Deep Learning-based Contactless PPG Methods

Contributor: Aoxin Ni

Physiological measurements are widely used to determine a person's health condition. Photoplethysmography (PPG) is a physiological measurement method that is used to detect volumetric changes in blood in vessels beneath the skin. Medical devices based on PPG have been introduced to measure different physiological measurements including heart rate (HR), respiratory rate, heart rate variability (HRV), oxyhemoglobin saturation, and blood pressure. Due to its low cost and non-invasive nature, PPG is utilized in many devices such as finger pulse oximeters, sports bands, and wearable sensors. PPG-based physiological measurements can be categorized into two types: contact-based and contactless.

1. Background

A common practice in the medical field to measure the heart rate is ECG or electrocardiography [1][2], where voltage changes in the heart electrical activity are detected using electrodes placed on the skin. In general, ECG provides a more reliable heart rate measurement compared to PPG [3][4]. Hence, ECG is often used as the reference for evaluation of PPG methods [1][2][3][4]. Typically, 10 electrodes of the ECG machine are attached to different parts of the body including the wrist and ankle. Different from ECG, PPG-based medical devices possess differing sensor shapes placed on different parts of the body such as rings, earpieces, and bands [1][5][6][7][8][9][10], and they all use a light source and a photodetector to detect the PPG signal with signal processing, see Figure 1. The signal processing is for the purpose of processing the reflected optical signal from the skin [11].

Figure 1. PPG signal processing framework.

Early research in this field concentrated on obtaining the PPG signal and ways to perform pulse wave analysis [12]. A comparison between ECG and PPG is discussed in [13][14]. There are survey papers covering different PPG applications that involve the use of wearable devices [15][16], atrial fibrillation detection [17], and blood pressure monitoring [18]. Papers have also been published which used deep learning for contact-based PPG, e.g., [19][20][21][22]. The previous survey papers on contact-based PPG methods are listed in Table 1.

Table 1. Previous survey papers on contact-based PPG methods.

| Emphasis	Ref	Year	Task
Contact	[12]	2007	Basic principle of PPG operation, pulse wave analysis, clinical applications
Contact ECG and PPG	[13]	2018	Breathing rate (BR) estimation from ECG and PPG, BR algorithms and its assessment
Contact	[17]	2020	Approaches for PPG-based atrial fibrillation detection
Contact Wearable device	[15]	2019	PPG acquisition, HR estimation algorithms, developments on wrist PPG applications, biometric identification
Contact ECG and PPG	[14]	2012	Accuracy of pulse rate variability (PRV) as an estimate of HRV

Emphasis	Ref	Year	Task
Contact Wearable device	[16]	2018	Current developments and challenges of wearable PPG-based monitoring technologies
Contact Blood pressure	[18]	2015	Approaches involving PPG for continuous and non-invasive monitoring of blood pressure

Although contact-based PPG methods are non-invasive, they can be restrictive due to the requirement of their contact with the skin. Contact-based methods can be irritating or distracting in some situations, for example, for newborn infants [23][24][25][26]. When a less restrictive approach is desired, contactless PPG methods are considered. The use of contactless PPG methods or remote PPG (rPPG) methods has been growing in recent years [27][28][29][30][31].

Contactless PPG methods usually utilize a video camera to capture images which are then processed by image processing algorithms [27][28][29][30][31]. The physics of rPPG is similar to contact-based PPG. In rPPG methods, the light-emitting diode in contact-based PPG methods is replaced with ambient illuminance, and the photodetector is replaced with a video camera, see Figure 2. The light reaching the camera sensor can be separated into static (DC) and dynamic (AC) components. The DC component corresponds to static elements including tissue, bone, and static blood, while the AC component corresponds to the variations in light absorption due to arterial blood volume changes. Figure 3 provides an illustration of the image processing framework in rPPG methods. The common image processing steps involved in the framework are illustrated in this figure. In the signal extraction part of the framework, a region of interest (ROI), normally on the face, is extracted.
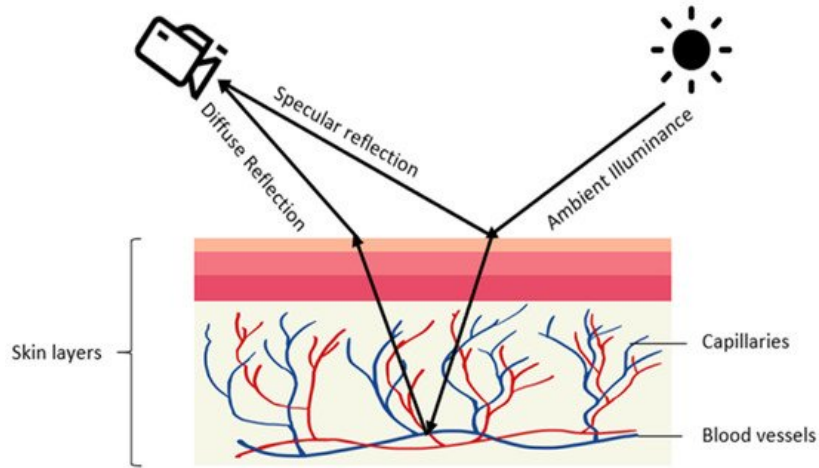
Figure 2. Illustration of rPPG generation: diffused and specular reflections of ambient illuminance are captured by a camera with the diffused reflection indicating volumetric changes in blood vessels.
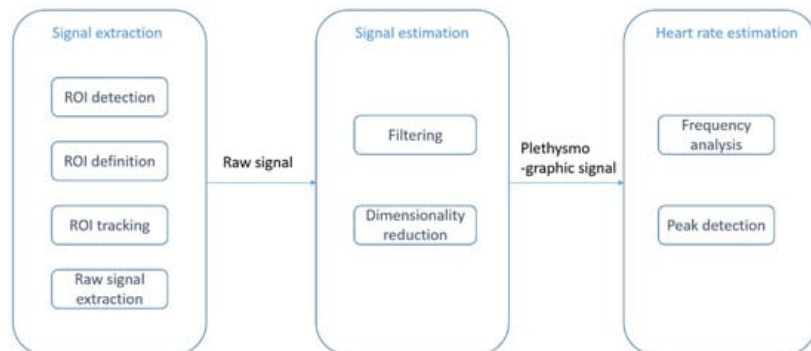
Figure 3. rPPG or contactless PPG image processing framework: signal extraction step (ROI detection and tracking), signal estimation step (filtering and dimensionality reduction), and heart rate estimation step (frequency analysis and peak detection).

