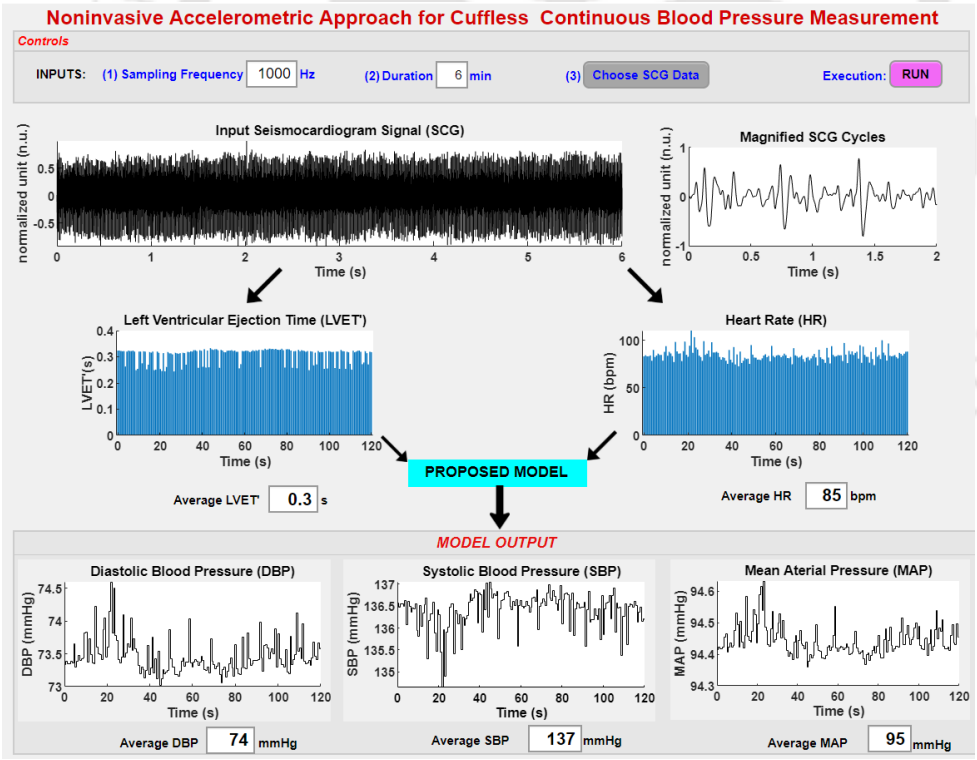


Figure A.1: GUI for our BP estimation model. MAP is computed as (SBP + 2DBP)/3.

A.2 Python based GUI for HR and HRV measurement

As a personalized cardiac healthcare application, a PyQT-based GUI is developed to measure the HR and HRV parameters using the facial video data. The system streams the video of a subject. It

segments and displays the facial skin pixels by using the method described in Chapter 4.2.1.1. Further raw RGB traces are obtained from the skin pixels and used to derive the scalogram-based feature maps (Chapter 4.2.1.1). The developed GUI uses the weights of the trained TFL network (Chapter 4.2.1.2) to estimate a rPPG signal. Subsequently, the system identifies the systolic peaks of the rPPG signal, enabling the measurement of HR and HRV indices. The system estimates following HRV parameters: meanNN, meanHR, SDSD, RMSSD, pNN50, and LF/HF. The designed GUI is shown in Figure A.2 which displays the raw RGB traces, estimated rPPG signal with detected systolic peaks, and the measured HR and HRV parameters of the subject.

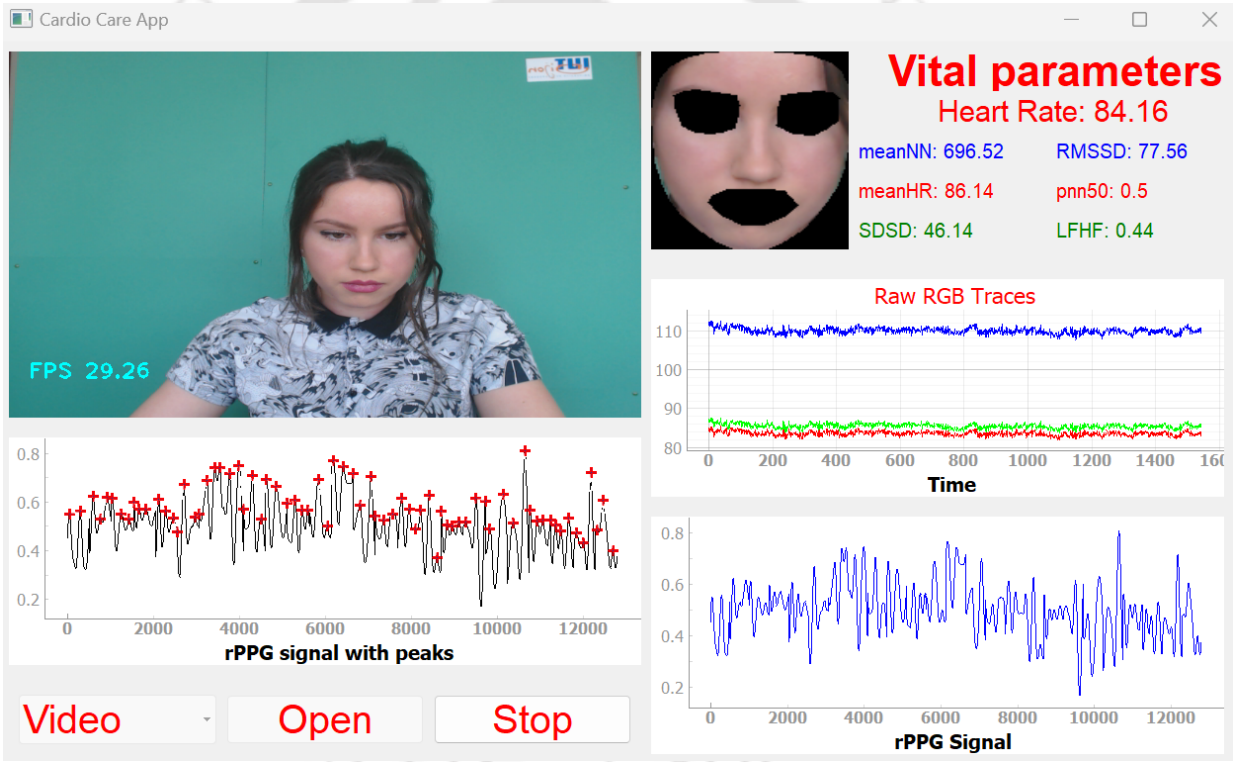


Figure A.2: GUI for cardiac parameter estimation using a TFL-based rPPG framework.

List of Publications

Journal Publications

1. M. Das, Tilendra Choudhary, L.N. Sharma, and M.K. Bhuyan, "Analyzing SCG methodology for identification of ventricular depolarization events," **Biomedical Signal Processing and Control**, vol. 100, pp. 106940, 2025.

2. M. Das, M. K. Bhuyan, and L. N. Sharma, "Time-Frequency Learning Framework for rPPG Signal Estimation Using Scalogram Based Feature Map of Facial Video Data," **IEEE Transactions on Instrumentation and Measurement**, 2023. doi: 10.1109/TIM.2023.3287243

3. M. Das, T. Choudhary, M. K. Bhuyan, and L. N. Sharma, "Non-Contact Heart Rate Measurement From Facial Video Data Using a 2D-VMD Scheme," **IEEE Sensors Journal**, vol. 22, no. 11, pp. 11153-–11161, 2022.

4. M. Das, T. Choudhary, L.N. Sharma, and M.K. Bhuyan, "Noninvasive Accelerometric Approach for Cuffless Continuous Blood Pressure Measurement", **IEEE Transactions on Instrumentation and Measurement**, vol. 70, pp. 1—9, 2021.

5. T. Choudhary, M. Das, M. K. Bhuyan, and L. N. Sharma, "Vibrocarotidography: A Novel Measurement Technique to Quantify Pulsations at Common Carotid Arteries," **IEEE Transactions on Instrumentation and Measurement,** vol. 70, pp. 1–8, 2021.

6. T. Choudhary, M. Das, L. N. Sharma, M. K. Bhuyan, "Analyzing seismocardiographic approach for heart rate variability measurement," **Biomedical Signal Processing and Control,** vol. 68, pp. 102793, 2021.

7. M. Das, M.K. Bhuyan, L.N. Sharma, Sultan Alfarhood, and Mejdl Safran "An End-to-End Bitask Network for Remote PPG and ECG Measurement: A Pilot Study," **Scientific Reports,** 2024 **(Under review)** .

Conference Publications

1. M. Das, T. Choudhary, M. K. Bhuyan, and L. N. Sharma, "A Multiresolution Method for Non-Contact Heart Rate Estimation Using Facial Video Frames," in **Proc. International Conference on Wireless Communications Signal Processing and Networking (WiSPNET),** 2022, pp. 115—119.

2. M. Das, M. K. Bhuyan, and L. N. Sharma, "DelNet based Systolic Peak Delineation in Remote PPG Signal from Facial Video Frames", in **Proc. 14th International Conference on Computing, Communication and Networking Technologies (ICCCNT),** 2023.

Bibliography

[1] T. Choudhary, M. Bhuyan, and L. Sharma, "A novel method for aortic valve opening phase detection using scg signal," *IEEE Sensors Journal*, vol. 20, no. 2, pp. 899–908, 2019.

[2] X. Ding and Y.-T. Zhang, "Pulse transit time technique for cuffless unobtrusive blood pressure measurement: from theory to algorithm," *Biomedical engineering letters*, vol. 9, pp. 37–52, 2019.

[3] Z. Tang, T. Tamura, M. Sekine, M. Huang, W. Chen, M. Yoshida, K. Sakatani, H. Kobayashi, and S. Kanaya, "A chair–based unobtrusive cuffless blood pressure monitoring system based on pulse arrival time," *IEEE Journal of biomedical and health informatics*, vol. 21, no. 5, pp. 1194–1205, 2016.

[4] 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.

[5] 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.

[6] 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.

[7] X. He, R. A. Goubran, and X. P. Liu, "Secondary peak detection of ppg signal for continuous cuffless arterial blood pressure measurement," *IEEE Transactions on Instrumentation and Measurement*, vol. 63, no. 6, pp. 1431–1439, 2014.

[8] K. Song, K.-y. Chung, and J.-H. Chang, "Cuffless deep learning-based blood pressure estimation for smart wristwatches," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 7, pp. 4292–4302, 2019.

[9] C. Landry, E. T. Hedge, R. L. Hughson, S. D. Peterson, and A. Arami, "Accurate blood pressure estimation during activities of daily living: A wearable cuffless solution," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 7, pp. 2510–2520, 2021.

[10] K. Soerensen, A. K. Verma, A. Blaber, J. Zanetti, S. E. Schmidt, J. J. Struijk, and K. Tavakolian, "Challenges in using seismocardiography for blood pressure monitoring," in *2017 Computing in Cardiology (CinC)*. IEEE, 2017, pp. 1–4.

[11] C. Yang and N. Tavassolian, "Combined seismo- and gyro-cardiography: A more comprehensive evaluation of heart-induced chest vibrations," *IEEE Journal of Biomedical and Health Informatics*, vol. 22, no. 5, pp. 1466–1475, Sep. 2018.

[12] 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.

[13] J. Park, S. Yang, J. Sohn, J. Lee, S. Lee, Y. Ku, and H. C. Kim, "Cuffless and continuous blood pressure monitoring using a single chest-worn device," *IEEE Access*, vol. 7, pp. 135 231–135 246, 2019.

