

<https://doi.org/10.1038/s41746-026-02715-1>

Roadmap of remote photoplethysmography from heart rate measurement toward clinical translation



Mohamed Elgendi^{1,2,3,9}✉, Shuo Li^{3,4,5,9}, Carlo Menon³, Edmund Y. Lam⁵, Basem Al-Omari⁶, Stefan Borik^{7,8}, Kirk H. Shelley⁸ & Aymen Alian⁸✉

Remote photoplethysmography (rPPG) enables non-contact heart rate (HR) and waveform measurement from facial video, offering advantages over contact methods while contending with motion, illumination changes, compression, and frame-rate variability. This roadmap integrates a systematic review, technical synthesis, and a *clinic-acquired* demonstration to accelerate rPPG research and translation. We outline the measurement pipeline and synthesize progress across eight domains: datasets, device factors, data scarcity, skin/ROI detection, model- and data-driven algorithms, filtering, and evaluation metrics. Strategies for improved robustness include targeted dataset expansion, augmentation, optimized ROI policies, motion and photometric normalization, and hardware configuration. A proof-of-feasibility using data collected in a routine clinical setting alongside reference monitors demonstrates high signal-level agreement under a multi-metric framework (correlation/concordance, absolute errors, and Bland–Altman bias/dispersion). We conclude with recommendations for accuracy, efficiency, and deployment in real-world and clinical contexts, providing a consolidated resource for researchers and clinicians and a pragmatic path toward reliable clinical adoption.

In recent years, with increasing public awareness of physical health, physiological signal monitoring technologies have received growing attention. Such technologies can facilitate the measurement of various physiological indicators, enabling health status assessment and proactive management. Among vital physiological indicators, heart rate (HR) holds particular significance. It serves as a valuable measure of fatigue levels¹, emotional state², and body functions³. Consequently, improving the accuracy and efficiency of HR measurement under different circumstances has become a key research focus in this field.

The widely accepted gold standard^{4,5} for HR measurement is the electrocardiogram (ECG). As shown in Fig. 2, ECG signals are recorded by placing electrodes on specific body locations and connecting wires to the recording machine⁶. Each heartbeat induces myocardial depolarization, altering skin surface potentials that can be captured by electrodes and

amplified into ECG signals. A typical ECG waveform contains three main components (P wave, QRS complex, and T wave), each reflecting specific phases of cardiac electrical activity^{7,8}. Since ECG signals contain abundant cardiac information, clinicians can evaluate HR, cardiac rhythm, PR interval, QT interval, and other aspects by analyzing the ECG waveform in order to determine the physical condition of the patient and potential heart diseases^{9–11}. However, the need for professional setup of traditional 12-lead ECG restricts its suitability for long-term monitoring outside clinical settings. Ambulatory ECG techniques such as Holter monitors or adhesive patch recorders have been developed to enable continuous cardiac rhythm monitoring, and ECG-enabled smartwatches allow users to obtain recordings without professional guidance. Nevertheless, all these approaches still rely on contact-based devices for signal acquisition, which could impose burdens on daily activities.

¹Department of Biomedical Engineering and Biotechnology, Khalifa University of Science and Technology, Abu Dhabi, United Arab Emirates. ²Healthcare Innovation Group (HEIG), Department of Biomedical Engineering, Khalifa University, Abu Dhabi, United Arab Emirates. ³Biomedical and Mobile Health Technology Lab, ETH Zürich, Zürich, Switzerland. ⁴Department of Information Technology and Electrical Engineering, ETH Zürich, Zürich, Switzerland. ⁵Department of Electrical and Electronic Engineering, The University of Hong Kong, Pokfulam, Hong Kong, China. ⁶Department of Public Health and Epidemiology, Khalifa University of Science and Technology, Abu Dhabi, United Arab Emirates. ⁷Department of Electromagnetic and Biomedical Engineering, Faculty of Electrical Engineering and Information Technology, University of Zilina, Zilina, Slovakia. ⁸Yale School of Medicine, Yale University, New Haven, CT, USA. ⁹These authors contributed equally: Mohamed Elgendi, Shuo Li. ✉ e-mail: mohamed.elgendi@ku.ac.ae; aymen.alian@yale.edu

Contact-based photoplethysmography (cPPG) offers a more convenient alternative by detecting blood volume changes through optical sensors placed on the skin¹². A typical setup for cPPG signal acquisition includes an LED light source and a photosensitive sensor. The light emitted by the LED is absorbed by blood vessels during the process of penetrating through the skin and tissue¹³. Variations in arterial blood flow caused by the rhythmic heartbeat lead to changes in the absorption intensity over time. The photosensitive sensor captures these periodic changes in light intensity and then converts them into an electronic signal, which is the resulting PPG signal and is also called blood volume pulse (BVP) signal. As presented in Fig. 2, cPPG techniques are widely used in smartwatches (reflective) and fingertip oximeters (transmissive)¹⁴. Compared with ECG, cPPG devices are easier to operate but provide less clinical information, making them suitable mainly for consumer health monitoring.

While contact-based methods such as ECG and cPPG have been widely used for HR monitoring, their reliance on skin contact can cause discomfort, limit long-term wearability, and restrict applications in certain scenarios such as telemedicine or large-scale public health screening. This limitation has motivated the development of fully non-contact alternatives. Remote photoplethysmography (rPPG) techniques were proposed to meet this need by estimating HR from videos of exposed skin regions, where blood flow induces subtle color variations that can be analyzed to extract the blood volume pulse (BVP) signal. The earliest exploration of imaging-based photoplethysmography was conducted by Wu, Blažek, and Schmitt in 2000¹⁵, who demonstrated that periodic changes in subcutaneous blood flow could produce measurable intensity variations in reflected light. Using a high-sensitivity video camera, they visualized dermal perfusion in a completely non-contact manner, laying the groundwork for later developments in rPPG. Building on this concept, Verkrusye et al.¹⁶ extended the approach

to facial videos illuminated by ambient light, marking the emergence of modern rPPG.

Since then, the field has rapidly evolved, with researchers developing increasingly sophisticated methods and frameworks to improve accuracy and robustness in HR estimation^{17,18}. Recent advances have expanded beyond algorithmic optimization to address fairness¹⁹, regional signal contributions^{20,21}, privacy-preserving signal removal²², clinical translation through healthcare kiosks²³, and improved signal quality assessment²⁴. Collectively, these developments highlight the rapid maturation of rPPG from experimental validation to real-world healthcare applications.

Results

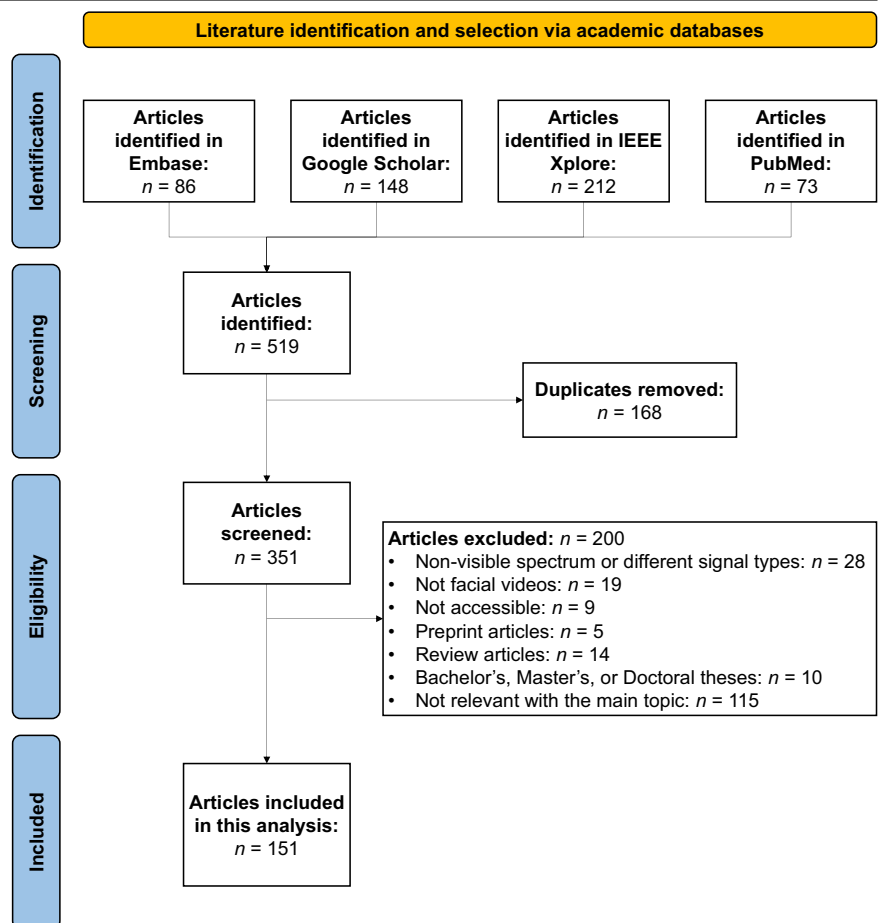
Study selection

As illustrated in Fig. 1, we conducted the initial identification across four academic databases. The initial search yielded the following results: 86 articles from Embase, 148 from Google Scholar, 212 from IEEE Xplore, and 73 from PubMed, totaling 519 articles. After removing 168 duplicates, 351 unique articles remained. Among them, 28 were excluded because they focus on non-visible spectrum (e.g., infrared) or other types of signals (e.g., ballistocardiography). Furthermore, 19 articles that used ROIs other than the human face were removed, as well as 9 articles that were not accessible. In addition, we excluded 5 preprint articles, 14 review papers, 10 theses (Bachelor's, Master's, or Doctoral), and 115 articles due to irrelevance to the main topic. After all exclusions, the number of articles included in our final review was 151.

rPPG pipeline

A typical rPPG-based HR estimation pipeline comprises several key steps (Fig. 2). First, a facial video is acquired using a standard camera. Next, the region of interest (ROI) is automatically identified; the forehead and cheeks

Fig. 1 | PRISMA flow diagram of the study selection process. After eligibility screening, 151 articles were included in this work from 519 initial search results.



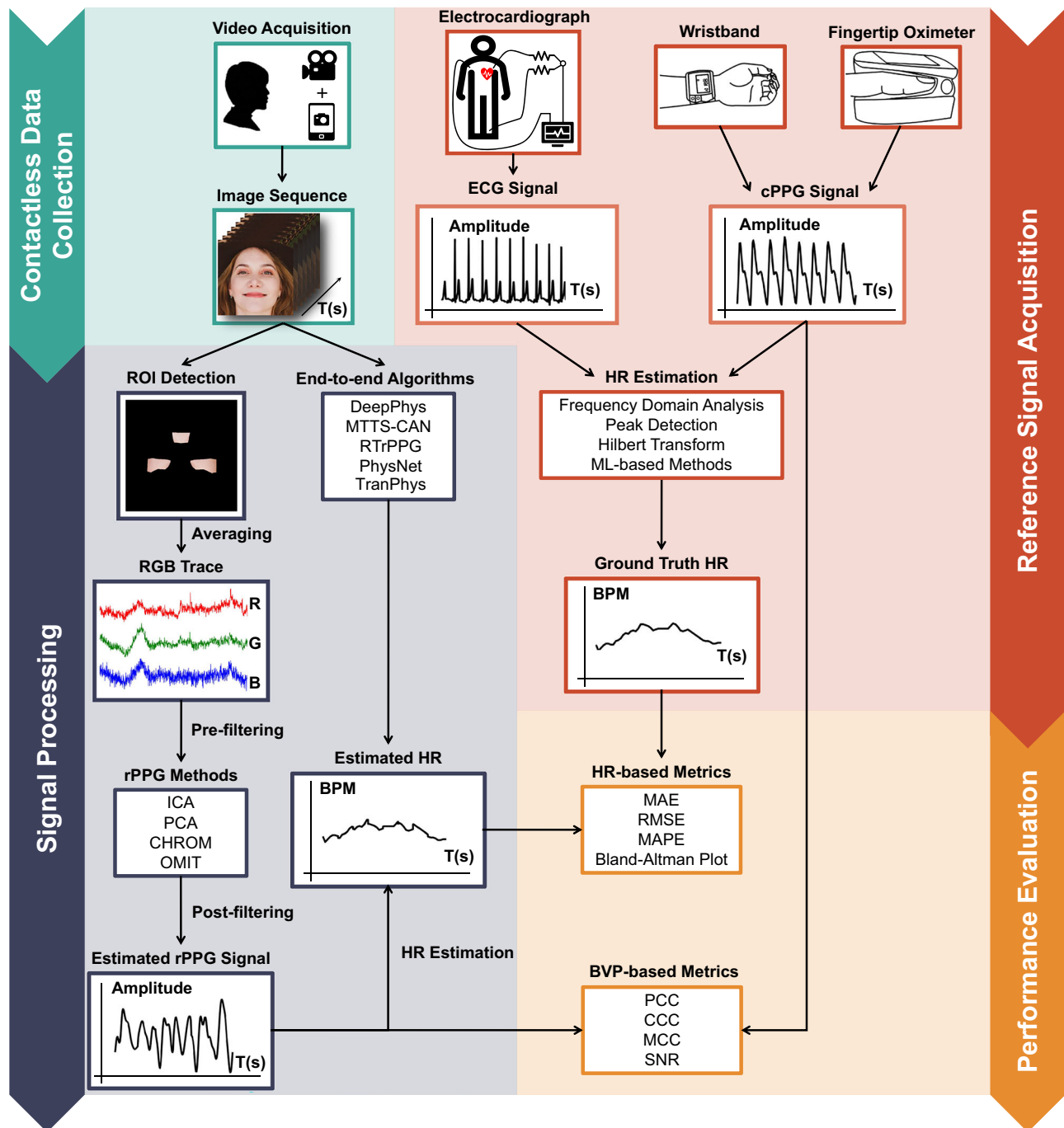


Fig. 2 | General pipeline for the use of rPPG algorithms in HR measurement and algorithm performance validation. RGB (red, green, and blue); ECG (electrocardiogram); cPPG (contact-based photoplethysmography); rPPG (remote photoplethysmography); HR (heart rate); BPM (beats per minute); ROI (region of interest); BVP (blood volume pulse). The pipeline contains three main steps: data collection, signal processing, and performance evaluation. The data collection step involves recording facial videos and collecting HR-related ground truth data. During the video acquisition process, webcams and smartphone cameras are commonly used as recording devices. Wristband oximeters and fingertip oximeters are used for cPPG signal recording. ECG signals are acquired via electrocardiographs. In the signal processing step, end-to-end algorithms like DeepPhys¹⁵⁴, MTTs-CAN (Multi-Task Temporal Shift Convolutional Attention Network)¹⁵⁵, RTrPPG (Real-Time remote Photoplethysmography)¹⁵⁹, PhysNet¹⁵⁶, and TranPhys¹⁶⁷ can be applied to raw videos to extract BVP signals or HR values. For model-driven rPPG methods like

ICA (Independent Component Analysis)¹³⁷, PCA (Principal Component Analysis)¹⁴⁰, CHROM (Chrominance-based method)¹⁴⁴, and OMIT (Orthogonal Matrix Image Transformation)¹⁴², RGB traces are extracted by averaging over selected ROIs and then used as algorithm inputs. In the performance evaluation step, estimated BVP signals or HR values are compared with the ground truth, which was previously collected. HR signals are often evaluated based on MAE (Mean Absolute Error), RMSE (Root Mean Square Error), MAPE (Mean Absolute Percentage Error), and Bland-Altman plot. Common BVP-based performance metrics are: PCC (Pearson Correlation Coefficient), CCC (Concordance Correlation Coefficient), MCC (Maximum Cross-Correlation), and SNR (Signal-to-Noise Ratio). The real human face image was sourced from the Real and Fake Face Detection Dataset (<https://www.kaggle.com/datasets/ciplab/real-and-fake-face-detection>), which is under the CC BY-NC-SA 4.0 License (<https://creativecommons.org/licenses/by-nc-sa/4.0>). Icons were adapted from the Microsoft 365 non-copyrighted icon library.

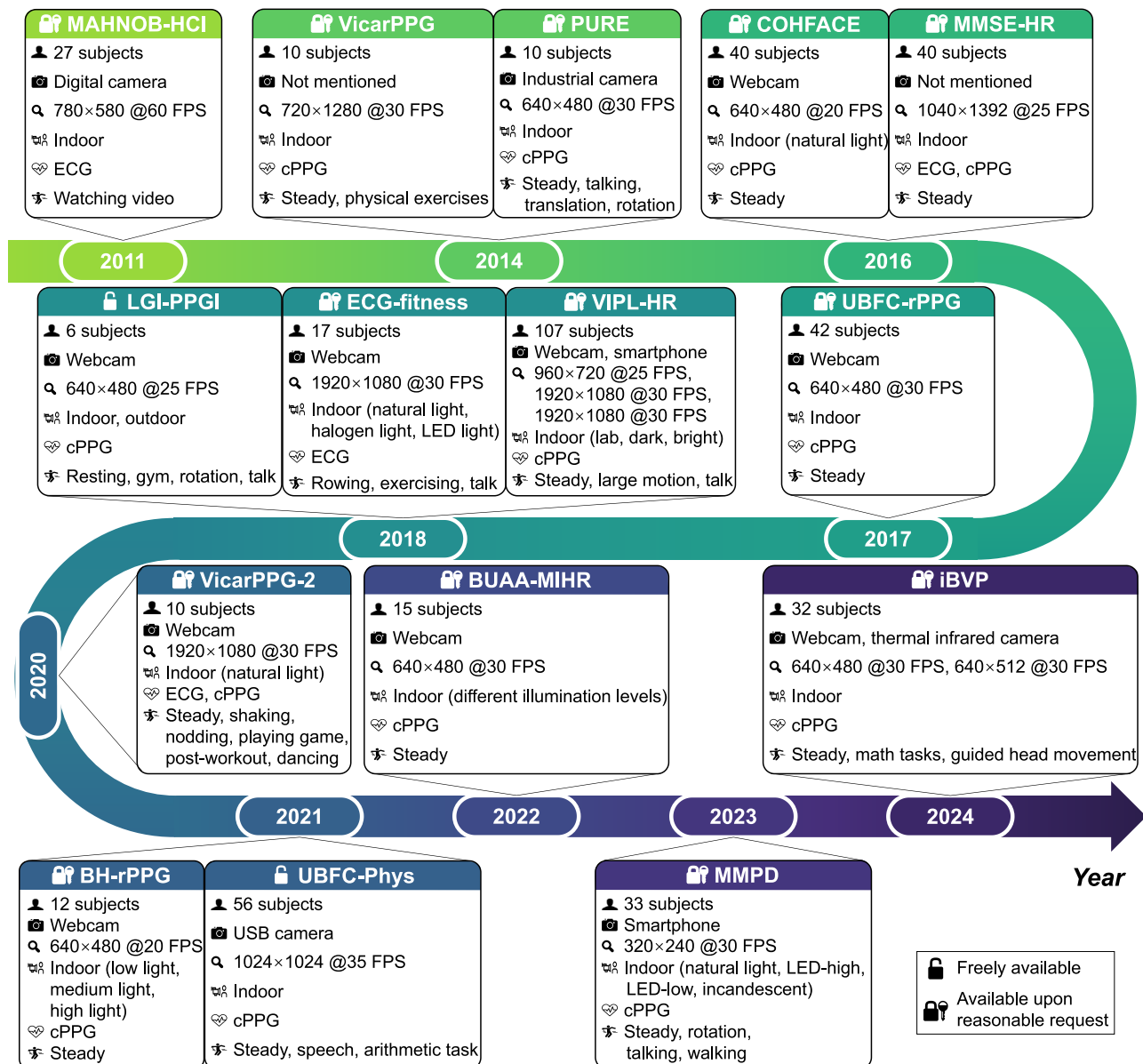


Fig. 3 | Timeline of current public datasets for rPPG algorithm validation. These datasets were developed for various research purposes. Each block provides the following information: number of included participants, camera model, resolution, frames per second (FPS), acquisition environment, ground truth signal modality,

and motion type. Icons were adapted from the Microsoft Fluent Emoji project (<https://github.com/microsoft/fluentui-emoji>), which is under the MIT License (<https://opensource.org/licenses/mit>).

are the most commonly chosen ROIs^{25,26}. Pixel values within these regions are then converted into red, green, and blue (RGB) traces, from which blood volume pulse (BVP) signals are derived using established algorithms. Before and/or after BVP extraction, temporal and spectral filtering can be leveraged to improve signal quality. HR is then estimated from the BVP using frequency-domain analysis²⁷, peak detection²⁸, Hilbert transform²⁹, or machine-learning methods that extract higher-order features³⁰. A windowing operation is often employed to select specific time segments during this process^{31–33}. Finally, standard performance metrics are used to evaluate the estimated BVP and HR signals against the ground truth.

Datasets

Throughout the evolution of rPPG techniques, several challenges related to input videos have emerged as central concerns. First and foremost, effectively mitigating motion artifacts remains a critical issue. Different types of subject movement, such as talking, head rotation, and nodding, introduce perturbations in skin pixel illumination, resulting in reduced measurement

accuracy. Second, achieving robust performance under varying illumination conditions presents another major challenge. Ideal lighting conditions, characterized by the proper placement of cameras^{34,35}, uniformly diffused facial illumination, and stable, moderate light intensity, are often unattainable in real-world applications. Datasets constructed for the development of rPPG techniques should serve as evaluation platforms to assess algorithm performance across a broad range of measurement scenarios. Moreover, demographic factors such as gender, age, and skin tone require careful consideration, as the underlying anatomical and physiological differences may potentially affect the measurement accuracy. In sum, during the formulation of an rPPG dataset, researchers ought to consider these facets and aim to enhance the representativeness of data distributions to better reflect real-world conditions. Several research groups have created and released rPPG datasets tailored to specific objectives. These public datasets serve as valuable resources for both algorithm development and performance validation. The 15 public rPPG datasets included in this work are summarized in Fig. 3.

