

Intelligent Remote Photoplethysmography-Based Methods for Heart Rate Estimation from Face Videos: A Survey

Smera Premkumar and Duraisamy Jude Hemanth *

Department of ECE, Karunya Institute of Technology and Sciences, Coimbatore 641 114, India

* Correspondence: judehemanth@karunya.edu

Abstract: Over the last few years, a rich amount of research has been conducted on remote vital sign monitoring of the human body. Remote photoplethysmography (rPPG) is a camera-based, unobtrusive technology that allows continuous monitoring of changes in vital signs and thereby helps to diagnose and treat diseases earlier in an effective manner. Recent advances in computer vision and its extensive applications have led to rPPG being in high demand. This paper specifically presents a survey on different remote photoplethysmography methods and investigates all facets of heart rate analysis. We explore the investigation of the challenges of the video-based rPPG method and extend it to the recent advancements in the literature. We discuss the gap within the literature and suggestions for future directions.

Keywords: remote photoplethysmography; face videos; deep learning; heart rate

Citation: Premkumar, S.; Hemanth, D.J. Intelligent Remote Photoplethysmography-Based Methods for Heart Rate Estimation from Face Videos: A Survey. *Informatics* **2022**, *9*, 57. <https://doi.org/10.3390/informatics9030057>

Academic Editor: Jiang Bian

Received: 24 May 2022

Accepted: 28 July 2022

Published: 7 August 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Heart rate (HR) is a performance indicator of a person's total cardiac output and a prospective clinical diagnosis tool. The gold standard for analyzing cardiac measurements is an electrocardiogram (ECG), which measures the electrical activity of the heart through sensors (called electrodes) attached to the skin. These electrodes are connected by wires to an ECG recording machine, and this type of contact measurement is appropriate for clinical setting. Another method is photoplethysmography (PPG), an optical and non-invasive technique that detects the changes in the blood volume pulse (BVP) in peripheral blood vessels via contact sensors attached to anatomical locations such as the wrists, fingers, and toes. Commercial wearable devices such as fitness trackers and smart-watches make use of this principle, where a sensor emits light to the skin and measures the reflected light intensity due to the optical absorption of blood [1]. Even though these methods are invasive, they require skin contact, which can cause discomfort, especially in neonates [2] and elderly care.

However, in recent years, remote measurement of HR has been a prominent research topic that measures the heart rate (HR) from face images and videos by analyzing tiny color variations or body movement [3]. This is a practical application of PPG technology in a completely non-invasive manner and is referred to as remote photoplethysmography (rPPG). It can predict not only the heart rate but also other vital information, such as heart rate variability and blood pressure, thereby inferring mental stress [4], variations in cardiovascular functions, quality of sleep [5], and drowsiness [6]. The advent of the digital camera brought this remote method to the masses. Remote heart rate monitoring applications have spread across the following fields:

- Hospital care [7];
- Telemedicine [8,9];
- Fitness assessment [10,11];
- Motion recognition [12];

- Automotive [13,14].

This paper aims to provide a critical review of state-of-art signal processing techniques and the learning-based algorithms in remote photoplethysmography. We discuss the challenges in rPPG measurements and recent advancements in the process.

2. Outline

Section 3 focuses on the motivations and problem statements of this work. The rest of this paper is organized as follows. Section 4 explains the rationale behind remote photoplethysmography and methods. We review the signal processing-based rPPG methods and learning-based methods in Section 5 and show publicly available datasets in the literature. Section 6 discusses the preceding sections and challenges. Finally, we conclude and confer with the research gap and future aspects in Section 7.

3. Motivations and Problem Statement

This paper, inspired by recent advances in remote photoplethysmography, allows for predicting the heart rate from skin color variations due to blood flow from face videos, as it is invisible to the human eye. A breakthrough development of signal processing and deep learning methods made rPPG an efflorescence in the current literature.

Although appreciable progress has made in rPPG methods in the last few years, few challenges still remain open, such as motion robustness, illumination, skin tone, and compression artifacts. Some relevant reviews of signal processing-based methods can be found in [15–20]. This paper focuses on improving understanding of the physiological phenomena represented within remote PPG. We set this paper's sights on two aims:

- To discuss rPPG measurement using signal processing methods as well as its recent furtherance in the deep learning environment;
- To harness the insight into the challenges on rPPG, and we anticipate some suggestions on the future direction.

4. Remote Photoplethysmography

Photoplethysmography (PPG) is a noninvasive optical technique that is used to detect volumetric changes in blood in the microvascular bed of tissue [21]. This method reckons with the principle that optical absorption of human skin varies with the blood volume pulse (BVP), which measures the amount of blood flowing through the tissues with each heartbeat. Human skin has three layers: the epidermis with capillaries, dermis with arterioles, and hypodermis with arteries [22]. When the skin is exhibited to light with a specific wavelength, the epidermis and dermis layer scatter light, whereas the hypodermis diffuses light [23,24]. In consonance with Lambert's law of light intensity [25], the light reflected through the skin can be contemplated in the process of diffusion and scattering.

Remote photoplethysmography (rPPG) is a contactless measurement that makes use of the PPG principle. It relies on a camera and then measures red, green, and blue light reflection changes from the skin as the contrast between specular and diffusion reflections, as demonstrated in Figure 1. Images or videos of human skin under ambient light sources or with dedicated illumination are recorded and processed to recover the plethysmography signal from which physiological parameters are extracted. A diffuse reflection component carries the information of PPG as it diffuses through the skin, whereas a specular reflection component is the one scattered by the surface of the skin. Even though the specular component has no pulse information, the total reflected light observed by the camera depends on the relative contribution of both components.

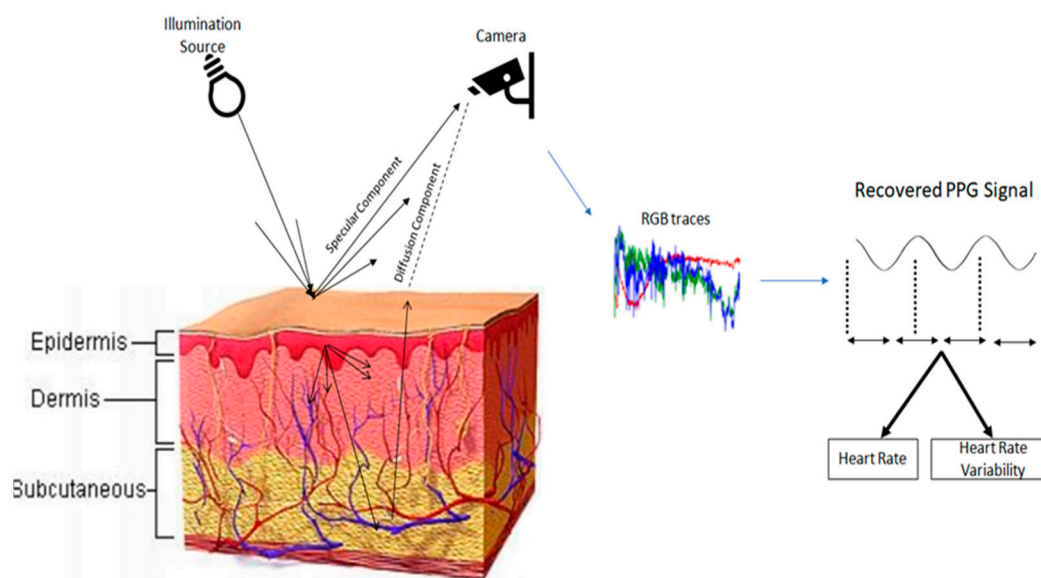


Figure 1. Illustration of remote photoplethysmography from face videos. This method quantifies the contrast between specular and diffusion reflection components and measures the changes in red, blue and green light reflected from the skin due to blood flow. A physiological signal is procured using suitable a computational approach and thereby predicting vital information.

In essence, the changes in blood volume during a cardiac cycle would cause minute color changes on the skin. Although these changes are invisible to the human eye, they could be captured by optical sensors. Accurate measurement of these changes generates a plethysmography signal, from which vital signs of the body such as the heart rate, heart rate variability, and respiration rate could be measured.

5. Remote Methods for HR Detection

A camera is capable of seizing the subtle pulsation of human skin due to blood circulation and could produce red, green, and blue raw signal traces by sampling different regions of the optical spectrum. These raw signals are then processed to obtain a plethysmography signal which contains physiological information. The existing remote PPG methods for HR measurement from human face videos can be classified as shown in Figure 2.

