A survey of blood oxygen saturation assessment from video

Alexis Wuyart, Laure Abensur Vuillaume, Choubeila Maaoui, Frédéric Bousefsaf

Biomedical Signal Processing and Control 110 (2025) 108069

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

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

A survey of blood oxygen saturation assessment from video

Alexis Wuyart [a], Laure Abensur Vuillaume [b], Choubeila Maaoui [a], Frédéric Bousefsaf [a,*]

[a] Université de Lorraine, LCOMS, F-57000 Metz, France
[b] CHR Metz-Thionville, Emergency department, Metz, France

ARTICLE INFO

Keywords:
Non-contact
Imaging photoplethysmography
Blood oxygenation
Pulse oximetry
Deep learning
Video-health monitoring

ABSTRACT

The combination of technical advances and the emergence of artificial intelligence in the medical field has given rise to new oxygen saturation (SpO_2) measurement processes. In recent years, camera-based estimation of SpO_2 has attracted increasing interest. This technique offers an optical way to acquire SpO_2 in a very practical way. This review provides an overview of the imaging photoplethysmography-based techniques developed to achieve SpO_2 prediction from video. A total of 81 articles have been assessed and are included in this review. The results obtained using artificial intelligence-based methods in particular are very promising in view of traditional methods. This suggests that AI-based methods will have an important future role to play in medical devices. The prospects are now focused on efforts to improve the robustness and performance of the models, and in fine, to stamp these medical devices in accordance with current standards (CE or FDA) to meet the demands of the medical profession and patient safety.

Contents

1. Introduction

Remote measurement of physiological signals from facial video has made significant progress the last past years [1]. Imaging (or remote) photoplethysmography (iPPG), the underlying principle, consists in measuring the subtle fluctuations of skin color that reflect complex light-tissue interactions. All types of camera, from simple webcams

* Corresponding author.
E-mail address: frederic.bousefsaf@univ-lorraine.fr (F. Bousefsaf).

https://doi.org/10.1016/j.bspc.2025.108069
Received 17 July 2024; Received in revised form 4 November 2024; Accepted 5 May 2025
Available online 23 May 2025

to professional cameras, can be employed to recover iPPG signals and estimate cardiovascular parameters. The face remains the most frequently observed area but different regions of interest have been studied over time [2]. Several studies demonstrated that pulse rate and its variability can be robustly and precisely estimated with conventional image processing techniques and, more recently, with deep learning solutions [3,4]. Current research in this field is now directed towards the measurement of new physiological parameters such as oxygen saturation (SpO_2) and blood pressure [5,6]. In the era of ubiquitous computing where mobile devices like smartphones and tablets are omnipresent, cameras correspond to sensors that are already available and, thus, that are particularly interesting for unobtrusively measuring vital signs.

Oxygen saturation is one of the five most important vital physiological parameters. It reflects the amount of oxygenated hemoglobin in blood. This is due to the fact that haemoglobin binds with free oxygen to form oxyhaemoglobin, leaving a relatively low percentage of free oxygen in the plasma. This difference is expressed in an oxygen-haemoglobin dissociation curve and a plot of the percentage saturation of haemoglobin as a function of the partial pressure of oxygen (PO_2). For a PO_2 of 100 mmHg, haemoglobin will be 100% saturated. It has been established that the oxyhaemoglobin level represents the oxygen content of the blood [7]. Measurement of oxygen saturation levels is part of the clinical examination of any acute situation or chronic lung or cardiac disease [8]. From a clinical standpoint, a drop in arterial oxygen saturation (SaO_2) can only be seen at 67%, when cyanosis is developing; the human eye is therefore not very discerning [9]. Measuring SaO_2 is therefore highly relevant for assessing hypoxaemia. The use of a pulse oximeter is therefore indicated whenever hypoxaemia is likely to occur, i.e. in almost any situation where a patient's state of health is of concern [9]. It is necessarily important for hypoxia/hypoxemia detection and is continuously monitored in many medical situations [10]. The commonly used method for the measurement of SpO_2 is pulse oximetry, a noninvasive technique that relies on a convenient and inexpensive contact device. The most common approach corresponds to transmissive oximetry, which is based on photoplethysmography. The device is here placed on a specific site, usually the fingertip or the earlobe, and emits two wavelengths of light through the body part. A photodetector measures the changing absorbance of each light. This allows determination of the absorbance of the pulsing arterial blood at each of the wavelengths [11]. Contact probes may have limitations in certain areas of application, requiring specialized installation and monitoring by experts [12]. They cannot be employed in situations involving various medical conditions, e.g. in the case of skin ulcer or burns [13]. In addition, it requires disinfecting medical devices with products that can deteriorate the equipment and thus contributes to the increase in maintenance costs [14]. They contribute to the reduction of transmissible diseases by offering non-contact measurements that reduce hand-carried contamination. The use of remote devices also offers the possibility of enhanced ambulatory monitoring in the context of telemedicine or self-monitoring, which has become more widespread in recent years in the context of Covid-19 and the general use of smartphones in the population. Non-contact monitoring tools are developing increasingly, mainly driven by the wellness industry [15]. The first proofs of concept for non-contact SpO_2 measurement were described in 2017 [16]. To date, only a few authors have examined the performance of these devices in relation to the performance expected in clinical practice, with a comparison made using a pulse oximeter [15]. Attempts to assess SpO_2 from video have been made but efforts must be undertaken to improve the reliability of the measurements. Conventional techniques that rely on image and signal processing as well as deep learning approaches have been proposed by the research community. In line with the exponential development of telemedicine, these advances will serve both the medical community and self-monitoring patients.

This article reviews the methods dedicated to SpO_2 assessment from video. We collected international journals and conference papers published between 1992 and september 2024. More than a hundred articles have been assessed. A total of 100 articles are included in this review, 43 being in direct relation with the main topic of this survey (remote blood oxygenation estimation from video). Because of the low volume of scientific articles dedicated to the remote SpO_2 measurement from video, we decided to not consider advanced automatic search strategies (e.g. PRISMA [17]).

We included three additional sections to the article. Section 2 presents the physiological and technical background of pulse oximetry as well as the inherent limitations. Section 3 reviews the different techniques devoted to SpO_2 estimation from video streams. We present a summary of the contributions in Section 4 along with potential developments.

2. Background

2.1. Photoplethysmography and imaging photoplethysmography

Photoplethysmography (PPG) corresponds to the indirect observation of blood volume variations by measuring absorption and reflection of light on skin tissues [18]. The incident light is absorbed by the skin, the bones and by the venous and arterial blood (Fig. 1, left illustration). Modifications in the pulsatile blood flow arise in the arteries and arterioles. These fluctuations in volume are therefore periodic and produced at each heartbeat: the volume of blood increases during systole (cardiac contraction) and decreases during diastole (cardiac relaxation). Fingertip oximeters are based on this principle.

Measuring PPG follows two distinct modes, transmission, and reflection (Fig. 1, right illustration). In transmission, the incident light emitted by a LED passes through the skin, blood, and bones before being detected by a photodetector (PD). In reflection, the LED and the PD are placed next to each other, the PD measuring the light reflected from the tissues.

Contact probes are conventionally employed to record physiological signals. They can however be limited in some scopes of application where a specialist must install and monitor them [12]. They cannot be employed in the case of trauma, skin ulcer, burns, congenital and contagious diseases [13]. Physiological signal measurement using contactless devices has gained vast attention [21,22]. Significant advancements have been reached over the last few years [23]. They demonstrated that standard video cameras are reliable devices that can be employed to measure a large set of biomedical parameters without any contact with the subject [2]. In the era of ubiquitous computing where smartphones are omnipresent, cameras are already available and particularly interesting for unobtrusively measuring vital signs.

Several biomedical parameters can be computed from PPG signals [1,2]: pulse rate and its variability [24], oxygen saturation, breathing rate, pulse transit time and blood pressure [25], peripheral vasomotor activity [26,27] and vascular occlusion. Imaging PPG have also been used to identify living skin in images [28,29]. A large ensemble of applications has been covered [30]: mixed reality [31], newborn health monitoring [32], physiological measurements of drivers [33], sleep monitoring [34], face anti-spoofing [35] and affective computing [1]. Standard (visual spectrum), near infrared and multispectral cameras have been employed to sense iPPG signals [1]. The availability of many databases has enabled the rapid development of iPPG techniques based on artificial intelligence and deep learning [3,4,30].

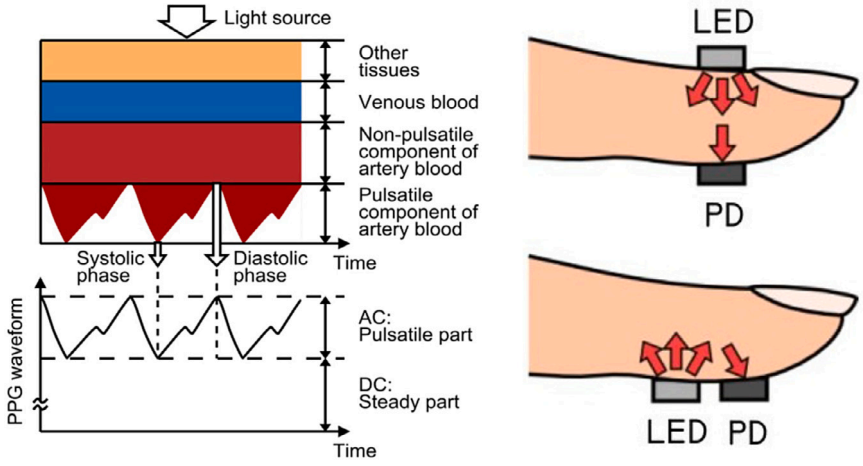

Fig. 1. Left: photoplethysmography and light attenuation in tissues and blood. Illustration taken from [19]. Right: PPG measurement modes, transmission (top figure) and reflection (bottom figure). PD stands for photodetector.
Source: Illustration modified from [20].

2.2. Contact oximetry

Regular assessments of oxygen saturation are of great importance in the monitoring of patients in intensive care units. Contact pulse oximeters have been widely used in routine and critical clinical applications since the development of pulse oximetry in the early 1980s [36]. Pulse oximeters are generally clipped to fingers or earlobes and take advantage of the fact that oxyhemoglobin and reduced hemoglobin absorb light differently at different wavelengths.

The principle of pulse oximetry is relatively simple and relies on PPG signals measured at two different wavelengths. Two LEDs emit light beams usually calibrated in red (660 nm) and near infrared (940 nm) [36,37]. The absorption of light by deoxygenated hemoglobin (Hb) and oxygenated hemoglobin (HbO_2) is proportional to their concentration (Fig. 2).