In earlier studies, video images from motionless faces were considered [32][33][34]. Several papers relate to exercising situations [35][36][37][38][39]. ROI detection and ROI tracking constitute two major image processing parts of the framework. The Viola and Jones (VJ) algorithm [40] is often used to detect face areas [41][42][43][44]. As an example of prior work on skin detection, a neural network classifier was used to detect skin-like pixels in [45]. In the signal estimation part, a bandpass filter is applied to eliminate undesired frequency components. A common choice for the frequency band is [0.7 Hz, 4 Hz],

which corresponds to an HR between 42 and 240 beats per minute (bpm) [45][46][47][48]. To separate a signal into uncorrelated components and to reduce dimensionality, independent component analysis (ICA) was utilized in [49][50][51][52] and principal component analysis (PCA) was utilized in [33][34][35][53][54]. In the heart rate estimation module, the dimensionality-reduced data will be mapped to certain levels using frequency analysis or peak detection methods. The survey papers on rPPG methods that have already appeared in the literature are listed in Table 2. These survey papers provide comparisons with contact-based PPG methods.

There are challenges in rPPG which include subject motion and ambient lighting variations [55][56][57]. Due to the success of deep learning in many computer vision and speech processing applications [58][59][60], deep learning methods have been considered for rPPG to deal with its challenges, for example, [39][44]. In deep learning methods, feature extraction and classification are carried out together within one network structure. The required datasets for deep learning models are collected using RGB cameras. As noted earlier, the focus of this review is on deep learning-based contactless heart rate measurement methods.

Table 2. Survey papers on conventional contactless methods previously reported in the literature.

Emphasis	Ref	Year	Task
Contactless	[61]	2018	Provides typical components of rPPG and notes the main challenges; groups published studies by their choice of algorithm.
Contactless	[62]	2012	Covers three main stages of monitoring physiological measurements based on photoplethysmographic imaging: image acquisition, data collection, and parameter extraction.
Contactless and contact	[63]	2016	States review of contact-based PPG and its limitations; introduces research activities on wearable and non-contact PPG.
Contactless and contact	[64]	2009	Reviews photoplethysmographic measurement techniques from contact sensing placement to non-contact sensing placement, and from point measurement to imaging measurement.
Contactless newborn infants	[23]	2013	Investigates the feasibility of camera-based PPG for contactless HR monitoring in newborn infants with ambient light.
Contactless newborn infants	[25]	2016	Comparative analysis to benchmark state-of-the-art video and image-guided noninvasive pulse rate (PR) detection.
Contactless and contact	[65]	2017	Heart rate measurement using facial videos based on photoplethysmography and ballistocardiography.
Contactless and contact	[66]	2014	Covers methods of non-contact HR measurement with capacitively coupled ECG, Doppler radar, optical vibrocardiography, thermal imaging, RGB camera, and HR from speech.
Contactless RR and contact	[67]	2011	Discusses respiration monitoring approaches (both contact and non-contact).
Contactless newborn infants	[26]	2019	Addresses HR measurement in babies.
Contactless	[68]	2019	Examines challenges associated with illumination variations and motion artifacts.
Contactless	[69]	2017	Covers HR measurement techniques including camera-based photoplethysmography, reflectance pulse oximetry, laser Doppler technology, capacitive sensors, piezoelectric sensors, electromyography, and a digital stethoscope.
Contactless Main challenges	[70]	2015	Covers issues in motion and ambient lighting tolerance, image optimization (including multi-spectral imaging), and region of interest optimization.

2. Contactless PPG Methods Based on Deep Learning

Previous works on deep learning-based contactless HR methods can be divided into two groups: combinations of conventional and deep learning methods, and end-to-end deep learning methods. In what follows, a review of these papers is provided.

2.1. Combination of Conventional and Deep Learning Methods

Li et al. 2021 [71] presented multi-modal machine learning techniques related to heart diseases. From Figure 3, it can be seen that one or more components of the contactless HR framework can be achieved by using deep learning. These components include ROI detection and tracking, signal estimation, and HR estimation.

2.1.1. Deep Learning Methods for Signal Estimation

Qiu et al. 2018 [72] developed a method called EVM-CNN. The pipeline of this method consists of three modules: face detection and tracking, feature extraction, and HR estimation. In the face detection and tracking module, 68 facial landmarks inside a bounding box are detected by using a regression local binary features-based approach [73]. Then, an ROI defined by eight points around the central part of a human face is automatically extracted and inputted into the next module. In the feature extraction module, spatial decomposition and temporal filtering are applied to obtain so-called feature images. The sequence of ROIs is down-sampled into several bands. The lowest bands are reshaped and concatenated into a new image. Three channels of this new image are transferred into the frequency domain; then, fast Fourier transform (FFT) is applied to remove the unwanted frequency bands. Finally, the bands are transferred back to the time domain by performing inverse FFT and merging into a feature image. In the HR estimation module, a convolutional neural network (CNN) is used to estimate HR from the feature image. The CNN used in this method has a simple structure with several convolution layers which uses depth-wise convolution and point-wise convolution to reduce the computational burden and model size.