[14] O. T. Inan, P. F. Migeotte, K. S. Park, M. Etemadi, K. Tavakolian, R. Casanella, J. Zanetti, J. Tank, I. Funtova, G. K. Prisk, and M. D. Rienzo, "Ballistocardiography and seismocardiography: A review of recent advances," *IEEE Journal of Biomedical and Health Informatics*, vol. 19, no. 4, pp. 1414–1427, July 2015.

[15] T. Choudhary, L. Sharma, and M. K. Bhuyan, "Automatic detection of aortic valve opening using seismocardiography in healthy individuals," *IEEE journal of biomedical and health informatics*, vol. 23, no. 3, pp. 1032–1040, 2018.

[16] F. Khosrow-khavar, K. Tavakolian, A. P. Blaber, J. M. Zanetti, R. Fazel-Rezai, and C. Menon, "Automatic annotation of seismocardiogram with high-frequency precordial accelerations," *IEEE Journal of Biomedical and Health Informatics*, vol. 19, no. 4, pp. 1428–1434, July 2015.

[17] K. Pandia, S. Ravindran, R. Cole, G. Kovacs, and L. Giovangrandi, "Motion artifact cancellation to obtain heart sounds from a single chest-worn accelerometer," in *2010 IEEE International Conference on Acoustics, Speech and Signal Processing*, March 2010, pp. 590–593.

[18] J. M. Zanetti and K. Tavakolian, "Seismocardiography: Past, present and future," in *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, July 2013, pp. 7004–7007.

[19] R. S. Crow, P. Hannan, D. Jacobs, L. He dquist, and D. M. Salerno, "Relationship between seismocardiogram and echocardiogram for events in the cardiac cycle," *American Journal of Noninvasive Cardiology*, vol. 8, no. 1, pp. 39–46, 1994.

[20] D. Rai, H. K. Thakkar, S. S. Rajput, J. Santamaria, C. Bhatt, and F. Roca, "A comprehensive review on seismocardiogram: current advancements on acquisition, annotation, and applications," *Mathematics*, vol. 9, no. 18, p. 2243, 2021.

[21] M. J. Tadi, E. Lehtonen, T. Hurnanen, J. Koskinen, J. Eriksson, M. Pänkäälä, M. Teräs, and T. Koivisto, "A real-time approach for heart rate monitoring using a hilbert transform in seismocardiograms," *Physiological measurement*, vol. 37, no. 11, p. 1885, 2016.

[22] A. Taebi, B. E. Solar, A. J. Bomar, R. H. Sandler, and H. A. Mansy, "Recent advances in seismocardiography," *Vibration*, vol. 2, no. 1, pp. 64–86, 2019. [Online]. Available: https://www.mdpi.com/2571-631X/2/1/5

[23] Y. Chuo, K. Tavakolian, and B. Kaminska, "Evaluation of a novel integrated sensor system for synchronous measurement of cardiac vibrations and cardiac potentials," *Journal of medical systems*, vol. 35, no. 4, pp. 445–455, 2011.

[24] P. Castiglioni, A. Faini, G. Parati, and M. D. Rienzo, "Wearable seismocardiography," in *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Aug 2007, pp. 3954–3957.

[25] B. Bozhenko, "Seismocardiography-a new method in the study of functional conditions of the heart," *Terapevticheskii arkhiv*, vol. 33, pp. 55–64, 1961.

[26] R. M. Baevskii, A. D. Egorov, and L. A. Kazarian, "The method of seismocardiography," *Kardiologiia*, vol. 18, pp. 87–89, 1964.

[27] P. F. Migeotte, S. D. Ridder, J. Tank, N. Pattyn, I. Funtova, R. Baevsky, X. Neyt, and G. K. Prisk, "Three dimensional ballisto- and seismo-cardiography: Hij wave amplitudes are poorly correlated to maximal systolic force vector," in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Aug 2012, pp. 5046–5049.

[28] D. Salerno and J. Zanetti, "Seismocardiography: A new technique for recording cardiac vibrations. concept, method, and initial observations," *Journal of cardiovascular technology*, vol. 9, no. 2, pp. 111–118, 1990.

[29] D. M. Salerno, J. M. Zanetti, L. A. Green, M. R. Mooney, J. D. Madison, and R. A. V. Tassel, "Seismocardiographic changes associated with obstruction of coronary blood flow during balloon angioplasty," *The American Journal of Cardiology*, vol. 68, no. 2, pp. 201 – 207, 1991. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0002914991907446

[30] R. A. Wilson, V. S. Bamrah, J. Lindsay, M. Schwaiger, and J. Morganroth, "Diagnostic accuracy of seismocardiography compared with electrocardiography for the anatomic and physiologic diagnosis of coronary artery disease during exercise testing," *The American Journal of Cardiology*, vol. 71, no. 7, pp. 536 – 545, 1993. [Online]. Available: http://www.sciencedirect.com/science/article/pii/000291499390508A

[31] M. Jerosch-Herold, J. Zanetti, H. Merkle, L. Poliac, H. Huang, A. Mansoor, F. Zhao, and N. Wilke, "The seismocardiogram as magnetic-field-compatible alternative to the electrocardiogram for cardiac stress monitoring," *The International Journal of Cardiac Imaging*, vol. 15, no. 6, pp. 523–531, 1999. [Online]. Available: http://dx.doi.org/10.1023/A:1006364518204

[32] M. Di Rienzo, F. Rizzo, G. Parati, G. Brambilla, M. Ferratini, and P. Castiglioni, "Magic system: a new textile-based wearable device for biological signal monitoring. applicability in daily life and clinical setting," in *2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*, Jan 2005, pp. 7167–7169.

[33] M. D. Rienzo, E. Vaini, P. Castiglioni, G. Merati, P. Meriggi, G. Parati, A. Faini, and F. Rizzo, "Wearable seismocardiography: Towards a beat-by-beat assessment of cardiac mechanics in ambulant subjects," *Autonomic Neuroscience*, vol. 178, no. 1–2, pp. 50 – 59, 2013, papers from {ESGCO} 2012 Meeting. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1566070213000805

[34] M. D. Rienzo, P. Meriggi, F. Rizzo, E. Vaini, A. Faini, G. Merati, G. Parati, and P. Castiglioni, "A wearable system for the seismocardiogram assessment in daily life conditions," in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Aug 2011, pp. 4263–4266.

[35] M. D. Rienzo, P. Meriggi, E. Vaini, P. Castiglioni, and F. Rizzo, "24h seismocardiogram monitoring in ambulant subjects," in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Aug 2012, pp. 5050–5053.

[36] A. M. Carek, J. Conant, A. Joshi, H. Kang, and O. T. Inan, "Seismowatch: Wearable cuffless blood pressure monitoring using pulse transit time," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 1, no. 3, pp. 40:1–40:16, Sep. 2017. [Online]. Available: http://doi.acm.org/10.1145/3130905

[37] T. Ha, J. Tran, S. Liu, H. Jang, H. Jeong, R. Mitbander, H. Huh, Y. Qiu, J. Duong, R. L. Wang *et al.*, "A chest-laminated ultrathin and stretchable e-tattoo for the measurement of electrocardiogram, seismocardiogram, and cardiac time intervals," *Advanced Science*, vol. 6, no. 14, p. 1900290, 2019.

[38] M. Di Rienzo, G. Rizzo, Z. M. IÅŸilay, and P. Lombardi, "Seismote: A multi-sensor wireless platform for cardiovascular monitoring in laboratory, daily life, and telemedicine," *Sensors*, vol. 20, no. 3, 2020. [Online]. Available: https://www.mdpi.com/1424-8220/20/3/680

[39] M. M. H. Shandhi, J. Fan, J. A. Heller, M. Etemadi, L. Klein, and O. T. Inan, "Estimation of changes in intracardiac hemodynamics using wearable seismocardiography and machine learning in patients with heart failure: a feasibility study," *IEEE Transactions on Biomedical Engineering*, vol. 69, no. 8, pp. 2443–2455, 2022.

[40] Y. Wang, J. Li, H. Wang, Z. Yan, Z. Xu, C. Li, Z. Zhao, and S. A. Raza, "Non-contact wearable synchronous measurement method of electrocardiogram and seismocardiogram signals," *Review of Scientific Instruments*, vol. 94, no. 3, 2023.

[41] M. Nikbakht, M. Chan, D. J. Lin, C. J. Nichols, M. Bibidakis, M. Soliman, and O. T. Inan, "Seismonet: A multi-node wireless wearable platform for enhanced physiological sensing," in *2023 IEEE 19th International Conference on Body Sensor Networks (BSN)*. IEEE, 2023, pp. 1–4.

[42] F. Khosrow-Khavar, K. Tavakolian, A. Blaber, and C. Menon, "Automatic and robust delineation of the fiducial points of the seismocardiogram signal for noninvasive estimation of cardiac time intervals," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 8, pp. 1701–1710, 2017.

[43] W. Lin, W. Chou, P. Chang, C. Chou, M. Wen, M. Ho, W. Lee, M. Hsieh, C. Lin, T. Tsai, and M. Lee, "Identification of location specific feature points in a cardiac cycle using a novel seismocardiogram spectrum system," *IEEE Journal of Biomedical and Health Informatics*, vol. 22, no. 2, pp. 442–449, March 2018.

[44] P. K. Sahoo, H. K. Thakkar, W.-Y. Lin, P.-C. Chang, and M.-Y. Lee, "On the design of an efficient cardiac health monitoring system through combined analysis of ecg and scg signals," *Sensors*, vol. 18, no. 2, p. 379, 2018.

[45] J. Yao, S. Tridandapani, C. A. Wick, and P. T. Bhatti, "Seismocardiography-based cardiac computed tomography gating using patient-specific template identification and detection," *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 5, pp. 1–14, 2017.

[46] V. Zakeri, A. Akhbardeh, N. Alamdari, R. Fazel-Rezai, M. Paukkunen, and K. Tavakolian, "Analyzing seismocardiogram cycles to identify the respiratory phases," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 8, pp. 1786–1792, Aug 2017.

[47] A. Taebi, B. E. Solar, and H. A. Mansy, "An adaptive feature extraction algorithm for classification of seismocardiographic signals," in *IEEE Conference SoutheastCon 2018*, April 2018, pp. 1–5.

[48] H. Ashouri and O. T. Inan, "Automatic detection of seismocardiogram sensor misplacement for robust pre-ejection period estimation in unsupervised settings," *IEEE Sensors Journal*, vol. 17, no. 12, pp. 3805–3813, June 2017.

[49] O. T. Inan, M. Baran Pouyan, A. Q. Javaid, S. Dowling, M. Etemadi, A. Dorier, J. A. Heller, A. O. Bicen, S. Roy, T. De Marco *et al.*, "Novel wearable seismocardiography and machine learning algorithms can assess clinical status of heart failure patients," *Circulation: Heart Failure*, vol. 11, no. 1, p. e004313, 2018.

[50] H. Ashouri, S. Hersek, and O. T. Inan, "Universal pre-ejection period estimation using seismocardiography: Quantifying the effects of sensor placement and regression algorithms," *IEEE Sensors Journal*, vol. 18, no. 4, pp. 1665–1674, Feb 2018.

[51] M. Di Rienzo, E. Vaini, and P. Lombardi, "An algorithm for the beat-to-beat assessment of cardiac mechanics during sleep on earth and in microgravity from the seismocardiogram," *Scientific Reports Nature*, vol. 7, no. 1, p. 15634, 2017.

[52] O. Lahdenoja, T. Hurnanen, Z. Iftikhar, S. Nieminen, T. Knuutila, A. Saraste, T. Kiviniemi, T. Vasankari, J. Airaksinen, M. Pänkäälä, and T. Koivisto, "Atrial fibrillation detection via accelerometer and gyroscope of a smartphone," *IEEE Journal of Biomedical and Health Informatics*, vol. 22, no. 1, pp. 108–118, Jan 2018.