The estimation of the ratio between relative amplitudes (i.e. AC component over DC component) of PPG signals measured at these two different wavelengths provides a quantity called ratio-of-ratios that relates to oxygen saturation. An empirically derived calibration curve is employed to translate the ratio to SpO_2 [38,39]. Empirical linear approximations are commonly formulated as:

$$SpO_2 = \alpha + \beta.R \tag{1}$$

Or else,

$$SpO_2 = \alpha - \beta.R \tag{2}$$

α and β here are constants extracted from the calibration curve obtained from a large data set. Consequently, each pulse oximeter manufacturer will have their own calibration curves and thus have their own constants.

Despite the efficiency of this technology, the devices must still be worn generally on the fingers or on the earlobes which leads to poor compliance and risks of infections and skin irritation, which can cause irritation and discomfort. For instance, the risk of not monitoring SpO_2 for very sensitive preterm infants is sometimes preferred to certain traumas caused by contact sensors. This is a serious problem, for which remote monitoring has recently been attempted.

2.3. Limitations of pulse oxymetry

One should also be aware of the limitations of the pulse oximeter. Many pitfalls have been identified in the scientific literature [40–42]. These limitations may initially be physiological. The haemoglobin dissociation curve is characterized by a sigmoid, so that high baseline partial pressure of oxygen (SpO_2) levels offer little variation in SpO_2, so the oximeter will have difficulty detecting potential hypoxaemia in patients with high SpO_2 levels. Hypoperfusion can also affect signal quality. It is a major problem in sepsis because oxygenation levels are difficult to assess [43]. Oxygenation ranges play an important role as they help to constitute the clinical assessment of the patient [40]. Jubran et al. indicated that numerous pathologies, such as carbon monoxide poisoning, cardiac arrhythmias, oedema or arteriovenous fistulas, have been identified as potentially affecting signal quality [42]. Substances have also been identified as interfering with the signal. Due to the fact that the oximeter only uses two wavelengths specifically to discriminate between HbO_2 and Hb, high levels of methaemoglobin and carboxyhaemoglobin can induce an inaccurate SpO_2 reading. Jubran et al. has also shown that the usual intravenous dyes (methylene blue, indigo carmine...) have the property of leveling down the SpO_2 levels measured up to 20 min after injection [42]. It is interesting to note that this author found an effect of skin tone on the quality of the observed signal. Dark pigmentation in sick individuals seems to bias the SpO_2 measurement more frequently. Other researchers, such as Sinex et al. suggested that darker skin tones result in more difficult light penetration, ultimately leading to overestimated SpO_2 values [41]. However, conflicting results in the literature have indicated that dark pigmentation does not cause an abnormal SpO_2 [40]. Consensus on this issue does not seem to be established and further investigation seems advisable.

Secondly, the limitations can be correlated to the measurement environment. Although corrected by oximeters, ambient light, through surgical lamps or bright lights, could bias the measurement results. Dirt on the sensor, mechanical interference (catheter, tourniquet...), electrical frequencies, cold temperatures that can cause shivering and tremors can all degrade the quality of the signal. Nail polish (if black, green or blue) is also recognized in the literature as interfering with the signal [40]. Artifacts stemming from noise and motion are arguably the most significant challenges, notwithstanding the fact that certain authors have developed methods to enhance tolerance towards them [44]. This is highlighted by the study investigated by Lutter et al. [45], where a significant number of pulse oximeter alarms were erroneous due to motion artifacts.

Fortunately, many filters are available to mitigate these artifacts. Among them, the Fast Fourier Transform (FFT), a technique used to decompose a non-periodic signal in the time domain into its different frequency components, is generally used [39]. Compressed Fast Fourier Series Approximation (CFSA) is used to analyze periodic signals by

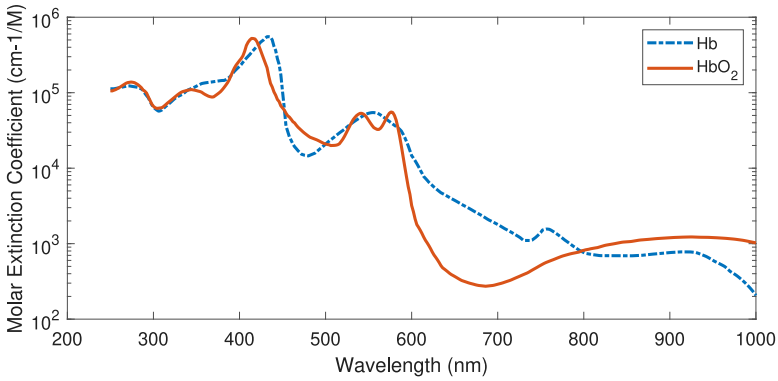

Fig. 2. Spectral absorption of deoxygenated hemoglobin (Hb) and oxygenated hemoglobin (HbO_2).

Table 1
Pulse Rate (PR) estimation through various filtering methods extracted from [50].

PPG signal	PR estimation	PR estimation error (%)
Clean PPG signal	61.357	–
Artifact PPG signal	78.188	27
ICA	73.561	19.9
Wavelet transform	69.307	12.9
ANC	66.338	8.13
CFSA	61.466	**0.19**

decomposing them into a sum of sinusoidal functions. Singular Value Decomposition (SVD) method can also be applied. This is based on decomposing the PPG signal matrix into singular values in order to eliminate the values corresponding to the artifacts [46]. Adaptive filters such as the Adaptive Noise Cancellation (ANC) method has been employed to clean up the PPG signal [47]. Similar to the SVD method, Independent Component Analysis (ICA), a blind source separation technique, can be employed to clean up signals. The main difference with the SVD method is that the ICA method can discriminate, separate, nonlinear signals [48]. Finally, the Wavelet Transform (WT), a technique relating to the time-frequency domain, is regularly found in the literature, particularly for the analysis of non-stationary signals [49]. As shown in Table 1, researchers compared these methods and concluded that the CFSA and SVD methods were the most efficient in reducing artifacts and restoring the signal [50].

That is why recorded PPG signals are processed by conventional filtering techniques, i.e. low-pass and high-pass filtering to remove trends high-frequency noise [19]. In addition, it is difficult to use this technology in contact for continuous monitoring despite its effectiveness. The devices must be worn on the fingers or earlobes, which can lead to the risk of infection and skin irritation. Oximeters should be cleaned after each use for hygienic reasons and to avoid potential contamination. Remote monitoring via cameras is therefore very promising in this context. Also, simultaneous monitoring of the physiological signals of several people by a single camera is possible [51], which would allow a large volume of patients to be monitored in real time.

However, a pitfall persists; few remote medical devices are currently CE or FDA stamped [52]. In Europe, for example, a CE mark must be affixed by the manufacturer before the product can be marketed. The requirements for obtaining this mark depend on the class of the medical device [53]. Medical devices are divided into four classes: class I, class IIa, class IIb and class III, depending on their level of risk. The level of risk depends in particular on the therapeutic or diagnostic purpose, the duration of use and the invasive nature of the device. Each research protocol must receive the favorable opinion of an ethics committee. These procedures are lengthy and require specific expertise. A non-contact device for measuring SpO_2 could be categorized as class IIa. This is a real obstacle for video measurements because a medical device requires a rigorous clinical evaluation with regard to numerous standards. Efforts must be made in this area to ensure the proper development of remote medical devices.

3. SpO_2 measurement from video

Existing works in the field of contactless blood oxygenation estimation can be categorized in 2 groups: conventional and deep learning methods. We first introduce the public datasets used in this field of research.

3.1. Public databases

To the best of our knowledge, only three datasets have been employed in the field of SpO_2 measurement from video: VIPL-HR [54], PURE [55], and UBFC-rPPG [29].

VIPL-HR [54] has originally been proposed for remote pulse rate estimation but SpO_2 readings were also recorded during the data collection. It can therefore be used for benchmarking contactless SpO_2 measurement methods. The full dataset contains 2451 RGB and 752 NIR facial videos of 107 subjects (79 males and 28 females) recorded by four acquisition devices (a web camera, a smartphone camera, a RGB-D camera and a NIR camera). Each video lasts around 30 s, with a frame rate of around 30 frames per second. The dataset has been employed in [56–58].

PURE dataset [55], contains 60 one-minute videos along with pulse signals recorded with a finger oximeter. The videos streams have not been compressed. The ground truth values of SpO_2 have been recorded with a finger clip pulse oximeter (pulox CMS50E). PURE has been employed in [58].

UBFC-RPPG [29] contains 43 uncompressed videos along with oximeter signals recorded by a finger clip sensor (Contec Medical CMS50E). Each video is about 2 min long and recorded with a webcam (Logitech C920 HD pro). The participants played a time-sensitive mathematical game that supposedly raises their pulse rate. The dataset has been used in the context of remote SpO_2 estimation by [56].

3.2. Conventional techniques

We will provide in a first section, a review of all the conventional methods proposed by the research community. They include adaptative PBV and ratio of ratios-based methods.

Inspired by contact pulse oximetry methods and to take advantage of the deeper wave penetrability in near infrared, the first camera-based methods [59] used a relatively complex acquisition system with multiple LED arrays in NIR synchronized with a camera. Researchers then proposed to use ambient light [60] and even cameras in the visible spectral range [61]. Conventional color cameras [62,63] as well as cameras with narrowband filters mounted on them [60] have been used to capture PPG signals and compute oxygen level.

3.2.1. Ratio of ratios method

Humphrey et al. [59] were the first to introduce a method for remote blood oxygenation assessment, based on the ratio of ratios conventional method. The components of the imaging system are a monochromatic camera, and a NIR bandpass glass filter. The light source used for the experiment is composed of a matrix of 36 LEDs, 18 LEDs with a peak wavelength of 760 nm and 18 LEDs with a peak wavelength of 880 nm. A control circuit energizes one type of LEDs at a type, such that the light source produces pulsed light centered at 760 nm and 880 nm with equal duration. An indicator light is visible in each frame, indicating if it was taken during 760 nm or 880 nm lighting. Using this lighting system, the experiment was conducted on 10 healthy subjects, and the forearm was chosen as a region of interest. After processing, a PPG signal is obtained for each lighting type. The ratio of ratios method has then been applied to estimate SpO_2 considering the 760 nm lighting as the red signal, and the 880 nm as the IR signal. Guazzi et al. [61] presented in 2015 a new method to track SpO_2 changes using only an RGB camera. This method carefully selects the regions of interest (ROI) by calculating the signal to noise ratios for each ROI, and was tested on subjects who had oxygen saturations comprised between 80 et 100%.

