Addressing Healthcare Needs of the Elderly in Contactless Blood Oxygen Saturation Estimation with rPPG

Yuqing Cao[1], Yuxing Li[1], Kai Hang Yiu[2, 3], and Edmund Y. Lam[1]

[1]Department of Electrical and Electronic Engineering, The University of Hong Kong, Pokfulam, Hong Kong SAR
[2]Department of Medicine, The University of Hong Kong, Pokfulam, Hong Kong SAR
[3]HKU-Shenzhen Hospital, 1 Haiyuan 1 Rd, Shenzhen, Guangdong Province, China, 518009

ABSTRACT

Remote photoplethysmography (rPPG) is an effective technique for non-contact estimation of vital signs using video-based detection of exposed skin. It enables contactless healthcare services for clinical measurements and medical diagnosis. While existing rPPG methods primarily focus on enhancing robustness against interference factors like motion artifacts and illumination changes, limited attention has been given to the influence of the age factor on rPPG model performance. This study explicitly analyzed the impact of the age factor on the application of an rPPG method to the blood oxygen saturation (SpO_2) estimation. Two key observations were made. First, it was more challenging to estimate SpO_2 with rPPG for the elderly than for young individuals. Second, the performance of SpO_2 estimation for the elderly could be improved by dividing the training data into different age groups and exclusively training the model on data collected from the elderly. These observations highlighted that the age factor had a significant impact on rPPG methods, emphasizing the need for explicit consideration of the age factor in rPPG method design.

Keywords: Remote Photoplethysmography (rPPG), Blood Oxygen Saturation Estimation, Elderly Healthcare

1. INTRODUCTION

The global community is currently grappling with the formidable challenge of population aging. According to the Government of Hong Kong, up to 20% of the population in this city is aged over 65 years old, and the number is expected to increase to 25% in 2028 and 35% in 2069.[1] Therefore, regular health surveillance for the elderly becomes increasingly more crucial for early detection of diseases and preemptive treatments to prevent severe complications.[2,3]

Accordingly, remote photoplethysmography (rPPG) offers a contactless and non-invasive approach to assess vital physiological parameters like blood oxygen saturation (SpO_2), by capturing blood flow changes in the skin resulting from cardiac activities.[4,5] Although the feasibility of rPPG in remote health monitoring has been demonstrated[6], little attention has been given to the impact of the age factor in datasets. Most state-of-the-art research primarily focuses on enhancing model architectures to mitigate various interference factors, such as illumination changes[7], skin tone variations[8], and body movements[9]. Nevertheless, a high-performing rPPG model trained and tested on public datasets might not generalize well to the elderly population. This limitation arises from the fact that public datasets[10] commonly overlook the age factor during data collection. However, the elderly constitute a distinct age group with facial features that differ from those of young individuals, leading to unique requirements for rPPG workflow designs.

This paper explicitly accounted for the age factor in the context of SpO_2 estimation with rPPG. Firstly, it demonstrated the heightened difficulty in estimating SpO_2 for the elderly compared to young individuals. Additionally, a practical and effective solution was proposed by refining the dataset, without modifying the model architecture. To be specific, the dataset was tailored to exclude data from young individuals and instead solely focus on elderly individuals. This work emphasized the significance of considering the age factor and provided valuable insights for future rPPG workflow designs to address the unique needs of geriatric healthcare.

Computational Optical Imaging and Artificial Intelligence in Biomedical Sciences, edited by Liang Gao, Guoan Zheng, Seung Ah Lee, Proc. of SPIE Vol. 12857, 128570C

