

DEVELOPMENT OF AN EFFICIENT ALGORITHM FOR HEART RATE ESTIMATION USING REMOTE PHOTOPLETHYSMOGRAPHY SIGNAL

**A Thesis Submitted
in Partial Fulfillment of the Requirements for the
Degree of**

MASTER OF TECHNOLOGY
in
Signal Processing and Digital Design
by

RITIKA GUPTA
(Roll No. 2K23/SPD/06)

Under the Supervision of

Dr. M. S. Choudhary
Professor, ECE, DTU

Dr. Manjeet Kumar
Assistant Professor, ECE, DTU



Department of Electronics and Communication Engineering

DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Main Bawana Road, Delhi-110042. India

May, 2025

ACKNOWLEDGEMENTS

I want to intimate my heartfelt thanks to my project guide, Dr. M. S. Choudhary, Professor and Dr. Manjeet Kumar, Assistant Professor of Department of Electronics and Communication Engineering of Delhi Technological University, for his tremendous support and assistance base on their knowledge. I am so grateful to them for assisting me with the all the necessary tools for the completion of the project. I also want to extend my heartfelt gratitude to all those who have supported my research on Heart rate variability of rPPG signal. I am grateful to the open-source community for developing and maintaining userfriendly deep learning frameworks for simplifying the implementation of the research. I specially feel very thankful for our parents, friends, and classmates for their support throughout my project period. Finally, I express my gratitude to everyone for supporting me directly or indirectly in completing this project successfully. Your support and inspiration have been truly invaluable, which encourages me.

Ritika Gupta
(2K23/SPD/06)

Dedicated

To

My Family & Friends



DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Shahbad Daultapur, Main Bawana Road, Delhi-42

CANDIDATE'S DECLARATION

I Ritika Gupta, hereby certify that the work which is being presented in the thesis entitled "Development of an Efficient Algorithm for Heart Rate Estimation using remote Photoplethysmography Signal" in partial fulfillment of the requirements for the award of the Degree of Master of Technology, submitted in the Department of Electronics and Communication Engineering, Delhi Technological University is an authentic record of my own work carried out during the period from August 2023 to May 2025 under the supervision of Dr. M. S. Choudhary, Professor and Dr. Manjeet Kumar, Assistant Professor of Department of Electronics and Communication Engineering, Delhi Technological University.

The matter presented in the thesis has not been submitted by me for the award of any other degree of this or any other Institute.

Candidate's Signature



DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Shahbad Daultapur, Main Bawana Road, Delhi-42

CERTIFICATE BY THE SUPERVISOR(s)

Certified that **Ritika Gupta** (2K23/SPD/06) has carried out their search work presented in this thesis entitled “**Development of an Efficient Algorithm for Heart Rate Estimation using remote Photoplethysmography signal**” for the award of **Master of Technology** from Department of Electronics and Communication Engineering, Delhi Technological University, Delhi, under our supervision. The thesis embodies results of original work, and studies are carried out by the student herself and the contents of the thesis do not form the basis for the award of any other degree to the candidate or to anybody else from this or any other University/Institution.

A handwritten signature in blue ink, appearing to read 'M. S. Choudhary', with a horizontal line underneath.

Signature

(Dr. M. S. Choudhary)

(Professor)

(DTU, Shahbad Daultapur,
Main Bawana Road,
Delhi-42)

Signature

(Dr. Manjeet Kumar)

(Assistant Professor)

(DTU, Shahbad Daultapur,
Main Bawana Road,
Delhi-42)

Date:

Development of an Efficient Algorithm for Heart Rate Estimation Using Remote Photoplethysmography Signal

Ritika Gupta

ABSTRACT

Accurate and non-contact heart rate monitoring has received substantial interest due to its usefulness in telemedicine, fitness tracking, and real-time health surveillance. Traditional contact-based methods, though reliable, pose challenges in terms of user ease, sensor placement, feasibility for remote monitoring. In light of these restrictions, this research explores a remote and scalable solution through Remote Photoplethysmography (rPPG)—a technique that estimates heart rate from face video recorded by a standard RGB camera.

The proposed method employs MediaPipe facial landmark detection to identify the facial region which are examples of stable Regions of Interest on the face that are known to have trustworthy blood volume pulse (BVP) signals. The collected RGB signals are converted into the HSV color space to increase the signal extractions's resilience, where the Saturation (S) channel is isolated due to its reduced sensitivity to illumination variations. This enhances the quality of the extracted signal. Further, a Fourier Decomposition Method (FDM) is applied to isolate heartbeat-related frequency components while reducing the impact of noise and motion artifacts. Experimental validation on publicly available UBFC-rPPG dataset demonstrates that the method performs competitively compared to existing rPPG techniques. The findings confirm that the proposed pipeline achieves consistent and reliable heart rate estimation even under realistic conditions involving movement and lighting changes.

This scientific exploration brings forth the growing body of non-contact physiological monitoring techniques, offering a non-intrusive and efficient alternative to traditional heart rate measurement tools. Future work may include enhancing motion resilience, evaluating performance across varying skin tones, integrating deep learning for automatic ROI selection and denoising, and extending the system to measure additional physiological parameters such as respiratory rate. The method holds promise for deployment in consumer health applications, mobile platforms, and large-scale telehealth systems.

LIST OF PUBLICATIONS

- [1] Ritika Gupta, Manjeet Kumar and Mahipal Singh Choudhary, “Heart Rate Estimation Using Remote PPG and Facial Landmark Features with Fourier Decomposition and ROI Optimization”, presented at the 3rd International Conference on Women Researchers in Electronics and Computing, 2025 (WREC'25), NIT Jalandhar.
(Received Best Paper Award).

- [2] Ritika Gupta, Manjeet Kumar and Mahipal Singh Choudhary, “Remote Photoplethysmography for Heart Rate Estimation: A Comprehensive Survey on Methods and Applications”, submitted at the 1st International Conference on Power and Intelligent Control System, 2025, NIT Hamirpur.
(Accepted).

TABLE OF CONTENTS

Title	Page No.
ACKNOWLEDGMENTS	ii
CANDIDATE's DECLARATION	iv
CERTIFICATE BY THE SUPERVISOR(s)	v
ABSTRACT	vi
LIST OF PUBLICATIONS	vii
TABLE OF CONTENTS	viii
LIST OF TABLES	x
LIST OF FIGURES	xi
LIST OF ABBREVIATIONS	xii
CHAPTER 1 INTRODUCTION	1
1.1 PRINCIPLE BEHIND rPPG	2
1.2 HEART RATE ESTIMATION USING FACIAL VIDEOS	3
1.3 APPLICATIONS	4
1.3.1 PEDIATRIC HEALTH MONITORING	4
1.3.2 SMARTPHONE-BASED VITAL MONITORING	5
1.3.3 SKIN MICROCIRCULATION ANALYSIS	5
1.3.4 TELEHEALTH AND REMOTE PATIENT MONITORING	5
1.3.5 MENTAL HEALTH AND STRESS MONITORING	6
1.3.6 SPORTS AND FITNESS MONITORING	6
CHAPTER 2 LITERATURE SURVEY	7
2.1 CONVENTIONAL SIGNAL PROCESSING TECHNIQUES	8
2.1.1 DISCRETE WAVELET TRANSFORM	8
2.1.2 ADAPTIVE FILTERING	8
2.1.3 KALMAN FILTERING	8
2.1.4 BANDPASS FILTERING	9
2.1.5 FAST FOURIER TRANSFORM FILTERING	9
2.1.6 ICA AND PCA FILTERS	9
2.1.7 SAVITZKY-GOLAY FILTERING	10
2.2 ADVANCEMENTS IN DEEP LEARNING TECHNIQUES FOR rPPG	10
CHAPTER 3 PROPOSED METHOD FOR REMOTE PHOTOPLETHYSMOGRAPHY BASED HEART RATE DETECTION	15
3.1 FACE DETECTION AND TRACKING	16
3.2 SKIN DETECTION AND SUB-ROIS EXTRACTION	16
3.3 COLOR SPACE TRANSFORMATION	18
3.3.1 RGB TO HSV CONVERSION	19
3.3.2 SIGNIFICANCE OF THE SATURATION (S) CHANNEL IN HEART RATE ESTIMATION	20

Title	Page No.
3.3.3 ADVANTAGES OF HSV OVER RGB IN HR ESTIMATION	20
3.4 FOURIER DECOMPOSITION METHOD	21
CHAPTER 4 EXPERIMENTAL ANALYSIS	25
4.1 DATABASE DESCRIPTION	25
4.1.1 DATASET COMPOSITION AND ACQUISITION SETUP	25
4.1.2 GROUND-TRUTH PPG SIGNAL ACQUISITION	26
4.1.3 FRAME EXTRACTION	26
4.1.4 SIGNIFICANCE OF THE UBFC-rPPG DATASET	27
4.2 PERFORMANCE METRICS	27
4.2.1 MEAN ABSOLUTE ERROR (MAE)	27
4.2.2 ROOT MEAN SQUARE ERROR (RMSE)	28
4.2.3 MEAN ERROR RATE (MER%)	28
4.2.4 PEARSON'S CORRELATION COEFFICIENT	28
4.2.5 WINDOWING TECHNIQUE FOR HEART RATE ESTIMATION	29
4.2.6 HEART RATE CALCULATION USING FFT	29
4.3 RESULTS	29
4.3.1 PEARSON'S CORRELATION COEFFICIENT (CC)	30
4.3.2 OUTLIER DETECTION AND ERROR CONSISTENCY	32
CHAPTER 5 FINAL ANALYSIS: DISCUSSION, CONCLUSION, AND FUTURE DIRECTIONS	33
5.1 DISCUSSION	33
5.1.1 ANALYSIS AND INTERPRETATION	33
5.1.2 COMPUTATIONAL COST ANALYSIS	34
5.2 CONCLUSIONS	35
5.3 FUTURE SCOPE	36
5.3.1 ROBUSTNESS AGAINST MOTION ARTIFACTS	36
5.3.2 MULTI-MODAL SIGNAL FUSION	36
5.3.3 EXPLORATION ACROSS DIFFERENT SKIN TONES AND ETHNICITIES	37
5.3.4 AUTOMATED SUB-ROI SELECTION AND NOISE REDUCTION	37
5.3.5 DEEP LEARNING BASED ARTIFACT REMOVAL	37
5.3.6 INTEGRATION WITH REAL-TIME APPLICATIONS	37
REFERENCES	39
LIST OF PUBLICATIONS AND THEIR PROOFS	44

LIST OF TABLES

Title	Page No.
Table 2.1 Analysis of various methods depicting the evolution of rPPG studies.....	13
Table 4.1 HR estimation outcomes of our approach and a few cutting-edge techniques.....	30

LIST OF FIGURES

Title	Page No.
Figure 1.1	The dichromatic reflection model depicts the concept of remote photoplethysmography (rPPG).....2
Figure 1.2	Emerging applications of rPPG technology.....4
Figure 2.1	The flowchart of rPPG signal processing.....8
Figure 2.2	The architecture of HR-CNN.....11
Figure 2.3	The architecture of SAM-rPPGNet.....11
Figure 2.4	The basic architecture of GAN.....12
Figure 2.5	The architecture of PhysFormer.....12
Figure 2.6	The architecture of RADIANT.....12
Figure 2.7	The architecture of SiNC.....13
Figure 3.1	A graphical representation of the proposed algorithm for non-contact HR estimation.....15
Figure 3.2	An illustration of the optimal ROI selection.....17
Figure 3.3	RGB to HSV colour space conversion for enhanced BVP signal extraction.....19
Figure 3.4	The proposed methodology in operation over a dataset recording....23
Figure 4.1	The UBFC-rPPG dataset.....25
Figure 4.2	The Ground Truth PPG signal.....26
Figure 4.3	Scatter plot between the <i>href</i> and <i>hest</i> values for different recordings from the dataset.....31
Figure 4.4	Band-Altman plot describing the agreement between <i>href</i> and <i>hest</i> values for different subjects.....31

LIST OF ABBREVIATIONS

Abbreviation	Full Form
HR	Heart Rate
PPG	Photoplethysmography
rPPG	Remote Photoplethysmography
BVP	Blood Volume Pulse
ROI	Region of Interest
FDM	Fourier Decomposition Method
FFT	Fast Fourier Transform
MAE	Mean Absolute Error
RMSE	Root Mean Square Error
CC	Correlation Coefficient
MER	Mean Error Rate
HSV	Hue, Saturation, Value (Colour Space)
RGB	Red, Green, Blue (Colour Model)
SNR	Signal-to-Noise Ratio
FPS	Frames Per Second
UBFC	University of Burgundy-Franche-Comté (Dataset)
iPPG	Imaging Photoplethysmography
DL	Deep Learning
ML	Machine Learning
SpO ₂	Oxygen Saturation
GAN	Generative Adversarial Network
CNN	Convolutional Neural Network
LSTM	Long Short-Term Memory
ViT	Vision Transformer
HCI	Human-Computer Interaction
SD	Standard Deviation
LOA	Limits of Agreement

CHAPTER 1

INTRODUCTION

Estimating heart rate is a crucial physiological measure extensively utilized in healthcare and wellness applications, encompassing cardiovascular monitoring, fitness assessment, and psychophysiological evaluation. Traditional HR measuring methods mainly utilize contact-based sensors, including electrocardiography (ECG) and photoplethysmography (PPG), both requiring direct skin contact for accurate readings. However, these methods can be inconvenient for persistent tracking and observation, such as in sleep apnea studies, as they constrain movement and may cause discomfort. Moreover, attaching sensors can be particularly challenging in certain medical conditions, such as wound assessment (burns, ulcers, or trauma), fragile or sensitive skin, and newborn medical care.

Widely adopted in both clinical and non-clinical environments, the technique offers a practical solution for tracking vital physiological signals in real time. It plays an essential part in the early identification of possible health concerns and supports ongoing wellness monitoring. The technique works by shining light onto the skin and using a photodetector to identify the alterations in light transmission or reflection. These modifications result from variations in the volume of blood within the underlying vessels. With each heartbeat, blood flow changes, and it subtly changes the way light behaves as it passes through the tissue, data that is monitored and converted into a physiological signal to be analyzed [1].

According to Beer-Lambert's rule, PPG uses light-emitting diodes and, photodetector to identify blood changes in perfusion brought on by heart activity, also known as the blood volumetric pulse (BVP) [2]. In a similar vein, non-contact techniques use cameras as sensors to capture BVP under natural light. Different optical attributes of oxygenated or deoxygenated blood form the basis of this practice since they have different absorption properties of light, and oxygenated blood is lighter in quality. A remotely located RGB camera can measure periodic changes in complexion based on the difference in blood oxygenation. A BVP-like signal can be reconstructed from analyzing these subtle changes in colour with a pulse across consecutive video frames for vital sign estimation [3]. Consequently, this method is widely recognized in the literature as remote PPG (rPPG) or image PPG (iPPG).

A standard camera can be used by rPPG to capture facial footage of the subject variations in the skin colour to produce the rPPG signal [1]. The method is based on directing light onto the skin and measuring the reflection or transmission changes of the illuminance by a photodetector. The changes are the result of changes in the blood volume of the underlying vessels. Blood flow changes with each heartbeat, and it makes the light behave slightly differently as it travels through the tissue, data

measured and translated into a physiological signal to study. Additionally, the signal strength in rPPG is generally weaker than contact-based methods, requiring precise processing to maintain accuracy.

In earlier research, Verkruyse et al. [4] initially suggested the application of consumer-level cameras as a breakthrough in the development of the rPPG technique. In their initial development, they concluded that the various channels of the RGB colour space convey different levels of photoplethysmographic information, with the green channel having the most heartbeat-related signal. This lines up with the biological principle that Hb is a more absorptive medium for green light and, thus, very sensitive to variation in blood volume. The study successfully validated the potential of extracting heart rate data from standard video footage without the need for specialized equipment. Since this breakthrough, numerous approaches have been proposed for remote heart rate monitoring using rPPG, and the field continues to attract active research and innovation.

However, many factors constrain the effectiveness of rPPG; BVP are weak signals by nature, and thus anything from ambient lighting to activities resulting in facial muscle contractions from the person in question may smear the noise across the target signal that seeks to retain accuracy and stability in rPPG measures. An effective technique in this direction is to select an ROI least affected by muscle movement to improve the resistance of the signal to noise. With the advancement of face detection and landmark-based algorithms, facial skin has become the area of interest most commonly investigated in rPPG HR estimation [5].

1.1 Principle behind rPPG

Remote photoplethysmography is an optical method that makes it possible to obtain physiological parameters, such as heart rate, in a non-contact manner by measuring small differences in skin colour due to blood volume changes during the cardiac cycle [6]. The principle of the rPPG states that, with every heartbeat, the cardiovascular system's blood propagation creates periodic waves by altering the

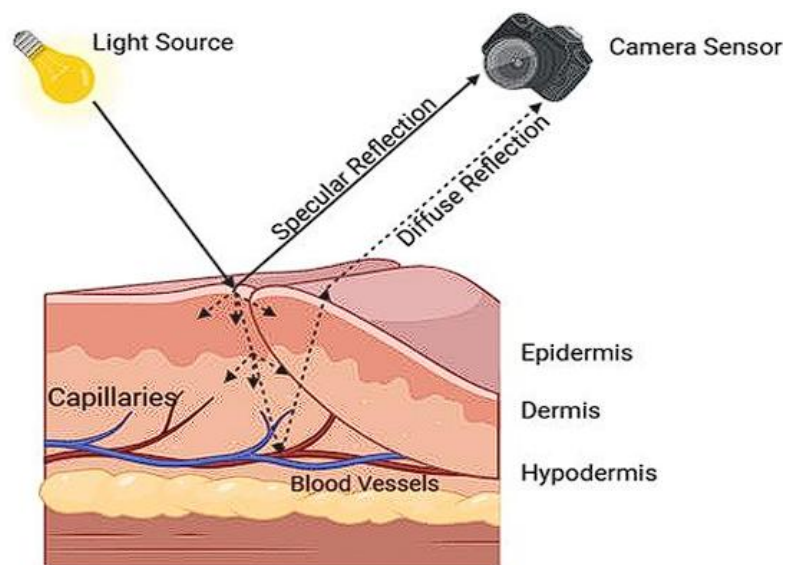


Fig. 1.1 The dichromatic reflection model depicts the concept of remote photoplethysmography (rPPG).

blood volume in the microvascular tissue bed under the skin. These variations in blood volume are brought on by pulsatile blood flow in the superficial microvascular tissues, and the modulation patterns show how light is absorbed and reflected on the skin's surface. As illustrated in Fig. 1.1, the dichromatic reflection model (DRM) [7] provides more insight into the rPPG concept. When light from the environment strikes the skin, it generates both specular and diffuse reflections. Specular reflection occurs at the surface level, reflecting light directly off the skin without carrying any physiological information. In contrast, diffuse reflection contains essential physiological signals by penetrating the skin and interacting with the blood vessels underneath. However, the video signal captured by a camera is a mixture of both reflection types. As a result, effective rPPG signal extraction relies on isolating the diffuse reflection component, which carries the meaningful data, while minimizing the influence of specular reflections. Standard cameras can capture such minor fluctuations in reflected light, which, after appropriate signal processing, can be analyzed and interpreted as meaningful physiological data [4]. Thus, it is a very efficacious technology for all applications in telemedicine, fitness, and remote health diagnostics, as it measures heart rate without contact.