In 2011, the multimodal database for affect recognition and implicit tagging (MAHNOB-HCI) was introduced³⁶. It became the first public dataset widely utilized for assessing the performance of rPPG techniques. However, as it was primarily designed for multimedia tagging research, the MAHNOB-HCI dataset was not specifically constructed for rPPG algorithm evaluation. In 2014, Stricker et al.³⁷ proposed the PURE dataset, which was used for pulse rate measurement on a mobile service robot. The VicarPPG dataset³⁸ was released in the same year. These two datasets provide various motion types to facilitate rPPG algorithm validation across a broad range of scenarios.

To support evaluation under different emotional states and skin tones, Zhang et al.³⁹ released the MMSE-HR dataset in 2016. In the same year, the COHFACE dataset⁴⁰ was introduced to promote reproducible research in rPPG techniques; it was collected under two lighting conditions: studio light and natural light. In 2017, the UBFC-rPPG dataset was constructed to assess an unsupervised skin tissue segmentation method⁴¹. In 2018, following the development of the local group invariance (LGI) algorithm⁴², the LGI-PPGI dataset was made publicly available. Additionally, the VIPL-HR⁴³ and ECG-fitness⁴⁴ datasets were also released to the public. The VIPL-HR dataset includes visible light videos of 107 subjects. To our knowledge, it is the largest public rPPG dataset, comprising 2378 videos across nine measurement conditions. The ECG-fitness database was created to comprise recordings of various physical activities on fitness machines. It was originally used to evaluate the performance of a convolutional neural network (CNN) for HR measurement in gym settings.

In 2020, the VicarPPG2 dataset⁴⁵ was introduced to expand the evaluation scope of rPPG technologies to a wider variety of subject motions. Nevertheless, datasets specifically designed to assess rPPG algorithms under different light intensities had remained notably absent. To address this gap, Yang et al.⁴⁶ created the BH-rPPG dataset in 2021, which includes recordings at three distinct illumination levels. Moreover, in 2022, Xi et al.⁴⁷ proposed the BUAA-MIHR dataset, defining 11 illumination levels through quantitative illuminance measurements. A more comprehensive dataset comprising multiple application domains across various skin tones, subject movement, illumination conditions, and video resolutions, namely the multi-domain mobile video physiology dataset (MMPD), was proposed by Tang et al.⁴⁸ in 2023. It was constructed based on common measurement scenarios of rPPG defined in previous studies. The most recent contribution is the imaging blood volume pulse (iBVP) dataset that was released by Joshi et al.⁴⁹ in 2024. The dataset construction offers novel insights from two perspectives: first, the ground truth cPPG signals were recorded from the ear, which are more synchronized with the facial rPPG signals; the other is the per-sample signal quality labels provided by the dataset to discard noisy segments of ground truth and video sequences.

Overall, existing public datasets have covered quite a few common scenarios, such as offices and fitness environments. It is worth noting that other domains such as remote medical consultation, in-vehicle monitoring, and gaming sessions, hold considerable commercial value but remain underexplored⁵⁰. As exhibited in Fig. 3, most public datasets are not freely available to the research community. With the consideration regarding subject privacy and the risk of data misuse, applicants are typically required to possess academic qualifications and contact the data distributor. For datasets with stricter access requirements, applicants may also need to sign an end-user license agreement (EULA) to formalize the application process. Currently, only the LGI-PPGI⁴² and UBFC-Phys⁵¹ datasets can be directly accessed via public online resources.

Device-related factors

In the context of rPPG algorithms, the method of data acquisition typically involves capturing facial videos using standard RGB cameras. Webcams are commonly employed due to their practicality and affordability, as opposed to more advanced cameras that are often prohibitively expensive. As shown in Fig. 3, the UBFC-rPPG, LGI-PPGI, COHFACE, BUAA-MIHR, VIPL-HR, ECG-fitness, VicarPPG-2, and BH-rPPG datasets selected webcams of the Logitech series, which have also been widely used in other self-

constructed datasets. Moreover, the landscape of video recording tools has expanded with the advent of mobile HR monitoring, prompting researchers to utilize the integrated cameras of smartphones for facial video acquisition⁵². For example, both the VIPL-HR and MMPD datasets employ standard smartphone cameras to record facial video sequences. In comparison to webcams, using smartphones for video collection introduces additional challenges. Mobile phone cameras often incorporate built-in algorithms and filters aimed at enhancing facial aesthetics. However, such enhancements may introduce unintended perturbations that affect video authenticity, making it necessary to disable them during recording. Rear cameras are preferred over front-facing cameras due to their generally higher image quality and resolution. Beyond these considerations, stabilizing the recording position using a tripod or a stand helps mitigate motion artifacts caused by camera shaking.

Resolution and FPS. Video resolution and frames per second (FPS) are two critical factors in determining the video quality, but the research community has different opinions on whether they significantly influence the quality of extracted rPPG signals. Some scholars argue that enhancing image resolution improves the signal-to-noise ratio (SNR), thereby yielding more accurate HR measurement^{53–55}. Notably, the majority of the included rPPG datasets, as depicted in Fig. 3, have a resolution of 640×480 . The highest resolution observed is 1920×1080 , which is commonly achieved via smartphone cameras. However, these resolutions are still considered low in the context of super-resolution video reconstruction. In addition to super-resolution techniques, a more straightforward approach is to use advanced cameras capable of capturing higher resolutions videos. However, the increasing cost related to hardware and storage pose significant barriers. Alongside resolution, FPS is another determinant that may affect measurement accuracy. According to the findings of Rouast et al.⁵⁶, a higher FPS rate results in smoother video sequences, thereby facilitating the extraction of HR-related signals. The frame rate should also remain stable to ensure the temporal consistency. This is one reason that rPPG algorithms sometimes exhibit significant errors in real-time video streaming scenarios^{57,58}. Álvarez Casado et al.⁵⁹ quantitatively analyzed the effects of reduced frame rate, video resolution degradation, and frame rate instability on rPPG signal quality under limited network bandwidth and computing resources. They also developed mitigation strategies from both the transmitter and receiver sides to maintain acceptable measurement accuracy even under resource-constrained conditions. Nevertheless, Blackford et al.⁶⁰ argued that the influence of resolution and FPS under multiple camera settings is actually very slight, probably due to the complementary information provided by the camera array. In their experiments, they reduced the FPS from 120 to 60 and 30, and the resolution from 658×492 to 329×246 . The results demonstrated that the HR estimation accuracy remained stable even after decreasing both resolution and FPS.

Video compression methods. In addition to resolution and FPS, video compression methods may have a greater impact on HR measurement accuracy. McDuff et al.⁶¹ conducted a systematic study of video compression methods, revealing a linear relationship between the SNR of extracted BVP signals and the bit rate. Hyanda et al.⁶² concluded that lossless compression methods, such as FFV1, HuffYUV, v410, and v210, tend to yield better HR estimation accuracy compared with x264, x265, MPEG-4, and VP9, which belong to lossy compression methods. The VP9 video codec was recommended among lossy compression methods for its ability to balance HR estimation accuracy and video size. Furthermore, different encoding modes like constant bitrate (CBR) and constant rate factor (CRF) have also been proven to affect the rPPG signal quality⁶³. Instead of relying on conventional video compression methods, Wang et al. developed a specialized compression scheme that can effectively preserve the rPPG signal quality while reducing video size to approximately 1/60 of the original⁶⁴. A recent study by Nguyen et al.⁶⁵ conducted a comprehensive evaluation of key factors influencing the

performance of rPPG algorithms, including video resolution, frame rate, and compression methods. Their results supported the significant impact of all three factors on the accuracy of remote HR measurement. They also proposed a range of strategies to mitigate these negative effects and enhance the robustness of rPPG algorithms.

Camera exposure. The influence of the camera exposure process is generally considered a minor factor. The impact of exposure time—also referred to as shutter time—has been demonstrated in several studies^{66–68}. These studies also reported that regulating the exposure time range can mitigate the occurrence of saturation within selected ROIs. Additionally, the rolling shutter effect stemming from the readout process of CMOS sensors can introduce distortions and phase shifts during video recording. Several experiments have provided evidence showing that this effect has an impact on rPPG signal quality⁶⁹.

Ground truth signal acquisition. The collection of synchronized ground truth signals is a crucial step during dataset construction. The ground truth signals are typically provided in the form of ECG or cPPG recordings, which are then used to compute ground truth HR values. The methods of ECG and cPPG signal acquisition have been schematically illustrated in Fig. 2. As presented in Fig. 3, the MAHNOB-HCI and ECG-fitness datasets, used ECG signals as the ground truth. In contrast, the VicarPPG, PURE, COHFACE, UBFC-rPPG, VIPL-HR, LGI-PPGI, BH-rPPG, UBFC-Phys, BUAA-MIHR, MMPD, and iBVP datasets employed cPPG signals. In order to provide more comprehensive ground truth data, the MMSE-HR and VicarPPG2 datasets collected both ECG and cPPG signals. To guarantee data fidelity, most physiological signal acquisition instruments were selected among mature products under the approval for clinical use from the U.S. Food and Drug Administration (FDA).

Strategies for addressing data scarcity

Over the years, deep learning (DL) algorithms have shown superior potential and significant performance compared to model-driven algorithms. Beyond the algorithm structure, the effectiveness of DL-based algorithms depends largely on the training dataset. A larger dataset often enables the algorithm to capture a broader array of patterns, leading to improved accuracy. Various strategies have been proposed to counterbalance the risk of overfitting and to improve the generalization ability of rPPG algorithms. In addition to data deficiency, the diversity of the dataset is another determinant. Some public datasets suffer from a lack of diversity in participant demographics, such as age, gender, and ethnicity⁷⁰.

To mitigate demographic imbalance and dataset-induced bias, future rPPG datasets should adopt standardized design and reporting protocols. At a minimum, datasets should (i) report participant age ranges, sex, and skin tone using standardized scales (e.g., Fitzpatrick skin type); (ii) ensure coverage across multiple demographic strata rather than single-group convenience sampling; (iii) document health status and exclusion criteria where applicable; (iv) provide transparent dataset access policies with clearly defined academic or ethical use conditions; and (v) report demographic stratification in performance evaluation to expose potential bias. Adoption of such protocols would facilitate fairness-aware benchmarking and improve the translational reliability of rPPG algorithms.

Without such standardized protocols, this lack of diversity could potentially lead to bias when applying DL algorithms to underrepresented groups. Furthermore, the dearth of application scenarios (e.g., varying lighting conditions and subject motions) may pose similar challenges.

Expand the dataset size. The most straightforward and effective solution to overcome the challenge of data scarcity is to collect an expanded dataset encompassing diverse facets; however, this approach is encumbered by practical constraints. Firstly, collecting facial videos raises significant privacy concerns. Ensuring compliance with ethical guidelines and secure data storage is challenging. Moreover, recruiting a

sufficient number of participants with diverse demographic features is difficult. The recruitment scope of participants may lead to potential biases and must be carefully considered. Building a comprehensive rPPG dataset is also highly resource-intensive. Costs associated with instrument maintenance, participant recruitment, data management, and infrastructural support tend to grow with the data volume.

Data augmentation. An alternative and pragmatic approach involves performing data augmentation based on existing datasets, i.e., generating new data from available resources. Cubic spline interpolation offers a simple yet effective solution for data augmentation⁷¹. This interpolation method is used to compress or stretch the extracted RGB traces and corresponding cPPG signals. Thus, multiple synthetic samples with different HR values can be generated. Similarly, Li et al.⁷² and Vance et al.⁷³ directly synthesize video frames and cPPG or HR signals through window stretching and shrinkage. In order to tackle varying illumination conditions, Yang et al.⁴⁶ proposed generating augmented lighting variation data during training, followed by image enhancement strategies. Moreover, to improve algorithm adaptability to diverse motion types, motion transfer techniques have been introduced⁷⁴. This approach embeds synthetic PPG signals within spatial masks, upon which first-order motion models are applied, thereby infusing motion into static facial videos. In addition to environmental illumination and subject movement, skin tone diversity is another critical factor^{75–77}. To address this, a joint optimization framework was designed to generate facial videos with various skin tones⁷⁸. This bio-realistic skin tone translation mechanism helps mitigate imbalances in skin color representation. However, recent studies have highlighted that synthetic data generation may introduce extraneous artifacts, leading to noticeable distribution shifts between synthetic and real-world data. Interpolation methods used for HR variation can generate noise with non-physiological features that might be erroneously captured by DL models⁷¹. For motion transfer techniques, the problem of the synthetic-to-real gap was also identified, where visual inconsistencies or boundary artifacts manifest in synthetic data, deviating from authentic video distributions⁷⁴. Furthermore, skin-tone translation mechanisms have the potential to introduce visual biases that compromise the integrity of the underlying physiological signals, thereby undermining the reliability and realism of the augmented datasets⁷⁸. To circumvent this issue, an alternative approach is to use labeled real-world images for pre-training⁷⁹. Particularly for Transformer-based architectures, which are extremely data-hungry, pre-training on a large-scale image dataset has been shown to help enhance the generalization ability⁸⁰.

Advanced techniques for algorithm generalization improvement.

The development of DL technologies has also enabled researchers to apply more advanced methods to leverage available real-world data. Gao et al.⁸¹ proposed an instance-based deep transfer learning technique that improves the generalization performance of HR estimation algorithms by detecting and emphasizing high-quality cPPG signals during the training process. Domain adaptation is another essential branch of transfer learning. Du et al.⁸² introduced a domain adaptation method based on a dual-bridging structure, which can generate spatiotemporal maps with different types of random noise. Through adversarial training, their method enables the algorithm to adapt to diverse noise patterns, thereby promoting its robustness. To further address the challenge of domain shifts during inference, Lee et al.⁸³ developed a continual test-time adaptation (TTA) method, which updates only the batch normalization (BN) layers and uses a non-contrastive unsupervised learning loss in the frequency domain to improve algorithm stability against catastrophic forgetting.

Skin detection methods

Before extracting RGB signals or other relevant features, skin detection is an essential step that can significantly affect the results, particularly in non-learning-based algorithms. The importance of automatic skin detection is

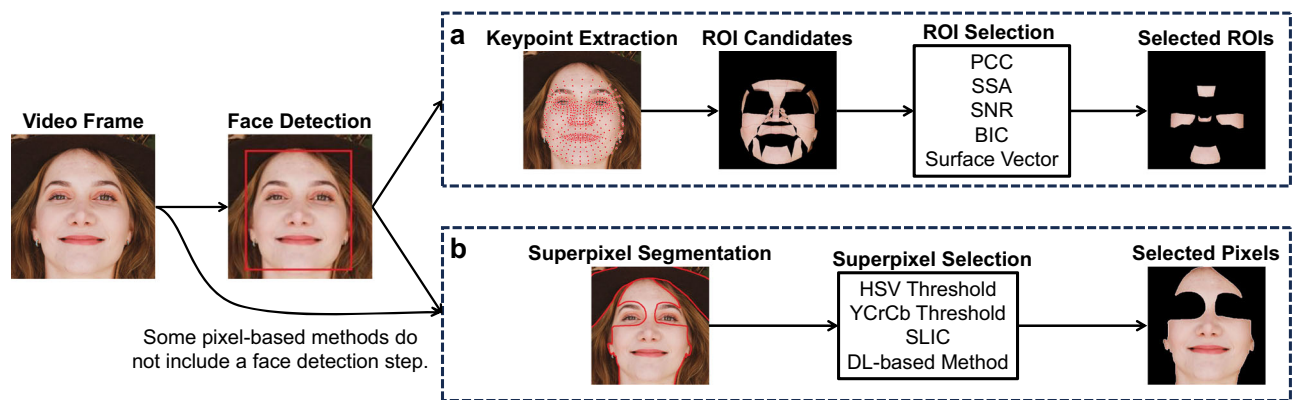


Fig. 4 | Two categories of skin detection methods for rPPG-based HR measurement. **a** Keypoint-based methods. Given the input frame, the face of the subject is detected, and the corresponding facial keypoints are extracted. ROI candidates are divided according to keypoint locations. **b** Pixel-based methods. The face detection and keypoint extraction steps are skipped. Instead, the input frame is directly

segmented into several superpixels. In both methods, a subsequent ROI selection process can be applied to identify high-quality skin regions. The real human face image was sourced from the Real and Fake Face Detection Dataset (<https://www.kaggle.com/datasets/ciplab/real-and-fake-face-detection>), which is under the CC BY-NC-SA 4.0 License (<https://creativecommons.org/licenses/by-nc-sa/4.0>).

further highlighted in challenging scenarios involving low-light conditions, subject movement, or facial occlusions. Several papers have underscored that segmenting facial skin and excluding non-skin pixels can improve the rPPG signal quality^{84,85}. An effective skin detection method can identify high-quality regions that are critical for accurate HR estimation. According to the reviewed studies, skin detection methods can be categorized into two main types: keypoint-based methods and pixel-based methods. The principles and workflows of these methods are shown in Fig. 4.

Keypoint-based methods. Keypoint-based methods generally encompass two main steps: face detection and keypoint extraction. The face detection step involves locating the subject's face and generating a bounding box for each frame. High-precision localization is important to suppress motion artifacts. The subsequent keypoint extraction step aims to segment and track facial ROIs with greater precision. In addition to accuracy, detection speed is also an important factor. Detection speed can directly influence the real-time performance during measurement, which is a basic requirement for practical applications. The Viola-Jones face detector^{86,87} is widely used in rPPG pipelines due to its recognized real-time performance. It utilizes Haar-like features and a cascaded classifier to efficiently detect human faces. Tracking algorithms such as the KLT tracker⁸⁸, Kalman filter^{89,90}, and MOSSE tracker⁹¹ are typically applied for continuous localization, and the stability of the tracking algorithm has been emphasized⁹². Multiple algorithms and packages are available for keypoint extraction. Classical feature point extractors (e.g., Shi-Tomasi corner detector⁹³ and SURF⁹⁴) were applied in earlier studies. However, such types of keypoints are not specifically designed for facial features. A multi-task convolutional neural network (MTCNN)⁹⁵ was utilized in several of the included articles for face detection and keypoint extraction^{78,96–99}. This model can help deal with the problem of occlusion and motion variations, which often cause failures when using the Viola-Jones face detector. MTCNN can extract five facial landmarks that aid in locating different ROIs. Some open-source libraries specifically built for face detection can extract more facial landmarks and thus have been intensively used in recent studies. Dlib¹⁰⁰, developed in C++, is a Python API for cross-platform development. The facial landmark extractor of dlib defines 68 keypoints, providing an improved ROI segmentation accuracy. Similar APIs, including FaceReader¹⁰¹, SeetaFace¹⁰², OMRON OKAO¹⁰³, CLNF OpenFace¹⁰⁴, and MediaPipe Face Mesh¹⁰⁵, were also employed in some papers. Among these APIs, MediaPipe Face Mesh offers 468 pre-defined facial landmarks and real-time detection speed, thus providing a representative face detection framework for rPPG applications. A comparative study¹⁰⁶ has shown that the accuracy,

processing speed, and flexibility of MediaPipe Face Mesh have all surpassed those of dlib, Viola-Jones, and MTCNN methods.

Many of the collected articles incorporated an ROI selection step after detecting ROI candidates to evaluate the signal quality of each candidate ROI and select one or more high-quality ROIs for extracting more informative BVP signals. Pixel-wise filtering offers a relatively straightforward approach. It involves computing the mean and standard deviation for pixel values within each ROI and subsequently removing outlier pixels that significantly deviate from the mean values³⁸. Kossack et al.¹⁰⁷ proposed a reliability score based on the Pearson correlation coefficient (PCC) and the maximum magnitude in the frequency domain. They calculated the reliability of each ROI and selected the one with the highest score for BVP extraction. Wang¹⁰⁸ ranked the candidate ROIs using a quality metric based on singular spectrum analysis (SSA). He recommended selecting one-half to one-fourth of the candidate patches. In addition, SNR, which is a commonly used measurement of signal quality, was also utilized for ROI selection^{109–112}. Nevertheless, the SNR calculation always includes an FFT process with a minimum complexity of $O(n \log n)$. Considering the high computing cost, the Bayesian information criterion (BIC) was introduced as a replacement¹¹³. With a complexity of $O(n)$, the BIC metric was shown to be faster and more resilient to subject movement, thereby providing superior selection capability. In a quantitative comparison of commonly used SQIs, the results evaluated on the PURE and LGI-PPGI datasets showed that SNR can more accurately detect high-quality cardiac signals (with a suggested threshold of 0.293), while kurtosis is more effective in distinguishing between acceptable and unfit signals²⁴. Rather than using a single pre-defined ROI, Boccignone et al. developed a method that improves ROI selection by clustering signals from multiple facial regions based on their spectral characteristics¹¹⁴. In the work of Zhang et al.¹¹⁵ and Zhao et al.¹¹⁶, the advantage of using multiple ROIs was also highlighted. At the physical modeling level, a signal quality metric based on surface normal vectors in 3D facial landmarks was developed¹¹⁷. This metric was inspired by the theory that the surface orientation of keypoints can significantly affect the ratio between specular reflection and diffuse reflection, and that only diffuse reflection carries physiological information. Their results showed that a small reflection angle can provide a higher rPPG signal quality. The experiments of Cantrill et al. also supported this phenomenon¹¹⁸. However, Fan et al.¹¹⁹ collected facial videos of different reflection angles and found that the highest HR estimation accuracy was at 60 degrees instead of 0 degrees. This suggests that an optimal range of reflection angles may exist, which should handle a tradeoff between reducing specular reflection and keeping enough diffuse reflection to capture BVP variations.

