Validating ViSCOPE: Vital Signs Contactless Estimation Pipeline for Robot-Aided Rehabilitation

Rita Molle[1], Christian Tamantini[1,2,*], Fabrizio Taffoni[1], Andrea Caroppo[3], Andrea Manni[3], Pietro Aleardo Siciliano[3], Loredana Zollo[1,†] and Alessandro Leone[3,†]

[1]*Research Unit of Advanced Robotics and Human-Centred Technologies, Università Campus Bio-Medico di Roma, Via Alvaro del Portillo 21, 00128, Rome, Italy*

[2]*Institute of Cognitive Sciences and Technologies, National Research Council of Italy, Via Giandomenico Romagnosi 18a, 00196, Rome, Italy*

[3]*Institute for Microelectronics and Microsystems, National Research Council of Italy, Via per Monteroni c/o Campus Universitario Palazzina A3, 73100, Lecce, Italy*

Abstract

In rehabilitation settings, accurate monitoring of physiological parameters is critical for tailoring therapeutic interventions and ensuring patient safety. This paper introduces ViSCOPE, a contactless vital sign estimation pipeline, and its integration into a robotic rehabilitation platform, aimed at providing non-invasive monitoring of heart rate, breathing rate, and oxygen saturation.

ViSCOPE was validated against gold-standard devices in both resting conditions and after physical exertion, simulating the demands of rehabilitation exercises. In resting conditions, the system achieved mean absolute errors of 5.50 ± 4.91 bpm for heart rate, 5.13 ± 2.86 bpm for breathing rate, and $1.23 \pm 0.43\%$ for oxygen saturation. However, after physical activity, the error committed in estimating the heart rate significantly increased up to 13.10 ± 8.79 bpm ($p < 1.00 \cdot 10^{-3}$), indicating a reduced accuracy.

While ViSCOPE provides reliable monitoring in resting conditions, further refinement is required to improve accuracy during physical exertion, a common scenario in rehabilitation. Future efforts will focus on optimizing the system for real-world clinical applications, ensuring accurate monitoring throughout dynamic therapy sessions.

Keywords

Remote PPG, Contactless Monitoring, Robot-aided Rehabilitation, Multimodal Monitoring

1. Introduction

Rehabilitation robots have become indispensable instruments in both physical and cognitive therapy [1]. Moreover, the introduction of multimodal monitoring systems enables the realization of sophisticated physiological monitoring systems [2]. Such systems facilitate the continuous assessment of vital parameters, thereby providing real-time insights into the patient's health status during therapy sessions [3]. This data-driven approach permits the continuous estimation of the complex user state to dynamically adjust the therapeutic interventions, thereby ensuring that treatments are personalized to the patient's functional capacity and psychophysiological responses [4].

Conventional techniques for measuring vital signs are typically contact-based. For example, commercial and low-cost wearable smart sensors enable signal processing algorithms to provide discrete Heart Rate (HR), Breathing Rate (BR), or blood oxygenation (SpO_2) values, becoming increasingly popular in the market. The main problem with wearable sensor-based monitoring is that end-users with different

Workshop on Advanced AI Methods and Interfaces for Human-Centered Assistive and Rehabilitation Robotics (a Fit4MedRob event) - AIxIA 2024

*Corresponding author.

†These authors contributed equally.

rita.molle@unicampus.it (R. Molle); christian.tamantini@cnr.it (C. Tamantini); f.taffoni@unicampus.it (F. Taffoni); andrea.caroppo@cnr.it (A. Caroppo); andrea.manni@cnr.it (A. Manni); pietroaleardo.siciliano@cnr.it (P. A. Siciliano); l.zollo@unicampus.it (L. Zollo); alessandro.leone@cnr.it (A. Leone)