In this study, the basic concepts of rPPG are exploited to obtain a very efficient heart-rate estimation technique to solve general issues in non-contact monitoring, like motion artifacts, lighting variations, and environmental noise. Thus, the algorithm makes the heart rate estimation more accurate and reliable and, hence, demonstrates the practicality of rPPG in real-world situations with efficient signal acquisition and processing techniques.

1.2 Heart rate estimation using facial videos

The enhancement in signal processing and computer vision techniques has made facial video heart rate estimation more secure in terms of accuracy and reliability. We will discuss the heart rate estimation system workflows based on rPPG. The capturing of any facial video corresponded with colour fluctuations due to incessant blood flow, while the extraction of rPPG signals was necessary to find the heart rate. For the initial face video, ROI was defined from which the temporal signals were extracted. Blind source separation (BSS) [8], periodicity measurements, or deep-learning (DL) architectures [9] are used to generate the rPPG signal from these data. Noise is reduced by filtering the generated rPPG signal [10]. HR is calculated by analyzing the frequency spectrum or R-R intervals of the rPPG signal.

Advancements in face detection and landmark algorithms have led to the inclusion of facial skin as the primary ROI in rPPG investigations [11]. Researchers are relentlessly improving ROI detection systems and investigating various ROI sites. However, several challenges have been faced by these processes, including sensitivity to facial movements, lighting variations, and different skin tones; all of these generate a considerable amount of noise and decrease the accuracy of heart rate detection. It is therefore important to ensure appropriate ROI stabilization and integrate robust signal extraction techniques for dealing with such challenges.

In response to these limitations, the proposed algorithm for facial tracking and optimized ROI extraction techniques ensures that the signal is consistently acquired even in unevenly dynamic conditions. On top of this, with better stabilization

of the pipeline for the processing of facial videos and better signal analysis techniques, this algorithm improves the performance of heart rate estimation, compared to traditional methods, to a higher extent in the case of environments with varying lighting and motion artifacts.

1.3 Applications

The evolution of rPPG has redefined the landscape of non-contact vital sign monitoring, enabling seamless measurement of physiological parameters like heart rate (HR), respiratory rate (RR), and blood oxygen saturation (SpO₂) through video-based analysis. This contactless technology leverages subtle skin colour changes captured by standard cameras, eliminating the discomfort and limitations of traditional sensors. Its versatility has paved the way for numerous real-world applications, extending from clinical health monitoring to fitness and mental health assessments. The various applications of rPPG technology are demonstrated in Fig. 1.2

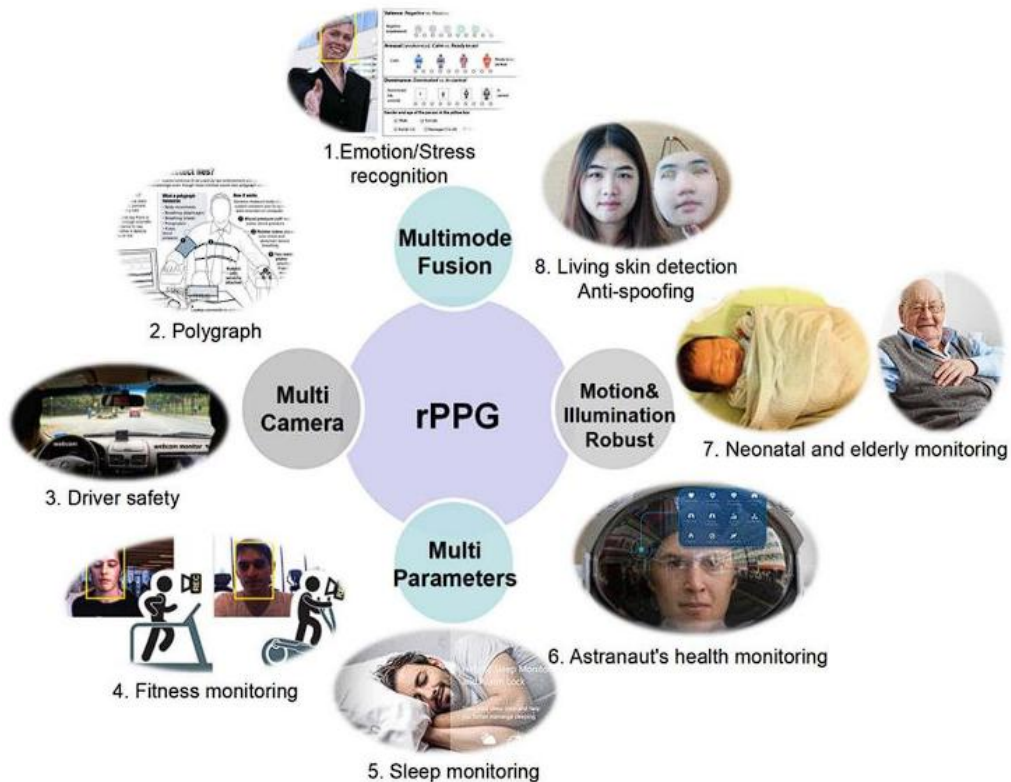


Fig. 1.2 Emerging applications of rPPG technology

1.3.1 Pediatric Health Monitoring

The rPPG technology is greatly employed in pediatric health monitoring. Usual vital sign-measuring techniques for neonates and infants tend to require uncomfortable contact sensors and sometimes, even invasive ones. On the other hand, rPPG provides contactless monitoring for heart rate, respiration rate, and oxygen saturation in children, including neonates from birth to 28 days. This approach is less

painful and purely non-invasive; hence, it reduces the likelihood of skin irritation and infection compared to conventional sensors. Besides, the acceptability and feasibility of rPPG to children and their parents point to this being an option of choice concerning continuous monitoring in NICUs and pediatric wards. Remote measurement of physiological parameters enables health professionals to observe changes in vital signs without the patients or clients experiencing physical disturbance, boosting patient comfort and allowing for frequent and standardized monitoring. This application also facilitates remote assessments, where pediatric patients can be monitored from home, reducing the need for frequent hospital visits while ensuring safety, health stability.

1.3.2 Smartphone-based Vital Monitoring

The implementation of rPPG in mobile-based applications is revolutionary in achieving pervasive health monitoring. There are indeed some good examples of the same, as we alluded to above with the WellFie app that is using the camera of your smartphone to do contactless sensing (like heart rate, respiratory rate and even blood pressure to an extent). Studies have confirmed that the readings from the app are just as accurate as readings from a conventional medical device. This innovation converts typical smartphones into very useful health monitors, and allows people to obtain his/her physiological information anywhere, anytime. This means individuals with chronic diseases, such as hypertension or cardiovascular ailments, can use real-time logging of data and prompt feedback without the hassle of wearable sensors or frequent visits to the clinic. Furthermore, the data collected can be easily shared with medical professionals for remote consultations, making it a cornerstone for mobile health (mHealth) initiatives.

1.3.3 Skin Microcirculation Analysis

Aside from breathing rate and heart rate, rPPG has also been applied to skin microcirculation research. Emerging developments led to the development of adaptive illumination and near-infrared light rPPG devices. These devices provide enhanced accuracy in capturing blood perfusion maps, which are crucial for assessing skin health and circulation. Adaptive illumination adjusts the lighting dynamically to optimize signal quality, even in suboptimal environmental conditions. Such a capability allows health practitioners to measure microvascular health quantitatively, detect pathology of blood flow, and determine the effectiveness of therapeutic intervention. Applications include diabetic foot care monitoring, pressure ulcer prevention, and skin graft monitoring, where real-time, noncontact measurement is highly beneficial.

1.3.4 Telehealth and Remote Patient Monitoring

rPPG's non-contact nature positions it as a highly valuable technology for telehealth and remote patient monitoring. As more and more healthcare becomes dependent on remote consultations, rPPG offers the capability to monitor cardiovascular health metrics in real time without the requirement of being present. This feature is particularly beneficial to the mobility-impaired, those in remote areas,

or even in the event of pandemics, when contact needs to be minimized to an absolute level. Real-time video processing allows doctors to monitor heart rate variability, identify arrhythmias, and monitor general cardiovascular health during telemedicine consultations. The simplicity of setting up a standard camera for monitoring eliminates the logistical constraints of wearable devices, offering patients a seamless experience while ensuring medical oversight remains consistent.

1.3.5 Mental Health and Stress Monitoring

In addition to physical wellness, rPPG has proven to be highly effective in the area of mental wellness and stress monitoring. Physiological indices and signals, such as heart rate variability (HRV), are highly related to the prevailing emotional states and the level of stress. Through the monitoring of changes in HRV, rPPG can serve as an indicator of psychological health. Mental well-being applications of rPPG are real-time stress measurement in stress environments, mood tracker and early detection of anxiety disorders. Wearables camera or smartphone app integration offers self-monitoring of emotional state change and adaptation of activities or routines. For clinicians, this non-contact monitoring provides a continuous stream of physiological data, offering insights into patients' mental health without intrusive procedures.

1.3.6 Sports and Fitness Monitoring

Another emerging use of rPPG is in sports and fitness applications. Conventional heart rate measurement is based on chest straps or wrist units, which are uncomfortable and cause inconvenience during stages of high-intensity activity. In contrast, rPPG enables athletes to monitor heart rate, respiratory rate and recovery time by using video-based method only. Cameras may be positioned to capture real-time performance measures that are not obtrusive to the athlete during training. This information will help you get the most out of your training, avoid overtraining, and improve your recovery techniques. Moreover, its non-contact technology is perfect for team sports where a large number of athletes are otherwise difficult to monitor at once. The versatility of remote Photoplethysmography (rPPG) extends far beyond traditional health monitoring, unlocking applications across pediatric care, telemedicine, mental health, and even fitness. Its non-contact nature, coupled with the ability to deliver real-time physiological insights, makes it a groundbreaking solution for modern health challenges. Future enhancements in multi-modal sensing, adaptive illumination, and motion stabilization are expected to further solidify its role in continuous and contactless health assessment.

CHAPTER 2

LITERATURE SURVEY

Traditional heart rate monitoring techniques, such as Electrocardiography (ECG) and Photoplethysmography (PPG), involve contact with skin, so it is not feasible in continuous and comfortable monitoring scenarios. In contrast, rPPG enables heart rate estimation by analyzing facial skin regions that reflect blood flow patterns without the need for physical sensors, making it highly suitable for telemedicine, fitness applications, and real-time health monitoring. Numerous techniques were developed to provide enhanced remote photoplethysmography (PPG) signal processing, assisting in reducing noise as well as compensating for motion artifacts. Most of the time, machine learning algorithms are used for the automatic detection of events [12]. Nevertheless, challenges continue to deter progress by a factor concerning the rPPG signal analysis since there are still many motion artifacts and ambient light within it. Thus, there is a need to address such challenges to enhance the rPPG-based health monitoring systems.

Various methods have been proposed for improving the accuracy and robustness of rPPG-based heart rate estimation, which range from elementary pixel intensity analysis to more advanced signal processing systems. Recent research has used far more sophisticated approaches to signal processing techniques, including deep learning methods, blind source separation (BSS), optical reflection models, and independent component analysis (ICA) [13], principal component analysis (PCA) [14]. All these are meant to help extract the pure rPPG signal while separating unwanted noise components. In simple terms, model-driven methods focus on breaking down RGB signals into a linear combination of the underlying blood volume pulse (BVP) and various noise elements. This approach helps to isolate noise based on certain predefined assumptions or constraints. In contrast, deep learning frameworks employ neural networks to create relationships between the video data and/or features extracted from a video frame with physiological targets such as Heart Rate (HR) or Blood Volume Pulse (BVP) waveforms. These models utilize the parameters learned from vast training datasets to predict HR components.

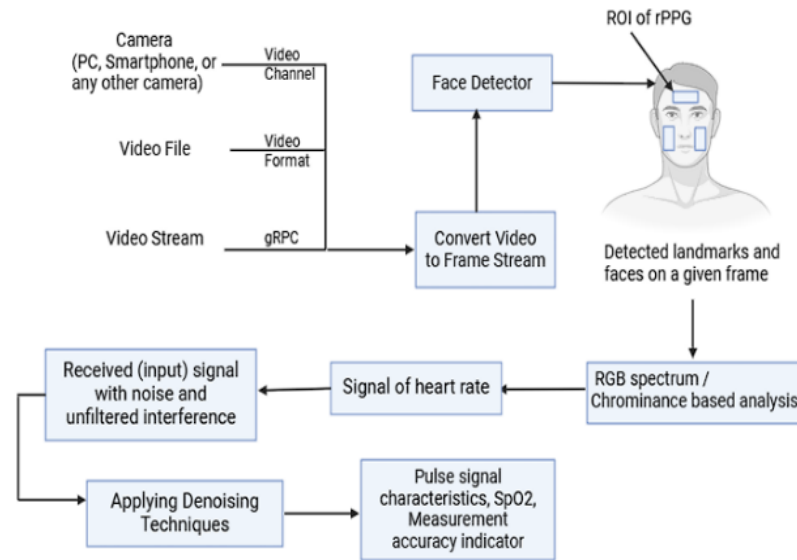


Fig 2.1 The flowchart of rPPG signal processing

2.1 Conventional Signal Processing Techniques

Under typical ambient lighting circumstances, early studies on rPPG-based heart rate measurement have produced encouraging results. However, issues including skin tone fluctuations, illumination variances, and motion disturbances can compromise measurement accuracy. Several signal-processing methods have been created to solve these issues, and their basic flow is shown in Fig. 2.1.

2.1.1 Discrete Wavelet Transform

DWT decomposes PPG signals into different frequency bands, facilitating the effective removal of motion artifacts through methods like soft thresholding [15].

2.1.2 Adaptive Filtering

By reducing the error between the PPG signal and a reference signal that correlates with motion artifacts, methods like the Least Mean Squares (LMS) adaptive filter are put into practice that not only identifies but also eliminate motion artifact [16].

2.1.3 Kalman Filtering

This recursive estimation technique adjusts filter parameters based on dynamic models of physiological processes and measurement noise, making it effective for denoising signals and estimating parameters in rPPG analysis [17].

2.1.4 Bandpass Filtering

Setting a bandpass filter is one of the core signal processing methods for analyzing rPPG signals. It helps focus on desirable frequency components, e.g., Blood Volume Pulse or Pulse Rate, while rejecting the noise from disturbance and changes in the illumination source. This, in turn, generally leads to a better estimate of physiological parameters, since it accentuates the pulsatile signal and blocks other frequencies that may contaminate its measurement. In contrast, the method will contrast with being effective since a mischoice between the cutoff frequencies will distort the signal or allow noise to enter. Also, given that it is a static filter, it is hard to incorporate adjustments when changes in environmental conditions take place very quickly, hence the need for some post-processing for artifact removal.

2.1.5 Fast Fourier Transform Filtering

Being a type of filter in the realm of signal processing, FFT helps to decompose the rPPG signals from time-domain data into their constitutive frequency components. Once the frequency peak associated with cardiac-related signals is located, parameters of the physical nature of BVP and PR can be derived accurately. However, leakage in the FFT may occur, which is highly noticeable in short-time signals or nonstationary signals. The spectral leakage mandates a proper window and resolution choice since, otherwise, artifacts could be introduced into analyses or results could be deformed. Nonetheless, FFTs are, without any question, a very powerful tool for frequency domain analysis of the rPPG [18].

2.1.6 ICA and PCA filters

Independent Component Analysis (ICA) serves as a highly competent method for decomposing multi-channel rPPG signals into statistically independent components and disentangling physiological signals from complex and noisy recordings. It stands out with the ability to separate sources, namely, hemodynamic changes and motion artifacts, based on the premise of statistical independence. Hence, ICA will provide a much stronger basis for signal extraction under conditions where sources overlap or amount to much noise. However, to a large extent, the performance of ICA depends on how accurately the independence assumption holds, whereas in cases of incorrect application, it might generate misleading components or discard relevant information. ICA can also, in high-dimensional data, be computationally intensive.

PCA, or Principal Component Analysis, is a method by which rPPG signals can be transformed into several orthogonal components that will allow the identification of dominant patterns of variation within the source. This method helps reduce dimensionality and cut down on noise while keeping important physiological references intact, making it a great option for extracting features and interpreting signals. PCA allows for the comparison of different components against a physiological parameter. However, its fundamental assumption of orthogonality doesn't apply to all rPPG signals, and using it incorrectly could result in losing some

valuable and subtle features of the signal. PCA continues to remain a handy method for improving signal clarity and interpretability despite these limitations [19].

2.1.7 Savitzky-Golay Filtering

Savitzky-Golay filtering is a broadly applied technique in rPPG analysis for the purpose of smoothing signals, attenuating high-frequency noise, and partly preserving the important characteristics of physiological signals. By applying local polynomial regression, it can efficiently enhance the signal-to-noise ratio, but is easily allowed to process data in real-time due to its low computational cost. However, it has a defined response, which can lead to phase distortion or smoothing of the signal associated with inappropriate filter parameters or non-stationary signals. Although by itself it can be effective, it will not eliminate all sources of noise, and thus will often need further pre-processing. Under careful tuning of the parameters, Savitzky-Golay filtering can indeed be considered a valuable processing tool for improving the quality of rPPG signals [20].