Table 1 | Overview of model-driven rPPG methods for HR measurement

Author (year)	Basic Principle
Verkruysse et al. ¹⁶	Spatial averaging, digital filtering, and spectral analysis
Poh et al. ¹³⁷	ICA
Lewandowska et al. ¹⁴⁰	PCA
de Haan et al. ¹⁴⁴	Chrominance-based algorithm to eliminate specular reflection artifacts.
de Haan et al. ¹⁴⁵	Linear combination of RGB traces, blood volume pulse signature
Feng et al. ⁸⁴	ICA with automatic sorting
Feng et al. ²⁴⁰	Optical skin modeling, adaptive green/red channel difference operation
Wang et al. ¹⁴¹	Spatial subspace estimation, temporal subspace rotation, eigen-decomposition of skin pixels, pulse extraction via rotation measurement
Wang et al. ¹⁴²	Dichromatic reflection model, projection onto skin-tone-orthogonal plane, alpha-tuning
Macwan et al. ^{138,139}	Periodicity constraint based on autocorrelation, chrominance constraint based on skin tone
Demirezen et al. ¹⁴⁹	Nonlinear mode decomposition
Pilz et al. ⁴²	Invariant feature extraction, Lie group of rigid transformations, stochastic modeling
Macwan et al. ¹²⁵	Multi-objective optimization based on ICA, autocorrelation, and negentropy
Zhou et al. ¹⁴⁸	Signal spectral characteristics analysis, optimal projection vector, constraint plane
Song et al. ¹⁵⁰	Empirical mode decomposition, multiset canonical correlation analysis
Álvarez Casado et al. ¹⁵²	Reduced QR factorization, Householder reflections
Haugg et al. ¹³⁶	Linear combination of RGB traces

The articles are sorted according to the year of publication. The detailed information on research emphasis and principles is provided in Supplementary Information (Table S1). Abbreviations: *RGB* (red, green, and blue); *ICA* (independent component analysis); *PCA* (principal component analysis); *KLT* (Kanade–Lucas–Tomasi).

Pixel-based methods. In contrast to keypoint-based approaches, pixel-based methods do not rely on facial landmark localization. As shown in Fig. 4b, the algorithms can directly classify image pixels into skin pixels and non-skin pixels. Thus, the entire image is divided into several large regions called superpixels (groups of adjacent pixels in an image that share similar properties, like color, texture, and intensity). This division process automatically skips the face-detection step, thus simplifying the detection process. Among pixel-based methods, thresholding is the simplest approach to distinguish skin from non-skin pixels. Some studies have demonstrated the effectiveness of using thresholding in the hue, saturation, and value (HSV) color space, as well as in the luma, red-difference chroma, and blue-difference chroma (YCrCb) color space^{120,121}. In addition to thresholding, clustering methods can also be applied for superpixel segmentation. Jaiswal et al.¹²² and Li et al.¹²³ utilized the simple linear iterative clustering (SLIC) method to segment each frame into superpixels. In the method of Jaiswal et al., the least mean squares (LMS) filter was applied in parallel to suppress the influence of illumination variation. However, SLIC methods may impose large computational burdens. This problem can be mitigated by implicitly identifying the superpixel boundaries and only processing a fraction of the image pixels¹²⁴. Certain studies^{85,125–127} adopted Conaire’s method¹²⁸ for skin detection, which maximizes the agreement between independent data sources in HSV space iteratively. Moreover, DL-based semantic segmentation methods have demonstrated strong performance for selecting skin pixels^{72,129}. In contrast to traditional segmentation algorithms, DL-based methods exhibit greater adaptability within unconstrained environments and demonstrate enhanced flexibility.

Model-driven algorithms

Model-driven algorithms refer to those developed based on prior knowledge or deterministic signal processing techniques. Through explicit mathematical and physical modeling, these algorithms extract physiological signals depending on theoretical foundations such as optical properties of skin, color space variations, and blind source separation. As illustrated in Fig. 2, after detecting the ROIs, RGB signals—also referred to as RGB traces—are computed by averaging the selected pixels in each color channel. Subsequently, dimensionality reduction techniques are employed to transform RGB traces into BVP signals, which contain more robust and precise HR-

related information. Most model-driven rPPG methods aim to enhance the dimensionality reduction process, serving as fusion strategies that integrate HR-related signals from RGB channels. An ideal rPPG method should extract high-quality signals that can capture the cardiac activity at a deeper level while suppressing noise and artifacts¹³⁰. In this roadmap, we conducted a comprehensive survey of model-driven rPPG methods. The key characteristics are summarized in Table 1. The details of the algorithm principles are provided in Supplementary Information (Table S1).

The simplest and most straightforward approach is to select the channel that exhibits the highest similarity with ground truth cPPG signals. Prior investigations^{16,131–134} consistently highlighted the alignment of the G channel with PPG signals, due to its superior correspondence compared to the R and B channels. This conclusion is further supported by the anatomical fact that hemoglobin shows better absorption of green light than red light¹³⁵. Nonetheless, relying solely on the G channel may not be optimal, as the R and B channels also contain complementary information useful for BVP signal extraction and should be appropriately utilized⁸⁴. A simple and efficient ratio-based algorithm called GRGB rPPG was proposed by Haugg et al.¹³⁶. In this method, rPPG signals are estimated by taking the sum of the ratios of the green-to-red (GR) channel and the green-to-blue (GB) channel. This algorithm was demonstrated to be effective and computationally efficient. Additionally, as a classical blind source separation (BSS) technique, an independent component analysis (ICA) algorithm was applied to extract BVP signals from raw RGB traces¹³⁷, where the normalized traces are decomposed into three independent source signals. As a rule of thumb, it is recommended that the second component be extracted, which commonly reveals a dominant HR-related signal. The ICA method can be further improved by adding two priors: one based on periodicity using autocorrelation, and the other based on the chrominance (CHROM)-based constraint of skin pixels^{138,139}. Similarly, principal component analysis (PCA) has also been applied as a BSS technique¹⁴⁰. Unlike ICA, PCA does not assume statistical independence or any specific underlying structure in the original features. The experimental accuracy is comparable to ICA, with the advantage of lower computational complexity. A spatial subspace rotation (2SR) method was presented by Wang et al.¹⁴¹ without the requirement of skin-tone or pulse-related priors. It extracts pulse signals by estimating the spatial subspace of skin pixels and then measuring the temporal rotation. They also developed the plane-orthogonal-to-skin (POS)

method¹⁴², which employs a projection plane orthogonal to the skin tone for pulse signal extraction. Departing from popular interpretations based on optical and spectral analysis, the local group invariance (LGI) method was designed to utilize invariant features, rearranging and concentrating the BVP signals in vector space⁴². Beyond the RGB space, the exploration of alternative color spaces reveals the efficacy of the H channel in the HSV color space and the Q channel in the luma, in-phase, and quadrature (YIQ) color space¹⁴³.

Mitigating artifacts resulting from motion and illumination variability is crucial for the development of more robust rPPG algorithms. De Haan et al.¹⁴⁴ proposed a robust CHROM-based rPPG method to address motion-induced distortions frequently encountered in BSS techniques. The algorithm demonstrated consistent performance under varying measurement conditions. To continuously improve motion robustness, the PBV method¹⁴⁵ was introduced after one year, showing superior performance over CHROM across diverse datasets. This method assumes that pulse signals can be reconstructed by a linear combination of normalized RGB traces. It utilizes the relative pulsatile properties of the color channels to extract HR-related signals and remove distortions. Feng et al.⁸⁴ devised an automatic selection procedure based on the ICA method to suppress motion artifacts. Rather than picking the second component by default, the independent component with the highest correlation to the reference sine function was recommended. Various studies proposed strategies to estimate subject movement and cancel out motion-induced artifacts. These motion compensation techniques include using facial optical flows¹⁴⁶ or tracking non-skin pixels that move in sync with skin pixels¹⁴⁷. Some researchers focused on refining existing algorithms by introducing a priori knowledge. For instance, the projection vector based on the spectral characteristics (PSC) algorithm¹⁴⁸, inspired by the POS algorithm, constrained the projection vector using channel signal variance, demonstrating improved robustness under strong disturbance cases. Macwan et al.¹²⁵ improved the ICA method by introducing multi-objective optimization with auto-correlation. The proposed method, MAICA, has shown outstanding performance in the presence of motion artifacts. ICA-based algorithms have also been improved by incorporating nonlinear components. In the nonlinear mode decomposition (NMD) technique, motion artifacts are suppressed by leveraging its inherent noise-robust characteristics¹⁴⁹. Focusing on uneven lighting interference, Song et al.¹⁵⁰ developed the ensemble empirical mode decomposition with multiset canonical correlation analysis (EEMD-MCCA) method. The EEMD is employed to decompose the G channel, and the MCCA is used to determine the underlying candidate pulses. A retinex-based method was employed to enhance the image quality of frames captured in dark illuminance environments. The experimental results have quantitatively confirmed that this image enhancement technique can improve the performance of common rPPG algorithms. Moreover, the study highlighted the potential reliability of the GREEN and ICA methods over POS under inadequate illumination conditions.

Besides noise caused by motion and illumination, video compression artifacts have been identified as another influential factor in rPPG algorithms. The video compression artifacts commonly appear in three forms: amplitude degradation, structured high-frequency noise, and trace interruptions¹⁵¹. To address these challenges, a framework based on SSA was devised by the researchers. This framework comprises three essential steps: single-channel band-pass filtering, SSA decomposition and reconstructed component (RC) selection, and spectral masking refinement. Furthermore, the orthogonal matrix image transformation (OMIT) method, which is specially tailored for heavily compressed videos, was demonstrated to be more robust than the GREEN and POS methods¹⁵². The OMIT algorithm extracts an orthogonal matrix with uncorrelated components. To improve computational efficiency, it employs thin QR factorization in place of full QR factorization. Additionally, Eulerian video magnification¹⁵³ can be applied to attenuate video compression artifacts and enhance imperceptible signals.

Data-driven algorithms

For data-driven algorithms, considerable amounts of data are generally required for the optimization of learnable parameters. Instead of relying on explicit signal modeling, ML and DL techniques are commonly leveraged to automatically obtain the mapping relationship between the input raw data and the output physiological signals. With the rapid advancement of artificial intelligence (AI) and associated hardware capacity, increasing attention has been paid toward data-driven rPPG methods. Compared to conventional approaches, data-driven methods demonstrate superior capabilities for extracting higher-order features and offer promising generalization ability across diverse application scenarios. These algorithms can be categorized into two major types based on their integration stage within the processing pipeline. Figure 2 illustrates that end-to-end algorithms can directly estimate HR from input video sequences that are preprocessed under simple operations such as resizing, cropping, frame difference, warping, and color space transformation. Some algorithms first estimate the BVP signals and then derive HR values using frequency spectrum analysis or peak detection methods. In this roadmap, such algorithms are also classified as end-to-end approaches. The second type is referred to as hybrid methods, which apply ML or DL techniques at one or multiple stages within the workflow. For a comprehensive overview, the basic principles of data-driven algorithms are summarized in Table 2. The detailed descriptions are elaborated in Supplementary Information (Table S2).

End-to-end structure. DeepPhys is the first end-to-end deep CNN-based algorithm developed for HR measurement from facial videos without relying on explicit signal processing stages¹⁵⁴. This framework employs the VGG architecture and adopts the temporal difference between consecutive frames as input. By jointly training motion and appearance branches, DeepPhys achieved enhanced performance across different illumination conditions and subject motion types. However, the utilization of VGG-based architectures in real-time applications is hindered by their large number of parameters. To improve on-device performance, the multi-task temporal shift convolutional attention network (MTTS-CAN) architecture was proposed, which can be deployed on mobile platforms with a real-time inference speed¹⁵⁵. In a similar vein, Comas et al.⁹⁹ presented a more lightweight architecture inspired by a Taylor series expansion and also introduced the temporal adaptive location shift (TALOS) loss to learn latent shift probabilities of the ground truth during training. Their method achieved comparable accuracy to DeepPhys, while the model size is 200 times smaller. Beyond the aforementioned 2D-CNN structure, 3D-CNN frameworks inherently capture temporal characteristics of input videos and are thus more powerful for extracting spatiotemporal features^{156–158}. Nonetheless, 3D-CNNs are more computationally intensive due to their increased dimension, resulting in greater computational and memory demands, as well as the need for more training data. In response to this problem, an ultralight 3D-CNN architecture named real-time rPPG (RT-rPPG) was proposed, which has a substantially faster inference speed compared with other state-of-the-art 3D-CNN methods¹⁵⁹. Speth et al.¹⁶⁰ developed a weakly supervised method to train 3D-CNNs using HR labels without the need for synchronized rPPG signals as the ground truth.

Beyond considerations of model size and inference speed, denoising ability is a critical aspect of rPPG algorithms. Liu et al.¹⁶¹ applied an adversarial correlation analysis learning strategy based on the background region for noise removal. Yue et al.⁵³ aimed to improve HR estimation accuracy for low-resolution facial videos using a two-stage network. Their strategy involves first enhancing input video resolution using an rPPG information recovery network (rPPGRNet), followed by HR estimation via a temporally aware HR measurement network (THRNet). Illumination noise can also be suppressed via temporal normalization on RGB channels and direct current (DC) removal¹⁶². To address the issue of phase delay when using ground truth cPPG signals collected using fingertip or wristband oximeters, recently proposed methods^{163,164} often incorporate the spectrum

Table 2 | Overview of data-driven rPPG methods for HR measurement

Author (year)	Structure	Basic principle
Hsu et al. ¹⁷³	Hybrid	Support vector regression, multiple feature fusion, and temporal segment fusion
Chen et al. ¹⁵⁴	End-to-end	VGG-style CNN, motion representation based on the skin reflection model, and attention mechanism
Yu et al. ¹⁵⁶	End-to-end	3D-CNN with encoder-decoder structure
Hu et al. ⁹⁸	End-to-end	Convolutional LSTM network, attention mechanism for temporal correlation
Wu et al. ¹⁷⁷	Hybrid	Adaptive neural network model selection
Wu et al. ¹⁷⁶	Hybrid	Multi-channel convolution autoencoder
Liu et al. ¹⁵⁵	End-to-end	Temporal shift module, attention mechanism, dual-branch network, multi-task learning
Song et al. ²⁴¹	Hybrid	GAN for pulse wave form generation, adversarial loss, time-frequency domain loss
Hu et al. ²⁴²	Hybrid	Attention module and temporal fusion module
Comas et al. ⁹⁹	End-to-end	Temporal derivative module, temporal adaptive location shift loss, spatial and temporal modeling
Botina-Monsalve et al. ¹⁵⁹	End-to-end	3D-CNN, YUV color space, temporal-frequency loss
Lokendra et al. ¹⁷⁵	Hybrid	Delaunay triangulation, temporal convolution network, action units for facial expression analysis
Jaiswal et al. ¹⁸⁰	Hybrid	Fuse multi-scale retinex with RGB signals, CNN with the spatiotemporal map as the input
Sun et al. ¹⁸²	Hybrid	ViT framework, spatiotemporal feature map, and attention mechanism to reduce computational complexity
Jaiswal et al. ¹⁷⁹	Hybrid	CNN, adaptive color space transformation
Yu et al. ¹⁸³	Hybrid	Multi-scale spatiotemporal map, virtual PPG as the ground truth, Lightweight U-Net architecture
Lee et al. ²⁰¹	End-to-end	A mix of VGG-style and ResNet-style networks, a one-branch convolutional attention network, temporal shift module
Lee et al. ¹⁵⁸	End-to-end	Long short-term convolutional network, temporal attention refinement module, 3D-CNN hourglass structure, hybrid loss
Hansen et al. ¹⁵⁷	End-to-end	Combination of multiple facial ROIs, 3D-CNN
Das et al. ²⁴³	Hybrid	2D-CNN architecture, scalogram-based feature map using wavelet transformation
Yang et al. ²⁴⁴	End-to-end	Bayesian neural network, uncertainty-guided self-training, regularization based on periodic priors
Shao et al. ¹⁶⁷	End-to-end	ViT, temporal pooling attention, self-supervised masked autoencoder
Liu et al. ¹⁶¹	End-to-end	Adversarial canonical correlation analysis learning, spatiotemporal rPPG aggregation, self-supervised noise extraction from background
Liu et al. ¹⁸⁴	Hybrid	Self-supervised learning, masked autoencoder, noise-insensitive spatiotemporal map augmentation
Chowdhury et al. ¹⁷⁴	Hybrid	Encoder-decoder architecture, combination of U-Net and LinkNet
Gao et al. ⁸¹	Hybrid	Convolutional and deconvolutional-based signal quality assessment, instance-based deep transfer learning
Qian et al. ¹⁸⁵	Hybrid	Transformer-based architecture, dual-path framework, spatial and temporal TokenLearner
Speth et al. ¹⁶⁰	End-to-end	Weakly supervised learning, 3D-CNN, using heart rate as labels
Zhu et al. ²³¹	End-to-end	Identity-specific facial contours, temporal compressor unit, cross-task feature combiner, multi-task learning
Xiao et al. ¹⁶³	End-to-end	Fusing original and differential frames, 3D-CNN with encoder-decoder architecture, self-similarity loss
Xiong et al. ¹⁷¹	End-to-end	Spatiotemporal graph denoise block, attention masks
Shao et al. ¹⁶⁸	End-to-end	Video swin Transformer, encoder-decoder structure, temporal difference supervision
Ma et al. ¹⁷⁸	Hybrid	Multi-scale spatiotemporal map, channel-temporal attention module, ShuffleNetV2
Zhang et al. ¹⁷⁰	End-to-end	Mamba, Kolmogorov–Arnold network, hybrid Eulerian video magnification
Zhai et al. ¹⁶⁴	End-to-end	Feature space interpolation, high and low frequency feature fusion, adversarial learning
Salah et al. ¹⁶⁹	End-to-end	3D-CNN, video ViT, Siamese architecture, multi-head self-attention mechanism
Liu et al. ¹⁷²	End-to-end	Spiking neural network, Transformer, inter-frame difference

The articles are sorted according to the year of publication. The detailed information on research emphasis and principles is provided in Supplementary Information (Table S2). Abbreviations: *CNN* (convolutional neural network); *LSTM* (long short-term memory); *GAN* (generative adversarial network); *ViT* (vision Transformer); *PPG* (photoplethysmography); *ROI* (region of interest).

distribution similarity supervision mechanism to mitigate the negative impact of phase shift during training.

In addition to CNN architectures, researchers have explored other DL structures to capture contextual features in both spatial and temporal domains. Long short-term memory (LSTM) modules can be integrated into CNN-based methods, and a mask branch can be applied for noise suppression⁹⁸. With the emergence and rapid advancement of the Transformer architecture in vision tasks since 2020¹⁶⁵, Transformer-based rPPG algorithms were developed as powerful solutions for HR measurement^{166–169}. However, as a common problem for Transformer-based architectures, the inference speed often falls short of real-time requirements. Recent studies introduced computationally efficient algorithm architectures or modules such as Mamba, Kolmogorov–Arnold network (KAN)¹⁷⁰, graph neural

network (GNN)¹⁷¹, and spiking neural network (SNN)¹⁷² to address this issue while maintaining superior noise suppression capability during measurement.

Hybrid structure. In hybrid methods, structures with learnable parameters are embedded into one or more specific stages of the rPPG workflow. In 2014, Hsu et al.¹⁷³ first introduced the concept of estimating HR using ML algorithms. They used a support vector regression (SVR) algorithm to leverage mid-level features and designed fusion strategies to exploit both feature and temporal redundancies. Chowdhury et al.¹⁷⁴ integrated ML techniques into the post-processing step for rPPG signal refinement. They employed the LGI algorithm for raw signal extraction and a segmentation technique to obtain the final

Table 3 | Summary of commonly used evaluation metrics for rPPG-based HR measurement

Metric	Signal type	Formula
Mean absolute error (MAE)	HR	$MAE = \frac{1}{m} \sum_{i=1}^m \hat{h}_i - h_i $
Root mean square error (RMSE)	HR	$RMSE = \frac{1}{m} \sqrt{\sum_{i=1}^m (\hat{h}_i - h_i)^2}$
Mean absolute percentage error (MAPE)	HR	$MAPE = \frac{100\%}{m} \sum_{i=1}^m \left \frac{\hat{h}_i - h_i}{h_i} \right $
Bland-Altman plot	HR	$AVG_i = \frac{\hat{h}_i + h_i}{2}, DIFF_i = \hat{h}_i - h_i$
Pearson correlation coefficient (PCC)	HR/ PPG	$PCC = \frac{m \sum_{i=1}^m x_i y_i - \sum_{i=1}^m x_i \sum_{i=1}^m y_i}{\sqrt{m \sum_{i=1}^m x_i^2 - (\sum_{i=1}^m x_i)^2} \sqrt{m \sum_{i=1}^m y_i^2 - (\sum_{i=1}^m y_i)^2}}$
Concordance correlation coefficient (CCC)	HR/ PPG	$CCC = \frac{2 \sum_{i=1}^m (x_i - \bar{x})(y_i - \bar{y})}{(m-1)(\bar{x} - \bar{y})^2 + \sum_{i=1}^m (x_i - \bar{x})^2 + \sum_{i=1}^m (y_i - \bar{y})^2}$
Cosine similarity	HR/ PPG	$DIST(x, y) = \cos(\theta) = \frac{\sum_{i=1}^m x_i y_i}{\sqrt{\sum_{i=1}^m x_i^2} \sqrt{\sum_{i=1}^m y_i^2}}$
Signal-to-noise ratio (SNR)	PPG	$SNR = \frac{1}{m} \sum_{i=1}^m 10 \log_{10} \left(\frac{\sum_{v \in \mathcal{V}} (u(v) F\{x_i\} F\{v\})^2}{\sum_{v \in \mathcal{V}} (1 - u(v)) F\{x_i\} F\{v\})^2} \right)$
Maximum cross-correlation (MCC)	PPG	$MCC = \max \left(\frac{F^{-1}\{BPass\{F\{x\} F\{y\}\}\}}{(\sum_{i=1}^m (x_i - \bar{x})) \cdot (\sum_{i=1}^m (y_i - \bar{y}))} \right)$

In the formulas, m represents the total number of samples, and i represents the i -th time window. $F\{\}$ and $F^{-1}\{\}$ represent FFT and inverse FFT operations, respectively. $u(v)$ is a binary mask in the frequency domain that selects the power contained within the pre-defined BPM range around the reference HR and its first harmonic. $\hat{h}$ and h represent estimated HR values and reference HR values. Similarly, x and y represent estimated and reference signals (e.g., PPG or HR time series, depending on the evaluation settings). $BPass\{\}$ represents a band-pass filter operation.

rPPG estimation. To mitigate the influence of facial expressions, which are recognized as the predominant source of temporal signal noise, the AND-rPPG approach was developed¹⁷⁵. This method integrates action units (AUs) for facial expression analysis and temporal convolution networks (TCNs) for denoising distinct facial regions. For in-vehicle monitoring scenarios with luminance variations, Wu et al.¹⁷⁶ proposed two ML-based approaches: the convolution autoencoder (CAE) and the multi-channel convolution autoencoder (Multi-CAE). These methods first transformed input videos into spectrograms via the short-time Fourier transform (STFT) and then processed spectrograms using the CAE modules for noise mitigation. In another paper, they developed an adaptive neural network model selection (ANNMS) method by personalizing for different drivers using an automatic model selection mechanism¹⁷⁷.