[53] Z. Iftikhar, O. Lahdenoja, M. J. Tadi, T. Hurnanen, T. Vasankari, T. Kiviniemi, J. Airaksinen, T. Koivisto, and M. Pänkäälä, "Multiclass classifier based cardiovascular condition detection using smartphone mechanocardiography," *Scientific reports*, vol. 8, no. 1, pp. 1–14, 2018.

[54] T. Hurnanen, E. Lehtonen, M. J. Tadi, T. Kuusela, T. Kiviniemi, A. Saraste, T. Vasankari, J. Airaksinen, T. Koivisto, and M. Pänkäälä, "Automated detection of atrial fibrillation based on time–frequency analysis of seismocardiograms," *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 5, pp. 1233–1241, Sep. 2017.

[55] H. Lee, H. Lee, and M. Whang, "An enhanced method to estimate heart rate from seismocardiography via ensemble averaging of body movements at six degrees of freedom," *Sensors*, vol. 18, no. 1, p. 238, 2018.

[56] P.-F. Migeotte, V. Mucci, Q. Delière, L. Lejeune, and P. Van De Borne, "Multi-dimensional kineticardiography a new approach for wearable cardiac monitoring through body acceleration recordings," in *XIV Mediterranean Conference on Medical and Biological Engineering and Computing 2016*. Springer, 2016, pp. 1125–1130.

[57] M. J. Tadi, E. Lehtonen, A. Saraste, J. Tuominen, J. Koskinen, M. Teräs, J. Airaksinen, M. Pänkäälä, and T. Koivisto, "Gyrocardiography: A new non-invasive monitoring method for the assessment of cardiac mechanics and the estimation of hemodynamic variables," *Scientific reports*, vol. 7, no. 1, pp. 1–11, 2017.

[58] C. Yang, S. Tang, and N. Tavassolian, "Utilizing gyroscopes towards the automatic annotation of seismocardiograms," *IEEE Sensors Journal*, vol. 17, no. 7, pp. 2129–2136, April 2017.

[59] C. Yang, S. Tang, and N. Tavassolian, "Annotation of seismocardiogram using gyroscopic recordings," in *2016 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, Oct 2016, pp. 204–207.

[60] F. Landreani, A. Martín-Yebra, C. Casellato, C. Frigo, E. Pavan, P.-F. Migeotte, and E. G. Caiani, "Beat-to-beat heart rate detection by smartphone's accelerometers: validation with ecg," in *2016 38th annual international conference of the IEEE engineering in medicine and biology society (EMBC)*. IEEE, 2016, pp. 525–528.

[61] O. Lahdenoja, T. Hurnanen, Z. Iftikhar, S. Nieminen, T. Knuutila, A. Saraste, T. Kiviniemi, T. Vasankari, J. Airaksinen, M. Pänkäälä *et al.*, "Atrial fibrillation detection via accelerometer and gyroscope of a smartphone," *IEEE Journal of Biomedical and Health Informatics*, vol. 22, no. 1, pp. 108–118, 2017.

[62] Z. Iftikhar, O. Lahdenoja, M. Jafari Tadi, T. Hurnanen, T. Vasankari, T. Kiviniemi, J. Airaksinen, T. Koivisto, and M. Pänkäälä, "Multiclass classifier based cardiovascular condition detection using smartphone mechanocardiography," *Scientific reports*, vol. 8, no. 1, p. 9344, 2018.

[63] J. Metzler, K. Kroschel, and D. Willersinn, "Automatic detection of measurement points for non-contact vibrometer-based diagnosis of cardiac arrhythmias," in *Medical Imaging 2017: Image-Guided Procedures, Robotic Interventions, and Modeling*, vol. 10135. International Society for Optics and Photonics, 2017, p. 101351S.

[64] Z. XIA, M. M. H. SHANDHI, O. T. INAN, and Y. ZHANG, "Interference and removal of respiration harmonics on noncontact seismo-cardiogram signals," *Structural Health Monitoring 2017*, no. shm, 2017.

[65] M. M. H. SHANDHI, Z. XIA, O. T. INAN, and Y. ZHANG, "Clutter effect on the noncontact seismocardiogram signals measured using microwave radars," *Structural Health Monitoring 2017*, no. shm, 2017.

[66] Z. Xia, M. M. H. Shandhi, O. T. Inan, and Y. Zhang, "Non-contact sensing of seismocardiogram signals using microwave doppler radar," *IEEE Sensors Journal*, vol. 18, no. 14, pp. 5956–5964, 2018.

[67] P. Shirkovskiy, A. Laurin, N. Jeger-Madiot, D. Chapelle, M. Fink, and R. Ing, "Airborne ultrasound surface motion camera: application to seismocardiography," *Applied Physics Letters*, vol. 112, no. 21, p. 213702, 2018.

[68] Y. Hu, E. G. Kim, G. Cao, S. Liu, and Y. Xu, "Physiological acoustic sensing based on accelerometers: A survey for mobile healthcare," *Annals of biomedical engineering*, vol. 42, pp. 2264–2277, 2014.

[69] C. Yang and N. Tavassolian, "Motion noise cancellation in seismocardiogram of ambulant subjects with dual sensors," in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Aug 2016, pp. 5881–5884.

[70] M. Paukkunen, P. Parkkila, T. Hurnanen, M. Pänkäälä, T. Koivisto, T. Nieminen, R. Kettunen, and R. Sepponen, "Beat-by-beat quantification of cardiac cycle events detected from three-dimensional precordial acceleration signals," *IEEE Journal of Biomedical and Health Informatics*, vol. 20, no. 2, pp. 435–439, March 2016.

[71] G. Shafiq, S. Tatinati, and K. C. Veluvolu, "Automatic annotation of peaks in seismocardiogram for systolic time intervals," in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Aug 2016, pp. 2672–2675.

[72] A. Taebi and H. A. Mansy, "Time-frequency distribution of seismocardiographic signals: A comparative study," *Bioengineering*, vol. 4, no. 2, p. 32, 2017.

[73] T. Choudhary, L. N. Sharma, and M. K. Bhuyan, "Suppression of artifacts from seismocardiogram signal using two-stage kalman filtering model," in *2018 International Conference on Signal Processing and Communications (SPCOM)*, July 2018, pp. 417–421.

[74] M. Malcangi, H. Quan, E. Vaini, P. Lombardi, and M. Di Rienzo, "Applying the efunn evolving paradigm to the recognition of artefactual beats in continuous seismocardiogram recordings," in *International Conference on Engineering Applications of Neural Networks*. Springer, 2017, pp. 256–264.

[75] C. Yang and N. Tavassolian, "Motion artifact cancellation of seismocardiographic recording from moving subjects," *IEEE Sensors Journal*, vol. 16, no. 14, pp. 5702–5708, July 2016.

[76] A. Q. Javaid, H. Ashouri, A. Dorier, M. Etemadi, J. A. Heller, S. Roy, and O. T. Inan, "Quantifying and reducing motion artifacts in wearable seismocardiogram measurements during walking to assess left ventricular health," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 6, pp. 1277–1286, June 2017.

[77] A. Taebi and H. Mansy, "Noise cancellation from vibrocardiographic signals based on the ensemble empirical mode decomposition," *J. Biotechnol. Bioeng*, vol. 2, no. 2, p. 00024, 2017.

[78] J. Burian, H. Toedtmann, M. Haescher, M. Aehnelt, and A. Kuijper, "Making noise-improving seismocardiography based heart analysis with denoising autoencoders," in *Proceedings of the 8th international Workshop on Sensor-Based Activity Recognition and Artificial Intelligence*, 2023, pp. 1–9.

[79] G. Cosoli, L. Casacanditella, E. Tomasini, and L. Scalise, "Heart rate assessment by means of a novel approach applied to signals of different nature," in *Journal of Physics: Conference Series*, vol. 778, no. 1. IOP Publishing, 2017, p. 012001.

[80] C. Yang and N. Tavassolian, "Combined seismo-and gyro-cardiography: A more comprehensive evaluation of heart-induced chest vibrations," *IEEE journal of biomedical and health informatics*, vol. 22, no. 5, pp. 1466–1475, 2017.

[81] F. Landreani, A. Martin-Yebra, C. Casellato, C. Frigo, E. Pavan, P. Migeotte, and E. G. Caiani, "Beat-to-beat heart rate detection by smartphone's accelerometers: Validation with ecg," in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Aug 2016, pp. 525–528.

[82] A. Dinh, F. M. Bui, and T. Nguyen, "An accelerometer based system to measure myocardial performance index during stress testing," in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Aug 2016, pp. 4877–4880.

[83] G. Zhang, A. C. Cottrell, I. C. Henry, and D. B. McCombie, "Assessment of pre-ejection period in ambulatory subjects using seismocardiogram in a wearable blood pressure monitor," in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Aug 2016, pp. 3386–3389.

[84] F. Khosrow-Khavar, "Automatic and non-invasive delineation of the seismocardiogram signal for the estimation of cardiac time intervals with applications in diastolic timed vibration and early stage hemorrhage detection," Ph.D. dissertation, Applied Sciences: School of Engineering Science, 2016.

[85] G. Shafiq, S. Tatinati, W. T. Ang, and K. C. Veluvolu, "Automatic identification of systolic time intervals in seismocardiogram," *Scientific reports, Nature*, vol. 6, p. 37524, 2016.

[86] A. Q. Javaid, S. Dowling, M. Etemadi, J. A. Heller, S. Roy, L. Klein, and O. T. Inan, "Quantification of posture induced changes in wearable seismocardiogram signals for heart failure patients," in *2016 Computing in Cardiology Conference (CinC)*, Sep. 2016, pp. 777–780.

[87] R. Abbasi-Kesbi, A. Valipour, and K. Imani, "Cardiorespiratory system monitoring using a developed acoustic sensor," *Healthcare Technology Letters*, vol. 5, no. 1, pp. 7–12, 2018.

[88] A. Sundar and V. Pahwa, "Evaluating the performance of state of the art algorithms for enhancement of seismocardiogram signals," in *Proceedings of the First International Conference on Intelligent Computing and Communication*. Springer, 2017, pp. 37–45.

[89] C. Yang and N. Tavassolian, "Motion noise cancellation in seismocardiographic monitoring of moving subjects," in *2015 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, Oct 2015, pp. 1–4.

[90] V. Velvizhi and E. Priya, "A preprocessing techniques for seismocardiogram signals in removing artifacts," in *International Conference on Futuristic Communication and Network Technologies*. Springer, 2020, pp. 845–853.

[91] D. Naufal, M. Pramudyo, T. L. E. Rajab, A. W. Setiawan, and T. Adiono, "The evaluation of seismocardiogram signal pre-processing using hybridized variational mode decomposition method," *Biomedical Engineering Letters*, vol. 12, no. 4, pp. 381–392, 2022.

[92] D. J. Lin, J. P. Kimball, J. Zia, V. G. Ganti, and O. T. Inan, "Reducing the impact of external vibrations on fiducial point detection in seismocardiogram signals," *IEEE Transactions on Biomedical Engineering*, vol. 69, no. 1, pp. 176–185, 2022.

[93] C. Yang and N. Tavassolian, "An independent component analysis approach to motion noise cancelation of cardio-mechanical signals," *IEEE Transactions on Biomedical Engineering*, vol. 66, no. 3, pp. 784–793, March 2019.

[94] T. Choudhary, L. N. Sharma, and M. K. Bhuyan, "Heart sound extraction from sternal seismocardiographic signal," *IEEE Signal Processing Letters*, vol. 25, no. 4, pp. 482–486, 2018.

[95] P. K. Jain and A. K. Tiwari, "An algorithm for automatic segmentation of heart sound signal acquired using seismocardiography," in *2016 International Conference on Systems in Medicine and Biology (ICSMB)*, Jan 2016, pp. 157–161.

