Computer Methods and Programs in Biomedicine 219 (2022) 106771

Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

journal homepage: www.elsevier.com/locate/cmpb

Availability and performance of face based non-contact methods for heart rate and oxygen saturation estimations: A systematic review

Ankit Gupta [a,c,*], Antonio G. Ravelo-García [a,b], Fernando Morgado Dias [a,c]

[a] Interactive Technologies Institute/Larsys/Madeira Interactive Technologies Institute, Caminho da Penteada, Funchal, 9020-105, Portugal
[b] Universidad de Las Palmas de Gran Canaria, C. Juan de Quesada, 30, Las Palmas, 35001, Spain
[c] Universidade da Madeira, Caminho da Penteada, Funchal, 9020-105, Portugal

ARTICLE INFO

Article history:
Received 8 October 2021
Revised 3 March 2022
Accepted 21 March 2022

Keywords:
Non-contact estimation approaches
Physiological parameters
Blood volume pulse
Heart rate
Oxygen saturation

ABSTRACT

Background: Consumer-level cameras have provided an advantage of designing cost-effective, non-contact physiological parameters estimation approaches which is not possible with gold standard estimation techniques. This encourages the development of non-contact estimation methods using camera technology. Therefore, this work aims to present a systematic review summarizing the currently existing face-based non-contact methods along with their performance.

Methods: This review includes all heart rate (HR) and oxygen saturation (SpO2) studies published in journals and a few reputed conferences, which have compared the proposed estimation methods with one or more standard reference devices. The articles were collected from the following research databases: Institute of Electrical and Electronics Engineers (IEEE), PubMed, Web of Science (WoS), Science Direct, and Association of Computer Machinery (ACM) digital library. All database searches were completed on May 20, 2021. Each study was assessed using a finite set of identified factors for reporting bias.

Results: Out of 332 identified studies, 32 studies were selected for the final review. Additionally, 18 studies were included by thoroughly checking these studies. 3 out of 50 (6%) studies were performed in clinical conditions, while the remaining studies were carried out on a healthy population. 42 out of 50 (84%) studies have estimated HR, while 5/50 (10%) studies have measured SpO_2 only. The remaining three studies have estimated both parameters. The majority of the studies have used 1–3 min videos for estimation. Among the estimation methods, Deep Learning and Independent component analysis (ICA) were used by 11/42 (26.19%) and 9/42 (21.42%) studies, respectively. According to the Bland-Altman analysis, only 8/45 (17.77%) HR studies achieved the clinically accepted error limits whereas, for SpO_2, 4/5 (80%) studies have matched the industry standards (±3%).

Discussion: Deep Learning and ICA have been predominantly used for HR estimations. Among deep learning estimation methods, convolutional neural networks have been employed till date due to their good generalization ability. Most non-contact HR estimation methods need significant improvements to implement these methods in a clinical environment. Furthermore, these methods need to be tested on the subjects suffering from any related disease. SpO_2 estimation studies are challenging and need to be tested by conducting hypoxemic events. The authors would encourage reporting the detailed information about the study population, the use of longer videos, and appropriate performance metrics and testing under abnormal HR and SpO_2 ranges for future estimation studies.

1. Introduction

Physiological parameters serve as reliable indicators of an individual's health. Five physiological parameters used for determining the individual's health status are Blood Pressure (BP), Heart Rate (HR), oxygen saturation (SpO_2), body temperature, and Breathing Rate (BR) [1]. Estimating these parameters requires sophisticated apparatus that involves the placement of electrodes or sensors in

* Corresponding author at: Interactive Technologies Institute/Larsys/Madeira Interactive Technologies Institute, Caminho da Penteada, 9020-105, Funchal, Portugal.
E-mail addresses: ankit.gupta@iti.larsys.pt (A. Gupta), antonio.ravelo@ulpgc.es (A.G. Ravelo-García), morgado@uma.pt (F.M. Dias).

https://doi.org/10.1016/j.cmpb.2022.106771

contact with the skin using adhesive gels. This approach of estimating parameters is called the contact-based approach, which has limitations such as discomfort or allergies due to the sensors or electrode placement. Therefore, it is not suitable in certains scenarios such as unobtrusive monitoring, sensitive or burnt skin, and neonatal intensive care unit (ICU) monitoring. This orignates the need for an alternative non-contact based estimation approach that is based on a non-contact variant of Photoplethysmography (PPG) known as remote PPG (rPPG) or imaging PPG (iPPG). rPPG or iPPG requires a camcorder with an illumination source to select the Region of Interest (ROI) appropriately, followed by a PPG signal extraction and physiological parameter estimations [2]. An appropriate ROI selection is critical for accurate PPG extraction, thereby estimating correct physiological parameters. Several studies have been performed to explore the potential ROIs for estimation. In this context, the first attempt of ROI selection was performed by Pavlidis et al. [3], which explored the potential of the face in estimating cardiac pulse, BR, and blood flow, using a thermal imaging model. The study concluded that the thin layer of tissues present in the facial region enables the estimation of the blood volume pulse accurately. Furthermore, Verkruysse et al. [4] proved the potential of the facial area for HR and respiratory rate monitoring.This was further validated in a study conducted by van der Kooij and Naber [5], which explored the potential of different body organs and found that the facial region can provide more accurate HR estimates than other body organs. SpO2 can also be accurately estimated using facial ROI [6]. Currently, three vital parameters, namely HR, SpO2, and BR, can be computed using PPG, hence rPPG. Among these, HR and SpO2 are estimated in various critical scenarios such as intensive care units, surgery, COVID diagnosis, and sleep quality analysis, [7,8]. Furthermore, breathing rate is not the most frequently used monitoring vital signs and has relatively limited applicability [9]. Therefore, this review focuses on face-based HR and SpO2 non-contact estimation studies.

Several aspects have already been explored in the existing reviews in the literature. Hassan et al. [10] and, Wang et al. [11] have explored the existing HR estimation methods followed by comparing them using benchmark datasets. Sun and Thakor [12] discussed the implications of imaging PPG (iPPG) and its technical limitations. Kranjec et al. [13] presented a review summarizing image and non-image based methods for HR estimation. Hardford et al. [14] presented a systematic review that discusses the clinical aspect of all image based non-contact approaches using a modified Guidelines for Reporting Reliability and Agreement Studies (GRRAS) criterion.

However, no systematic review presented the non-contact estimation approaches from the engineering specific context. Therefore, this review aims to provide a detailed review of non-contact facial video-based physiological parameter estimation methods in clinical and non-clinical settings. This is the first systematic review summarizing the technical and non-technical elements of face-based non-contact estimation approaches to the best of our knowledge. This will provide a basis for further research in this thriving domain and help improve the quality of future studies.The contributions of this systematic review for standardizing and improving the non-contact estimation studies are as follows:

- This review strictly follows newly updated PRISMA guidelines [15] and answers all the relevant questions regarding designing a non-contact estimation study, such as selecting ROI, estimation methods, and performance metrics.
- It summarizes and analyzes the existing HR and SpO2 non-contact estimation studies published in journals, collected using five research databases. It also proposes a novel protocol for "risk of bias" analysis by identifying crucial factors from the existing studies.
- It presents the statistical analysis of all the performance metrics used in the existing studies. This will allow a baseline to compare the newly proposed methods with existing state-of-the-art studies' already reported performance metrics.
- It also presents the advantages and disadvantages of the various parameters, estimation methods, and data acquisition methods, enabling the identification of a suitable regime for conducting a physiological parameter estimation study.

2. Methods

2.1. Eligibility criteria

This systematic review is prepared following the new Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement [15]. The PRISMA checklist for this review is available as supplementary material Annex 1.

It includes all the studies related to facial video-based HR and SpO_2 estimation in comparison with the appropriate reference devices. Few HR estimation studies have included other physiological parameters: breathing rate, step count, eye blink, and HRV. It is important to note that no explicit searches have been performed for the estimation of these parameters. Instead, the search is limited to journal papers only. Conference papers and book chapters were not considered due to limited novelty in the studies. However, papers published in reputed international conferences such as the conference on computer vision and pattern recognition (CVPR), International Conference on Computer Vision, and European Conference on Computer Vision (ICCV) were included in this review. Additionally, patents were also not included since these do not align with the protocol used in this systematic review.

2.2. Information sources

The search has been performed in the following databases: Institute of Electrical and Electronics Engineers (IEEE), PubMed, Web of Science (WoS), Science Direct and, Association of Computer Machinery (ACM) digital library. The database search was completed on May 20, 2021.

2.3. Search strategy

The articles search was conducted between December 2008 to May 2021, since the first attempt for Heart rate and SpO_2 estimation using the face region was published in December 2008 by Verkryusse et al. [16]. Furthermore, the studies from the engineering specific context were considered for this review. Specifically, studies demonstrating the novel approaches for any of the components/modules (Region of Interest selection, BVP/PPG signal extraction, and Heart rate calculation) of estimation using face videos using any imaging modality were only considered for this review. Specifically, this review includes the studies conducted using various image color models such as RGB, CIE, LAB, YCbCr, YUV, NIR, and monochrome color filters(e.g., magenta, orange). However, studies demonstrating the implications of existing non-contact methods in the scenarios such as driving were not included. Related estimation studies using other organs of the human body such as arms, hand palms, etc., were excluded. Moreover, the studies relating to fetal heart rate monitoring, patents related to HR or SpO_2 estimation methods or devices, use of smartphone sensors by contact, PPG signals or contact based approaches, and animal physiological parameters estimations were also excluded. The sole reason for excluding these studies is that they are beyond the scope of this review. The details of the search strategy are presented in Table 1.