Beyond the RGB space, transforming RGB traces into other color spaces has shown tangible benefits¹⁷⁸. An adaptive color space transformation (CoST) layer featuring learnable parameters for color channel transformation offers a viable alternative¹⁷⁹. The pre-trained CoST layer transforms RGB traces into an optimal color space and thus provides more informative signals under various lighting conditions. To eliminate the effect of illumination variations, Jaiswal et al. incorporated multi-scale retinex (MSR) as a complementary color space alongside RGB traces¹⁸⁰. Subsequent to this preprocessing step, the proposed rPPG-FuseNet takes the fused spatiotemporal feature map as input and estimates the HR values. The retinex theory can also be combined with DL architectures to estimate the illumination and reflectance components¹⁸¹. In fact, it is a common approach for many data-driven algorithms of hybrid structure to first transform facial video sequences into different kinds of spatiotemporal feature maps as the input of subsequent networks^{182–185}.

Filtering techniques

The choice of filters plays a crucial role in enhancing the quality of extracted BVP signals in rPPG-based HR measurement and thus requires careful consideration. They can be applied either as pre-filtering to the RGB traces or as post-filtering to the estimated rPPG signals. Filters used in rPPG algorithms are typically designed based on prior knowledge of HR frequency ranges and are employed to suppress noise caused by subject movement and complex lighting conditions.

In our roadmap, band-pass Butterworth filters of different orders were found to be the most commonly used. The Butterworth filter¹⁸⁶ is a type of linear filter recognized for its smooth frequency response and efficient processing speed. It is designed to offer a maximally flat frequency response within the passband. This frequency characteristic helps minimize amplitude and phase fluctuations. According to a recent comparative analysis of different common unsophisticated filters, Butterworth filters may be particularly suitable for rPPG methods compared to cPPG approaches¹⁸⁷. Apart from Butterworth, elliptic, and Chebyshev filters, with their sharp transition properties, yield superior SNR scores. In contrast, window-based filters, such as moving averages, median filters, and Hamming windows, tend to perform less effectively, especially when processing low-quality raw signals. Their experiments also highlighted the impact of filter order. In complex scenarios requiring dynamic adjustment of filter parameters, the infinite impulse response (IIR) filter was recommended for robust rPPG-based HR estimation¹⁸⁸. Raw RGB traces or estimated rPPG signals without post-filtering often exhibit baseline drifts and low-frequency fluctuations. Tarvainen et al.¹⁸⁹ designed the detrending filter to eliminate long-term variations that are unrelated to cardiac activities, such as the sensor time drift and illumination changes. This approach is valued for its ability to efficiently suppress distortion effects.

Several studies have proposed more advanced filtering methods. A color-distortion filter (CDF) was designed based on the reflection spectra differences between pulsatile and non-pulsatile components¹²¹. In challenging fitness scenarios, the CDF outperformed the band-pass filter (BPF), confirming its superior performance. The CDF can also be integrated with an amplitude-selective filter (ASF) to jointly process RGB traces¹⁹⁰. The ASF can serve to select RGB signals containing HR information, while the CDF removes color distortion. Beyond temporal domain filtering, filter banks might be more effective for processing frequency components⁹⁷. Different frequency components can be adaptively adjusted by constructing a set of exponential smoothing filters with various frequency bins. Moreover, wavelet transform techniques have also exhibited application values to suppress motion artifacts and other types of noise¹⁹¹. In data-driven approaches, it is recommended that filters be embedded within the algorithm architecture, serving as learnable modules. For instance, multilayer perceptron (MLP)⁹⁶, 1D-CNN¹⁹², and LSTM structures¹²⁷ can all serve as filters with learnable parameters.

Evaluation metrics

Assessing the performance of rPPG methods is essential for evaluating measurement reliability and facilitating quantitative comparisons. A variety of evaluation metrics have been adopted to assess the HR estimation accuracy of rPPG methods. These metrics can generally be categorized into two groups. One group directly compares the estimated HR with the ground truth HR. The other group measures the dissimilarity between rPPG signals and ground truth cPPG signals. The commonly employed evaluation metrics are summarized in Table 3.

The mean absolute error (MAE), also known as the difference in beats per minute (Δ BPM), and the root mean square error (RMSE) are among the metrics most frequently used in HR values. These metrics quantify overall prediction errors, with lower values indicating higher measurement accuracy. To provide more statistical insight, standard deviations (STDs) of these metrics are often reported alongside them to establish confidence intervals for BPM estimates. In comparison, the mean absolute percentage error (MAPE) metric provides a normalized and dimensionless metric to reflect algorithm performance across different HR ranges and subjects. In some

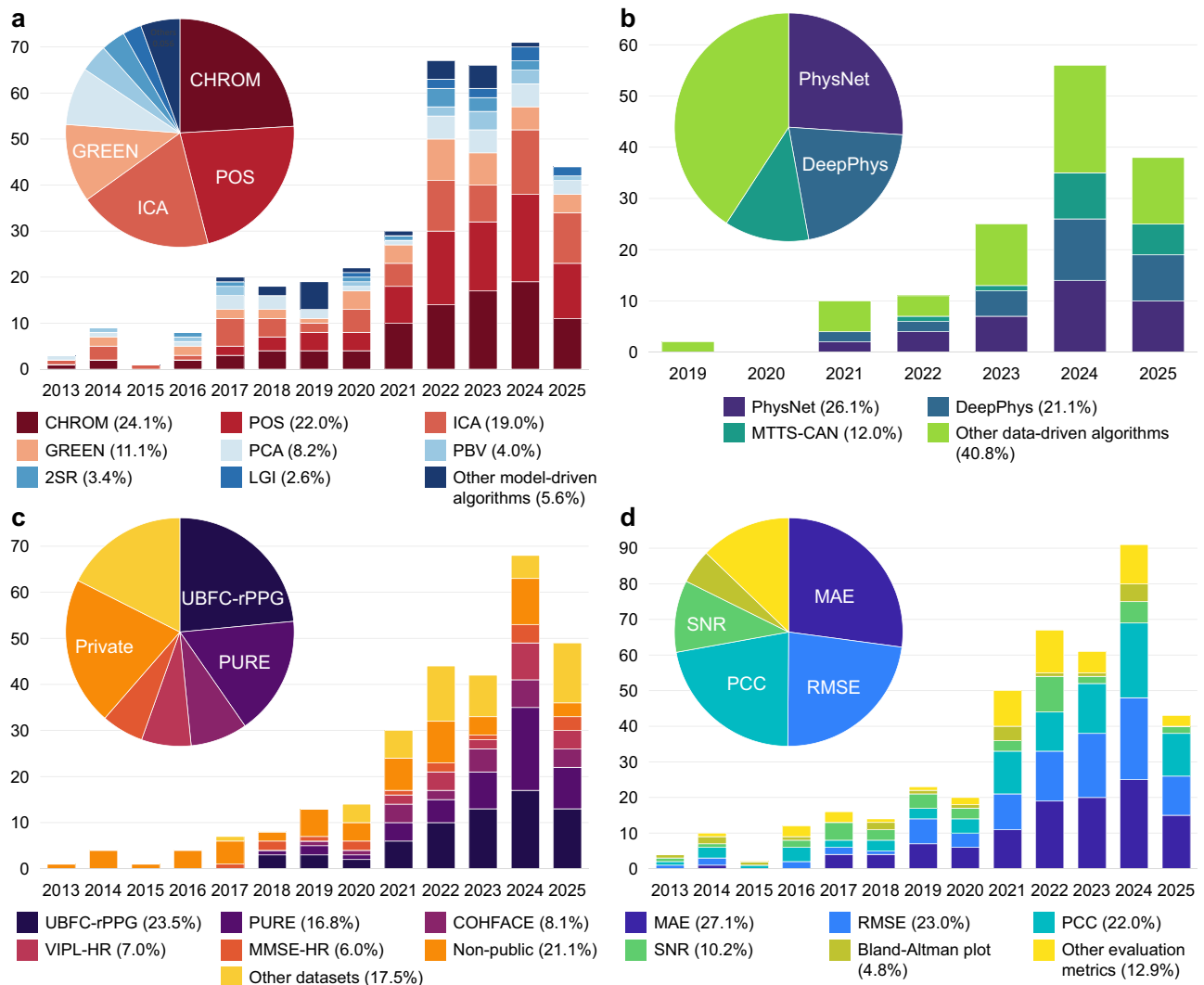


Fig. 5 | Statistics of usage frequency. The statistics are conducted based on all included articles during the literature search. Usage count less than 10 are categorized as “Others”. **a** Model-driven rPPG algorithms: CHROM¹⁴⁴, POS¹⁴², ICA¹³⁷, GREEN¹⁶, PCA¹⁴⁰, PBV¹⁴⁵, 2SR¹⁴¹, LGI⁴², and others. **b** Data-driven rPPG algorithms:

PhysNet¹⁵⁶, DeepPhys¹⁵⁴, MTTs-CAN¹⁵⁵, and others. **c** Datasets for algorithm performance evaluation: UBFC-rPPG⁴¹, PURE³⁷, COHFACE⁴⁰, VIPL-HR⁴³, MMSE-HR³⁹, non-public datasets, and others. **d** Performance evaluation metrics: MAE, RMSE, PCC, SNR, Bland-Altman plot, and others.

studies, researchers set a threshold for MAPE to indicate the practical applicability of the proposed algorithms. Additionally, the Bland–Altman analysis is commonly used as a visualization tool to assess the measurement agreement between the proposed algorithm and standard instruments.

Besides these error-based metrics, correlation-based metrics have also been leveraged for evaluating HR estimation performance. In the pyVHR framework^{193,194}, Pearson correlation coefficient (PCC) and concordance correlation coefficient (CCC) are recommended to measure the agreement between estimated and reference HR values. Template matching correlation (TMC) has been proposed to measure correlation in a fragmentary manner. Dividing the complete signal into multiple equal-length segments allows detection of regularity in a time window, thus mitigating the impact of abnormal values at specific time points. These metrics provide complementary perspectives by assessing the trajectory agreement and are particularly useful for reflecting measurement error in HR dynamics under large-scale variation. It is worth noting that PCC, CCC, and cosine similarity are sometimes leveraged to capture the discrepancies between estimated rPPG and ground truth cPPG signals, in order to provide a PPG signal-level performance reflection. Maximum cross-correlation (MCC) has also been introduced as both a loss function and evaluation metric for rPPG-supervised learning¹⁹⁵. Integrating MCC as a term of the loss function can

enhance its robustness against random temporal offsets in the ground truth. FFT-based techniques are often used to compute MCC more efficiently. In addition to formulated metrics, DL-based approaches have also been explored to quantify rPPG signal quality. Song et al.¹⁹⁶ proposed and compared both Bayesian neural networks (BNNs) and deep ensemble methods to evaluate the uncertainty of rPPG signals in HR estimation. These methods can be useful for providing confidence estimates in reference-free real-world applications. To complement these metric-based evaluations, we include a brief illustration in the following section on clinical translation.

Statistical data synthesis

In Fig. 5, we present the overall and year-wise usage frequency analysis regarding model-driven algorithms, data-driven algorithms, datasets, and evaluation metrics based on the results of the literature search. A clear rising tendency in research activity related to rPPG-based HR measurement over the years is evident, reflecting the increasing attention from the research community. Among model-driven algorithms, CHROM, POS, and ICA are the most frequently adopted, while PhysNet and DeepPhys are the mainstream representatives of data-driven approaches. Despite being proposed for a certain period of time, they still account for a significant proportion, even in studies of recent years. Notably, data-driven algorithms with usage

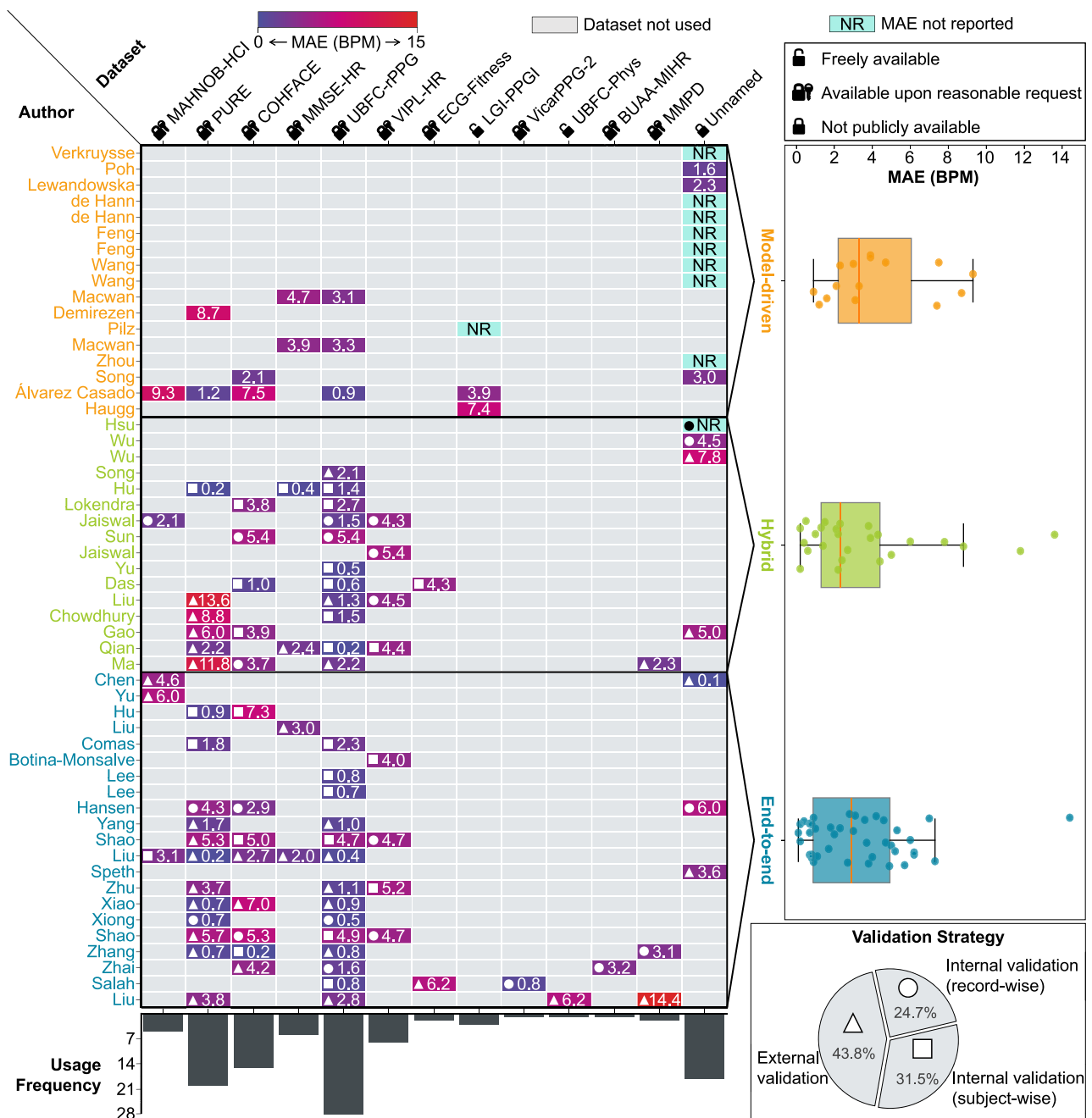


Fig. 6 | Performance statistics of rPPG algorithms across different datasets. The evaluation metric is mean absolute error (MAE). Each rPPG algorithm is labeled using the corresponding author's last name, and the order is consistent with Tables 1 and 2. The algorithms are categorized into three types: model-driven, hybrid, and end-to-end. For data-driven approaches (i.e., hybrid and end-to-end), the employed validation strategies are annotated using different symbols, defined in the order of evaluation strictness: external validation, where training and test data are sampled from different datasets; internal subject-wise validation, where the original dataset is split into training and test sets such that data from the same subject do not appear in both training and test sets; and internal record-wise validation, where training and

test sets may contain data sequences from the same subjects or even the overlapping recording sessions. Boxplots on the right display the MAE distributions of the three algorithm categories. Notably, internal record-wise validation is susceptible to performance overestimation due to subject overlap between training and test sets. To ensure a rigorous performance assessment, evaluation results obtained from internal record-wise validation were strictly excluded from the right-side boxplots. At the bottom, the dataset usage frequency across all evaluated algorithms is illustrated with barplots. Icons were adapted from the Microsoft Fluent Emoji project (<https://github.com/microsoft/fluentui-emoji>), under the MIT License (<https://opensource.org/licenses/mit>).

frequencies below 10 account for 40.8%, demonstrating the algorithm diversity and rapid evolution of this category. There are 23.5% papers using the relatively simple UBFC-rPPG dataset and 21.1% using private datasets. This highlights the need for developing more challenging and open-source datasets in future research. In terms of evaluation metrics, MAE, RMSE, as well as PCC and SNR, are the most frequently used HR-based and PPG-

based metrics, respectively. These should serve as references for future algorithm development and assessment.

According to the differences in principle, rPPG algorithms can generally be categorized into two types: model-driven and data-driven. Data-driven algorithms can be further divided into end-to-end and hybrid structures. In Fig. 6, we summarize the algorithm performances across

different datasets and display the group-wise performance comparison. It can be observed that although algorithms of end-to-end structure are in the majority of data-driven algorithms, the development of hybrid structure has persisted. The boxplot results show that there are no significant performance differences among the three types of algorithm structures. Considering that the gold standard for consumer-grade devices in MAE is 5 BPM¹⁹⁷, these three types of algorithm structures are all able to achieve acceptable measurement accuracy. Additionally, the UBFC-rPPG, PURE, and COHFACE datasets are the most commonly used public datasets for rPPG-based HR estimation algorithm development. This statistical result is generally consistent with the overall dataset usage statistics derived from the complete results of our literature search, as shown in Fig. 5. For data-driven algorithms, 43.8% of the relevant studies utilized the most rigorous external validation, and 31.5% adopted subject-wise internal validation. Of note, 24.7% of studies used internal record-wise validation. This approach is highly susceptible to data leakage, as subject-specific features from the same individual may appear in both training and test sets. Consequently, these studies may report inflated MAE performances. To ensure a rigorous comparison of actual algorithm capabilities, these potentially biased results were strictly excluded from the aggregated statistics in the right-side boxplots of Fig. 6.

Translation in clinical settings

According to our literature search, only three articles (3/151, 1.99%) specifically addressed clinically related topics. Among them, two focused on paediatric care and one on long-term, non-contact monitoring during liquid bolus administration. Hatib et al.¹⁹⁸ conducted a two-phase validation study in children and neonates. Their research suggested that the measurement results obtained on children under 10 years old (including newborns) did not reach clinically acceptable accuracy; by contrast, older paediatric cohorts exhibited higher measurement reliability, with the subgroup of 12–16 years old achieving a Spearman correlation of 0.82 with reference measures of HR recordings. The study conducted by Wegerif et al.¹⁹⁹ presented a similar phenomenon of age-dependent performance characteristics in paediatric care settings. They collected facial videos from 346 children across diverse health conditions and applied the POS algorithm¹⁴² for rPPG signal extraction. Their method achieved RMSEs of 15.4 BPM for subjects below 5 years old, 7.5 BPM for those aged 5–11 years, and 5.4 BPM in adolescents older than 11 years.

To bridge the performance gap in younger cohorts, these studies proposed concrete data collection and algorithmic strategies. These included the implementation of interactive gamification (e.g., animations to encourage the compliance of children), specialized overhead recording hardware setups for supine neonates, and explicit privacy/safety protocols (e.g., offline data processing and addressing concerns regarding infrared radiation). At the algorithmic level, optimizations included recommending the mid-face as a superior ROI over the forehead, expanding the HR search space up to 200 BPM, and redesigning screening filters to mitigate systematic bias arising from the mismatch between the adult and paediatric distributions.