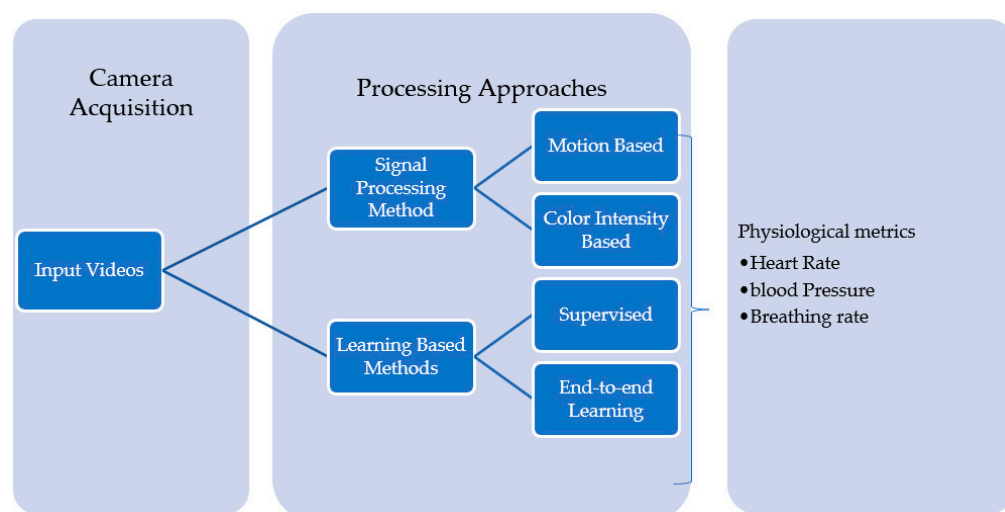


Figure 2. Classification of computational techniques used in remote photoplethysmography for recovering physiological information from videos. From the input video stream, each frame is analyzed with different computational approaches to obtain vital information. The signal processing

method is an unsupervised approach to processing input video frames, and it is classified into motion-based and color intensity-based methods. The learning-based approach is the recent trend in technology, and based on the perspective of workflow, it could be classified as a supervised (hybrid) approach and end-to-end learning.

The motion-based method for detecting a heart rate (HR) emanated from the ballistocardiogram [26]. This explains the relation between cardiac output and the amplitude of human body movements. Later, heart rate measurement using ballistocardiography (BCG) motion of the head with a wearable device was explained in [27]. Sooner, the possibility of heart rate detection from face videos by measuring subtle head motion due to the influx of blood at each beat was shown in [28]. In this method, a combination of principal component analysis (PCA) and the filtering method was used to identify the individual beats and evaluate them in 18 subjects. A Viola-Jones face detector [29] is used for region of interest (ROI) detection.

A motion-based method was explained in [30], using a single ROI and independent component analysis (ICA) subsequently. The technological improvements using BCG methods were scrutinized in [31], and it was concluded that more studies were needed to mitigate motion artifact challenges. Although these motion-based methods are invariant to illumination, the voluntary head motion and complex facial expressions could degrade the reliability of this method.

In this paper, we focus on color intensity-based methods because of their increasing attention in the literature, since they enable heart rate detection from a simple camera with ambient light as an illumination source. These methods detect heart rates from camera recordings with the help of different image and signal processing techniques. The possibility of non-contact physiological computation using a thermal camera was introduced in [32], and it was demonstrated that plethysmography signals could be measured from the human face from simple consumer-level camera recordings with ambient light conditions [33]. Since then, a substantial amount of research has been conducted in remote photoplethysmography. The rPPG methods can be split into two categories according to the previous works: signal processing-based methods and learning-based methods.

5.1. Signal Processing Methods

This method is a color intensity-based approach to measuring PPG from face videos. First, a region of interest (ROI) of each frame of the input video is detected, and then the red, green, and blue channels are spatially averaged to form raw signal traces. These traces are then processed by different signal processing techniques to recover the physiological signal. The entire process can be divided into three stages as demonstrated below. An overview of the general steps in the signal processing-based approach for recovering the heart rate is illustrated in Figure 3:

- (1). Pre-processing;
- (2). Signal extraction;
- (3). Heart rate estimation (post-processing).

5.1.1. Pre-Processing

Face Detection and ROI Tracking

Since heart rate detection is based on the photoplethysmography signals, which are derived from imperceptible skin color variations caused by pulsatile flow, it is essential to process the video frames. The process starts with the extraction of the face and localizes the measurement region of interest (ROI) for each video frame. In some of the previous works, face detection has been explained manually, with a subject standing stock-still. However, most of the works have performed face detection automatically by using the Viola-Jones algorithm was explained in [29], which is based on a machine learning approach that provides a bounding box of the subject as a result. This algorithm is a

bookmark in rPPG methods, as it possesses a high detection rate and is available in the computer vision library of OpenCV and MATLAB.

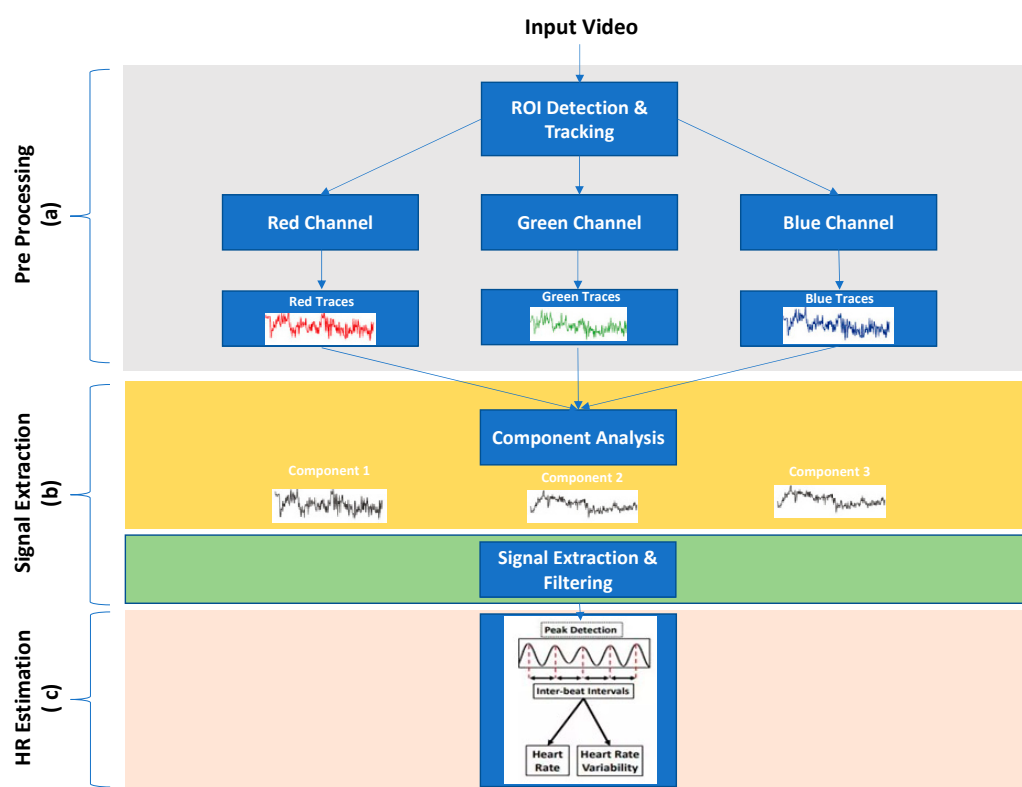


Figure 3. Workflow of signal processing method. It includes 3 stages. (a) Pre-processing is needed to obtain red, green, and blue traces from input video frames. This stage includes face tracking and ROI detection. (b) Signal extraction is performed using different signal processing algorithms, and it includes a filtering process to obtain a good quality physiological signal. (c) Heart rate estimation is the final step, where the physiological signal is processed using peak detection or frequency analysis to obtain required vital information.

Other popular algorithms used for face detection are active appearance models (AAM), a statistical model that provides facial landmarks [34], dlib [35], mtcnn [36], and the Kanade–Lucas–Tomasi approach [37,38], which provides limited assumptions about the image and possess high accuracy.

Selecting a suitable region of interest (ROI) is the next challenging step, as it has a direct impact on the accuracy and reliability of the general algorithm. ROI detection finds a set of pixels that has the most significant PPG information, and these pixels are spatially averaged to obtain the plethysmography signal [39].

Several studies were explained the quality of the ROI having a direct influence on the quality of the signal. Heart rate estimation utilizing the whole face has been proposed in some of previous works, although eye movements near the eye area may cause artifacts. Due to the high amount of light absorption, the skin regions with capillaries would produce a strong signal [40]. However, many researchers selected the forehead and cheeks [41–43] as the most significant ROI areas, as they are less susceptible to muscle movements compared with other regions of the face. Table 1 describes the summary of different methods of face selection and ROI detection. The authors of [44] were proposed that the forehead and cheeks would be computationally efficient ROIs. They divided and analyzed different face regions and evaluated the quality by using evaluation matrices.