As shown in Figure 4, in this method, the first two modules which are face detection/tracking and feature extraction are conventional rPPG approaches, whereas the HR estimation module uses deep learning to improve performance for HR estimation.
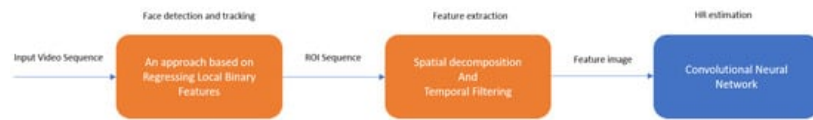

Figure 4. EVM-CNN modules.

2.1.2. Deep Learning Methods for Signal Extraction

Luguev et al. 2020 [74] established a framework which uses deep spatial-temporal networks for contactless HRV measurements from raw facial videos. In this method, a 3D convolutional neural network is used for pulse signal extraction. As for the computation of HRV features, conventional signal processing methods including frequency domain analysis and peak detection are used. More specifically, raw video sequences are inputted into the 3D-CNN without any skin segmentation. Several convolution operations with rectified linear units (ReLU) are used as activation functions together with pooling operations to produce spatiotemporal features. In the end, a pulse signal is generated by a channel-wise convolution operation. The mean absolute error is used as the loss function of the model.

Paracchini et al. 2020 [75] implemented rPPG based on a single-photon avalanche diode (SPAD) camera. This method combines deep learning and conventional signal processing to extract and examine the pulse signal. The main advantage of using a SPAD camera is its superior performance in dark environments compared with CCD or CMOS cameras. Its framework is shown in Figure 5. The signal extraction part has two components which are facial skin detection and signal creation. A U-shape network is then used to perform skin detection including all visible facial skin surface areas rather than a specific skin area. The output of the network is a binary skin mask. Then, a raw pulse signal is obtained by averaging the intensity values of all the pixels inside the binary mask. As for signal estimation, this is achieved by filtering, FFT, and peak detection. The experimental results include HR, respiration rate, and tachogram measurements.

Figure 5. rPPG using SPAD camera.

In another work from Zhan et al. 2020 [76], the focus was placed on understanding the CNN-based PPG signal extraction. Four questions were addressed: (1) Does the CNN learn PPG, BCG, or a combination of both? (2) Can a finger oximeter be directly used as a reference for CNN training? (3) Does the CNN learn the spatial context information of the measured skin? (4) Is the CNN robust to motion, and how is this motion robustness achieved? To answer these four questions, a CNN-PPG framework and four experiments were designed. The results of these experiments indicate the availability of

multiple convolutional kernels is necessary for a CNN to arrive at a flexible channel combination through the spatial operation but may not provide the same motion robustness as a multi-site measurement. Another conclusion reached is that the PPG-related prior knowledge may still be helpful for the CNN-based PPG extraction.

2.2. End-to-End Deep Learning Methods

In this section, end-to-end deep learning systems are stated which take video as the input and use different network architectures to generate a physiological signal as the output.

2.2.1. VGG-Style CNN

Chen and Mcduff 2018 [77] developed an end-to-end method for video-based heart and breathing rates using a deep convolutional network named DeepPhys. To address the issue caused by subject motion, the proposed method uses a motion representation algorithm based on a skin reflection model. As a result, motions are captured more effectively. To guide the motion estimation, an attention mechanism using appearance information was designed. It was shown that the motion representation model and the attention mechanism used enable robust measurements under heterogeneous lighting and motions.

The model is based on a VGG-style CNN for estimating the physiological signal derived under motion [78]. VGG is an object recognition model that supports up to 19 layers. Built as a deep CNN, VGG is shown to outperform baselines in many image processing tasks. Figure 6 illustrates the architecture of this end-to-end convolutional attention network. A current video frame at time t and a normalized difference between frames at t and t + 1 constitute the inputs to the appearance and motion models, respectively. The network learns spatial masks, which are shared between the models, and extracts features for recovering the blood volume pulse (BVP) and respiration signals.
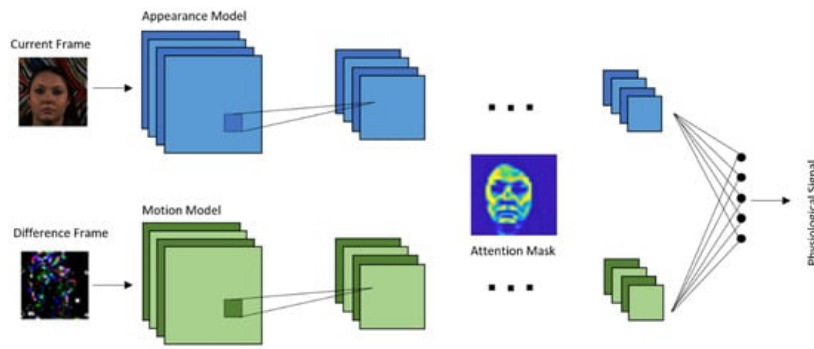

Figure 6. DeepPhys architecture.

Deep PPG proposed by Reiss et al. 2019 [79] addresses three shortcomings of the existing datasets. First is the dataset size. While the number of subjects can be considered as sufficient (8–24 participants in each dataset), the length of each session's recording can be rather short. Second is the small numbers of activities. The publicly available datasets include data from only two–three different activities. Additionally, third is data recording in laboratory settings rather than in real-world environments.

