

# Toward Motion Robustness: A masked attention regularization framework in remote photoplethysmography

## Supplementary Material

### 1. Metric details

We follow the approach outlined in [3] by using mean absolute error (MAE), root mean square error (RMSE), mean absolute percentage error (MAPE), and Pearson’s correlation coefficient ( $r$ ) as evaluation metrics for heart rate (HR). The details of related equations are as follows.

$$MAE = \frac{1}{N} \sum_{n=1}^N |Z - Y| \quad (1)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{n=1}^N (Z - Y)^2} \quad (2)$$

$$MAPE = \frac{1}{N} \sum_{n=1}^N \left| \frac{Z - Y}{Z} \right| \quad (3)$$

$$r = \frac{\sum_{n=1}^N (Z_n - \bar{Z})(Y_n - \bar{Y})}{\sqrt{(\sum_{n=1}^N Z_n - \bar{Z})^2 (\sum_{n=1}^N Y_n - \bar{Y})^2}} \quad (4)$$

### 2. Ablation study

#### 2.1. HRV and RF evaluation

The face is a complex structure that contains numerous blood vessels, which play a crucial role in maintaining the health and function of the skin and underlying tissues. Fig. 1 is an overview of the distribution of blood vessels in the face. The facial veins are located just beneath the skin and drain blood from the surface of the face, running throughout forehead, cheek, chin, eyebrow, and nose. The alteration in light absorption resulting from blood flow is crucially significant for rPPG measurements. Our novel approach, MAR-rPPG, is designed to capture these variations and provide precise heart rate assessments.

Following the method [11], we evaluate the heart rate variability (HRV) and respiratory frequency (RF) performance of our proposed MAR-rPPG on the UBFC-rPPG dataset. For HRV, we assess three attributes: the low frequency (LF), high frequency (HF), and LF/HF ratio, and LF and HF are determined using the interbeat intervals in the low-frequency (0.04 to 0.15 Hz) and high-frequency (0.15 to 0.4 Hz) ranges of rPPG signals. To comprehensively report the performance, we utilize three metrics: standard deviation (Std), RMSE, and Pearson’s correlation coefficient  $r$ . We conduct a five-fold cross-validation experiment and compare our approach with some famous methods, such as

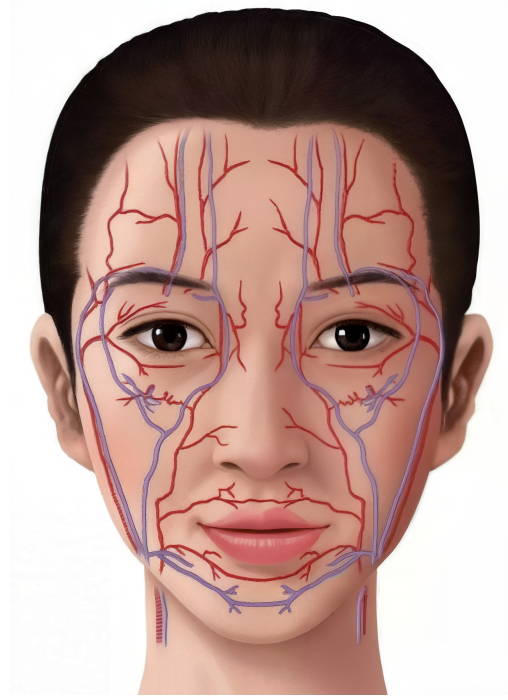


Figure 1. The overview of superficial facial veins.

POS [8], CHROM [1], GREEN [7], CVD [5], rPPGNet [9], Dual-Gan [4], Physformer [10], Gideon *et al.* [2], REA-LFA [11]. We utilize the toolkit HeartPy [6] to calculate HRV and RF.

As shown in Tab. 1, we can see that MAR-rPPG surpasses all traditional ones and many deep learning ones. Besides, our approach offers improved accuracy in estimating RF and LF/HF ratio when compared to other advanced methods. For LF and HF, PhysFormer and Dual-GAN show better performance. This suggests that the proposed technique has the potential to excel not only in heart rate estimation tasks but also in predicting rPPG signals for RF measurements and heart rate variability analysis.

#### 2.2. Loss hyperparameter selection

As illustrated in Tab. 2, we can find that the hyperparameter  $\alpha$  and  $\beta$  are retain consistant and accurate rPPG estimation on the PURE dataset, while only one parameter setting with  $\alpha = 0.1$  and  $\beta = 0.9$  cannot converge and no metric outputs. We select  $\alpha = 0.3$  and  $\beta = 0.5$  for all experiments in this paper.

