

Estimating Heart Rate Variability from wrist-worn photoplethysmography devices in daily activities: a preliminary Convolutional Denoising Autoencoder approach

1st Gianluca Rho

*Dip. di Ingegneria dell'Informazione
Research Center "E.Piaggio"
University of Pisa
Pisa, Italy
gianluca.rho@phd.unipi.it*

2nd Nicola Carbonaro

*Dip. di Ingegneria dell'Informazione
University of Pisa
Pisa, Italy
nicola.carbonaro@unipi.it*

3rd Marco Laurino

*Institute of Clinical Physiology
National Research Council
Pisa, Italy
marco.laurino@cnr.it*

4th Alessandro Tognetti

*Dip. di Ingegneria dell'Informazione
University of Pisa
Pisa, Italy
alessandro.tognetti@unipi.it*

5th Alberto Greco

*Dip. di Ingegneria dell'Informazione
University of Pisa
Pisa, Italy
alberto.greco@unipi.it*

Abstract—Wrist photoplethysmography (PPG) devices are gaining popularity as a non-invasive means to monitor pulse rate variability (PRV) in daily-life settings. Yet, movement artifacts make reliable estimation of PRV challenging during physical activities even if not very intense. Various approaches based on spectral analysis and deep learning (DL) have provided mean HR over time with low estimation errors. However, mean HR dynamics cannot be adopted to derive detailed information about autonomic activity, for which PRV time series is necessary. In this preliminary work, we propose a novel approach combining a convolutional denoising autoencoder (CNN-DAE) with a physiologically-constrained custom loss function, which leverages synchronous electrocardiographic (ECG) recordings and inter-beat interval (IBI) information to reconstruct the PPG signal, free from artifacts, and obtain reliable PRV. The reconstructed PRV has been averaged across time windows to estimate the mean HR and compare it against those obtained from standard bandpass filtering procedures of PPG and ECG's HRV, which was used as the gold standard reference. Our preliminary results suggest that our method can accurately estimate PRV, providing mean HR unbiased estimates with significantly lower error rates than conventional approaches. This suggests that the proposed methodology could be adopted to denoise PPG time series in uncontrolled environments.

Keywords—photoplethysmography, heart rate variability, deep learning, autoencoders, wearable devices

Funded by the European Union. Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them. The work is supported by European Union's Horizon Europe Research and Innovation Programme under grant agreement No. 101057103 – project TOLIFE. This research has also received partial funding from the Italian Ministry of Education and Research (MIUR) in the framework of the ForeLab project (Departments of Excellence).

I. INTRODUCTION

Monitoring heart rate variability (HRV) allows for the assessment of autonomic imbalances associated with both physiological and pathological conditions, providing information to users about their lifestyle and assisting clinicians in personalizing treatments. Although wearable electrocardiographic (ECG) holters are considered the gold standard for estimating HRV in unconstrained settings, they may feel uncomfortable to wear over long periods. In this context, several studies indicated wrist photoplethysmography (PPG) sensors as an easy-to-use and non-invasive means to estimate mean heart rate (HR) and, to some extent, the pulse rate variability (PRV), i.e., a surrogate of HRV, during daily-life activities [1].

Wrist devices can provide accurate HR estimates during resting activities [2], [3]. However, motion artifacts (MA) can severely deteriorate the signal quality during physical activities even if not very intense, affecting the reliability of HR estimates [2], [3]. In this context, several signal processing techniques have been proposed to denoise and reconstruct PPG time series, including adaptive and Kalman filtering [4], [5], wavelet analysis [6], time-frequency analysis [7], and signal decomposition [8], [9]. However, these approaches involve the tuning of thresholds and parameters which may vary according with the context of the recordings, potentially making their application to free-living conditions not practical.

Recently, solutions based on deep learning (DL) have been proposed for denoising PPG time series recorded in daily-life settings and extracting useful parameters from them [10]–[14]. Among them, solutions based on convolutional neural

network (CNNs) have showed good denoising performances, recovering an estimate of the mean HR over time with high accuracy [12]–[14]. Yet, recovering PRV time series is a more complex task, as it requires the accurate denoising and reconstruction of individual PPG pulses to derive the distance between consecutive peaks reliably. Furthermore, mean HR dynamics cannot be used as a surrogate of PRV when more specific parameters about autonomic balance need to be estimated. In this light, a methodology capable of providing a reliable estimate of PRV could be of great interest.