Implementing these advanced signal processing methods enhances the accuracy and reliability of rPPG-based HR measurements, addressing key challenges in non-invasive physiological monitoring.

2.2 Advancements in Deep Learning techniques for rPPG

Innovations in deep learning have reached the level of utmost progress for extracting PPG signals. However, most models are not generalised across skin tones or environments because the datasets are not diverse enough. Some efforts have been tried out, such as developing effective solutions using convolutional neural networks, building 3D models, and using self-attention transformers like EfficientPhys [21]. Nevertheless, remote PPG applications with deep learning are still very complicated with architecture design and data curation for reliable heart rate measurements. Table 2.1 depicts the evolution of distinct rPPG methods for HR estimation.

While rPPG methods have seen significant advancements, stability remains a challenge. Several current methodologies fail to account for the physiological aspects of facial BVP signal assessment, limiting their practical application. Variations in the geographical distribution of facial blood vessels and the temporal consistency of HR signals are of particular importance. As a result, many current methods are tailored for specific video conditions, which makes them less effective when dealing with footage that features a variety of facial characteristics, different demographic backgrounds, or real-life situations. This limitation is clear in the experimental results shared in previous studies. Different areas of the face show varying strengths of blood volume pulse (BVP) signals, and inconsistent lighting can throw off the denoising process, leading to poorer signal quality. Yet, most rPPG techniques focus on just one facial area of interest (ROI), which can compromise the reliability and precision of signal extraction.

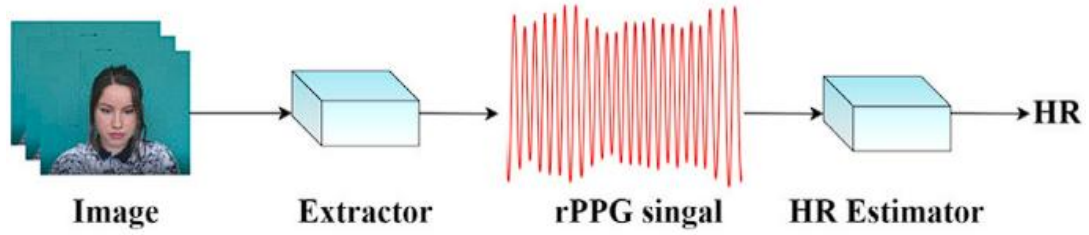


Fig. 2.2 The framework of HR-CNN

In recent years, deep learning (DL) techniques have really taken off in the realm of imaging photoplethysmography (iPPG), mainly because they excel at capturing complex spatial and temporal dynamics. These techniques generally fall into two main categories: end-to-end and modular (non-end-to-end) learning architectures. As highlighted in [23], non-end-to-end methods usually rely on spatiotemporal signal representations to pull out iPPG-related features from specific facial areas. A prime example of this approach is a two-phase convolutional neural network (CNN), which first extracts features and then estimates heart rate (HR), as shown in Fig. 2.2. Building on this idea, model called SAM-rPPGNet has been created, which includes attention mechanisms to boost accuracy even further. This architecture is made up of three main parts: a feature detection system for facial inputs, an iPPG signal derivation block, and a signal refinement unit, as depicted in Fig. 2.3.

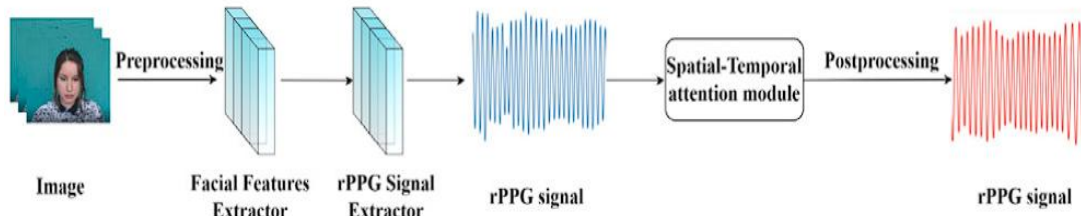


Fig. 2.3 The architecture of SAM-rPPGNet

Several advanced methods, including EVM-CNN [24], AND-rPPG [25], and DT-rPPG [26], have successfully utilized spatio-temporal maps as effective representations of heart rate, which are used as inputs to train deep learning models. While non-end-to-end models benefit from rich iPPG data and perform reliably under realistic conditions, they typically require extensive preprocessing of temporal HR features. In contrast, end-to-end models bypass this step by learning directly from raw facial video sequences. Notably, Chen and McDuff [27] introduced Deepphys—Using normalized frame differences, the first end-to-end convolutional attention model calculates pulse signals. Building on this, newer models like HR-CNN [28], PhysNet, and PhysFormer++ [29] have emerged for direct heart rate extraction from video inputs.

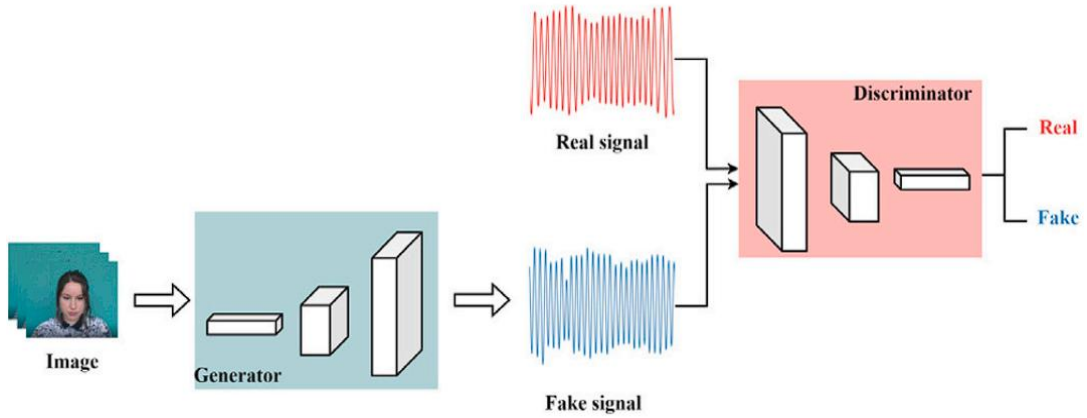


Fig. 2.4 The basic architecture of GAN

Since they first appeared in 2014, Generative Adversarial Networks (GANs) have become incredibly influential in the realms of image analysis and computer vision, thanks to their knack for creating high-quality synthetic data. A standard GAN setup consists of two main parts: a generator and a discriminator. These two networks are trained at the same time but with opposite objectives—the generator's job is to create data that closely resembles real samples, while the discriminator works to tell the difference between real and generated inputs. This ongoing rivalry drives the generator to continually enhance its outputs, leading to results that look more and more lifelike, as shown in Fig. 2.4.

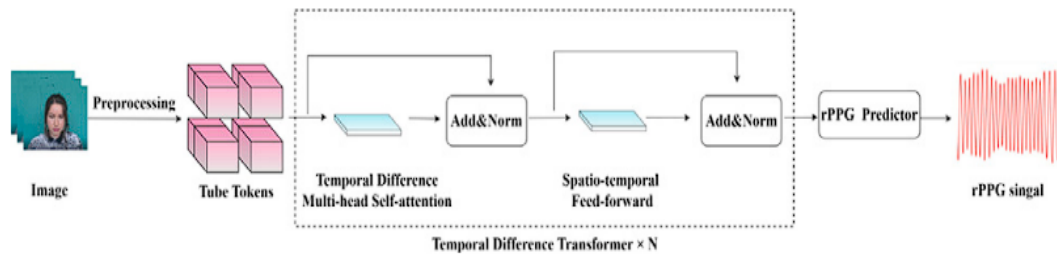


Fig. 2.5 The architecture of PhysFormer

In their research, Yu et al. [30] introduced a groundbreaking method called Physformer for estimating heart rate (HR). This approach leverages the transformer model to deliver top-notch results in analyzing remote photoplethysmography (rPPG) signals, as shown in Fig. 2.5. The heart of Physformer is its time-difference transformer mechanism, which skillfully captures long-range spatiotemporal relationships. This architecture is particularly adept at understanding both local temporal changes and wider spatiotemporal trends that are essential for precise HR estimation.

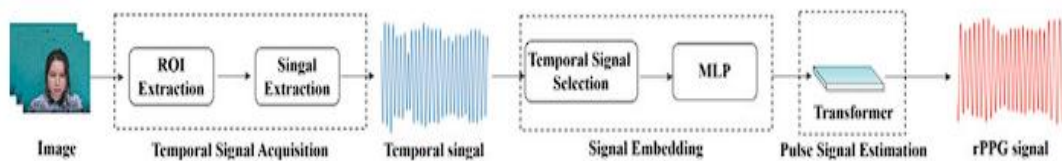


Fig. 2.6 The architecture of RADIANT

Recently, Gupta et al. [31] introduced an exciting new framework called RADIANT. This innovative approach merges signal embedding techniques with a transformer-based architecture, as illustrated in Fig. 2.6. Signal embedding enhances the portrayal of iPPG features while reducing noise interference. Their process starts with traditional methods to extract initial rPPG signals from chosen facial ROIs. These signals are then processed through Multi-Layer Perceptron (MLP) layers for embedding, which helps preserve more detailed rPPG information and improves signal quality. You can find a summary of the evolution of various rPPG-based heart rate estimation methods in Table 2.1.

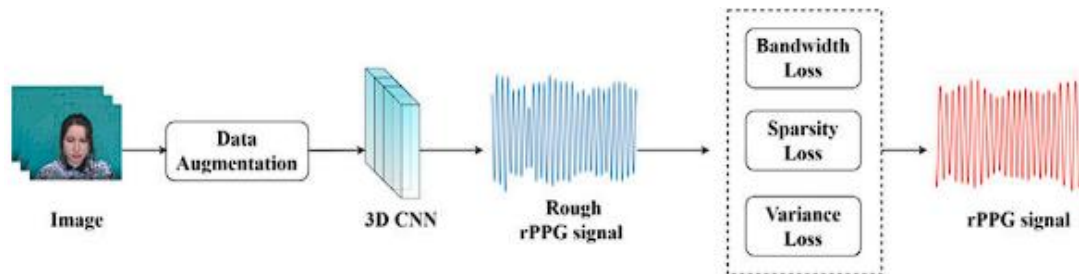


Fig. 2.7 The architecture of SiNC

Contrastive learning has proven to be highly effective for unsupervised methods; however, it also presents certain drawbacks, including high computational complexity and substantial processing costs due to the need for large sample comparisons. Addressing these limitations, Speth et al. [32] introduced SiNC, as illustrated in Fig. 2.7. Unlike traditional contrastive learning, which relies on comparing a large number of samples, SiNC leverages prior knowledge of periodic signals to achieve efficient learning without the overhead of extensive sample comparisons.

In real-world settings, DL-based approaches must account for variations in ethnicity, lighting conditions, and other environmental factors to ensure accurate results. However, capturing such diverse data is challenging due to the wide range of real-life scenarios. Additionally, physiological states vary over time and differ significantly between individuals, making it difficult to develop models that generalize well across the entire population.

Table 2.1. Analysis of various methods depicting the evolution of rPPG studies

Sr	Model	Year	Method	Description
1	MTTS-CAN [33]	2020	Attention + TSM (2D CNN)	Using an attention mechanism to steer the motion modeling process, paired with a Temporal Shift Module (TSM).
2	SAM-rPPGNet [34]	2021	Attention	Introduced a SAM for signal estimation that uses 3D CNN to learn prominent features and reduce head motion noise.

3	PulseGAN [35]	2021	CHROM (GAN)	Using GAN for signal filtering, a coarse rPPG signal can be converted into a high-fidelity rPPG signal.
4	PhysFormer++ [29]	2023	Temporal-Difference Learning + SlowFast	On top of PhysFormer, the intricate cross-speed relation in the dual-channel slowFast architectural concept is included for more dependable head motion.
5	RADIANT [36]	2023	Signal Embedding	A domain-generalized rPPG network that utilizes decoupled feature learning.
6	Multi-Task [37]	2021	Data Augmentation	A multi-task approach for extraction of rPPG signals and enhancement of data through augmentation techniques.
7	DG-rPPGNet [38]	2022	Disentangled Feature Learning	Domain generalized iPPG network that leverages disentangled feature learning to improve cross-domain performance.
8	SimPer [39]	2023	Data Augmentation + Contrastive Learning	Using Generalized Contrastive Loss and Relative Sampling Rates to Learn Robust and Effective Periodic Representations.
9	rPPG-MAE [40]	2023	Spatial-Temporal Map + MAE	The best unsupervised performance was attained by the first rPPG approach to use MAE in conjunction with the creation of a new PC-STMap.
10	APNET [41]	2022	Axis Projection	The idea of APNET is put forth, which projects films into several axes to gather data from every direction.
11	PRN augmented [32]	2022	Data Augmentation + Style Transfer	A 3D CNN-based skin tone generator that creates a uniformly dark tone from facial photos with varying skin tones.
12	SiNC [42]	2023	Data Augmentation + Penalized Regression	As the first unsupervised technique without contrastive learning, penalized regression was used to create the loss function.

CHAPTER 3

Proposed Method for Remote Photoplethysmography-based Heart Rate Detection

The key purpose of heart rate monitoring is to achieve accurate and reliable measurements in real-time, which is crucial for health monitoring, fitness tracking, and medical diagnostics. Traditional methods like ECG and PPG require direct contact with the skin, making them inconvenient for continuous and non-intrusive monitoring applications. This limitation has encouraged further investigation in remote photoplethysmography, which allows for heart rate estimation using standard RGB cameras through the detection of minor variations in skin tone on the face. However, existing rPPG methods face significant challenges, including sensitivity to motion artifacts, lighting variations, and skin tone differences, which compromise signal quality and accuracy. Additionally, the lack of robust Region of Interest (ROI) optimization and noise-filtering mechanisms further hampers real-world application. Hence, the problem is formulated as follows: How can non-contact heart rate estimation be optimized for accuracy and robustness in dynamic real-world scenarios? To address this, a novel algorithm is proposed that leverages MediaPipe facial landmark detection, HSV color space transformation, and the Fourier Decomposition Method (FDM) to enhance signal clarity, reduce noise, and provide reliable heart rate estimation across varying conditions. Fig. 3.1 shows a visual illustration of the proposed algorithm for non-contact based HR estimation.

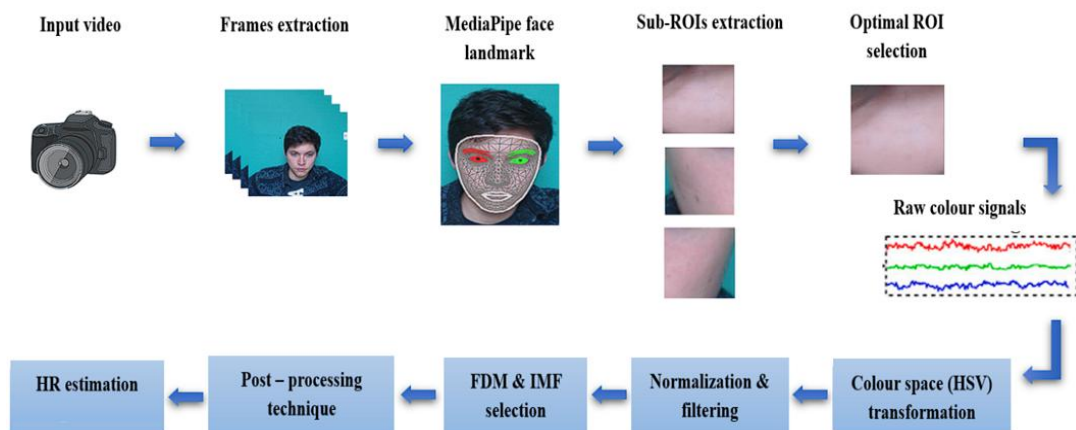


Fig. 3.1 A graphical representation of the proposed algorithm for non-contact HR estimation.

3.1 Face detection and tracking

Since, the skin pixels in each frame give pulsating cardiac information, they are included in the spatial domain of the rPPG HR estimate method. The face is typically seen in the facial recordings as an unhindered, well-lit body part, which makes it a suitable region of interest for BVP signal extraction. In this regard, the facial area in each frame must be localized. To retrieve BVP-related information, just the skin pixels are required, as the non-skin (zero intensity) pixels may induce noise to the BVP signal. This is achieved by MediaPipe Face Landmark [43], introduced by Karttynnik et al. as a technique for identifying geometry points of the face and expressions in videos and images of the face. The geometric facial mapping model employs a recurrent deep neural framework to provide ultra-fast inference performance on mobile GPUs. Fig. 3.2 (b) demonstrates that the model can infer 468 points, making it suitable for virtual try-ons, augmented reality effects, and cosmetics. Screen coordinates (x- and y-) are used to represent the position of the m^{th} landmark point, l_m . The landmark point in the t^{th} frame $I(t)$ situated at $l_m(t) = [x_m(t), y_m(t)]$. (1) shows that Inference is performed on a per-frame basis and generates a collection of landmarks, $L(t) = [l_1(t), l_2(t), \dots, l_{468}(t)]$. The m^{th} landmark detected per frame creates the m^{th} trajectory $L_m = [l_m(1), l_m(2), \dots, l_m(n)]$ of the recorded video $V = [I(1), I(2), \dots, I(n)]$. The signal's x- and y-components are built using x-coordinate signals $V = [I(1), I(2), \dots, I(n)]$, and y-coordinate signals $y = [y_m(1), y_m(2), \dots, y_m(n)]$, respectively.

$$L(t) = \text{Landmark}(I(t)) \quad (3.1)$$

ROI extraction involves identifying face landmarks and then extracting sub-ROIs from the forehead, left cheek, and right cheek regions.