Raw Signal Trace Extraction

To obtain the raw signal traces, the detected ROIs were separated into RGB channels. Then, the three channels were averaged spatially over all the pixels to obtain the red, green, and blue signal traces. Subsequent processing would be performed on these raw traces.

5.1.2. Signal Extraction

This stage includes filtering and dimensionality reduction. The raw signal obtained from the ROI might have unwanted noise due to motion or illumination. To remove the noise, a filtering process was performed on the raw RGB traces, and thereby the signal-to-noise ratio (SNR) would be increased. An increased SNR value provides a good quality plethysmography signal.

Filtering

Filtering is the process in which digital filters were applied to the raw signal traces based on some prior knowledge of HR frequencies. Before applying dimensionality reduction, a filtering process would be performed on the raw signals to achieve a good signal-to-noise ratio. A frequency band (0.7 Hz–4 Hz) is normally selected which leads to 42–240 beats per minute HR [45]. The filtered signal can be directly used for plethysmography signal detection [46]. According to [47], the green channel signal carries more PPG information compared with the other channels. However, the red and blue channels also carry some complementary information. In the green channel approach, the filtered green channel component is taken for further processing to obtain a PPG signal. It uses the spatially averaged pixel value of green traces and then normalizes the traces. Then, it performs an FFT to transform the signal from the spatial domain to the frequency domain and calculate the power spectral density (PSD) distribution.

Dimensionality Reduction

Dimensionality reduction methods are used to minimize the dimensionality from raw signals to achieve a more accurate and robust PPG information signal. The major classifications of the rPPG methods are based on how they extract plethysmography signals from the raw traces. The signal extraction methods can be classified broadly into three categories [48]:

- Blind source separation;
- Model-based methods;
- Design-based methods.

A PPG signal is considered a one-dimensional signal which is represented as a linear combination of the weighted sum of the raw signals, and it is taxing to estimate their weights [49]. Blind source separation (BSS) algorithms were introduced in [50], and the purpose of BSS algorithms is to separate the desired PPG signal from noise and artifacts due to statistical independence and correlation. Principle component analysis (PCA) and independent component analysis (ICA) are typical BSS techniques that are widely applied for dimensionality reduction.

An ICA-based algorithm was explained in [51] as an optimal combination of the raw signals, in which the raw signals are separated into independent non-Gaussian channels. In this method, the authors arbitrated that the second component produced after the ICA is considered a periodic one used for further processing. Several authors adopted this method in their works.

Principal component analysis (PCA) has been proposed [52], and these authors claimed the effectiveness of their approach on ICA, which may lead to the same result in some applications. Later, different methods for rPPG investigated in [53] and deferred to ICA to yield better accuracy and reliability. This BSS approach was further investigated and adopted in the literature [54], explaining the performance limits of ICA-based

techniques down the line. In the BSS method, the raw signal traces are combined, and the most periodic independent signal selected is the PP signal. The main drawback of this method is that this does not also consider motion in the given periodic signal. Thus, the major limitation of BSS can be concluded to be motion intolerance.

Chrominance-based (CHROM) algorithms [55], which belong to the model-based approach, mitigates the subject motion issues in the BSS algorithm. The authors proposed a method in which the RGB pixels in each frame of the input video have been identified using a color filter method and claimed that white illumination is successfully eliminated by the proposed skin tone standardization approach. CHROM eliminates the specular reflection component by using a color difference chrominance signal and taking advantage of the BSS method. However, both methods still do not consider illumination, as it is a significant noise source in the recovered signal.

To overcome this, the spatial subspace rotation (2SR) method was proposed in [56], which exploits the benefits of statistical measurement of multiple pixel sensors in a camera. This method is performed in both the spatial and temporal domains. First, a subspace of skin pixels is constructed, and then the rotation angle between the frames is measured to determine the PPG information. The authors claimed the 2SR method outperformed ICA and CHROM.

5.1.3. Heart Rate Estimation

The heart rate (HR) is evaluated from the recovered PPG signal either by peak detection or frequency analysis. In the peak detection approach, individual peaks are used to extract the heart rate. Later, the authors of [57] showed the physiological measurements using five-band camera sensors. Based on the error range of the reliable methods of heart rate detection, the medically tolerable accuracy is set to three beats per minute (BPM), which represents the accuracy of the rPPG method to be the same as traditional contact methods. A photoplethysmography signal is considered a time-varying intensity signal. From the resulting physiological signal, the heart rate (HR) is the inverse of the average time difference between two consecutive beats in the time domain. However, in the frequency domain, the HR is extracted with the highest energy power spectrum of the physiological signal. We could calculate the instantaneous HR by measuring the beat-to-beat HR, and this is more informative, but this requires accurate peak detection.

An automated method has been proposed to detect the peak-to-peak time between systolic and diastolic inflexion using the second-order derivative of the recovered signal. An analysis of HR detection methods was performed [58] based on the variations of the inter-beat intervals. A short-time Fourier transfer (STFT) method for HR detection was proposed in [59], and it is more effective when the heart rate pattern changes rapidly. A predictive model was also developed using workout video frames, and it would be more productive under real-time scenarios.

However, frequency analysis is the commonly adopted method in the literature. In this method, the extracted PPG signal is converted to the frequency domain using an FFT [60] or DCT [61], where Welch's method is used for density estimation. The strongest periodic signal within the frequency band is considered the signal with PPG information and computes the main heart rate over a particular period. Later, the authors of [62] introduced a generative adversarial network (GAN), a deep learning-based technique to learn rPPG noise impacts. An analysis of some of the relevant signal processing-based rPPG methods can be found in Table 1.

Table 1. Analysis of some signal processing-based methods for rPPG.

Publication	Preprocessing	Signal Extraction and HR Estimation Methods	Database	Performance and Comments
(Verkruysse, 2008)	Manual	Bandpass filter, FFT	Self-collected	Recorded with a simple digital camera and ambient light, performance measured qualitatively
(Poh, 2010)	Automated face tracker faces (Viola and Jones (VJ), Lienhart and Mad)	ICA, FFT	Self-collected	With movement artifacts, root mean square deviation (RMSE): 4.63 bpm
(Poh, 2011)	Automated face tracker	ICA, five-points moving average filter, and bandpass filter	Self-collected	RMSE: 1.24 bpm Correlation coefficient: 1.00
(Lewandowska, 2011)	Manual	Principle component analysis (PCA)	Self-collected	Pulse rate from two-color channels
(Haan, 2013)	Automatic face detection	Chrominance-based approach (fixed signal combination, FFT)	Self-collected	RMSE: 0.4 bpm
(Mannaperuma, 2015)	Automatic face detection	ICA, the channel with the strongest blood volume pulse signal is selected, inverted, and interpolated, and then peaks are detected	Self-collected	Find band camera sensor (correlation: 1.00)
(Wang, 2015)	Face detection by VJ	FFT, bandpass, and adaptive bandpass, motion-resistant Remote PPG method	Self-collected	Peak detection performance compared with ICA method using bland Altman plot
(Wang, 2015)	Manual	Spatial pruning + temporal filtering	Self-collected	SNR improvement 3.34 to 6.74 dB on state-of-the-art methods
(Wang, 2016)	Spatial distribution of skin pixel	2SR algorithm	Self-collected	Results compared with ICA, CHROM, AND PBV SNR-6.55

(Yu, 2019)	Spatial ROI selection and tracking	Novel semi-blind source extraction method, MAICA	UBFC-rPPG MMSE-HR	UBFC-rPPG (MAE-0.55BPM) MMSE-HR (MAE-3.91)
(Fouad, 2019)	Automatic face tracking	Uses BSS algorithm, FT	Self-collected	Studied factors affecting accuracy
(Gudi, 2020)	Active appearance model (AAM) Head orientation	Unsupervised method operates in real time FFT	PURE VIPL-HR CO-HAFACE	0.34 bpm 0.57 bpm 0.46 bpm

5.2. Learning-Based Methods

Signal processing-based rPPG methods were explained in the previous sections. In the literature, recent trends include learning-based PPG measurements. The major benefit is that they could detect the heart rate directly from video input, and the system learns the rPPG mechanism from the beginning. Learning-based techniques can be divided into two categories for better understanding: supervised learning methods and end-to-end learning methods. An illustration of the workflow can be seen in Figure 4. With the supervised learning approach, the feature extraction should be performed manually, whereas deep learning methods extract features directly from the input video without any human intervention.