Recent studies show that measuring SpO_2 with standard visible (VIS) cameras is feasible [61,64]. Conventional methods that study the measurement of SpO_2 from VIS video employ a pipeline-based strategy where the relative PPG amplitudes at two wavelengths are first estimated and then scaled using calibration constants. These hand-crafted signal processing pipelines make it challenging to capture the complexity of the phenomenon and the inclusion of other influencing factors (e.g. lighting condition or skin color). Results obtained so far show that there is room for improvement [65]. Moreover, limited validation (e.g. small number of subjects and small variation of SpO_2 values [66]) and use of specific or NIR hardware [67] must be pointed out.

All the camera-based methods use two specific wavelength bands but substantial progress has been made in contact pulse oximetry using more than two bands [68–70] to make the measurement more reliable. Since the seminal work of Humphreys et al. [59], progress has been made in measuring physiological parameters by video analysis [37]. They all mostly employ a pipeline-based strategy where the relative PPG amplitudes, obtained via classical remote PPG techniques at two wavelengths are first estimated and then scaled using calibration coefficients [71,72]. Scully et al. [62] and Casalino et al. [10] estimate SpO_2 by comparing the red and blue bands. The blue band is here a representative of the infrared band used in traditional pulse oximetry. Takahashi et al. proposed to automatically find optimal bands using a regression model fit on dedicated features before computing the ratio of ratios using these selected bands [73].

Sun et al. [74] based their work on this last observation - the main challenge being the restriction to wavelengths in the visible range. They used a smartphone (iPhone X from Apple Inc.). After illuminating the ROI (back of the hand) with the smartphone's flash, the video sequence was processed to capture physiological data at 60 fps via the RGB camera for a few seconds. The reflectance images recorded on the green and red channels are chosen to calculate the ratio of the ratios. The pipeline of PPG signal extraction is shown in Fig. 3.

Similarly, the continuous components were calculated by passing time-traced signals through a low-pass filter at a cut-off frequency of 0.3 Hz. Once the pulsatile and continuous components have been acquired, the ratio of ratios, calculated as follows, can be deduced.

$$R = \frac{\frac{AC_R}{DC_R}}{\frac{AC_G}{DC_G}} \tag{3}$$

According to the non-linear variation of pulsatile blood volume, the SpO_2 is actually the function of R and ΔR [75]:

$$SpO_2 = R + \Delta R \tag{4}$$

Where ΔR is the deviation of the ratio of ratios caused by the change in light scattering. Consequently, ΔR can be represented mainly by the variation of the gray value recorded by the RGB sensors, which can be used to improve the estimation of SpO_2. The estimated overall error between the smartphone and the pulse oximeter in predicting SpO_2 via the conventional ratio method was $\pm 2.008\%$ [74]. Although encouraging for this innovative method, this value can be improved.

Under controlled conditions, SpO_2 can be calibrated with red and green light, but the accuracy is lower than the usual red-NIR window [76]. Indeed, the errors for red and green increase for lower temperatures or when the measurement is made on a different skin site (cheek instead of forehead for example), due to probing the vascular system at different depths. Other problems with the use of an RGB camera are the addition of noise (relatively low signal-to-noise ratio) or measurement bias such as dependence on spectral content of ambient lighting, reflectance of the spectrum modifying the perceived SpO_2, unequal penetration depths of the wavelengths in the skin or even spectral sensitivities of the RGB channels that depend on the camera [77].

Rather than using the standard deviations of the AC components for the calculations, Wei et al. [64] performed a blind source separation of the AC components and then used coefficients retained in the mixed matrix to replace the required variables (energy distribution of the source signals) in the algorithm. In this way, stable data were selected to compensate for potential defects in the combination of RGB signals. The denoising power of the algorithm was later significantly improved [64]. The experimental results indicated that the proposed method produced a reliable estimate of SpO_2 that could potentially be used in real applications. In a similar direction, Al-Naji et al. [78] proposed to decompose red and green PPG signals computed from video into a collection of signals at different frequencies using the Ensemble Empirical Mode Decomposition followed by an Independent Component Analysis. The component that presented the highest signal-to-noise ratio is employed to compute SpO_2 using a conventional calibration model. The system includes a preprocessing step (i.e. skin detection) and has been validated with 14 patients of different ages and with different skin tones. The results exhibit good agreement with those computed from a pulse oximeter. Finally, although highly conceptual and not very pragmatic, researchers have theorized multi-wavelength pulse oximeters [68,79]. They have established that multiple wavelengths improved accuracy and solved the problems of motion artifacts and errors using reflection oximetry.

Lan et al. [70] introduced the dynamic spectrum (DS) method. This is a signal processing technique that allows information to be extracted from pulsatile signals used to measure oxygen saturation. The DS method entails taking the Fourier transform of the PPG signal after applying an wavelet transform to remove low-frequency noise. The resulting spectrum is then depicted as a function of time and wavelength. At each wavelength, the signal's harmonic amplitudes are measured, resulting in an assortment of data known as the dynamic spectrum. The pulse signal is then dissociated from the low-frequency noise and the measurement of the signal's fluctuations at various wavelengths is effected. This method enables the elimination of environmental interferences, individual differences, and static fabric interferences.

Standard signal and image pre-processing techniques are based on the spatial averaging of raw PPG signals (such as red, green, blue and in the near-infrared) of the pixels in defined ROIs [76,80]. The signals are then frequently normalized using a technique known as AC/DC normalization, which divides each signal by its low-frequency component to make the measurements stable to changes in external illumination or motion. Filters such as the Butterworth filter can be applied to these signals [64,76]. More advanced pre-processing methods have been used by Vogels et al. [34]. The two steps of the preprocessing method proposed in the study are Gaussian smoothing and temporal rigid blocks. Gaussian smoothing is used for the input images to mitigate the effects of sharp edges in the images, in particular the boundaries between skin and non-skin areas. To identify regions of interest in

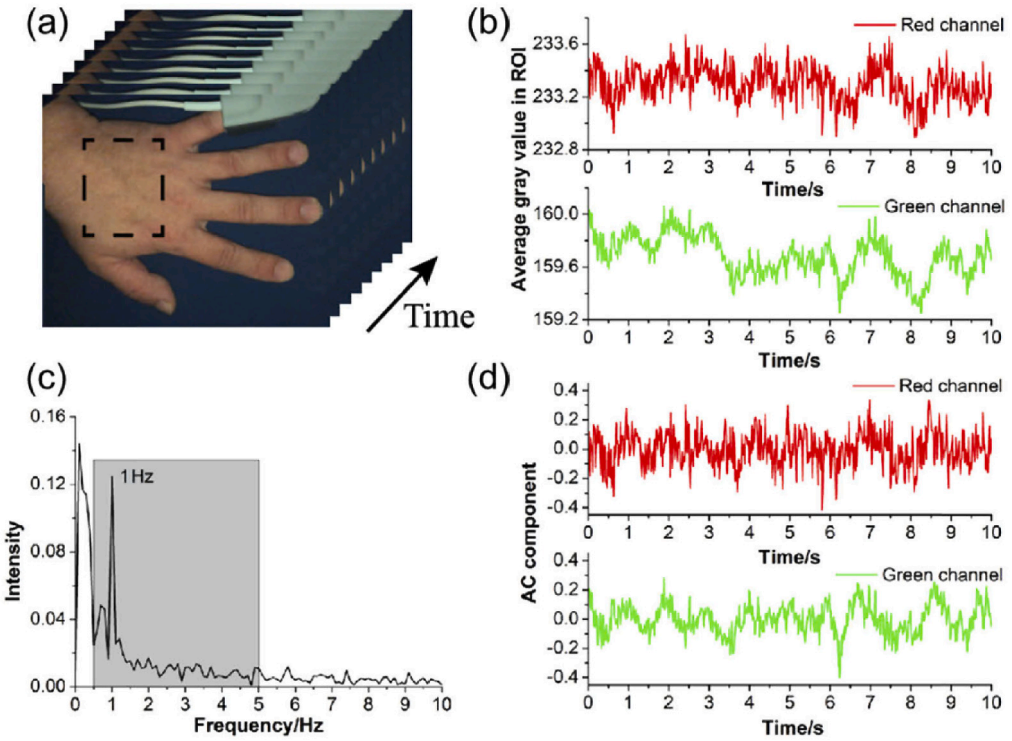

Fig. 3. Process of extracting PPG signals from a video sequence recorded at 60 FPS. (a) A video sequence is captured with the ROI highlighted. (b) PPG signals are obtained by averaging the gray levels for each channel. (c) An FFT is used to switch from the time domain to the frequency domain, highlighting the FC at 1 Hz. A bandpass filter (0.5 to 5 Hz) is used to extract the pulsatile components (AC) from the time series. (d) Pulsatile components of the red and green channels.
Source: Illustration extracted from [74].

an image, rigid temporal segmentation divides each image into fixed size blocks, each representing a 29 × 29 pixel value. This subdivision reduces camera measurement error and sensor noise. The values from each image are then combined to create a 3D data cube which is used to extract the pulse signal. The time window used in this study for pulse signal extraction is 10 s, with an overlap of 8 s. Overall, Gaussian smoothing reduces noise while temporal rigid blocks correct unintentional camera movements and improve the accuracy of motion detection.

Polarization-Dependent Optical Sectioning uses the polarization of light to create a three-dimensional image of the skin microvascularization [81]. This simply involves shining a beam of polarized light onto the skin and measuring the polarization of the reflected light. In a different degree, Rosa et al. [82] used the Eulerian Video Magnification technique. This method takes a typical video sequence and applies spatial decomposition followed by temporal filtering of the images. The resulting signal is then amplified to reveal hidden information. This technique allows us to visualize the blood flow that fills the face and to amplify small movements.

These pre-processing techniques are essential to reduce side effects, improve overall data quality and model's performance.