3.2 Skin detection and sub-ROIs extraction

When it comes to estimating heart rate through rPPG, getting accurate skin pixel information is essential for capturing a clear and dependable Blood Volume Pulse (BVP) signal. To detect faces, we use the MediaPipe Landmark Detection Model, which skillfully identifies 468 key facial points to define different areas of the face. From these points, we focus on three main sub-ROIs: the forehead, the left cheek, and the right cheek. These specific regions are strategically chosen because they are generally less prone to movement artifacts and provide clearer signals related to blood flow variations. The forehead is especially effective due to its smooth surface and minimal muscle movement, while the cheek regions capture strong pulsatile signals from underlying vasculature.

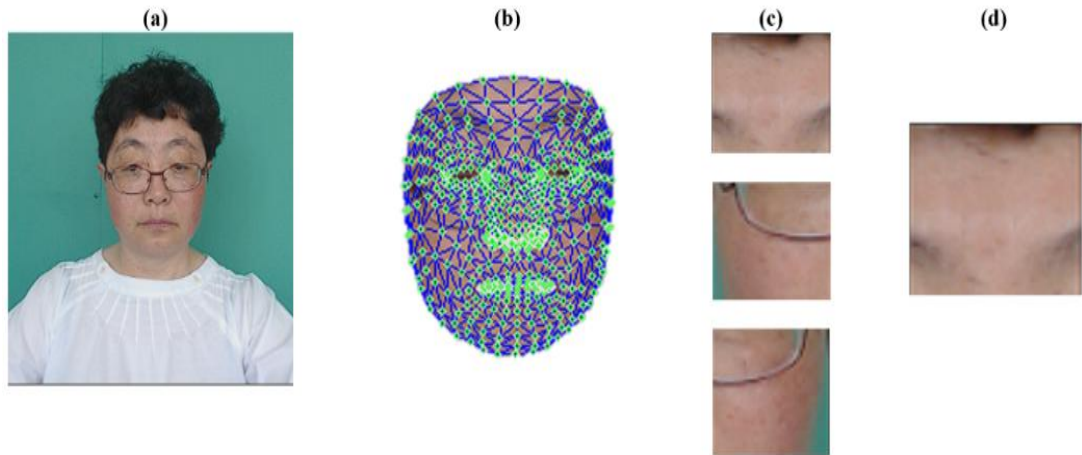


Fig. 3.2 An illustration of the optimal ROI selection. a) The subject's raw frame. b) The MediaPipe face landmark's outcome. c) Sub-ROIs were derived for the forehead, left cheek, and right cheek areas, respectively, using landmark points. d) Among the three extracted sub-ROIs, the ROI with the lowest MAE values is chosen as the best one.

Fig. 3.2 (c) shows the sub-ROI extractions utilizing landmark points. Once these sub-ROIs are extracted, it is essential to filter out non-skin elements that do not contribute to BVP signal extraction. These non-skin elements may include hair strands, spectacle frames, facial scars, and even portions of the background that are mistakenly captured during face detection.

Even with a careful choice of sub-ROIs, there might still be some non-skin pixels hanging around in the defined areas, like those pesky beard spots, reflections from glasses, or random background clutter. To tackle this issue, we use a smart strategy for selecting the best sub-ROIs. We assess how well each sub-ROI captures a clear BVP signal by looking at the Mean Absolute Error (MAE) analysis. MAE serves as our go-to performance metric to measure the noise and error levels in the signals extracted from each sub-ROI. Among the three regions, the one with the lowest MAE is selected as the optimal sub-ROI, as it represents the most noise-free and stable signal for heart rate estimation, as shown in Fig. 3.2 (d).

Once the sub-ROI is fixed, one can calculate the RGB signals— $r(t)$, $g(t)$, and $b(t)$ —for each frame by spatial averaging over pixel values within the region. Spatial averaging is often considered a noise reduction procedure since it uniformly blends any colour intensity present over the entire sub-ROI area. Meanwhile, this averaging is performed separately on red, green, and blue channel values to capture very fine changes in skin reflectance due to pulsatile blood flow. Later, these averaged RGB signals become very important for processing, such as conversion of signals into various colour spaces and signal decomposition, which are necessary for making an accurate heart rate detection.

This methodical approach—melding facial landmark detection, skin segmentation, sub-ROI optimization, and spatial averaging—ensures that only the best quality signals are sent through for heart rate estimation. By cutting down on noise and honing in on clean BVP signals, this technique greatly enhances the reliability and effectiveness of non-contact heart rate monitoring.

3.3 Color Space Transformation

Red, green, and blue signals, $r(t)$, $g(t)$, and $b(t)$, respectively, being derived from the chosen sub-ROI, the next important step deals with improving the clarity and robustness of the Blood Volume Pulse (BVP) signals that are extracted from the skin surface. The RGB colour space, while able to reasonably capture skin reflectance variations, nevertheless offers a significant disadvantage in handling fluctuations brought about by lighting conditions and movement, which could compromise the quality of the BVP signal. To counteract this problem, the RGB signals are transformed into the HSV (Hue, Saturation, Value) color space. This alternate skin-colour representation leads to a more robust skin colour variation description along with better signal quality.

The HSV color space is designed to be perceptually uniform, separating color (chromatic content) from brightness (intensity). This separation makes it easier to manipulate and analyze color information in a more intuitive way [44]. Unlike the RGB model, which intertwines color and intensity in a linear space, HSV represents colors through three separate components:

- Hue (H) : This represents the type of color (e.g., red, blue, green) and is measured in degrees on a 360° color wheel. For instance, red is at 0° , green at 120° , and blue at 240° . It is visualized as an angular dimension in the HSV cylinder, indicating the basic wavelength of light.
- Saturation (S): This characterizes the intensity or purity of the colour, ranging from 0% (completely desaturated, appearing grey) to 100% (fully saturated, vivid colour). Saturation measures how much white light is mixed with the colour—lower saturation indicates more whiteness, while higher saturation represents a purer colour.
- Value (V): This defines the brightness or luminance of the colour, varying from 0% (black) to 100% (full brightness). The Value component is analogous to light intensity, with higher values indicating lighter shades of the colour.

Geometrically, the HSV colour space is represented as a hexagonal cone (hexcone), where the Value (V) ranges vertically from black at the bottom to white at the top, representing varying intensities. The Hue (H) is arranged circularly around the cone's axis, while Saturation (S) extends radially from the centre (grey axis) to the outer boundary. As the value increases, the hexagonal cross-sections expand, representing brighter colours. The geometry of HSV allows for easier manipulation of colour properties, making it more suitable for detecting subtle skin reflectance changes caused by blood flow. RGB to HSV transformation is shown below in Fig. 3.3.

3.3.1 RGB to HSV Conversion

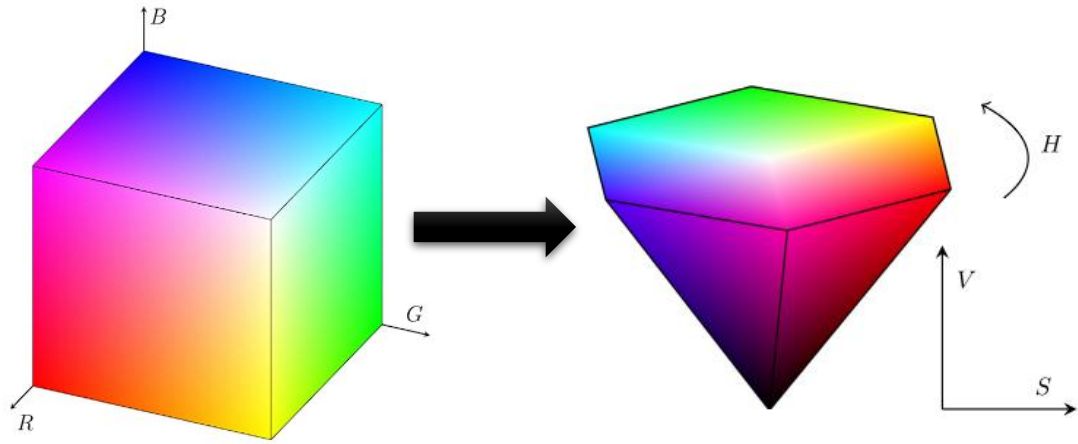


Fig. 3.3 RGB to HSV color space conversion for enhanced BVP signal extraction.

The conversion from RGB to HSV is mathematically represented as follows:

First, the RGB values are normalized to a $[0, 1]$ range:

$$R' = \frac{r}{255}; G' = \frac{g}{255}; B' = \frac{b}{255}$$

Next, the greatest and smallest values among the normalized components are identified:

$$\begin{aligned} C_{max} &= \text{MAX}(R', G', B') \\ C_{min} &= \text{MIN}(R', G', B') \end{aligned}$$

The variation between the greatest and smallest values is calculated as:

$$\Delta = C_{max} - C_{min}$$

The hue component (H) is then computed based on which of the R', G', or B' channels is dominant:

$$H = \begin{cases} 60^\circ \times \left(\frac{G' - B'}{\Delta} \bmod 6 \right) & , \quad C_{max} = R' \\ 60^\circ \times \left(\frac{B' - R'}{\Delta} + 2 \right) & , \quad C_{max} = G' \\ 60^\circ \times \left(\frac{R' - G'}{\Delta} + 4 \right) & , \quad C_{max} = B' \end{cases} \quad (3.2)$$

The saturation component (S) is calculated as follows:

$$S = \begin{cases} 0, & C_{\max} = 0 \\ \frac{\Delta}{C_{\max}}, & \text{otherwise} \end{cases} \quad (3.3)$$

Finally, the value component (V) is simply the maximum value among the three channels;

$$V = C_{\max} \quad (3.4)$$

This conversion process results in a representation where the S channel effectively isolates blood flow variations, minimizing the influence of ambient light changes.

3.3.2 Significance of the Saturation (S) Channel in Heart Rate Estimation

When it comes to the three components, the Saturation (S) channel really shines for estimating heart rate. Unlike the RGB model, where color and brightness are all tangled up together, the S channel does a great job of separating color purity from light intensity. This makes it much more reliable when the lighting conditions are all over the place, which is often the case in real life. Each heartbeat causes a slight change in the saturation of the skin color due to the pulsatile blood flow, and the S channel captures these changes beautifully. These fluctuations are directly linked to variations in blood volume, which helps us get a much clearer Blood Volume Pulse (BVP) signal.

To make sure we get accurate heart rate readings, we separate color intensity from brightness. This helps reduce the effects of changing lighting conditions, which is why the S channel is a solid option for non-contact heart rate estimation. We then run this channel through a 4th-order Butterworth bandpass filter, set to a frequency range of 0.7 to 3 Hz—perfectly aligned with typical human heart rates, which range from 42 to 180 beats per minute. This filtering process helps cut out any noise that falls outside the expected heart rate range, making the signal much clearer.

3.3.3 Advantages of HSV over RGB in HR Estimation

- **Lighting Robustness:** HSV's separation of color and brightness reduces sensitivity to illumination changes.
- **Noise Reduction:** The S channel isolates pulsatile signals effectively, filtering out irrelevant brightness changes.
- **Enhanced Signal Clarity:** The transformation improves the Signal-to-Noise Ratio (SNR), making the BVP signal more distinct.
- **Geometric Simplicity:** The cylindrical representation of HSV simplifies the processing of skin reflectance.

3.4 Fourier Decomposition Method

At this point, the selected color signal contains not just BVP information but also some unwanted components caused by time invariant and time varying sources. In our work, to denoise the BVP signal, we have used the Fourier decomposition method, which analyzes and extracts the frequency components of the colour signals. Although the Empirical Mode Decomposition (EMD) algorithm has been widely applied to non-stationary and non-linear signals [45], suffering from limitations such as mode mixing and end effects, as well as a lack of mathematical clarity, it pinpoints difficulty from such effects in the accuracy of heart rate estimates from rPPG signals. The Fourier Decomposition Method supersedes EMD by encompassing all these deficiencies. FDM provides a better adaptive and more robust approach for HR estimation since it gives much more reliable decomposition and the ability to capture all relevant frequency components that constitute rPPGs. Fourier Decomposition is a commonly used technique for adaptive narrow-band filtering of non-stationary and nonlinear signals [46]. The FDM decomposes the input data, real values and time-limited signal $x(t)$ in the interval $[t_1, t_1 + T_0]$ into a set of limited-bandwidth Fourier intrinsic band functions (FIBFs), which can then be utilized to decompose multi-component and nonstationary signals in both low and high frequency.

$$x(t) = \sum_{i=1}^M y_i(t) + a_0 \quad (3.5)$$

Here, $y_i(t)$ is the i^{th} FIBF, a_0 being the mean value of $x(t)$, respectively. A FIBF $y_i(t)$ meets the following conditions :

- Zero mean: $\int_{t_1}^{t_1+T_0} y_i(t) dt = 0$, for $\forall i$;
- Orthogonal: $\int_{t_1}^{t_1+T_0} y_i(t) y_j(t) dt = 0$, for $i \neq j$;
- Analytic representation: (h_{ref}) , where $\hat{y}_i(t)$ represents the Hilbert transform of $y_i(t)$.

In order to apply FDM, the signal $x(t)$ is periodically extended into $x_{T_0}(t) = x(t - kT_0)$, $k = 1, 2, \dots, \infty$. As such, $x(t)$ can be expressed as :

$$x(t) = x_{T_0}(t)w(t) \quad (3.6)$$

$$w(t) = \begin{cases} 1, & t_1 \leq t \leq t_1 + T_0 \\ 0, & \text{otherwise,} \end{cases}$$

Then, signal $x_{T_0}(t)$, that is periodic, is analyzed through Fourier series expansion,

i.e.

$$x_{T_0}(t) = a_0 + \sum_{k=1}^{\infty} [a_k \cos(k\omega_0 t) + b_k \sin(k\omega_0 t)] \quad (3.7)$$

Here,

$$\omega_0 = \frac{2\pi}{T_0},$$

$$a_k = \frac{2}{T_0} \int_{t_1}^{t_1+T_0} x_{T_0}(t) \cos(k\omega_0 t) dt,$$

$$b_k = \frac{2}{T_0} \int_{t_1}^{t_1+T_0} x_{T_0}(t) \sin(k\omega_0 t) dt,$$

$$a_0 = \frac{1}{T_0} \int_{t_1}^{t_1+T_0} x_{T_0}(t) dt$$

Using Euler's formula,

$$x_{T_0}(t) = a_0 + \operatorname{Re}\{z_{T_0}(t)\} \quad (3.8)$$

where ,

$$z_{T_0}(t) = \sum_{k=1}^{+\infty} c_k \exp(jk\omega_0 t) \quad (3.9)$$

Here, $c_k = a_k - jb_k$. Afterwards, for obtaining a set of FIBFs, (9) is rewritten as

$$z_{T_0}(t) = \sum_{i=1}^M a_i(t) \exp[j\phi_i(t)] \quad (3.10)$$

In our work, the selection of the most suitable Intrinsic Mode Functions (IMFs) is performed from the forward (i.e. low to high frequency scan) search of AFIBFs, aiming to enhance the extraction of relevant information from the original signal while minimizing noise. The underlying assumption is that the low-frequency IMFs most likely capture the dominant periodic components of the physiological signal, such as BVP corresponding to the heart rate, whereas the high-frequency IMFs are predominantly constituted of noise or fast oscillations that are least relevant for heart rate estimation. The forward search of AFIBFs is evaluated as follows:

$$a_i(t) \exp(j\phi_i(t)) = \sum_{k=N_{i-1}+1}^{N_i} c_k \exp(jk\omega_0 t) \quad (3.11)$$

with $N_0 = 0$ and $N_M = \infty$. The FIBFs are the real part of AFIBFs presented in (3.11).

As for the formulation of the reconstructed signal in this work, it is a combination of the recovered low-to-high IMFs, and only those IMFs carrying the essential heart-rate-related frequencies are retained. This way, the temporal and frequency properties of the original signal can be preserved, while noise and non-stationary components that could hinder the heart rate extraction task are filtered out. By virtue of the low-to-high IMFs, the proposed method efficiently removes non-stationary and noise contaminants and thereby provides a clean and reliable

representation of the heart rate signal.