[96] A. Taebi, "Characterization, classification, and genesis of seismocardiographic signals," 2018. [Online]. Available: https://stars.library.ucf.edu/etd/5832

[97] F. Santucci, D. Lo Presti, C. Massaroni, E. Schena, and R. Setola, "Precordial vibrations: A review of wearable systems, signal processing techniques, and main applications," *Sensors*, vol. 22, no. 15, p. 5805, 2022.

[98] F. Khosrow-khavar, K. Tavakolian, and C. Menon, "Moving toward automatic and standalone delineation of seismocardiogram signal," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Aug 2015, pp. 7163–7166.

[99] M. J. Tadi, E. Lehtonen, T. Koivisto, M. Pänkäälä, A. Paasio, and M. Teräs, "Seismocardiography: Toward heart rate variability (hrv) estimation," in *Medical Measurements and Applications (MeMeA), 2015 IEEE International Symposium on*, May 2015, pp. 261–266.

[100] A. Laurin, F. Khosrow-Khavar, A. P. Blaber, and K. Tavakolian, "Accurate and consistent automatic seismocardiogram annotation without concurrent ecg," *Physiological Measurement*, vol. 37, no. 9, pp. 1588–1604, 2016. [Online]. Available: http://stacks.iop.org/0967-3334/37/i=9/a=1588

[101] T. Choudhary, L. N. Sharma, and M. K. Bhuyan, "Automatic detection of aortic valve opening using seismocardiography in healthy individuals," *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 3, pp. 1032–1040, May 2019.

[102] T. Choudhary, M. K. Bhuyan, and L. N. Sharma, "A novel method for aortic valve opening phase detection using scg signal," *IEEE Sensors Journal*, pp. 1–1, 2019.

[103] T. Choudhary, M. Bhuyan, and L. Sharma, "Orthogonal subspace projection based framework to extract heart cycles from scg signal," *Biomedical Signal Processing and Control*, vol. 50, pp. 45 – 51, 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1746809419300059

[104] T. Choudhary, M. K. Bhuyan, and L. N. Sharma, "Delineation and analysis of seismocardiographic systole and diastole profiles," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–8, 2021.

[105] M. J. Singh, L. Sharma, and S. Dandapat, "Hilbert vibration decomposition of seismocardiogram for hr and hrv estimation," in *2022 IEEE International Conference on Signal Processing and Communications (SPCOM)*, 2022, pp. 1–5.

[106] F. Khosrow-Khavar, K. Tavakolian, and C. Menon, "Moving toward automatic and standalone delineation of seismocardiogram signal," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2015, pp. 7163–7166.

[107] H. K. Thakkar and P. K. Sahoo, "Towards automatic and fast annotation of seismocardiogram signals using machine learning," *IEEE Sensors Journal*, vol. 20, no. 5, pp. 2578–2589, 2020.

[108] V. Zakeri and K. Tavakolian, "Identification of respiratory phases using seismocardiogram: A machine learning approach," in *2015 Computing in Cardiology Conference (CinC)*, Sept 2015, pp. 305–308.

[109] V. Zakeri, A. Akhbardeh, N. Alamdari, R. Fazel-Rezai, M. Paukkunen, and K. Tavakolian, "Analyzing seismocardiogram cycles to identify the respiratory phases," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 8, pp. 1786–1792, 2016.

[110] B. E. Solar, A. Taebi, and H. A. Mansy, "Classification of seismocardiographic cycles into lung volume phases," in *2017 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*. IEEE, 2017, pp. 1–2.

[111] A. Taebi, B. E. Solar, and H. A. Mansy, "An adaptive feature extraction algorithm for classification of seismocardiographic signals," in *SoutheastCon 2018*. IEEE, 2018, pp. 1–5.

[112] O. Lahdenoja, T. Hurnanen, Z. Iftikhar, S. Nieminen, T. Knuutila, A. Saraste, T. Kiviniemi, T. Vasankari, J. Airaksinen, M. Pänkäälä, and T. Koivisto, "Atrial fibrillation detection via accelerometer and gyroscope of a smartphone," *IEEE Journal of Biomedical and Health Informatics*, vol. 22, no. 1, pp. 108–118, 2018.

[113] M. O. Poliac, J. M. Zanetty, D. Salerno, and G. L. Wilcox, "Seismocardiogram (scg) interpretation using neural networks," in *Computer-Based Medical Systems-Proceedings of the Fourth Annual IEEE Symposium*. IEEE Computer Society, 1991, pp. 288–289.

[114] M. Malcangi, H. Quan, E. Vaini, P. Lombardi, and M. Di Rienzo, "Applying the efunn evolving paradigm to the recognition of artefactual beats in continuous seismocardiogram recordings," in *Engineering Applications of Neural Networks: 18th International Conference, EANN 2017, Athens, Greece, August 25–27, 2017, Proceedings*. Springer, 2017, pp. 256–264.

[115] D. M. Salerno and J. Zanetti, "Seismocardiography for monitoring changes in left ventricular function during ischemia," *Chest*, vol. 100, no. 4, pp. 991 – 993, 1991. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0012369216324904

[116] A. W. Scherhag, S. Pfleger, C. Ceconi, W. Voelker, J. Gehring, U. Staedt, and D. L. Heene, "Evaluation of signal-averaged cardiokymography for the detection of ischemic left ventricular dysfunction," *International Journal of Cardiology*, vol. 59, no. 3, pp. 305 – 312, 1997. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S016752739702963X

[117] J. Wahlström, I. Skog, P. Händel, F. Khosrow-khavar, K. Tavakolian, P. K. Stein, and A. Nehorai, "A hidden markov model for seismocardiography," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 10, pp. 2361–2372, Oct 2017.

[118] S. Siecinski, E. J. Tkacz, and P. S. Kostka, "Comparison of hrv indices obtained from ecg and scg signals from cebs database," *Biomedical engineering online*, vol. 18, pp. 1–15, 2019.

[119] A. Taebi, A. J. Bomar, R. H. Sandler, and H. A. Mansy, "Heart rate monitoring during different lung volume phases using seismocardiography," in *SoutheastCon 2018*. IEEE, 2018, pp. 1–5.

[120] F. Cocconcelli, N. Mora, G. Matrella, and P. Ciampolini, "High-accuracy, unsupervised annotation of seismocardiogram traces for heart rate monitoring," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 9, pp. 6372–6380, 2020.

[121] S. Sieciński, E. J. Tkacz, and P. S. Kostka, "Heart rate variability analysis on electrocardiograms, seismocardiograms and gyrocardiograms of healthy volunteers and patients with valvular heart diseases," *Sensors*, vol. 23, no. 4, p. 2152, 2023.

[122] P. Dehkordi, F. Khosrow-Khavar, M. Di Rienzo, O. T. Inan, S. E. Schmidt, A. P. Blaber, K. Sørensen, J. J. Struijk, V. Zakeri, P. Lombardi *et al.*, "Comparison of different methods for estimating cardiac timings: a comprehensive multimodal echocardiography investigation," *Frontiers in physiology*, vol. 10, p. 1057, 2019.

[123] M. M. H. Shandhi, B. Semiz, S. Hersek, N. Goller, F. Ayazi, and O. T. Inan, "Performance analysis of gyroscope and accelerometer sensors for seismocardiography-based wearable pre-ejection period estimation," *IEEE journal of biomedical and health informatics*, vol. 23, no. 6, pp. 2365–2374, 2019.

[124] I. Korzeniowska-Kubacka, B. Kuśmierczyk-Droszcz, M. Bilińska, B. Dobraszkiewicz-Wasilewska, K. Mazurek, and R. Piotrowicz, "Seismocardiography-a non-invasive method of assessing systolic and diastolic left ventricular function in ischaemic heart disease," *Cardiology Journal*, vol. 13, no. 4, pp. 319–325, 2006.

[125] K. Tavakolian, "Systolic time intervals and new measurement methods," *Cardiovascular Engineering and Technology*, vol. 7, no. 2, pp. 118–125, 2016. [Online]. Available: http://dx.doi.org/10.1007/s13239-016-0262-1

[126] K. Tavakolian, F. Khosrow-Khavar, B. Kajbafzadeh, M. Marzencki, S. Rohani, B. Kaminska, and C. Menon, "Seismocardiographic adjustment of diastolic timed vibrations," in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 2012, pp. 3797–3800.

[127] M. Pänkäälä, T. Koivisto, O. Lahdenoja, T. Kiviniemi, A. Saraste, T. Vasankari, and J. Airaksinen, "Detection of atrial fibrillation with seismocardiography," in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Aug 2016, pp. 4369–4374.

[128] T. Koivisto, M. Pänkäälä, T. Hurnanen, T. Vasankari, T. Kiviniemi, A. Saraste, and J. Airaksinen, "Automatic detection of atrial fibrillation using mems accelerometer," in *2015 Computing in Cardiology Conference (CinC)*, Sept 2015, pp. 829–832.

[129] K. Tavakolian, G. A. Dumont, and A. P. Blaber, "Analysis of seismocardiogram capability for prediction of mild to moderate haemorrhage; preliminary results," in *Computing in Cardiology 2013 (CinC)*, Sept 2013, pp. 1107–1110.

[130] K. Pandia, O. T. Inan, and G. T. A. Kovacs, "A frequency domain analysis of respiratory variations in the seismocardiogram signal," in *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, July 2013, pp. 6881–6884.

[131] N. Alamdari, K. Tavakolian, V. Zakeri, R. Fazel-Rezai, and A. Akhbardeh, "A morphological approach to detect respiratory phases of seismocardiogram," in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Aug 2016, pp. 4272–4275.

[132] T. Choudhary, M. Bhuyan, and L. Sharma, "Effect of respiratory effort levels on scg signals," in *The IEEE Region 10 Symposium (TENSYMP)*, 2019.

[133] D. Shao, C. Liu, and F. Tsow, "Noncontact physiological measurement using a camera: a technical review and future directions," *ACS sensors*, vol. 6, no. 2, pp. 321–334, 2020.

[134] J. Park, H. S. Seok, S.-S. Kim, and H. Shin, "Photoplethysmogram analysis and applications: An integrative review," *Frontiers in Physiology*, vol. 12, p. 808451, 2022.

[135] Y. Sun and N. Thakor, "Photoplethysmography revisited: from contact to noncontact, from point to imaging," *IEEE transactions on biomedical engineering*, vol. 63, no. 3, pp. 463–477, 2015.

[136] H. H. Asada, P. Shaltis, A. Reisner, S. Rhee, and R. C. Hutchinson, "Mobile monitoring with wearable photoplethysmographic biosensors," *IEEE engineering in medicine and biology magazine*, vol. 22, no. 3, pp. 28–40, 2003.

[137] W. B. Murray and P. A. Foster, "The peripheral pulse wave: information overlooked," *Journal of clinical monitoring*, vol. 12, pp. 365–377, 1996.

[138] J. Dorlas and J. Nijboer, "Photo-electric plethysmography as a monitoring device in anaesthesia: application and interpretation," *British journal of anaesthesia*, vol. 57, no. 5, pp. 524–530, 1985.

[139] A. A. Awad, A. S. Haddadin, H. Tantawy, T. M. Badr, R. G. Stout, D. G. Silverman, and K. H. Shelley, "The relationship between the photoplethysmographic waveform and systemic vascular resistance," *Journal of clinical monitoring and computing*, vol. 21, pp. 365–372, 2007.

[140] E. Seitsonen, I. Korhonen, M. Van Gils, M. Huiku, J. Lötjönen, K. Korttila, and A. Yli-Hankala, "Eeg spectral entropy, heart rate, photoplethysmography and motor responses to skin incision during sevoflurane anaesthesia," *Acta Anaesthesiologica Scandinavica*, vol. 49, no. 3, pp. 284–292, 2005.

[141] 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.