A new dataset, called PPG-DaLiA [80], was thus introduced in this paper: a PPG dataset for motion compensation and heart rate estimation in daily living activities. Figure 7 illustrates the architecture of the VGG-like CNN used, where the time–frequency spectra of PPG signals are used as the input to estimate the heart rate.
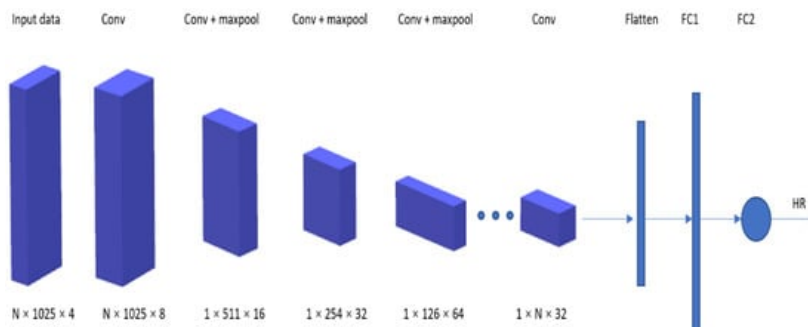

Figure 7. Deep PPG architecture.

2.2.2. CNN-LSTM Network

Long short-term memory (LSTM) is a recurrent neural network (RNN) architecture which allows only process handling a single data point (such as images), but also an entire sequence of data points (such as speech or video). It has been previously used for various tasks such as connected handwriting recognition, speech recognition, and anomaly detection in network traffic [81][82][83].

rPPG signals are usually collected using a video camera with a limitation of being sensitive to multiple contributing factors, which include variation in skin tone, lighting condition, and facial structure. Meta-rPPG [84] is an end-to-end supervised learning approach which performs well when training data are abundant with a distribution that does not deviate too much from the testing data distribution. To cope with the unforeseeable changes during testing, a transductive meta-learner that takes unlabeled samples during testing for a self-supervised weight adjustment is used to provide fast adaptation to the changes. The network proposed in this paper is split into two parts: a feature extractor and an rPPG estimator modeled by a CNN and an LSTM network, respectively.

2.2.3. 3D-CNN Network

A 3D convolutional neural network is a type of network with kernel sliding in three dimensions. 3D-CNN is shown to have better performance in spatiotemporal information learning than 2DCNN [85].

Špetlík et al. 2018 [41] proposed a two-step convolutional neural network to estimate the heart rate from a sequence of facial images, see Figure 8. The proposed architecture has two components: an extractor and an HR estimator. The extractor component is run over a temporal image sequence of faces. The signal is then fed to the HR estimator to predict the heart rate.
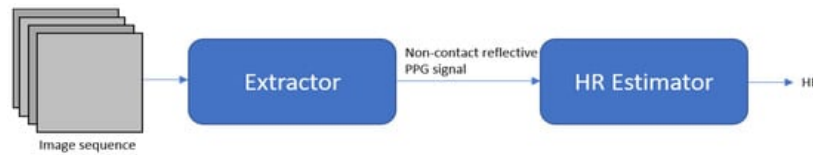
Figure 8. HR-CNN modules.

In the work from Yu et al. 2019 [86], a two-stage end-to-end method was proposed. This work deals with video compression loss and recovers the rPPG signal from highly compressed videos. It consists of two parts: (1) a spatiotemporal video enhancement network (STVEN) for video enhancement, and (2) an rPPG network (rPPGNet) for rPPG signal recovery. rPPGNet can work on its own for obtaining rPPG measurements. The STVEN network can be added and jointly trained to further boost the performance, particularly on highly compressed videos.

Another method from Yu et al. 2019 [87] provides the use of deep spatiotemporal networks for reconstructing precise rPPG signals from raw facial videos. With the constraint of trend consistency in ground truth pulse curves, this method is able to recover rPPG signals with accurate pulse peaks. The heartbeat peaks of the measured rPPG signal are located at the corresponding R peaks of the ground truth ECG signal.

To address the issue of a lack of training data, a heart track convolutional neural network was developed by Rerepelkina et al. 2020 [88] for remote video-based heart rate tracking. This learning-based method is trained on synthetic data to accurately estimate the heart rate in different conditions. Synthetic data do not include video and include only PPG curves. To select the most suitable parts of the face for pulse tracking at each particular moment, an attention mechanism is used.

Similar to the previous methods, the method proposed by Bousefsaf et al. 2019 [89] also uses synthetic data. Figure 9 illustrates the process of how synthetic data are generated. A 3D-CNN classifier structure was developed for both extraction and classification of unprocessed video streams. The CNN acts as a feature extractor. Its final activations are fed into two dense layers (multilayer perceptron) that are used to classify the pulse rate. The network ensures concurrent mapping by producing a prediction for each local group of pixels.

Figure 9. Process of generating synthetic data.

Liu et al. 2020 [90] developed a lightweight rPPG estimation network, named DeeprPPG, based on spatiotemporal convolutions for utilization involving different types of input skin. To further boost the robustness, a spatiotemporal rPPG aggregation strategy was designed to adaptively aggregate rPPG signals from multiple skin regions into a final one.

Extensive experimental studies were conducted to show its robustness when facing unseen skin regions in unseen scenarios. Table 3 lists the contactless HR methods that use deep learning.

Table 3. Deep learning-based contactless PPG methods.