In this work, we propose a preliminary approach based on a convolutional denoising autoencoder (CNN-DAE) for denoising PPG time series in the context of high-intensity activity recordings and improve the reliability of HR estimates from PRV [9]. In particular, we implement a novel training strategy introducing a physiologically constrained penalty to the model’s loss function.

DAEs learn to reconstruct PPG signal by minimizing the distance between the reconstruction of input noisy time series and their clean reference counterpart. However, recording two identical PPGs, where one of them is not susceptible to any artifact, is a complex task. To derive a clean reference signal to train the network with, we applied a PPG simulation model which uses RR intervals of the synchronously acquired ECG as an input [15], [16]. We built the training loss as the mean squared error (MSE) between the model output and the reference time series, and we combined it with a custom penalty term based on the difference between the inter-beat interval (IBI) time series obtained from the model output and reference PPGs, respectively. A custom penalty factor is introduced in the calculation to also account for spurious and missing peaks observed in the model output. This allows the network to learn realistic artifact dynamics while reconstructing PPG time series with higher physiological plausibility.

We derived the mean HR as the average of the PRV time series within non-overlapping windows of 30 seconds. Finally, the results of this processing chain were compared with the mean HR estimated from the original PPG series applying a standard spectral analysis. The estimates obtained from both approaches were then compared with those obtained from ECG-derived PPG time series, considered as the gold standard reference in this work.

II. DENOISING ALGORITHM

A schematic overview of the denoising algorithm is depicted in Fig.1. A CNN-DAE is designed to denoise and reconstruct PPG time series. The model is trained based on a custom loss function which combines the error between the predicted output and clean reference PPG time series, as well as the error between their respective predicted and reference RR time series. Reference data is derived through a PPG simulation model which takes the RR information obtained from synchronous ECG recordings as input.

A. CNN-DAE model

The model is based on the structure of a DAE, consisting of an encoder, a bottleneck, and a decoder. The encoder is made of five blocks. Each block consists of a convolutive layer, an exponential linear unit (ELU) activation function, batch normalization, and an average pooling layer of size 2. We adopted 32 output kernels for the first convolutive layer, and 64 output kernels for the others. We set the alpha parameter of ELU layers at its default value of 1.

The bottleneck is made of four blocks including a convolutive layer, ELU function, and batch normalization each. The number of output kernels at each CNN block is 128, 256, 128, and 64 respectively.

The decoder is designed to mirror the encoder, with the last block including only a convolutive layer with one output kernel and a linear activation function. We set a kernel size of 32 for all the CNN blocks of the network.

B. Reference PPG data generation

We adopted a phenomenological model of PPG [15], [16] to generate reference counterparts of the noisy input time series. This model leverages RR information of synchronous low-noise ECG to derive an accurate surrogate of the ideally artifact-free PPG. More specifically, single PPG pulses are obtained as the linear combination of three functions: a log-normal and two Gaussians. Notably, the parameters of these functions can be adjusted to model five different types of PPG pulses. Hence, it is possible to account for morphological differences of the pulse shape due to the particular PPG sensor adopted. The width of both the systolic and diastolic parts of each pulse depends on the adjacent RR intervals through two inversely proportional time scale factors. Similarly, the peak amplitude is scaled to account for the variation of ventricular filling time, which could not be negligible for episodes such as ectopic beats [15].

The complete modeled PPG signal $y(t)$ is obtained by placing the pulses at the occurrence of each ECG R peak, and adding a noise term $\epsilon(t)$:

$$y(t) = \sum_{i=1}^N p_i(t - \delta_i) + v(t) \quad (1)$$

where $p_i(t - \delta_i)$ is the i -th PPG peak, δ_i is the time instant at which the corresponding R peak occurs, and N is the number of detected R peaks. The noise term $v(t)$ is obtained by convolving white Gaussian noise with a filter whose spectral response resemble the properties of motion artifacts (see [15] for more details), and makes the generated PPG signal more realistic. Since our aim was to obtain an artifact-free reference to train the network with, we did not include the additive error term to the model equation.