Table 1
Search Strategy.

	Search Strategy
1	(Contactless OR Noncontact OR non-contact OR non-invasive OR non-invasive OR remote OR contact free OR contact-free).ti
2	(Oximetry).MeSH OR (SpO_2 OR "blood oxygen* OR blood saturation OR oxygen saturation).ti
3	(Heart Rate Determination).MeSH OR (Heart rate OR Pulse rate NOT(Variability)).ti
4	(face AND video* OR webcam OR camera OR smartphone OR mobile camera).ti
5	(iPPG OR rPPG OR Physiological).ti OR Photoplethysmography).MeSH
6	2 AND 3
7	3 OR 4
8	(1 OR 7) AND 6

2.4. Study selection and data collection

All the articles resulting from the implemented search strategy were consolidated to the Endnote reference manager. First, the set of duplicate conference articles and patents were screened out using the reference manager. The researcher Ankit Gupta (AG) performed the title, abstract and full text screening. Two researchers Antonio G. Ravelo-García (AR) and Fernando Morgado Dias (FMD), were consulted to avoid doubts or confusion in the article screening process.

2.5. Data collection and outcomes

The data are collected by thoroughly reading the article's full text. Consequently, AG collected study title, estimated physiological parameters, reference device(s) used, video length, frame rate (fps), camera resolution, shooting distance, databases used, number of subjects, age range, gender, clinical/normal study, skin (types or color) and ethnicity information, ROI used, ROI selection method, BVP/PPG extraction method, color channels, lighting source, performance metrics, Bland-Altman analysis information, type of artifacts (Motion, Illumination or both) addressed.

The potential outcome of this study is to present the available face-based non-contact HR and SpO_2 estimation methods using various image modalities in comparison with the respective reference device(s). The secondary outcome is to assess their performance defined by the inclusion/exclusion of important parameters included in every study and to analyze the corresponding reported performance metrics, their practical implications, and limitations. Furthermore, the common performance metrics reported for all studies were selected and used for study categorization using a threshold, which could mitigate the effect of low and high extreme values. Due to heterogeneity among HR and SpO_2 estimation studies, different performance metrics were independently chosen for both types of studies.

2.6. Study quality assessment

AG performed the quality assessment of the included articles using a proposed scoring scheme, supervised by AR and FMD. The scores were given based on the inclusion of selected parameters as well as performance metrics for HR and SpO_2 studies.

The proposed scoring protocol uses a two-step scoring process: scoring based on the inclusion and exclusion of the following factors: camera resolution, shooting distance, number of participants, Bland-Altman analysis, motion or illumination artifacts (HR studies), ethnicity; analyzing the performance metrics (RMSE and correlation for HR and R-squared for SpO_2) and accuracy (error$< \pm 5$ beats per minute) (bpm) justifying the method's clinical relevance. For inclusion of each parameter, a score of one is assigned and zero otherwise, except artifacts. The studies addressing motion and illumination artifacts simultaneously, are assigned a score of 2, whereas, in the case of none or one of the artifacts, a score of 0 or 1 is assigned to each study, respectively. For quantitative values such as the number of participants, RMSE, correlation, and R-squared values, the median is calculated to mitigate the effect of lower and higher extremes. Additionally, in some cases, a single score is also calculated by combining the score of the individual parameters. For instance, the "AND" operation for video resolution and camera distance and the "OR" operation for RMSE and Correlation were used for HR estimation studies.

Individual scores are summed up for overall quality assessment. For HR estimation studies, these scores are finally used to categorize studies into three categories: weak, fair, and strong. The studies with the score less than 2 will be categorized as weak, since in almost all the reported studies, camera characteristics (camera shooting distance, video resolution), and performance metrics (RMSE, and correlation) were reported summing up to a score of 2. Studies with a score between 3 and 5 would be categorized as fair, whereas a score ranging from 6 to 8 would signify a strong study. On the other hand, SpO_2 studies with a score of 1 would be categorized as weak, since majority of studies used regression, while the studies with the score of 3 and 4 would be considered fair and strong studies.

2.7. Visual interpretation and tabulation of results

This review uses box-whisker plots and bar and error charts. Box-whisker plots will be used to examine the distribution of performance metrics used for estimation studies, while bar charts will be used to visualize the study categorization, distributions of regions of interest, and the estimation methods for HR and SpO_2 studies. Error charts will be used for analyzing the upper and lower statistical limits of Bland-Altman plots. Additionally, tables will also be used for presenting statistical values (mean $\pm$ standard deviation) of all performance metrics used in the estimation studies in the review and identified factors collection and scoring in supplementary files.

2.8. Heterogeneity, missing data, and subgroup analysis

Since all the studies have reported different performance metrics to support the efficacy of methods, calculating a single statistical metric would be infeasible and insufficient. Alternatively, similar numerical metrics were compared to provide a descriptive summary.

The proposed protocol penalizes for missing data, which is required for the study's quality assessment.

Since the authors have collected the data from individual studies and assessed the studies' quality based on the data, they have not performed any subgroup analysis. Furthermore, this review aimed at providing a narrative summary based on the data collected from the included studies for HR and SpO_2 estimations independently.

3. Results

3.1. Study screening results

Out of 332 articles retrieved from the search strategy presented in Table 1 using multiple databases, 32 articles were included, followed by data collection and analysis. While screening these articles, 18 more studies were included by thoroughly checking the references list of every included article. This technique is called

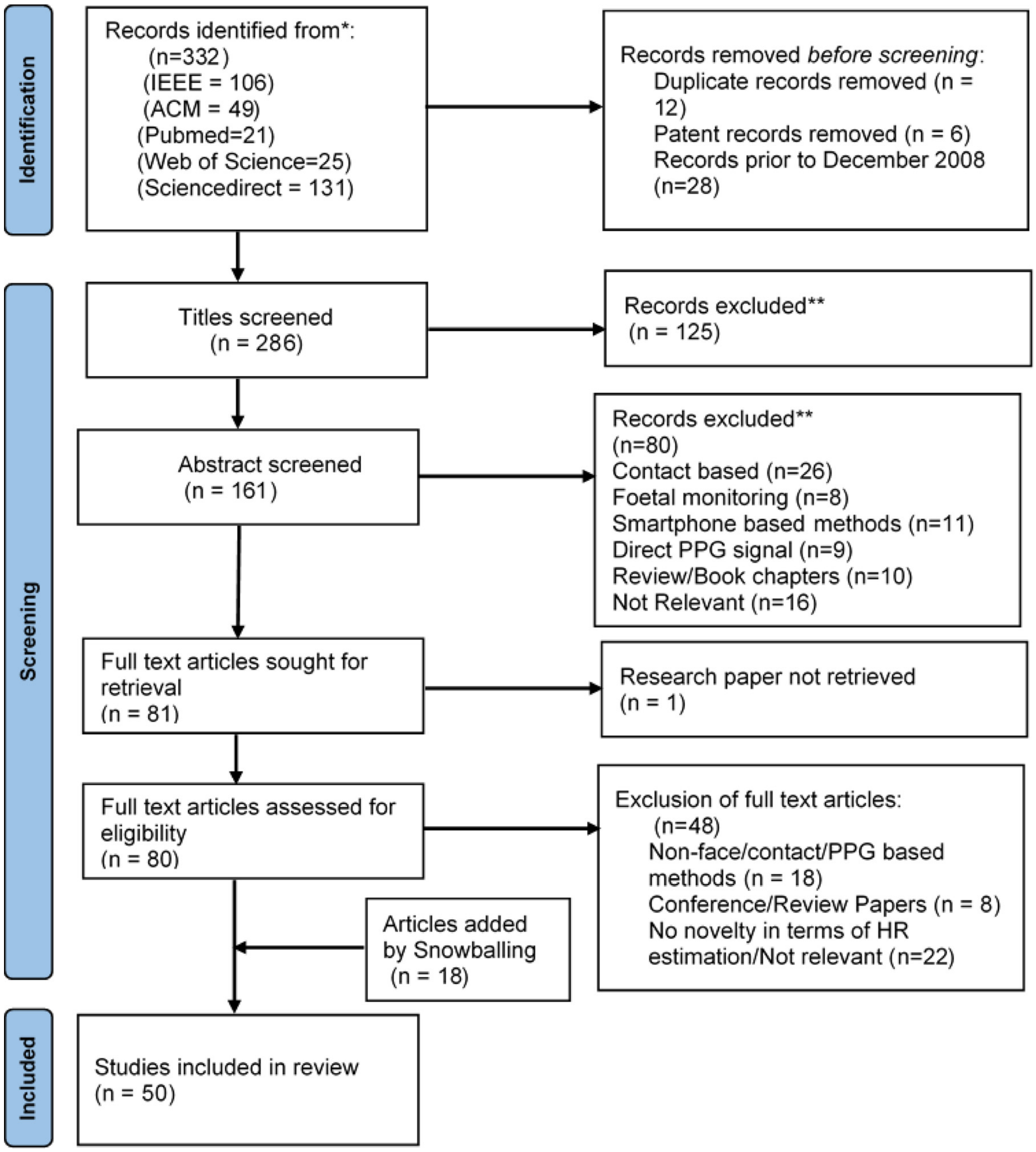


Fig. 1. Prisma Flow Diagram.

snowballing. Fig. 1 depicts the PRISMA flow diagram illustrating the article's screening process. Consequently, 50 articles were included in the final review, out of which 38/50 (76%) studies estimated, 10/50 (20%) studies estimated two, and the remaining estimated three physiological parameters. It is important to note that no explicit search was performed for other parameters except for HR and SpO_2, although other parameters were also estimated as a part of HR estimation studies.