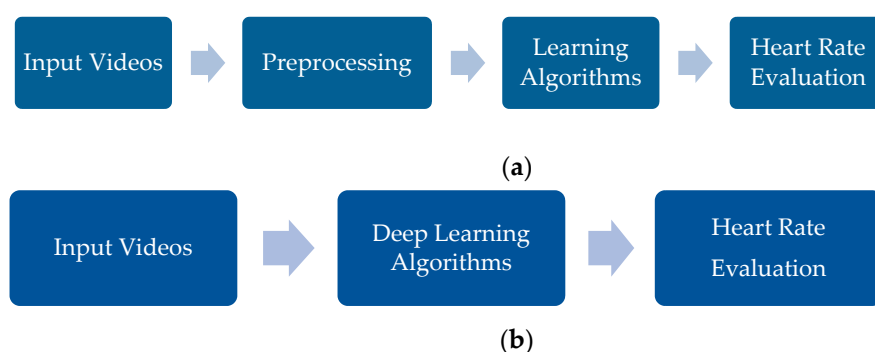


Figure 4. Functional diagram of learning-based methods. (a) The supervised method is a hybrid method where feature extraction is manually performed before feeding the input frames into the learning algorithm. (b) The end-to-end method is an unsupervised method, with no manual pre-processing needed.

5.2.1. Supervised Learning Methods

This method is a combination of both the manual and learning-based approaches, in which the preprocessing part is performed manually and the result feeds into the learning networks. The motivation to develop this algorithm is to mitigate the issues of signal processing-based methods, and it was a successful strategy to a certain extent.

A machine learning approach was proposed in [63] to improve the accuracy of the conventional method, which evaluated and compared the ICA method with two machine learning techniques: linear regression and the k-nearest neighbor (kNN) classifier in a controlled situation. Linear regression is a model between a dependent variable and explanatory variables, whereas kNN is a learning-based approach [64] that measures the training instance closest to the known test instance. The kNN takes the average heart rate of the k-nearest neighbors, and the results have shown that it would outperform the ICA method. Later on, more advanced machine learning techniques such as convolutional neural networks [65,66] and temporal neural networks were proposed.

A two-layer LSTM was explained in [67] and showed that noise signals can preserve functional signals. Synthetic signals are used to train the model, and the results are analyzed on a public domain database. A feature extraction stream can be observed in [68], which learned a robust feature representation and developed a complementary stream to extract reliable vital signals. A unified neural network was reported for estimating the HR, and performance analysis was performed using the COHFACE dataset.

A single-photon avalanche diode (SPAD) camera-based method was introduced in [69] and provided a hybrid method that analyzed the frame stream with a neural network followed by signal processing techniques for HR detection, and it showed its effectiveness in unrestrained illumination. A deep HR method was proposed in [70], and the authors also explained a machine learning approach with a frontend and backend component. The front end learns independently from training video samples, whereas the back end is a fully connected neural network for HR estimation and evaluated on two different datasets.

A Siamese rPPG network [71] proposed feature learning from two facial regions simultaneously. A two-branch model was trained jointly, and the results were evaluated on three benchmark datasets and shown to surpass the results of the existing methods.

5.2.2. End-to-End Learning-Based Approach

With the emergence of the deep learning end-to-end method, extensive opportunities are opening up for performing tasks more efficiently in a better way. The first end-to-end learning model ‘DeepPhys’ was introduced in [72], which is based on a convolutional attention network (CAN) and enables spatiotemporal visualization of the signals. This paper proposed a skin reflection model that is exceptionally robust in different illumination conditions. Since it is an end-to-end system, the intermediate steps in the state-of-the-art method could be removed successfully. The authors evaluated the proposed method on three different datasets and have shown surpassing results when compared with the state-of-the-art approaches.

Subsequently, SynRhythm was proposed [73] for HR estimations and it is an unsupervised learning based approach. Two successive convolutional neural networks (CNN) are used to extract the blood volume pulse from a sequence of images and thereby the heart rate. RhythmNet [74] exploits the CNN and gated recurrent units to form a spatiotemporal representation. A VPL-HR database [75] containing 2378 visible light subjects was introduced to study the algorithm’s robustness with motion and illumination variance. Nonetheless, a compression artifact challenge has yet to be investigated. Belatedly, a deep spatiotemporal network for regenerating the HR from videos was proposed in [76] and used the MAHNOB HCI and OBF databases for experiments. The results were evaluated and compared with RNN and 3DCNN-based PhysNet algorithms and showed better performance. Three signal processing methods, including the CHROM and POS methods, were replicated, and the results were compared with the proposed algorithm. The main advantage is that the proposed algorithm allows HRV features, and it would be a beneficial method in realistic situations.

A two-step convolutional network was introduced in [77], where it was trained by alternating optimization, and the results were validated on three publicly available datasets as well as on a newly collected dataset of 204 fitness-themed videos. However, compression is still a challenging scenario. The authors of [78] proposed a transfer learning strategy from a limited number of face videos and used a deep HR estimator from synthetic rhythm signals. This algorithm uses a sine function to represent the periodic part of the synthetic signal and limit the frequency to overcome the challenges, such as a large volume of training data and illumination. Even if the proposed approach showed effectiveness with the state-of-the-art methods, it still needs a large database for a more accurate HR.

A neural architecture called AutoHR was proposed in [79], which evaluated the convolution difference in the spatial domain. Subsequently, the authors of [80] performed a

comparative evaluation and showed the learning-based method to achieve better performance in the signal processing methods. They also showed a low error rate, which makes learning-based methods applicable in real-time scenarios. Some relevant papers on deep learning-based rPPG can be found in Table 2.

Table 2. Analysis of some learning-based methods for rPPG.

Serial No	Paper	Network	Description	Datasets
1	(Weixuan, 2018)	DeepPhys	First end-to-end network	RGB VIDEO 1 RGB VIDEO 11 MAHNOB-HCI IR VIDEO
2	(Niu, 2018)	SynRhythm	Transfer learning strategy and synthetic rhythm signals	MAHNOB-HCI MMSE-HR
3	(Spetlik. R, 2018)	HR-CNN	Uses 2-step CNN	MAHNOB-HCI PURE COHFACE
4	(Wang, 2020)	Two stream CNN	Two-stream end-to-end network	COHFACE
5	(Niu.X, 2020)	RhythmNet	End-to-end spatial-temporal representation	MAHNOB-HCI MMSE-HR VIPL-HR
6	(Yu.Z, 2020)	AutoHR	Neural architecture search (NAS)	MAHNOB-HCI MMSE-HR VIPL-HR
7	(Min Hu, 2021)	ETA-rPPGNet	Time domain attention mechanism	PURE MMSE-HR COHFACE UBFC-rPPG
8	(Hao LU, 2021)	NAS-HR	A neural network-based method	PURE VIPL-HR

An end-to-end three-dimensional (3D) spatiotemporal convolutional network was introduced in [81] which used a multi-hierarchical feature fusion-based attention module. It efficiently minimized the impact of motion and noise. Two publicly available datasets were used for evaluation, and it reconstructed the physiological signals accurately.

A three-domain segment network, ETA-rPPG Net, was illustrated in [82] along with a time domain attention module that used a convolutional kernel. A two-part loss function was proposed for supervised training, and it could effectively reduce the noise interference from illumination variation. However, despite showing better results, more robust models in low-constraint environments are still needed.

A major drawback of the learning-based approach is the large amount of data needed for training the network to achieve robustness and accuracy. To overcome this difficulty, the authors of [83] proposed an approach to training a deep HR estimator from synthetic PPG signals and a limited number of available face data. The authors showed the effectiveness of their approach using public datasets. The authors explained the effectiveness of extracting the HR from face videos deprived of video processing.

Later, the authors of [84] came up with a meta-learning approach (Meta-rPPG) that focuses on using a synthetic gradient generator, and it requires several transductive inference steps and achieves a greater accuracy than the state-of-the-art methods. A metaphysical model that works well with supervised and unsupervised models was proposed in [85] and evaluated on two different datasets. However, the performance degraded when

the subject was darker. This paper demonstrated better performance than the state-of-the-art approaches. Even if it outperformed the results of the state-of-the-art signal processing methods, it still needs manual feature extraction. The main challenges still need to be mitigated are the following:

- It requires a large volume of training data;
- Poor performance under realistic conditions;
- Low accuracy due to compression;
- Complexity due to intermediate steps.

An end-to-end model proposed in [85] using undercomplete independent component analysis U-LMA was tested under three scenarios to estimate the nonlinear cumulative density function (CDF). Another skin segmentation method was introduced in [86] to process low-resolution inputs, make use of depth-wise convolutional layers, and localize skin pixels. The authors proved the real-time better performance on a small IoT device.