· 1605-7422 · doi: 10.1117/12.3000619

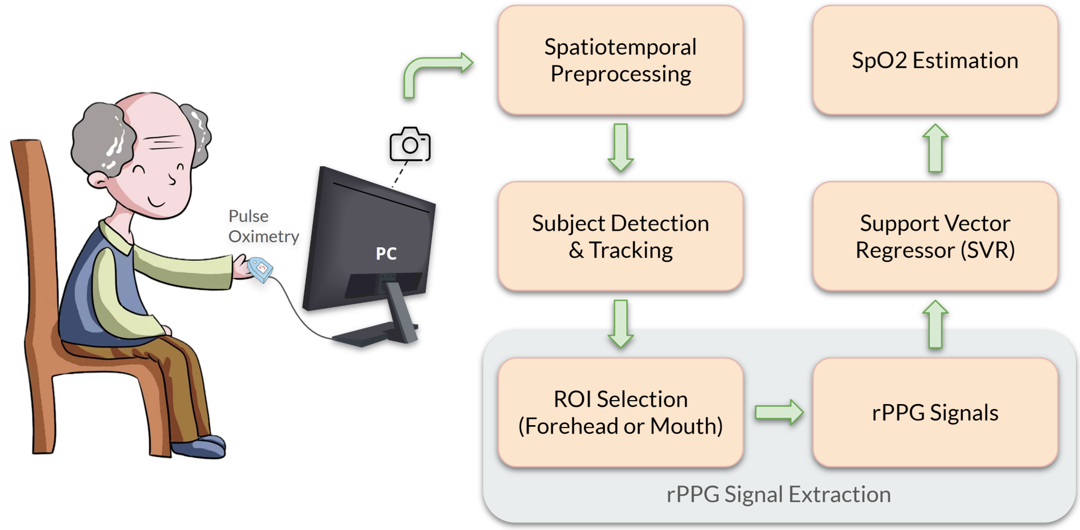


Figure 1. Schematic of data collection setup and rPPG estimation workflow.

2. METHODS

2.1 Data Collection Setup

A total of 84 patients in Hong Kong University Shenzhen (HKU-SZ) hospital volunteered for this study. For each participant, one facial video along with its corresponding SpO_2 value was recorded, resulting in 73 pairs of effective raw data. The remaining 11 videos were deemed ineffective due to facial occlusions or subjects moving out of frame. The effective dataset consisted of 44 male and 29 female participants, with ages ranging from 28 to 80 years, a mean age of 56.96 years, and a standard deviation of 14.39 years.

All data were recorded in a controlled environment with a consistent lighting condition and room temperature. As depicted in Fig. 1, facial videos of the volunteers were captured by a webcam embedded in a personal computer (PC). Camera settings including exposure, white balance, and focus were kept consistent. Videos were recorded at a resolution of 1080×1920 and a rate of 30 frames per second, saved in MPEG-4 format with eight-bit RGB quality. Additionally, pulse oximetry was attached to the finger of the participant to simultaneously record the SpO_2 value, serving as the ground truth (GT) for the rPPG based SpO_2 estimation.

2.2 rPPG Workflow

A step-by-step rPPG workflow was designed for the task of SpO_2 estimation.

i. The captured videos were first preprocessed in both spatial and temporal domains. In the spatial domain, brightness adjustment and denoising techniques were applied to each video frame to enhance the signal-to-noise ratio (SNR). In the temporal domain, the first and last three seconds of the videos were removed to mitigate potential motion artifacts.

ii. A facial detection and tracking algorithm was utilized to locate the subject. The tracked faces contained regions of interest (ROI) from which rPPG signals could be extracted.

iii. Two areas, forehead and mouth, were selected as ROIs from each frame. The extracted rPPG signals from these two ROIs were processed independently.

iv. Subsequently, an rPPG signal extraction procedure was performed when sufficient visible feature points were detected, as indicated by the gray rectangle box in Fig. 1. The signal extraction involved averaging pixel values within the selected ROI for each RGB channel at each frame t, resulting in $RED(t)$, $GREEN(t)$, and $BLUE(t)$. According to the dichromatic reflection model (DRM)[11], the difference between the red and green channels could be utilized for SpO_2 prediction. As a result, the value of SpO_2 could be derived from the light reflection theory and the Beer-Lambert law[12], expressed as

$$SpO_2 = \alpha + \beta \frac{RED_{std}/RED_{mean}}{GREEN_{std}/GREEN_{mean}} = \alpha + \beta \cdot R\,, \tag{1}$$

where α and β were parameters to be determined by the model. The subscripts *std* and *mean* corresponded to the standard deviation and mean value, respectively.

v. Equation (1) indicated that SpO_2 estimation is a regression task for R given $RED(t)$ and $GREEN(t)$. Therefore, a support vector regressor (SVR)[13] model was employed to predict the final SpO_2 values.

3. EXPERIMENTS AND RESULTS

3.1 Experiment Design

To investigate the impact of the age factor on the model performance, the collected dataset was divided into two groups: 'young' (< 60 years old) and 'elderly' (≥ 60 years old). This division was based on the age definition provided by the United Nations.[14] Consequently, there were 39 subjects in the 'young' group and 34 subjects in the 'elderly' group. The entire group of 73 subjects was labeled as 'mixed' for a notation purpose.

Three experiments were designed to explore the influence of the age factor on the model performance. All experiments followed the same workflow outlined in Sec. 2.2 with consistent SVR model parameters and train-test-split policy. For each experiment, 25 training videos and 5 testing videos were randomly selected from the respective target group using a random seed generated by a random integer between 0 to 999. The first two experiments utilized the 'mixed' group as the training dataset. In the first experiment, the testing dataset exclusively involved the 'young' group, while the second experiment focused solely on the 'elderly' group. In the third experiment, both the training and testing datasets were from the 'elderly' group. The testing dataset remained the same as that in the second experiment.