Focus	Ref	Year	Feature	Dataset
End-to-end system Robust to illumination changes and subject's motion	[41]	2018	A two-step convolutional neural network composed of an extractor and HR estimator	COHFACE PURE MAHNOB-HCI
Signal estimation enhancement	[72]	2019	Eulerian video magnification (EVM) to extract face color changes and using CNN to estimate heart rate	MMSE-HR
3D-CNN for signal extraction	[74]	2020	Using deep spatiotemporal networks for contactless HRV measurements from raw facial videos; employing data augmentation	MAHNOB-HCI
Single-photon camera	[75]	2020	Neural network for skin detection	N/A
Understanding of CNN-based PPG methods	[76]	2020	Analysis of CNN-based remote PPG to understand limitations and sensitivities	HNU PURE
End-to-end system Attention mechanism	[77]	2018	Robust measurement under heterogeneous lighting and motions	MAHNOB-HCI
End-to-end system Real-life conditions dataset	[79]	2019	Major shortcoming of existing datasets: dataset size, small number of activities, data recording in laboratory setting	PPG-DaLiA
Synthetic training data Attention mechanism	[88]	2020	CNN training with synthetic data to accurately estimate HR in different conditions	UBFC-RPPG MoLi-ppg-1 MoLi-ppg-2
Synthetic training data	[89]	2019	Automatic 3D-CNN training process with synthetic data with no image processing	UBFC-RPPG
End-to-end supervised learning approach Meta-learning	[84]	2017	Meta-rPPG for abundant training data with a distribution not deviating too much from distribution of testing data	MAHNOB-HCI UBFC-RPPG
Counter video compression loss	[86]	2019	STEVEN for video quality enhancement rPPGNet for signal recovery	MAHNOB-HCI
Spatiotemporal network	[87]	2019	Measuring rPPG signal from raw facial video; taking temporal context into account	MAHNOB-HCI
Spatiotemporal network	[90]	2020	Spatiotemporal convolution network, different types of input skin	MAHNOB-HCI PURE

References

1. Kinnunen, H.; Rantanen, A.; Kenttä, T.; Koskimäki, H. Feasible assessment of recovery and cardiovascular health: Accuracy of nocturnal HR and HRV assessed via ring PPG in comparison to medical grade ECG. Physiol. Meas. 2020, 41, 04NT01.

2. Orphanidou, C. Derivation of respiration rate from ambulatory ECG and PPG using ensemble empirical mode decomposition: Comparison and fusion. Comput. Biol. Med. 2017, 81, 45–54.

3. Clifton, D.A.; Meredith, D.; Villarroel, M.; Tarassenko, L. Home monitoring: Breathing rate from PPG and ECG. Inst. Biomed. Eng. 2012. Available online: (accessed on 26 May 2021).

4. Madhav, K.V.; Raghuram, M.; Krishna, E.H.; Komalla, N.R.; Reddy, K.A. Extraction of respiratory activity from ECG and PPG signals using vector autoregressive model. In Proceedings of the 2012 IEEE International Symposium on Medical Measurements and Applications Proceedings, Budapest, Hungary, 18–19 May 2012; pp. 1–4.

5. Gu, W.B.; Poon, C.C.; Leung, H.K.; Sy, M.Y.; Wong, M.Y.; Zhang, Y.T. A novel method for the contactless and continuous measurement of arterial blood pressure on a sleeping bed. In Proceedings of the 2009 Annual International

Conference of the IEEE Engineering in Medicine and Biology Society, Minneapolis, MN, USA, 3–6 September 2009; pp. 6084–6086.

6. Wang, L.; Lo BP, L.; Yang, G.Z. Multichannel reflective PPG earpiece sensor with passive motion cancellation. IEEE Transact. Biomed. Circ. Syst. 2007, 1, 235–241.

7. Kabiri Ameri, S.; Ho, R.; Jang, H.; Tao, L.; Wang, Y.; Wang, L.; Schnyer, D.M.; Akinwande, D.; Lu, N. Graphene electronic tattoo sensors. ACS Nano 2017, 11, 7634–7641.

8. Nardelli, M.; Vanello, N.; Galperti, G.; Greco, A.; Scilingo, E.P. Assessing the Quality of Heart Rate Variability Estimated from Wrist and Finger PPG: A Novel Approach Based on Cross-Mapping Method. Sensors 2020, 20, 3156.

9. Phan, D.; Siong, L.Y.; Pathirana, P.N. Smartwatch: Performance evaluation for long-term heart rate monitoring. In Proceedings of the 2015 International Symposium on Bioelectronics and Bioinformatics (ISBB), Beijing, China, 14–17 October 2015; pp. 144–147.

10. Wong, M.Y.; Leung, H.K.; Pickwell-MacPherson, E.; Gu, W.B.; Zhang, Y.T. Contactless recording of photoplethysmogram on a sleeping bed. In Proceedings of the 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Beijing, China, 14–17 October 2009; pp. 907–910.

11. Challoner, A.V.; Ramsay, C.A. A photoelectric plethysmograph for the measurement of cutaneous blood flow. Phys. Med. Biol. 1974, 19, 317–328.

12. Allen, J. Photoplethysmography and its application in clinical physiological measurement. Physiol. Meas. 2007, 28, R1.

13. Charlton, P.H.; Birrenkott, D.A.; Bonnici, T.; Pimentel, M.A.; Johnson, A.E.; Alastruey, J.; Tarassenko, L.; Watkinson, P.J.; Beale, R.; Clifton, D.A. Breathing rate estimation from the electrocardiogram and photoplethys-mogram: A review. IEEE Rev. Biomed. Eng. 2017, 11, 2–20.