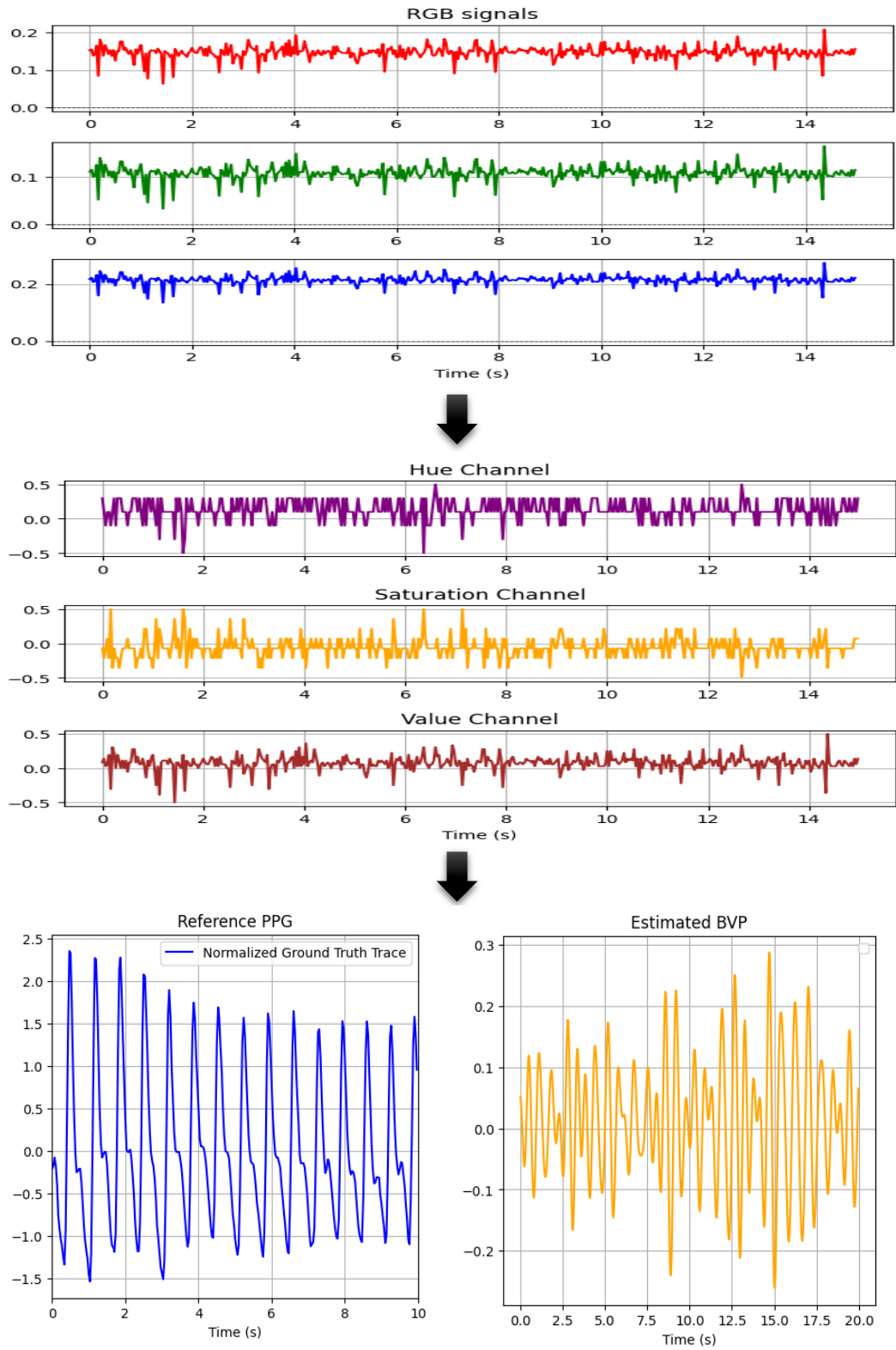


Fig. 3.4 The proposed methodology in operation over a dataset recording

Fig. 3.4 illustrates the introduced approach for heart rate estimation scheme applied to a sample video from the UBFC dataset. The first column of subplots presents the initial RGB signals (R, G, and B) extracted from the optimally selected sub-ROIs—specifically, the forehead, left cheek, and right cheek regions. These raw signals are heavily contaminated with noise and artifacts, making blood volume changes difficult to distinguish. Such distortions stem from ambient lighting fluctuations, motion artifacts, and non-skin pixels that interfere with the true photoplethysmographic signals. Consequently, the characteristic pulsatile nature of the Blood Volume Pulse (BVP) is not clearly visible in the raw RGB channels, highlighting the need for effective signal enhancement.

To address this, the RGB signals undergo a Color Space Transformation into the HSV (Hue, Saturation, Value) space. This transformation decouples color intensity from brightness, effectively separating the pulsatile components of blood flow into distinct channels. Notably, the Saturation (S) channel captures these periodic-like variations with higher clarity, as it is less sensitive to illumination changes and better at preserving blood volume pulse information. This substantiates the hypothesis that the BVP signal, which is diffused across the raw RGB channels, can be consolidated and amplified in the HSV space, enhancing its detectability.

After transforming the color space, we apply Fourier Decomposition to pinpoint the key frequencies that correspond to heartbeats. This decomposition process enhances the Blood Volume Pulse (BVP) signal by breaking it down into its core frequency bands, effectively filtering out any noise and highlighting the heart rate component. As shown in Fig. 3.4, the estimated iPPG closely mirrors the reference PPG waveform, which confirms that our method is effective.

To measure heart rate, we process the estimated BVP signal using a Fast Fourier Transform (FFT) along with a Kaiser window. The FFT process converts a signal from the time domain into its constituent frequencies by decomposing it into frequency components, allowing us to spot distinct peaks that correspond to heartbeats. In the resulting spectrum, we identify the highest peak within the physiological heart rate range of 0.7–3.5 Hz, which we label as the Heart Rate (HR) frequency. This final step enables us to estimate heart rate accurately and without any physical contact, showing a strong correlation with the actual PPG readings.

CHAPTER 4

Experimental Analysis

4.1 Database Description

The proposed heart rate estimation algorithm has been put to the test using the UBFC-rPPG dataset [47], which is a publicly accessible benchmark specifically created for analyzing heart rates through rPPG. This dataset was gathered at the University of Burgundy-Franche-Comté and includes 50 facial video recordings, divided into two main categories: SIMPLE and REALISTIC.

4.1.1 Dataset Composition and Acquisition Setup

The SIMPLE subset consists of 8 video recordings taken in a controlled environment. In this scenario, participants were asked to sit still with their eyes closed, which helped reduce head movement and outside distractions. This subset is mainly used for baseline testing, where the effects of motion artifacts and lighting changes are minimal. In contrast, the REALISTIC subset includes 42 video recordings that mimic more natural and interactive human-computer interactions. Here, participants took part in a time-sensitive math game designed to create heart rate variability. This gaming experience not only triggers physiological changes in heart rate but also encourages spontaneous head movements, making the dataset perfect for evaluating the effectiveness of non-contact heart rate estimation algorithms in real-life situations.



Fig. 4.1 The UBFC-rPPG dataset. (a) SIMPLE subset. (b) REALISTIC subset

Dataset for both subsets was collected using a Logitech C920 HD Prowebcam, set about 1 meter away from the participants. The camera recorded facial videos at a resolution of 640×480 pixels in an uncompressed 8-bit RGB format, ensuring high-quality visuals for rPPG analysis. Each video was filmed at 30 frames per second (fps), providing enough temporal resolution to capture the subtle shifts in skin color linked to blood volume pulses. The SIMPLE subset recordings lasted around 60 seconds, while the REALISTIC subset videos were up to two minutes long, allowing ample time for heart rate variations during gameplay. An example video from each of the two subsets of the UBFC-rPPG dataset can be seen in Fig. 4.1.

4.1.2 Ground-Truth PPG Signal Acquisition

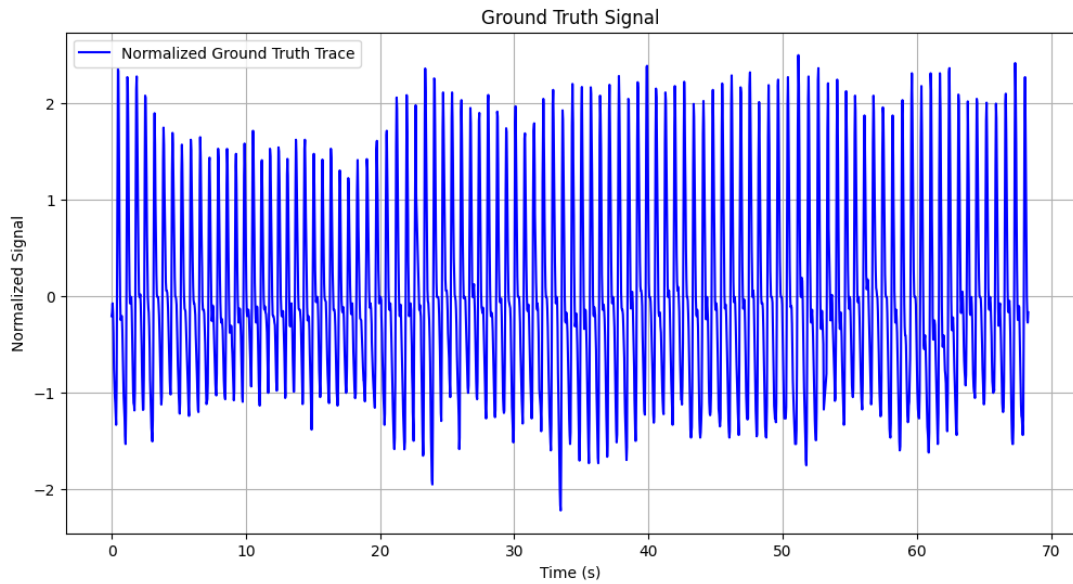


Fig. 4.2 The Ground Truth PPG signal

To ensure we have a solid reference for estimating heart rates, we recorded ground-truth Photoplethysmography (PPG) signals during each session. We used a Contec Medical CMS50E pulse oximeter to gather the PPG data at a sampling rate of 60 Hz. The device was placed on the participant's fingertip, allowing us to capture the blood volume pulse information right from the skin's surface. This ground-truth signal, illustrated in Fig. 4.2, acts as a baseline for assessing how accurate our rPPG-based heart rate estimation is. By comparing the heart rate values we extracted from the facial videos with the PPG measurements, we can effectively validate the reliability and robustness of the proposed algorithm.

4.1.3 Frame Extraction

Before diving into the analysis, we first get each video recording ready for frame extraction at the native 30 fps rate. This process breaks down the continuous video stream into individual frames, with each frame capturing a specific moment of facial skin reflectance. Next, we use the MediaPipe Landmark Detection Model to identify and extract Regions of Interest (ROIs)—specifically focusing on the forehead,

left cheek, and right cheek. Once we have those frames, we convert them into the HSV color space, isolating the Saturation (S) channel for further analysis.

The preprocessing phase also takes care of illumination normalization and motion stabilization to minimize any artifacts. Since the lighting conditions can vary quite a bit during the recordings, especially in the REALISTIC subset, normalization helps maintain a consistent signal quality across all frames. Plus, we filter out any temporal noise caused by head movements, which really boosts the clarity of the BVP signal we extract from the facial skin regions.

4.1.4 Significance of the UBFC-rPPG Dataset

The UBFC-rPPG dataset serves as a comprehensive baseline for examining the performance of algorithms built on rPPG-based heart rate estimation. In this dataset the two-part setup allows for validation in both controlled (SIMPLE) and real-world interactive (REALISTIC) scenarios assuring that comprehensive validation can be performed in ideal and practical scenarios. Furthermore, this dataset accompanies the synchronized, ground-truth PPG signals that serve as ideal metrics for accuracy and provides important benchmarks from the estimated heart rate values from rPPG when validating the estimated values. The structure of the dataset additionally including lighting variations and natural head movement will allow us to assess the ability of the proposed algorithm to cope with changes in environments and motion artifacts.

4.2 Performance Metrics

To assess the performance and reliability of the proposed heart rate estimation algorithm, several standard metrics are utilized. Additionally, a windowing method has been employed to estimate the heart rate (HR) in real time using a 20-second window updated every 2 seconds. Each metric and methods will serve as a useful indicator to evaluate the accuracy and reliability of the estimated HR signal compared to the corresponding reference HR from the ground-truth PPG signals.

4.2.1. Mean Absolute Error (MAE)

The Mean Absolute Error (MAE) calculates the average magnitude of the absolute differences between the estimated heart rate (h_{est}) and the reference heart rate (h_{ref}) derived from the ground-truth PPG signal. It is calculated using the formula:

where,

$$MAE = \frac{1}{N} \sum_{i=1}^N |h_{est,i} - h_{ref,i}| \quad (4.1)$$

h_{est} represents the estimated heart rate, and h_{ref} represents the reference heart rate.

MAE provides an intuitive measure of the error in beats per minute (bpm), reflecting the average deviation of the estimated HR from the true HR. The key advantage of MAE is its simplicity and straightforward interpretation, making it one

of the most widely used error metrics in heart rate estimation.

4.2.2. Root Mean Square Error (RMSE)

The Root Mean Square Error (RMSE) is another commonly used metric that quantifies the square root of the average squared differences between the estimated and reference heart rates. It is mathematically represented as:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (h_{est,i} - h_{ref,i})^2} \quad (4.2)$$

RMSE is more sensitive to larger errors compared to MAE because it squares the differences before averaging. This property makes it particularly useful for identifying significant deviations in heart rate estimation. A lower RMSE value indicates better model performance with minimal large deviations from the actual HR.

4.2.3. Mean Error Rate (MER%)

The Mean Error Rate (MER%) expresses the average error as a percentage of the reference heart rate, providing a normalized view of the estimation accuracy. It is calculated as follows:

$$MER\% = \left(\frac{1}{N} \right) \sum_{i=1}^N \left| \frac{(h_{est,i} - h_{ref,i})}{h_{ref,i}} \right| \times 100 \quad (4.3)$$

MER allows for a more interpretable assessment of the algorithm's performance by representing the deviation in relative terms rather than absolute terms. This is particularly useful when comparing error rates across varying heart rate ranges.

4.2.4. Pearson's Correlation Coefficient (CC)

Pearson's Correlation Coefficient calculates the linear correlation between the estimated heart rate (h_{est}) and the reference heart rate (h_{ref}), quantifying the degree to which the two variables move together. It is defined as:

where:

$$CC = \frac{(\sum (h_{est,i} - \mu_{est})(h_{ref,i} - \mu_{ref}))}{\sqrt{[\sum (h_{est,i} - \mu_{est})^2 \sum (h_{ref,i} - \mu_{ref})^2]}} \quad (4.4)$$

h_{est} and h_{ref} are the mean values of the estimated and reference heart rates, respectively.

The value of CC ranges from -1 to +1: +1 being a perfect positive correlation, 0 states no correlation, while -1 says a perfect negative correlation. A high CC value close to +1 implies that the estimated HR closely follows the fluctuations of the ground-truth HR, indicating accurate prediction.

4.2.5. Windowing Technique for Heart Rate Estimation

To achieve smooth and real-time HR estimation, the proposed method employs a windowing technique. This technique calculates the heart rate in overlapping time intervals, ensuring continuous monitoring and reducing latency. The specific configuration used is:

- Window Length: 20 seconds
- Window Update Interval: 2 seconds

The overlapping window mechanism allows the algorithm to extract heart rate estimates at fine intervals, providing real-time feedback while smoothing out short-term noise. For each 20-second window, the BVP signal is analyzed in the frequency domain.

4.2.6 Heart Rate Calculation Using FFT

Within each window segment, a Fast Fourier Transform (FFT) is applied to convert the BVP signal from the time domain to the frequency domain. This transformation reveals the dominant frequency components corresponding to heartbeats. The frequency spectrum is searched within the physiological range of 0.7 Hz to 3.5 Hz (42–210 bpm). The highest peak within this range is selected as the dominant heartbeat frequency:

$$HR (bpm) = f_{peak} \times 60 \quad (4.5)$$

where, f_{peak} is the frequency of the dominant peak detected by FFT.

To accurately identify heartbeat intervals, we use the `find_peaks` function from the SciPy library in Python. Once we detect these peaks, we can calculate the average interval between consecutive beats, which we then convert into beats per minute (bpm).

The heart rate estimation method we're proposing relies on solid evaluation metrics like MAE, RMSE, MER%, and CC to measure accuracy and reliability. Plus, we've enhanced real-time monitoring with a windowing strategy that ensures smooth heart rate detection every 2 seconds using a 20-second rolling window.

4.3 Results

The heart rate estimation method we proposed underwent a thorough evaluation using the UBFC-rPPG dataset, which is a well-known benchmark tailored for non-contact heart rate monitoring. To achieve precise and real-time heart rate detection, we utilized a sliding window approach, breaking down each recording into 20-second segments with an 18-second overlap. This overlapping structure enabled the extraction of HR values every 2 seconds, providing high temporal resolution and smooth tracking of heart rate fluctuations throughout the video sequences.

Table 4.1 HR estimation outcomes of our approach and a few cutting-edge techniques

Methods	MAE(bpm)	RMSE(bpm)	MER(%)	ρ
POS	8.35	10.00	9.85	0.24
CHROM	8.20	9.92	9.17	0.27
Green	6.01	7.87	6.48	0.29
SynRhythm	5.59	6.82	5.5	0.72
DeepPhys	4.58	14.76	4.86	0.78
RhythmNet	3.91	7.72	4.06	0.93
PhysNet	3.0	10.05	3.52	0.75
Proposed Method	1.31	1.96	1.50	0.78

The quantitative evaluation of the proposed method demonstrated a Mean Absolute Error (MAE) of 1.31 bpm and a Root Mean Square Error (RMSE) of 1.96 bpm. These results signify a high degree of accuracy, achieved improvements of 6.89 bpm, 3.27 bpm, and 1.69 bpm over CHROM, DeepPhys, and PhysNet, [48] respectively as shown in Table 4.1. The robustness of the proposed technique is attributed to the Fourier Decomposition Method (FDM), combined with HSV color space transformation, which effectively captures subtle blood volume changes while minimizing noise and motion artifacts. These values indicate precise heart rate estimation with minimal deviation from the ground-truth values obtained from PPG signals. The low MAE value reflects the method's ability to consistently produce estimates close to the actual heart rate, while the RMSE highlights its resilience in handling larger errors, which are effectively minimized through signal enhancement techniques.

4.3.1 Pearson's Correlation Coefficient (CC)

The Pearson's Correlation Coefficient (CC) was measured at 0.78, signifying a strong positive correlation between the estimated heart rate h_{est} and the reference heart rate h_{ref} . This correlation suggests that the estimated HR closely follows the fluctuations and variations observed in the reference HR. While slightly lower than ideal, the CC value demonstrates reliability across different segments of the video.

To visualize this correlation, a scatter plot as shown in Fig. 4.3 was generated, where the calculated HR values are plotted against the ground truth HR values. In the ideal scenario, all points would align perfectly along the diagonal identity line $h_{est} = h_{ref}$. The plot illustrated that the points are densely clustered around the diagonal, indicating strong agreement. Minor deviations were observed, primarily during high-motion artifacts or sudden illumination shifts, but the overall alignment validates the effectiveness of the proposed method.