Beyond paediatrics, Momeni et al.²⁰⁰ reported the first feasibility exploration of rPPG-based cardiac activity monitoring during long-term fluid bolus administration. Their developed pipeline incorporated MTCNN⁹⁵ for face and keypoint detection, with the forehead and cheeks selected as ROIs for RGB signal extraction, and applied both CHROM^{138,139} and POS algorithms for rPPG signal conversion. The HR measurement results showed their rPPG-based methods achieved MAEs of 4.27 BPM (CHROM) and 4.28 BPM (POS), which are close to clinical standards. Importantly, the measurement setup enabled synchronous monitoring of hemodynamic parameters, including stroke volume (SV) and systemic vascular resistance (SVR). The authors also underscored that the vasoconstriction induced by cold fluid administration (15 °C) can potentially result in increased measurement error compared with warm fluid administration (37 °C).

To illustrate the capability of video-based rPPG in capturing HR dynamics under clinically relevant stress conditions, we compared cPPG

and rPPG signals acquired during a lower body negative pressure (LBNP) protocol (Fig. 7). This protocol is used to simulate central hypovolemia and assess cardiovascular tolerance in human subjects.

In Fig. 7a, the rPPG closely tracks the reference cPPG, showing *strong agreement* across a broad HR range. The correlation-based metrics were exceptionally high (PCC = 99.85%, CCC = 99.85%, and cosine similarity = 99.99%), while amplitude-based metrics also confirmed high accuracy (MAE = 0.60 BPM, RMSE = 1.16 BPM, MAPE = 0.76%). The Bland–Altman analysis indicated minimal bias and narrow limits of agreement, confirming that rPPG reliably captures rapid HR changes under varying physiological conditions.

In Fig. 7b, the comparison demonstrates *good agreement* over a narrower HR range. The correlation remained high (PCC = 92.05%, CCC = 91.97%, cosine = 99.98%), with low absolute errors (MAE = 0.54 BPM, RMSE = 1.02 BPM, MAPE = 0.92%). The Bland–Altman plot again revealed minimal bias and tight dispersion, confirming that even under low-amplitude or motion-stable conditions, rPPG yields consistent HR estimates.

Together, these results illustrate the ability of the CHROM-based rPPG method to reproduce cPPG-based HR recordings across both dynamic and stable scenarios within this experimental setting. The combination of correlation, concordance, and absolute error metrics underscores the importance of a multi-metric evaluation framework: while correlation quantifies waveform similarity, absolute metrics (MAE, RMSE) capture physiological accuracy. When interpreted jointly, they provide a comprehensive depiction of agreement that is useful for methodological evaluation.

However, these strong agreement metrics were obtained under deliberately controlled experimental conditions—including fixed camera–subject geometry, disabled automatic camera settings, standardized cross-polarized illumination, supine positioning with minimal motion, and a predefined staged LBNP protocol in a homogeneous cohort of healthy volunteers—and should therefore be interpreted as upper-bound estimates under idealized conditions rather than representative of expected performance in unconstrained clinical or real-world environments.

This proof-of-feasibility demonstration is intended solely to illustrate the practical acquisition and signal-level agreement of facial rPPG recordings in a controlled clinical research environment. Given the limited sample size, these results should not be interpreted as evidence of clinical efficacy, diagnostic accuracy, or readiness for clinical deployment. Large-scale, prospective clinical studies across diverse patient populations are required to establish clinical validity and regulatory suitability.

Discussion

Despite rapid technical advances, direct clinical evidence for rPPG remains limited. Among the 151 reviewed studies, only a small subset explicitly involves clinical populations or clinically grounded endpoints. Consequently, the current body of evidence primarily supports technical feasibility rather than validated clinical utility. This roadmap therefore, emphasizes the identification of translational gaps and the methodological prerequisites required to progress from laboratory validation toward rigorous clinical evaluation.

In this paper, we conducted a literature review that focuses on rPPG techniques that use a single RGB camera to record facial videos for HR measurement. To gain a comprehensive understanding of this field, we examined research advances in eight components of the rPPG pipeline and conducted extensive searches on mainstream academic databases. Although rPPG technology has developed over just a decade, it has evolved from low initial accuracy to performance comparable to cPPG technology in simple scenarios, demonstrating its significant development potential and application prospects. Compared to ECG and cPPG, rPPG enables non-contact measurement, substantially reducing both measurement costs and operational complexity. However, the non-contact feature of rPPG technology also makes it susceptible to various noise sources, including complex lighting conditions, subject movement, and video compression artifacts. As mentioned in this article, the main challenges encountered in the

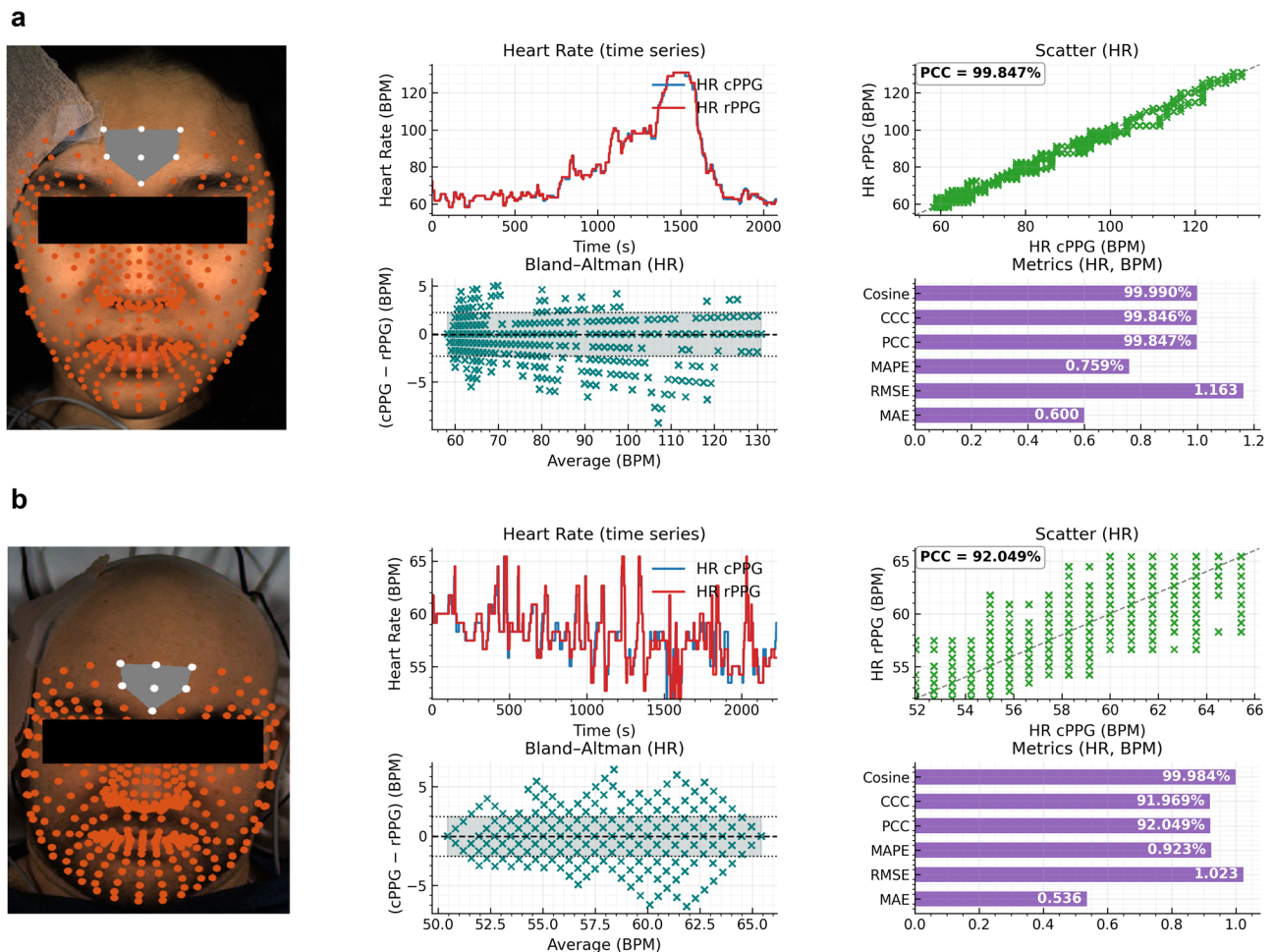


Fig. 7 | Agreement between camera-based rPPG and reference contact PPG for heart rate (HR). *Left column:* An anonymized video frame with MediaPipe facial landmarks (orange dots) and the forehead region of interest (ROI) used for rPPG extraction (gray polygon). The white dots indicate the landmarks defining the ROI boundary. *Middle column:* Synchronized HR time series from contact PPG (cPPG, blue) and rPPG (red). *Right column (top):* Scatter plot of HR_{rPPG} versus HR_{cPPG} with the identity line and annotated Pearson correlation coefficient (PCC). *Right column (bottom):* Bland-Altman plot of (cPPG - rPPG) HR with mean bias (dashed line)

and 95% limits of agreement (dotted lines). *Far right:* Bar summary of Cosine similarity, concordance correlation coefficient (CCC), PCC, mean absolute percentage error (MAPE, %), root-mean-square error (RMSE, BPM), and mean absolute error (MAE, BPM). rPPG signals were extracted using the chrominance (CHROM) method^{144,239} with cPPG as the reference. **a**, Example showing strong agreement across a wide HR range (PCC $\approx$ 99.85%, MAE $\approx$ 0.60 BPM, RMSE $\approx$ 1.16 BPM, MAPE $\approx$ 0.76%). **b** Example showing good agreement across a narrower HR range (PCC $\approx$ 92.05%, MAE $\approx$ 0.54 BPM, RMSE $\approx$ 1.02 BPM, MAPE $\approx$ 0.92%).

development of rPPG algorithms can be summarized in terms of four aspects: algorithm effectiveness, application environment, hardware restrictions, and insufficient data. Over the past decade, researchers have overcome many issues related to these challenges through a variety of approaches. Nonetheless, moving towards real-world applications will require continuous efforts from the research, industrial, and clinical communities.

In practical scenarios that involve deploying algorithms on hardware terminals, certain requirements for computation speed arise. This is particularly true in some use cases that have strict safety requirements, like automatic driving and intelligent surveillance. These applications often demand processing speeds at the millisecond level. However, a certain degree of latency is regarded as acceptable in the field of HR measurement; thus, the inference speed is less tightly constrained. Nonetheless, the real-time performance of rPPG algorithms still attracted significant attention from the research community. Some model-driven rPPG methods focus on reducing the operation time while maintaining measurement accuracy. According to the experimental results of Lewandowska et al.¹⁴⁰, it took 223 ms and 94 ms for the ICA algorithm to process the full-face ROI and the forehead ROI, respectively. Their proposed PCA algorithm achieved improved computing times of 1.4 ms and 1.2 ms. The OMIT algorithm also

achieved a computation speed of 17 ms per frame¹⁵². In terms of algorithmic complexity, the GRGB algorithm¹³⁶ with a time complexity of $O(1)$ outperforms the POS algorithm, which has a complexity of $O(n^2)$. These results indicate that model-driven rPPG algorithms have evolved in both measurement accuracy and computation speed, reaching millisecond-level performance.

Data-driven methods, especially those based on DL approaches, often require substantial computational resources due to their large number of parameters. Therefore, improving inference speed through lightweight architectures has become an important research direction. While most end-to-end architectures use 3D-CNNs to process facial videos, Comas et al.⁹⁹ employed a 2D-CNN as the spatial encoder. Their algorithm achieved an HR measurement accuracy comparable to that of PhysNet¹⁵⁶ while containing only 0.7% of its parameters. Lee et al.²⁰¹ compared the inference time and HR measurement accuracy of several data-driven methods. Their results show that DeepPhys¹⁵⁴, MTTs-CAN¹⁵⁵, PhysNet¹⁵⁶, EfficientPhys²⁰², as well as their proposed algorithm, were all able to maintain an MAE within 2.1 BPM while having an average computation speed of less than 70 ms per frame. These results indicate that state-of-the-art DL-based rPPG algorithms can meet the general requirement when adequate computational resources are available.

From the perspective of algorithm structure, end-to-end models directly take facial videos as input and output estimated rPPG signals or HR values, simplifying the pipeline and eliminating multiple intermediate steps. Such a design streamlines the process and reduces the complexity without the need to decompose the main task into multiple substeps. However, end-to-end algorithms typically involve a larger number of parameters, requiring substantial data and computational resources during both training and real-world deployment. Additionally, as “black boxes”, the interpretability of end-to-end algorithms is relatively poor, thus making it difficult to adapt performance under specific challenging conditions. To address this, future research should consider integrating attention-map visualization²⁰³ or physically-constrained techniques, ensuring model decisions are consistent with known physiological and optical principles. In contrast, a hybrid structure involves separate processing steps that can serve as structured regularization with prior knowledge, leading to improved interpretability and robustness. The research of Zhan et al.²⁰⁴ has demonstrated this point by incorporating priors of knowledge-based algorithms such as CHROM and POS into CNN architectures to enhance their robustness against motion artifacts. Moreover, the smaller size allows hybrid algorithms to achieve a faster training speed. Nevertheless, this may limit their model capabilities compared to end-to-end models given sufficient training data, especially as more well-established datasets have become available in recent years. On the other hand, most end-to-end models involve a large number of parameters and require parallel computing resources like graphics processing units (GPUs). In resource-constrained environments, deploying model-driven or hybrid algorithms might serve as a more realistic solution. Figure 7 provides a synthesis view of how HR agreement between camera-derived rPPG and reference contact PPG should be evaluated and reported. The four complementary analytical panels—time-series overlay, scatter plot, Bland–Altman comparison, and the multi-metric summary—together reveal dimensions of fidelity that no single statistic can capture. Correlation-based indices (PCC, CCC, cosine similarity) quantify linear and concordant association between HR estimates, but remain sensitive to polarity, scaling, and temporal misalignment. In contrast, magnitude-based metrics (MAE, RMSE) directly quantify absolute differences in BPM and are invariant to sign. The Bland–Altman analysis complements both by characterizing bias and dispersion across the HR operating range, thereby exposing systematic under- or over-estimation that may otherwise be hidden.

In Fig. 7a, rPPG achieves near-perfect correspondence with contact PPG (PCC = 99.85%, CCC = 99.85%, MAE = 0.60 BPM, RMSE = 1.16 BPM, MAPE = 0.76%), demonstrating that under well-controlled illumination and motion-stable conditions, video-based HR estimation can closely reproduce reference values across a wide dynamic range. In Fig. 7b, a narrower HR range still yields high concordance (PCC = 92.05%, CCC = 91.97%, MAE = 0.54 BPM, RMSE = 1.02 BPM, MAPE = 0.92%), confirming that the approach remains robust even when HR variability is minimal. These examples together highlight that high correlation does not always coincide with minimal error, and conversely, low absolute errors can persist even when correlation is reduced by constrained HR variability. Hence, both association- and error-based metrics are indispensable for a complete interpretation.

Beyond numerical agreement, the figure underscores how acquisition and preprocessing decisions fundamentally shape rPPG fidelity. Region-of-interest (ROI) selection—whether focusing on the forehead, cheeks, or fused multi-ROI signals—along with illumination stability, compression, frame rate, and motion artifacts, all influence the extracted waveform prior to HR computation. A rigorous pipeline should therefore combine photometric normalization, motion compensation, and multi-ROI fusion with outlier rejection, followed by signal-quality gating using spectral or harmonic consistency criteria. Explicit documentation of every stage—color-space transform (e.g., CHROM, POS), temporal filtering, detrending, synchronization with reference signals, and sign-alignment policy—remains critical to ensure reproducibility and comparability across studies. Specifically, regarding the phase delay between the measured and reference signals, we recommend a two-fold protocol: (i) for dataset developers, providing

synchronized timestamps for each video frame and reference signal sample to prevent hardware-level time offsets; and (ii) for algorithm developers, explicitly reporting the alignment strategies or compensation methods for the pulse transit time (PTT) shift (e.g., cross-correlation based alignment¹⁹⁵ or spectrum distribution similarity supervision mechanisms^{163,164}) to ensure fair comparisons across different studies. For actionable reproducibility, Supplementary Information (Table S3) provides a reproducibility checklist covering acquisition details, preprocessing steps, synchronization and alignment with reference signals, validation design, and code/model availability, together with minimum and recommended hardware conditions for deployment.

Overall, Fig. 7 serves as a template for standardized reporting of PPG–rPPG agreement in HR estimation. Beyond presenting summary statistics, a comprehensive evaluation should include: (i) waveform overlays to visualize temporal alignment, (ii) synchronous scatter plots to assess monotonicity and polarity, (iii) Bland–Altman analysis for bias and limits of agreement, and (iv) a balanced metric set spanning correlation, concordance, and absolute errors. Reporting both percentage-based (e.g., MAPE) and unit-based errors (e.g., MAE, RMSE) alongside Bland–Altman results enables a more transparent and physiologically interpretable assessment. Such multi-metric frameworks are essential for bridging technical validation and clinical translation of rPPG-based monitoring systems.

Non-contact rPPG can improve comfort and compliance—particularly for paediatric and elderly patients and for prolonged observation—by reducing adhesive sensors and cabling. While the present illustration is limited to two subjects and HR trends, it motivates prospective studies comparing rPPG-derived HR trajectories against reference monitors across diverse skin tones, lighting, and motion, and testing whether rPPG trends can flag decompensation earlier or with fewer artifacts in real-world workflows. These considerations are especially salient in neonatal and paediatric care, where minimizing sensor burden is a priority. In addition to adult monitoring, rPPG has shown promising potential in neonatal and paediatric environments. Hatib et al.¹⁹⁸ conducted a two-phased clinical study evaluating rPPG deployment in paediatric wards, demonstrating its ability to reduce sensor burden in neonates. Similarly, Wegerif et al.¹⁹⁹ compared rPPG-based pulse rate estimation methods in infants and confirmed its feasibility in clinical paediatric settings, including intensive care units. These findings suggest that rPPG can complement existing practices in neonatal ICUs by enabling continuous, non-contact monitoring without disturbing fragile patients.

Furthermore, with the increasing trend towards telemedicine, rPPG offers a viable solution for remote HR monitoring of patients, enabling timely interventions and reducing the need for frequent hospital visits^{205,206}. The insights provided by our clinical partners were instrumental in highlighting these applications and ensuring that our roadmap addresses the practical needs of healthcare providers. A comprehensive review²⁰⁷ highlighted the challenges and prospects of applying computer vision and rPPG-based monitoring in clinical environments, while a recent clinical study²⁰⁸ demonstrated the feasibility of such techniques for diagnosing respiratory infections in paediatric emergency room settings. Together, these works reinforce the applicability of rPPG in acute care scenarios where conventional monitoring is often impractical or resource-limited.

While rPPG holds promise for integration into telemedicine and remote monitoring workflows, translation into regulated clinical use requires adherence to established medical device pathways. Depending on intended use, rPPG systems may fall under wellness monitoring or regulated medical device categories, each associated with distinct evidence requirements. Beyond algorithmic validation on public datasets, clinical translation necessitates staged evaluation, including (i) technical verification under controlled conditions, (ii) pilot feasibility studies in representative clinical environments, and (iii) large-scale prospective clinical trials demonstrating safety, accuracy, and clinical benefit relative to reference standards. Engagement with regulatory frameworks (e.g., FDA, CE marking) and alignment with data governance, privacy, and interoperability standards are essential prerequisites for approval and adoption in clinical practice.

Beyond physical health, rPPG has been increasingly applied in affective computing and mental health assessments. Several studies have demonstrated that rPPG-derived metrics, such as heart rate variability and signal amplitude patterns, can help detect emotional states, including stress, anxiety, and depression^{209,210}. These applications open the door for rPPG to support mental health monitoring in home-based and outpatient settings. One of the key factors driving the growth of rPPG applications in public health is its integration into smartphones, security cameras, and other visual inspection devices²¹¹. Several studies have suggested that such noninvasive remote monitoring significantly reduces hospitalizations in chronic heart failure patients and may benefit other chronic disease populations^{212,213}. A recent field validation concluded that wireless health monitoring devices are convenient, user friendly, and effective for mass screening in community settings²¹⁴. To maximize the global impact of remote monitoring and mobile health technologies, it is vital to address remaining technical challenges, foster standards for data interoperability, and promote equitable access in low resource environments^{215,216}.

Although the technical and clinical potential of rPPG is increasingly demonstrated in research, its adoption in regulated clinical practice remains limited. For rPPG-based systems to be approved for medical use, they must undergo rigorous validation and meet regulatory requirements. In the United States, such systems would typically require FDA clearance as Class II medical devices, demonstrating substantial equivalence to predicate technologies through clinical trials and safety evaluations. Similarly, Conformité Européenne (CE) marking under the European Medical Device Regulation requires conformity with safety and performance standards. Clinical validation studies comparing rPPG output with gold-standard ECG or cPPG readings across diverse populations and use conditions are essential.

Moreover, adherence to data interoperability frameworks and privacy regulations (e.g., HIPAA, GDPR) is critical for deployment in clinical and telehealth environments. Beyond regulatory compliance, large-scale deployment of facial video-based physiological monitoring requires explicit governance and risk-mitigation strategies. Privacy-preserving design principles—including data minimization, purpose limitation, and preferential use of on-device processing—should be adopted to reduce reliance on centralized storage of identifiable facial video. Where feasible, raw video frames should be treated as transient input, with only de-identified physiological signals or derived features retained. Clear consent frameworks, role-based access control, encryption, retention limits, and audit mechanisms are essential to mitigate misuse and support trustworthy deployment. Importantly, governance strategies should remain context-dependent, reflecting whether an rPPG system is intended for wellness monitoring, research, or regulated clinical decision support.

