Supplementary Material : TRUST: Time-domain Residual Unsupervised Stability Technique for Improved Heart Rate Estimation:

Shahzad Ahmad[1] Sania Bano[2] Sukalpa Chanda[1]
shahzaa@hiof.no sania.22eez0012@iitrpr.ac.in sukalpa@ieee.org
Santosh Kumar Vipparthi[2] Subrahmanyam Murala[3]
skvipparthi@iitrpr.ac.in muralas@tcd.ie
[1] Østfold University College, Norway [2] Indian Institute of Technology Ropar, India [3] Trinity College Dublin, Ireland

1. Supplementary Material

We present supplementary material that enhances the understanding of our main paper through additional details and in-depth qualitative analysis. This supplementary content is structured as follows:

- Error metric calculation
- Waveform
- Feature heat maps

1.1. Error metric calculation

We follow the metrics to calculate the measurement accuracy.
Here $\hat{y}$ = heart rate, y = ground truth, error $e = \hat{y} - y$

1.1.1 Pearson Correlation Coefficient (r)

The value of the Pearson correlation coefficient is between -1 to +1.

$$r = \frac{\frac{1}{N}\sum_{i=1}^{N}\hat{y_i}y_i - \frac{1}{N}\sum_{i=1}^{N}\hat{y_i}\frac{1}{N}\sum_{i=1}^{N}y_i}{\sqrt{\frac{1}{N}\sum_{i=1}^{N}\hat{y_i}^2 - (\frac{1}{N}\sum_{i=1}^{N}\hat{y_i})^2}\sqrt{\frac{1}{N}\sum_{i=1}^{N}y_i^2 - (\frac{1}{N}\sum_{i=1}^{N}y_i)^2}} \tag{1}$$

1.1.2 Mean Absolute Error (MAE)

$$MAE = \sum_{i=1}^{N}\frac{|e_i|}{N} \tag{2}$$

1.1.3 Root Mean Squared Error (RMSE)

$$RMSE = \sqrt{\sum_{i=1}^{N}\frac{e_i^2}{N}} \tag{3}$$

1.1.4 Standard deviation (STD)

$$STD = \sqrt{\frac{1}{N}\sum_{i=1}^{N}e_i^2 - \left(\frac{1}{N}\sum_{i=1}^{N}e_i\right)^2} \tag{4}$$

1.2. Feature heat maps

In this visual representation, as shown in Fig.1, feature heat maps captured at different spatial resolutions unveil the impact of video movements. The TRUST framework demonstrates its effectiveness by first stabilizing features at low-level spatial resolutions, ensuring stability even in dynamic scenarios such as head movements, body movements, and lighting conditions. The observed feature stability at various levels showcases TRUST's ability to maintain coherence amidst video motion. This unique approach of stabilizing and then extracting features significantly enhances the resulting rPPG signal quality. The visual evidence underscores TRUST's robustness, highlighting its potential for accurate and stable camera-based vital signs estimation in real-world applications.

1.3. Waveform

Additional waveforms as shown in Fig. 2. from all four datasets are provided for a better understanding of predicted signal characteristics.

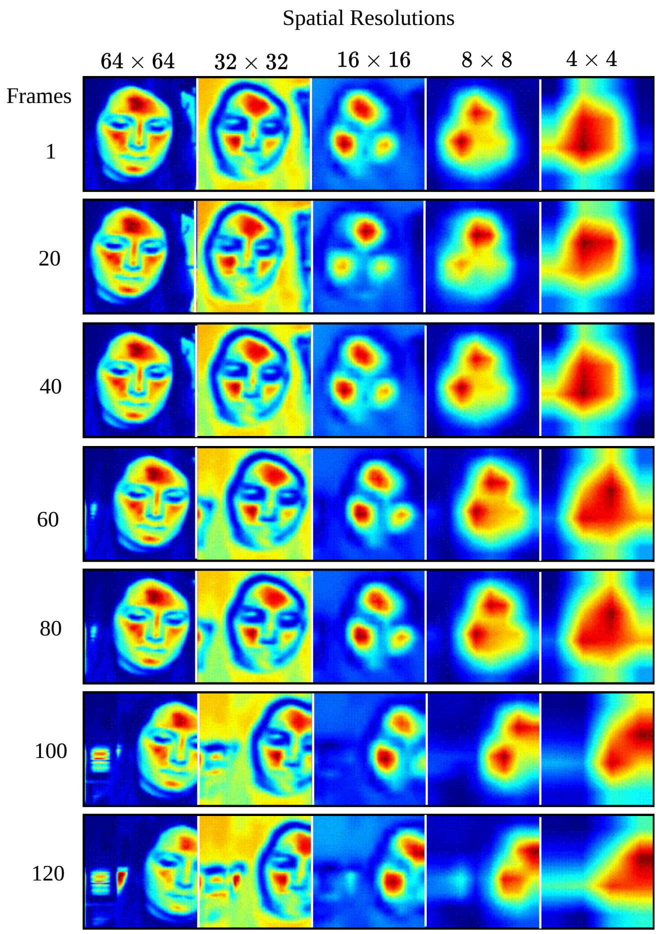

Figure 1. Feature heat maps captured at varying spatial resolution levels with a frame interval of 20 reveal stable features despite video movement. The stability is particularly noticeable in low-level spatial resolutions (16×16, 8×8, and 4×4), emphasizing the robust and consistent features of our method in addressing dynamic conditions.

Figure 2. The comparison between the predicted waveform (blue curve) and their corresponding ground truth waveform (red curve) for all four datasets.