To further evaluate the agreement between the estimated and reference heart rates, a Bland-Altman analysis was performed. The result is as shown in Fig. 4.4. This analysis is instrumental in identifying systematic biases and understanding the spread of error. The Bland-Altman plot revealed an average bias of -1.62 bpm,

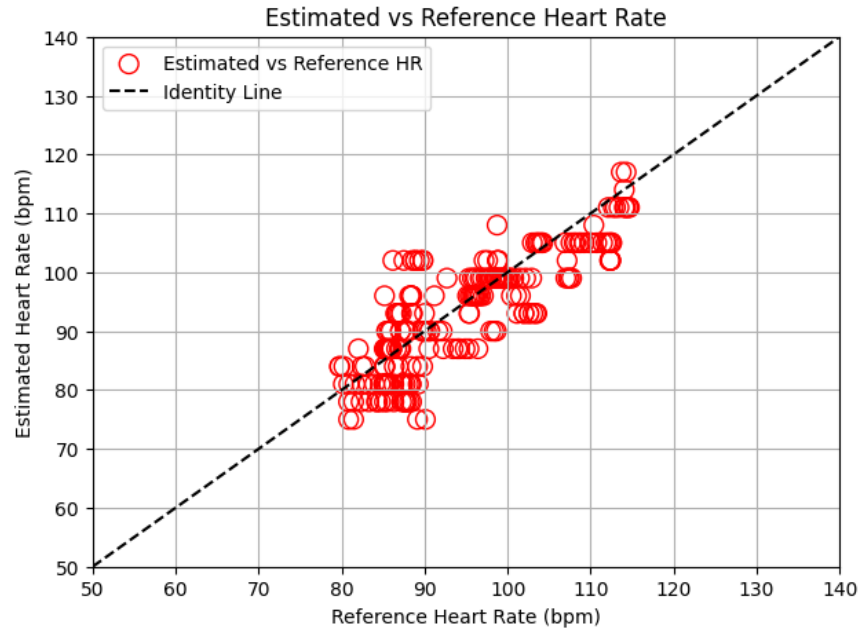


Fig. 4.3 Scatter plot between the *href* and *hest* values for different recordings from the dataset

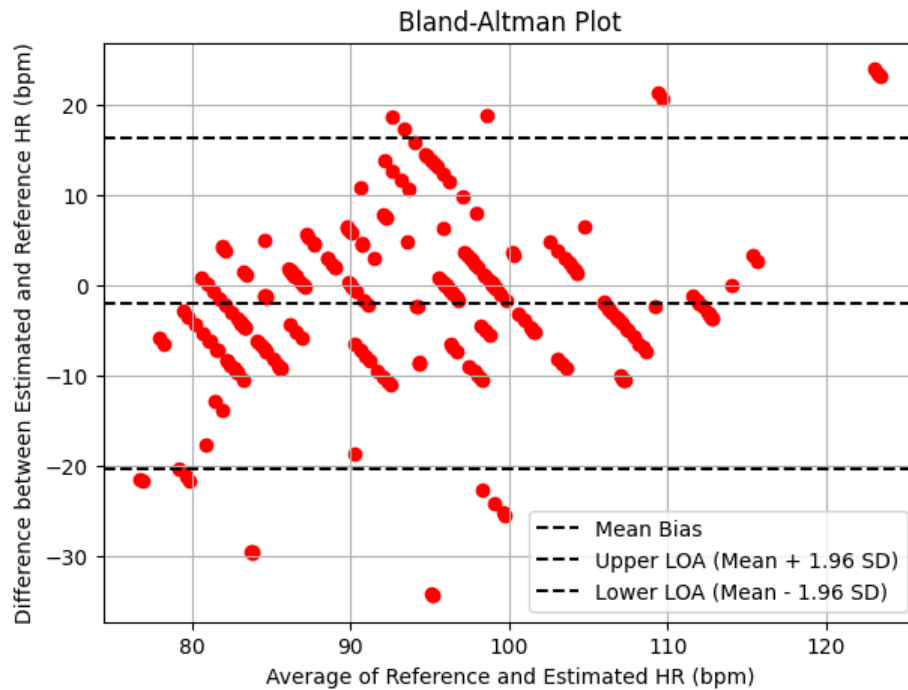


Fig. 4.4 Band-Altman plot describing the agreement between *href* and *hest* values for different subjects.

indicating a slight underestimation of heart rate on average.

The Limits of Agreement (LOA), calculated as $\text{Mean} \pm 1.96 \text{ SD}$, were found to range from -3.35 bpm to 4.19 bpm. This tight range indicates a small error band, which shows that the method can reliably provide estimates even when conditions change. Most of the data points fell within these limits, highlighting a strong

level of consistency. Additionally, the spread didn't reveal any significant proportional bias, suggesting that the error distribution remained fairly stable across both low and high heart rate ranges. This reliability is essential for telemedicine applications, where accurate monitoring is vital for effective patient care.

4.3.2 Outlier detection and error consistency

To enhance the accuracy of heart rate estimation, we implemented a post-processing step that included outlier detection using a carefully chosen threshold mechanism. This process involved examining the estimated heart rate values to spot and eliminate any unusual peaks that strayed far from the expected physiological range. We established an optimal threshold based on the statistical distribution of heart rate variations, which helped us effectively filter out outliers caused by sudden movements, changes in lighting, or noise interference.

By applying this outlier detection method, we significantly improved the overall stability of heart rate estimation, as it helped to weed out misleading signals that could distort the average heart rate. This optimization was especially beneficial during periods of rapid movement or fluctuating lighting, where raw estimates often spiked inaccurately. The post-processing phase played a crucial role in minimizing noise-induced deviations, ultimately leading to a smoother and more reliable final heart rate output.

The analysis of error distribution showed that the estimation error stayed within ± 2 bpm for most of the video frames, demonstrating its reliability against small physiological and environmental changes. Additionally, the post-processed heart rate values remained accurate within the clinically acceptable range of ± 5 bpm for nearly all segments. This underscores the method's practicality for continuous health monitoring, where real-time accuracy is essential.

The results reveal that the heart rate estimation algorithm proposed here achieves impressive accuracy and aligns well with the actual PPG signals from the UBFC-rPPG dataset. By utilizing color space transformation, Fourier Decomposition, and a sliding window approach, the method consistently provides accurate heart rate readings. Plus, the addition of post-processing with outlier detection boosts reliability, ensuring strong performance even in real-world scenarios with motion artifacts and varying lighting conditions.

The method's capability to keep error rates low, maintain a strong correlation, and exhibit minimal bias highlights its potential for non-contact health monitoring in telemedicine and fitness applications.

CHAPTER 5

Final Analysis : Discussion, Conclusion and Future Directions

5.1 Discussion

5.1.1 Analysis and Interpretation

The heart rate estimation method we've proposed marks a significant leap forward in non-contact photoplethysmography (rPPG). It does this by utilizing MediaPipe facial landmark detection and fine-tuning the extraction of Regions of Interest (ROI). The main goal here was to pinpoint the best sub-ROIs from facial landmarks to boost signal stability and cut down on noise artifacts, especially from non-skin areas like hair, eyes, and the background. By carefully selecting these facial regions, we were able to lessen the effects of spatial variability and motion-induced distortions, resulting in a clearer and more consistent blood volume pulse (BVP) signal.

One of the key changes in our process was switching from RGB to HSV color space, with a particular emphasis on the Saturation (S) channel. We chose this channel because it holds up well against changes in lighting, which is a common hurdle in non-contact heart rate monitoring. Traditional RGB analysis often struggles with fluctuations caused by varying light conditions; however, the HSV transformation effectively separates color information (H), saturation (S), and brightness (V), making it easier to isolate the pulsatile signals. This conversion not only enhanced the BVP data but also reduced the influence of environmental lighting, leading to more dependable heart rate detection in different settings.

For precise heart rate extraction, the method employed the Fourier Decomposition Method (FDM), which efficiently decomposed the raw signals into frequency components. The heart rate was identified within the physiological range of 0.7–3 Hz, where the dominant peak corresponded to the pulsatile heartbeat. Unlike traditional Fast Fourier Transform (FFT), FDM allowed for finer resolution in the frequency domain, capturing subtle variations in the BVP signal that are indicative of heartbeats.

While this method has its advantages, it's not completely free from motion artifacts, especially when someone is moving their face or talking. These artifacts can create noise in the heart rate's dominant frequency bands, which might throw off the accuracy of the FDM extraction. To tackle this issue, we added a post-processing outlier detection mechanism that uses an optimal threshold to filter out any unusual peaks and stabilize the heart rate signal. This extra step really boosted the reliability of

heart rate estimates, particularly during those moments when motion-induced noise is more likely to occur.

The experimental results from the UBFC-rPPG dataset showcase how effective this method is, with a Mean Absolute Error (MAE) of 1.31 bpm, a Root Mean Square Error (RMSE) of 1.96 bpm, and a Pearson's Correlation Coefficient (CC) of 0.78. The scatter plots showed a strong correlation with the reference heart rate values, and the Bland-Altman analysis revealed minimal bias with tight Limits of Agreement (LOA). These results suggest that this method can accurately monitor heart rate in real-time, even under less-than-ideal conditions.

5.1.2 Computational Cost Analysis

One of the defining advantages of the proposed method is its computational efficiency. The experimentation was conducted using Google Colab, a cloud-based platform that provides an isolated Python 3 runtime. The implementation was performed entirely on a CPU-based hardware accelerator, without the need for dedicated GPUs. Despite this, the method achieved real-time processing capabilities, demonstrating its lightweight computational footprint. The primary libraries used included:

- OpenCV for image processing and frame manipulation,
- NumPy for efficient numerical operations,
- SciPy for signal processing,
- Matplotlib for visualization of HR estimations and error plots,
- Mediapipe for precise facial landmark detection and Region of Interest (ROI) extraction,
- os for handling directory paths and file operations during data processing.

The sliding window framework, which uses 20-second windows with an 18-second overlap, allowed for continuous heart rate detection, updating HR values every 2 seconds. This approach not only improved the temporal resolution but also made memory usage more efficient by processing only small segments of video frames during each iteration. By employing Fourier Decomposition, we could streamline signal extraction, concentrating on specific frequency bands and lightening the computational load that usually comes with full-spectrum FFT analysis.

Mediapipe was particularly important in identifying stable facial landmarks, which helped in accurately extracting regions of interest (ROI). This process removed non-skin areas from the analysis, boosting signal clarity and cutting down on noise. Additionally, the integration with OS simplified file management during dataset processing, making it easier to navigate through video frames and manage data storage.

Choosing Google Colab as the execution environment offered easy access to cloud resources without the limitations of local machines. Even when relying solely on CPU processing, the method kept an optimal runtime, making it suitable for deployment on low-power devices like mobile health monitors or IoT-based wearables.

In summary, the design principles of this method focused on not just accuracy and robustness, but also on computational efficiency, enabling smooth real-time heart rate monitoring even with limited hardware resources..

5.2 Conclusion

The heart rate estimation method proposed in this study marks a major leap forward in non-contact physiological monitoring. By harnessing the capabilities of Remote Photoplethysmography (rPPG), this technique estimates heart rate (HR) using facial videos taken with standard RGB cameras. Unlike traditional methods that require physical contact, like ECG and finger-based PPG, this approach is completely non-invasive, making it perfect for telehealth, remote patient monitoring, and fitness applications where discreet and continuous HR tracking is essential.

To enhance robustness, the RGB signals are transformed into the HSV color space, with a particular emphasis on the Saturation (S) channel. This channel is notably resistant to fluctuations in ambient lighting, a common hurdle in video-based heart rate estimation. This transformation allows for better isolation of pulsatile blood flow signals compared to traditional RGB analysis, leading to clearer BVP waveforms. By tackling the issues caused by light-induced artifacts, this method ensures consistent performance across different lighting conditions, which is vital for practical applications.

The extracted signals undergo processing through the Fourier Decomposition Method (FDM). This step is pivotal in filtering out noise and isolating heartbeat-related frequencies within the 0.7–3 Hz range, corresponding to typical human heart rates. Unlike standard FFT, FDM offers a more refined frequency analysis that enhances the signal-to-noise ratio (SNR), allowing the subtle pulsations of heartbeats to emerge more distinctly. This contributes directly to the low error metrics achieved in the experimentation, with a Mean Absolute Error (MAE) of 1.31 bpm, a Root Mean Square Error (RMSE) of 1.96 bpm, and a Pearson's Correlation Coefficient (CC) of 0.78, all indicating a strong agreement with ground-truth PPG values.

To further enhance reliability, the method incorporates a post-processing technique for outlier detection, employing an optimal threshold mechanism. This stage is crucial for eliminating erroneous spikes in HR estimation caused by rapid head movements or lighting fluctuations. By smoothing the output signal, the method ensures that the final HR estimations are both stable and realistic, aligning closely with physiological expectations.

The experimental validation conducted on the UBFC-rPPG dataset showcased the method's effectiveness in real-world conditions, where artificial lighting variations and minor head movements were present. Despite the inherent challenges of non-contact monitoring, the proposed approach maintained high accuracy and reliability, validating its potential for deployment in telemedicine and continuous health monitoring scenarios. Furthermore, the scatter plots and Bland-Altman analysis provided visual confirmation of the method's alignment with reference HR values, demonstrating its capacity to accurately track real-time heart rate changes.

However, it is acknowledged that motion artifacts still pose a challenge to the method's robustness, particularly during intense facial movements or conversations. While the post-processing outlier detection mitigates some of these effects, future iterations of the method could benefit from advanced denoising techniques, possibly through machine learning-based artifact removal. Integrating

adaptive filtering mechanisms may further enhance its capability to differentiate between genuine BVP signals and noise-induced variations.

In conclusion, the proposed heart rate estimation method introduces a powerful, non-contact solution for real-time health monitoring using standard RGB cameras. Its reliance on landmark-based ROI extraction, HSV transformation, and Fourier Decomposition enables precise HR tracking with minimal error. The integration of outlier detection further strengthens its reliability, paving the way for its application in telehealth, fitness monitoring, and remote patient care. Future improvements focusing on multi-modal data integration and machine learning-driven noise suppression are expected to extend its usability to diverse environmental conditions and broader physiological metrics, including respiratory rate estimation and stress monitoring.

5.3 Future scope

The promising results achieved by the proposed heart rate estimation method highlight its potential for real-world applications in telemedicine, fitness monitoring, and continuous health assessment. However, there remain several avenues for further improvement and expansion to enhance its robustness, applicability, and global usability.

5.3.1 Robustness Against Motion Artifacts

Although the proposed method incorporates post-processing with outlier detection, significant head movements and facial expressions still introduce noise that affects heart rate accuracy. Future work can focus on integrating motion compensation techniques that dynamically adjust the Region of Interest (ROI) based on head motion. Advanced optical flow algorithms or deep-learning-based motion stabilization could be explored to minimize the impact of non-rigid facial deformations during heart rate monitoring.

5.3.2 Multi-Modal Signal Fusion

Currently, the method relies solely on RGB facial videos for heart rate estimation. A potential enhancement is the integration of multi-modal signals such as thermal imaging, infrared (IR) sensing, and depth cameras. Thermal cameras can capture heat variations corresponding to blood flow, while IR sensors are less affected by ambient light, making them ideal for low-light conditions. Depth sensors could improve the isolation of facial regions, further minimizing noise from the background.

The fusion of these signals with the existing rPPG pipeline could be implemented using multi-modal learning frameworks, allowing for more robust HR estimation even in challenging environments. This would also pave the way for the simultaneous estimation of other vital signs, such as respiratory rate and oxygen saturation (SpO₂), expanding its clinical relevance.

5.3.3 Exploration Across Different Skin Tones and Ethnicities

Another thing to consider for future evaluation is to study the method's performance across various skin tones and ethnicities. Skin reflectance varies significantly with pigment levels of the skin and probably might reduce the signal quality of rPPG-based monitoring. The future work may incorporate a wider variety of dataset skin tones of lighting conditions for algorithm robustness evaluation.

Moreover, adaptive calibration methods that vary color space transformations based on skin reflectance could further enhance accuracy along all demographics. Deep learning methods launched on ethnically variant training data could learn such invariant features so that heart rate estimation would be reliable independent of skin tone.

5.3.4 Automated Sub-ROI Selection and Noise Reduction

Currently, an optimally fixed sub-ROI based on MAE is used. Further improvements can be pursued through dynamic sub-ROI selection, which would be algorithm-driven. A self-learning model may analyze signal quality in real time and adjust the sub-ROI to the region where the signal has the best integrity, thus making the HR estimation method more resilient.

In addition, more advanced denoising techniques, including Wavelet Transform-based filtering or an RNN, could be used to separate true BVP signals from noise. When varying ambient light conditions or subtle head movements hamper the detection of these signals, these advanced methods might outperform traditional filtering techniques.

5.3.5 Deep Learning-Based Artifact Removal

The Fourier Decomposition with thresholding for outlier removal is the current method, whereas the deep learning-based artifact removal approach seems to be a promising option. The CNNs and GANs would be used to automatically discriminate between noise and true pulse signals even in high-motion segments. These networks could learn very complex temporal patterns, thus providing a clear separability for hr estimation.

Then, use of transformers, such as ViTs, could further align the method's attention to segments of clean signals and away from noise-filled segments. This would improve robustness and reduce post-processing overhead.

5.3.6 Integration with Real-Time Applications

To develop into a more practical system for real-life scenarios, the method could be altered for a real-time deployment on wearable devices and mobile apps. Certain little models like quantization or model pruning might change their algorithm to a form executable on low-power embedded systems. Direct interfacing with edge computing infrastructures will enable real-time monitoring minus the cloud dependency, thus keeping unwanted latency at bay and privacy intact. Real-time monitoring applications can be:

- Telemedicine for continuous remote patient monitoring.
- Fitness applications studying real-time heart rate during workouts.
- Mental health assessment wherein HRV is the prime indicator of stress

levels.

The encouraging results obtained in this research open the pathway for several other research avenues both for furthering the scope of user adaptability and robustness of the proposed heart rate estimation method. Using its strength as a base, future research aims at further smoothing some current hurdles as well as extending the approach to several other more challenging and practical real-world scenarios.