6. Datasets

To evaluate the rPPG algorithms, most of the authors used privately recorded datasets which are not available publicly. DEAP is a multimodal dataset which was put forward in [87] for human emotion analysis. The authors made the dataset available publicly with the physiological signals of 32 participants and 40 videos. Later, in [88], MAHNOB-HCI, a dataset with a large collection of modalities was recorded and made open to the public. High synchronization accuracy makes this database beneficial for researchers who need to assess their methods and algorithms in challenging databases. The authors of [89] conducted analysis and evaluated different public datasets. They also introduced a cleaner PPG set with a collection of truth peaks for 13 major datasets to overcome the noise and miscalculations in public datasets.

Practically, the main challenge regarding datasets is the lack of publicly available datasets under realistic conditions. Most of the papers in the literature were assessed on privately owned databases, which makes it difficult to generalize the algorithms. Selections of datasets that are publicly available are shown in Table 3.

Table 3. Publicly available datasets.

Dataset	Subject	Camera	Physiological Signal	
PURE	10 Subjects 59 Videos	480p@30fps Lossless PNG images	Ground truth PPG @60 Hz	Recorded Movement such as talking, rotation, translation
MAHNOB HCI	27 Subjects 627 Videos	780 × 580P@51fps H.264 format	Ground Truth PPG @256 Hz	Subject recorded while watching video stimuli
COHFACE	40 Subjects 164 Videos	480p@20fps MPEG4 Part 2format	Ground Truth PPG @256 Hz	Subject recorded illuminated by a spotlight and natural light
MMSE-HR	40 Subjects 102 Videos	1040 × 1392@25fps JPEG Images	Instantaneous HR@1 kHz	Part of a large multimodal corpus, subject exhibit facial expressions
Vicar PPG	10 Subjects 20 Videos	720p@30fps H.264 format	Ground Truth PPG @30 Hz	Subject recorded before and after workout

UBFC – RPPG	42 Subjects 42 Videos	480p@30fps Raw video format (lossless)	Ground Truth PPG @30/60 Hz	Subject recorded while playing game
--------------------	--------------------------	---	-------------------------------	---

7. Challenges

The preceding sections explained different approaches to HR detection from face videos. From the literature, it is clear that learning-based methods are robust and flexible and work better in practical applications. Since remote photoplethysmography is a camera-based technology, certain challenges such as skin melanin tone, illumination conditions, subject motion, and compression impacts need to be addressed for accurate measurement of heart rates. In the literature, we could find different works carried out to overcome these challenges. Deep learning networks can overcome these limitations to an extent by training large datasets.

The influence of the compression schemes of motion in different video formats has been investigated, and the quality loss against compression artifacts was investigated [90], addressing the compression problem in detail and evaluating the significant decrease in performance of rPPG algorithms with the increase in compression. The authors observed the compression to degrade the accuracy of the measured physiological signals in real-time processing. Since most of the datasets were recorded under laboratory settings with good conditions, the rPPG-based HR gives better results compared with the traditional contact-based techniques. However, in real-world applications, video compression is inevitable, as it helps to reduce storage, transmission time, and bandwidth. The videos captured through the commercial cameras undergo different compression codecs and bitrates, and so the frames observed from a camera significantly affected the compression artifacts. Since compression plays an important role in signal detection, the compression artifact impact remains open, and only a few pieces of research have been carried out in this area.

In [91], the authors explained the types of compression artifacts and proposed a single-channel framework to reduce the effects of compression. They claimed that the red and blue color components are the ones most affected by video compression due to the low bit rate. The authors of [85] developed a STVEN autoencoder to convert video from one bit rate to another. They performed an image enhancement procedure to overcome the compression effects. Subsequently, the authors of [92] proposed a deep super-resolution network for low-resolution video which enhanced the rPPG method in compressed video and conducted a performance analysis at varying compression levels and in different formats. The authors proposed an approach to recover PPG signals from compressed videos rather than enhancing them and also evaluated the effects of compression on different skin types. However, the authors did not consider the effects of compression from motion.

To sum up, video compression degrades the quality of PPG measurements, since it relies upon subtle changes in the signal from the camera. However, the compression does not affect the quality of the videos, as it is typically optimized for visual quality. Since the remote methods consider minute changes in the signal, it is important to develop methods that can mitigate compression loss. Other significant gaps can be seen in the data and privacy concerns.

7.1. Data Implication

Different datasets contain different amounts of motion, resulting in the difficulty of generalizing an algorithm. It is important to have benchmarks to evaluate the efficiency of different approaches [93]. A public benchmark dataset, Vicar Vision, has been developed to overcome the reproducibility problem in rPPG research, which defines the illumination and motion challenges. There is no benchmark dataset available to address the challenges in the rPPG environment. Another issue is skin tone, as greater amounts of

melanin absorb more light than other skin types, and thus the pixels may become saturated. This results in a weaker physiological signal measurement.

Most of the datasets contain lighter skin tone participants because they were collected from European countries and the United States of America. A meta-analysis method explained the significant drop in performance for darker skin tones. To study the impact, the authors combined three datasets with different participants and concluded that datasets with better representation are needed for more accurate vital sign measurements using rPPG. The skin tone biases in the rPPG environment were investigated in [94], and a physically driven approach was proposed in [95].

7.2. Privacy Concern

Since this is camera-based technology, there is potential risk in terms of ethics and the privacy of the subject. Researchers proposed innovative methods to mitigate this concern. The Privacy-Phys model was proposed in [96] based on a pretrained model of a 3D CNN. A novel algorithm, pulse edit, was proposed in [97] to edit the facial video physiological signal to protect the subject's privacy disclosure.

8. Conclusions

In this paper, we performed a critical review of different remote photoplethysmography methods for heart rate detection from facial videos. This survey also aids in highlighting the advantages and disadvantages of different techniques and approaches to HR detection. Additionally, we observed the impact of compression artifacts on rPPG methods and reviewed some works that took video compression into account. A significant research gap can be seen in the literature for taking compression into consideration. Another crucial challenge that needs to be addressed is the performance gap between skin color tones, as this plays a key role in real-time scenarios. We hope that recent advancements in neural networks can help to mitigate the current issues. In our future work, we would like to develop some hybrid approaches to increase the accuracy and investigate the possibilities of advancing remote methods by using neural models to alleviate the existing challenges.

Author Contributions: Conceptualization, S.P.; Supervision, D.J.H.; Review and Editing: S.P. and D.J.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Biswas, D.; Simoes-Capela, N.; Van Hoof, C.; Van Helleputte, N. Heart Rate Estimation from Wrist-Worn Photoplethysmography: A Review. *IEEE Sens. J.* **2019**, *19*, 6560–6570. <https://doi.org/10.1109/JSEN.2019.2914166>.
2. Scalise, L.; Bernacchia, N.; Ercoli, I.; Marchionni, P. Heart rate measurement in neonatal patients using a web camera. In Proceedings of the MeMeA 2012—2012 IEEE Symposium on Medical Measurements and Applications, Budapest, Hungary, 18–19 May 2012; pp. 6–9. <https://doi.org/10.1109>.
3. Benedetto, S.; Caldato, C.; Greenwood, D.C.; Bartoli, N.; Pensabene, V.; Actis, P. Remote heart rate monitoring—Assessment of the Face reader rPPg by Noldus. *PLoS ONE* **2019**, *14*, e0225592. <https://doi.org/10.1371/journal.pone.0225592>.
4. Kuncoro, C.B.D.; Luo, W.-J.; Kuan, Y.-D. Wireless Photoplethysmography Sensor for Continuous Blood Pressure Bio signal Shape Acquisition. *J. Sens.* **2020**, *2020*, 7192015. <https://doi.org/10.1155/2020/7192015>.
5. Hilmisson, H.; Berman, S.; Magnusdottir, S. Sleep apnea diagnosis in children using software-generated apnea-hypopnea index (AHI) derived from data recorded with a single photoplethysmogram sensor (PPG): Results from the Childhood Adenotonsillectomy Study (CHAT) based on cardiopulmonary coupling analysis. *Sleep Breath.* **2020**, *24*, 1739–1749. <https://doi.org/10.1007/s11325-020-02049-6>.