[142] W. M. Jubadi and S. F. A. M. Sahak, "Heartbeat monitoring alert via sms," in *2009 IEEE Symposium on Industrial Electronics & Applications*, vol. 1. IEEE, 2009, pp. 1–5.

[143] E. Gil, M. Orini, R. Bailon, J. M. Vergara, L. Mainardi, and P. Laguna, "Photoplethysmography pulse rate variability as a surrogate measurement of heart rate variability during non-stationary conditions," *Physiological measurement*, vol. 31, no. 9, p. 1271, 2010.

[144] C. C. Poon, X. Teng, Y. Wong, C. Zhang, and Y. Zhang, "Changes in the photoplethysmogram waveform after exercise," in *2004 2nd IEEE/EMBS International Summer School on Medical Devices and Biosensors*. IEEE, 2004, pp. 115–118.

[145] A. Hasanin, S. A. R. Mohamed, and A. El-Adawy, "Evaluation of perfusion index as a tool for pain assessment in critically ill patients," *Journal of clinical monitoring and computing*, vol. 31, pp. 961–965, 2017.

[146] C.-L. Chu, Y.-Y. Huang, Y.-H. Chen, L.-P. Lai, and H.-M. Yeh, "An observational study: The utility of perfusion index as a discharge criterion for pain assessment in the postanesthesia care unit," *PloS one*, vol. 13, no. 5, p. e0197630, 2018.

[147] Q. Yousef, M. Reaz, and M. A. M. Ali, "The analysis of ppg morphology: investigating the effects of aging on arterial compliance," *Measurement Science Review*, vol. 12, no. 6, pp. 266–271, 2012.

[148] U. Rubins, A. Grabovskis, J. Grube, and I. Kukulis, "Photoplethysmography analysis of artery properties in patients with cardiovascular diseases," in *14th Nordic-Baltic Conference on Biomedical Engineering and Medical Physics: NBC 2008 16–20 June 2008 Riga, Latvia*. Springer, 2008, pp. 319–322.

[149] J. Liu, B. P. Yan, Y.-T. Zhang, X.-R. Ding, P. Su, and N. Zhao, "Multi-wavelength photoplethysmography enabling continuous blood pressure measurement with compact wearable electronics," *IEEE Transactions on Biomedical Engineering*, vol. 66, no. 6, pp. 1514–1525, 2018.

[150] J. B. Dillon and A. B. Hertzman, "The form of the volume pulse in the finger pad in health, arteriosclerosis, and hypertension," *American Heart Journal*, vol. 21, no. 2, pp. 172–190, 1941.

[151] S. R. Alty, N. Angarita-Jaimes, S. C. Millasseau, and P. J. Chowienczyk, "Predicting arterial stiffness from the digital volume pulse waveform," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 12, pp. 2268–2275, 2007.

[152] H. J. Baek, J. S. Kim, Y. S. Kim, H. B. Lee, and K. S. Park, "Second derivative of photoplethysmography for estimating vascular aging," in *2007 6th International Special Topic Conference on Information Technology Applications in Biomedicine*. IEEE, 2007, pp. 70–72.

[153] Y. Aiba, S. Ohshiba, S. Horiguchi, I. MORIOKA, K. MIYASHITA, I. KIYOTA, G. ENDO, H. TAKADA, and H. IWATA, "Peripheral hemodynamics evaluated by acceleration plethysmography in workers exposed to lead," *Industrial health*, vol. 37, no. 1, pp. 3–8, 1999.

[154] T. Otsuka, T. Kawada, M. Katsumata, and C. Ibuki, "Utility of second derivative of the finger photoplethysmogram for the estimation of the risk of coronary heart disease in the general population," *Circulation Journal*, vol. 70, no. 3, pp. 304–310, 2006.

[155] J. Simek, D. Wichterle, V. Melenovsky, J. Malik, S. Svacina, and J. Widimsky, "Second derivative of the finger arterial pressure waveform: an insight into dynamics of the peripheral arterial pressure pulse," *Physiological research*, vol. 54, no. 5, p. 505, 2005.

[156] K. Takazawa, "Clinical usefulness of the second derivative of a plethysmogram (acceleration plethysmogram)," *J Cardiol*, vol. 23, pp. 207–217, 1993.

[157] G. Dahlgren, F. Veintemilla, G. Settergren, and J. Liska, "Left ventricular end-systolic pressure estimated from measurements in a peripheral artery," *Journal of cardiothoracic and vascular anesthesia*, vol. 5, no. 6, pp. 551–553, 1991.

[158] J.-L. He, Y. Lecarpentier, K. Zamani, C. Coirault, G. Daccache, D. Chemla, N. Wuilliez, L. Larsonneur *et al.*, "Relation between aortic dicrotic notch pressure and mean aortic pressure in adults," *The American journal of cardiology*, vol. 76, no. 4, pp. 301–306, 1995.

[159] S. C. Millasseau, R. Kelly, J. Ritter, and P. Chowienczyk, "Determination of age-related increases in large artery stiffness by digital pulse contour analysis," *Clinical science*, vol. 103, no. 4, pp. 371–377, 2002.

[160] D. McDuff, "Camera measurement of physiological vital signs," *ACM Computing Surveys*, vol. 55, no. 9, pp. 1–40, 2023.

[161] X. Chen, J. Cheng, R. Song, Y. Liu, R. Ward, and Z. J. Wang, "Video-based heart rate measurement: Recent advances and future prospects," *IEEE Transactions on Instrumentation and Measurement*, vol. 68, no. 10, pp. 3600–3615, 2018.

[162] V. Blazek, T. Wu, and D. Hoelscher, "Near-infrared ccd imaging: Possibilities for noninvasive and contactless 2d mapping of dermal venous hemodynamics," in *Optical Diagnostics of Biological Fluids V*, vol. 3923. SPIE, 2000, pp. 2–9.

[163] W. Verkruysse, L. O. Svaasand, and J. S. Nelson, "Remote plethysmographic imaging using ambient light." *Optics express*, vol. 16, no. 26, pp. 21 434–21 445, 2008.

[164] A. Lam and Y. Kuno, "Robust heart rate measurement from video using select random patches," in *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 3640–3648.

[165] S. Xu, L. Sun, and G. K. Rohde, "Robust efficient estimation of heart rate pulse from video," *Biomedical optics express*, vol. 5, no. 4, pp. 1124–1135, 2014.

[166] W. Wang, A. C. Den Brinker, S. Stuijk, and G. De Haan, "Algorithmic principles of remote ppg," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 7, pp. 1479–1491, 2016.

[167] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition. CVPR 2001*, vol. 1. Ieee, 2001, pp. I–I.

[168] D. E. King, "Dlib-ml: A machine learning toolkit," *The Journal of Machine Learning Research*, vol. 10, pp. 1755–1758, 2009.

[169] A. Asthana, S. Zafeiriou, S. Cheng, and M. Pantic, "Robust discriminative response map fitting with constrained local models," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2013, pp. 3444–3451.

[170] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, "Joint face detection and alignment using multitask cascaded convolutional networks," *IEEE signal processing letters*, vol. 23, no. 10, pp. 1499–1503, 2016.

[171] C. Lugaresi, J. Tang, H. Nash, C. McClanahan, E. Uboweja, M. Hays, F. Zhang, C.-L. Chang, M. G. Yong, J. Lee *et al.*, "Mediapipe: A framework for building perception pipelines," *arXiv preprint arXiv:1906.08172*, 2019.

[172] J. Shi *et al.*, "Good features to track," in *1994 Proceedings of IEEE conference on computer vision and pattern recognition*. IEEE, 1994, pp. 593–600.

[173] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *IJCAI'81: 7th international joint conference on Artificial intelligence*, vol. 2, 1981, pp. 674–679.

[174] D. Comaniciu, V. Ramesh, and P. Meer, "Kernel-based object tracking," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 25, no. 5, pp. 564–577, 2003.

[175] J. F. Henriques, R. Caseiro, P. Martins, and J. Batista, "Exploiting the circulant structure of tracking-by-detection with kernels," in *Computer Vision–ECCV 2012: 12th European Conference on Computer Vision, Florence, Italy, October 7-13, 2012, Proceedings, Part IV 12*. Springer, 2012, pp. 702–715.

[176] Y. Sun and N. Thakor, "Photoplethysmography revisited: from contact to noncontact, from point to imaging," *IEEE transactions on biomedical engineering*, vol. 63, no. 3, pp. 463–477, 2015.

[177] M.-Z. Poh, D. J. McDuff, and R. W. Picard, "Non-contact, automated cardiac pulse measurements using video imaging and blind source separation." *Optics express*, vol. 18, no. 10, pp. 10 762–10 774, 2010.

[178] M. Z. Poh, D. J. McDuff, and R. W. Picard, "Advancements in noncontact, multiparameter physiological measurements using a webcam," *IEEE transactions on biomedical engineering*, vol. 58, no. 1, pp. 7–11, 2010.

[179] G. De Haan and A. Van Leest, "Improved motion robustness of remote-ppg by using the blood volume pulse signature," *Physiological measurement*, vol. 35, no. 9, p. 1913, 2014.

[180] M. Lewandowska and J. Nowak, "Measuring pulse rate with a webcam," *Journal of Medical Imaging and Health Informatics*, vol. 2, no. 1, pp. 87–92, 2012.

[181] X. Li, J. Chen, G. Zhao, and M. Pietikainen, "Remote heart rate measurement from face videos under realistic situations," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 4264–4271.

[182] L. Feng, L.-M. Po, X. Xu, Y. Li, and R. Ma, "Motion-resistant remote imaging photoplethysmography based on the optical properties of skin," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 25, no. 5, pp. 879–891, 2014.

[183] W. Wang, S. Stuijk, and G. De Haan, "Exploiting spatial redundancy of image sensor for motion robust rppg," *IEEE transactions on Biomedical Engineering*, vol. 62, no. 2, pp. 415–425, 2014.

[184] H. Rahman, M. U. Ahmed, and S. Begum, "Non-contact physiological parameters extraction using facial video considering illumination, motion, movement and vibration," *IEEE Transactions on Biomedical Engineering*, vol. 67, no. 1, pp. 88–98, 2019.

[185] L. Shan and M. Yu, "Video-based heart rate measurement using head motion tracking and ica," in *2013 6th International Congress on Image and Signal Processing (CISP)*, vol. 1. IEEE, 2013, pp. 160–164.

[186] R. Irani, K. Nasrollahi, and T. B. Moeslund, "Improved pulse detection from head motions using dct," in *2014 international conference on computer vision theory and applications (VISAPP)*, vol. 3. IEEE, 2014, pp. 118–124.

[187] M. Lewandowska, J. Rumiński, T. Kocejko, and J. Nowak, "Measuring pulse rate with a webcam—a non-contact method for evaluating cardiac activity," in *2011 federated conference on computer science and information systems (FedCSIS)*. IEEE, 2011, pp. 405–410.

[188] D. Wedekind, A. Trumpp, F. Gaetjen, S. Rasche, K. Matschke, H. Malberg, and S. Zaunseder, "Assessment of blind source separation techniques for video-based cardiac pulse extraction," *Journal of biomedical optics*, vol. 22, no. 3, pp. 035 002–035 002, 2017.

[189] G. De Haan and V. Jeanne, "Robust pulse rate from chrominance-based rppg," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 10, pp. 2878–2886, 2013.

[190] C. A. Casado and M. B. López, "Face2ppg: An unsupervised pipeline for blood volume pulse extraction from faces," *IEEE Journal of Biomedical and Health Informatics*, 2023.

[191] S. Kwon, H. Kim, and K. S. Park, "Validation of heart rate extraction using video imaging on a built-in camera system of a smartphone," in *2012 annual international conference of the IEEE engineering in medicine and biology society*. IEEE, 2012, pp. 2174–2177.