However, the study did not include non-contact methods using chest, arm, palm, or finger for HR estimations since this review is constrained to face-based methods only. Furthermore, studies constituting fetal heart rate monitoring were not considered since HR estimation was performed using the lower abdominal area.

One article's full text by Zhang et al. [17] could not be retrieved for which the authors were contacted, but no responses were received.

3.2. Population characteristics

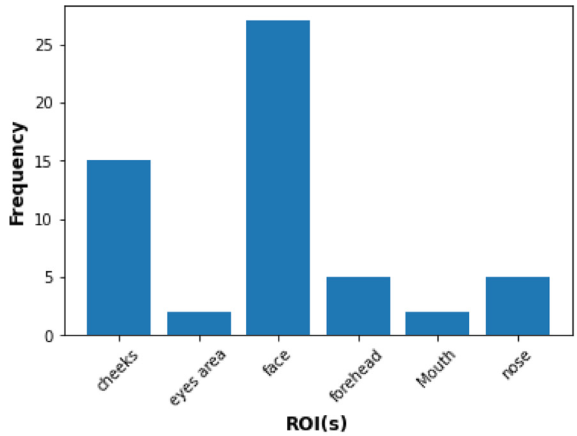
3.2.1. Age and gender

21/50 (42%) did not report the age, while the remaining 29/50 (58%) studies reported the age range. The minimum and maximum age range for all studies lie between 18 and 80. Furthermore, it is difficult to plot the distribution of age ranges due to the considerable heterogeneity.

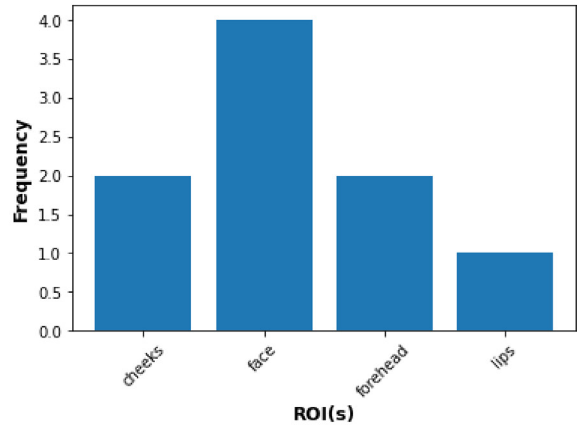
Gender has been reported by 34/50 (68%) studies, while 16 studies did not provide any relevant information. Apart from 4/50 (8%) studies [18–21], all the study samples were male dominant. However, these studies have collected data from a relatively lower number of participants.

3.2.2. Ethnicity and skin color

Numerous studies proved the importance of considering the ethnicity or skin color for HR and SpO_2 estimation studies since



(a) HR estimation Studies.



(b) SpO_2 estimation studies.

Fig. 2. ROI distribution with HR and SpO_2 studies.

darker skin tone poses more challenges for estimation tasks than white skin. 14/50 (28%) studies reported the subjects' ethnicity, skin color, or tone information for estimation studies. Furthermore, 8/14 (57.14%) [8,22–27] studies used the Fitzpatrick scale to define the subjects' skin tone. The study conducted by Haan and Jeanne [28] included the subjects from i-vi (all scales), while 2 studies conducted by Wang et al. [22,23] included subjects with the i-v scale. The remaining studies included subjects with a scale ranging from ii-iv. On the other hand, 6/14 (42.85%) studies instead mentioned the ethnicities of the subjects: 3 studies [2,29,30] considered two or more,2 studies [18,23] considered subjects with Asian ethnicity, and one study conducted by Kumar et al. [31] has considered skin color.

3.3. Study design

3.3.1. Physiological parameters

Out of the total included studies, 42/50 (84%) studies belong to HR estimations, while 5/50 (10%) to SpO_2 estimations. Additionally, 3/50 (6%) studies estimated both physiological parameters simultaneously. Furthermore, out of 42 studies, HRV, breathing rate, eye blink and step counts were estimated in 4/50 (8%) [20,30–32], 4/50 (8%) [32–35], and 1/50 (2%) studies, respectively.

3.3.2. Databases used

For HR studies, 34/45 (75.55%) studies used self-created databases with the number of participants ranging from 4 to 117 with $25.82 \pm 25.11 (mean \pm std)$. In contrast, the remaining studies used benchmark databases to prove the efficacy of their respective HR estimation methods. Moreover, 24/34 (35.29%) studies have only used their databases, while the rest have used self-created, as well as publicly available databases. 3/11 (27.27%) studies [21,36,37] have used a single database, while 8/11 (72.72%) [38–45] studies have employed more than one database for performance analysis of their HR estimation algorithms. All SpO_2 studies created their databases with the number of participants ranging from 4 to 46 with 20.5 ± 16.78.

3.3.3. Region of interest selection

The face-based physiological parameters estimation needs an ROI, which will be used to extract the source signal. In total, all HR estimation studies used six ROIs, namely, face, cheeks, nose, forehead, areas near eyes, and mouth. 7/45 (15.5%) [25,30,34,46–49] studies have used two or more ROIs from the face region, while face and cheeks were used by 27/45 (60%) and 15/45 (33.33%) studies, respectively. On the other hand, nose and forehead have been used by five studies each while the remaining used mouth and areas near eyes. The ROI distribution for HR and SpO_2 studies is shown in Fig. 2.

3.3.4. Artifacts

Artifacts removal is an important step for BVP or PPG signal extraction using facial videos since the PPG component has relatively weaker strength and amplitude than the artifacts. HR studies have addressed two types of artifacts, namely motion and illumination artifacts. 4/45 (8.88%) HR studies have addressed and alleviated the effect of illumination, while motion artifacts have been addressed and mitigated in 16/45 (35.55%) studies. On the other hand, 16/45 (35.55%) studies [19,23,24,28,29,31,35,39,40,42,45,49–52] have addressed and proposed strategies to lessen the effect of both artifacts. 9/45 (20%) HR and all SpO_2 estimation studies did not address any artifact.

3.3.5. Estimation methods

As mentioned before, conventional HR estimation methods deal with extracting the BVP/PPG signal from the RGB signal traces followed by calculating the highest frequency and multiplying it with 60 for frequency to bpm conversion. Among conventional HR estimation methods, 9/45 (20%) studies [2,30,32,34,42,51,53–55] used Independent Component Analysis (ICA), 6/45 (13.33%) studies [20,23,28,29,41,48] used color subspace transformations (CHROM/POS), and 6/45 (13.33%) studies [18,50,56–59] used Ensemble Mode Decomposition (EMD) and its variants. However, with the advent of deep learning, several end-to-end HR estimation methods have also been proposed, which use different types of neural networks architectures for estimating the heart rate using a facial video. 11/45 (24.44%) [24,34,35,37–41,44,45,52] studies utilized neural networks and its variants for HR estimation. Other PPG extraction methods used by HR estimation studies are wavelet transforms [19,25,60], filtering based methods [47], autoregressive models [29], Gaussian mixture models [61], Eulerian video magnification [33], independent vector analysis [18,21,44], maximum ratio diversity [31], sub-band selection [19], multiple linear regression [59], and Second Order Blind Identification (SOBI) [34,55]. Fig. 3 depicts the distribution of the estimation methods used in the literature, whereas Table 2 presents the classification of estimation methods as the feature and classifier based methods. On the other hand, all SpO_2 studies used the ratio of ratios method followed by regression, except the study conducted by Van Gastel et al. [26]. This study proposed an adaptive PBV method (APBV) for BVP signals extraction from multiple ROIs, followed by their