6. Wilson, N.; Guragain, B.; Verma, A.; Archer, L.; Tavakolian, K. Blending Human and Machine: Feasibility of Measuring Fatigue Through the Aviation Headset. *Hum. Factors* **2020**, *62*, 553–564. <https://doi.org/10.1177/0018720819849783>.
7. Yu, X.; Laurentius, T.; Bollheimer, C.; Leonhardt, S.; Antink, C.H. Noncontact Monitoring of Heart Rate and Heart Rate Variability in Geriatric Patients Using Photoplethysmography Imaging. *IEEE J. Biomed. Health Inform.* **2021**, *25*, 1781–1792. <https://doi.org/10.1109/JBHI.2020.3018394>.
8. Sasangohar, F.; Davis, E.; Kash, B.A.; Shah, S.R. Remote patient monitoring and telemedicine in neonatal and pediatric settings: Scoping literature review. *J. Med. Internet Res.* **2018**, *20*, e295. <https://doi.org/10.2196/jmir.9403>.
9. Poh, M.Z.; McDuff, D.J.; Picard, R.W. Advancements in noncontact, multiparameter physiological measurements using a webcam. *IEEE Trans. Biomed. Eng.* **2011**, *58*, 7–11. <https://doi.org/10.1109/TBME.2010.2086456>.
10. Sinhal, R.; Singh, K.; Raghuwanshi, M.M. An Overview of Remote Photoplethysmography Methods for Vital Sign Monitoring. *Adv. Intell. Syst. Comput.* **2020**, *992*, 21–31. https://doi.org/10.1007/978-981-13-8798-2_3.
11. Chang, M.; Hung, C.-C.; Zhao, C.; Lin, C.-L.; Hsu, B.-Y. Learning based Remote Photoplethysmography for Physiological Signal Feedback Control in Fitness Training. In Proceedings of the 15th IEEE Conference on Industrial Electronics and Applications (ICIEA), 2020; pp. 1663–1668. <https://doi.org/10.1109/ICIEA48937.2020.9248164>.
12. Zaunseder, S.; Trumpp, A.; Wedekind, D.; Malberg, H. Cardiovascular assessment by imaging photoplethysmography—a review. *Biomed. Tech.* **2018**, *63*, 529–535. <https://doi.org/10.1515/bmt-2017-0119>.
13. Huang, P.W.; Wu, B.J.; Wu, B.F. A Heart Rate Monitoring Framework for Real-World Drivers Using Remote Photoplethysmography. *IEEE J. Biomed. Health Inform.* **2021**, *25*, 1397–1408. <https://doi.org/10.1109/JBHI.2020.3026481>.
14. Wu, B.F.; Chu, Y.W.; Huang, P.W.; Chung, M.L. Neural Network Based Luminance Variation Resistant Remote-Photoplethysmography for Driver's Heart Rate Monitoring. *IEEE Access* **2019**, *7*, 57210–57225. <https://doi.org/10.1109/ACCESS.2019.2913664>.
15. Hoffman, W.F.C.; Lakens, D. Addressing Reproducibility Issues in Remote Photoplethysmography (rPPG) Research: An Investigation of Current Challenges and Release of a Public Algorithm Benchmarking Dataset. 25 June 2021. <https://doi.org/10.17605/OSF.IO/XJF7U>. Available online: <https://data.4tu.nl/repository/uuid:2ac74fbd-2276-44ad-aff1-2f68972b7b51> (accessed on 20 March 2022).
16. Gupta, Y.; Kaur, A.; Arora, A.; Kapoor, S.; Gupta, M. Heart-Rate Evaluation Using Remote Photoplethysmography—A Case Study. In Proceedings of the International Conference on Innovative Computing & Communications (ICICC), 18 May 2020. <https://doi.org/10.2139/ssrn.3604112>.
17. McDuff, D.J.; Estep, J.R.; Piasecki, A.M.; Blackford, E.B. A survey of remote optical photoplethysmography imaging methods. In Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, November 2015; pp. 6398–6404. <https://doi.org/10.1109/EMBC.2015.7319857>.
18. Rouast, P.V.; Adam, M.T.P.; Chiong, R.; Cornforth, D.; Lux, E. Remote heart rate measurement using low-cost RGB face video: A technical literature review. *Front. Comput. Sci.* **2018**, *12*, 858–872. <https://doi.org/10.1007/s11704-016-6243-6>.
19. van der Kooij, K.M.; Naber, M. An open-source remote heart rate imaging method with practical apparatus and algorithms. *Behav. Res. Methods* **2019**, *51*, 2106–2119. <https://doi.org/10.3758/s13428-019-01256-8>.
20. Wang, C.; Pun, T.; Chanel, G. A comparative survey of methods for remote heart rate detection from frontal face videos. *Front. Bioeng. Biotechnol.* **2018**, *6*, 1–16. <https://doi.org/10.3389/fbioe.2018.00033>.
21. Allen, J. Photoplethysmography and its application in clinical physiological measurement. *Physiol. Meas.* **2007**, *28*, R1. <https://doi.org/10.1088/0967-3334/28/3/R01>.
22. Liu, J.; Yan, B.P.-Y.; Dai, W.-X.; Ding, X.-R.; Zhang, Y.-T.; Zhao, N. Multi-wavelength photoplethysmography method for skin arterial pulse extraction. *Biomed. Opt. Express* **2016**, *7*, 4313. <https://doi.org/10.1364/boe.7.004313>.
23. De Haan, G.; Van Leest, A. Improved motion robustness of remote-PPG by using the blood volume pulse signature. *Physiol. Meas.* **2014**, *35*, 1913–1926. <https://doi.org/10.1088/0967-3334/35/9/1913>.
24. Huelsbusch, M.; Blazek, V. Contactless mapping of rhythmical phenomena in tissue perfusion using PPGI. In Proceedings of the Medical Imaging 2002: Physiology and Function from Multidimensional Images, 2002; Volume 4683, p. 110. <https://doi.org/10.1117/12.463573>.
25. Kevin Zhou, S.; Chellappa, R.; Ramanathan, N. (2009). Unconstrained Face Recognition from a Single Image. In *The Essential Guide to Image Processing*, 1st ed.; Elsevier: Amsterdam, The Netherlands, 2009. <https://doi.org/10.1016/B978-0-12-374457-9.00024-X>.
26. Starr, I.; Rawson, A.J.; Schroeder, H.A.; Joseph, N.R. Studies on the estimation of cardiac output in man, and of abnormalities in cardiac function, from the heart's recoil and the blood's impacts; the ballistocardiogram. *Am. J. Physiol. -Leg. Content* **1939**, *127*, 1–28. <https://doi.org/10.1152/ajplegacy.1939.127.1.1>.
27. Da He, D.; Winokur, E.S.; Sodini, C.G. A continuous, wearable, and wireless heart monitor using head ballistocardiogram (BCG) and head electrocardiogram (ECG). In Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, 2011; pp. 4729–4732. <https://doi.org/10.1109/IEMBS.2011.6091171>.
28. Balakrishnan, G.; Durand, F.; Guttag, J. Detecting pulse from head motions in video. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2013; pp. 3430–3437. <https://doi.org/10.1109/CVPR.2013.440>.
29. Viola, P.; Jones, M.J. Robust Real-Time Face Detection. *Int. J. Comput. Vis.* **2004**, *57*, 137–154. <https://doi.org/10.1023/B:VISI.0000013087.49260.fb>.