3.2.2. Adaptative PBV method

Van Gastel et al. [83] proposed a method that is not directly based on the ratio-of-ratios. To introduce this method, the Pulse Blood Volume (PBV) method should be explained first [84]. As the pulsatility amplitudes are relatively small, they will be subject to measurement errors during movement-induced variations of the subject. The ratio method is therefore quite sensitive to this. To overcome this limitation, de Haan et al. [84] proposed the PBV method, which is based on the unique and intrinsic signature of the blood volume pulse signal. This innovative method uses the various pulsatile amplitudes relative to the RGB channels. This allows for the discrimination of intensity variations caused by changes in blood volume from dissimilar variations. A Pulse Blood Volume vector, $\overrightarrow{P}_{\mathrm{bv}}$, represents the direction in RGB color space of PPG signal variations. This vector can be calculated as follows:

$$\overrightarrow{P}_{bv} = \frac{[\sigma(\overrightarrow{R}_n), \sigma(\overrightarrow{G}_n), \sigma(\overrightarrow{B}_n)]}{\sqrt{\sigma^2(\overrightarrow{R}_n) + \sigma^2(\overrightarrow{G}_n) + \sigma^2(\overrightarrow{B}_n)}} \tag{5}$$

The signals $\overrightarrow{R}_{\mathrm{n}}$, $\overrightarrow{G}_{\mathrm{n}}$, and $\overrightarrow{B}_{\mathrm{n}}$ represent the three PPG signals corresponding to the red, green, and blue components respectively. This guide vector will be used as a parameter for combining the RGB signals as shown in Fig. 4. It will be adjusted according to the input color signals in order to obtain the optimal combination vector. This $\overrightarrow{P}_{\mathrm{bv}}$ vector is static, its value does not change over time or for different people under identical recording conditions. This vector is used to calculate a weight matrix W, which transforms the three PPG signals into a single, much less noisy pulse signal such as:

$$\overrightarrow{W} = k.\overrightarrow{P}_{bv}.Q^{-1} \tag{6}$$

Considering k as a fixed scalar to normalize $\overrightarrow{W}$, such that the Euclidean norm of the vector $\overrightarrow{W}$ is equal to one.

$$Q = C_n.C_n^T \tag{7}$$

C_{n} is a matrix of size 3*N containing the three RGB signals, where N is the number of samples. $\overrightarrow{S}$, the cleaned pulse signal, can be computed using $\overrightarrow{W}$ as follows:

$$\overrightarrow{S} = \overrightarrow{W}.C_n \tag{8}$$

The PBV process is particularly effective in removing distortions in PPG signals, as shown in Fig. 5. It is now possible to take advantage of this method to determine SpO_2 indirectly while avoiding potential artifacts; this is what the APBV method is all about. The method consists of a training and model building phase based on the

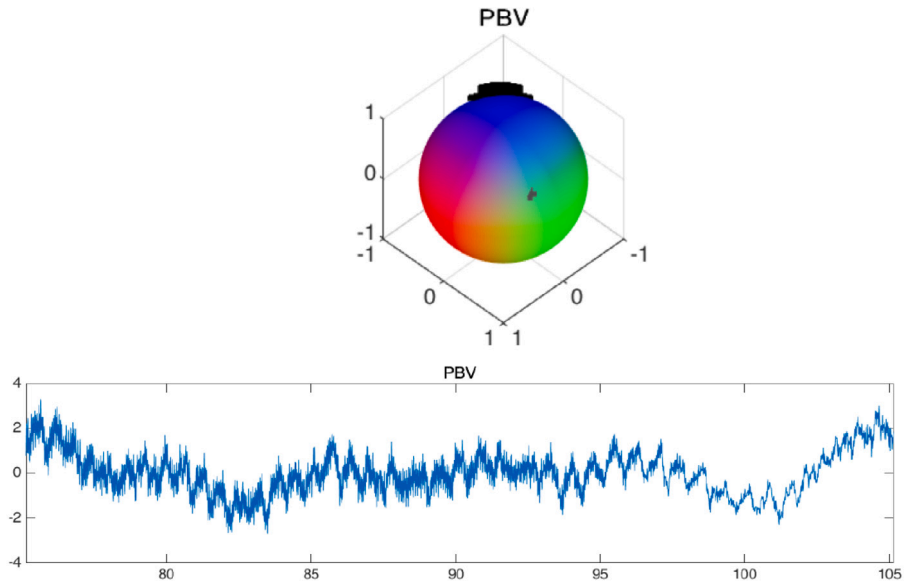

Fig. 4. RGB sphere representing the vector at the output of PBV ([0.33 0.78 0.53]) materialized by the gray cross and the corresponding maximized signal. *Source:* Illustration extracted from [85].

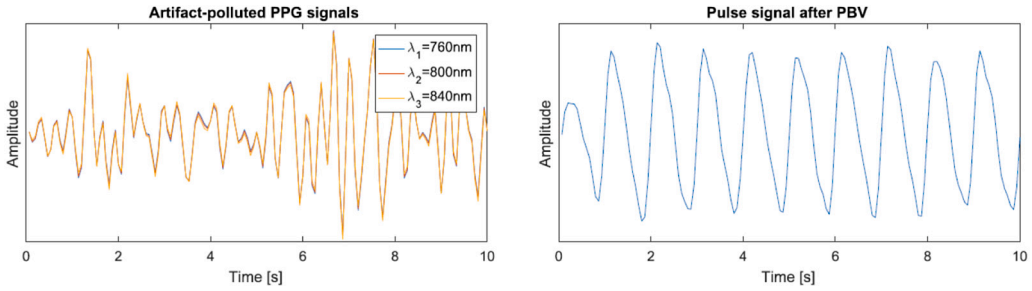

Fig. 5. Ability of the PBV method to remove artifacts in a PPG signal. Respectively a PPG signal heavily polluted by artifacts and the resulting pulse signal. *Source:* Illustration modified from [83].

calculation of a collection of Signal-to-Noise Ratios (SNR) for different SpO_2 ranges. The estimation of an unknown SpO_2 from a PPG signal is done by observing the SNR response. This is made possible because when a "wrong" PBV is applied to a PPG signal, the quality of the resulting pulse signal decreases. By reverse analogy, the higher SNR most accurately captures the data and can therefore be related to the correct SpO_2 value. Fig. 6 illustrates how this works.

The APBV method can be summarized mathematically as follows [77,83]:

$$SpO_2 = \text{argmax} SNR\left(k.\vec{P}_{bv}\left(SpO_2\right)\left|C_n C_n^T\right|^{-1}.C_n\right) \quad (9)$$

Van et al. [83] represented this method via synthetic data with added noise, simulating linear de-saturation with oxygenation levels between 60 and 100% and a regular HR. This is the first method in a remote configuration to obtain satisfactory results in case of significant subject movement.

The APBV method has therefore been shown to outperform conventional ratio-based methods. However, the addition of wavelengths complicates the calibration. It was suggested to use a data-driven approach to increase the accuracy of the calibration. The results showed that performance improved over a calibration model based on opto-physiological modeling, with an average reduction in RMSE of 1.84%. Furthermore, it was shown that the use of near infrared wavelengths alone, as opposed to a combination of visible red and NIR wavelengths, which is typically used in pulse oximeters, only slightly increases the error by 0.29% [66].

Deep learning strategies can implicitly integrate the different hand-crafted steps by directly mapping the RGB signals to SpO_2 values. It is important to note that this strategy has recently been explored for this physiological parameter, probably due to the ability to obtain a public dataset even if this remains a major obstacle.

3.3. AI-based techniques

Artificial intelligence techniques are frequently used in different fields of science (agriculture [86], medicine [87,88] etc.). Convolutional neural networks (CNN) have become a popular method for estimating pulse rate [89] or respiratory rate [90]. It is therefore natural that AI-based framework has been proposed by the research community to automatically estimate SpO_2 from video steams. Ding et al. [91] were the first to use a CNN to estimate SpO_2 from a video stream. Using data cleaned of motion artifacts by the SVD method (see Table 1) and filtered either via a Butterworth bandpass filter (noted BW, 0.7 to 4 Hz), and/or via a Savitsky-Golay filter (10 s, third order). The architecture proposed by Ding et al. [91] is presented in Fig. 7.

This convolutional neural network is composed of two temporal convolutional layers including successively a max pooling and dropout layer. Note that the dropout is a technique that aims mainly to prevent overlearning by randomly evicting neurons. This is followed by a flatten layer which converts the multidimensional feature maps into a one-dimensional vector. The network includes, at its end, a fully connected layer that allows classification or regression of the data - in this case

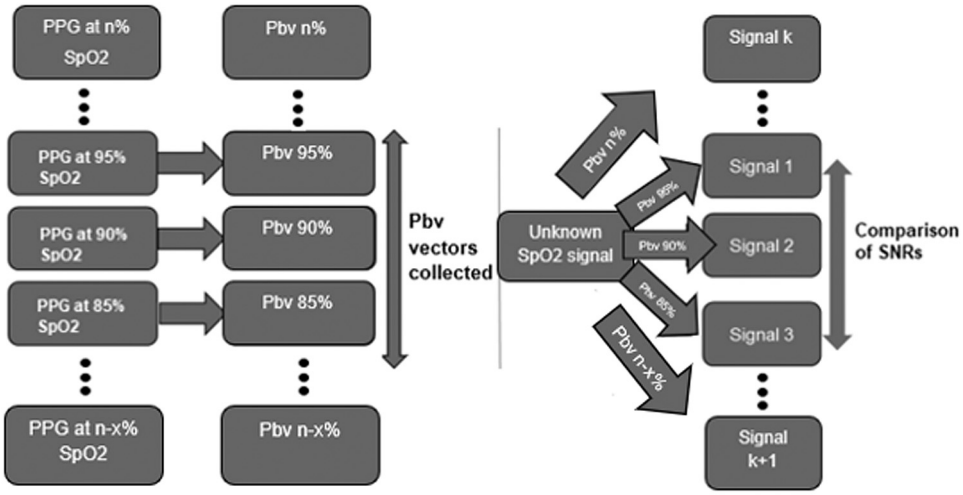

Fig. 6. Principle of APBV [83].

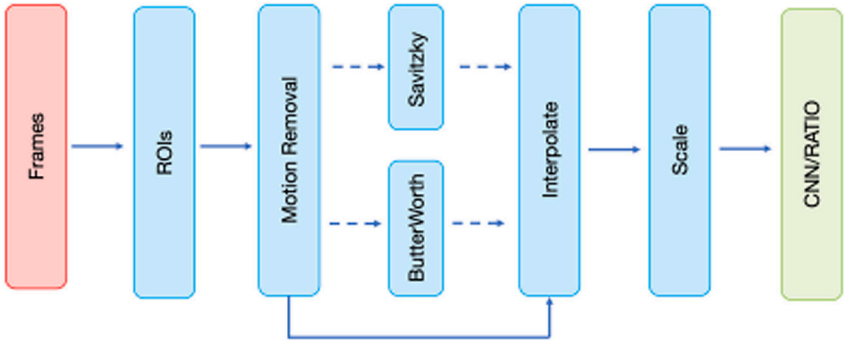

Fig. 7. Data processing flowchart presented in [91].

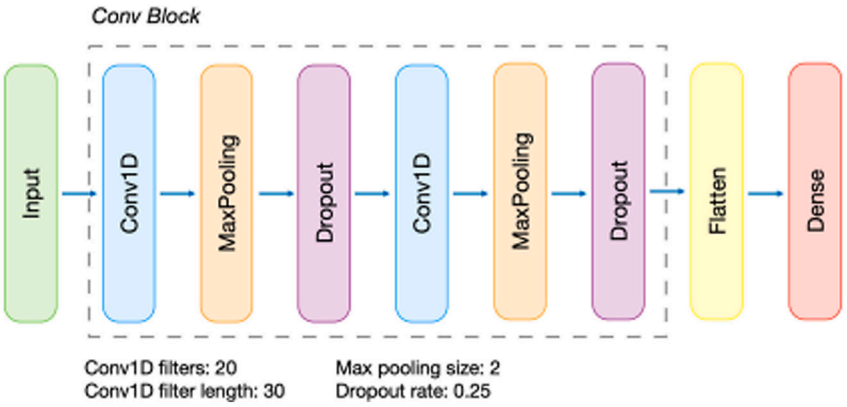

Fig. 8. Artificial neural architecture presented in [91].

a regression that determines SpO_2. The architecture of the CNN is schematically represented in Fig. 8.