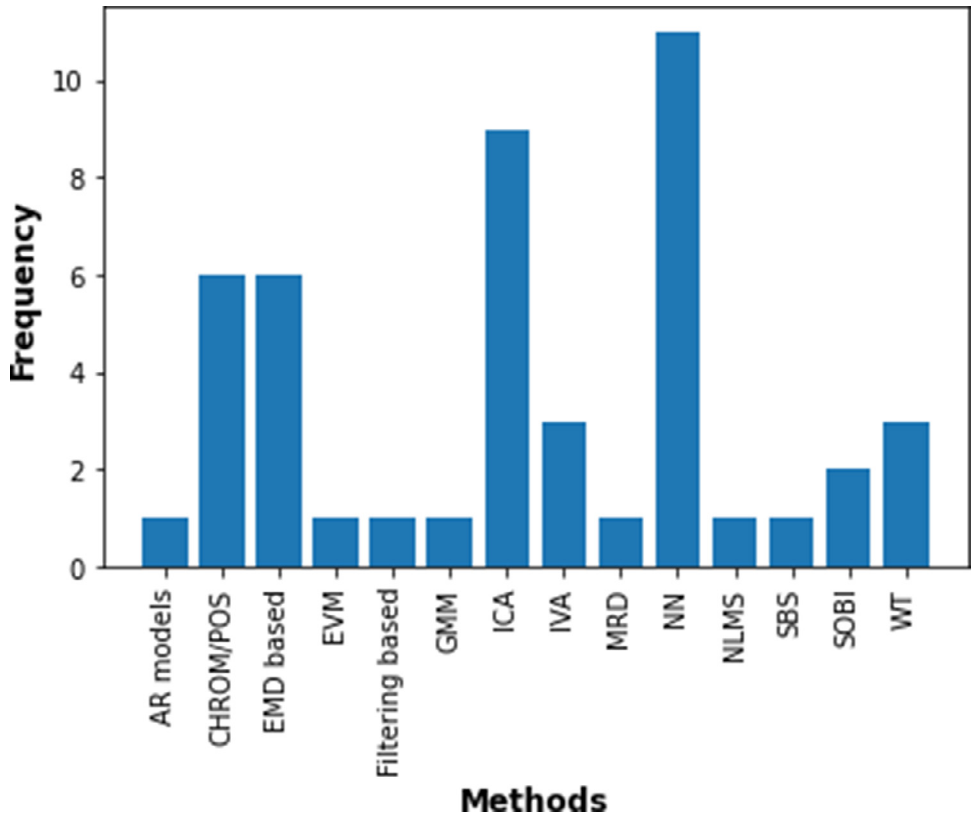


Fig. 3. Various estimation methods used for estimation of HR studies:AR-Autoregressive,EVM- Eulerian Video Magnification, GMM-Gaussian Mixture Model, ICA-Independent Component Analysis, IVA- Independent Vector Analysis, MRD-Maximum Ratio Diversity, NN-Neural Networks, NLMS- Non-Linear mean square filter, SBS-Sub-Band Selection, SOBI-Second Order Blind Identification, WT-Wavelet Transform.

Table 2
Estimation Methods classification.

Feature based methods	Classifier based methods
Color subspace transformations (CHROM [20,28,41], POS [23,29,48]), EMD based (EEMD [18,50,57–59], CEEDMAN [56]), EVM [36], JBSS (IVA [18,21,44]), MRD [31], SOBI [34,55], SBS [22], WT [19,25,60], ANN [24], ICA [2,30,32,34,42,51,53–55], Filtering Based methods [47], and GMM [61].	Convolutional 3D Networks [37], VGG-16 [39]), CNN [36,40], CNN with attention based mechanisms [35,38,45], ResNet-18 [41], Autoregressive Models [33], and Mulitple Regression [59].

mapping to different SpO_2 levels ranging from 65 to 100% with an interval of 5%. The proposed APBV method is based on the PBV method proposed by De Haan and Van Leest [62], which extracts the blood volume pulse vector (PBV) for BVP signal extraction.

3.4. Instruments used

3.4.1. Reference devices

Four reference devices have been used for comparing the estimated values with ground truth. This review has only reported the use of reference devices for the self-created databases since reference devices information for the benchmark databases could be easily extracted from the respective articles. As mentioned before, 34 HR estimation studies have created their databases. 8/34 (23.52%) studies have used electrocardiogram as a reference device, 22/34 (64.70%) studies have used pulse oximeters as a reference device, while one study conducted by Wang et al. [23] used both reference devices. Few HR estimation studies used Arm Band HR monitor [24] and sphygmomanometer [57,59]. On the other hand, SpO_2 ground truth acquisition was carried out using a pulse oximeter only.

3.4.2. Camera characteristics

Non-contact studies require a camera for video acquisition of ROI; hence camera characteristics such as camera type, video resolution, frame rate, and shooting distance affect the accuracy of parameter estimation. However, the camera type was not included in the study's quality assessment, but it is a critical parameter since it defines the quantity and quality of PPG information for physiological parameters estimation.

The camera type is defined by the color channels used for video acquisition in a non-contact study. The color channel selection is crucial for extracting accurate PPG information. For instance, the RGB spectra have stronger pulsatile strength than the infrared spectra. Therefore, this section summarizes the respective studies' color channel distribution and camera types for the included studies. 40/50 (80%) studies used red, green, and blue channels, followed by NIR used by 6/50 (10%) studies. Other color channels used in the existing non-contact studies include monochrome color filters [30,63,64], YCbCr [19,60], YUV [52], and LAB [50]. Furthermore, a few studies have used more than one spectra, e.g., Kado et al. [51] and Yu et al. [20] used a combination of RGB and NIR spectra. The details regarding the color channels are presented in supplementary files (Table 1). Additionally, YUV, YCbCr, and LAB color channels can be deduced from the RGB image model; therefore, 42 studies have used RGB cameras. Four studies used NIR cameras, whereas five used customized camera setups, with three studies using monochrome and two using RGB-NIR spectra combination.

The distance between the subject's face and the camera is another important parameter since a larger distance between the face and the camera deteriorates the strength of PPG signal information. Hence, it is necessary to identify a suitable shooting distance for a cleaner PPG signal. The shooting distance for HR and SpO_2 estimations ranged between 0.3 and 2 m, respectively, while, the widely used shooting distances were 0.5 and 1.0 m, used by 11/50 (22%) and (13/50) (26%), respectively. However, 4/50 (8%) studies used less than 0.5 m shooting distance, while it is 1.5 m or greater for 10/50 (20%) studies. Furthermore, a few studies have acquired the videos using more than one shooting distance. Specifically, the studies conducted by Song et al. [53] and Tran et al. [29] tested the effect of video shooting distances on HR estimation, whereas different shooting distances for different activities were also used by Li et al. [49]. 12/50 (24%) HR estimation studies did not report camera shooting distances.

Camera resolutions also play a vital role in accurate HR estimation by providing finer details from individual image frames, which are crucial to detect subtle color changes for extracting PPG information. A higher camera resolution provides more information but also needs intense computations. Hence, identifying a camera resolution with minimal information loss is a non-trivial task for accurate HR estimation. A diversified range of video camera resolutions has been employed to estimate accurate HR and SpO_2 estimations 19/45 (42.22%) HR estimation studies used cameras with a resolution of 640 × 480. In contrast, the twelve studies used 320 × 240, 1280 × 720, and 1920 × 1080, each employed by four HR studies.

A higher frame rate provides a larger number of contiguous images for a video, thereby providing more information to detect the blood volume pulse from raw RGB signal traces. Similar to resolution, a higher frame rate has more computational requirements. Hence, it is necessary to use a frame rate that not only ensures minimal loss but also provides a noise free PPG signal. All estimation studies, except in one, by Song et al. [41] have reported the frame rates for video acquisition with a range of 12–120 frames per second (fps). 29/50 (58%) estimation studies used 30 frames per second (fps) for video acquisition. Other frame rates used by estimation studies were 15 fps [2,26,32,46,54], 20 fps [22,23,28,44,45], and 25 fps [22,23,28,45,63]. However, numerous studies have also gathered the video samples at a higher sampling rate, for instance, 50 fps [20], 60 fps [38], 100 fps [20], etc.

3.5. Clinical studies

Although most estimation studies were conducted on healthy individuals, three clinical studies [20,33,60] have been included using the search strategy mentioned in Table 1. Most importantly, these studies have estimated two or more physiological parameters. Yu et al. [20] conducted a study on geriatric patients which aimed at estimating heart rate and heart rate variability, while the study conducted by Tarassenko et al. [33] estimated heart rate, SpO_2, and breathing rate of the patients undergoing dialysis. The study conducted by Bal [60] aimed at estimating the heart rate and SpO_2 in the pediatric intensive care limit.

3.6. Performance metrics

3.6.1. HR estimation studies

The performance analysis for HR estimation studies utilized 5 metrics, namely, mean and standard deviation error (SD), root mean square error (RMSE), mean of error-rate percentage (MER), signal to noise (SNR) ratio, and correlation. Among all of them, the majority of studies used RMSE and correlation. The mean and standard deviations of all metrics are provided in Table 3. Few studies have tested their estimation algorithms under different application scenarios or using multiple databases wherein average RMSE or correlation values were calculated for the analysis. Additionally, accuracy and Bland-Altman analysis were also included. The details of all metrics are found in the supplementary file (Table 6-Table 12). 25/45 (55.55%) studies reported RMSE, out of which 10/25 (40%) have achieved the RMSE within ±5 bpm, whereas the remaining studies have the mean RMSE of 2.73 bpm and standard deviation of 1.32 bpm. A box-whisker plot in Fig. 4a depicts the error distribution of error metrics. As shown in the figure, higher RMSE values for the two studies correspond to testing the proposed methods under challenging conditions such as fitness exercise and cold pressure test (in which subjects were asked to put hands in very cold water), while collecting the video and the ground truth HR values.