0009-0004-4603-2654 (R. Molle); 0000-0001-6238-2241 (C. Tamantini); 0000-0003-3215-1375 (F. Taffoni); 0000-0003-0318-8347 (A. Caroppo); 0000-0001-5716-5824 (A. Manni); 0000-0002-1312-4593 (P. A. Siciliano); 0000-0002-8015-010X (L. Zollo); 0000-0002-8970-3313 (A. Leone)

levels of cognitive function must remember to wear the devices and charge them frequently, restricting their use. Additionally, data quality can be compromised due to movement artifacts, environmental factors, or improper sensor placement, further restricting the reliability of the monitoring. Alternatively, contactless sensors are less invasive and can track vital signs in authentic, lifelike settings without interfering with a person's regular activities. It is important to note, also, that the COVID-19 pandemic has led to a rise in the use of noncontact technology for vital sign monitoring. The literature review revealed that there are currently two primary categories of contactless vital sign estimation methods from RGB images: motion-based methods and methods that assess color intensity changes. The latter are the most widely used and effective techniques. Their foundation lies in the analysis of the remote plethysmographic signal (rPPG) obtained from the vision sensor [5, 6]. The integration of such technology on robotic platforms is widely investigated in the literature. For example, in [7] the authors developed a camera system consisting of one Infrared camera and three monochrome cameras to reliably facilitate the contactless acquisition of vital sign parameters. This camera system was mounted on a teleoperated robot that can successfully and reliably deliver vital sign measurements while navigating complex clinical environments and maintaining social distancing. Takir et al. [8] published a study in which children with autism spectrum disorder interacted and played games with a robot. rPPG signals were extracted using face images captured through the camera during the interaction between the robot and the children. The study aimed to use rPPG signals in emotion recognition as an alternative method since other emotion recognition modalities face challenges during robot-child interaction. The authors of [9] proposed an agile quadruped robotic system that comprises a set of contactless monitoring systems for measuring vital signs (skin temperature, BR, HR, and SpO_2) and a tablet computer to enable face-to-face medical interviewing. In the context of robotic platforms, the topic under consideration (contactless vital signs evaluation) presents critical issues that have been only partially addressed in scientific works, such as subject motion, ambient light illumination, distance from the camera, estimation of non-at-rest vital signs, and motion artifacts. Also, most of the previous studies of non-contact video-based vital signs monitoring have been on healthy volunteers, sometimes in sunlight only, and have usually concentrated on HR estimation or BR estimation, without considering other important vital signs such as SpO_2.

Previous studies exhibited some limitations. Firstly, to the best of the authors' knowledge, there are currently no contactless approaches that simultaneously estimate HR, BR, and SpO_2 from RGB images, as most techniques focus on one or, at best, two parameters. Secondly, the accuracy of contactless systems has never been validated under conditions of physical exertion, which is commonly present during rehabilitation sessions, where patients perform exercises that directly impact physiological parameters. These gaps limit the applicability of existing systems in dynamic and intensive clinical settings like rehabilitation.

Therefore, this paper aims to propose the Vital Signs Contactless Estimation Pipeline (ViSCOPE) and its integration into a service robot employed in rehabilitative healthcare settings. Specifically, the robot can act as a robotic coach in the context of robot-aided rehabilitation, administering rehabilitation exercises, providing corrective feedback, and even physically interacting with the patient. Moreover, the robotic therapist can guide the patient through exercises, offering corrective feedback, and assisting in movement execution when needed [10]. Equipping robots used for rehabilitative coaching with physiological monitoring tools like ViSCOPE simplifies the setup, ensuring physiological monitoring throughout the therapy sessions. Moreover, the accuracy of the proposed ViSCOPE system is validated with respect to gold-standard measurements, both in resting conditions and after physical activity, resembling the altered physiological status of patients engaged in physical therapy.

2. Materials and Methods

2.1. Vital Sign Contactless estimation Pipeline

Our proposed ViSCOPE input is represented by a stream of images obtained from an RGB image, collected by a service robot, from which the rPPG is extracted and processed for the subsequent vital

signs estimation. Figure 1 depicts the algorithmic pipeline through logical blocks. Regarding the pre-processing phase, face detection is one of the most crucial algorithmic steps. Here, the Mediapipe library [11] was used to make face detection as independent of orientation and/or distance as feasible. Mediapipe uses Machine Learning to create 3D surface geometry with only one camera input. Subsequently, for effective signal extraction, a facial image-based remote rPPG algorithm also needs to choose a Region of Interest (ROI) within the identified facial region. Numerous studies have demonstrated that the facial regions where the rPPG signal is strongest are the forehead and the cheeks. Accordingly, in the present work, the facial regions are identified by the corresponding facial landmarks returned by the Mediapipe library and the rPPG signal was processed from the R, G, and B color information extracted from such ROIs.