[192] L. Kong, Y. Zhao, L. Dong, Y. Jian, X. Jin, B. Li, Y. Feng, M. Liu, X. Liu, and H. Wu, "Non-contact detection of oxygen saturation based on visible light imaging device using ambient light," *Optics express*, vol. 21, no. 15, pp. 17 464–17 471, 2013.

[193] L. A. Aarts, V. Jeanne, J. P. Cleary, C. Lieber, J. S. Nelson, S. B. Oetomo, and W. Verkruysse, "Non-contact heart rate monitoring utilizing camera photoplethysmography in the neonatal intensive care unit—a pilot study," *Early human development*, vol. 89, no. 12, pp. 943–948, 2013.

[194] F. Zhao, M. Li, Y. Qian, and J. Z. Tsien, "Remote measurements of heart and respiration rates for telemedicine," *PloS one*, vol. 8, no. 10, p. e71384, 2013.

[195] D. McDuff, S. Gontarek, and R. W. Picard, "Improvements in remote cardiopulmonary measurement using a five band digital camera," *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 10, pp. 2593–2601, 2014.

[196] J. Kuo, S. Koppel, J. L. Charlton, and C. M. Rudin-Brown, "Evaluation of a video-based measure of driver heart rate," *Journal of safety research*, vol. 54, pp. 55–e29, 2015.

[197] M. Kumar, A. Veeraraghavan, and A. Sabharwal, "Distanceppg: Robust non-contact vital signs monitoring using a camera," *Biomedical optics express*, vol. 6, no. 5, pp. 1565–1588, 2015.

[198] W. Wang, S. Stuijk, and G. de Haan, "A novel algorithm for remote photoplethysmography: Spatial subspace rotation," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 9, pp. 1974–1984, 2016.

[199] D. Shao, F. Tsow, C. Liu, Y. Yang, and N. Tao, "Simultaneous monitoring of ballistocardiogram and photoplethysmogram using a camera," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 5, pp. 1003–1010, 2016.

[200] W. Wang, A. C. den Brinker, S. Stuijk, and G. de Haan, "Robust heart rate from fitness videos," *Physiological measurement*, vol. 38, no. 6, p. 1023, 2017.

[201] C. B. Pereira, X. Yu, T. Goos, I. Reiss, T. Orlikowsky, K. Heimann, B. Venema, V. Blazek, S. Leonhardt, and D. Teichmann, "Noncontact monitoring of respiratory rate in newborn infants using thermal imaging," *IEEE transactions on Biomedical Engineering*, vol. 66, no. 4, pp. 1105–1114, 2018.

[202] K. Lee, D. K. Han, and H. Ko, "Video analytic based health monitoring for driver in moving vehicle by extracting effective heart rate inducing features," *Journal of Advanced Transportation*, vol. 2018, pp. 1–9, 2018.

[203] Q. Zhang, Y. Zhou, S. Song, G. Liang, and H. Ni, "Heart rate extraction based on near-infrared camera: Towards driver state monitoring," *IEEE Access*, vol. 6, pp. 33 076–33 087, 2018.

[204] K. Xie, C.-H. Fu, H. Liang, H. Hong, and X. Zhu, "Non-contact heart rate monitoring for intensive exercise based on singular spectrum analysis," in *2019 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR)*. IEEE, 2019, pp. 228–233.

[205] O. R. Patil, W. Wang, Y. Gao, and Z. Jin, "A camera-based pulse transit time estimation approach towards non-intrusive blood pressure monitoring," in *2019 IEEE International Conference on Healthcare Informatics (ICHI)*. IEEE, 2019, pp. 1–10.

[206] Y. Zhou, H. Ni, Q. Zhang, and Q. Wu, "The noninvasive blood pressure measurement based on facial images processing," *IEEE Sensors Journal*, vol. 19, no. 22, pp. 10 624–10 634, 2019.

[207] M. van Gastel, S. Stuijk, S. Overeem, J. P. van Dijk, M. M. van Gilst, and G. de Haan, "Camera-based vital signs monitoring during sleep–a proof of concept study," *IEEE journal of biomedical and health informatics*, vol. 25, no. 5, pp. 1409–1418, 2020.

[208] G. Scebba, G. Da Poian, and W. Karlen, "Multispectral video fusion for non-contact monitoring of respiratory rate and apnea," *IEEE Transactions on Biomedical Engineering*, vol. 68, no. 1, pp. 350–359, 2020.

[209] P.-W. Huang, B.-J. Wu, and B.-F. Wu, "A heart rate monitoring framework for real-world drivers using remote photoplethysmography," *IEEE journal of biomedical and health informatics*, vol. 25, no. 5, pp. 1397–1408, 2020.

[210] X. Yu, T. Laurentius, C. Bollheimer, S. Leonhardt, and C. H. Antink, "Noncontact monitoring of heart rate and heart rate variability in geriatric patients using photoplethysmography imaging," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 5, pp. 1781–1792, 2021.

[211] Y.-P. Yu, B.-H. Kwan, C.-L. Lim, S.-L. Wong, and P. Raveendran, "Video-based heart rate measurement using short-time fourier transform," in *2013 International Symposium on Intelligent Signal Processing and Communication Systems*. IEEE, 2013, pp. 704–707.

[212] D. McDuff, S. Gontarek, and R. W. Picard, "Remote detection of photoplethysmographic systolic and diastolic peaks using a digital camera," *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 12, pp. 2948–2954, 2014.

[213] W. Chen and D. McDuff, "Deepphys: Video-based physiological measurement using convolutional attention networks," in *Proceedings of the european conference on computer vision (ECCV)*, 2018, pp. 349–365.

[214] X. Niu, H. Han, S. Shan, and X. Chen, "Synrhythm: Learning a deep heart rate estimator from general to specific," in *2018 24th international conference on pattern recognition (ICPR)*. IEEE, 2018, pp. 3580–3585.

[215] X. Niu, X. Zhao, H. Han, A. Das, A. Dantcheva, S. Shan, and X. Chen, "Robust remote heart rate estimation from face utilizing spatial-temporal attention," in *2019 14th IEEE international conference on automatic face & gesture recognition (FG 2019)*. IEEE, 2019, pp. 1–8.

[216] Y. Qiu, Y. Liu, J. Arteaga-Falconi, H. Dong, and A. El Saddik, "Evm-cnn: Real-time contactless heart rate estimation from facial video," *IEEE transactions on multimedia*, vol. 21, no. 7, pp. 1778–1787, 2018.

[217] R. Song, S. Zhang, C. Li, Y. Zhang, J. Cheng, and X. Chen, "Heart rate estimation from facial videos using a spatiotemporal representation with convolutional neural networks," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 10, pp. 7411–7421, 2020.

[218] X. Liu, B. Hill, Z. Jiang, S. Patel, and D. McDuff, "Efficientphys: Enabling simple, fast and accurate camera-based cardiac measurement," in *Proceedings of the IEEE/CVF winter conference on applications of computer vision*, 2023, pp. 5008–5017.

[219] A. Kwasniewska, M. Szankin, J. Ruminski, A. Sarah, and D. Gamba, "Improving accuracy of respiratory rate estimation by restoring high resolution features with transformers and recursive convolutional models," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 3857–3867.

[220] Z. Yu, Y. Shen, J. Shi, H. Zhao, P. H. Torr, and G. Zhao, "Physformer: Facial video-based physiological measurement with temporal difference transformer," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2022, pp. 4186–4196.

[221] H. Lu, H. Han, and S. K. Zhou, "Dual-gan: Joint bvp and noise modeling for remote physiological measurement," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021, pp. 12 404–12 413.

[222] R. Song, H. Chen, J. Cheng, C. Li, Y. Liu, and X. Chen, "Pulsegan: Learning to generate realistic pulse waveforms in remote photoplethysmography," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 5, pp. 1373–1384, 2021.

[223] J. Gideon and S. Stent, "The way to my heart is through contrastive learning: Remote photoplethysmography from unlabelled video," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2021, pp. 3995–4004.

[224] Z. Sun and X. Li, "Contrast-phys: Unsupervised video-based remote physiological measurement via spatiotemporal contrast," in *European Conference on Computer Vision*. Springer, 2022, pp. 492–510.

[225] H.-Y. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand, and W. Freeman, "Eulerian video magnification for revealing subtle changes in the world," *ACM transactions on graphics (TOG)*, vol. 31, no. 4, pp. 1–8, 2012.

[226] H. Lu and H. Han, "Nas-hr: Neural architecture search for heart rate estimation from face videos," *Virtual Reality & Intelligent Hardware*, vol. 3, no. 1, pp. 33–42, 2021.

[227] W. Sun, Q. Sun, H.-M. Sun, Q. Sun, and R.-S. Jia, "Vit-rppg: a vision transformer-based network for remote heart rate estimation," *Journal of Electronic Imaging*, vol. 32, no. 2, pp. 023 024–023 024, 2023.

[228] X. Zhang, Z. Xia, J. Dai, L. Liu, J. Peng, and X. Feng, "Msdn: A multi-stage deep network for heart-rate estimation from facial videos," *IEEE Transactions on Instrumentation and Measurement*, 2023.

[229] S. Ding, Z. Ke, Z. Yue, C. Song, and L. Lu, "Noncontact multiphysiological signals estimation via visible and infrared facial features fusion," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1–13, 2022.

[230] R. Špetlík, V. Franc, and J. Matas, "Visual heart rate estimation with convolutional neural network," in *Proceedings of the british machine vision conference, Newcastle, UK*, 2018, pp. 3–6.

[231] Y.-Y. Tsou, Y.-A. Lee, C.-T. Hsu, and S.-H. Chang, "Siamese-rppg network: Remote photoplethysmography signal estimation from face videos," in *Proceedings of the 35th annual ACM symposium on applied computing*, 2020, pp. 2066–2073.

[232] J. Xiong, W. Ou, Z. Liu, J. Gou, W. Xiao, and H. Liu, "Graphphys: Facial video-based physiological measurement with graph neural network," *Computers and Electrical Engineering*, vol. 113, p. 109022, 2024.

[233] C. Zhao, H. Wang, and Y. Feng, "Msstnet: Multi-scale facial videos pulse extraction network based on separable spatiotemporal convolution and dimension separable attention," *Virtual Reality & Intelligent Hardware*, vol. 5, no. 2, pp. 124–141, 2023.

[234] M. Hu, F. Qian, D. Guo, X. Wang, L. He, and F. Ren, "Eta-rppgnet: Effective time-domain attention network for remote heart rate measurement," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–12, 2021.

[235] J. S. Lee, G. Hwang, M. Ryu, and S. J. Lee, "Lstc-rppg: Long short-term convolutional network for remote photoplethysmography," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 6014–6022.

[236] H. Shao, L. Luo, J. Qian, S. Chen, C. Hu, and J. Yang, "Tranphys: Spatiotemporal masked transformer steered remote photoplethysmography estimation," *IEEE Transactions on Circuits and Systems for Video Technology*, 2023.

[237] H. Kuang, C. Ao, X. Ma, and X. Liu, "Shuffle-rppgnet: Efficient network with global context for remote heart rate variability measurement," *IEEE Sensors Journal*, 2023.

[238] J. Peng, W. Su, H. Chen, J. Sun, and Z. Tian, "Cl-spo2net: Contrastive learning spatiotemporal attention network for non-contact video-based spo2 estimation," *Bioengineering*, vol. 11, no. 2, p. 113, 2024.

[239] S. K. R. Manne, S. Zhu, S. Ostadabbas, and M. Wan, "Automatic infant respiration estimation from video: A deep flow-based algorithm and a novel public benchmark," in *International Workshop on Preterm, Perinatal and Paediatric Image Analysis*. Springer, 2023, pp. 111–120.