The results (Table 2) are highly interesting and show firstly that all the processes tested - on the basis of root mean square error (RMSE) and mean absolute error (MAE) - achieved better accuracy than the traditional ratio-of-ratios method. The latter is based as a benchmark via pulse oximetry. All convolution approaches appear to perform similarly although the Raw PPG approach has better consistency as evidenced by the results. The strategies put in place to filter PPG signals (BW and Savitsky-Golay filters) did not perform as intuitively expected. Augmentation of the training data based on an oversampling technique also did not show satisfactory results. In fine, these results indicate that it is preferable to use raw PPG signals (although noise is removed from the input) to reduce computational complexity and let the CNN optimize an automatic feature extraction strategy.

Mathew et al. [92] have gone further and not only explored the prediction of SpO_2, but have tried to provide some answers as to the interpretability of the CNN model. They proposed three "physiologically motivated" neural network structures. The different models proposed differ in their arrangements for combining channels and extracting temporal features. The first model first combines the color channels, then the temporal features are extracted using convolutional and pooling layers. Finally, a single node is used to represent the predicted SpO_2 level. The second model proposes to reverse the order of channel

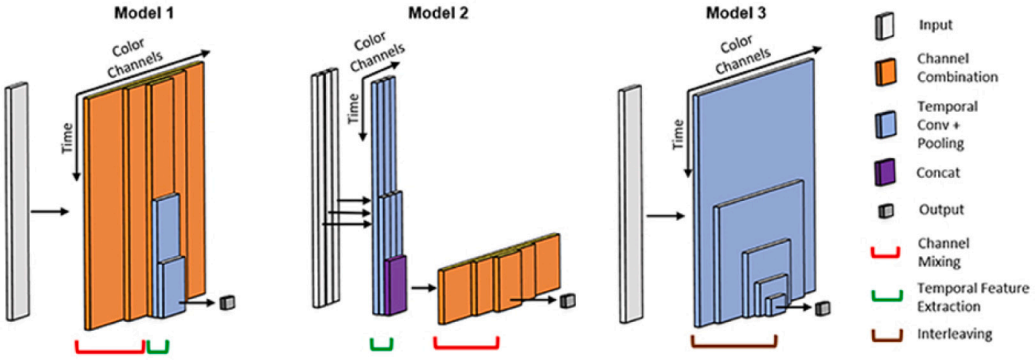

Fig. 9. Structures of CNNs used to predict SpO_2 level from iPPG signals. These models were proposed by Mathew et al. [92].

Table 2
Overall comparison of the different processes proposed in [91]. The augmentation corresponds to an oversampling procedure that artificially increases the number of training data.

Model	iPhone 6S		iPhone 7 Plus	
	RMSE	MAE	RMSE	MAE
Ratio-of-ratios	3.07	2.52	3.40	2.80
Conv (Raw PPG)	**2.54**	**2.02**	**3.01**	2.38
Conv (BW)	2.74	2.08	3.07	2.40
Conv (BW + Bias)	2.68	2.05	3.03	**2.33**
Conv (Raw PPG) + Augmentation	2.81	2.19	3.20	2.51

mixing and temporal feature extraction compared to the first model. The RGB channels are used separately for temporal feature extraction, allowing the convolutional layers to learn the different features for each channel. As with the first model, the color channels are then blended before the SpO_2 value is extracted. Finally, the third and last model allows to explore the possibility of mixing the color channel mixing step and the temporal feature extraction. Note that the number of filters is reduced at each successive convolutional layer, thus progressively reducing the number of combined channels. The structures of the CNNs used are shown in Fig. 9.

The results are presented as the median of the values obtained from all participants and are listed in Table 3. All of the established CNNs outperformed the model previously presented by Ding et al. [91], with significant improvement in MAE and RMSE. The third model seems to perform best on these indices, suggesting that it is *a priori* the best structure for individualized learning. It is also interesting to highlight that the second model is the most efficient in terms of correlation in the palm up and palm down recording situations.

The study proposed by Kedir et al. [93] can also be mentioned. The authors used a neural network called Conv3D for feature extraction, initially trained on RGB video sequences for human action recognition. They initialized their network from scratch using multispectral data to extract features. Multispectral imaging is a method for measuring different narrow spectral bands compared to RGB imaging. Then, temporarily, each 16-channel input is supplied to a Conv3D block. The feature maps are passed through a spatial and temporal attentional channel module after feature extraction from the multispectral data. The attentional channel module selects from a collection of feature maps provided by Conv3D those that are most relevant. Then, the most accurate feature maps are introduced into a bidirectional long-term memory (LSTM), which is a recurrent neural network architecture, that incorporates spatial and temporal information for the final prediction of the oxygenation curve.

Hu et al. [58] proposed a hybrid multi-model that takes advantage of the ratio of ratios principle. It comprises two modules (see Fig. 10). The first is named Residuals and Coordinate Attention (RCA in Fig. 10) and corresponds to a feature extraction network (a deep CNN) that extracts feature signals with physiological information from face videos. The first module is connected to a SpO_2 estimation module constituted by 2 sub-models: a color channel model that computes SpO_2 from blue and red PPG signal by the ratio of ratios equation (like in Eq. (3) but with blue wavelengths instead of green wavelengths) and a network-based model that predicts SpO_2 using a CNN. The color channel model is then combined with the network-based model using a dedicated loss function (SmoothL1Loss) to output the final SpO_2 estimation. The model has been tested on two datasets (VIPL-HR [54] and PURE [55]) and exhibits mean absolute errors $\leq 2\%$.

We can also mention the studies of Hamoud et al. [56] and Akamatsu et al. [94]. Hamoud et al. used a specific regressor method (XGBoost) to measure SpO_2 using the features extracted by a pre-trained CNN. The final models achieved good accuracy on VIPL-HR [54] and UBFC-rPPG [29] with a MAE of 1.17 and 0.84, respectively. Akamatsu et al. proposed to extract the DC and AC components from spatiotemporal maps using either conventional filtering techniques or convolutional layers before feeding a 18-layers ResNet [95] with these components (see Fig. 11). The validation has been conducted on a self-collected dataset recorded using a webcam (Logitech C920n HD PRO) and a pulse oximeter (OxyTrue A). 50 subjects participated, they were asked to hold their breath three time to display lower SpO_2. The full experiment lasts three minutes. Their results exhibit reasonable accuracy increases against conventional techniques (e.g. ratio of ratios).

Finally, Cheng et al. [57] proposed to input spatiotemporal maps computed from RGB cameras to convolutional neural networks to predict SpO_2. These maps have originally been proposed by Niu et al. [89] and are representations that embed the physiological information of the original RGB video. Several standard deep learning architectures (e.g. ResNet [95] and DenseNet [96]) have been tested on the VIPL-HR dataset [54]. The best performing model achieved 1.27% in mean absolute error and 1.71% in root mean squared error. These results outperform the conventional ratio of ratios method for contactless SpO_2 measurement. Because most of the subjects that participated in the VIPL-HR dataset are Asians with Fitzpatrick Scale [97] comprised between III and IV, the method proposed by the authors may be biased to people with these skin types and may not perform as well on darker skin tones (type V to VI). This is a common concern in remote vital sign monitoring [98].

The advances provided by convolutional neural networks are considerable, already performing to some extent as well as traditional techniques. There are many opportunities for optimization and there is every reason to believe that AI will play a major role in the measurement of remote physiological signals, especially SpO_2. One potential limitation could be the lack of training data for CNNs to train properly. As mentioned by Ding et al. [91], it is not at all clear that a CNN model is generalizable to all acquisition devices, which would exacerbate this limitation. We would then likely talk about a personalization strategy by smartphone type. It should be noted that the studies reported in this subsection were based on SpO_2 measurements in normal ranges

Table 3
Comparison of the performance of each structure of the proposed CNNs. PD and PU indicate that the palm of the hand was facing downwards and upwards, respectively, during the recording [92].

	Hand	Correlation			MAE [%]			RMSE [%]		
	mode	Train	Val	Test	Train	Val	Test	Train	Val	Test
Model 1	PD	0.86	0.75	0.41	1.90	1.52	2.21	2.26	1.94	2.51
[92]	PU	0.78	0.82	0.39	1.32	1.26	2.16	1.54	1.60	2.70
Model 2	PD	0.79	0.74	**0.46**	1.55	1.63	2.09	1.91	1.98	2.52
[92]	PU	0.86	0.77	**0.41**	1.08	1.52	1.96	1.34	1.70	2.48
Model 3	PD	0.81	0.77	0.44	1.64	1.27	**1.93**	1.99	1.59	**2.48**
[92]	PU	0.93	0.80	**0.41**	1.50	1.25	**1.81**	1.72	1.47	**2.43**
Ding et al.	PD	0.82	0.71	0.38	1.75	1.39	3.25	2.09	1.73	3.83
[91]	PU	0.83	0.72	0.34	1.60	1.26	3.40	1.93	1.58	4.58

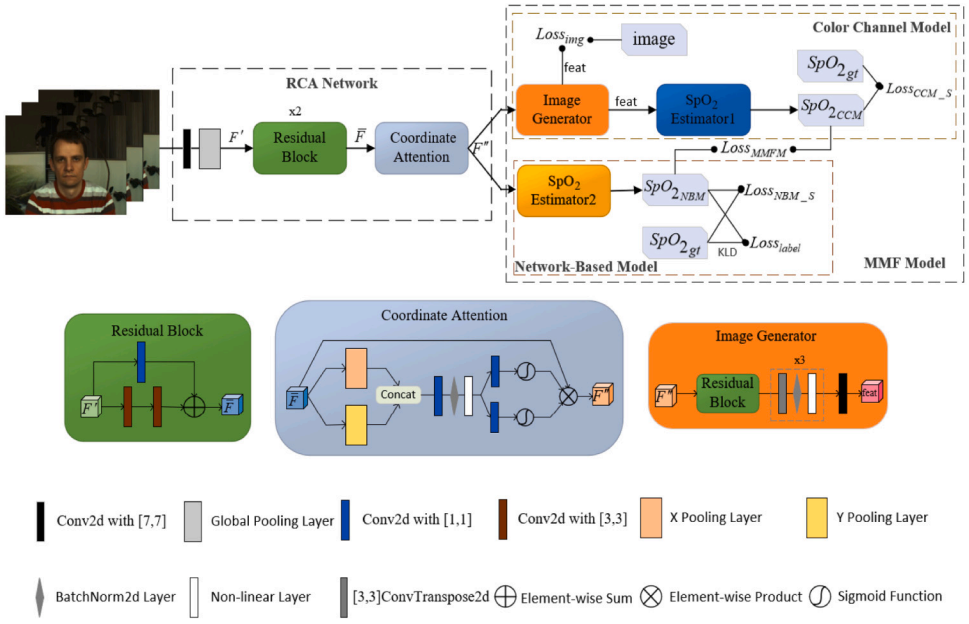

Fig. 10. Architecture proposed by Hu et al. [58] to estimate SpO_2 level from video.