14. Schäfer, A.; Vagedes, J. How accurate is pulse rate variability as an estimate of heart rate variability?: A review on studies comparing photoplethysmographic technology with an electrocardiogram. Int. J. Cardiol. 2013, 166, 15–29.

15. 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.

16. Castaneda, D.; Esparza, A.; Ghamari, M.; Soltanpur, C. A review on wearable photoplethysmography sensors and their potential future applications in health care. Int. J. Biosensors Bioelectron. 2018, 4, 195.

17. Pereira, T.; Tran, N.; Gadhoumi, K.; Pelter, M.M.; Do, D.H.; Lee, R.J.; Colorado, R.; Meisel, K.; Hu, X. Photoplethysmography based atrial fibrillation detection: A review. npj Digit. Med. 2020, 3, 1–12.

18. Nye, R.; Zhang, Z.; Fang, Q. Continuous non-invasive blood pressure monitoring using photoplethysmography: A review. In Proceedings of the 2015 International Symposium on Bioelectronics and Bioinformatics (ISBB), Beijing, China, 14–17 October 2015; pp. 176–179.

19. Johansson, A. Neural network for photoplethysmographic respiratory rate monitoring. Med. Biol. Eng. Comput. 2003, 41, 242–248.

20. Panwar, M.; Gautam, A.; Biswas, D.; Acharyya, A. PP-Net: A Deep Learning Framework for PPG-Based Blood Pressure and Heart Rate Estimation. IEEE Sens. J. 2020, 20, 10000–10011.

21. Biswas, D.; Everson, L.; Liu, M.; Panwar, M.; Verhoef, B.-E.; Patki, S.; Kim, C.H.; Acharyya, A.; Van Hoof, C.; Konijnenburg, M.; et al. CorNET: Deep Learning Framework for PPG-Based Heart Rate Estimation and Biometric Identification in Ambulant Environment. IEEE Trans. Biomed. Circ. Syst. 2019, 13, 282–291.

22. Chang, X.; Li, G.; Xing, G.; Zhu, K.; Tu, L. DeepHeart. ACM Trans. Sens. Netw. 2021, 17, 1–18.

23. Aarts, L.A.; Jeanne, V.; Cleary, J.P.; Lieber, C.; Nelson, J.S.; Oetomo, S.B.; Verkruysse, W. Non-contact heart rate monitoring utilizing camera photoplethysmography in the neonatal intensive care unit—A pilot study. Early Hum. Dev. 2013, 89, 943–948.

24. Villarroel, M.; Chaichulee, S.; Jorge, J.; Davis, S.; Green, G.; Arteta, C.; Zisserman, A.; McCormick, K.; Watkinson, P.; Tarassenko, L. Non-contact physiological monitoring of preterm infants in the Neonatal Intensive Care Unit. NPJ Digit. Med. 2019, 2, 128.

25. Sikdar, A.; Behera, S.K.; Dogra, D.P. Computer-vision-guided human pulse rate estimation: A review. IEEE Rev. Bio Med. Eng. 2016, 9, 91–105.

26. Anton, O.; Fernandez, R.; Rendon-Morales, E.; Aviles-Espinosa, R.; Jordan, H.; Rabe, H. Heart Rate Monitoring in Newborn Babies: A Systematic Review. Neonatology 2019, 116, 199–210.

27. Fernández, A.; Carús, J.L.; Usamentiaga, R.; Alvarez, E.; Casado, R. Unobtrusive health monitoring system using video-based physiological in-formation and activity measurements. IEEE 2013, 89, 943–948.

28. Haque, M.A.; Irani, R.; Nasrollahi, K.; Moeslund, T.B. Heartbeat rate measurement from facial video. IEEE Intell. Syst. 2016, 31, 40–48.

29. Kumar, M.; Veeraraghavan, A.; Sabharwal, A. DistancePPG: Robust non-contact vital signs monitoring using a camera. Biomed. Opt. Express 2015, 6, 1565–1588.

30. Liu, S.Q.; Lan, X.; Yuen, P.C. Remote photoplethysmography correspondence feature for 3D mask face presentation attack detection. In Proceedings of the European Conference on Computer Vision, ECCV Papers, Munich, Germany, 8–14 September 2018; pp. 558–573.

31. Gudi, A.; Bittner, M.; Lochmans, R.; van Gemert, J. Efficient real-time camera based estimation of heart rate and its variability. In Proceedings of the IEEE International Conference on Computer Vision Workshops, Seoul, Korea, 27–28 October 2019.

32. Verkruysse, W.; Svaasand, L.O.; Nelson, J.S. Remote plethysmographic imaging using ambient light. Opt. Express 2008, 16, 21434–21445.

33. Balakrishnan, G.; Durand, F.; Guttag, J. Detecting Pulse from Head Motions in Video. In Proceedings of the 2013 IEEE Conference on Computer Vision and Pattern Recognition, Portland, OR, USA, 25–27 June 2013; pp. 3430–3437.

34. 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), Szczecin, Poland, 18–21 September 2011; pp. 405–410.

35. De Haan, G.; Jeanne, V. Robust pulse rate from chrominance-based rPPG. IEEE Transact. Biomed. Eng. 2013, 60, 2878–2886.

36. Tasli, H.E.; Gudi, A.; den Uyl, M. Remote PPG based vital sign measurement using adaptive facial regions. In Proceedings of the 2014 IEEE International Conference on Image Processing (ICIP), Paris, France, 27–30 October 2014; pp. 1410–1414.

37. Yu, Y.P.; Kwan, B.H.; Lim, C.L.; Wong, S.L. Video-based heart rate measurement using short-time Fourier transform. In Proceedings of the 2013 International Symposium on Intelligent Signal Processing and Communication Systems, Naha, Japan, 12–15 November 2013; pp. 704–707.