Table 3
Performance Metrics Statistics.

Metric	Number of Studies	*Mean ± SD*
Mean Error	9	0.57 ± 1.49
Standard Deviation	17	4.91 ± 2.41
RMSE	28	5.15 ± 3.24
MER	8	6.03 ± 1.26
Correlation	26	0.88 ± 0.11
Signal to Noise ratio	5	3.17 ± 1.75

Pearson correlation values have been reported in 26/45 (57.77%) of the HR estimation studies. Among them, 15/26 (57.69%) [2,19,28,32,34,36,38,42,45,49,50,53,57,59,60] studies achieved the correlation value of 0.90 or more, while the average correlation value for the remaining studies is 0.77 with the standard deviation of 0.093.

10/45 (22.22%) [19,21,24,29,43,49,51,59,60,65] studies have reported accuracy, which is, calculated as the percentage of samples with the error difference $\pm 5bpm$ between ground truth and estimated HR values. This metric is used to justify the clinical relevance, since the clinically accepted error between reference device measurement and estimation is ±5bpm [66]. The Pearson correlation values and accuracy distribution are shown in Fig. 4b.

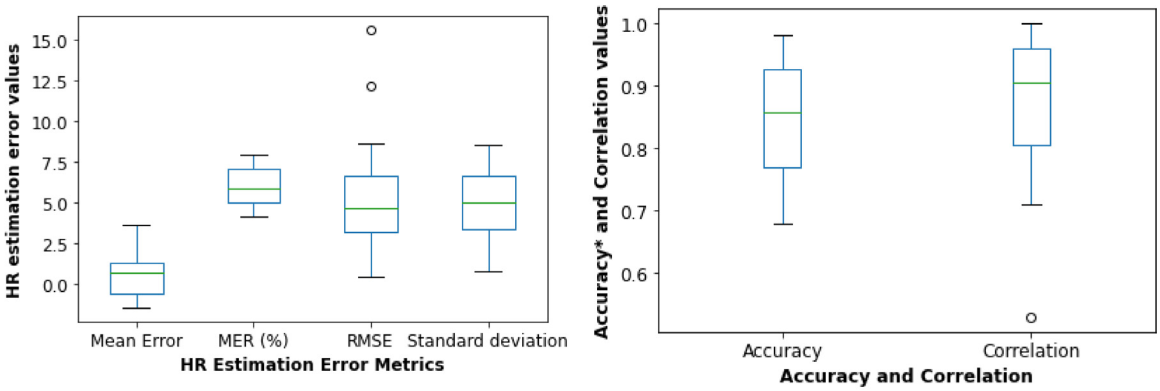
23/45 (51.11%) studies have included B-A plots in their analysis. Additionally, one study by Lin and Lin [47] did not present the B-A plot, but rather the level of agreements. Fig. 5a depicts the mean bias and upper and lower levels of agreement for HR estimation studies in chronological order. 8/23 (34.78%) studies achieved the mean difference within the clinically accepted range, while the rest may need significant improvements in the future. The detailed limits of agreement of all studies are presented in the supplementary file (Table 4 and Table 5).

3.6.2. SpO_2 estimation studies

As mentioned before, 7/8 (87.5%) non-contact estimation studies used regression analysis for SpO_2 (with range 80–100%) calculations, utilizing the ratio of ratios method using two wavelength light intensities. 5/8 (62.5%) [26,27,46,63,64] studies have reported R^2 values. Furthermore, the root mean squared metric (A-rms %) was calculated for two studies [8,46], while the same number of studies [60,63] have used the Pearson correlation value to test the algorithm's performance. 3/5 (60%) studies [26,27,63] have achieved an R^2 value of 0.8 or more, while the remaining two studies achieved relatively lower values, 0.65 and 0.58, respectively. Overall, the mean R2 value is 0.78, with a standard deviation of 0.14. Furthermore, 5/8 (62.5%) studies have used B-A plots to showcase the performance of the proposed methods, as depicted in Fig. 5b.

3.7. Challenges

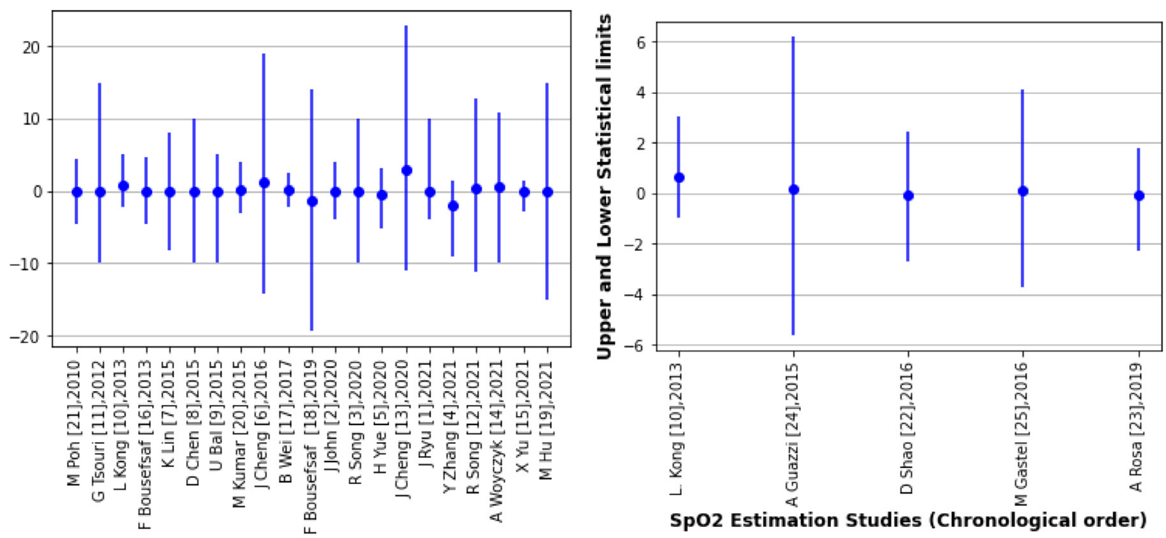
Non-contact methods for HR estimation studies deal with three types of noises: camera quantization, motion, and illumination



(a) Error Metrics Distribution

(b) Accuracy and Correlation Distribution

Fig. 4. HR Performance metrics distribution of HR estimation studies.



(a) Error bar charts for HR studies.

(b) Error bar charts for SpO_2 studies.

Fig. 5. A summary of Bland-Altman analysis for HR and SpO_2 estimation studies.

noise. Almost all the studies have assumed a constant light illumination incident on every region of the face, which does not comply with the real-time situations. Secondly, the low strength of the PPG signal as compared with the noise due to motion and illumination artifacts poses a major challenge for extracting the cleaner PPG signal. Low resolution and camera shooting distance further degrade the quality of the acquired PPG signal. Furthermore, a temperature colder than the ambient may increase the blood viscosity, which reduces the blood flow, thereby resulting inaccurate PPG signal extraction. Therefore, camera quantization noise, motion and illumination variation artifacts, extreme colder temperature conditions, camera characteristics such as the camera shooting distance and resolution are critical challenges that need to be addressed while designing the non-contact estimation study.

3.8. Applications

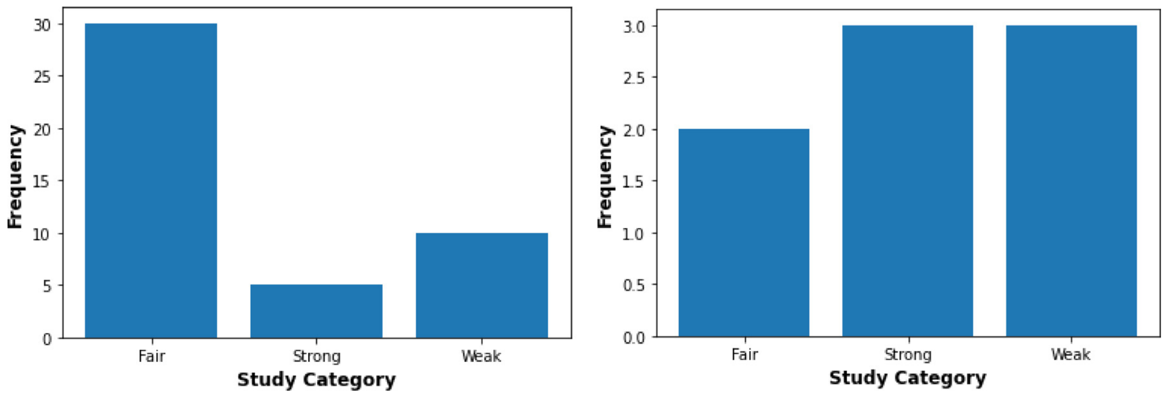
The majority of existing non-contact methods for physiological parameters estimations are in a proof of concept stage. However, several studies have deployed their methods for real-time applications. 8/50 (18%) studies have deployed their proposed non-contact estimation methods in real-time situations like fitness exercise, driving, and clinical conditions.