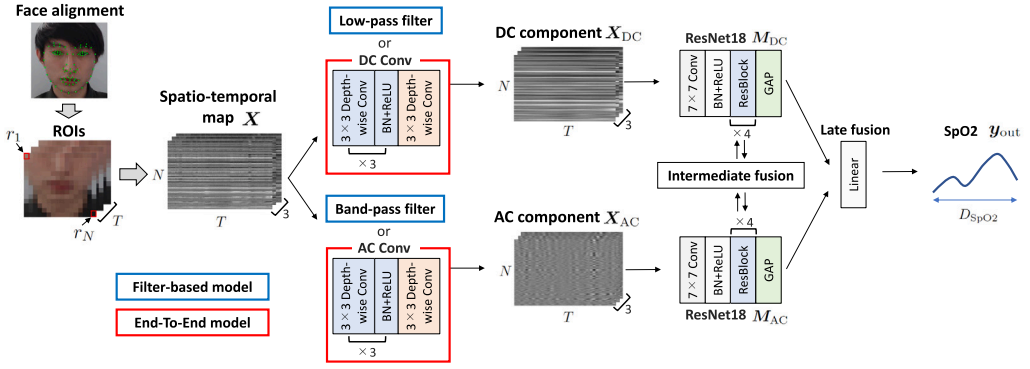

Fig. 11. Overview of the method proposed by Akamatsu et al. [94].

without de-saturation. The subjects were free of pathology and the measurements were taken in a non-emergency environment. In sum, the backgrounds were quite remote from the classical clinical framework in which the use of this artificial intelligence could be very efficient.

3.4. Overview of video-based SpO_2 techniques and outlook

Imaging photoplethysmography is rapidly gaining overall traction as a low-cost solution for remotely monitoring vital signs such as SpO_2. In comparison to contact-based devices, iPPG technology offers significant benefits during clinical trials, including improved patient mobility and avoidance of skin irritation. Additionally, the clinical setting offers a controlled environment where patients remain mostly still, and lighting remains constant.

Despite considerable advancements, the current methodologies encounter a few stumbling blocks highlighted in the previous sections and summarized in the Table 4. Conventional techniques often find themselves tripped up by tricky conditions like motion artifacts. Ratio of ratios method can stumble when faced with variations in skin tone and lighting. Even the adaptive PBV method can struggle to deal with rapid SpO_2 changes. AI-based approaches, while promising, face their

Table 4
Comparison of methods determining SpO_2 by video analysis.

Method	Strengths	Weaknesses
Ratio of ratios	- Simplicity and ease of implementation: the straightforward algorithmic approach allows for easy integration into various systems. - No calibration required: does not necessitate calibration with known SpO_2 values, simplifying deployment.	- Sensitivity to motion and light variability: performance can degrade significantly with motion or changes in ambient lighting conditions. - Limited accuracy: it is generally less accurate, especially in non-ideal conditions, compared to more sophisticated methods. - Noise sensitivity: performance may be adversely affected by noise in the video signal or suboptimal image quality.
APBV	- Robustness to artifacts: this method demonstrates strong performance in mitigating motion artifacts and adapting to physiological variations. - Advanced signal processing: it employs sophisticated techniques that greatly enhance signal accuracy and quality.	- Implementation complexity: more challenging to deploy compared to the ratio of ratios method, requiring more advanced algorithms. - Calibration requirement: may necessitate initial calibration or tuning for optimal performance in specific scenarios.
AI-based	- High accuracy and adaptability: machine learning models leverage extensive training data to achieve superior performance and adaptability across a wide range of conditions. - Continuous learning: these models can continuously improve over time with exposure to diverse datasets, enhancing their effectiveness in varied environments.	- Resource intensive: requiring substantial computational resources for both training and real-time processing, which can limit their deployment on devices with constrained capabilities. - Data dependency: performance also heavily depends on the quality and volume of the training data, which can be challenging to obtain.

own challenges with data demands and computational complexity. It is also worth noting that data collection in hospital environments presents distinct privacy challenges

Looking ahead, research efforts should focus on refining signal processing algorithms to efficiently sidestep motion artifacts, integrating machine learning models to adapt to individual nuances, enhancing sensitivity to SpO_2 variations, and emerging practical AI models for clinical use. The corollary of this observation is, moreover, partially reflected in specialized reviews from recent years [99,100].

4. Conclusion

4.1. Summary and general remarks

After a brief review of the concept of PPG and its relationship to the contact oximetry and the natural limitations that arise from it, we propose in this article a review of methods dedicated to the assessment of SpO_2 from video. These methods include traditional methods such as the ratio of ratios and adaptive PBV as well as AI-based techniques. The emergence of artificial intelligence in the medical field, particularly in SpO_2 prediction, indicates a substantial potential for surpassing standard models. It is worth mentioning that the current ISO standard for pulse oximeters (i.e. ISO 80601-2-61:2017) suggests a tolerable error of up to 3%; this is already the case for the remote SpO_2 estimates via convolutional neural networks presented so far.

4.2. Limitations

Although the use of contactless devices is currently limited to the general public, their use by healthcare professionals appears to be of interest on several fronts. On one hand, contactless measurement is an undeniable advantage in terms of health risks and disinfection. On the other hand, it is sometimes impossible to measure SpO_2 (like hypotension or Raynaud's syndrome). However, efforts are still needed to improve the performance of these techniques. The validation of a remote medical device through CE or FDA standards must be an absolute necessity. This will establish the trust that is required in a medical context for use in hospitals, doctors' practices and by individuals, especially in teleconsultation. In addition to the technical and developmental aspects, the very development of these mass-market tools can lead to questioning by clinicians. The algorithms are developed against a gold standard whose lack of reliability remains one of its main limitations, and not in a global clinical context.

Limited validation with small number of subjects and restricted variations of SpO_2 and blood pressure must particularly be pointed out. Color cameras, which are widely used, present a better sensitivity in the visible spectral bands. This, however, implies specific limitations because SpO_2 is conventionally determined by measuring the relative pulsatile amplitudes of arterial blood at different wavelengths. First, it is mainly the red wavelengths that are affected by changes in SpO_2 while the shorter wavelengths vary only slightly. Moreover and because of its shallow penetration depth, blue light is more impacted by specular reflections than longer wavelengths. Measurement of remote PPG signals from RGB video recordings are likely to be corrupted by these artifacts, thus compromising the reliability of SpO_2 estimates. Second, the temperature dependency of the relative PPG amplitudes is a factor that makes calibration for visible light measurements much more difficult. Handcrafted signal processing pipelines make it challenging to capture the complexity of the phenomenon and the inclusion of other influencing factors (e.g. lighting condition or skin color). While the availability of many databases has enabled the rapid development of remote PPG techniques based on deep learning, it is undoubtedly a barrier to camera-based pulse oximetry because, to date, no public dataset is commonly used in this field of research.

4.3. Medical applications

Oxygen saturation is one of the main physiological indices used to evaluate the medical condition of a person. An abnormal level of oxygen saturation is often associated with severe medical condition like hypoxia or chronic obstructive pulmonary disease [74].

Diagnosis of systemic scleroderma, changes in cerebral blood flow, migraines, and allergies on the skin surface are among the areas of medical application of SpO_2 estimation from video [37]. Digital video cameras have also been employed to detect oxygen de-saturation in healthy infants [69]. Reliable monitoring of heart rate and oxygen saturation is a critical component of newborn care. In neonatal intensive care units, the low birth-weight skin of newborns is particularly susceptible to trauma caused by the adhesives used to monitor them. Employing cameras for remote measurement of oxygen saturation would provide a less intrusive means to safely monitor infants without the risks associated with traditional monitoring devices.

Regarding the recent SARS-CoV-2 (COVID-19) pandemic, blood oxygen saturation is a relevant parameter to monitor as it may reflect compromised oxygen intakes [74], thus be alarming for a suspected infection of the coronavirus. In this scenario, oxygen saturation drops without the person being aware of it (silent hypoxia), which may lead to severe symptoms [73]. Automatic determination of the level of SpO_2 by non-contact means becomes particularly relevant here [37], in particular because the vulnerable population is recommended to monitor their oxygen status continuously for early COVID-19 detection [79].

CRediT authorship contribution statement

Alexis Wuyart: Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis. **Laure Abensur Vuillaume:** Writing – review & editing, Writing – original draft. **Choubeila Maaoui:** Supervision. **Frédéric Bousefsaf:** Writing – review & editing, Writing – original draft, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

References

[1] D. McDuff, Camera measurement of physiological vital signs, 2021, arXiv preprint arXiv:2111.11547.

[2] S. Zaunseder, A. Trumpp, D. Wedekind, H. Malberg, Cardiovascular assessment by imaging photoplethysmography–a review, Biomed. Tech. (Biomed. Engineering) (2018).

[3] A. Ni, A. Azarang, N. Kehtarnavaz, A review of deep learning-based contactless heart rate measurement methods, Sensors 21 (11) (2021) 3719, http://dx.doi.org/10.3390/s21113719, URL https://www.mdpi.com/1424-8220/21/11/3719.

[4] C.-H. Cheng, K.-L. Wong, J.-W. Chin, T.-T. Chan, R.H.Y. So, Deep learning methods for remote heart rate measurement: A review and future research agenda, Sensors 21 (18) (2021) 6296, http://dx.doi.org/10.3390/s21186296, URL https://www.mdpi.com/1424-8220/21/18/6296.

[5] L. Zhao, C. Liang, Y. Huang, G. Zhou, Y. Xiao, N. Ji, Y.-T. Zhang, N. Zhao, Emerging sensing and modeling technologies for wearable and cuffless blood pressure monitoring, Npj Digit. Med. 6 (1) (2023) 93.