REFERENCES

- [1] H. Xiao, T. Liu, Y. Sun, Y. Li, S. Zhao, and A. Avolio, "Remote photoplethysmography for heart rate measurement: A review," *Biomedical Signal Processing and Control*, vol. 88, p. 105608, Feb. 2024, doi: 10.1016/j.bspc.2023.105608.
- [2] Y. Sun and N. Thakor, "Photoplethysmography Revisited: From Contact to Noncontact, From Point to Imaging," *IEEE Trans Biomed Eng*, vol. 63, no. 3, pp. 463–477, Mar. 2016, doi: 10.1109/TBME.2015.2476337.
- [3] A. D. Mehta and H. Sharma, "Heart Rate Estimation from RGB Facial Videos Using Robust Face Demarcation and VMD," in *2021 National Conference on Communications (NCC)*, Jul. 2021, pp. 1–6. doi: 10.1109/NCC52529.2021.9530067.
- [4] W. Verkruijsse, L. O. Svaasand, and J. S. Nelson, "Remote plethysmographic imaging using ambient light," *Opt Express*, vol. 16, no. 26, pp. 21434–21445, Dec. 2008, doi: 10.1364/oe.16.021434.
- [5] I. Odinaev, K. L. Wong, J. W. Chin, R. Goyal, T. T. Chan, and R. H. Y. So, "Robust Heart Rate Variability Measurement from Facial Videos," *Bioengineering (Basel)*, vol. 10, no. 7, p. 851, Jul. 2023, doi: 10.3390/bioengineering10070851.
- [6] D. N. Tran, H. Lee, and C. Kim, "A robust real time system for remote heart rate measurement via camera," in *2015 IEEE International Conference on Multimedia and Expo (ICME)*, Jun. 2015, pp. 1–6. doi: 10.1109/ICME.2015.7177484.
- [7] G. De Haan and V. Jeanne, "Robust Pulse Rate From Chrominance-Based rPPG," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 10, pp. 2878–2886, Oct. 2013, doi: 10.1109/TBME.2013.2266196.
- [8] 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, Jan. 2017, doi: 10.1109/TAFFC.2016.2515084.
- [9] 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, Oct. 2019, doi: 10.1109/TIM.2018.2879706.
- [10] 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)*, Jul. 2019, pp. 6521–6524. doi: 10.1109/EMBC.2019.8857012.
- [11] G. Lempe, S. Zaunseder, T. Wirthgen, S. Zipser, and H. Malberg, "ROI Selection for Remote Photoplethysmography," in *Bildverarbeitung für die Medizin 2013*, H.-P. Meinzer, T.

M. Deserno, H. Handels, and T. Tolxdorff, Eds., Berlin, Heidelberg: Springer, 2013, pp. 99–103. doi: 10.1007/978-3-642-36480-8_19.

[12] Y. Tong *et al.*, “Detail-preserving arterial pulse wave measurement based Biorthogonal wavelet decomposition from remote RGB observations,” *Measurement*, vol. 222, p. 113605, Nov. 2023, doi: 10.1016/j.measurement.2023.113605.

[13] M.-Z. Poh, D. J. McDuff, and R. W. Picard, “Non-contact, automated cardiac pulse measurements using video imaging and blind source separation,” *Opt. Express, OE*, vol. 18, no. 10, pp. 10762–10774, May 2010, doi: 10.1364/OE.18.010762.

[14] 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)*, Sep. 2011, pp. 405–410. Accessed: Jul. 04, 2024. [Online]. Available: <https://ieeexplore.ieee.org/abstract/document/6078233>

[15] A. Biswas, M. Singha Roy, and R. Gupta, “Motion Artifact Reduction from Finger Photoplethysmogram Using Discrete Wavelet Transform,” in *Recent Trends in Signal and Image Processing*, S. Bhattacharyya, A. Mukherjee, H. Bhaumik, S. Das, and K. Yoshida, Eds., Singapore: Springer, 2019, pp. 89–98. doi: 10.1007/978-981-10-8863-6_10.

[16] D. Pollreisz and N. TaheriNejad, “Detection and Removal of Motion Artifacts in PPG Signals,” *Mobile Netw Appl*, vol. 27, no. 2, pp. 728–738, Apr. 2022, doi: 10.1007/s11036-019-01323-6.

[17] J. F. Thayer, S. S. Yamamoto, and J. F. Brosschot, “The relationship of autonomic imbalance, heart rate variability and cardiovascular disease risk factors,” *International Journal of Cardiology*, vol. 141, no. 2, pp. 122–131, May 2010, doi: 10.1016/j.ijcard.2009.09.543.

[18] X. Liu, J. Fromm, S. Patel, and D. McDuff, “Multi-Task Temporal Shift Attention Networks for On-Device Contactless Vitals Measurement,” Feb. 28, 2021, *arXiv:arXiv:2006.03790*. doi: 10.48550/arXiv.2006.03790.

[19] F. Anowar, S. Sadaoui, and B. Selim, “Conceptual and empirical comparison of dimensionality reduction algorithms (PCA, KPCA, LDA, MDS, SVD, LLE, ISOMAP, LE, ICA, t-SNE),” *Computer Science Review*, vol. 40, p. 100378, May 2021, doi: 10.1016/j.cosrev.2021.100378.

[20] M. Schmid, D. Rath, and U. Diebold, “Why and How Savitzky–Golay Filters Should Be Replaced,” *ACS Meas. Sci. Au*, vol. 2, no. 2, pp. 185–196, Apr. 2022, doi: 10.1021/acsmeasuresciau.1c00054.

[21] M. Savic and G. Zhao, “RS-rPPG: Robust Self-Supervised Learning for rPPG,” in *2024 IEEE 18th International Conference on Automatic Face and Gesture Recognition (FG)*, May 2024, pp. 1–10. doi: 10.1109/FG59268.2024.10581991.

[22] R. Song, X. Sun, J. Cheng, X. Yang, and X. Chen, “Video-Based Heart Rate Measurement Against Uneven Illuminations Using Multivariate Singular Spectrum Analysis,” *IEEE Signal Processing Letters*, vol. 29, pp. 2223–2227, 2022, doi: 10.1109/LSP.2022.3215112.

- [23] G.-S. J. Hsu, R.-C. Xie, A. Ambikapathi, and K.-J. Chou, “A deep learning framework for heart rate estimation from facial videos,” *Neurocomputing*, vol. 417, pp. 155–166, Dec. 2020, doi: 10.1016/j.neucom.2020.07.012.
- [24] Y. Qiu, Y. Liu, J. Arteaga-Falconi, H. Dong, and A. E. Saddik, “EVM-CNN: Real-Time Contactless Heart Rate Estimation From Facial Video,” *IEEE Transactions on Multimedia*, vol. 21, no. 7, pp. 1778–1787, Jul. 2019, doi: 10.1109/TMM.2018.2883866.
- [25] B. Lokendra and G. Puneet, “AND-rPPG: A novel denoising-rPPG network for improving remote heart rate estimation,” *Computers in Biology and Medicine*, vol. 141, p. 105146, Feb. 2022, doi: 10.1016/j.compbiomed.2021.105146.
- [26] X. Zhang, Z. Xia, L. Liu, and X. Feng, “Demodulation Based Transformer for rPPG Generation and Heart Rate Estimation,” *IEEE Signal Processing Letters*, vol. 30, pp. 1042–1046, 2023, doi: 10.1109/LSP.2023.3302697.
- [27] W. Chen and D. McDuff, “DeepPhys: Video-Based Physiological Measurement Using Convolutional Attention Networks,” in *Computer Vision – ECCV 2018*, vol. 11206, V. Ferrari, M. Hebert, C. Sminchisescu, and Y. Weiss, Eds., in Lecture Notes in Computer Science, vol. 11206, Cham: Springer International Publishing, 2018, pp. 356–373. doi: 10.1007/978-3-030-01216-8_22.
- [28] R. Špetlík, “Visual Heart Rate Estimation with Convolutional Neural Network”.
- [29] Z. Yu *et al.*, “PhysFormer++: Facial Video-based Physiological Measurement with SlowFast Temporal Difference Transformer,” Feb. 07, 2023, *arXiv*: arXiv:2302.03548. doi: 10.48550/arXiv.2302.03548.
- [30] Z. Yu, Y. Shen, J. Shi, H. Zhao, P. Torr, and G. Zhao, “PhysFormer: Facial Video-based Physiological Measurement with Temporal Difference Transformer,” in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, New Orleans, LA, USA: IEEE, Jun. 2022, pp. 4176–4186. doi: 10.1109/CVPR52688.2022.00415.
- [31] A. K. Gupta, R. Kumar, L. Birla, and P. Gupta, “RADIANT: Better rPPG estimation using signal embeddings and Transformer,” in *2023 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, Jan. 2023, pp. 4965–4975. doi: 10.1109/WACV56688.2023.00495.
- [32] J. Speth, N. Vance, P. Flynn, and A. Czajka, “Non-Contrastive Unsupervised Learning of Physiological Signals from Video,” in *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2023, pp. 14464–14474. doi: 10.1109/CVPR52729.2023.01390.
- [33] X. Liu, J. Fromm, S. Patel, and D. McDuff, “Multi-Task Temporal Shift Attention Networks for On-Device Contactless Vitals Measurement,” presented at the Neural Information Processing Systems (NeurIPS), Dec. 2020. Accessed: May 20, 2025. [Online]. Available: <https://www.microsoft.com/en-us/research/publication/multi-task-temporal-shift-attention-networks-for-on-device-contactless-vitals-measurement/>
- [34] M. Hu, F. Qian, X. Wang, L. He, D. Guo, and F. Ren, “Robust Heart Rate Estimation

With Spatial–Temporal Attention Network From Facial Videos,” *IEEE Transactions on Cognitive and Developmental Systems*, vol. 14, no. 2, pp. 639–647, Jun. 2022, doi: 10.1109/TCDS.2021.3062370.

[35] 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, May 2021, doi: 10.1109/JBHI.2021.3051176.

[36] A. K. Gupta, R. Kumar, L. Birla, and P. Gupta, “RADIANT: Better rPPG estimation using signal embeddings and Transformer,” in *2023 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, Jan. 2023, pp. 4965–4975. doi: 10.1109/WACV56688.2023.00495.

[37] Y.-Y. Tsou, Y.-A. Lee, and C.-T. Hsu, “Multi-task Learning for Simultaneous Video Generation and Remote Photoplethysmography Estimation,” in *Computer Vision – ACCV 2020*, H. Ishikawa, C.-L. Liu, T. Pajdla, and J. Shi, Eds., Cham: Springer International Publishing, 2021, pp. 392–407. doi: 10.1007/978-3-030-69541-5_24.

[38] W.-H. Chung, C.-J. Hsieh, S.-H. Liu, and C.-T. Hsu, “Domain Generalized RPPG Network: Disentangled Feature Learning with Domain Permutation and Domain Augmentation,” in *Computer Vision – ACCV 2022*, L. Wang, J. Gall, T.-J. Chin, I. Sato, and R. Chellappa, Eds., Cham: Springer Nature Switzerland, 2023, pp. 41–57. doi: 10.1007/978-3-031-26284-5_3.

[39] Y. Yang *et al.*, *SimPer: Simple Self-Supervised Learning of Periodic Targets*. 2022. doi: 10.48550/arXiv.2210.03115.

[40] D.-Y. Kim, S.-Y. Cho, K. Lee, and C.-B. Sohn, “A Study of Projection-Based Attentive Spatial–Temporal Map for Remote Photoplethysmography Measurement,” *Bioengineering*, vol. 9, p. 638, Nov. 2022, doi: 10.3390/bioengineering9110638.

[41] Y. Ba, Z. Wang, K. D. Karınca, O. D. Bozkurt, and A. Kadambi, “Style Transfer with Bio-realistic Appearance Manipulation for Skin-tone Inclusive rPPG,” in *2022 IEEE International Conference on Computational Photography (ICCP)*, Aug. 2022, pp. 1–12. doi: 10.1109/ICCP54855.2022.9887649.

[42] W. Wang, A. C. Den Brinker, S. Stuijk, and G. De Haan, “Algorithmic Principles of Remote PPG,” *IEEE Trans. Biomed. Eng.*, vol. 64, no. 7, pp. 1479–1491, Jul. 2017, doi: 10.1109/TBME.2016.2609282.

[43] Y. Kartynnik, A. Ablavatski, I. Grishchenko, and M. Grundmann, “Real-time Facial Surface Geometry from Monocular Video on Mobile GPUs,” Jul. 15, 2019, *arXiv: arXiv:1907.06724*. doi: 10.48550/arXiv.1907.06724.

[44] H. G. K. -, Z. B. -, and M. S. S. -, “Optimal Color Image Enhancement Using Wavelet and K-means Clustering,” *JDCTA*, vol. 5, no. 1, pp. 112–122, Jan. 2011, doi: 10.4156/jdcta.vol5.issue1.13.

[45] “Empirical mode decomposition for respiratory and heart rate estimation from the photoplethysmogram | IEEE Conference Publication | IEEE Xplore.” Accessed: Feb. 19, 2025.

[Online]. Available: <https://ieeexplore.ieee.org/document/6713498>

[46] A. Garde, W. Karlen, P. Dehkordi, J. Ansermino, and G. Dumont, “Empirical mode decomposition for respiratory and heart rate estimation from the photoplethysmogram,” in *Computing in Cardiology 2013*, Sep. 2013, pp. 799–802. Accessed: May 20, 2025. [Online]. Available: <https://ieeexplore.ieee.org/document/6713498>

[47] S. Bobbia, R. Macwan, Y. Benezeth, A. Mansouri, and J. Dubois, “Unsupervised skin tissue segmentation for remote photoplethysmography,” *Pattern Recognition Letters*, vol. 124, pp. 82–90, Jun. 2019, doi: 10.1016/j.patrec.2017.10.017.

[48] H. Lu, H. Han, and S. K. Zhou, “Dual-GAN: Joint BVP and Noise Modeling for Remote Physiological Measurement,” in *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, TN, USA: IEEE, Jun. 2021, pp. 12399–12408. doi: 10.1109/CVPR46437.2021.01222.

LIST OF PUBLICATIONS AND THEIR PROOFS

- [1] Ritika Gupta, Manjeet Kumar and Mahipal Singh Choudhary, “Heart Rate Estimation Using Remote PPG and Facial Landmark Features with Fourier Decomposition and ROI Optimization”, presented at the 3rd International Conference on Women Researchers in Electronics and Computing, 2025 (WREC'25), NIT Jalandhar.
(Received Best Paper Award).
- [2] Ritika Gupta, Manjeet Kumar and Mahipal Singh Choudhary, “Remote Photoplethysmography for Heart Rate Estimation: A Comprehensive Survey on Methods and Applications”, submitted at the 1st International Conference on Power and Intelligent Control System, 2025, NIT Hamirpur.
(Accepted).

Paper ID 324 - WREC'25

1 message

Microsoft CMT <email@msr-cmt.org>
To: Ritika Gupta <ritikagupta_23spd06@dtu.ac.in>

Tue, Apr 8, 2025 at 1:11 PM

Dear Dr. Ritika Gupta,

Paper ID: 324

Heart Rate Estimation Using Remote PPG and Facial Landmark Features with Fourier Decomposition and ROI Optimization

** Please ignore this email if you have already registered for the conference.**

Your paper is accepted. Now you can log in to your CMT account to view the detailed reviews. Please ensure that you submit the formatted manuscript in the Camera-Ready section by today, following our formatting guidelines and incorporating all reviewer comments. If your paper requires revisions, please submit the response sheet as a supplementary file along with the revised manuscript in the Camera-Ready section. Make sure the paper's plagiarism level stays below 20%, as exceeding this limit will result in rejection at any stage of evaluation/publication.

Additionally, you can complete your conference registration by paying the required fee at: https://v1.nitj.ac.in/events_registration/ic_wrec_2025/login. We encourage you to register at your earliest convenience.

For documentation and certification purposes, please fill out the Google form: <https://docs.google.com/forms/d/1R1MnCreWujqMZ7cm5rPgqywlqf1FV8WJ20kl3gxT6-U/edit?ts=67dcfdfa>

We sincerely appreciate your contribution to WREC'25 and thank you for the opportunity to consider your work.

Kind regards,
Dr. Urvashi
Chair (WREC'25)

To stop receiving conference emails, you can check the 'Do not send me conference email' box from your User Profile.

Microsoft respects your privacy. To learn more, please read our [Privacy Statement](#).

Microsoft Corporation
One [Microsoft Way](#)
[Redmond, WA 98052](#)



Certificate No. NITJ/WREC_2025/ OL_BP_10
3rd International Conference

Women Researchers in Electronics and Computing

WREC'25



Theme: Sustainable Development Goals

(18-19 April, 2025)

Certificate of Appreciation

This is to certify that paper titled

Heart Rate Estimation Using Remote PPG and Facial Landmark Features with Fourier Decomposition and ROI

Optimization

presented by

Ritika Gupta

&

authored by

Ritika Gupta, Manjeet Kumar, Mahipal Singh Choudhary

has been awarded as **Best Paper** for the Track **T5: Data Science for Sustainable Development** during 3rd International Conference “Women Researchers in Electronics and Computing” on Sustainable Development Goals held on 18-19 April 2025 organized by **Dr B R Ambedkar National Institute of Technology Jalandhar, Punjab, India.**

Prof. Mamta Khosla
Executive General Chair

Dr. Indu Saini
Executive General Chair



Dr B R Ambedkar
National Institute of Technology
Jalandhar - 144008

**International Conference on Women Researchers in
Electronics and Computing (WREC 2025)**
(April 18 - 20, 2025)

Registration Fee e-Receipt

Participant's Details

Registration ID : 2025081101
Name : RITIKA GUPTA
Category : External (Outside NITJ)
Participant Type : Student
Nationality : India
Mobile No. : 09129464720
Email ID : ritikagupta_23spd06@dtu.ac.in

Fee Details

Amount (GST 18%) : **INR 3540.00**
Transaction ID : 113713776139
Transaction Date : 05-Apr-2025 20:38:52
Payment Status : **SUCCESS**

Status of Conference Paper Submitted to PICS-2025 (Revision)

1 message

Microsoft CMT <noreply@msr-cmt.org>
To: Ritika Gupta <ritikagupta_23spd06@dtu.ac.in>

Mon, May 26, 2025 at 7:52 PM

Dear Authors,

We are pleased to inform you that your paper titled "Remote Photoplethysmography for Heart Rate Estimation: A Comprehensive Survey on Methods and Applications" (Submission ID: #258), has been provisionally accepted for presentation at the PICS-2025 conference, scheduled to take place from July 4-5, 2025, at NIT Hamirpur, (H.P.), India.

Kindly submit the Camera-Ready Paper by 31 May 2025, after addressing all the reviewer comments by logging in to your CMT account.