In the feasibility example presented in this work, facial video data were collected under IRB approval with explicit participant consent, stored on restricted local systems, and processed separately from raw video storage, illustrating practical safeguards for handling identifiable visual data in controlled research settings.

Emphasizing clinical safety requires not only minimizing false alarms but also implementing robust “fail-safe” mechanisms that trigger contact-based verification when rPPG signals are deemed unreliable for clinical decision-making. Beyond improving signal quality, future work must prioritize the development of standardized validation frameworks that enforce rigorous evaluation strategies (e.g., external validation or subject-wise partitioning) to mitigate data leakage and prevent overoptimistic performance reporting. Such rigor is essential to support regulatory approval and seamless integration into clinical and electronic health systems. Importantly, given the heterogeneity in datasets, acquisition protocols, and validation strategies, cross-study MAE comparisons should be interpreted as descriptive summaries rather than directly comparable indicators of algorithm performance. HR measurement devices based on cPPG algorithms have been widely adopted in consumer electronics, including smart wristbands and portable pulse oximeters. In contrast, techniques based on rPPG remain largely confined to laboratory environments, and their

accuracy often lags behind cPPG systems, particularly under complex real-world conditions involving motion, illumination variability, and heterogeneous user populations. Future research should therefore prioritize bridging the performance gap between rPPG and established contact-based technologies while leveraging the unique advantages of rPPG, including its non-contact nature, low cost, and potential for passive monitoring using widely available imaging devices.

Drawing on the insights from this roadmap, including the literature review, methodological analyses, and emerging clinical demonstrations, we outline a prioritized translational pathway for advancing rPPG-based heart rate (HR) monitoring toward clinical and regulatory validation. In the near term, the most critical priority is the construction of large, standardized public datasets with synchronized reference electrocardiogram (ECG) signals, demographic diversity, and controlled acquisition conditions. Such datasets are essential to enable robust subject-wise validation, reduce bias associated with skin tone, age, and environmental variability, and facilitate fair benchmarking across algorithms. In particular, datasets capturing dynamic physiological conditions, such as rapid HR fluctuations during exercise, will be necessary to support emerging real-world applications²¹⁷. In high-impact domains, including telemedicine, in-vehicle monitoring, and virtual reality, publicly available datasets will also accelerate the transition from laboratory research to scalable deployment by enabling consistent evaluation protocols and reproducibility. Furthermore, large-scale datasets are indispensable for training data-intensive deep learning models, for which existing validation frameworks remain incomplete and require further development²¹⁸.

In the intermediate stage, improving algorithm robustness and generalizability remains a central challenge. This includes the development of standardized preprocessing pipelines, motion-resilient signal extraction methods, and rigorous cross-dataset validation strategies. In addition to real data collection, systematic data augmentation frameworks will play a critical role in enhancing model performance. Future work should move beyond simple interpolation-based approaches and instead incorporate standardized validation criteria that ensure both physiological fidelity and visual realism. For example, synthetic waveform generation should preserve pulse morphology and temporal dynamics, while image-based augmentation should maintain realistic skin texture characteristics. A dual-validation strategy combining physiological consistency metrics and visual fidelity measures may help bridge the synthetic-to-real gap and ensure that augmented datasets improve generalization without introducing artifacts^{78,219,220}.

Further improvements are also required in the stages preceding rPPG signal estimation and following initial HR extraction. In particular, the selection and tracking of regions of interest (ROIs) remain critical determinants of measurement accuracy, as performance varies across facial regions under different illumination conditions and motion levels^{221,222}. While prior work has introduced signal quality index (SQI) benchmarks²⁴, additional validation under diverse real-world scenarios is necessary to support adaptive and real-time ROI selection. Emerging vision foundation models, such as Sapiens²²³, offer promising opportunities to enhance robustness through improved face detection, keypoint localization, and surface modeling. In parallel, incorporating uncertainty quantification and structured failure-mode analysis may enable systems to identify unreliable measurements and provide low-confidence warnings under adverse conditions such as excessive motion or poor illumination¹⁹⁶. The integration of auxiliary signals, such as inertial measurements, may further support motion artifact compensation, which remains underexplored in current rPPG pipelines^{224–226}.

Clarifying the role of rPPG within the broader ecosystem of cardiac monitoring technologies is another essential step toward clinical translation. While 12-lead ECG remains the gold standard for diagnosis, and wearable contact PPG enables continuous monitoring, rPPG offers a distinct advantage as a non-contact modality. Its potential lies in large-scale screening, telehealth integration, and passive monitoring in everyday environments. Identifying the appropriate “ecological niche” for rPPG,

where its unique advantages outweigh its current limitations, will be critical for guiding both research priorities and clinical adoption.

In the longer term, the deployment of rPPG algorithms on mobile and edge platforms represents a key opportunity for real-world impact. Achieving this will require lightweight models, efficient inference, and robustness to heterogeneous user conditions. Model compression and pruning techniques may help reduce computational complexity while maintaining performance^{227–229}. Integration with multimodal sensing frameworks, such as combined RGB and mmWave systems, may further improve reliability in challenging environments and address limitations related to motion, privacy, and signal variability²³⁰. Importantly, prospective real-world studies will be required to evaluate system performance in practical deployment scenarios, including telemedicine platforms, mobile devices, and ambient monitoring systems, while reporting clinically meaningful metrics aligned with emerging regulatory standards for AI-enabled medical technologies.

Beyond healthcare settings, incorporating rPPG into real-world application platforms may unlock additional use cases. For example, integration into virtual reality environments can enable adaptive biofeedback systems that respond dynamically to users' physiological states, enhancing immersion and user experience. Similarly, multi-task learning frameworks that jointly perform identity recognition and physiological estimation have demonstrated potential for improving signal robustness and reliability²³¹.

Future research should explore unsupervised and self-supervised learning paradigms to reduce reliance on labeled datasets, which remain scarce and difficult to acquire. Recent studies have demonstrated that deep neural networks can estimate HR with minimal or no annotations, suggesting a promising direction for improving generalization across diverse environments and populations^{195,232,233}. Leveraging large-scale unlabeled video data may therefore play a key role in advancing rPPG toward robust, scalable, and clinically applicable solutions.

There is no doubt that we are entering a new era characterized by more advanced information technologies. Driven by the synchronous progress in both software and hardware, the real-world application of more flexible in-situ health monitoring has become increasingly feasible. With the support of rPPG techniques, extracting cardiac rhythm from a person only requires using a standard RGB camera, which is widely available today. The scientific principles behind the measurement process have exhibited significant research value. While several trial-stage platforms have emerged, the rPPG technology remains confined to laboratory settings, with a considerable gap to large-scale productization. Within the remote HR estimation pipeline, each substep can exert substantial influence on the final measurement accuracy. It is therefore necessary to systematically review and summarize different components within the rPPG pipeline. Except for algorithmic improvements, deeper considerations are required regarding noise mitigation from the signal source, efficient utilization of available datasets, and deployment at the edge-computing side. This article has methodological limitations, including database and language restrictions and manual inclusion decisions that may introduce selection bias. Additionally, further exploration of clinical applications requires closer cooperation among industry, academia, and clinical practitioners. Any aspect that can achieve a greater breakthrough would accelerate the adoption of more cost-effective HR measurement tools, eliminating the need for purchasing dedicated devices.

Methods

Eligibility criteria

In this study, we summarize the key challenges encountered during the development of rPPG technology for HR measurement into four aspects. These challenges served as the main references for including and excluding literature search results based on their research focus. The descriptions are shown as follows:

- *Algorithm effectiveness.* The effectiveness of an rPPG algorithm depends on its ability to accurately extract subtle variations in skin

color associated with the cardiac cycle. However, many algorithms rely on strong assumptions that might not be suitable in some scenarios and thus lead to poor adaptability. More advanced algorithms demonstrate better resilience to motion artifacts and noise, resulting in improved measurement accuracy.

- *Application environment.* While the ideal goal is to apply rPPG algorithms under dynamic and unconstrained conditions, in practice, factors such as uneven illumination, insufficient lighting, and subject movement often result in degraded performance²³⁴. These challenges make it difficult for rPPG algorithms to maintain high accuracy across application scenarios.
- *Hardware restriction.* During the deployment phase of rPPG algorithms, hardware limitations can significantly impact algorithm performance. Camera capability, computation speed, and memory capacity are essential factors that need to be considered.
- *Insufficient data.* The development and evaluation of rPPG algorithms depend heavily on high-quality datasets. In particular, data scarcity and imbalance are prevalent and noteworthy concerns. Demographic bias related to age, gender, skin tone, and health status can substantially affect algorithm validation^{235,236}. This problem is specifically severe for data-driven algorithms, where the data distribution of the training set directly affects the algorithm's generalization and performance.

In addition to these requirements related to research emphasis, included studies were restricted to experimental and observational articles written in English and published between 2011 and 2025. Studies were excluded if they focused on the non-visible spectrum, used signal types other than PPG, or did not use facial videos as the input. Publications that were inaccessible, preprints, reviews, or theses not published in peer-reviewed academic journals were also excluded from this work.

Information sources and search strategy

Following the PRISMA guideline^{237,238}, we conducted a comprehensive literature search across four major academic databases—Embase, Google Scholar, IEEE Xplore, and PubMed—restricted to studies published in English between 2011 and 2025. The search was completed on May 22, 2025. The complete search strings are provided below:

- *Embase:* (“rPPG”:ti OR “remote PPG”:ti OR “remote photoplethysmography”:ti) AND (“heart” OR “pulse” OR “heartbeat” OR “cardiac” OR “HRV”)
- *Google Scholar:* allintitle:(rPPG OR “remote PPG” OR “remote photoplethysmography”) + (heart OR pulse OR heartbeat OR HRV OR cardiac)
- *IEEE Xplore:* (“Document Title”: rPPG OR “Document Title”: remote PPG OR “Document Title”: remote photoplethysmography) AND (“Full Text Only”: heart OR pulse OR heartbeat OR cardiac OR HRV)
- *PubMed:* (rPPG[Title] OR remote PPG[Title] OR remote photoplethysmography[Title]) AND (heart OR pulse OR heartbeat OR cardiac OR HRV)

Synthetic analysis

In this article, we conducted synthetic analyses based on the studies in the field of rPPG-based HR measurement, which can be summarized as below:

- Statistics of usage frequency of different model-driven rPPG algorithms, data-driven rPPG algorithms, datasets, and performance evaluation metrics.
- Performance comparison for model-driven and data-driven algorithms (further divided into end-to-end and hybrid structures) across different datasets.

The synthetic analyses were based on the set of studies identified through our literature search. For a specific analysis, articles that lacked relevant characteristics were excluded. When reporting rPPG algorithm

performance for HR measurement, we used MAE as the primary evaluation metric, which is the most frequently reported across studies. No conversions from alternative evaluation metrics were performed.

Clinical demonstration in simulated hypovolemia: data and processing

We analyzed video and reference signals from the Yale lower-body negative pressure dataset (IRB No. 2000020057; 24 healthy adults). The reference for rPPG evaluation was the cPPG signal, from which ground-truth HR was obtained using the physiological monitor. rPPG signals were extracted from the forehead regions of interest using facial landmark-based tracking. The signals were detrended and processed with a wavelet-based method to derive HR. In the Results, we present data from two representative subjects to illustrate the clinical application of rPPG, alongside the simultaneously recorded ground-truth HR for comparison. Camera specifications and signal-processing parameters are detailed in Supplementary Information (Table S4).

Ethics approval and informed consent

This analysis was approved by the Yale University Institutional Review Board (IRB No. 2000020057) and conducted in accordance with the Declaration of Helsinki. Written informed consent to participate was obtained from all participants. Written consent for publication of identifiable facial images was obtained from the two participants depicted in Fig. 7.

Data availability

No new datasets were generated in this study. The recordings analyzed were obtained from a previously approved protocol incorporating rPPG measurements during lower body negative pressure (LBNP)-induced hypovolemia (IRB No. 2000020057, Yale University)²³⁹. De-identified derived data underlying Fig. 7, including rPPG-derived heart rate time series and associated analysis outputs, are available from the corresponding authors upon reasonable request, subject to Yale University Institutional Review Board (IRB) approval and a signed data use agreement. All other data supporting the findings of this study are included within the article.

Code availability

Custom analysis scripts were developed to extract facial regions of interest, compute rPPG signals using the chrominance-based (CHROM) method¹⁴⁴, and estimate heart rate from video recordings. The implementation was performed in Python (version 3.10) using standard scientific libraries including NumPy (v1.24), SciPy (v1.10), and OpenCV (v4.8). Signal processing steps included temporal detrending, band-pass filtering (0.7–4 Hz), and frequency- and peak-based heart rate estimation. The code used to generate the results in Fig. 7, together with minimal example inputs required for reproducibility, is available from A.A. and S.B. upon reasonable request. Access may be subject to institutional data-sharing agreements due to the use of IRB-governed data.

Received: 7 March 2025; Accepted: 26 April 2026;

Published online: 09 May 2026

References

1. Segerstrom, S. C. & Nes, L. S. Heart rate variability reflects self-regulatory strength, effort, and fatigue. *Psychol. Sci.* **18**, 275–281 (2007).
2. Appelhans, B. M. & Luecken, L. J. Heart rate variability as an index of regulated emotional responding. *Rev. Gen. Psychol.* **10**, 229–240 (2006).
3. Antelmi, I. et al. Influence of age, gender, body mass index, and functional capacity on heart rate variability in a cohort of subjects without heart disease. *Am. J. Cardiol.* **93**, 381–385 (2004).
4. Stracina, T., Ronzhina, M., Redina, R. & Novakova, M. Golden standard or obsolete method? review of ECG applications in clinical and experimental context. *Front. Physiol.* **13**, 867033 (2022).
5. Garvey, J. L. ECG techniques and technologies. *Emerg. Med. Clin.* **24**, 209–225 (2006).
6. Ramasamy, S. & Balan, A. Wearable sensors for ECG measurement: a review. *Sens. Rev.* **38**, 412–419 (2018).
7. Yanowitz, F. G. Introduction to ECG interpretation. in *LDS Hospital and Intermountain Medical Center* (Intermountain Healthcare, 2012).
8. Wasilewski, J. & Poloński, L. An introduction to ECG interpretation. in *ECG Signal Processing, Classification and Interpretation: A Comprehensive Framework of Computational Intelligence*, 1–20 (Springer, 2011).
9. Auer, R. et al. Association of major and minor ECG abnormalities with coronary heart disease events. *Jama* **307**, 1497–1505 (2012).
10. Babusiak, B. & Gala, M. Detection of abnormalities in ECG. In *Proc. Information Technologies in Biomedicine: Third International Conference, ITIB 2012*, 161–171 (Springer, 2012).
11. De Luna, A. B. *Basic Electrocardiography: Normal and Abnormal ECG patterns* (John Wiley & Sons, 2008).
12. Přibíl, J., Přibílová, A. & Frollo, I. Comparative measurement of the PPG signal on different human body positions by sensors working in reflection and transmission modes. *Eng. Proc.* **2**, 69 (2020).
13. Lopez, S. Pulse Oximeter Fundamentals and Design. Freescale Semiconductor Application Note AN4327 Rev. 2 (Freescale Semiconductor, 2012).
14. Constant, N., Wang, T. & Mankodiya, K. Pulseband: a hands-on tutorial on how to design a smart wristband to monitor heart-rate. In *Proc. IEEE Virtual Conference on Applications of Commercial Sensors (VCACS)*, 1–3 (IEEE, 2015).
15. Wu, T., Blažek, V. & Schmitt, H. J. Photoplethysmography imaging: a new noninvasive and noncontact method for mapping of the dermal perfusion changes. In *Optical Techniques and Instrumentation for the Measurement of Blood Composition, Structure, and Dynamics*, Vol 4163, 62–70 (SPIE, 2000).
16. Verkruysse, W., Svaasand, L. O. & Nelson, J. S. Remote plethysmographic imaging using ambient light. *Opt. Express* **16**, 21434–21445 (2008).
17. Castellano Ontiveros, R., Elgendi, M. & Menon, C. A machine learning-based approach for constructing remote photoplethysmogram signals from video cameras. *Commun. Med.* **4**, 109 (2024).
18. Frey, L., Menon, C. & Elgendi, M. Blood pressure measurement using only a smartphone. *npj Digit. Med.* **5**, 1–14 (2024).
19. Bondarenko, M., Menon, C. & Elgendi, M. Demographic bias in public remote photoplethysmography datasets. *npj Digit. Med.* **8**, 593 (2024).
20. Bondarenko, M., Menon, C. & Elgendi, M. The role of face regions in remote photoplethysmography for contactless heart rate monitoring. *npj Digit. Med.* **8**, 1–12 (2024).
21. Li, S., Elgendi, M. & Menon, C. Optimal facial regions for remote heart rate measurement during physical and cognitive activities. *npj Cardiovasc. Health* **1**, 1–12 (2025).
22. Bhutani, S., Elgendi, M. & Menon, C. Preserving privacy and video quality through remote physiological signal removal. *Commun. Eng.* **4**, 66 (2024).
23. Bhutani, S. et al. Vital signs-based healthcare kiosks for screening chronic and infectious diseases: a systematic review. *Commun. Med.* **5**, 28 (2024).
24. Elgendi, M., Martinelli, I. & Menon, C. Optimal signal quality index for remote photoplethysmogram sensing. *npj Biosensing* **1**, 5 (2024).
25. Takeuchi, H., Ohsuga, M. & Kamakura, Y. A study on region of interest in remote PPG and an attempt to eliminate false positive results using svm classification. In *Proc. IEEE International Conference on Artificial Intelligence in Engineering and Technology (IICAIET)*, 1–5 (IEEE, 2021).
26. Haugg, F., Elgendi, M. & Menon, C. Effectiveness of remote PPG construction methods: a preliminary analysis. *Bioengineering* **9**, 485 (2022).