[240] A. Jain, S. Sphurthy, P. Bharadwaj, and L. Krishna, "Evm-ippg: Design and evaluation of a vital monitoring system in nicu," in *2023 10th International Conference on Signal Processing and Integrated Networks (SPIN)*. IEEE, 2023, pp. 484–489.

[241] K. Gibson, A. Al-Naji, J. Fleet, M. Steen, A. Esterman, J. Chahl, J. Huynh, and S. Morris, "Non-contact heart and respiratory rate monitoring of preterm infants based on a computer vision system: A method comparison study," *Pediatric research*, vol. 86, no. 6, pp. 738–741, 2019.

[242] I. Lorato, S. Stuijk, M. Meftah, D. Kommers, P. Andriessen, C. van Pul, and G. de Haan, "Towards continuous camera-based respiration monitoring in infants," *Sensors*, vol. 21, no. 7, p. 2268, 2021.

[243] K. Murakami, M. Yoshioka, and J. Ozawa, "Non-contact pulse transit time measurement using imaging camera, and its relation to blood pressure," in *2015 14th IAPR International Conference on Machine Vision Applications (MVA)*. IEEE, 2015, pp. 414–417.

[244] P. Zurek, O. Krejcar, M. Penhaker, M. Cerny, and R. Frischer, "Continuous noninvasive blood pressure measurement by near infra red ccd camera and pulse transmit time systems," in *2010 Second International Conference on Computer Engineering and Applications*, vol. 2. IEEE, 2010, pp. 449–453.

[245] D. Shao, Y. Yang, C. Liu, F. Tsow, H. Yu, and N. Tao, "Noncontact monitoring breathing pattern, exhalation flow rate and pulse transit time," *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 11, pp. 2760–2767, 2014.

[246] V. Chandrasekaran, R. Dantu, S. Jonnada, S. Thiyagaraja, and K. P. Subbu, "Cuffless differential blood pressure estimation using smart phones," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 4, pp. 1080–1089, 2012.

[247] Y. Adachi, Y. Edo, R. Ogawa, R. Tomizawa, Y. Iwai, and T. Okumura, "Noncontact blood pressure monitoring technology using facial photoplethysmograms," in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2019, pp. 2411–2415.

[248] C. Lumyong, N. Yodrabum, K. Winaikosol, and T. Titijaroonroj, "Skin video-based blood pressure approximation using chrom with lstm-nn," in *2023 15th International Conference on Knowledge and Smart Technology (KST)*. IEEE, 2023, pp. 1–6.

[249] L. D. van Putten and K. E. Bamford, "Improving systolic blood pressure prediction from remote photoplethysmography using a stacked ensemble regressor," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 5956–5963.

[250] P. K. Parashiva, R. Damodaran, M. Kareem, and N. S. Narayan, "Device agnostic measurement of blood pressure from rppg signals," in *2023 31st European Signal Processing Conference (EUSIPCO)*. IEEE, 2023, pp. 1015–1019.

[251] S. Schnaubelt, M.-K. Breyer, J. Siller-Matula, and H. Domanovits, "Atrial fibrillation: a risk factor for unfavourable outcome in covid-19? a case report," *European Heart Journal-Case Reports*, vol. 4, no. FI1, pp. 1–6, 2020.

[252] Z. Yu, X. Li, and G. Zhao, "Remote photoplethysmograph signal measurement from facial videos using spatio-temporal networks," *arXiv preprint arXiv:1905.02419*, 2019.

[253] L. Tarassenko, M. Villarroel, A. Guazzi, J. Jorge, D. Clifton, and C. Pugh, "Non-contact video-based vital sign monitoring using ambient light and auto-regressive models," *Physiological measurement*, vol. 35, no. 5, p. 807, 2014.

[254] T. Vogels, M. Van Gastel, W. Wang, and G. De Haan, "Fully-automatic camera-based pulse-oximetry during sleep," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2018, pp. 1349–1357.

[255] Y. Wang, W. Wang, M. van Gastel, and G. de Haan, "Modeling on the feasibility of camera-based blood glucose measurement," in *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops*, 2019, pp. 0–0.

[256] T. T. Chowdhury, T. Mishma, S. Osman, and T. Rahman, "Estimation of blood glucose level of type-2 diabetes patients using smartphone video through pca-da," in *Proceedings of the 6th International Conference on Networking, Systems and Security*, 2019, pp. 104–108.

[257] S.-E. Kim, S.-G. Yu, N. H. Kim, K. H. Suh, and E. C. Lee, "Restoration of remote ppg signal through correspondence with contact sensor signal," *Sensors*, vol. 21, no. 17, p. 5910, 2021.

[258] D. Djeldjli, F. Bousefsaf, C. Maaoui, and F. Bereksi-Reguig, "Imaging photoplethysmography: signal waveform analysis," in *2019 10th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)*, vol. 2. IEEE, 2019, pp. 830–834.

[259] R. V. Tuckson, M. Edmunds, and M. L. Hodgkins, "Telehealth," *New England Journal of Medicine*, vol. 377, no. 16, pp. 1585–1592, 2017.

[260] O. Anton, H. Dore, E. Rendon-Morales, R. Aviles-Espinosa, P. Seddon, D. Wertheim, R. Fernandez, and H. Rabe, "Non-invasive sensor methods used in monitoring newborn babies after birth, a clinical perspective," *Maternal Health, Neonatology and Perinatology*, vol. 8, no. 1, p. 9, 2022.

[261] N. N. Sahoo, B. Murugesan, A. Das, S. Karthik, K. Ram, S. Leonhardt, J. Joseph, and M. Sivaprakasam, "Deep learning based non-contact physiological monitoring in neonatal intensive care unit," in *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE, 2022, pp. 1327–1330.

[262] N. Blanik, K. Heimann, C. Pereira, M. Paul, V. Blazek, B. Venema, T. Orlikowsky, and S. Leonhardt, "Remote vital parameter monitoring in neonatology–robust, unobtrusive heart rate detection in a realistic clinical scenario," *Biomedical Engineering/Biomedizinische Technik*, vol. 61, no. 6, pp. 631–643, 2016.

[263] R. Amelard, K. J. Pfisterer, S. Jagani, D. A. Clausi, and A. Wong, "Non-contact assessment of obstructive sleep apnea cardiovascular biomarkers using photoplethysmography imaging," in *Optical Diagnostics and Sensing XVIII: Toward Point-of-Care Diagnostics*, vol. 10501. SPIE, 2018, pp. 225–229.

[264] G. Sun, T. Negishi, T. Kirimoto, T. Matsui, and S. Abe, "Noncontact monitoring of vital signs with rgb and infrared camera and its application to screening of potential infection," *Non-invasive diagnostic methods-image processing*, 2018.

[265] R. Song, J. Li, M. Wang, J. Cheng, C. Li, and X. Chen, "Remote photoplethysmography with an eemd-mcca method robust against spatially uneven illuminations," *IEEE Sensors Journal*, vol. 21, no. 12, pp. 13 484–13 494, 2021.

[266] P. Gupta, B. Bhowmick, and A. Pal, "Exploring the feasibility of face video based instantaneous heart-rate for micro-expression spotting," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2018, pp. 1316–1323.

[267] V. Kessler, P. Thiam, M. Amirian, and F. Schwenker, "Pain recognition with camera photoplethysmography," in *2017 Seventh International Conference on Image Processing Theory, Tools and Applications (IPTA)*. IEEE, 2017, pp. 1–5.

[268] D. Huang, X. Feng, H. Zhang, Z. Yu, J. Peng, G. Zhao, and Z. Xia, "Spatio-temporal pain estimation network with measuring pseudo heart rate gain," *IEEE Transactions on Multimedia*, vol. 24, pp. 3300–3313, 2021.

[269] R. Yang, Z. Guan, Z. Yu, X. Feng, J. Peng, and G. Zhao, "Non-contact pain recognition from video sequences with remote physiological measurements prediction," *arXiv preprint arXiv:2105.08822*, 2021.

[270] R. M. Sabour, Y. Benezeth, P. De Oliveira, J. Chappe, and F. Yang, "Ubfc-phys: A multimodal database for psychophysiological studies of social stress," *IEEE Transactions on Affective Computing*, 2021.

[271] H. Monkaresi, N. Bosch, R. A. Calvo, and S. K. D'Mello, "Automated detection of engagement using video-based estimation of facial expressions and heart rate," *IEEE Transactions on Affective Computing*, vol. 8, no. 1, pp. 15–28, 2016.

[272] E. Magdalena Nowara, T. K. Marks, H. Mansour, and A. Veeraraghavan, "Sparseppg: Towards driver monitoring using camera-based vital signs estimation in near-infrared," in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, 2018, pp. 1272–1281.

[273] Q. Zhang, G.-q. Xu, M. Wang, Y. Zhou, and W. Feng, "Webcam based non-contact real-time monitoring for the physiological parameters of drivers," in *The 4th Annual IEEE International Conference on Cyber Technology in Automation, Control and Intelligent*. IEEE, 2014, pp. 648–652.

[274] B.-F. Wu, Y.-W. Chu, P.-W. Huang, and M.-L. Chung, "Neural network based luminance variation resistant remote-photoplethysmography for driver's heart rate monitoring," *IEEE Access*, vol. 7, pp. 57 210–57 225, 2019.

[275] Y.-C. Tsai, P.-W. Lai, P.-W. Huang, T.-M. Lin, and B.-F. Wu, "Vision-based instant measurement system for driver fatigue monitoring," *IEEE Access*, vol. 8, pp. 67 342–67 353, 2020.

[276] B.-F. Wu, Y.-W. Chu, P.-W. Huang, M.-L. Chung, and T.-M. Lin, "A motion robust remote-ppg approach to driver's health state monitoring," in *Computer Vision–ACCV 2016 Workshops: ACCV 2016 International Workshops, Taipei, Taiwan, November 20-24, 2016, Revised Selected Papers, Part I 13*. Springer, 2017, pp. 463–476.

[277] J. Hernandez-Ortega, S. Nagae, J. Fierrez, and A. Morales, "Quality-based pulse estimation from nir face video with application to driver monitoring," in *Pattern Recognition and Image Analysis: 9th Iberian Conference, IbPRIA 2019, Madrid, Spain, July 1–4, 2019, Proceedings, Part II 9*. Springer, 2019, pp. 108–119.

[278] U. A. Ciftci, I. Demir, and L. Yin, "Fakecatcher: Detection of synthetic portrait videos using biological signals," *IEEE transactions on pattern analysis and machine intelligence*, 2020.

[279] H. Qi, Q. Guo, F. Juefei-Xu, X. Xie, L. Ma, W. Feng, Y. Liu, and J. Zhao, "Deeprhythm: Exposing deepfakes with attentional visual heartbeat rhythms," in *Proceedings of the 28th ACM international conference on multimedia*, 2020, pp. 4318–4327.

[280] S. Fernandes, S. Raj, E. Ortiz, I. Vintila, M. Salter, G. Urosevic, and S. Jha, "Predicting heart rate variations of deepfake videos using neural ode," in *Proceedings of the IEEE/CVF international conference on computer vision workshops*, 2019, pp. 0–0.

[281] S. Liu, P. C. Yuen, S. Zhang, and G. Zhao, "3d mask face anti-spoofing with remote photoplethysmography," in *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part VII 14*. Springer, 2016, pp. 85–100.

[282] Y. Liu, A. Jourabloo, and X. Liu, "Learning deep models for face anti-spoofing: Binary or auxiliary supervision," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 389–398.

[283] S.-Q. Liu, X. Lan, and P. C. Yuen, "Remote photoplethysmography correspondence feature for 3d mask face presentation attack detection," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018, pp. 558–573.

[284] X. Li, J. Komulainen, G. Zhao, P.-C. Yuen, and M. Pietikäinen, "Generalized face anti-spoofing by detecting pulse from face videos," in *2016 23rd International Conference on Pattern Recognition (ICPR)*. IEEE, 2016, pp. 4244–4249.