38. Monkaresi, H.; Hussain, M.S.; Calvo, R.A. Using Remote Heart Rate Measurement for Affect Detection. In Proceedings of the FLAIRS Conference, Sydney, Australia, 3 May 2014.

39. Monkaresi, H.; Calvo, R.A.; Yan, H. A Machine Learning Approach to Improve Contactless Heart Rate Monitoring Using a Webcam. IEEE J. Biomed. Health Inf. 2013, 18, 1153–1160.

40. Viola, P.; Jones, M.J.C. Rapid object detection using a boosted cascade of simple features. In Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2001), Kauai, HI, USA, 8–14 December 2001; Volume 1, p. 3.

41. Špetlík, R.; Franc, V.; Matas, J. Visual heart rate estimation with convolutional neural network. In Proceedings of the British Machine Vision Conference, Newcastle, UK, 3–6 September 2018; pp. 3–6.

42. Wei, L.; Tian, Y.; Wang, Y.; Ebrahimi, T.; Huang, T. Automatic webcam-based human heart rate measurements using laplacian eigenmap. In Proceedings of the Asian Conference on Computer Vision, Daejeon, Korea, 5–9 November 2012; Springer: Berlin/Heidelberg, Germany, 2012; pp. 281–292.

43. Xu, S.; Sun, L.; Rohde, G.K. Robust efficient estimation of heart rate pulse from video. Biomed. Opt. Express 2014, 5, 1124–1135.

44. Hsu, Y.C.; Lin, Y.L.; Hsu, W. Learning-based heart rate detection from remote photoplethysmography features. In Proceedings of the 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Florence, Italy, 4–9 May 2014; pp. 4433–4437.

45. Lee, K.Z.; Hung, P.C.; Tsai, L.W. Contact-free heart rate measurement using a camera. In Proceedings of the 2012 Ninth Conference on Computer and Robot Vision, Toronto, ON, Canada, 28–30 May 2012; pp. 147–152.

46. 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.

47. Poh, M.Z.; McDuff, D.J.; Picard, R.W. Advancements in noncontact, multiparameter physiological measurements using a webcam. IEEE Transact. Biomed. Eng. 2010, 58, 7–11.

48. Li, X.; Chen, J.; Zhao, G.; Pietikainen, M. Remote heart rate measurement from face videos under realistic situations. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; pp. 4264–4271.

49. Lee, K.; Lee, J.; Ha, C.; Han, M.; Ko, H. Video-Based Contactless Heart-Rate Detection and Counting via Joint Blind Source Separation with Adaptive Noise Canceller. Appl. Sci. 2019, 9, 4349.

50. Kwon, S.; Kim, H.; Park, K.S. Validation of heart rate extraction using video imaging on a built-in camera system of a smartphone. In Proceedings of the 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, San Diego, CA, USA, 28 August–1 September 2012; pp. 2174–2177.

51. Datcu, D.; Cidota, M.; Lukosch, S.; Rothkrantz, L. Noncontact automatic heart rate analysis in visible spectrum by specific face regions. In Proceedings of the 14th International Conference on Computer Systems and Technologies, Ruse, Bulgaria, 28–29 June 2013; pp. 120–127.

52. Holton, B.D.; Mannapperuma, K.; Lesniewski, P.J.; Thomas, J.C. Signal recovery in imaging photoplethysmography. Physiol. Meas. 2013, 34, 1499.

53. 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), Lisbon, Portugal, 5–8 January 2014; Volume 3, pp. 118–124.

54. Wang, W.W.; Stuijk, S.S.; De Haan, G.G. Exploiting Spatial Redundancy of Image Sensor for Motion Robust rPPG. IEEE Trans. Biomed. Eng. 2015, 62, 415–425.

55. Feng, L.; Po, L.M.; Xu, X.; Li, Y. Motion artifacts suppression for remote imaging photoplethysmography. In Proceedings of the 2014 19th International Conference on Digital Signal Processing, Hong Kong, China, 20–23 August 2014; pp. 18–23.

56. Tran, D.N.; Lee, H.; Kim, C. A robust real time system for remote heart rate measurement via camera. In Proceedings of the 2015 IEEE International Conference on Multimedia and Expo (ICME), Turin, Italy, 29 June–3 July 2015; pp. 1–6.

57. McDuff, D. Deep super resolution for recovering physiological information from videos. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Salt Lake City, UT, USA, 18–22 June 2018; pp. 1367–1374.

58. Hu, J.; Niu, H.; Carrasco, J.; Lennox, B.; Arvin, F. Voronoi-Based Multi-Robot Autonomous Exploration in Unknown Environments via Deep Reinforcement Learning. IEEE Trans. Veh. Technol. 2020, 69, 14413–14423.

59. Ciregan, D.; Meier, U.; Schmidhuber, J. Multi-column deep neural networks for image classification. In Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, USA, 16–21 June 2012; pp. 3642–3649.

60. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. Commun. ACM 2012, 60, 1097–1105.

61. Rouast, P.V.; Adam MT, P.; Chiong, R.; Cornforth, D.; Lux, E. Remote heart rate measurement using low-cost RGB face video: A technical lit-erature review. Front. Comput. Sci. 2018, 12, 858–872.

62. Liu, H.; Wang, Y.; Wang, L. A review of non-contact, low-cost physiological information measurement based on photople-thysmographic imaging. In Proceedings of the 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, San Diego, CA, USA, 28 August–1 September 2012; pp. 2088–2091.