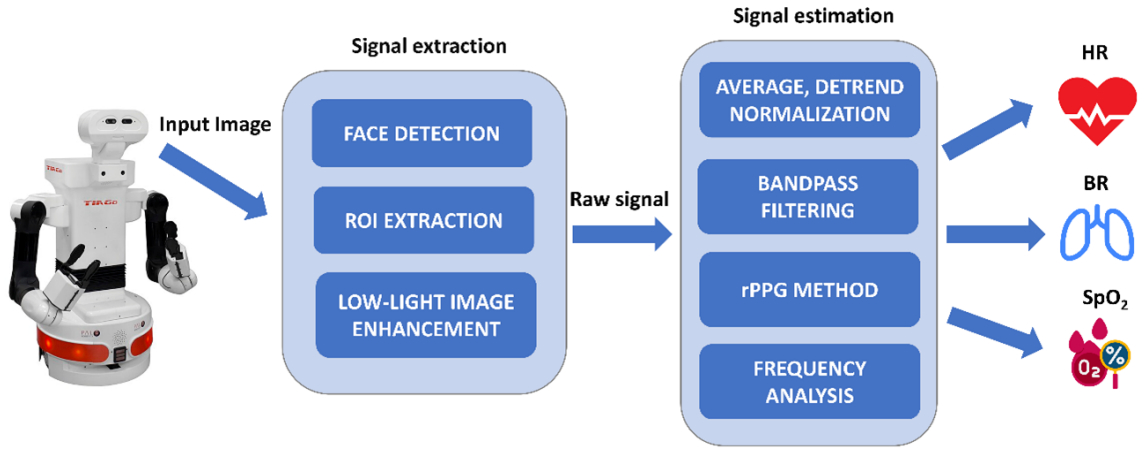

Figure 1: Graphical representation through block diagram of the pipeline designed and implemented for vital signs estimation by rPPG signal extracted from the facial ROI

Subsequently, an algorithm for low-light image enhancement was employed for improving the brightness of the chosen ROIs. The algorithm is designed to balance ROI brightness while preserving the details, such as color variations within the forehead and cheeks that are related to blood volume changes. Unlike traditional histogram equalization methods, the algorithm implements an approach based on an upgraded version of Cuckoo Search methodology [12]. After applying the algorithmic blocks included in the pre-processing step, the subject's head motion and ambient lighting variations are significantly less disruptive to the values in the RGB components.

Considering a continuous monitoring scenario, the video segment that is used for vital sign estimation is represented by a sliding time window. This is the portion of the video through which the discrete value of HR, BR, and SpO_2 is computed. In our work, the raw RGB signals are segmented using a 30-second moving window, with a 1-second scroll between consecutive windows. After extracting the raw signals for each frame belonging to the sliding window, specific signal processing techniques are employed to improve signal quality for the subsequent feature extraction steps. Consequently, detrending is applied to remove linear trends from the raw signal and, since the interest is in the periodicity of the signal, the resulting raw signal is normalized by dividing it by its maximum absolute value and smoothed using a sliding average filter. To obtain accurate measurements of HR and BR, it is essential to process the raw RGB signals by filtering out unrealistic frequencies. To this end, a third-order band-pass filter with optimal characteristics is employed to eliminate both high and low-frequency noise. This filtering process specifically targets frequency components outside the designated ranges: [0.75Hz – 3.5Hz] for HR and [0.15Hz - 0.5Hz] for BR.

Next, in the proposed pipeline, after having assessed the latest findings in the scientific literature, the chrominance-based method [13] is used for the temporal reconstruction of the rPPG signal. At this point, given an estimate of the plethysmography signal, HR and BR can be estimated using frequency analysis. For this purpose, this signal, which contains a distinct periodicity, is converted to the frequency

domain using the Fast Fourier Transform. To calculate the average estimations for HR and BR, the frequencies associated with the peaks of the power spectrum are analyzed. For HR estimation, we focus on the frequency that corresponds to the highest peak intensity within the range of [0.75Hz - 3.5Hz], whereas for BR estimation, we examine the frequency with the maximum peak intensity in the range of [0.15Hz - 0.5Hz]. The HR and BR values, represented as average beats per minute, are derived by multiplying the obtained frequency values by 60. Finally, to evaluate the values of SpO_2 we have reproduced the approach described in [14] and based on the calculation of the ratio of the concentration of oxygenated hemoglobin to the total concentration of hemoglobin present in the blood.