3/8 (33.33%) studies have tested their methods in clinical conditions. Specifically, Tarassenko et al. [33] estimated HR, SpO2, and respiratory rate (RR) of the patients undergoing dialysis, Bal [60] monitored the health status of patients in pediatric intensive care units by estimating HR and SpO2, and Yu et al. [20] estimated HR and HRV for Geriatric patients undergoing physiotherapy treatment. Furthermore, 4/8 (55.55%) [23,28,44,48] studies have estimated heart rate during fitness exercises, while one study by Wu et al. [24] used a non-contact approach for HR monitoring of drivers. This proves the effectiveness of non-contact methods in real-time applications, where it is infeasible to use conventional or gold standard physiological parameters estimation techniques.

On the other hand, few studies such as [55] and [47] have also estimated parameters such as step count and eye blink. Step count can be used to monitor fitness, whereas eye blink can be used to analyze sleep quality.

3.9. Studies quality assessment results

This study have identified seven vital parameters for HR estimation studies, namely: camera characteristics (camera resolution and shooting distances), Bland-Altman analysis, results score performance metrics (RMSE and correlation), artifacts, accuracy ($error < \pm 5bpm$), number of subjects used for the study, and inclusion/exclusion of ethnicity. On the other hand, SpO_2 estimation studies quality was assessed using the following four parameters: camera characteristics, number of subjects, inclusion/exclusion of



(a) HR (b) SpO_2

Fig. 6. Study Categorization Results.

Bland-Altman analysis, and the coefficient of determination (R^2). Other parameters similar to HR studies such as artifacts, accuracy, RMSE, correlation, and ethnicity were not reported in most studies, and hence not included in this analysis.

The details of the studies and their analysis based on these parameters are presented in a supplementary file (Tables 1, 2, and 3). The studies were categorized into three categories: strong, fair, and weak, as depicted in Fig. 6. Based on the proposed protocol, 5 HR studies were identified as "strong", reporting maximum identified parameters, while the number of "fair" and "weak" studies were found to be 30 and 11, respectively. On the other hand, 3 SpO_2 studies are categorized as weak, 2 studies as fair and 3 studies are categorized as strong.

4. Discussion

This review is intended to search and summarize the currently existing facial video-based non-contact methods for estimating two widely used physiological parameters, HR and SpO_2, respectively. Monitoring these parameters under real-time environments such as clinical conditions, driving, or fitness exercise will only be feasible using non-contact approaches since it allows higher degrees of freedom, unlike contact based approaches. The analysis and comparison of multi-factors opted for diverse studies would enable researchers to wisely select the important parameters for designing the study and assist them in quantifying their respective studies based on the distribution of various error metrics, correlation, and accuracy using this review.

4.1. Context of evidence and limitations

The non-contact estimation approaches summarized in this review are at the proof of concept stage with a few shortcomings. These include relatively constrained video acquisition settings, smaller sample sizes, and a limited clinical context. The reference devices used for most of the studies are pulse oximeters, while some have also used ECG. While other studies have used other devices, whose accuracy can be questionable compared with a standard HR monitoring device (ECG or PPG). Moreover, the selection of a valid reference device plays a key role in assessing the applicability of the proposed method in comparison to it. Additionally, a valid reference device could also address the limitations of the ECG or PPG in interpreting the results.

Most studies have used lower camera resolution, making it a cost-effective solution for real-time monitoring such as driving, fitness exercises, and clinical monitoring. However, selecting the appropriate video resolution is challenging and also affected by the distance between the camera and the subject's face. A study conducted by Song et al. [53] attempted to find the optimal resolution and camera shooting distance. It was concluded that higher resolution enhances the quality of the PPG signal whereas a distance of more than 1 m will deteriorate the HR estimations, which is consistent with the findings of this review. However, high-resolution cameras are computationally intensive for estimating physiological parameters, and camera distance of 1 m or less limits the applicability of non-contact approaches for clinical or sleep settings. Additionally, the frame rate plays a key role in tracking tiny variations present in the image sequences, which ensures an accurate PPG signal. A few attempts have been made with higher sampling rates, but the performance of HR estimation with a higher frame rate in comparison with 30 fps did not show significant performance improvement [67]. Furthermore, a frame rate of 30 fps worked well with most of the studies. Except for clinical and SpO_2 estimation studies, most studies have acquired data for about 1 min to 3 min, which may limit the legitimacy of the methods. Moreover, a shorter interval may hinder the robustness of the proposed method under different estimation conditions. The SpO_2 estimation studies have considered relatively longer duration subject's videos. The limitation of the prolonged video acquisition is the presence of artifacts due to movement and uneven illuminations. Besides, image quantization can also produce undesirable noise, but this effect can be mitigated by assuming constant light over the region of interest. The presence of these artifacts deteriorates the estimations of the physiological parameters, as the PPG signal (from the RGB color channel) is very weak and is difficult to extract from the artifacts' corrupted signals. An obvious approach to mitigate this problem would be to convert RGB to other movement or illumination artifact resistant color channels such as YUV [52], LAB [50], etc. For dark scenarios, an infrared channel could be a better alternative, but the only problem is that the strength of the PPG signal is relatively weaker than the signal from the RGB channel. A combination of RGB with IR, which is similar to the one conducted by Kado et al. [51] or other color models, may produce promising results but at the cost of increasing the problem's complexity and while also being computationally intensive.

A vast category of rPPG extraction methods has been used in the literature. Neural networks and their variants have been used extensively for HR estimation studies. The neural networks based methods performed relatively better than other conventional rPPG estimation methods. Additionally, there has been extensive use of transfer learning for HR estimation methods. Most importantly, neural networks do not need assumptions to process the data, which was the case with the existing state-of-the-art methods. On the other hand, SpO_2 studies employed regression using the ratio of ratios method with an exception [26], which used BVP signals to map to SpO_2 levels ranging from 65 to 100%.

8/45 (17.78%) HR estimation studies [2,20,25,31,34,56,64,68] have achieved clinically acceptable error differences as depicted in Fig. 3, whereas other studies might need significant improvements in the future. Furthermore, several studies have justified their method's clinical relevance by reporting the accuracy, which is calculated as the percentage of study samples having an error less than ± 5 bpm. However, it is worth noting that these studies have predominantly used their self-created databases under well-constrained laboratory conditions to test their method in the normal HR and SpO_2 ranges. The performance of these methods may deteriorate for abnormal HR parameter ranges. In addition, the accuracy in this scenario will not be sufficient to justify their clinical relevance and therefore needs further analysis. On the other hand, the performance of SpO_2 estimation under extreme conditions is difficult to test since it needs multiple breath holding events, which is not always possible for individuals. Consequently, developing a robust SpO_2 estimation method proves to be difficult due to the need to measure the subtle changes in the saturated blood. Therefore, there are limited SpO_2 estimation studies in the literature. There is a need to devise methods for estimating SpO_2 values from a single PPG signal extracted from the facial ROI, similar to other physiological parameters.

Another limitation within almost all studies is that the parameters were estimated for healthy individuals, limiting the estimation methods' ability for diseased people with conditions such as hypoxemia, bradycardia, or tachycardia.

Finally, this study found some common factors applicable for all non-contact approaches to estimate both physiological parameters. Ethnicity, movement, illumination, and clinical relevance depicted by accuracy are the common factors used by all studies. Although none of the SpO_2 estimation studies have reported movement and illumination artifacts, but these factors are worth considering while developing non-contact approaches for physiological parameter estimations.

4.2. Limitations of the review process

There are certain limitations and challenges associated with synthesizing this review. The search strategy aimed to summarize and analyze novel methodologies for HR and SpO_2 estimations using only facial videos. This excludes the estimation studies from other body parts, which may limit our findings, and analysis to face region. Being a review, we excluded patents and commercial applications in these areas. This review did not include unpublished studies or conference papers (except for a few discussed in" eligibility criteria"). Therefore, this review may have publication bias. Furthermore, reporting and lead-time bias may be possible since databases were used to collect research articles using our search strategy. In addition, it was difficult to compare the studies with missing information such as an insufficient description of population, highly diversified error metrics. These factors might affect the assessment of the "risk of bias" among studies. Additionally, some studies have presented their results using visual representation; it is challenging to extract numerical values from them for analysis. Overall, there is a high level of heterogeneity among studies, which was difficult to tackle for studies quality assessment.