In brief, the first two experiments assessed the performance of the same model on the 'young' and 'elderly' groups, respectively. The last two experiments compared two models trained by the 'mixed' and 'elderly' groups, respectively. Models performance were evaluated on the same subgroup within the 'elderly' group.

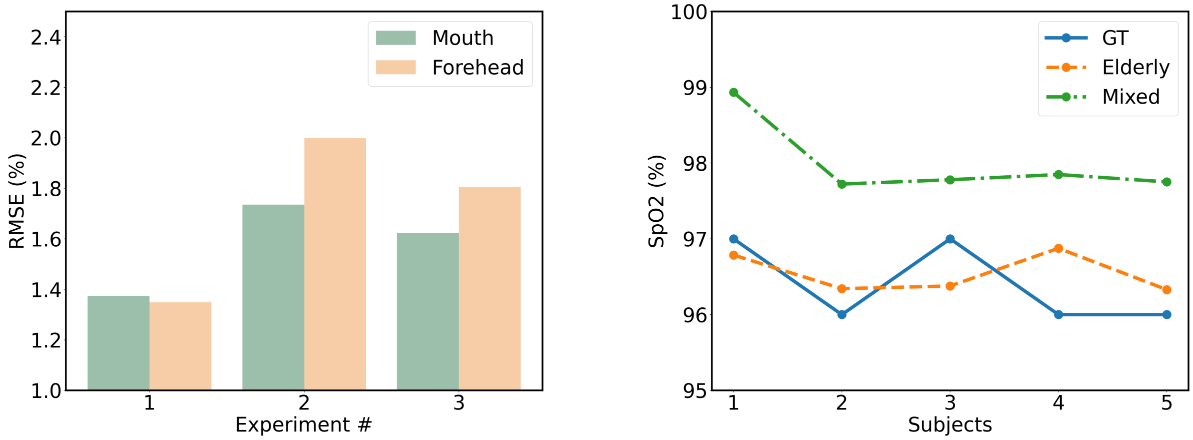
3.2 Experimental Results and Visualization

Table 1 presented the arithmetic mean of metrics of 500 repeated trials with varying random integers generated per trial. The 'Train' and 'Test' columns indicated different dataset configurations used in the three experiments to study the impact of the age factor. Evaluation metrics for the model performance include root mean squared error (RMSE), mean squared error (MSE), and mean absolute error (MAE). Lower values of these metrics indicated better performance of the rPPG model. Two ROI selection methods, namely 'Forehead' and 'Mouth', were utilized for each experiment. To be more intuitive, the column of 'RMSE' in the table 1 was visualized as a bar chart in Fig. 2(a). To further compare the performance of the last two experiments, Fig. 2(b) depicted a plot of SpO_2 estimations on the same subjects from the 'elderly' group. These estimations were made by models trained from the last two experiments using the 'Forehead' method.

The results obtained from the experiments support two claims regarding the impact of the age factor on SpO_2 estimation with rPPG. Firstly, it was evident that the RMSE values for the first experiment were significantly

Table 1. Metrics of experiments on the age factor

#	Train	Test	RMSE		MSE		MAE	
			Forehead	Mouth	Forehead	Mouth	Forehead	Mouth
1	Mixed	Young	1.35	1.37	2.02	2.15	1.13	1.20
2	Mixed	Elderly	2.00	1.74	4.83	3.95	1.60	1.34
3	Elderly	Elderly	1.81	1.62	4.19	3.72	1.29	1.27



(a) **(b)**

Figure 2. **(a)** Bar chart of the root mean squared errors (RMSE) of SpO_2 predictions (%) using forehead and mouth as ROI respectively. Three sets of experiments involved various training and testing datasets from 'young', 'elderly', and 'mixed' groups. **(b)** Plot of SpO_2 values (%) predicted from models trained on different age groups. The testing dataset was the same subjects from the 'elderly' group. The dashed and dash-dotted lines represented the ones trained on the 'elderly' and 'mixed' groups respectively. The solid line represented the ground truth (GT) recorded from pulse oximetry.

lower than those for the second experiment, regardless of the ROI selection method. This finding substantiated the claim that estimating SpO_2 was more difficult for the elderly than for the young. The rPPG model, which performed well on young individuals, may struggle to generalize effectively to the elderly population. This discrepancy could potentially be attributed to the presence of wrinkles and age-related skin changes exhibited on the elderly, which may introduce interference with rPPG signals[15]. Secondly, it was noteworthy that the third experiment trained exclusively on the 'elderly' group outperformed the second one trained on the 'mixed' group. This observation supported the claim that SpO_2 estimation with rPPG could be enhanced by training models on a dataset exclusively containing data collected from the elderly. Similar trends could be observed when analyzing the MSE and MAE results, further reinforcing the proposed claims. Furthermore, Fig. 2(b) demonstrated that the predicted SpO_2 values from the model trained on the 'elderly' group were very close to GT, with errors less than 1%. On the other hand, the model trained on the 'mixed' group failed to accurately predict SpO_2 for subjects from the 'elderly' group, indicating a clear bias. Therefore, refining the training dataset to focus on the elderly population leads to a significant improvement in the accuracy of SpO_2 estimation, without requiring any changes to the rPPG model itself.