27. Tran, T.-T., Pham Van, T. & Phan Dai D. A remote PPG based approach for heart rate estimation. *J. Mil. Sci. Technol.* **79**, 31–40 (2022).
28. Argüello Prada, E. J. & Serna Maldonado, R. D. A novel and low-complexity peak detection algorithm for heart rate estimation from low-amplitude photoplethysmographic (PPG) signals. *J. Med. Eng. Technol.* **42**, 569–577 (2018).
29. Labunets, L. V. & Ryakhina, M. Y. Heart rate estimation based on remote photoplethysmography signal hilbert transform. In *Proc. 26th International Conference on Digital Signal Processing and its Applications (DSPA)*, 1–5 (IEEE, 2024).
30. Bashar, S. S., Miah, M. S., Karim, A. Z., Al Mahmud, M. A. & Hasan, Z. A machine learning approach for heart rate estimation from PPG signal using random forest regression algorithm. In *Proc. International Conference on Electrical, Computer and Communication Engineering (ECCE)*, 1–5 (IEEE, 2019).
31. Karim, A. Z., Miah, M. S., Jamal, G. A., Fahima, R. A. & Rahman, M. T. Application of chrominance based rPPG in estimation of heart rate from video signal. In *Proc. 24th International Conference on Computer and Information Technology (ICCIT)*, 1–6 (IEEE, 2021).
32. Lyu, F., Zhu, Y., Chen, S. & Wang, L. Video-based heart rate measurement via the anti-shake remote photoplethysmography (rPPG) pulse extraction. In *International Conference on Computer Application and Proc. Information Security (ICCAIS 2021)*, vol. 12260, 149–157 (SPIE, 2022).
33. Hosni, A. & Atef, M. Remote real-time heart rate monitoring with recursive motion artifact removal using PPG signals from a smartphone camera. *Multimed. Tools Appl.* **82**, 20571–20588 (2023).
34. Guler, S., Ozturk, O., Golparvar, A., Dogan, H. & Yapici, M. K. Effects of illuminance intensity on the green channel of remote photoplethysmography (rPPG) signals. *Phys. Eng. Sci. Med.* **45**, 1317–1323 (2022).
35. Rosa, R., Floriano, A., Côgo Lampiery, L. & Freire, T. Evaluation of body positions for heart rate measurement using remote photoplethysmography: towards a mobile system for vital signs detection in unconscious people. In *Anais do 14º Simpósio Brasileiro de Automação Inteligente* (Galoa, 2019).
36. Soleymani, M., Lichtenauer, J., Pun, T. & Pantic, M. A multimodal database for affect recognition and implicit tagging. *IEEE Trans. Affect. Comput.* **3**, 42–55 (2011).
37. Stricker, R., Müller, S. & Gross, H.-M. Non-contact video-based pulse rate measurement on a mobile service robot. In *Proc. 23rd IEEE International Symposium on Robot and Human Interactive Communication*, 1056–1062 (IEEE, 2014).
38. Tasli, H. E., Gudi, A. & Den Uyl, M. Remote PPG based vital sign measurement using adaptive facial regions. In *Proc. IEEE International Conference on Image Processing (ICIP)*, 1410–1414 (IEEE, 2014).
39. Zhang, Z. et al. Multimodal spontaneous emotion corpus for human behavior analysis. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 3438–3446 (IEEE, 2016).
40. Heusch, G., Anjos, A. & Marcel, S. A reproducible study on remote heart rate measurement. Preprint at *arXiv* <https://doi.org/10.48550/arXiv.1709.00962> (2017).
41. Bobbia, S., Macwan, R., Benezeth, Y., Mansouri, A. & Dubois, J. Unsupervised skin tissue segmentation for remote photoplethysmography. *Pattern Recognit. Lett.* **124**, 82–90 (2019).
42. Pilz, C. S., Zaunseder, S., Krajewski, J. & Blazek, V. Local group invariance for heart rate estimation from face videos in the wild. In *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 1335–13358 (IEEE, 2018).
43. Niu, X., Han, H., Shan, S. & Chen, X. VIPL-HR: a multi-modal database for pulse estimation from less-constrained face video. In *Proc. Computer Vision–ACCV 2018: 14th Asian Conference on Computer Vision*, 562–576 (Springer, 2019).
44. Špetlík, R., Franc, V. & Matas, J. Visual heart rate estimation with convolutional neural network. In *Proc. British Machine Vision Conference*, 3–6 (BMVA, 2018).
45. Gudi, A., Bittner, M. & van Gemert, J. Real-time webcam heart-rate and variability estimation with clean ground truth for evaluation. *Appl. Sci.* **10**, 8630 (2020).
46. Yang, Z., Wang, H. & Lu, F. Assessment of deep learning-based heart rate estimation using remote photoplethysmography under different illuminations. *IEEE Trans. Hum. Mach. Syst.* **52**, 1236–1246 (2022).
47. Xi, L., Chen, W., Zhao, C., Wu, X. & Wang, J. Image enhancement for remote photoplethysmography in a low-light environment. In *Proc. 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020)*, 1–7 (IEEE, 2020).
48. Tang, J. et al. MMPD: multi-domain mobile video physiology dataset. In *Proc. 45th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 1–5 (IEEE, 2023).
49. Joshi, J. & Cho, Y. IBVP dataset: RGB-thermal rPPG dataset with high resolution signal quality labels. *Electronics* **13**, 1334 (2024).
50. Bevilacqua, F., Engström, H. & Backlund, P. Accuracy evaluation of remote photoplethysmography estimations of heart rate in gaming sessions with natural behavior. In *Proc. International Conference on Advances in Computer Entertainment*, 508–530 (Springer, 2017).
51. Sabour, R. M., Benezeth, Y., De Oliveira, P., Chappé, J. & Yang, F. UBFC-Phys: a multimodal database for psychophysiological studies of social stress. *IEEE Trans. Affect. Comput.* **14**, 622–636 (2023).
52. Di Lernia, D., Finotti, G., Tsakiris, M., Riva, G. & Naber, M. Remote photoplethysmography (rPPG) in the wild: remote heart rate imaging via online webcams. *Behav. Res. Methods* **56**, 6904–6914 (2024).
53. Yue, Z. et al. Deep super-resolution network for rPPG information recovery and noncontact heart rate estimation. *IEEE Trans. Instrum. Meas.* **70**, 1–11 (2021).
54. Premkumar, K. S., Angelopoulou, A., Kapetanios, E., Chausaulet, T. & Hemanth, D. J. Super-resolution convolutional network for image quality enhancement in remote photoplethysmography based heart rate estimation. In *Proc. International Conference on Computational Science*, 157–163 (Springer, 2022).
55. Zhai, D., Chen, W., Li, Q., Yu, M. & Wu, H. Adversarial learning network for recovering rPPG signals from low-resolution images for remote heart rate measurement. *Biomed. Signal Process. Control* **102**, 107382 (2025).
56. Rouast, P. V., Adam, M., Dörner, V. & Lux, E. Remote photoplethysmography: evaluation of contactless heart rate measurement in an information systems setting. In *Proc. Applied Informatics and Technology Innovation Conference* (Springer, 2016).
57. Zou, B., Lv, D. & Lyu, X. Feature signal resampling for rPPG-based remote cardiac pulse measurement with streaming video. In *Proc. IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 3916–3918 (IEEE, 2022).
58. Chen, C.-C., Lin, S.-X. & Jeong, H. Low-complexity timing correction methods for heart rate estimation using remote photoplethysmography. *Sensors* **25**, 588 (2025).
59. Álvarez Casado, C., Nguyen, L., Silvén, O. & Bordallo López, M. Assessing the feasibility of remote photoplethysmography through videocalls: A study of network and computing constraints. In *Scandinavian Conference on Image Analysis*, 586–598 (Springer, 2023).
60. Blackford, E. B. & Estepp, J. R. Effects of frame rate and image resolution on pulse rate measured using multiple camera imaging photoplethysmography. In (eds Gimi, B. & Molthen, R. C.) *Medical*

- Imaging 2015: Biomedical Applications in Molecular, Structural, and Functional Imaging*, Vol. 9417, 94172D (SPIE, 2015).
61. McDuff, D. J., Blackford, E. B. & Estep, J. R. The impact of video compression on remote cardiac pulse measurement using imaging photoplethysmography. In *Proc. 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)*, 63–70 (IEEE, 2017).
62. Hyanda, M. H. et al. A comparative evaluation of video codecs for rPPG-based heart rate estimation. In *Proc. Asia Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, 243–247 (IEEE, 2023).
63. Tilbury, B., Arevalillo-Herráez, M. & Ramzan, N. Comparison of constant rate factor and constant bitrate mode encoding for rPPG detection. In *Proc. 15th International Conference on Quality of Multimedia Experience (QoMEX)*, 205–208 (IEEE, 2023).
64. Wang, J., Shan, C., Liu, Z., Zhou, S. & Shu, M. Physiological information preserving video compression for rPPG. *IEEE J. Biomed. Health Inform.* **29**, 3563–3575 (2025).
65. Nguyen, N., Nguyen, L., Li, H., López, M. B. & Casado, C. Á. Evaluation of video-based rPPG in challenging environments: Artifact mitigation and network resilience. *Comput. Biol. Med.* **179**, 108873 (2024).
66. Laurie, J., Higgins, N., Peynot, T. & Roberts, J. Dedicated exposure control for remote photoplethysmography. *IEEE Access* **8**, 116642–116652 (2020).
67. van Esch, R. et al. Remote PPG for heart rate monitoring: lighting conditions and camera shutter time. In *Proc. 43rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Virtual*, 1–5 (IEEE, 2021).
68. Xiao, C., Yi, S., Yu, D. & Wang, W. Dynamic exposure controlling and suppression for highly robust remote-PPG in fitness. In *Proc. 46th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 1–4 (IEEE, 2024).
69. Mironenko, Y., Kalinin, K., Kopelovich, M. & Petrushan, M. Remote photoplethysmography: rarely considered factors. In *Proc. IEEE/CVF conference on computer vision and pattern recognition workshops*, 296–297 (IEEE, 2020).
70. Dasari, A., Prakash, S. K. A., Jeni, L. A. & Tucker, C. S. Evaluation of biases in remote photoplethysmography methods. *npj Digit. Med.* **4**, 91 (2021).
71. Lampier, L. C. et al. A deep learning approach to estimate pulse rate by remote photoplethysmography. *Physiol. Meas.* **43**, 075012 (2022).
72. Li, B., Jiang, W., Peng, J. & Li, X. Deep learning-based remote-photoplethysmography measurement from short-time facial video. *Physiol. Meas.* **43**, 115003 (2022).
73. Vance, N., Speth, J., Sporrer, B. & Flynn, P. Promoting generalization in cross-dataset remote photoplethysmography. In *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 5984–5992 (IEEE, 2023).
74. Perche, S. et al. Data-augmentation for deep learning based remote photoplethysmography methods. In *Proc. International Conference on e-Health and Bioengineering (EHB)*, 1–4 (IEEE, 2021).
75. Jayasinghe, J. & Malasinghe, L. ML and DL model comparison with skin tone and dynamic range analysis for rPPG. In *2024 6th International Conference on Advancements in Computing (ICAC)*, 510–515 (IEEE, 2024).
76. Savic, M. & Zhao, G. RS+rPPG: robust strongly self-supervised learning for rPPG. In *Proc. IEEE Transactions on Circuits and Systems for Video Technology* (IEEE, 2025).
77. Shirke, A. S., Arunachala, C., Tiwari, N. & Nataraj, K. Eulerian magnification and transformers for enhanced rPPG-based pulse rate estimation in Indian populations. In *Proc. IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI)*, Vol 3, 1–5 (IEEE, 2025).
78. Ba, Y., Wang, Z., Karınca, K. D., Bozkurt, O. D. & Kadambi, A. Style transfer with bio-realistic appearance manipulation for skin-tone inclusive rPPG. In *Proc. IEEE International Conference on Computational Photography (ICCP)*, 1–12 (IEEE, 2022).
79. Gupta, A. K., Kumar, R., Birla, L. & Gupta, P. RADIANT: better rPPG estimation using signal embeddings and transformer. In *Proc. IEEE/CVF Winter Conference on Applications of Computer Vision*, 4976–4986 (IEEE, 2023).
80. Russakovsky, O. et al. Imagenet large scale visual recognition challenge. *Int. J. Comput. Vis.* **115**, 211–252 (2015).
81. Gao, H., Zhang, C., Pei, S. & Wu, X. IDTL-rPPG: remote heart rate estimation using instance-based deep transfer learning. *Biomed. Signal Process. Control* **95**, 106416 (2024).
82. Du, J., Liu, S.-Q., Zhang, B. & Yuen, P. C. Dual-bridging with adversarial noise generation for domain adaptive rPPG estimation. In *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 10355–10364 (2023).
83. Lee, H., Lee, H., Lee, B. H. & Chun, S. Y. Continual test-time adaptation for robust remote photoplethysmography estimation. *IEEE Access* **13**, 55049–55060 (2025).
84. Feng, L., Po, L.-M., Xu, X. & Li, Y. Motion artifacts suppression for remote imaging photoplethysmography. In *Proc. 19th International Conference on Digital Signal Processing*, 18–23 (IEEE, 2014).
85. Fouad, R., Omer, O. A. & Aly, M. H. Optimizing remote photoplethysmography using adaptive skin segmentation for real-time heart rate monitoring. *IEEE Access* **7**, 76513–76528 (2019).
86. Viola, P. & Jones, M. Rapid object detection using a boosted cascade of simple features. In *Proc. 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, Vol 1, 1–I (IEEE, 2001).
87. Viola, P., Jones, M. et al. Robust real-time object detection. *Int. J. Comput. Vis.* **4**, 4 (2001).
88. Lucas, B. D. & Kanade, T. An iterative image registration technique with an application to stereo vision. In *IJCAI'81: 7th International Joint Conference on Artificial Intelligence*, Vol 2, 674–679 (Morgan Kaufmann Publishers Inc., 1981).
89. Kalman, R. E. A new approach to linear filtering and prediction problems. *J. Basic Eng.* **82**, 35–45 (1960).
90. Kalman, R. E. & Bucy, R. S. New results in linear filtering and prediction theory. *J. Basic Eng.* **83**, 95–108 (1961).
91. Bolme, D. S., Beveridge, J. R., Draper, B. A. & Lui, Y. M. Visual object tracking using adaptive correlation filters. In *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2544–2550 (IEEE, 2010).
92. Zhao, C., Mei, P., Xu, S., Li, Y. & Feng, Y. Performance evaluation of visual object detection and tracking algorithms used in remote photoplethysmography. In *2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)*, 1646–1655 (IEEE, 2019).
93. Shi, J. et al. Good features to track. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 593–600 (IEEE, 1994).
94. Bay, H., Ess, A., Tuytelaars, T. & Van Gool, L. Speeded-up robust features (SURF). *Comput. Vis. Image Underst.* **110**, 346–359 (2008).
95. Zhang, K., Zhang, Z., Li, Z. & Qiao, Y. Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE Signal. Process. Lett.* **23**, 1499–1503 (2016).
96. Wu, B.-F., Wu, Y.-C. & Chou, Y.-W. A compensation network with error mapping for robust remote photoplethysmography in noise-heavy conditions. *IEEE Trans. Instrum. Meas.* **71**, 1–11 (2022).
97. Huang, P.-W., Wu, B.-J. & Wu, B.-F. A heart rate monitoring framework for real-world drivers using remote photoplethysmography. *IEEE J. Biomed. Health Inform.* **25**, 1397–1408 (2021).
98. Hu, M., Guo, D., Wang, X., Ge, P. & Chu, Q. A novel spatial-temporal convolutional neural network for remote photoplethysmography. In *Proc. 12th International Congress on Image and Signal Processing*,

- BioMedical Engineering and Informatics (CISP-BMEI)*, 1–6 (IEEE, 2019).
99. Comas, J., Ruiz, A. & Sukno, F. Efficient remote photoplethysmography with temporal derivative modules and time-shift invariant loss. In *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2182–2191 (IEEE, 2022).
 100. King, D. E. Dlib-ml: a machine learning toolkit. *J. Mach. Learn. Res.* **10**, 1755–1758 (2009).
 101. Den Uyl, M. & Van Kuilenburg, H. The FaceReader: online facial expression recognition. In *Proc. Measuring Behavior*, Vol 30, 589–590 (Citeseer, 2005).
 102. Wu, S., Kan, M., He, Z., Shan, S. & Chen, X. Funnel-structured cascade for multi-view face detection with alignment-awareness. *Neurocomputing* **221**, 138–145 (2017).
 103. OMRON OKAO Vision software. http://www.omron.com/r_d/coretech/vision/okao.html.
 104. Zadeh, A., Chong Lim, Y., Baltusaitis, T. & Morency, L.-P. Convolutional experts constrained local model for 3D facial landmark detection. In *Proc. IEEE International Conference on Computer Vision Workshops*, 2519–2528 (IEEE, 2017).
 105. Kartynnik, Y., Ablavatski, A., Grishchenko, I. & Grundmann, M. Real-time facial surface geometry from monocular video on mobile GPUs. In *CVPR Workshop on Computer Vision for Augmented and Virtual Reality 2019* (Long Beach, CA, 2019).
 106. Jayasinghe, J., Katsigiannis, S. & Malasinghe, L. Comparative study of face tracking algorithms for remote photoplethysmography. In *Proc. International Conference on Electrical, Computer and Energy Technologies (ICECET)*, 1–7 (IEEE, 2023).
 107. Kossack, B., Wisotzky, E., Hilsmann, A. & Eisert, P. Automatic region-based heart rate measurement using remote photoplethysmography. In *Proc. IEEE/CVF International Conference on Computer Vision*, 2755–2759 (IEEE, 2021).
 108. Wang, G. Influence of ROI selection for remote photoplethysmography with singular spectrum analysis. In *Proc. IEEE International Conference on Artificial Intelligence and Industrial Design (AIID)*, 416–420 (IEEE, 2021).
 109. Bobbia, S., Benezeth, Y. & Dubois, J. Remote photoplethysmography based on implicit living skin tissue segmentation. In *Proc. 23rd International Conference on Pattern Recognition (ICPR)*, 361–365 (IEEE, 2016).
 110. Wei, W., Vatanparvar, K., Zhu, L., Kuang, J. & Gao, A. Remote photoplethysmography and heart rate estimation by dynamic region of interest tracking. In *Proc. 44th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 3243–3248 (IEEE, 2022).
 111. Yang, Y. et al. Motion robust remote photoplethysmography in CIElab color space. *J. Biomed. Opt.* **21**, 117001 (2016).
 112. Feng, L. et al. Dynamic ROI based on K-means for remote photoplethysmography. In *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 1310–1314 (IEEE, 2015).
 113. Benezeth, Y., Bobbia, S., Nakamura, K., Gomez, R. & Dubois, J. Probabilistic signal quality metric for reduced complexity unsupervised remote photoplethysmography. In *Proc. 13th International Symposium on Medical Information and Communication Technology (ISMICT)*, 1–5 (IEEE, 2019).
 114. Boccignone, G. et al. Enhancing rPPG pulse-signal recovery by facial sampling and PSD clustering. *Biomed. Signal Process. Control* **101**, 107158 (2025).
 115. Zhang, Y. et al. MaskFusionNet: a dual-stream fusion model with masked pre-training mechanism for rPPG measurement. In *IEEE Trans. Circuits Syst. Video Technol.* **34**, 11521–11534 (2024).
 116. Zhao, P. et al. Toward motion robustness: a masked attention regularization framework in remote photoplethysmography. In *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 7829–7838 (IEEE, 2024).
 117. Wong, K. L. et al. Optimising rPPG signal extraction by exploiting facial surface orientation. In *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2165–2171 (IEEE, 2022).
 118. Cantrill, S., Ahmedt-Aristizabal, D., Petersson, L., Suominen, H. & Armin, M. A. Orientation-conditioned facial texture mapping for video-based facial remote photoplethysmography estimation. In *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 354–363 (IEEE, 2024).
 119. Fan, X. et al. Remote photoplethysmography based on reflected light angle estimation. *Physiol. Meas.* **45**, 035005 (2024).
 120. Unakafov, A. M. Pulse rate estimation using imaging photoplethysmography: generic framework and comparison of methods on a publicly available dataset. *Biomed. Phys. Eng. Express* **4**, 045001 (2018).
 121. Wang, W., den Brinker, A. C., Stuijk, S. & de Haan, G. Color-distortion filtering for remote photoplethysmography. In *Proc. 12th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2017)*, 71–78 (IEEE, 2017).
 122. Jaiswal, K. B. & Meenpal, T. Continuous pulse rate monitoring from facial video using rPPG. In *Proc. 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, 1–5 (IEEE, 2020).
 123. Li, J., Vatanparvar, K., Zhu, L., Kuang, J. & Gao, A. Enhancement of remote PPG and heart rate estimation with optimal signal quality index. In *Proc. IEEE-EMBS International Conference on Wearable and Implantable Body Sensor Networks (BSN)*, 1–4 (IEEE, 2022).
 124. Bobbia, S. et al. Real-time temporal superpixels for unsupervised remote photoplethysmography. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 1341–1348 (IEEE, 2018).
 125. Macwan, R., Benezeth, Y. & Mansouri, A. Heart rate estimation using remote photoplethysmography with multi-objective optimization. *Biomed. Signal Process. Control* **49**, 24–33 (2019).
 126. Botina-Monsalve, D. et al. Long short-term memory deep-filter in remote photoplethysmography. In *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 1242–1249 (IEEE, 2020).
 127. Botina-Monsalve, D., Benezeth, Y. & Miteran, J. Performance analysis of remote photoplethysmography deep filtering using long short-term memory neural network. *Biomed. Eng. OnLine* **21**, 69 (2022).
 128. Conaire, C. O., O'Connor, N. E. & Smeaton, A. F. Detector adaptation by maximising agreement between independent data sources. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 1–6 (IEEE, 2007).
 129. Tang, C., Lu, J. & Liu, J. Non-contact heart rate monitoring by combining convolutional neural network skin detection and remote photoplethysmography via a low-cost camera. In *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 1390–13906 (IEEE, 2018).
 130. Kim, S.-E., Yu, S.-G., Kim, N. H., Suh, K. H. & Lee, E. C. Restoration of remote PPG signal through correspondence with contact sensor signal. *Sensors* **21**, 5910 (2021).
 131. Lin, Y.-C. & Lin, Y.-H. A study of color illumination effect on the SNR of rPPG signals. In *Proc. 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 4301–4304 (IEEE, 2017).
 132. Sinhal, R., Singh, K. R. & Gupta, K. Color intensity: a study of rPPG algorithm for heart rate estimation. In *Proc. International Conference on Computational Performance Evaluation (ComPE)*, 580–584 (IEEE, 2021).
 133. Xi, L., Wu, X., Chen, W., Wang, J. & Zhao, C. Weighted combination and singular spectrum analysis based remote