[6] W. Chen, Z. Yi, L.J.R. Lim, R.Q.R. Lim, A. Zhang, Z. Qian, J. Huang, J. He, B. Liu, Deep learning and remote photoplethysmography powered advancements in contactless physiological measurement, Front. Bioeng. Biotechnol. 12 (2024) http://dx.doi.org/10.3389/fbioe.2024.1420100, URL https://www.frontiersin.org/journals/bioengineering-and-biotechnology/articles/10.3389/fbioe.2024.1420100.

[7] C.V. Preuss, A. Kalava, K.C. King, Prescription of Controlled Substances: Benefits and Risks, StatPearls Publishing, Treasure Island (FL), 2023, URL http://europepmc.org/books/NBK537318.

[8] P. Rees, F. Dudley, Oxygen therapy in chronic lung disease, Bmj 317 (7162) (1998) 871–874.

[9] C. Hanning, J. Alexander-Williams, Fortnightly review: pulse oximetry: a practical review, Bmj 311 (7001) (1995) 367–370.

[10] G. Casalino, G. Castellano, G. Zaza, A mhealth solution for contact-less self-monitoring of blood oxygen saturation, in: 2020 IEEE Symposium on Computers and Communications, ISCC, IEEE, 2020, pp. 1–7.

[11] F. Wieringa, Pulse oxigraphy: and other new perspectives through the near infrared window, ISBN: 978-90-5986-233-3, 2007, OCLC 150208347.

[12] 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.

[13] Y. Sun, N. Thakor, Photoplethysmography revisited: from contact to noncontact, from point to imaging, IEEE Trans. Biomed. Eng. 63 (3) (2016) 463–477.

[14] 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.

[15] A. Gupta, A.G. Ravelo-García, F.M. Dias, Availability and performance of face based non-contact methods for heart rate and oxygen saturation estimations: A systematic review, Comput. Methods Programs Biomed. (2022) 106771.

[16] W. Verkruysse, M. Bartula, E. Bresch, M. Rocque, M. Meftah, I. Kirenko, Calibration of contactless pulse oximetry, Anesth. Analg. 124 (1) (2017) 136.

[17] D. Moher, A. Liberati, J. Tetzlaff, D.G. Altman, t. PRISMA Group*, Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement, Ann. Intern. Med. 151 (4) (2009) 264–269.

[18] J. Allen, Photoplethysmography and its application in clinical physiological measurement, Physiol. Meas. 28 (3) (2007) R1–R39.

[19] T. Tamura, Current progress of photoplethysmography and SPO2 for health monitoring, Biomed. Eng. Lett. 9 (1) (2019) 21–36, http://dx.doi.org/10.1007/s13534-019-00097-w, URL http://link.springer.com/10.1007/s13534-019-00097-w.

[20] T. Tamura, Y. Maeda, M. Sekine, M. Yoshida, Wearable photoplethysmographic sensors—Past and present, Electronics 3 (2) (2014) 282–302, http://dx.doi.org/10.3390/electronics3020282, URL http://www.mdpi.com/2079-9292/3/2/282.

[21] A. Al-Naji, K. Gibson, S.-H. Lee, J. Chahl, Monitoring of cardiorespiratory signal: Principles of remote measurements and review of methods, IEEE Access (2017).

[22] A. Qayyum, A.S. Malik, A.N. Shuaibu, N. Nasir, Estimation of non-contact smartphone video-based vital sign monitoring using filtering and standard color conversion techniques, in: 2017 IEEE Life Sciences Conference, LSC, IEEE, 2017, pp. 202–205.

[23] A.A. Kamshilin, O.V. Mamontov, Physiological origin of camera-based PPG imaging, in: Contactless Vital Signs Monitoring, Elsevier, 2022, pp. 27–50, http://dx.doi.org/10.1016/B978-0-12-822281-2.00010-X, URL https://linkinghub.elsevier.com/retrieve/pii/B978012822281200010X.

[24] S. Premkumar, D.J. Hemanth, Intelligent remote photoplethysmography-based methods for heart rate estimation from face videos: A survey, Informatics 9 (3) (2022) 57, http://dx.doi.org/10.3390/informatics9030057, URL https://www.mdpi.com/2227-9709/9/3/57.

[25] J. Steinman, A. Barszczyk, H.-S. Sun, K. Lee, Z.-P. Feng, Smartphones and video cameras: Future methods for blood pressure measurement, Front. Digit. Heal. 3 (2021) 770096, http://dx.doi.org/10.3389/fdgth.2021.770096, URL https://www.frontiersin.org/articles/10.3389/fdgth.2021.770096/full.

[26] F. Bousefsaf, C. Maaoui, A. Pruski, Peripheral vasomotor activity assessment using a continuous wavelet analysis on webcam photoplethysmographic signals, Bio- Med. Mater. Eng. 27 (5) (2016) 527–538.

[27] A. Trumpp, J. Schell, H. Malberg, S. Zaunseder, Vasomotor assessment by camera-based photoplethysmography, Curr. Dir. Biomed. Eng. 2 (1) (2016) 199–202.

[28] W. Wang, S. Stuijk, G. de Haan, Living-skin classification via remote-PPG, IEEE Trans. Biomed. Eng. (2017).

[29] S. Bobbia, R. Macwan, Y. Benezeth, A. Mansouri, J. Dubois, Unsupervised skin tissue segmentation for remote photoplethysmography, Pattern Recognit. Lett. (2017).

[30] T. Shaik, X. Tao, N. Higgins, L. Li, R. Gururajan, X. Zhou, U.R. Acharya, Remote patient monitoring using artificial intelligence: Current state, applications, and challenges, WIREs Data Min. Knowl. Discov. (2023) http://dx.doi.org/10.1002/widm.1485, URL https://onlinelibrary.wiley.com/doi/10.1002/widm.1485.

[31] C. Hurter, D. McDuff, Cardiolens: remote physiological monitoring in a mixed reality environment, in: ACM SIGGRAPH 2017 Emerging Technologies, ACM, 2017, p. 6.

[32] M. Paul, S.C. Behr, C. Weiss, K. Heimann, T. Orlikowsky, S. Leonhardt, Spatio-temporal and -spectral feature maps in photoplethysmography imaging and infrared thermograph, BioMedical Eng. OnLine 20 (1) (2021) 8, http://dx.doi.org/10.1186/s12938-020-00841-9, URL https://biomedical-engineering-online.biomedcentral.com/articles/10.1186/s12938-020-00841-9.

[33] Q. Zhang, Y. Zhou, S. Song, G. Liang, H. Ni, Heart rate extraction based on near-infrared camera: Towards driver state monitoring, IEEE Access 6 (2018) 33076–33087.

[34] T. Vogels, M. van Gastel, W. Wang, G. de Haan, Fully-automatic camera-based pulse-oximetry during sleep, Sleep 1 (2) (2018) 3.

[35] S. Liu, P.C. Yuen, S. Zhang, G. Zhao, 3D mask face anti-spoofing with remote photoplethysmography, in: European Conference on Computer Vision, Springer, 2016, pp. 85–100.

[36] J.W. Severinghaus, Takuo aoyagi: discovery of pulse oximetry, Anesth. Analg. 105 (6) (2007) S1–S4, Publisher: LWW.

[37] I.Y. Volkov, A.A. Sagaidachnyi, A.V. Fomin, Photoplethysmographic imaging of hemodynamics and two-dimensional oximetry, Opt. Spectrosc. (2022) http://dx.doi.org/10.1134/S0030400X22080057, URL https://link.springer.com/10.1134/S0030400X22080057.

[38] T. Rusch, R. Sankar, J. Scharf, Signal processing methods for pulse oximetry, Comput. Biol. Med. 26 (2) (1996) 143–159.

[39] K.A. Reddy, B. George, N.M. Mohan, V.J. Kumar, A novel calibration-free method of measurement of oxygen saturation in arterial blood, IEEE Trans. Instrum. Meas. 58 (5) (2009) 1699–1705.

[40] S. DeMeulenaere, Pulse oximetry: uses and limitations, J. Nurse Pr. 3 (5) (2007) 312–317.

[41] J.E. Sinex, Pulse oximetry: principles and limitations, Am. J. Emerg. Med. 17 (1) (1999) 59–66.

[42] A. Jubran, Pulse oximetry, Intensive Care Med. 30 (2004) 2017–2020.

[43] B. Casserly, R. Read, M.M. Levy, Hemodynamic monitoring in sepsis, Crit. Care Clin. 25 (4) (2009) 803–823.

[44] J.-C. Cobos-Torres, M. Abderrahim, Simple measurement of pulse oximetry using a standard color camera, in: 2017 40th International Conference on Telecommunications and Signal Processing, TSP, IEEE, 2017, pp. 452–455.

[45] N.O. Lutter, S. Urankar, S. Kroeber, False alarm rates of three third-generation pulse oximeters in PACU, ICU and IABP patients, Anesth. Analg. 94 (1 Suppl) (2002) S69–75.

[46] S. Ismail, U. Akram, I. Siddiqi, Heart rate tracking in photoplethysmography signals affected by motion artifacts: a review, EURASIP J. Adv. Signal Process. 2021 (1) (2021) 1–27.

[47] S. Kunchon, T. Desudchit, C. Chinrungrueng, Comparative evaluation of adaptive filters in motion artifact cancellation for pulse oximetry, in: 2009 5th International Colloquium on Signal Processing & Its Applications, IEEE, 2009, pp. 307–311.

[48] K. Alghoul, S. Alharthi, H. Al Osman, A. El Saddik, Heart rate variability extraction from videos signals: ICA vs. EVM comparison, IEEE Access 5 (2017) 4711–4719.

[49] O. Rioul, P. Duhamel, Fast algorithms for discrete and continuous wavelet transforms, IEEE Trans. Inform. Theory 38 (2) (1992) 569–586.

[50] K.V.P. Naraharisetti, M. Bawa, Comparison of different signal processing methods for reducing artifacts from photoplethysmograph signal, in: 2011 IEEE International Conference on Electro/Information Technology, IEEE, 2011, pp. 1–8.

[51] M.-Z. Poh, D.J. McDuff, R.W. Picard, Advancements in noncontact, multiparameter physiological measurements using a webcam, IEEE Trans. Biomed. Eng. 58 (1) (2011) 7–11.

[52] N.G. Cortez, I.G. Cohen, A.S. Kesselheim, FDA regulation of mobile health technologies, N. Engl. J. Med. 371 (4) (2014) 372–379.

[53] European Union Law, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC, Regulation (EU) 2017/745 of the European parliament and of the council on medical devices, 2017, http://data.europa.eu/eli/reg/2017/745/2024-07-09.

[54] X. Niu, H. Han, S. Shan, X. Chen, VIPL-HR: A multi-modal database for pulse estimation from less-constrained face video, in: Asian Conference on Computer Vision, Springer, 2018, pp. 562–576.