30. Shan, L.; Yu, M. Video-based heart rate measurement using head motion tracking and ICA. In Proceedings of the 2013 6th International Congress on Image and Signal Processing, CISP, 2013; Volume 1, pp. 160–164. <https://doi.org/10.1109/CISP.2013.6743978>.
31. Inan, O.T. (2012). Recent advances in cardiovascular monitoring using ballistocardiography. In Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, 2012; pp. 5038–5041. <https://doi.org/10.1109/EMBC.2012.6347125>.
32. Pavlidis, I.; Dowdall, J.; Sun, N.; Puri, C.; Fei, J.; Garbey, M. Interacting with human physiology. *Comput. Vis. Image Underst.* **2007**, *108*, 150–170. <https://doi.org/10.1016/j.cviu.2006.11.018>.
33. Verkrusse, W.; Svaasand, L.O.; Nelson, J.S. Remoteplethysmography imaging using ambient light. *Opt. Express* **2008**, *16*, 21434–21445.
34. Cootes, T.F.; Edwards, G.J.; Taylor, C.J. Active appearance models. In Proceedings of the European Conference on Computer Vision (ICCV), 1998; pp. 484–498.
35. Kazemi, V.; Sullivan, J. One millisecond face alignment with an ensemble of regression trees. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2014; pp. 1867–1874. <https://doi.org/10.1109/CVPR.2014.241>.
36. Zhang, K.; Zhang, Z.; Li, Z.; Qiao, Y. Joint Face Detection and Alignment Using Multi-Task Cascaded Convolutional Networks. *IEEE Signal Process. Lett.* **2016**, *23*, 1499–1503. <https://doi.org/10.1109/LSP.2016.2603342>.
37. Tomasi, C.; Kanade, T. Detection and tracking of point features. *Tech. Rep. Int. J. Comput. Vision* 1991, *9*, 137–154.
38. Qian, R.J.; Sezan, M.I.; Matthews, K.E. A robust real-time face tracking algorithm. In Proceedings of the International Conference on Image Processing (ICIP), 1998; Volume 1, pp. 131–135.
39. Kwon, S.; Kim, J.; Lee, D.; Park, K. ROI analysis for remote photoplethysmography on facial video. In Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, November 2015; pp. 4938–4941. <https://doi.org/10.1109/EMBC.2015.7319499>.
40. Van Gastel, M.; Stuijk, S.; de Haan, G. Robust respiration detection from remote photoplethysmography. *Biomed. Opt. Express* **2016**, *7*, 4941. <https://doi.org/10.1364/boe.7.004941>.
41. Hassan, M.A.; Malik, G.S.; Saad, N.; Karasfi, B.; Ali, Y.S.; Fofi, D. Optimal source selection for image photoplethysmography. In Proceedings of the IEEE International Instrumentation and Measurement Technology Conference (I2MTC), Taipei, Taiwan, 2016; pp. 1–5. <https://doi.org/10.1109/I2MTC.2016.7520406>.
42. Tasli, H.E.; Gudi, A.; Uyl, M.D. Remote ppg based vital sign measurement using adaptive facial regions Vicarious Perception Technologies Intelligent Systems Lab Amsterdam, University of Amsterdam, The Netherlands. In Proceedings of the International Conference on Image Processing (ICIP), 2014; pp. 1410–1414.
43. ElMaghraby, A.; Abdalla, M.; Enany, O.; El Nahas, M.Y. Detect and Analyze Face Parts Information using Viola- Jones and Geometric Approaches. *Int. J. Comput. Appl.* **2014**, *101*, 23–28. <https://doi.org/10.5120/17667-8494>.
44. Holton, B.D.; Mannapperuma, K.; Lesniewski, P.J.; Thomas, J.C. Signal recovery in imaging photoplethysmography. *Physiol. Meas.* **2013**, *34*, 1499–1511. <https://doi.org/10.1088/0967-3334/34/11/1499>.
45. Bousefsaf, F.; Maaoui, C.; Pruski, A. Continuous wavelet filtering on webcam photoplethysmographic signals to remotely assess the instantaneous heart rate. *Biomed. Signal Process. Control* **2013**, *8*, 568–574. <https://doi.org/10.1016/j.bspc.2013.05.010>.
46. Qi, H.; Wang, Z.J.; Miao, C. Non-contact driver cardiac physiological monitoring using video data. In Proceedings of the 2015 IEEE China Summit and International Conference on Signal and Information Processing, ChinaSIP 2015, 2015; pp. 418–422. <https://doi.org/10.1109/ChinaSIP.2015.7230436>.
47. Poh, M.Z.; McDuff, D.J.; Picard, R.W. Non-contact, automated cardiac pulse measurements using video imaging and blind source separation. *Opt. Express* **2010**, *18*, 10762–10774. <https://doi.org/10.1364/OE.18.010762>.
48. Takano, C.; Ohta, Y. Heart rate measurement based on a time-lapse image. *Med. Eng. Phys.* **2007**, *29*, 853–857. <https://doi.org/10.1016/j.medengphy.2006.09.006>.
49. Djeldjli, D.; Bousefsaf, F.; Maaoui, C.; Bereksi-Reguig, F. Imaging Photoplethysmography: Signal Waveform Analysis. In Proceedings of the 2019 10th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, IDAACS 2019, 2019; Volume 2, pp. 830–834. <https://doi.org/10.1109/IDAACS.2019.8924239>.
50. Wedekind, D.; Trumpp, A.; Gaetjen, F.; Rasche, S.; Matschke, K.; Malberg, H.; Zaunseder, S. Assessment of blind source separation techniques for video-based cardiac pulse extraction. *J. Biomed. Opt.* **2017**, *22*, 035002. <https://doi.org/10.1117/1.jbo.22.3.035002>.
51. Mannapperuma, K.; Holton, B.D.; Lesniewski, P.J.; Thomas, J.C. Performance limits of ICA-based heart rate identification techniques in imaging photoplethysmography. *Physiol. Meas.* **2015**, *36*, 67–83. <https://doi.org/10.1088/0967-3334/36/1/67>.
52. Lewandowska, M.; Rumiński, J.; Kocejko, T.; Nowak, J. Measuring pulse rate with a webcam—A non-contact method for evaluating cardiac activity. In Proceedings of the 2011 Federated Conference on Computer Science and Information Systems, FedCSIS, 2011; pp. 405–410.
53. Li, X.; Chen, J.; Zhao, G.; Pietikäinen, M. Remote heart rate measurement from face videos under realistic situations. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2014; pp. 4264–4271. <https://doi.org/10.1109/CVPR.2014.543>.
54. Singh, K.R.; Gupta, K. Color Intensity: A Study of RPPG Algorithm for Heart Rate Estimation Color Intensity : A Study of RPPG Algorithm for Heart Rate Estimation. December. 2020. <https://doi.org/10.51201/10288>