- photoplethysmography pulse extraction in low-light environments. *Med. Eng. Phys.* **105**, 103822 (2022).
134. Ontiveros, R. C., Elgendi, M., Missale, G. & Menon, C. Evaluating RGB channels in remote photoplethysmography: a comparative study with contact-based PPG. *Front. Physiol.* **14**, 1296277 (2023).
 135. Horecker, B. L. The absorption spectra of hemoglobin and its derivatives in the visible and near infra-red regions. *J. Biol. Chem.* **148**, 173–183 (1943).
 136. Hugg, F., Elgendi, M. & Menon, C. GRGB rPPG: an efficient low-complexity remote photoplethysmography-based algorithm for heart rate estimation. *Bioengineering* **10**, 243 (2023).
 137. Poh, M.-Z., McDuff, D. J. & Picard, R. W. Non-contact, automated cardiac pulse measurements using video imaging and blind source separation. *Opt. Express* **18**, 10762–10774 (2010).
 138. Macwan, R., Benezeth, Y., Mansouri, A., Nakamura, K. & Gomez, R. Remote photoplethysmography measurement using constrained ICA. In *Proc. E-Health and Bioengineering Conference (EHB)*, 430–433 (IEEE, 2017).
 139. Macwan, R., Benezeth, Y. & Mansouri, A. Remote photoplethysmography with constrained ICA using periodicity and chrominance constraints. *Biomed. Eng. OnLine* **17**, 22 (2018).
 140. Lewandowska, M., Rumiński, J., Kocejko, T. & Nowak, J. Measuring pulse rate with a webcam - a non-contact method for evaluating cardiac activity. In *Proc. Federated Conference on Computer Science and Information Systems (FedCSIS)* 405–410 (IEEE, 2011).
 141. Wang, W., Stuijk, S. & De Haan, G. A novel algorithm for remote photoplethysmography: spatial subspace rotation. *IEEE Trans. Biomed. Eng.* **63**, 1974–1984 (2016).
 142. Wang, W., Den Brinker, A. C., Stuijk, S. & De Haan, G. Algorithmic principles of remote PPG. *IEEE Trans. Biomed. Eng.* **64**, 1479–1491 (2017).
 143. Ernst, H., Scherpf, M., Malberg, H. & Schmidt, M. Optimal color channel combination across skin tones for remote heart rate measurement in camera-based photoplethysmography. *Biomed. Signal Process. Control* **68**, 102644 (2021).
 144. De Haan, G. & Jeanne, V. Robust pulse rate from chrominance-based rPPG. *IEEE Trans. Biomed. Eng.* **60**, 2878–2886 (2013).
 145. De Haan, G. & Van Leest, A. Improved motion robustness of remote-PPG by using the blood volume pulse signature. *Physiol. Meas.* **35**, 1913–1926 (2014).
 146. Zhu, Q. et al. A comparative study of principled rPPG-based pulse rate tracking algorithms for fitness activities. *IEEE Trans. Biomed. Eng.* **72**, 152–165 (2024).
 147. Hebbbar, S. & Sato, T. Motion robust remote photoplethysmography via frequency domain motion artifact reduction. In *Proc. IEEE Biomedical Circuits and Systems Conference (BioCAS)*, 1–4 (IEEE, 2021).
 148. Zhou, K., Krause, S., Blocher, T. & Stork, W. Enhancing remote-PPG pulse extraction in disturbance scenarios utilizing spectral characteristics. In *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 280–281 (IEEE, 2020).
 149. Demirezen, H. & Erdem, C. E. Remote photoplethysmography using nonlinear mode decomposition. In *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 1060–1064 (IEEE, 2018).
 150. Song, R. et al. Remote photoplethysmography with an EEMD-MCCA method robust against spatially uneven illuminations. *IEEE Sens. J.* **21**, 13484–13494 (2021).
 151. Zhao, C., Lin, C.-L., Chen, W. & Li, Z. A novel framework for remote photoplethysmography pulse extraction on compressed videos. In *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 1380–138009 (IEEE, 2018).
 152. Álvarez Casado, C. & Bordallo López, M. Face2PPG: an unsupervised pipeline for blood volume pulse extraction from faces. *IEEE J. Biomed. Health Inform.* **27**, 5530–5541 (2023).
 153. Karmuse, S. M., Kakhandki, A. L. & Anandhalli, M. A robust rPPG approach for continuous heart rate measurement based on face. *J. Inst. Eng.* **106**, 849–860 (2022).
 154. Chen, W. & McDuff, D. DeepPhys: video-based physiological measurement using convolutional attention networks. In *Proc. European Conference on Computer Vision (ECCV)*, 349–365 (Springer, 2018).
 155. Liu, X., Fromm, J., Patel, S. & McDuff, D. Multi-task temporal shift attention networks for on-device contactless vitals measurement. In (eds Larochelle, H., Ranzato, M., Hadsell, R., Balcan, M. & Lin, H.) *Advances in Neural Information Processing Systems*, Vol 33, 19400–19411 (Curran Associates, Inc., 2020).
 156. Yu, Z., Li, X. & Zhao, G. Remote photoplethysmograph signal measurement from facial videos using spatio-temporal networks. In *Proc. 30th British Machine Vision Conference 2019, BMVC 2019*, 277 (BMVA Press, 2019).
 157. Hansen, P., Lozano, M. G., Kamrani, F. & Brynielsson, J. Real-time estimation of heart rate in situations characterized by dynamic illumination using remote photoplethysmography. In *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 6094–6103 (IEEE, 2023).
 158. Lee, J. S., Hwang, G., Ryu, M. & Lee, S. J. LSTC-rPPG: long short-term convolutional network for remote photoplethysmography. In *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 6014–6022 (IEEE, 2023).
 159. Botina-Monsalve, D., Benezeth, Y. & Miteran, J. RT-rPPG: an ultra light 3DCNN for real-time remote photoplethysmography. In *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2146–2154 (IEEE, 2022).
 160. Speth, J., Vatanparvar, K., Zhu, L., Kuang, J. & Gao, A. Freq2Time: weakly supervised learning of camera-based rPPG from heart rate. In *ICASSP 2024-2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 6880–6884 (IEEE, 2024).
 161. Liu, S.-Q. & Yuen, P. C. Robust remote photoplethysmography estimation with environmental noise disentanglement. *IEEE Trans. Image Process.* **33**, 27–41 (2024).
 162. Huang, Z., Fu, C.-H., Zhang, L. & Hong, H. Deep learning-based robust heart rate estimation using remote photoplethysmography under different illuminations. In *Proc. Fifteenth International Conference on Signal Processing Systems (ICSPS 2023)*, Vol 13091, 183–191 (SPIE, 2024).
 163. Xiao, H. et al. SimFuPulse: a self-similarity supervised model for remote photoplethysmography extraction from facial videos. *Biomed. Signal Process. Control* **98**, 106736 (2024).
 164. Zhai, D. et al. Research on robust measurement method of heart rate using remote photoplethysmography based on adversarial learning network with high and low frequency features. *IEEE Trans. Circuits Syst. Video Technol.* **35**, 5208–5222 (2025).
 165. Han, K. et al. A survey on vision transformer. *IEEE Trans. Pattern Anal. Mach. Intell.* **45**, 87–110 (2022).
 166. Zhang, X., Xia, Z., Liu, L. & Feng, X. Demodulation based transformer for rPPG generation and heart rate estimation. *IEEE Signal Process. Lett.* **30**, 1042–1046 (2023).
 167. Shao, H. et al. TranPhys: spatiotemporal masked transformer steered remote photoplethysmography estimation. *IEEE Trans. Circuits Syst. Video Technol.* **34**, 3030–3042 (2024).
 168. Shao, H. et al. TranPulse: remote photoplethysmography estimation with time-varying supervision to disentangle multi-physiologically interference. *IEEE Trans. Instrum. Meas.* **73**, 1–11 (2024).
 169. Salah, M. K. B., Jouvet, P. & Noumeir, R. rPPG estimation: vision transformer with 3D temporal central difference. *IEEE Trans. Instrum. Meas.* **74**, 1–13 (2025).
 170. Zhang, H. et al. MaKAN-Mixer: Channel interaction-based Mamba method for rPPG extraction. *IEEE J. Biomed. Health Inform.* <https://doi.org/10.1109/JBHI.2025.3532488> (2025).

171. Xiong, J. et al. STGNet: spatio-temporal graph neural networks considering inherent properties of physiological signals for camera-based remote photoplethysmography. *Biomed. Signal Process. Control* **98**, 106690 (2024).
172. Liu, M. et al. Spiking-PhysFormer: camera-based remote photoplethysmography with parallel spike-driven transformer. *Neural Netw.* **185**, 107128 (2025).
173. Hsu, Y., Lin, Y.-L. & Hsu, W. Learning-based heart rate detection from remote photoplethysmography features. In *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 4433–4437 (IEEE, 2014).
174. Chowdhury, M. H. et al. LGI-rPPG-Net: a shallow encoder-decoder model for rPPG signal estimation from facial video streams. *Biomed. Signal Process. Control* **89**, 105687 (2024).
175. Lokendra, B. & Puneet, G. AND-rPPG: a novel denoising-rPPG network for improving remote heart rate estimation. *Comput. Biol. Med.* **141**, 105146 (2022).
176. Wu, B.-F., Huang, P.-W., He, D.-H., Lin, C.-H. & Chen, K.-H. Remote photoplethysmography enhancement with machine learning methods. In *Proc. IEEE International Conference on Systems, Man and Cybernetics (SMC)*, 2466–2471 (IEEE, 2019).
177. Wu, B.-F., Chu, Y.-W., Huang, P.-W. & Chung, M.-L. Neural network based luminance variation resistant remote-photoplethysmography for driver's heart rate monitoring. *IEEE Access* **7**, 57210–57225 (2019).
178. Ma, J.-R., Jia, R.-S., Zhang, M.-Q., Jiang, M.-X. & Sun, H.-M. CTA-Net: a lightweight network for remote photoplethysmography signal estimation with channel-temporal attention. *IEEE Internet Things J.* **12**, 26855–26867 (2025).
179. Jaiswal, K. B., Choubey, M., Mishra, A., Singh, D. & Meenpal, T. Color space analysis for improvement in rPPG. In *Proc. International Conference for Advancement in Technology (ICONAT)*, 1–7 (IEEE, 2023).
180. Jaiswal, K. B. & Meenpal, T. rPPG-FuseNet: non-contact heart rate estimation from facial video via RGB/MSR signal fusion. *Biomed. Signal Process. Control* **78**, 104002 (2022).
181. Chen, S. et al. Deep learning-based image enhancement for robust remote photoplethysmography in various illumination scenarios. In *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 6076–6084 (IEEE, 2023).
182. Sun, W., Sun, Q., Sun, H.-M., Sun, Q. & Jia, R.-S. ViT-rPPG: a vision transformer-based network for remote heart rate estimation. *J. Electron. Imaging* **32**, 023024–023024 (2023).
183. Yu, S.-N., Wang, C.-S. & Chang, Y. P. Heart rate estimation from remote photoplethysmography based on light-weight U-Net and attention modules. *IEEE Access* **11** (2023).
184. Liu, X. et al. rPPG-MAE: self-supervised pretraining with masked autoencoders for remote physiological measurements. *IEEE Trans. Multimed.* **26**, 7278–7293 (2024).
185. Qian, W. et al. Dual-path tokenlearner for remote photoplethysmography-based physiological measurement with facial videos. *IEEE Trans. Comput. Soc. Syst.* **11**, 4465–4477 (2024).
186. Butterworth, S. et al. On the theory of filter amplifiers. *Wirel. Eng.* **7**, 536–541 (1930).
187. Guler, S., Golparvar, A., Ozturk, O., Dogan, H. & Kaya Yapici, M. Optimal digital filter selection for remote photoplethysmography (rPPG) signal conditioning. *Biomed. Phys. Eng. Express* **9**, 027001 (2023).
188. Gupta, K., Sinhal, R. & Badhiye, S. S. Remote photoplethysmography-based human vital sign prediction using cyclical algorithm. *J. Biophotonics* **17**, e202300286 (2024).
189. Tarvainen, M. P., Ranta-Aho, P. O. & Karjalainen, P. A. An advanced detrending method with application to HRV analysis. *IEEE Trans. Biomed. Eng.* **49**, 172–175 (2002).
190. Ahmadi, N. et al. Development and evaluation of a contactless heart rate measurement device based on rPPG. In *Proc. 29th IEEE International Conference on Electronics, Circuits and Systems (ICECS)*, 1–4 (IEEE, 2022).
191. Abdulrahman, L. Q. Two-stage motion artifact reduction algorithm for rPPG signals obtained from facial video recordings. *Arab. J. Sci. Eng.* **49**, 2925–2933 (2024).
192. Kopeliovich, M., Mironenko, Y. & Petrushan, M. Architectural tricks for deep learning in remote photoplethysmography. In *Proc. IEEE/CVF International Conference on Computer Vision Workshops*, 0–0 (2019).
193. Boccignone, G. et al. An open framework for remote-PPG methods and their assessment. *IEEE Access* **8**, 216083–216103 (2020).
194. Boccignone, G. et al. pyVHR: a Python framework for remote photoplethysmography. *PeerJ Comput. Sci.* **8**, e929 (2022).
195. Gideon, J. & Stent, S. The way to my heart is through contrastive learning: remote photoplethysmography from unlabelled video. In *Proc. IEEE/CVF International Conference on Computer Vision (ICCV)*, 3975–3984 (IEEE, 2021).
196. Song, R. et al. Uncertainty quantification for deep learning-based remote photoplethysmography. *IEEE Trans. Instrum. Meas.* **72**, 1–12 (2023).
197. Consumer Technology Association. Physical activity monitoring for heart rate - real-world analysis. Technical Report. ANSI/CTA-2065.1 (Consumer Technology Association, 2023).
198. Hatib, N. A. A. et al. A two-phased study on the use of remote photoplethysmography (rppg) in paediatric care. *Ann. Transl. Med.* **12**, 46 (2024).
199. Wegerif, S. et al. Paediatric pulse rate measurements: a comparison of methods using remote photoplethysmography. In *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 364–370 (IEEE, 2024).
200. Momeni, M. et al. Facial remote photoplethysmography for continuous heart rate, stroke volume, and systemic vascular resistance monitoring during prolonged warm/cold fluid bolus administration. *IEEE Sens. J.* **25**, 16151–16169 (2025).
201. Lee, R. J., Sivakumar, S., Lim, K. H., Chee, A. L. T. & Yon, J. L. S. Heart rate estimation using remote photoplethysmography and deep learning. In *Proc. International Conference on Digital Applications, Transformation & Economy (ICDATE)*, 1–5 (IEEE, 2023).
202. Liu, X., Hill, B., Jiang, Z., Patel, S. & McDuff, D. Efficientphys: enabling simple, fast and accurate camera-based cardiac measurement. In *Proc. IEEE/CVF Winter Conference on Applications of Computer Vision*, 5008–5017 (IEEE, 2023).
203. Chung, M., Won, J. B., Kim, G., Kim, Y. & Ozbulak, U. Evaluating visual explanations of attention maps for transformer-based medical imaging. In *Proc. International Conference on Medical Image Computing and Computer-Assisted Intervention*, 110–120 (Springer, 2024).
204. Zhan, Q., Wang, W. & De Haan, G. Analysis of CNN-based remote-PPG to understand limitations and sensitivities. *Biomed. Opt. Express* **11**, 1268–1283 (2020).
205. Tohma, A., Nishikawa, M., Hashimoto, T., Yamazaki, Y. & Sun, G. Evaluation of remote photoplethysmography measurement conditions toward telemedicine applications. *Sensors* **21**, 8357 (2021).
206. Curran, T. et al. Camera-based remote photoplethysmography for blood pressure measurement: current evidence, clinical perspectives, and future applications. *Connect. Health Telemed.* **2**, N-A (2023).
207. Huang, B. et al. Challenges and prospects of visual contactless physiological monitoring in clinical study. *npj Digit. Med.* **6**, 231 (2023).
208. Jokinen, E. et al. Understanding the feasibility of computer vision in diagnosing respiratory infections in pediatric emergency rooms. Preprint at medRxiv <https://doi.org/10.1101/2025.04.17.25325898> (2025).

209. Koelstra, S. et al. Deap: a database for emotion analysis; using physiological signals. *IEEE Trans. Affect. Comput.* **3**, 18–31 (2011).
210. Yu, Z., Li, X. & Zhao, G. Facial-video-based physiological signal measurement: recent advances and affective applications. *IEEE Signal Process. Mag.* **38**, 50–58 (2021).
211. Chen, W. et al. Deep learning and remote photoplethysmography powered advancements in contactless physiological measurement. *Front. Bioeng. Biotechnol.* **12**, 1420100 (2024).
212. Stevenson, L. W., Ross, H. J., Rathman, L. D. & Boehmer, J. P. Remote monitoring for heart failure management at home: Jacc scientific statement. *J. Am. Coll. Cardiol.* **81**, 2272–2291 (2023).
213. Ong, M. K. et al. Effectiveness of remote patient monitoring after discharge of hospitalized patients with heart failure: the better effectiveness after transition–heart failure (beat-hf) randomized clinical trial. *JAMA Intern. Med.* **176**, 310–318 (2016).
214. Wu, J.-Y., Wang, Y., Ching, C. T. S., Wang, H.-M. D. & Liao, L.-D. lot-based wearable health monitoring device and its validation for potential critical and emergency applications. *Front. Public Health* **11**, 1188304 (2023).
215. McCool, J., Dobson, R., Whittaker, R. & Paton, C. Mobile health (mhealth) in low-and middle-income countries. *Annu. Rev. Public Health* **43**, 525–539 (2022).
216. Canali, S., Schiaffonati, V. & Aliverti, A. Challenges and recommendations for wearable devices in digital health: Data quality, interoperability, health equity, fairness. *PLoS Digit. Health* **1**, e0000104 (2022).
217. Wu, Y.-C. et al. Motion robust remote photoplethysmography measurement during exercise for contactless physical activity intensity detection. *IEEE Trans. Instrum. Meas.* **72**, 1–14 (2023).
218. Dawoodjee, I. & Ghahramani, A. A flexible framework for design and validation of rPPG methods. *IEEE Access* **13**, 50059–50076 (2025).
219. Jayasumana, S. et al. Rethinking fid: towards a better evaluation metric for image generation. In *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 9307–9315 (IEEE, 2024).
220. Yan, B.-R., Rios, E. A., Lee, W.-H. & Lai, B.-C. DLPrPPG: development and design of deep learning platform for remote photoplethysmography. In *Proc. IEEE International Symposium on Circuits and Systems (ISCAS)*, 697–701 (IEEE, 2022).
221. Kim, D.-Y., Lee, K. & Sohn, C.-B. Assessment of ROI selection for facial video-based rPPG. *Sensors* **21**, 7923 (2021).
222. Li, S., Elgendi, M. & Menon, C. Optimal facial regions for remote heart rate measurement during physical and cognitive activities. *npj Cardiovasc. Health* **1**, 33 (2024).
223. Khiroudkar, R. et al. Sapiens: Foundation for human vision models. In *Proc. European Conference on Computer Vision*, 206–228 (Springer, 2024).
224. Niu, X., Han, H., Shan, S. & Chen, X. Continuous heart rate measurement from face: a robust rPPG approach with distribution learning. In *Proc. IEEE International Joint Conference on Biometrics (IJCB)*, 642–650 (IEEE, 2017).
225. Macwan, R., Bobbia, S., Benezeth, Y., Dubois, J. & Mansouri, A. Periodic variance maximization using generalized eigenvalue decomposition applied to remote photoplethysmography estimation. In *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 1413–14138 (IEEE, 2018).
226. Li, J. et al. Enhance heart rate measurement from remote PPG with head motion awareness from image. In *Proc. 46th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 1–4 (IEEE, 2024).
227. Vance, N. & Flynn, P. Refining remote photoplethysmography architectures using CKA and empirical methods. In *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 380–388 (IEEE, 2024).
228. Zhao, C., Cao, P., Xu, S., Li, Z. & Feng, Y. Pruning rPPG networks: toward small dense network with limited number of training samples. In *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2055–2064 (IEEE, 2022).
229. Chen, Y., Gou, X. & Yang, C. ZYNQ-based real-time heart rate estimation system using remote photoplethysmography. In *Proc. IEEE 19th Conference on Industrial Electronics and Applications (ICIEA)*, 1–6 (IEEE, 2024).
230. Álvarez Casado, C. et al. A distributed framework for remote multimodal biosignal acquisition and analysis. In *Proc. Nordic Conference on Digital Health and Wireless Solutions*, 127–146 (Springer, 2024).
231. Zhu, D. et al. Leveraging identity-specific facial contours for enhanced heart rate estimation in remote photoplethysmography. In *Proc. IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 4123–4128 (IEEE, 2024).
232. Lee, E., Chen, E. & Lee, C.-Y. Meta-rPPG: remote heart rate estimation using a transductive meta-learner. In *Proc. Computer Vision–ECCV 2020: 16th European Conference*, 392–409 (Springer, 2020).
233. Wu, B. et al. Semi-rPPG: semi-supervised remote physiological measurement with curriculum pseudo-labeling. *IEEE Trans. Instrum. Meas.* **74**, 1–11 (2025).
234. Cho, D., Kim, J., Lee, K. J. & Kim, S. Reduction of motion artifacts from remote photoplethysmography using adaptive noise cancellation and modified HSI model. *IEEE Access* **9**, 122655–122667 (2021).
235. Aslami, F. & Sripian, P. Comparative analysis of rPPG algorithms across diverse countries of origin. In *Proc. International Seminar on Application for Technology of Information and Communication (iSemantic)*, 290–293 (IEEE, 2024).
236. Chen, S. et al. An image enhancement based method for improving rppg extraction under low-light illumination. *Biomed. Signal Process. Control* **100**, 106963 (2025).
237. Page, M. J. et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* **372**, n71 (2021).
238. Rethlefsen, M. L. & Page, M. J. PRISMA 2020 and PRISMA-S: common questions on tracking records and the flow diagram. *J. Med. Libr. Assoc.* **110**, 253 (2022).
239. Borik, S. et al. Photoplethysmography imaging to assess facial perfusion under simulated hypovolemia. *Physiol. Meas.* **46**, 075008 (2025).
240. Litong, F., Lai-Man, P., Xuyuan, X., Yuming, L. & Ruiyi, M. Motion-resistant remote imaging photoplethysmography based on the optical properties of skin. *IEEE Trans. Circuits Syst. Video Technol.* **25**, 879–891 (2015).
241. Song, R. et al. PulseGAN: learning to generate realistic pulse waveforms in remote photoplethysmography. *IEEE J. Biomed. Health Inform.* **25**, 1373–1384 (2021).
242. Hu, M. et al. rPPG-based heart rate estimation using spatial-temporal attention network. *IEEE Trans. Cogn. Dev. Syst.* **14**, 1630–1641 (2021).
243. Das, M., Bhuyan, M. & Sharma, L. Time-frequency learning framework for rPPG signal estimation using scalogram-based feature map of facial video data. *IEEE Trans. Instrum. Meas.* **72**, 1–10 (2023).
244. Yang, Z., Wang, H., Lu, F. & Zhao, Q. Towards practical facial video-based remote heart rate estimation via cross domain rPPG adaptation. In *Proc. 2023 10th International Conference on Biomedical and Bioinformatics Engineering*, 154–161 (ACM, 2023).

Acknowledgements

M.E. acknowledges the support from Khalifa University through the Research Grant FSU-2025-001.

Author contributions

M.E. designed and led the study. The literature search was conducted by two reviewers: S.L. and M.E. Both reviewers collaborated in constructing the search strategy. S.L. conducted the initial literature search, and M.E. independently verified the search results and confirmed the extracted data. C.M. contributed to the study's conceptualization and provided translational input. A.A. conceived the experiments. S.B., K.H.S., and A.A. conducted the clinical experiments and provided clinical input. S.B. implemented and adapted the algorithmic pipeline for processing the rPPG data showcased in the study. E.Y.L. contributed narrative statistical analyses. B.A. provided public health input. All authors reviewed the manuscript.

Competing interests

M.E. is an Editorial Board Member of npj Biosensing but was not involved in the editorial handling or peer review of this manuscript. The other authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41746-026-02715-1>.

Correspondence and requests for materials should be addressed to Mohamed Elgendi or Aymen Alian.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2026