2.2. Experimental Evaluation

The experimental setup used for the software validation is shown in Fig. 2 and includes:

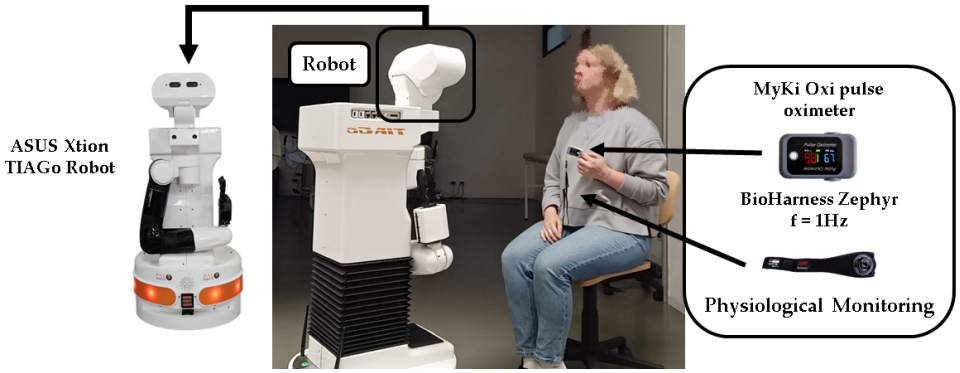

Figure 2: Experimental setup.

- the robot TIAGo (PAL Robotics S.L., Barcelona, Spain), whose head-mounted ASUS Xtion camera works at 30 Hz and has a resolution of 640 × 480;
- the BioHarness 3.0 chest belt, developed by Zephyr^{TM} Technology, to record the cardiorespiratory activity (i.e. HR and BR). The chest belt is worn on the skin at the sternum level and works at a frequency of 1 Hz;
- the MyKi Oxy pulse oximeter to record the oxygen saturation levels in the arterial blood (SpO_2). It works with a frequency of 1 Hz.

The overall systems were developed and integrated in Robot Operating System (ROS) Melodic middleware on Ubuntu 18.04 LTS.

Six healthy subjects (4 males, 2 females, with an average age of 27.67 ± 2.58) were enrolled for the ViSCOPE validation. The experimental procedure involved two recording sessions: one under resting condition (denoted as Rest, R) and another following one minute of physical activity (denoted as Physical Activity, PA) performed by the subject. The physical activity consisted of performing jumping jacks for one minute to elevate the subject's physiological parameters. Each subject underwent both conditions, each repeated five times and lasting 30 seconds, resulting in ten recordings per participant. The ambient lighting was maintained between 300 and 500 lux throughout the experiment to ensure consistency.

During each recording session, the subject sat stationary in front of the TIAGo robot, positioned at a fixed distance of 50 cm. The subject was instructed to maintain a steady gaze at the robot, avoiding any movement, for the entire 30-second duration of the recording.

While the recordings were ongoing, physiological parameters were estimated by the software (HR_{vs}, BR_{vs}, $SpO_{2,vs}$), and a value was given at the end of each 30-second acquisition period. Simultaneously, the gold-standard (GS) devices, BioHarness and Miki Oxy (as previously defined), continuously recorded physiological parameters throughout the 30 seconds, providing measurements at their respective sampling frequencies (HR_{gs}, BR_{gs}, $SpO_{2,gs}$).

To validate the ViSCOPE software, the Absolute Error (AE) is calculated to assess how closely the estimated values from the proposed approach match the values obtained from the GS devices. The purpose of this comparison is to evaluate the accuracy of the ViSCOPE software in estimating the three physiological parameters. The AE of the physiological parameters was computed with the following equation:

$$|\Delta X| = |X_{vs} - X_{gs}| \tag{1}$$

where X represents the physiological parameter (i.e. HR, BR, SpO_2). Specifically, X_{gs} is the last value recorded by the corresponding GS device, thus to the value acquired at the 30^{th} second.

The closer the AE is to zero, the more reliable the ViSCOPE software is in reproducing the GS measurements. This approach helps to demonstrate whether the proposed system can be trusted for monitoring physiological data and how accurately it can replicate the true values obtained from clinically validated devices. Therefore, AE serves as a critical metric for validating the performance of the software in real-world scenarios.

Moreover, the Mann-Whitney U test with a p-value of 0.05 was applied on the AE of each parameter to evaluate if there were statistically significant differences between the two conditions (i.e. R and PA).