Method	RF			HRV: LF			HRV: HF			HRV: LF/HF		
	Std↓	RMSE↓	$r$ ↑	Std↓	RMSE↓	$r$ ↑	Std↓	RMSE↓	$r$ ↑	Std↓	RMSE↓	$r$ ↑
POS [8]	0.109	0.107	0.087	0.171	0.169	0.479	0.171	0.169	0.479	0.405	0.399	0.518
CHROM [1]	0.086	0.089	0.102	0.243	0.240	0.159	0.243	0.240	0.159	0.655	0.645	0.266
GREEN [7]	0.087	0.086	0.111	0.186	0.186	0.280	0.186	0.186	0.280	0.361	0.365	0.492
CVD [5]	0.017	0.018	0.252	0.053	0.065	0.740	0.053	0.065	0.740	0.169	0.168	0.812
rPPGNet [9]	0.030	0.034	0.233	0.071	0.070	0.686	0.071	0.070	0.686	0.212	0.208	0.744
Dual-Gan [4]	0.010	0.010	0.395	0.034	0.035	0.891	0.034	0.035	0.891	0.131	0.136	0.881
Physformer [10]	0.009	<b>0.009</b>	0.413	<b>0.030</b>	<b>0.032</b>	<b>0.895</b>	<b>0.030</b>	<b>0.032</b>	<b>0.895</b>	0.126	0.130	0.893
Gideon <i>et al.</i> [2]	0.061	0.098	0.103	0.091	0.139	0.694	0.091	0.139	0.694	0.525	0.691	0.684
REA-LFA [11]	0.023	0.028	0.351	0.047	0.062	0.769	0.047	0.062	0.769	0.160	0.164	0.831
MAR-rPPG(Ours)	<b>0.008</b>	0.031	<b>0.838</b>	0.065	0.283	0.856	0.065	0.283	0.856	<b>0.029</b>	<b>0.126</b>	<b>0.925</b>

Table 1. Comparison of RF and HRV estimations on the UBFC-rPPG dataset. The best results are in bold. Std: standard deviation, RMSE: Root Mean Square Error,  $r$ : Pearson correlation coefficient.

$\alpha$	$\beta$	MAE ↓	RMSE ↓	MAPE ↓	$r$ ↑
0.1	0.1	0.250	0.645	0.273	1.000
0.1	0.3	0.083	0.288	0.106	1.000
0.1	0.5	0.083	0.288	0.106	1.000
0.1	0.7	0.083	0.288	0.106	1.000
0.1	0.9	-	-	-	-
0.3	0.1	0.166	0.408	0.267	1.000
0.3	0.3	0.250	0.865	0.324	0.999
<b>0.3</b>	<b>0.5</b>	<b>0.083</b>	<b>0.288</b>	<b>0.106</b>	<b>1.00</b>
0.3	0.7	0.333	0.706	0.440	0.999
0.3	0.9	0.083	0.288	0.106	1.000
0.5	0.1	0.083	0.288	0.106	1.000
0.5	0.3	0.083	0.288	0.106	1.000
0.5	0.5	0.083	0.288	0.106	1.000
0.5	0.7	0.083	0.288	0.106	1.000
0.5	0.9	0.083	0.288	0.106	1.000
0.7	0.1	0.166	0.577	0.166	1.000
0.7	0.3	0.083	0.288	0.106	1.000
0.7	0.5	0.083	0.288	0.159	1.000
0.7	0.7	0.083	0.288	0.106	1.000
0.7	0.9	0.083	0.288	0.106	1.000
0.9	0.1	0.166	0.577	0.166	1.000
0.9	0.3	0.083	0.288	0.106	1.000
0.9	0.5	0.166	0.577	0.166	1.000
0.9	0.7	0.083	0.288	0.106	1.000
0.9	0.9	0.083	0.288	0.106	1.000

Table 2. Ablation study for loss hyperparameter  $\alpha$  and  $\beta$ . Our implementation is marked in **shadow**, ‘-’ indicates that the relevant parameters cannot converge.

## References

- [1] Gerard de Haan and Vincent Jeanne. Robust pulse rate from chrominance-based rppg. *IEEE Transactions on Biomedical Engineering*, 60(10):2878–2886, 2013.
- [2] John Gideon and Simon Stent. The way to my heart is through contrastive learning: Remote photoplethysmography from unlabelled video. In *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 3975–3984,

2021.

- [3] Xin Liu, Xiaoyu Zhang, Girish Narayanswamy, Yuzhe Zhang, Yuntao Wang, Shwetak Patel, and Daniel McDuff. Deep physiological sensing toolbox. *arXiv preprint arXiv:2210.00716*, 2022.
- [4] Hao Lu, Hu Han, and S. Kevin Zhou. Dual-gan: Joint bvp and noise modeling for remote physiological measurement. In *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 12399–12408, 2021.
- [5] Xuesong Niu, Zitong Yu, Hu Han, Xiaobai Li, Shiguang Shan, and Guoying Zhao. Video-based remote physiological measurement via cross-verified feature disentangling. In *Computer Vision – ECCV 2020*, pages 295–310, 2020.
- [6] Paul van Gent, Haneen Farah, Nicole van Nes, and Bart van Arem. Analysing noisy driver physiology real-time using off-the-shelf sensors: Heart rate analysis software from the taking the fast lane project. *Journal of Open Research Software*, 2019.
- [7] Wim Verkruysse, Lars O Svaasand, and J Stuart Nelson. Remote plethysmographic imaging using ambient light. *Opt Express*, 16:21434–45, 2008.
- [8] Wenjin Wang, Albertus C. den Brinker, Sander Stuijk, and Gerard de Haan. Algorithmic principles of remote ppg. *IEEE Transactions on Biomedical Engineering*, 64(7):1479–1491, 2017.
- [9] Zitong Yu, Wei Peng, Xiaobai Li, Xiaopeng Hong, and Guoying Zhao. Remote heart rate measurement from highly compressed facial videos: An end-to-end deep learning solution with video enhancement. In *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 151–160, 2019.
- [10] Zitong Yu, Yuming Shen, Jingang Shi, Hengshuang Zhao, Philip Torr, and Guoying Zhao. Physformer: Facial video-based physiological measurement with temporal difference transformer. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4176–4186, 2022.
- [11] Zijie Yue, Miaojing Shi, and Shuai Ding. Facial video-based remote physiological measurement via self-supervised learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(11):13844–13859, 2023.