63. Sun, Y.; Thakor, N. Photoplethysmography Revisited: From Contact to Noncontact, From Point to Imaging. IEEE Trans. Biomed. Eng. 2016, 63, 463–477.

64. Hu, S.; Peris, V.A.; Echiadis, A.; Zheng, J.; Shi, P. Development of effective photoplethysmographic measurement techniques: From contact to non-contact and from point to imaging. In Proceedings of the 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Minneapolis, MN, USA, 3–6 September 2009; pp. 6550–6553.

65. Hassan, M.A.; Malik, A.S.; Fofi, D.; Saad, N.; Karasfi, B.; Ali, Y.S.; Meriaudeau, F. Heart rate estimation using facial video: A review. Biomed. Signal Process. Control 2017, 38, 346–360.

66. Kranjec, J.; Beguš, S.; Geršak, G.; Drnovšek, J. Non-contact heart rate and heart rate variability measurements: A review. Biomed. Signal Process. Control 2014, 13, 102–112.

67. AL-Khalidi, F.Q.; Saatchi, R.; Burke, D.; Elphick, H.; Tan, S. Respiration rate monitoring methods: A review. Pediatr. Pulmonol. 2011, 46, 523–529.

68. Chen, X.; Cheng, J.; Song, R.; Liu, Y.; Ward, R.; Wang, Z.J. Video-Based Heart Rate Measurement: Recent Advances and Future Prospects. IEEE Trans. Instrum. Meas. 2019, 68, 3600–3615.

69. Kevat, A.C.; Bullen DV, R.; Davis, P.G.; Kamlin, C.O.F. A systematic review of novel technology for monitoring infant and newborn heart rate. Acta Paediatr. 2017, 106, 710–720.

70. McDuff, D.J.; Estepp, J.R.; Piasecki, A.M.; Blackford, E.B. A survey of remote optical photoplethysmographic imaging methods. In Proceedings of the 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Milan, Italy, 25–29 August 2015; pp. 6398–6404.

71. Li, P.; Hu, Y.; Liu, Z.-P. Prediction of cardiovascular diseases by integrating multi-modal features with machine learning methods. Biomed. Signal Process. Control 2021, 66, 102474.

72. Qiu, Y.; Liu, Y.; Arteaga-Falconi, J.; Dong, H.; El Saddik, A. EVM-CNN: Real-Time Contactless Heart Rate Estimation From Facial Video. IEEE Trans. Multimedia 2018, 21, 1778–1787.

73. Ren, S.; Cao, X.; Wei, Y.; Sun, J. Face alignment at 3000 fps via regressing local binary features. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 24–27 June 2014; pp. 1685–1692.

74. Luguev, T.; Seuß, D.; Garbas, J.U. Deep Learning based Affective Sensing with Remote Photoplethysmography. In Proceedings of the 2020 54th Annual Conference on Information Sciences and Systems (CISS), Princeton, NJ, USA, 18–20 March 2020; pp. 1–4.

75. Paracchini, M.; Marcon, M.; Villa, F.; Zappa, F.; Tubaro, S. Biometric Signals Estimation Using Single Photon Camera and Deep Learning. Sensors 2020, 20, 6102.

76. Zhan, Q.; Wang, W.; De Haan, G. Analysis of CNN-based remote-PPG to understand limitations and sensitivities. Biomed. Opt. Express 2020, 11, 1268–1283.

77. Chen, W.; McDuff, D. Deepphys: Video-based physiological measurement using convolutional attention networks. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September2018; pp. 349–365.

78. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. arXiv 2014, arXiv:1409.1556.

79. Reiss, A.; Indlekofer, I.; Schmidt, P.; Van Laerhoven, K. Deep PPG: Large-Scale Heart Rate Estimation with Convolutional Neural Networks. Sensors 2019, 19, 3079.

80. Available online: (accessed on 26 May 2021).

81. Fernandez, A.; Bunke RB, H.; Schmiduber, J. A novel connectionist system for improved unconstrained handwriting recog-nition. IEEE Transact. Pattern Anal. Mach. Intell. 2009, 31.

82. Sak, H.; Senior, A.W.; Beaufays, F. Long Short-Term Memory Recurrent Neural Network Architectures for Large Scale Acoustic Modeling. Available online: (accessed on 26 May 2021).

83. Li, X.; Wu, X. Constructing long short-term memory based deep recurrent neural networks for large vocabulary speech recognition. In Proceedings of the 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), South Brisbane, Australia, 19–24 April 2015; pp. 4520–4524.

84. Lee, E.; Chen, E.; Lee, C.Y. Meta-rppg: Remote heart rate estimation using a transductive meta-learner. arXiv 2020, arXiv:2007.06786.

85. Tran, D.; Bourdev, L.; Fergus, R.; Torresani, L.; Paluri, M. Learning spatiotemporal features with 3d convolutional networks. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015; pp. 4489–4497.

86. 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 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Korea, 27 October–2 November 2019; pp. 151–160.

87. Yu, Z.; Li, X.; Zhao, G. Remote photoplethysmograph signal measurement from facial videos using spatio-temporal networks. arXiv 2019, arXiv:1905.02419.

88. Perepelkina, O.; Artemyev, M.; Churikova, M.; Grinenko, M. HeartTrack: Convolutional Neural Network for Remote Video-Based Heart Rate Monitoring. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, Seattle, WA, USA, 14–19 June 2020; pp. 288–289.

89. 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.

90. Liu, S.-Q.; Yuen, P.C. A General Remote Photoplethysmography Estimator with Spatiotemporal Convolutional Network. In Proceedings of the 2020 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020), Buenos Aires, Argentina, 16–20 November 2020; pp. 481–488.