3. Results and Discussion

The results obtained in terms of AE and with the statistical test application are shown in Fig. 3.

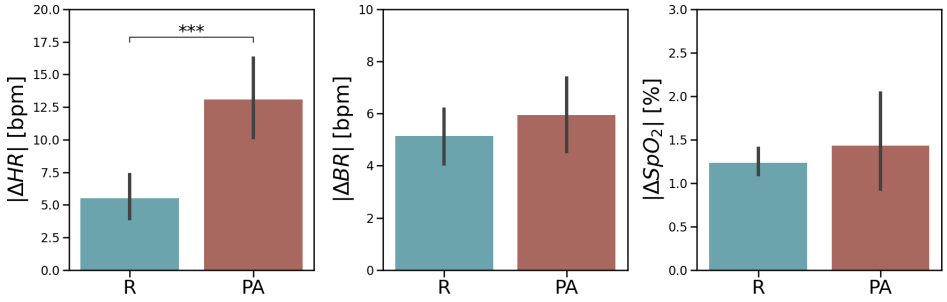

Figure 3: Results in terms of absolute error between the value estimated by the software and the gold-standard value for each physiological parameter (i.e. $|\Delta HR|$, $|\Delta BR|$, $|\Delta SpO_2|$) under both condition (Rest R and after Physical Activity PA). Moreover, statistical results are shown: applying Mann-Withney U test are shown. The statistical test was applied between the R and PA conditions. *** means that the statistical difference between those groups has achieved a $1.00 \cdot 10^{-4} < p\text{-value} \leq 1.00 \cdot 10^{-3}$.

For the R condition, the mean AE for HR (i.e. $|\Delta HR|$) is 5.50 ± 4.91 bpm, indicating relatively low error and variability in HR estimation during rest. The $|\Delta BR|$ shows a mean error of 5.13 ± 2.86 bpm, demonstrating a moderate error and lower variability in BR measurements. Lastly, the $|\Delta SpO_2|$ has a mean of $1.23 \pm 0.43\%$, indicating a small and consistent error in SpO_2 estimation.

For the PA condition, the $|\Delta HR|$ increases significantly with respect to the same metric in the R condition, with a mean of 13.10 ± 8.79 bpm, reflecting a statistically significant higher error and greater variability during physical exertion. The $|\Delta BR|$ shows a slightly higher error than the one obtained in R, pointing to more variability in BR measurements during activity (5.93 ± 3.96 bpm). The $|\Delta SpO_2|$ also increases slightly, with a mean of $1.43 \pm 1.59\%$, indicating more variability but still relatively low error in oxygen saturation estimation.

These results suggest that while the ViSCOPE system performs reasonably well in the R condition, with results comparable to previous studies [6], there is a notable increase in error, particularly in HR, in the acquisition made after PA. Importantly, the PA condition represents a novelty in this study and can not be directly compared to earlier research, as similar conditions have not been previously addressed. Since the ViSCOPE acquisitions will be performed using the robot in a rehabilitation context

for hospitalized patients, it is crucial to optimize the software to ensure more accurate monitoring of physiological parameters after subjects have undergone rehabilitation (i.e. in PA condition). Enhancing the accuracy of the software in these scenarios will improve the reliability of the data collected during rehabilitation sessions.

4. Conclusions

This study introduced ViSCOPE, a contactless vital sign estimation pipeline integrated into a robotic platform, aimed at enhancing physiological monitoring during rehabilitation sessions. The system was validated by comparing its accuracy against GS devices in both resting and post-physical activity conditions. Results indicate that ViSCOPE performs reliably in resting conditions, with a mean absolute error of 5.50 ± 4.91 bpm for HR, 5.13 ± 2.86 bpm for BR, and $1.23 \pm 0.43\%$ for SpO_2, demonstrating low error and variability. However, a notable increase in error was observed in post-physical activity conditions, with heart rate error rising to 13.10 ± 8.79 bpm, highlighting a statistically significant difference from resting values ($p < 1.00 \cdot 10^{-3}$).

These findings suggest that while ViSCOPE provides reliable estimates in resting conditions, further refinement is needed to improve its accuracy during and after physical exertion, a scenario common in rehabilitation sessions. Future works will be devoted to enhancing the system performance in presence of physical exertion to provide accurate estimations in real clinical settings. Additionally, increasing the participant pool can be useful to strengthen the validity of the results, and comparing outcomes from the current environmental conditions with those from a brighter lighting environment will provide insights into the impact of environmental factors on accuracy.