4.3. Future research and recommendations

This review aims to provide insight into this rapidly growing domain of developing non-contact approaches for face-based physiological parameters estimations. We have identified seven key factors for HR estimations and four for SpO_2 estimations, which should be addressed while designing an estimation study. The data collection is a crucial step, which might affect the efficacy of the proposed method. The study population should be properly described for better representation and result analysis. For instance, information such as age, gender, video resolution, and camera shooting distance should be provided to compare two or more studies. Furthermore, the selection of valid reference devices needs to be done for ground truth data collection. Traditional HR estimation is a complicated process and requires certain conventional assumptions with limited generalizability, while neural networks based methods are less dependent on these assumptions with good generalizing ability for highly diverse study samples [41]. Hence, the applicability of neural network based methods is recommended for physiological parameters estimation. Comparing estimated values with ground truth should be done using appropriate performance metrics. Although it is challenging to have a defined set of metrics for performance analysis of the proposed methods, reporting the following metrics: RMSE, correlation, accuracy, and Bland-Altman plots is highly recommended. Furthermore, for SpO_2, all studies used the ratio of ratios method considering two color channels; an alternative approach utilizing cleaner BVP/PPG signals like [26] could be beneficial in ensuring accurate estimations. Designing a SpO_2 estimation study is challenging, especially under hypoxemic events or severely infected Covid patients, since it is difficult to collect the data aligned with these conditions. Moreover, the ratio of ratios method have certain limitations in tracking minute variations from saturated blood. Finally, all studies have used a shorter time frame for physiological parameters estimation in healthy individuals. Future studies should focus on methods, which takes less time to estimate parameters using longer video sequences, handling challenges like motion and illumination artifacts, and performing under different ethnic groups for real-time scenarios. Furthermore, future estimation studies should also focus on estimating under the abnormal parameter ranges or during a cardiopulmonary or related diseased condition.

Registration and Protocol

This systematic review is not registered, and the protocol is not registered.

Declaration of Competing Interest

None of the authors declared any competing interests.

Acknowledgments

This work is supported by the LARSyS Project-UIDB/50009/2020, co-financed by Regional Development European Funds for the "Operational Programme Madeira 14-20"—Priority Axis 1 of the Autonomous Region of Madeira, number M1420-01-0145-FEDER-000002 and M1420-01-0247-FEDER-000033 -MTL-Maritimo Training lab project.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.cmpb.2022.106771.

References

[1] W.Q. Mok, W. Wang, S.Y. Liaw, Vital signs monitoring to detect patient deterioration: an integrative literature review, Int. J. Nurs. Pract. 21 (2015) 91–98.

[2] M.-Z. Poh, D.J. McDuff, R.W. Picard, Non-contact, automated cardiac pulse measurements using video imaging and blind source separation, Opt. Express 18 (10) (2010) 10762–10774.

[3] I. Pavlidis, J. Dowdall, N. Sun, C. Puri, J. Fei, M. Garbey, Interacting with human physiology, Comput. Vision Image Understanding 108 (1–2) (2007) 150–170.
[4] W. Verkruysse, L.O. Svaasand, J.S. Nelson, Remote plethysmographic imaging using ambient light, Opt. Express 16 (26) (2008) 21434–21445.
[5] K.M. van der Kooij, M. Naber, An open-source remote heart rate imaging method with practical apparatus and algorithms, Behav. Res. Methods 51 (5) (2019) 2106–2119.
[6] W. Verkruysse, M. Bartula, E. Bresch, M. Rocque, M. Meftah, I. Kirenko, Calibration of contactless pulse oximetry, Anesth. Analg. 124 (1) (2017) 136.
[7] B. Jafari, V. Mohsenin, Polysomnography, Clin. Chest Med. 31 (2) (2010) 287–297.
[8] A.d.F.G. Rosa, R.C. Betini, Noncontact SpO2 measurement using Eulerian video magnification, IEEE Trans. Instrum. Meas. 69 (5) (2019) 2120–2130.
[9] A. Nicolò, C. Massaroni, E. Schena, M. Sacchetti, The importance of respiratory rate monitoring: from healthcare to sport and exercise, Sensors 20 (21) (2020) 6396.
[10] M.A. Hassan, A.S. Malik, D. Fofi, N. Saad, B. Karasfi, Y.S. Ali, F. Meriaudeau, Heart rate estimation using facial video: a review, Biomed. Signal Process. Control 38 (2017) 346–360.
[11] C. Wang, T. Pun, G. Chanel, A comparative survey of methods for remote heart rate detection from frontal face videos, Front. Bioeng. Biotechnol. 6 (2018) 33.
[12] Y. Sun, N. Thakor, Photoplethysmography revisited: from contact to noncontact, from point to imaging, IEEE Trans. Biomed. Eng. 63 (3) (2015) 463–477.
[13] J. Kranjec, S. Beguš, G. Geršak, J. Drnovšek, Non-contact heart rate and heart rate variability measurements: a review, Biomed. Signal Process. Control 13 (2014) 102–112.
[14] M. Harford, J. Catherall, S. Gerry, J.D. Young, P. Watkinson, Availability and performance of image-based, non-contact methods of monitoring heart rate, blood pressure, respiratory rate, and oxygen saturation: a systematic review, Physiol. Meas. 40 (6) (2019) 06TR01.
[15] M.J. Page, J.E. McKenzie, P.M. Bossuyt, I. Boutron, T.C. Hoffmann, C.D. Mulrow, L. Shamseer, J.M. Tetzlaff, E.A. Akl, S.E. Brennan, The PRISMA 2020 statement: an updated guideline for reporting systematic reviews, BMJ 372 (2021).
[16] W. Verkruysse, L.O. Svaasand, J.S. Nelson, Remote plethysmographic imaging using ambient light, Opt. Express 16 (26) (2008) 21434–21445.
[17] Y. Zhang, Z. Chen, H.I. Hee, Noninvasive measurement of heart rate and respiratory rate for perioperative infants, J. Lightwave Technol. 37 (11) (2019) 2807–2814.
[18] J. Cheng, X. Chen, L. Xu, Z.J. Wang, Illumination variation-resistant video-based heart rate measurement using joint blind source separation and ensemble empirical mode decomposition, IEEE J. Biomed. Health Inform. 21 (5) (2016) 1422–1433.
[19] J. Ryu, S. Hong, S. Liang, S. Pak, Q. Chen, S. Yan, A measurement of illumination variation-resistant noncontact heart rate based on the combination of singular spectrum analysis and sub-band method, Comput. Methods Programs Biomed. 200 (2021) 105824, doi:10.1016/j.cmpb.2020.105824.
[20] X. Yu, T. Laurentius, C. Bollheimer, S. Leonhardt, C.H. Antink, Noncontact monitoring of heart rate and heart rate variability in geriatric patients using photoplethysmography imaging, IEEE J. Biomed. Health Inform. 25 (5) (2021) 1781–1792, doi:10.1109/JBHI.2020.3018394.
[21] H. Qi, Z. Guo, X. Chen, Z. Shen, Z. Jane Wang, Video-based human heart rate measurement using joint blind source separation, Biomed. Signal Process. Control 31 (2017) 309–320, doi:10.1016/j.bspc.2016.08.020.
[22] W. Wang, A.C. den Brinker, S. Stuijk, G. de Haan, Robust heart rate from fitness videos, Physiol. Meas. 38 (6) (2017) 1023.
[23] W. Wang, A.C. den Brinker, S. Stuijk, G. De Haan, Algorithmic principles of remote PPG, IEEE Trans. Biomed. Eng. 64 (7) (2016) 1479–1491.
[24] B.F. Wu, Y.W. Chu, P.W. Huang, M.L. Chung, Neural network based luminance variation resistant remote-photoplethysmography for driver's heart rate monitoring, IEEE Access 7 (2019) 57210–57225, doi:10.1109/ACCESS.2019.2913664.
[25] F. Bousefsaf, C. Maaoui, A. Pruski, Continuous wavelet filtering on webcam photoplethysmographic signals to remotely assess the instantaneous heart rate, Biomed. Signal Process. Control 8 (6) (2013) 568–574.
[26] M. Van Gastel, S. Stuijk, G. De Haan, New principle for measuring arterial blood oxygenation, enabling motion-robust remote monitoring, Sci. Rep. 6 (1) (2016) 1–16.
[27] A.R. Guazzi, M. Villarroel, J. Jorge, J. Daly, M.C. Frise, P.A. Robbins, L. Tarassenko, Non-contact measurement of oxygen saturation with an RGB camera, Biomed. Opt. Express 6 (9) (2015) 3320–3338.
[28] G.d. Haan, V. Jeanne, Robust pulse rate from chrominance-based rPPG, IEEE Trans. Biomed. Eng. 60 (10) (2013) 2878–2886, doi:10.1109/TBME.2013.2266196.
[29] Q. Tran, S. Su, W. Sun, M. Tran, Adaptive pulsatile plane for robust noncontact heart rate monitoring, IEEE Trans. Syst. Man Cybern. (2019) 1–13, doi:10.1109/TSMC.2019.2957159.
[30] O. Gupta, D. McDuff, R. Raskar, Real-time physiological measurement and visualization using a synchronized multi-camera system, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2016, pp. 46–53.
[31] M. Kumar, A. Veeraraghavan, A. Sabharwal, DistancePPG: robust non-contact vital signs monitoring using a camera, Biomed. Opt. Express 6 (5) (2015) 1565–1588.
[32] M.-Z. Poh, D.J. McDuff, R.W. Picard, Advancements in noncontact, multiparameter physiological measurements using a webcam, IEEE Trans. Biomed. Eng. 58 (1) (2010) 7–11.
[33] L. Tarassenko, M. Villarroel, A. Guazzi, J. Jorge, D. Clifton, C. Pugh, Non-contact video-based vital sign monitoring using ambient light and auto-regressive models, Physiol. Meas. 35 (5) (2014) 807.
[34] B. Wei, X. He, C. Zhang, X. Wu, Non-contact, synchronous dynamic measurement of respiratory rate and heart rate based on dual sensitive regions, Biomed. Eng. Online 16 (1) (2017) 1–21.
[35] W. Chen, D. McDuff, DeepPhys: video-based physiological measurement using convolutional attention networks, in: Proceedings of the European Conference on Computer Vision (ECCV), 2018, pp. 349–365.
[36] Y. Qiu, Y. Liu, J. Arteaga-Falconi, H. Dong, A.E. Saddik, EVM-CNN: real-time contactless heart rate estimation from facial video, IEEE Trans. Multimedia 21 (7) (2019) 1778–1787, doi:10.1109/TMM.2018.2883866.
[37] F. Bousefsaf, A. Pruski, C. Maaoui, 3D convolutional neural networks for remote pulse rate measurement and mapping from facial video, Appl. Sci. 9 (20) (2019) 4364.
[38] Z. Yu, W. Peng, X. Li, X. Hong, G. Zhao, Remote heart rate measurement from highly compressed facial videos: an end-to-end deep learning solution with video enhancement, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2020, pp. 151–160.
[39] G.-S.J. Hsu, R.-C. Xie, A. Ambikapathi, K.-J. Chou, A deep learning framework for heart rate estimation from facial videos, Neurocomputing 417 (2020) 155–166.
[40] Z. Yu, X. Li, X. Niu, J. Shi, G. Zhao, AutoHR: a strong end-to-end baseline for remote heart rate measurement with neural searching, IEEE Signal Process. Lett. 27 (2020) 1245–1249, doi:10.1109/LSP.2020.3007086.
[41] R. Song, S. Zhang, C. Li, Y. Zhang, J. Cheng, X. Chen, Heart rate estimation from facial videos using a spatiotemporal representation with convolutional neural networks, IEEE Trans. Instrum. Meas. 69 (10) (2020) 7411–7421, doi:10.1109/TIM.2020.2984168.
[42] R. Macwan, Y. Benezeth, A. Mansouri, Heart rate estimation using remote photoplethysmography with multi-objective optimization, Biomed. Signal Process. Control 49 (2019) 24–33, doi:10.1016/j.bspc.2018.10.012.
[43] J. Cheng, P. Wang, R. Song, Y. Liu, C. Li, Y. Liu, X. Chen, Remote heart rate measurement from near-infrared videos based on joint blind source separation with delay-coordinate transformation, IEEE Trans. Instrum. Meas. 70 (2021) 1–13, doi:10.1109/TIM.2020.3041083.
[44] M. Hu, F. Qian, X. Wang, L. He, D. Guo, F. Ren, Robust heart rate estimation with spatial-temporal attention network from facial videos, IEEE Trans. Cognit. Dev.Syst. (2021), doi:10.1109/TCDS.2021.3062370.
[45] M. Hu, F. Qian, D. Guo, X. Wang, L. He, F. Ren, ETA-rPPGNet: effective time-domain attention network for remote heart rate measurement, IEEE Trans. Instrum. Meas. 70 (2021) 1–12, doi:10.1109/TIM.2021.3058983.
[46] A. Moço, W. Verkruysse, Pulse oximetry based on photoplethysmography imaging with red and green light, J. Clin. Monit. Comput. 35 (1) (2021) 123–133.
[47] Y. Lin, Y. Lin, Step count and pulse rate detection based on the contactless image measurement method, IEEE Trans. Multimedia 20 (8) (2018) 2223–2231, doi:10.1109/TMM.2018.2790172.
[48] C. Zhao, C.-L. Lin, W. Chen, M.-K. Chen, J. Wang, Visual heart rate estimation and negative feedback control for fitness exercise, Biomed. Signal Process. Control 56 (2020) 101680, doi:10.1016/j.bspc.2019.101680.
[49] X. Li, J. Chen, G. Zhao, M. Pietikainen, Remote heart rate measurement from face videos under realistic situations, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 4264–4271.
[50] Y. Zhang, Z. Dong, K. Zhang, S. Shu, F. Lu, J. Chen, Illumination variation-resistant video-based heart rate monitoring using lab color space, Opt. Lasers Eng. 136 (2021) 106328, doi:10.1016/j.optlaseng.2020.106328.
[51] S. Kado, Y. Monno, K. Yoshizaki, M. Tanaka, M. Okutomi, Spatial-spectral-temporal fusion for remote heart rate estimation, IEEE Sens. J. 20 (19) (2020) 11688–11697, doi:10.1109/JSEN.2020.2997785.
[52] X. Niu, S. Shan, H. Han, X. Chen, RhythmNet: end-to-end heart rate estimation from face via spatial-temporal representation, IEEE Trans. Image Process. 29 (2020) 2409–2423, doi:10.1109/TIP.2019.2947204.
[53] R. Song, S. Zhang, J. Cheng, C. Li, X. Chen, New insights on super-high resolution for video-based heart rate estimation with a semi-blind source separation method, Comput. Biol. Med. 116 (2020) 103535, doi:10.1016/j.compbiomed.2019.103535.
[54] G.R. Tsouri, S. Kyal, S. Dianat, L.K. Mestha, Constrained independent component analysis approach to nonobtrusive pulse rate measurements, J. Biomed. Opt. 17 (7) (2012) 077011, doi:10.1117/1.jbo.17.7.077011.
[55] C. Zhang, X. Wu, L. Zhang, X. He, Z. Lv, Simultaneous detection of blink and heart rate using multi-channel ICA from smart phone videos, Biomed. Signal Process. Control 33 (2017) 189–200, doi:10.1016/j.bspc.2016.11.022.
[56] H. Yue, X. Li, K. Cai, H. Chen, S. Liang, T. Wang, W. Huang, Non-contact heart rate detection by combining empirical mode decomposition and permutation entropy under non-cooperative face shake, Neurocomputing 392 (2020) 142–152.
[57] D. Chen, J. Wang, K. Lin, H. Chang, H. Wu, Y. Chen, S. Lee, Image sensor-based heart rate evaluation from face reflectance using Hilbert-Huang transform, IEEE Sens. J. 15 (1) (2015) 618–627, doi:10.1109/JSEN.2014.2347397.
[58] R. Song, J. Li, M. Wang, J. Cheng, C. Li, X. Chen, Remote photoplethysmography with an EEMD-MCCA method robust against spatially uneven illuminations, IEEE Sens. J. (2021).
[59] K. Lin, D. Chen, W. Tsai, Face-based heart rate signal decomposition and evaluation using multiple linear regression, IEEE Sens. J. 16 (5) (2016) 1351–1360, doi:10.1109/JSEN.2015.2500032.