[55] R. Stricker, S. Müller, H.-M. Gross, Non-contact video-based pulse rate measurement on a mobile service robot, in: Robot and Human Interactive Communication, 2014 RO-MAN: the 23rd IEEE International Symposium on, IEEE, 2014, pp. 1056–1062.

[56] B. Hamoud, W. Othman, N. Shilov, A. Kashevnik, Contactless oxygen saturation detection based on face analysis: An approach and case study, in: 2023 33rd Conference of Open Innovations Association, FRUCT, IEEE, Zilina, Slovakia, 2023, pp. 54–62, http://dx.doi.org/10.23919/FRUCT58615.2023.10143059, URL https://ieeexplore.ieee.org/document/10143059/.

[57] C.-H. Cheng, Z. Yuen, S. Chen, K.-L. Wong, J.-W. Chin, T.-T. Chan, R.H.Y. So, Contactless blood oxygen saturation estimation from facial videos using deep learning, Bioengineering 11 (3) (2024) 251, http://dx.doi.org/10.3390/bioengineering11030251, URL https://www.mdpi.com/2306-5354/11/3/251.

[58] M. Hu, X. Wu, X. Wang, Y. Xing, N. An, P. Shi, Contactless blood oxygen estimation from face videos: A multi-model fusion method based on deep learning, Biomed. Signal Process. Control. 81 (2023) 104487, http://dx.doi.org/10.1016/j.bspc.2022.104487, URL https://linkinghub.elsevier.com/retrieve/pii/S1746809422009417.

[59] K. Humphreys, T. Ward, C. Markham, Noncontact simultaneous dual wavelength photoplethysmography: a further step toward noncontact pulse oximetry, Rev. Sci. Instrum. 78 (4) (2007) 044304.

[60] 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.

[61] 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.

[62] C.G. Scully, Jinseok Lee, J. Meyer, A.M. Gorbach, D. Granquist-Fraser, Y. Mendelson, K.H. Chon, Physiological parameter monitoring from optical recordings with a mobile phone, IEEE Trans. Biomed. Eng. 59 (2) (2012) 303–306, http://dx.doi.org/10.1109/TBME.2011.2163157, URL http://ieeexplore.ieee.org/document/5963704/.

[63] U. Bal, Non-contact estimation of heart rate and oxygen saturation using ambient light, Biomed. Opt. Express 6 (1) (2015) 86–97.

[64] B. Wei, X. Wu, C. Zhang, Z. Lv, Analysis and improvement of non-contact SpO2 extraction using an RGB webcam, Biomed. Opt. Express 12 (8) (2021) 5227, http://dx.doi.org/10.1364/BOE.423508, URL https://www.osapublishing.org/abstract.cfm?URI=boe-12-8-5227.

[65] A. Moço, W. Verkruysse, Pulse oximetry based on photoplethysmography imaging with red and green light, J. Clin. Monit. Comput. (2020) 1–11, Publisher: Springer.

[66] M. van Gastel, W. Verkruysse, G. de Haan, Data-driven calibration estimation for robust remote pulse-oximetry, Appl. Sci. 9 (18) (2019) 3857, Publisher: Multidisciplinary Digital Publishing Institute.

[67] 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) (2015) 1091–1098.

[68] T. Aoyagi, M. Fuse, N. Kobayashi, K. Machida, K. Miyasaka, Multiwavelength pulse oximetry: theory for the future, Anesth. Analg. 105 (6) (2007) S53–S58, Publisher: LWW.

[69] M.E. Wieler, T.G. Murphy, M. Blecherman, H. Mehta, G.J. Bender, Infant heart-rate measurement and oxygen desaturation detection with a digital video camera using imaging photoplethysmography, J. Perinatol. 41 (7) (2021) 1725–1731, http://dx.doi.org/10.1038/s41372-021-00967-1, URL http://www.nature.com/articles/s41372-021-00967-1.

[70] T. Lan, G. Li, L. Lin, A non-contact oxygen saturation detection method based on dynamic spectrum, Infrared Phys. Technol. 127 (2022) 104421, http://dx.doi.org/10.1016/j.infrared.2022.104421, URL https://linkinghub.elsevier.com/retrieve/pii/S1350449522004029.

[71] G. de Haan, V. Jeanne, Robust pulse rate from chrominance-based rPPG, IEEE Trans. Biomed. Eng. 60 (10) (2013) 2878–2886.

[72] W. Wang, A.C. den Brinker, S. Stuijk, G. de Haan, Algorithmic principles of remote PPG, IEEE Trans. Biomed. Eng. 64 (7) (2017) 1479–1491.

[73] R. Takahashi, K. Ashida, Y. Kobayashi, R. Tokunaga, S. Kodama, N. Tsumura, Oxygen saturation estimation based on optimal band selection from multi-band video, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 3850–3856.

[74] Z. Sun, Q. He, Y. Li, W. Wang, R.K. Wang, Robust non-contact peripheral oxygenation saturation measurement using smartphone-enabled imaging photoplethysmography, Biomed. Opt. Express 12 (3) (2021) 1746, http://dx.doi.org/10.1364/BOE.419268, URL https://www.osapublishing.org/abstract.cfm?URI=boe-12-3-1746.

[75] I. Fine, A. Weinreb, Multiple scattering effect in transmission pulse oximetry, Med. Biol. Eng. Comput. 33 (1995) 709–712.

[76] A. Moço, W. Verkruysse, Pulse oximetry based on photoplethysmography imaging with red and green light: Calibratability and challenges, J. Clin. Monit. Comput. 35 (1) (2021) 123–133.

[77] M. van Gastel, W. Verkruysse, Contactless SpO_2 with an RGB camera: experimental proof of calibrated SpO_2, Biomed. Opt. Express 13 (12) (2022) 6791, http://dx.doi.org/10.1364/BOE.471332, URL https://opg.optica.org/abstract.cfm?URI=boe-13-12-6791.

[78] A. Al-Naji, G.A. Khalid, J.F. Mahdi, J. Chahl, Non-contact SpO2 prediction system based on a digital camera, Appl. Sci. 11 (9) (2021) 4255, http://dx.doi.org/10.3390/app11094255, URL https://www.mdpi.com/2076-3417/11/9/4255.

[79] X. Tian, C.-W. Wong, S.M. Ranadive, M. Wu, A multi-channel ratio-of-ratios method for noncontact hand video based SpO$_2$ Monitoring using smartphone cameras, IEEE J. Sel. Top. Signal Process. 16 (2) (2022) 197–207, http://dx.doi.org/10.1109/JSTSP.2022.3152352, URL https://ieeexplore.ieee.org/document/9718200/.

[80] H. Liu, K. Ivanov, Y. Wang, L. Wang, A novel method based on two cameras for accurate estimation of arterial oxygen saturation, Biomed. Eng. Online 14 (1) (2015) 52.

[81] D. Mishra, N. Priyadarshini, S. Chakraborty, M. Sarkar, Blood oxygen saturation measurement using polarization-dependent optical sectioning, IEEE Sensors J. 17 (12) (2017) 3900–3908.

[82] A.d.F.G. Rosa, R.C. Betini, Noncontact spo 2 measurement using Eulerian video magnification, IEEE Trans. Instrum. Meas. 69 (5) (2019) 2120–2130, Publisher: IEEE.

[83] M. Van Gastel, S. Stuijk, G. De Haan, New principle for measuring arterial blood oxygenation, enabling motion-robust remote monitoring, Sci. Rep. 6 (2016).

[84] 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.

[85] D. Luguern, Nouvelle approche pour l'estimation du rythme respiratoire basée sur la photopléthysmographie sans contact (Ph.D. thesis), Bourgogne Franche-Comté, 2021.

[86] A. Alkan, M.U. Abdullah, H.O. Abdullah, M. Assaf, H. Zhou, A smart agricultural application: automated detection of diseases in vine leaves usinghybrid deep learning, Turk. J. Agric. For. 45 (6) (2021) 717–729.

[87] K.M. Sünnetci, A. Alkan, Lung cancer detection by using probabilistic majority voting and optimization techniques, Int. J. Imaging Syst. Technol. 32 (6) (2022) 2049–2065.

[88] K.M. Sunnetci, E. Kaba, F. Beyazal Çeliker, A. Alkan, Comparative parotid gland segmentation by using ResNet-18 and MobileNetV2 based DeepLab v3+ architectures from magnetic resonance images, Concurr. Comput.: Pr. Exp. 35 (1) (2023) e7405.

[89] 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. (2019) Publisher: IEEE.

[90] W. Chen, D. McDuff, DeepMag: Source specific motion magnification using gradient ascent, 2018, arXiv preprint arXiv:1808.03338.

[91] X. Ding, D. Nassehi, E.C. Larson, Measuring oxygen saturation with smartphone cameras using convolutional neural networks, IEEE J. Biomed. Heal. Informatics 23 (6) (2018) 2603–2610, Publisher: IEEE.

[92] J. Mathew, X. Tian, C.-W. Wong, S. Ho, D.K. Milton, M. Wu, Remote blood oxygen estimation from videos using neural networks, IEEE J. Biomed. Heal. Informatics (2023).

[93] A. Kedir, M. Ullah, J.R. Bauer, Spectranet: A deep model for skin oxygenation measurement from multi-spectral data, Electron. Imaging 32 (15) (2020) http://dx.doi.org/10.2352/ISSN.2470-1173.2020.15.COLOR-083, https://library.imaging.org/ei/articles/32/15/art00002, 83-1–83-1.

[94] Y. Akamatsu, Y. Onishi, H. Imaoka, Blood oxygen saturation estimation from facial video via dc and ac components of spatio-temporal map, in: ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP, IEEE, 2023, pp. 1–5.

[95] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778.
[96] G. Huang, Z. Liu, L. Van Der Maaten, K.Q. Weinberger, Densely connected convolutional networks, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 4700–4708.
[97] T.B. Fitzpatrick, The validity and practicality of sun-reactive skin types I through VI, Arch. Dermatol. 124 (6) (1988) 869, http://dx.doi.org/10.1001/archderm.1988.01670060015008, URL http://archderm.jamanetwork.com/article.aspx?doi=10.1001/archderm.1988.01670060015008.
[98] E.M. Nowara, D. McDuff, A. Veeraraghavan, A meta-analysis of the impact of skin tone and gender on non-contact photoplethysmography measurements, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2020, pp. 284–285.
[99] J. Wang, C. Shan, L. Liu, Z. Hou, Camera-based physiological measurement: Recent advances and future prospects, Neurocomputing (2024) 127282.
[100] L. Saikevičius, V. Raudonis, G. Dervinis, V. Baranauskas, Non-contact vision-based techniques of vital sign monitoring: Systematic review, Sensors 24 (12) (2024) 3963.