55. De Haan, G.; Jeanne, V. Robust pulse rate from chrominance-based rPPG. *IEEE Trans. Biomed. Eng.* **2013**, *60*, 2878–2886.
56. Wang, W.; Stuijk, S.; De Haan, G. A Novel Algorithm for Remote Photoplethysmography: Spatial Subspace Rotation. *IEEE Trans. Biomed. Eng.* **2016**, *63*, 1974–1984. <https://doi.org/10.1109/TBME.2015.2508602>.
57. McDuff, D.; Gontarek, S.; Picard, R.W. Improvements in remote cardiopulmonary measurement using a five-band digital camera. *IEEE Trans. Biomed. Eng.* **2014**, *61*, 2593–2601. <https://doi.org/10.1109/TBME.2014.2323695>.
58. McDuff, D.; Gontarek, S.; Picard, R.W. Remote detection of photoplethysmographic systolic and diastolic peaks using a digital camera. *IEEE Trans. Biomed. Eng.* **2014**, *61*, 2948–2954. <https://doi.org/10.1109/TBME.2014.2340991>.
59. Yu, Y.; Kwan, B.; Lim, C.; Wong, S.; Paramesran, R. Video-based heart rate measurement using a short-time Fourier transform. In Proceedings of the International Symposium on Intelligent Signal Processing and Communication Systems, 2013; pp. 704–707.
60. Feng, L.; Po, L.; Xu, X.; Li, Y.; Ma, R. Motion-Resistant Remote Imaging Photoplethysmography Based on the Optical Properties of Skin. *IEEE Trans. Circuits Syst. Video Technol.* **2015**, *25*, 879–891.
61. Irani, R.; Nasrollahi, K.; Moeslund, T.B. Improved pulse detection from head motions using DCT. In Proceedings of the 2014 International Conference on Computer Vision Theory and Applications (VISAPP), 2014; IEEE: Lisbon, Portugal, 2014; Volume 3.
62. Song, J.; He, T.; Gao, L.; Xu, X.; Hanjalic, A.; Shen, H.T. Unified Binary Generative Adversarial Network for Image Retrieval and Compression. *Int. J. Comput. Vis.* **2020**, *128*, 2243–2264. <https://doi.org/10.1007/s11263-020-01305-2>.
63. Monkaresi, H.; Calvo, R.A.; Yan, H. A machine learning approach to improve contactless heart rate monitoring using a webcam. *IEEE J. Biomed. Health Inform.* **2014**, *18*, 1153–1160. <https://doi.org/10.1109/JBHI.2013.2291900>.
64. Aha, D.W.; Kibler, D.; Albert, M.K. Instance-Based Learning Algorithms. *Mach. Learn.* **1991**, *6*, 37–66. <https://doi.org/10.1023/A:1022689900470>.
65. Spetlik, R.; Franc, V.; Cech, J.; Matas, J. Visual heart rate estimation with convolutional neural network. In Proceedings of the British Machine Vision Conference 2018, BMVC 2018, August 2018.
66. Estep, J.R.; Blackford, E.B.; Meier, C.M. Recovering pulse rate during motion artifact with a multi-imager array for non-contact imaging photoplethysmography. In Proceedings of the IEEE International Conference on Systems, Man and Cybernetics, January 2014; pp. 1462–1469. <https://doi.org/10.1109/smc.2014.6974121>.
67. Bian, M.; Peng, B.; Wang, W.; Dong, J. An Accurate LSTM Based Video Heart Rate Estimation Method. In Proceedings of the Chinese Conference on Pattern Recognition and Computer Vision (PRCV), Xi'an, China, 8–11 November, 2019.
68. Wang, Z.; Kao, Y.; Hsu, C. Vision-Based Heart Rate Estimation via a Two-Stream Cnn Zhi-Kuan Wang Ying Kao Chiou-Ting Hsu Department of Computer Science, National Tsing Hua University. In Proceedings of the 2019 IEEE International Conference on Image Processing (ICIP), 2019; pp. 3327–3331.
69. Paracchini, M.; Marcon, M.; Villa, F.; Zappa, F.; Tubaro, S. Biometric Signals Estimation Using Single Photon Camera and Deep Learning. *Sensors* **2020**, *20*, 6102. <https://doi.org/10.3390/s20216102>.
70. Sabokrou, M.; Pourreza, M.; Li, X.; Fathy, M.; Zhao, G. Deep-HR: Fast heart rate estimation from face video under realistic conditions. *Expert Syst. Appl.* **2021**, *186*, 115596. <https://doi.org/10.1016/j.eswa.2021.115596>.
71. Tsou, Y.Y.; Lee, Y.A.; Hsu, C.T.; Chang, S.H. Siamese-rPPG network: Remote photoplethysmography signal estimation from face videos. In Proceedings of the ACM Symposium on Applied Computing, 2020; pp. 2066–2073. <https://doi.org/10.1145/3341105.3373905>.
72. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; Springer: Berlin/Heidelberg, Germany, 2018; Volume 11206, pp. 356–373. https://doi.org/10.1007/978-3-030-01216-8_22.
73. Niu, X.; Han, H.; Shan, S.; Chen, X. SynRhythm: Learning a Deep Heart Rate Estimator from General to Specific. In Proceedings of the International Conference on Pattern Recognition, August 2018; pp. 3580–3585. <https://doi.org/10.1109/ICPR.2018.8546321>.
74. Niu, X.; Shan, S.; Han, H.; Chen, X. RhythmNet: End-to-End Heart Rate Estimation from Face via Spatial-Temporal Representation. *IEEE Trans. Image Process.* **2020**, *29*, 2409–2423. <https://doi.org/10.1109/TIP.2019.2947204>.
75. Niu, X.; Han, H.; Shan, S.; Chen, X. VIPL-HR: A Multi-modal Database for Pulse Estimation from Less-Constrained Face Video. In *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; Springer: Berlin/Heidelberg, Germany, 2019; Volume 11365, pp. 562–576. https://doi.org/10.1007/978-3-030-20873-8_36.
76. Yu, Z.; Li, X.; Zhao, G. Remote photoplethysmography signal measurement from facial videos using spatio-temporal networks. In Proceedings of the 30th British Machine Vision Conference 2019, BMVC 2019, 2020.
77. Liu, X.; Jiang, Z.; Fromm, J.; Xu, X.; Patel, S.; McDuff, D. (2021). MetaPhys: Few-shot adaptation for non-contact physiological measurement. In Proceedings of the ACM CHIL 2021—2021 ACM Conference on Health, Inference, and Learning, 2021; Association for Computing Machinery: New York, NY, USA; Volume 1. <https://doi.org/10.1145/3450439.3451870>.
78. Yu, Z.; Peng, W.; Li, X.; Hong, X.; Zhao, G. Remote heart rate measurement from highly compressed facial videos: An end-to-end deep learning solution with video enhancement. In Proceedings of the IEEE International Conference on Computer Vision, October 2019; pp. 151–160. <https://doi.org/10.1109/ICCV.2019.00024>.
79. Yu, Z.; Li, X.; Niu, X.; Shi, J.; Zhao, G. AutoHR: A Strong End-to-End Baseline for Remote Heart Rate Measurement with Neural Searching. *IEEE Signal Process. Lett.* **2020**, *27*, 1245–1249. <https://doi.org/10.1109/LSP.2020.3007086>.
80. Hernandez-Ortega, J.; Fierrez, J.; Morales, A.; Diaz, D. A Comparative Evaluation of Heart Rate Estimation Methods using Face Videos. In Proceedings of the 2020 IEEE 44th Annual Computers, Software, and Applications Conference, COMPSAC 2020, 2020; pp. 1438–1443. <https://doi.org/10.1109/COMPSAC48688.2020.00-53>.

81. Bousefsaf, F.; Pruski, A.; Maaoui, C. 3D convolutional neural networks for remote pulse rate measurement and mapping from facial video. *Appl. Sci.* **2019**, *9*, 4364. <https://doi.org/10.3390/app9204364>
82. Hu, M.; Qian, F.; Guo, D.; Wang, X.; He, L.; Ren, F. ETA-rPPGNet: Effective Time-Domain Attention Network for Remote Heart Rate Measurement. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 2506212. <https://doi.org/10.1109/TIM.2021.3058983>.
83. McDuff, D. Deep super-resolution for recovering physiological information from videos. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, June 2018; pp. 1448–1455. <https://doi.org/10.1109/CVPRW.2018.00185>.
84. Nowara, E.M.; McDuff, D.; Veeraraghavan, A. Systematic analysis of video-based pulse measurement from compressed videos. *Biomed. Opt. Express* **2021**, *12*, 494. <https://doi.org/10.1364/boe.408471>
85. Gupta, A.; Ravelo-Garcia, A.G.; Morgado-Dias, F. A Motion and Illumination Resistant Non-contact Method using Undercomplete Independent Component Analysis and Levenberg-Marquardt Algorithm. In *IEEE Journal of Biomedical and Health Informatics*; IEEE: Manhattan, NY, USA, 2022. <https://doi.org/10.1109/JBHI.2022.3144677>.
86. Paracchini, M.; Marcon, M.; Villa, F.; Cusini, I.; Tubaro, S. Fast Skin Segmentation on Low-Resolution Grayscale Images for Remote PhotoPlethysmography. *IEEE MultiMedia* **2022**, *29*, 28–35. <https://doi.org/10.1109/MMUL.2022.3152087>.
87. Koelstra, S.; Muhl, C.; Soleymani, M.; Lee, J.; Yazdani, A.; Ebrahimi, T.; Pun, T.; Nijholt, A.; Patras, I. DEAP: A Database for Emotion Analysis; Using Physiological Signals. *IEEE Trans. Affect. Comput.* **2012**, *3*, 18–31.
88. Soleymani, M.; Lichtenauer, J.; Pun, T.; Pantic, M. A multimodal database for affect recognition and implicit tagging. *IEEE Trans. Affect. Comput.* **2012**, *3*, 42–55. <https://doi.org/10.1109/T-AFFC.2011.25>.
89. Gudi, A.; Bittner, M.; van Gemert, J. Real-time webcam heart rate and variability estimation with clean ground truth for evaluation. *Appl. Sci.* **2020**, *10*, 8630. <https://doi.org/10.3390/app10238630>.
90. Hanfland, S.; Paul, M. (2016). Video Format Dependency of PPGI Signals. Poster, 1–6. Available online: http://poseidon2.feld.cvut.cz/conf/poster/poster2016/proceedings/Section_BI/BI_007_Hanfland.pdf (accessed on 25 March 2022).
91. Zhao, C.; Lin, C.L.; Chen, W.; Li, Z. A novel framework for remote photoplethysmography pulse extraction on compressed videos. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, June 2018; pp. 1380–1389. <https://doi.org/10.1109/CVPRW.2018.00177>
92. Nowara, E.M.; McDuff, D.; Veeraraghavan, A. A meta-analysis of the impact of skin tone and gender on non-contact photoplethysmography measurements. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2020; pp. 284–285.
93. Dasari, A.; Prakash, S.K.A.; Jeni, L.A.; Tucker, C.S. Evaluation of biases in remote photoplethysmography methods. *NPJ Digit. Med.* **2021**, *4*, 91.
94. Chari, P.; Kabra, K.; Karinca, D.; Lahiri, S.; Srivastava, D.; Kulkarni, K.; Chen, T.; Cannesson, M.; Jalilian, L.; Kadambi, A. Diverse R-PPG: Camera-based heart rate estimation for diverse subject skin-tones and scenes. *arXiv* **2020**, arXiv:2010.12769.
95. Zhang, P.; Li, B.; Peng, J.; Jiang, W. Multi-hierarchical Convolutional Network for Efficient Remote Photoplethysmograph Signal and Heart Rate Estimation from Face Video Clips. *arXiv* **2021**, arXiv:2104.02260.
96. Sun, Z.; Li, X. Privacy-Phys: Facial Video-Based Physiological Modification for Privacy Protection. *IEEE Signal Process. Lett.* **2022**, *29*, 1507–1511. <https://doi.org/10.1109/LSP.2022.3185964>.
97. Chen, M.; Liao, X.; Wu, M. PulseEdit: Editing Physiological Signals in Facial Videos for Privacy Protection. *IEEE Trans. Inf. Forensics Secur.* **2022**, *17*, 457–471. <https://doi.org/10.1109/TIFS.2022.3142993>.