[60] U. Bal, Non-contact estimation of heart rate and oxygen saturation using ambient light, Biomed. Opt. Express 6 (1) (2015) 86–97, doi:10.1364/boe.6.000086.
[61] A. Woyczyk, V. Fleischhauer, S. Zaunseder, Adaptive Gaussian mixture model driven level set segmentation for remote pulse rate detection, IEEE J. Biomed. Health Inform. 25 (5) (2021) 1361–1372, doi:10.1109/JBHI.2021.3054779.
[62] G. De Haan, A. Van Leest, Improved motion robustness of remote-PPG by using the blood volume pulse signature, Physiol. Meas. 35 (9) (2014) 1913.
[63] D. Shao, C. Liu, F. Tsow, Y. Yang, Z. Du, R. Iriya, H. Yu, N. Tao, Noncontact monitoring of blood oxygen saturation using camera and dual-wavelength imaging system, IEEE Trans. Biomed. Eng. 63 (6) (2016) 1091–1098, doi:10.1109/TBME.2015.2481896.
[64] L. Kong, Y. Zhao, L. Dong, Y. Jian, X. Jin, B. Li, Y. Feng, M. Liu, X. Liu, H. Wu, Non-contact detection of oxygen saturation based on visible light imaging device using ambient light, Opt. Express 21 (15) (2013) 17464–17471.
[65] P. Gupta, B. Bhowmick, A. Pal, MOMBAT: heart rate monitoring from face video using pulse modeling and Bayesian tracking, Comput. Biol. Med. 121 (2020) 103813, doi:10.1016/j.compbiomed.2020.103813.
[66] , Association for the Advancement of Medical Instrumentation. Cardiac monitors, heart rate meters, and alarms, in: American National Standard (ANSI/AAMI EC13: 2002) Arlington, VA, 2002, pp. 1–87.
[67] E.B. Blackford, J.R. Estepp, Effects of frame rate and image resolution on pulse rate measured using multiple camera imaging photoplethysmography, in: Medical Imaging 2015: Biomedical Applications in Molecular, Structural, and Functional Imaging, vol. 9417, International Society for Optics and Photonics, 2015, p. 94172D.
[68] J. John, S. Krishna, R.R. Galigekere, Automatic and adaptive signal- and background-ROIs with analytic-representation-based processing for robust webcam-based heart-rate estimation, IEEE Access 8 (2020) 34728–34736, doi:10.1109/ACCESS.2020.2973815.