Acknowledgments

This work was supported by the Italian Ministry of Research, under the complementary actions to the NRRP "Fit4MedRob - Fit for Medical Robotics" Grant PNC0000007, (CUP: B53C22006990001). Rita Molle is a PhD student enrolled in the National PhD in Artificial Intelligence, XXXVIII cycle, course on Health and life sciences, organized by Università Campus Bio-Medico di Roma.

References

[1] C. Tamantini, F. Cordella, C. Lauretti, F. S. di Luzio, M. Bravi, F. Bressi, F. Draicchio, S. Sterzi, L. Zollo, Patient-tailored adaptive control for robot-aided orthopaedic rehabilitation, in: 2022 international conference on robotics and automation (ICRA), IEEE, 2022, pp. 5434–5440.

[2] A. Mohebbi, Human-robot interaction in rehabilitation and assistance: a review, Current Robotics Reports 1 (2020) 131–144.

[3] C. Tamantini, F. Cordella, F. Scotto di Luzio, C. Lauretti, B. Campagnola, F. Santacaterina, M. Bravi, F. Bressi, F. Draicchio, S. Miccinilli, et al., A fuzzy-logic approach for longitudinal assessment of patients' psychophysiological state: an application to upper-limb orthopedic robot-aided rehabilitation, Journal of NeuroEngineering and Rehabilitation 21 (2024) 202.

[4] C. Tamantini, F. Cordella, N. L. Tagliamonte, I. Pecoraro, I. Pisotta, A. Bigioni, F. Tamburella, M. Lorusso, M. Molinari, L. Zollo, A data-driven fuzzy logic method for psychophysiological assessment: An application to exoskeleton-assisted walking, IEEE Transactions on Medical Robotics and Bionics (2024).

[5] P. V. Rouast, M. T. Adam, R. Chiong, D. Cornforth, E. Lux, Remote heart rate measurement using low-cost rgb face video: a technical literature review, Frontiers of Computer Science 12 (2018) 858–872.

[6] R. Cittadini, L. R. Buonocore, E. Matheson, M. Di Castro, L. Zollo, Robot-aided contactless monitoring of workers' cardiac activity in hazardous environment, IEEE Access 10 (2022) 133427–133438.

[7] H. W. Huang, J. Chen, P. R. Chai, C. Ehmke, P. Rupp, F. Z. Dadabhoy, et al., Mobile robotic platform for contactless vital sign monitoring, Cyborg and Bionic Systems (2022).
[8] S. Takır, H. Kose, B. Coskun, D. E. Barkana, rppg detection in children with autism spectrum disorder during robot-child interaction studies, In 2022 International Conference on Digital Image Computing: Techniques and Applications (DICTA) (2022).
[9] H. W. Huang, C. Ehmke, G. Merewether, F. Dadabhoy, A. Feng, A. J. Thomas, et al., Agile mobile robotic platform for contactless vital signs monitoring, Authorea Preprints (2023).
[10] L. Cristofori, C. D. Hromei, F. S. di Luzio, C. Tamantini, F. Cordella, D. Croce, L. Zollo, R. Basili, et al., Heal9000: an intelligent rehabilitation robot., in: SMARTERCARE@ AI* IA, 2021, pp. 29–41.
[11] C. Lugaresi, J. Tang, H. Nash, C. McClanahan, E. Uboweja, M. Hays, F. Zhang, C.-L. Chang, M. G. Yong, J. Lee, et al., Mediapipe: A framework for building perception pipelines, arXiv preprint arXiv:1906.08172 (2019).
[12] A. H. Gandomi, X.-S. Yang, A. H. Alavi, Cuckoo search algorithm: a metaheuristic approach to solve structural optimization problems, Engineering with computers 29 (2013) 17–35.
[13] G. De Haan, V. Jeanne, Robust pulse rate from chrominance-based rppg, IEEE transactions on biomedical engineering 60 (2013) 2878–2886.
[14] A. Caroppo, A. Manni, G. Rescio, P. Siciliano, A. Leone, Vital signs estimation in elderly using camera-based photoplethysmography, Multimedia Tools and Applications (2024) 1–24.