[285] Z. Yu, X. Li, P. Wang, and G. Zhao, "Transrppg: Remote photoplethysmography transformer for 3d mask face presentation attack detection," *IEEE Signal Processing Letters*, vol. 28, pp. 1290–1294, 2021.

[286] B. Lin, X. Li, Z. Yu, and G. Zhao, "Face liveness detection by rppg features and contextual patch-based cnn," in *Proceedings of the 2019 3rd international conference on biometric engineering and applications*, 2019, pp. 61–68.

[287] R. M. Seepers, W. Wang, G. De Haan, I. Sourdis, and C. Strydis, "Attacks on heartbeat-based security using remote photoplethysmography," *IEEE journal of biomedical and health informatics*, vol. 22, no. 3, pp. 714–721, 2017.

[288] C.-M. Chang, C.-C. Hung, C. Zhao, C.-L. Lin, and B.-Y. Hsu, "Learning-based remote photoplethysmography for physiological signal feedback control in fitness training," in *2020 15th IEEE Conference on Industrial Electronics and Applications (ICIEA)*. IEEE, 2020, pp. 1663–1668.

[289] C. Zhao, C.-L. Lin, W. Chen, M.-K. Chen, and J. Wang, "Visual heart rate estimation and negative feedback control for fitness exercise," *Biomedical Signal Processing and Control*, vol. 56, p. 101680, 2020.

[290] (Accessed: Mar. 2019) Accelerometer data-sheet (adxl335). [Online]. Available: https://www.analog.com/media/en/technical-documentation/data-sheets/ADXL335.pdf. [Online]. Available: https://www.analog.com/media/en/technical-documentation/data-sheets/ADXL335.pdf

[291] K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," *IEEE Transactions on Signal Processing*, vol. 62, no. 3, pp. 531–544, Feb 2014.

[292] P. S. Collaboration *et al.*, "Age-specific relevance of usual blood pressure to vascular mortality: a meta-analysis of individual data for one million adults in 61 prospective studies," *The Lancet*, vol. 360, no. 9349, pp. 1903–1913, 2002.

[293] D. L. Presti, C. Massaroni, C. S. J. Leitão, M. D. F. Domingues, M. Sypabekova, D. Barrera, I. Floris, L. Massari, C. M. Oddo, S. Sales *et al.*, "Fiber bragg gratings for medical applications and future challenges: A review," *Ieee Access*, vol. 8, pp. 156 863–156 888, 2020.

[294] T.-W. Wang, W.-X. Chen, H.-W. Chu, and S.-F. Lin, "Single-channel bioimpedance measurement for wearable continuous blood pressure monitoring," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–9, 2020.

[295] J. Liu, S. Qiu, N. Luo, S.-K. Lau, H. Yu, T. Kwok, Y.-T. Zhang, and N. Zhao, "Pca-based multi-wavelength photoplethysmography algorithm for cuffless blood pressure measurement on elderly subjects," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 3, pp. 663–673, 2020.

[296] V. P. Rachim, S. Kang, J.-H. Baek, and S.-M. Park, "Unobtrusive, cuffless blood pressure monitoring using a soft polymer sensor array with flexible hybrid electronics," *IEEE Sensors Journal*, vol. 21, no. 8, pp. 10 132–10 142, 2021.

[297] M. S. Imtiaz, R. Shrestha, T. Dhillon, K. A. Yousuf, B. Saeed, A. Dinh, and K. Wahid, "Correlation between seismocardiogram and systolic blood pressure," in *Electrical and Computer Engineering (CCECE), 2013 26th Annual IEEE Canadian Conference on*, May 2013, pp. 1–4.

[298] A. M. Weissler, R. G. Peeler, and W. H. Roehll Jr, "Relationships between left ventricular ejection time, stroke volume, and heart rate in normal individuals and patients with cardiovascular disease," *American heart journal*, vol. 62, no. 3, pp. 367–378, 1961.

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

[300] M. P. Tarvainen, J.-P. Niskanen, J. A. Lipponen, P. O. Ranta-aho, and P. A. Karjalainen, "Kubios hrv - heart rate variability analysis software," *Computer Methods and Programs in Biomedicine*, vol. 113, no. 1, pp. 210 – 220, 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0169260713002599

[301] J. Kautzner and A. John Camm, "Clinical relevance of heart rate variability," *Clinical cardiology*, vol. 20, no. 2, pp. 162–168, 1997.

[302] M. Vollmer, "Hrvtool–an open-source matlab toolbox for analyzing heart rate variability," in *2019 Computing in Cardiology (CinC)*. IEEE, 2019, pp. Page–1.

[303] A. Laurin, A. Blaber, and K. Tavakolian, "Seismocardiograms return valid heart rate variability indices," in *Computing in Cardiology 2013*, Sept 2013, pp. 413–416.

[304] J. Pan and W. J. Tompkins, "A real-time qrs detection algorithm," *IEEE transactions on biomedical engineering*, no. 3, pp. 230–236, 1985.

[305] 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.

[306] A. Bhandari, "Feature scaling for machine learning: Understanding the difference between normalization vs. standardization," *Analytics Vidhya*, vol. 34, 2020.

[307] A. M. Sameh, M. Abbas, M. Hazem, and M. Abd Elazeem, "Automativ assessment of systolic cardiac performance using pep/lvet index," in *IOP Conference Series: Materials Science and Engineering*, vol. 1051, no. 1. IOP Publishing, 2021, p. 012017.

[308] O. M. Corici, C.-A. Tănasie, D. O. Alexandru, M. C. Florescu, M. V. Comănescu, K. C. Kamal, T. Ş. Ţenea-Cojan, M. Iancău, and S. N. Dinescu, "A morpho-functional study using pep/lvet ratio and global longitudinal strain in patients with dilated cardiomyopathy," *Rom J Morphol Emryol*, vol. 59, pp. 93–103, 2018.

[309] R. P. Paiva, P. Carvalho, X. Aubert, J. Muehlsteff, J. Henriques, and M. Antunes, "Assessing pep and lvet from heart sounds: algorithms and evaluation," in *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 2009, pp. 3129–3133.

[310] P. Reant, M. Dijos, E. Donal, A. Mignot, P. Ritter, P. Bordachar, P. Dos Santos, C. Leclercq, R. Roudaut, G. Habib *et al.*, "Systolic time intervals as simple echocardiographic parameters of left ventricular systolic performance: correlation with ejection fraction and longitudinal two-dimensional strain," *European Journal of Echocardiography*, vol. 11, no. 10, pp. 834–844, 2010.

[311] H.-M. Su, T.-H. Lin, P.-C. Hsu, C.-Y. Chu, W.-H. Lee, S.-C. Chen, C.-S. Lee, W.-C. Voon, W.-T. Lai, and S.-H. Sheu, "A comparison between brachial and echocardiographic systolic time intervals," *PLoS One*, vol. 8, no. 2, p. e55840, 2013.

[312] S. Lee, D. Park, and K. H. Park, "Qrs complex detection based on primitive," *Journal of communications and networks*, vol. 19, no. 5, pp. 442–450, 2017.

[313] T. K. Ho, "Random decision forests," in *Proceedings of 3rd international conference on document analysis and recognition*, vol. 1. IEEE, 1995, pp. 278–282.

[314] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 2016, pp. 785–794.

[315] J. E. Hall and A. C. Guyton, *Guyton and Hall Textbook of Medical Physiology*. Elsevier: ISBN 10;81-8147-920-3, 2010.

[316] S. S. Ulyanov and V. V. Tuchin, "Pulse-wave monitoring by means of focused laser beams scattered by skin surface and membranes," in *Static and Dynamic Light Scattering in Medicine and Biology*, vol. 1884. Spie, 1993, pp. 160–167.

[317] E. Greneker, "Radar sensing of heartbeat and respiration at a distance with applications of the technology," in *Radar 97 (Conf. Publ. No. 449)*. IET, 1997, pp. 150–154.

[318] D. Konstantin and D. Zosso, "Two-dimensional variational mode decomposition," in *Energy Minimization Methods in Computer Vision and Pattern Recognition*, vol. 8932, 2015, pp. 197–208.

[319] A. Gudigar, U. Raghavendra, E. J. Ciaccio, N. Arunkumar, E. Abdulhay, and U. R. Acharya, "Automated categorization of multi-class brain abnormalities using decomposition techniques with mri images: A comparative study," *IEEE Access*, vol. 7, pp. 28 498–28 509, 2019.

[320] Q. Xiao, J. Li, S. Wu, W. Li, L. Yang, M. Dong, and Z. Zeng, "Adaptive dspi phase denoising using mutual information and 2d variational mode decomposition," *Measurement Science and Technology*, vol. 29, no. 4, p. 045203, 2018.

[321] S. Lahmiri and M. Boukadoum, "Biomedical image denoising using variational mode decomposition," in *2014 IEEE biomedical circuits and systems conference (BioCAS) proceedings*. IEEE, 2014, pp. 340–343.

[322] R. Bhagavatula and M. Savvides, "Analyzing facial images using empirical mode decomposition for illumination artifact removal and improved face recognition," in *2007 IEEE International Conference on Acoustics, Speech and Signal Processing-ICASSP'07*, vol. 1. IEEE, 2007, pp. I–505.

[323] C. Nastar, B. Moghaddam, and A. Pentland, "Flexible images: matching and recognition using learned deformations," *Computer Vision and Image Understanding*, vol. 65, no. 2, pp. 179–191, 1997.

[324] H. Gao, L. Ma, H. Dong, J. Lu, and G. Li, "An improved two-dimensional variational mode decomposition algorithm and its application in oil pipeline image," *Systems Science & Control Engineering*, vol. 8, no. 1, pp. 297–307, 2020.

[325] E. Sidick, "Power spectral density specification and analysis of large optical surfaces," in *Modeling aspects in optical metrology II*, vol. 7390. SPIE, 2009, pp. 195–206.

[326] L. Sharma, S. Dandapat, and A. Mahanta, "Ecg signal denoising using higher order statistics in wavelet subbands," *Biomedical Signal Processing and Control*, vol. 5, no. 3, pp. 214 – 222, 2010. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1746809410000145

[327] S. Kolkur, D. Kalbande, P. Shimpi, C. Bapat, and J. Jatakia, "Human skin detection using rgb, hsv and ycbcr color models. arxiv 2017," *arXiv preprint arXiv:1708.02694*.

[328] 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.

[329] P. Van Gent, H. Farah, N. Nes, and B. van Arem, "Heart rate analysis for human factors: Development and validation of an open source toolkit for noisy naturalistic heart rate data," in *Proceedings of the 6th HUMANIST Conference*, 2018, pp. 173–178.

[330] Y.-H. Byeon, S.-B. Pan, and K.-C. Kwak, "Intelligent deep models based on scalograms of electrocardiogram signals for biometrics," *Sensors*, vol. 19, no. 4, p. 935, 2019.

[331] D. McDuff and E. Blackford, "iphys: An open non-contact imaging-based physiological measurement toolbox," in *2019 41st annual international conference of the IEEE engineering in medicine and biology society (EMBC)*. IEEE, 2019, pp. 6521–6524.

[332] H.-Y. Chiu, H.-H. Shuai, and P. C.-P. Chao, "Reconstructing qrs complex from ppg by transformed attentional neural networks," *IEEE Sensors Journal*, vol. 20, no. 20, pp. 12 374–12 383, 2020.

[333] S. Wang, L. Huang, D. Jiang, Y. Sun, G. Jiang, J. Li, C. Zou, H. Fan, Y. Xie, H. Xiong *et al.*, "Improved multi-stream convolutional block attention module for semg-based gesture recognition," *Frontiers in Bioengineering and Biotechnology*, vol. 10, p. 909023, 2022.