Important Notes:

[1] The Camera-Ready Paper must be strictly in the Springer Conference Paper Format (figures captions, tables, headings, references, etc.). The templates are available in the following links:

MS Word: <https://drive.google.com/file/d/1tYPp2cGK2BS4f4fB8EqpTLRvzEqIXaxz/view>;

Latex: <https://www.overleaf.com/latex/templates/springer-conference-proceedings-template-updated-2022-01-12/wcvbtmwtqkqj>

[2] Proofread your manuscript thoroughly to confirm that it will require no revision. The acceptance is subject to the final plagiarism and quality check and satisfactory incorporation of reviewer comments (available on the Microsoft CMT Portal) and the submission will be checked for a similarity score below 20%, as determined by plagiarism detection tools.

[3] Please note that the paper will be considered "Accepted" only after completion of the Registration and submission of the Camera-Ready Paper. The registration procedure is available in the following link: <https://pics2025nith.com/Registration>. We have added an option of hybrid mode of presentation (offline/in-person) in the registration form.

[4] Prepare a response sheet detailing how each of the reviewers' comments have been addressed. Create a file named after your Submission ID (e.g.SID402) containing the compliance report to reviewers' comments. Finally, upload this file under the Supplementary File Upload section. Kindly upload the camera-ready paper and supplementary file by 31 May 2025.

In case of any query, kindly reach out to any of the Conference Organizing Committee members.

Regards,
PICS-2025 Organizing Committee

Please do not reply to this email as it was generated from an email account that is not monitored.

To stop receiving conference emails, you can check the 'Do not send me conference email' box from your User Profile.

Microsoft respects your privacy. To learn more, please read our [Privacy Statement](#).

Microsoft Corporation
One [Microsoft Way](#)
Redmond, WA 98052

e-Receipt for State Bank Collect Payment



NIT HAMIRPUR

NIT INSTITUTE HAMIRPUR , , HAMIRPUR-177005

Date: 30-May-2025

SBCollect Reference Number :	DU01309302
Category :	WORKSHOP STC FDP CONFERENCE
Amount :	₹3000
NAME OF THE PARTICIPANT :	Ritika Gupta
DESIGNATION :	PG Student
MOBILE NUMBER :	9129464720
e-mail :	ritikag22499@gmail.com
ADDRESS :	NIT Hamirpur
NAME OF INSTITUTE/ORGANIZATION :	National Institute of Technology Hamirpur
TITLE OF STC/FDP/WORKSHOP/CONF :	PICS-2025 Conference
DATES OF PROPOSED PROGRAMMES :	04-05 July 2025
ORGANIZING DEPARTMENT :	Department of Electrical Engineering
AMOUNT OF FEE PAID :	3000
REMARKS :	NA
COURSE FEE :	3000

Transaction charge : 0.00

Total Amount (In Figures) : 3,000.00

Total Amount (In words) : Rupees Three Thousand Only

Remarks : Conference Paper

Notification 1:

Notification 2:

Thank you for choosing SB Collect. If you have any query / grievances regarding the transaction, please contact us

Toll-free helpline number i.e. 1800-1111-09 / 1800 - 1234/1800 2100

Email -: sbcollect@sbi.co.in



DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Shahbad Daultapur, Main Bawana Road, Delhi-42

PLAGIARISM VERIFICATION

Title of the Thesis : “Development Of An Efficient Algorithm For Heart Rate Estimation Using Remote Photoplethysmography Signal”

Total Pages _____

Name of the Scholar : Ritika Gupta

Supervisor (s) :

(1) Dr. M. S. Choudhary, Professor, ECE, DTU

(2) Dr. Manjeet Kumar, Assistant Professor, ECE, DTU

Department of Electronics and Communication Engineering

This is to report that the above thesis was scanned for similarity detection. Process and outcome is given below:

Software used: _____

Similarity Index: _____,

Total Word Count: _____

Date: _____

A handwritten signature in blue ink, which appears to read 'Manjeet Kumar', is written over a horizontal line.

Candidate's Signature

Signature of Supervisor(s)

Manjeet Kumar

Thesis_final_draft_24-05-2025.pdf



Delhi Technological University

Document Details

Submission ID

trn:oid:::27535:97520556

Submission Date

May 24, 2025, 7:05 PM GMT+5:30

Download Date

May 24, 2025, 7:08 PM GMT+5:30

File Name

Thesis_final_draft_24-05-2025.pdf

File Size

1.4 MB

58 Pages

15,921 Words

90,971 Characters





16% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.




Filtered from the Report

- Bibliography
- Quoted Text

Match Groups

-  **220** Not Cited or Quoted 16%
Matches with neither in-text citation nor quotation marks
-  **8** Missing Quotations 0%
Matches that are still very similar to source material
-  **0** Missing Citation 0%
Matches that have quotation marks, but no in-text citation
-  **0** Cited and Quoted 0%
Matches with in-text citation present, but no quotation marks

Top Sources

- 9%  Internet sources
- 10%  Publications
- 11%  Submitted works (Student Papers)

Integrity Flags

0 Integrity Flags for Review

No suspicious text manipulations found.

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.

Match Groups

- 220** Not Cited or Quoted 16%
Matches with neither in-text citation nor quotation marks
- 8** Missing Quotations 0%
Matches that are still very similar to source material
- 0** Missing Citation 0%
Matches that have quotation marks, but no in-text citation
- 0** Cited and Quoted 0%
Matches with in-text citation present, but no quotation marks

Top Sources

- 9% Internet sources
- 10% Publications
- 11% Submitted works (Student Papers)

Top Sources

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	Submitted works	Indian School of Mines on 2025-05-23	<1%
2	Internet	dtu.ac.in	<1%
3	Publication	Hanguang Xiao, Tianqi Liu, Yisha Sun, Yulin Li, Shiyi Zhao, Alberto Avolio. "Remot...	<1%
4	Publication	Arpita Panigrahi, Hemant Sharma, Atin Mukherjee. "Video-based HR measureme...	<1%
5	Internet	arxiv.org	<1%
6	Internet	dspace.dtu.ac.in:8080	<1%
7	Submitted works	Hong Kong University of Science and Technology on 2025-05-02	<1%
8	Submitted works	University of Liverpool on 2025-04-29	<1%
9	Publication	Shuai Ding, Zhen Ke, Zijie Yue, Cheng Song, Lu Lu. "Non-contact Multi-physiologic...	<1%
10	Internet	www.semanticscholar.org	<1%

11	Publication	Guoliang Xiang, Song Yao, Yong Peng, Hanwen Deng, Xianhui Wu, Kui Wang, Ying...	<1%
12	Publication	Ru Jing Lee, Saaveethya Sivakumar, King Hann Lim. "Review on remote heart rate...	<1%
13	Publication	Haoyuan Gao, Chao Zhang, Shengbing Pei, Xiaopei Wu. "Region of Interest Analys...	<1%
14	Internet	www.mdpi.com	<1%
15	Submitted works	University of Sheffield on 2024-09-09	<1%
16	Internet	www.nature.com	<1%
17	Internet	jultika.oulu.fi	<1%
18	Submitted works	University of Lancaster on 2022-06-02	<1%
19	Internet	lrcdrs.bennett.edu.in	<1%
20	Submitted works	National University of Singapore on 2025-03-07	<1%
21	Internet	fastercapital.com	<1%
22	Publication	Trishna Saikia, Lokendra Birla, Anup Kumar Gupta, Puneet Gupta. "HREADAI: Hea...	<1%
23	Internet	export.arxiv.org	<1%
24	Internet	sites.google.com	<1%

25	Publication	Poonam Nandal, Mamta Dahiya, Meeta Singh, Arvind Dagur, Brijesh Kumar. "Pro...	<1%
26	Internet	renchengsong.github.io	<1%
27	Submitted works	The University of Manchester on 2024-08-28	<1%
28	Submitted works	University of Wollongong on 2024-03-26	<1%
29	Internet	digilib.unimed.ac.id	<1%
30	Internet	docslib.org	<1%
31	Submitted works	University of Liverpool on 2024-10-11	<1%
32	Submitted works	University of Westminster on 2025-04-16	<1%
33	Submitted works	Heriot-Watt University on 2025-02-14	<1%
34	Submitted works	University of Hertfordshire on 2023-09-17	<1%
35	Internet	www.mouthshut.com	<1%
36	Internet	www.researchgate.net	<1%
37	Publication	"Computer Vision – ECCV 2024", Springer Science and Business Media LLC, 2025	<1%
38	Submitted works	Hong Kong University of Science and Technology on 2021-12-03	<1%

39	Submitted works	University of Limerick on 2019-08-19	<1%
40	Submitted works	universiteknologimara on 2025-02-12	<1%
41	Publication	"Global Trends in Information Systems and Software Applications", Springer Scie...	<1%
42	Publication	Daniel McDuff. "Camera Measurement of Physiological Vital Signs", ACM Computi...	<1%
43	Publication	L Tarassenko, M Villarroel, A Guazzi, J Jorge, D A Clifton, C Pugh. "Non-contact vid...	<1%
44	Submitted works	Liverpool John Moores University on 2023-06-14	<1%
45	Submitted works	Universiti Teknologi Petronas on 2017-07-03	<1%
46	Submitted works	University of East London on 2025-05-09	<1%
47	Submitted works	University of Glasgow on 2024-04-25	<1%
48	Submitted works	University of Liverpool on 2025-05-05	<1%
49	Submitted works	University of Westminster on 2024-11-11	<1%
50	Submitted works	Loughborough University on 2023-09-01	<1%
51	Submitted works	University of Prince Mugrin on 2024-05-22	<1%
52	Internet	jmr.cmsjournal.net	<1%

53	Publication	Nicky Nirlipta Sahoo, VS Sachidanand, Matcha Naga Gayathri, Balamurali Muruge...	<1%
54	Submitted works	University of Surrey on 2023-05-17	<1%
55	Internet	accedacris.ulpgc.es	<1%
56	Internet	livrepository.liverpool.ac.uk	<1%
57	Internet	mSPACE.lib.umanitoba.ca	<1%
58	Internet	www.mccc.edu	<1%
59	Internet	www.tandfonline.com	<1%
60	Publication	Miha Finžgar, Primož Podržaj. "A wavelet-based decomposition method for a rob...	<1%
61	Submitted works	University of Glasgow on 2021-09-16	<1%
62	Submitted works	University of Sheffield on 2014-08-27	<1%
63	Submitted works	University of Sydney on 2024-05-23	<1%
64	Internet	cardiovascularultrasound.biomedcentral.com	<1%
65	Internet	nemo.asee.org	<1%
66	Internet	www.jatit.org	<1%

67	Internet	www.ncbi.nlm.nih.gov	<1%
68	Publication	Dezhao Zhai, Wei Chen, Qinwei Li, Ming Yu, Hang Wu. "Adversarial learning netw...	<1%
69	Submitted works	Monash University on 2021-09-22	<1%
70	Internet	alumni.media.mit.edu	<1%
71	Internet	www.tara.tcd.ie	<1%
72	Publication	"Advances in Visual Computing", Springer Science and Business Media LLC, 2019	<1%
73	Publication	"Computer Vision – ACCV 2020", Springer Science and Business Media LLC, 2021	<1%
74	Publication	Gasper Slapnicar, Erik Dovgan, Pia Cuk, Mitja Lustrek. "Contact-Free Monitoring o...	<1%
75	Publication	Nor Surayahani Suriani, Nur Adlina Jumain, Abdalla Abdurahman Ali, Norzali Hj M...	<1%
76	Submitted works	Nottingham Trent University on 2024-12-06	<1%
77	Publication	Ronny Stricker, Steffen Muller, Horst-Michael Gross. "Non-contact video-based pu...	<1%
78	Submitted works	University of Bristol on 2022-09-12	<1%
79	Submitted works	University of Liverpool on 2024-04-15	<1%
80	Submitted works	University of Oxford on 2025-05-21	<1%

81	Submitted works	University of St Andrews on 2024-02-19	<1%
82	Submitted works	University of Warwick on 2022-03-24	<1%
83	Submitted works	University of Westminster on 2025-04-16	<1%
84	Publication	Walter Karlen, Ainara Garde, Dorothy Myers, Cornie Scheffer, J Mark Ansermino, ...	<1%
85	Publication	Xuenan Liu, Xuezhi Yang, Dingliang Wang, Alexander Wong. "Detecting Pulse Rat...	<1%
86	Publication	Xuenan Liu, Xuezhi Yang, Rencheng Song, Jie Zhang, Longwei Li. "VideoCAD: An U...	<1%
87	Publication	Zhipeng Li, Hanguang Xiao, Ziyi Xia, Tianqi Liu, Xiaoxuan Huangang, Feizhong Zh...	<1%
88	Internet	dokumen.pub	<1%
89	Internet	mobt3ath.com	<1%
90	Internet	proceedings.science	<1%
91	Internet	repository.library.carleton.ca	<1%
92	Internet	tud.qucosa.de	<1%
93	Internet	vdoc.pub	<1%
94	Publication	"EMBECC & NBC 2017", Springer Science and Business Media LLC, 2018	<1%

95	Publication	"Emergent Converging Technologies and Biomedical Systems", Springer Science ...	<1%
96	Publication	Aarts, Lonneke A.M., Vincent Jeanne, John P. Cleary, C. Lieber, J. Stuart Nelson, Sid...	<1%
97	Submitted works	Amrita Vishwa Vidyapeetham on 2023-06-21	<1%
98	Publication	Arya Deo Mehta, Hemant Sharma. "OPOIRES: A robust non-contact respiratory ra...	<1%
99	Publication	Bixiao Ling, Pengfei Zhang, Jianjun Qian, Jian Yang. "Chapter 37 Spatial-Channel ...	<1%
100	Publication	Cheng-Yan Guo, Chi-Hung Huang, Chia-Chi Chang, Kuan-Jen Wang, Tung-Li Hsieh....	<1%
101	Submitted works	Curtin University of Technology on 2024-10-29	<1%
102	Publication	Edwin Arkel Rios, Chih-Chieh Lai, Bo-Rong Yan, Bo-Cheng Lai. "Parametric Study o...	<1%
103	Publication	H L Gururaj, Francesco Flammini, V Ravi Kumar, N S Prema. "Recent Trends in He...	<1%
104	Submitted works	Indian Institute of Technology, Madras on 2024-07-02	<1%
105	Submitted works	Jeju National University on 2023-04-03	<1%
106	Publication	Lecture Notes in Computer Science, 2015.	<1%
107	Submitted works	Liverpool Hope on 2023-09-23	<1%
108	Publication	Mikhail Kopeliovich, Yuriy Mironenko, Mikhail Petrushan. "Architectural Tricks fo...	<1%

109	Submitted works	Napier University on 2024-04-17	<1%
110	Publication	Rencheng Song, Senle Zhang, Chang Li, Yunfei Zhang, Juan Cheng, Xun Chen. "He...	<1%
111	Submitted works	Universiti Teknologi Petronas on 2018-03-27	<1%
112	Submitted works	University of Greenwich on 2018-04-19	<1%
113	Submitted works	University of Northumbria at Newcastle on 2024-01-14	<1%
114	Submitted works	University of St Andrews on 2025-04-04	<1%
115	Submitted works	University of Sydney on 2023-11-27	<1%
116	Submitted works	University of York on 2024-03-25	<1%
117	Publication	Wellington Pinheiro dos Santos, Juliana Carneiro Gomes, Maíra Araújo de Santan...	<1%
118	Publication	Yang Liu, Xiang Guo, Yuzhong Zhang. "Lightweight and interpretable convolution...	<1%
119	Publication	Yassine Ouzar, Djamaledine Djeldjli, Frédéric Bousefsaf, Choubeila Maaoui. "X-iP...	<1%
120	Publication	Zexing Zhang, Huimin Lu, Zhihai He, Adil Al-Azzawi, Songzhe Ma, Chenglin Lin. "S...	<1%
121	Publication	Zihao Hao, Jinnan Wang, Ge Zhang, Lizhen Gao, Xiaoming Zhang, Jun Liu, Xiaolian...	<1%
122	Internet	comparativethesis.blog.fc2.com	<1%

123	Internet	durham-repository.worktribe.com	<1%
124	Internet	escholarship.org	<1%
125	Internet	iris.polito.it	<1%
126	Internet	mafiadoc.com	<1%
127	Internet	repository.lib.ncsu.edu	<1%
128	Internet	stax.strath.ac.uk	<1%
129	Internet	uia.brage.unit.no	<1%
130	Submitted works	universiteteknologimara on 2025-02-17	<1%
131	Internet	www.frontiersin.org	<1%
132	Internet	www.mathicando.com	<1%
133	Publication	Jun Seong Lee, Gyutae Hwang, Moonwook Ryu, Sang Jun Lee. "LSTC-rPPG: Long S...	<1%
134	Submitted works	Monash University on 2024-10-28	<1%
135	Publication	Roberto Cittadini, Luca Rosario Buonocore, Eloise Matheson, Mario Di Castro, Lor...	<1%
136	Submitted works	University College London on 2018-09-07	<1%

137	Submitted works	University of Northumbria at Newcastle on 2021-05-27	<1%
138	Submitted works	University of Strathclyde on 2016-04-01	<1%
139	Submitted works	Delhi Technological University on 2025-05-06	<1%
140	Submitted works	Imperial College of Science, Technology and Medicine on 2021-09-02	<1%
141	Publication	JongSong Ryu, SunChol Hong, Shili Liang, SinIl Pak, Qingyue Chen, Shifeng Yan. "...	<1%
142	Publication	Kulkarni, Kimaya Milind. "Deep Dive into Mitigating Bias in Remote Plethysmogra...	<1%
143	Publication	Shalli Rani, Ayush Dogra, Ashu Taneja. "Smart Computing and Communication fo...	<1%
144	Publication	Thangaprakash Sengodan, Sanjay Misra, M Murugappan. "Advances in Electrical ...	<1%
145	Publication	Tomasi, Julia. "Genetic Investigation of Physiological Phenotypes to Understand A...	<1%
146	Submitted works	University College London on 2023-10-10	<1%
147	Submitted works	University of Sydney on 2023-05-11	<1%
148	Submitted works	University of the West Indies on 2022-03-31	<1%