4. DISCUSSIONS AND CONCLUSIONS

This paper studied the influence of the age factor on SpO_2 estimation with rPPG. A data collection setup was established to capture facial videos and assess SpO_2 values through pulse oximetry. A step-by-step rPPG workflow was implemented in several comparative experiments with different configurations of training and testing datasets. The experimental results led to two main conclusions. Firstly, estimating SpO_2 with rPPG was more challenging for the elderly compared to the young. Secondly, a solution was proposed wherein training exclusively on an elderly dataset can enhance the estimation performance. This highlighted the necessity of collecting datasets specifically tailored for the elderly, considering the limited attention given to this demographic in state-of-the-art public datasets.

Therefore, this study provided valuable insights for addressing the unique needs of geriatric healthcare by explicitly accounting for the age factor. Further research could be conducted to investigate the challenges in designing rPPG models for the elderly and develop enhanced algorithms accordingly. For instance, exploring the impact of facial textures like wrinkles and speckles on the elderly could lead to the development of customized rPPG approaches for the elderly. By advancing these state-of-the-art techniques, we can effectively enhance the well-being of the elderly society, ensuring that the benefits extend from research and development to practical implementation in the field of elderly healthcare.

REFERENCES

[1] "Hong Kong population projections for 2022-2046 released." https://www.censtatd.gov.hk/en/press_release_detail.html?id=5368 (2023).

[2] Kong, L., Zhao, Y., Dong, L., Jian, Y., Jin, X., Li, B., Feng, Y., Liu, M., Liu, X., and Wu, H., "Non-contact detection of oxygen saturation based on visible light imaging device using ambient light," *Optics Express* **21**(15), 17464–17471 (2013).

[3] Huang, B., Hu, S., Liu, Z., Lin, C.-L., Su, J., Zhao, C., Wang, L., and Wang, W., "Challenges and prospects of visual contactless physiological monitoring in clinical study," *NPJ Digital Medicine* **6**(1), 231 (2023).

[4] Casalino, G., Castellano, G., and Zaza, G., "A mHealth solution for contact-less self-monitoring of blood oxygen saturation," in [*2020 IEEE Symposium on Computers and Communications (ISCC)*], 1–7, IEEE (2020).

[5] Mathew, J., Tian, X., Wong, C.-W., Ho, S., Milton, D. K., and Wu, M., "Remote blood oxygen estimation from videos using neural networks," *IEEE Journal of Biomedical and Health Informatics* **27** (2023).

[6] Verkruysse, W., Svaasand, L. O., and Nelson, J. S., "Remote plethysmographic imaging using ambient light," *Opt. Express* **16**, 21434–21445 (Dec 2008).

[7] Lin, Y.-C. and Lin, Y.-H., "A study of color illumination effect on the snr of rppg signals," in [*2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*], 4301–4304, IEEE (2017).

[8] Nowara, E. M., McDuff, D., and Veeraraghavan, A., "A meta-analysis of the impact of skin tone and gender on non-contact photoplethysmography measurements," in [*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*], 284–285 (2020).

[9] Maity, A. K., Wang, J., Sabharwal, A., and Nayar, S. K., "RobustPPG: camera-based robust heart rate estimation using motion cancellation," *Biomedical Optics Express* **13**(10), 5447–5467 (2022).

[10] Cheng, C.-H., Wong, K.-L., Chin, J.-W., Chan, T.-T., and So, R. H., "Deep learning methods for remote heart rate measurement: A review and future research agenda," *Sensors* **21**(18), 6296 (2021).

[11] Shafer, S. A., "Using color to separate reflection components," *Color Research & Application* **10**(4), 210–218 (1985).

[12] Lambert, J. H., [*Photometria sive de mensura et gradibus luminis, colorum et umbrae*], sumptibus vidvae E. Klett, typis CP Detleffsen (1760).

[13] Drucker, H., Burges, C. J., Kaufman, L., Smola, A., and Vapnik, V., "Support vector regression machines," *Advances in Neural Information Processing Systems* **9** (1996).

[14] "Older persons." https://emergency.unhcr.org/protection/persons-risk/older-persons (2020).

[15] Batool, N. and Chellappa, R., "Detection and inpainting of facial wrinkles using texture orientation fields and markov random field modeling," *IEEE Transactions on Image Processing* **23**(9), 3773–3788 (2014).