C. Loss function

We designed a physiologically-constrained loss function which exploits the the reference PPG and its IBI time series:

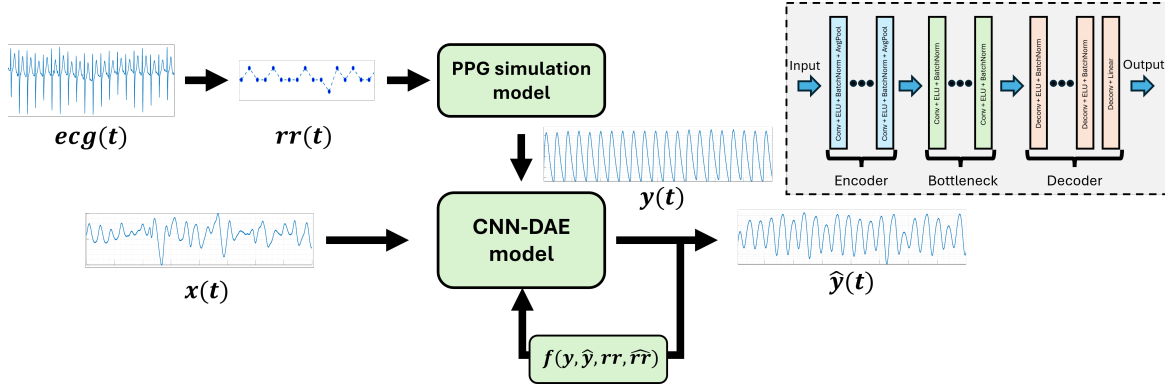


Fig. 1. Schematic representation of the proposed denoising algorithm. Noisy PPG time series $x(t)$ are provided as input to a convolutional-denoising autoencoder (CNN-DAE) to remove motion artifacts and reconstruct pulse peaks $\hat{y}(t)$. The network is also provided with a synthetic clean reference PPG $y(t)$ of the input. This is obtained through a simulation model which leverages the distance between consecutive ECG R peaks (rr). The model is trained based on a custom loss function $f(y, \hat{y}, rr, \hat{r}\hat{r})$ which takes into account both the mean squared error (MSE) between predicted and reference PPGs (i.e., $\hat{y}(t)$ and $y(t)$), and the MSE between their inter-beat intervals (i.e., $\hat{r}\hat{r}$ and rr). A block diagram representing the CNN-DAE structure is depicted in the gray box at the top-right of the figure.

$$f(y, \hat{y}, rr, \hat{r}\hat{r}) = MSE(y, \hat{y}) + \lambda \cdot MSE(rr, \hat{r}\hat{r}) \quad (2)$$

the first term evaluates the MSE between the reference PPG y and the predicted output $\hat{y}$. The second term instead evaluates the MSE between the reference IBI (rr) and the IBI estimated from the model output PPG ($\hat{r}\hat{r}$), and combines with the first term through the scalar parameter λ . While the first term allows for the network to learn how to reproduce the patterns of clean PPG signal, the second term confers knowledge to the network about plausible distances between consecutive peaks, acting as a fine tuning over the first term. Moreover, it provides a regularization mechanism through which spurious peaks or physiological peaks not properly reconstructed are penalized.

To achieve this, at each training step of the model, we identified pulse peaks from both y and $\hat{y}$ through an implementation of the validated multiscale peaks and troughs detection (MSPTD) algorithm [17]. We then derived the predicted and reference IBIs (i.e., $\hat{r}\hat{r}$ and rr) as the temporal distance between adjacent PPG pulse peaks. Since the CNN-DAE model may fail to provide an accurate reconstruction of each PPG pulse, additional spurious pulses and missing true peaks may occur. Accordingly, the length of $\hat{r}\hat{r}$ may differ from that of rr . To account for these scenarios, we adjusted the length of the two IBI time series as follows:

$$\hat{r}\hat{r} = \begin{cases} [\hat{r}\hat{r}_1, \dots, \hat{r}\hat{r}_N, \mathbf{0}] & \text{if } N' > N \\ [\hat{r}\hat{r}_1, \dots, \hat{r}\hat{r}_{N'}, \mathbf{a}] & \text{if } N' < N \\ [\hat{r}\hat{r}_1, \dots, \hat{r}\hat{r}'_N] & \text{if } N' = N \end{cases} \quad (3)$$

$$rr = \begin{cases} [rr_1, \dots, rr_N, \mathbf{a}] & \text{if } N' > N \\ [rr_1, \dots, rr'_N] & \text{if } N' \leq N \end{cases} \quad (4)$$

where $\mathbf{0} \in \mathbb{R}^{1 \times |N' - N|}$ is a vector of zeros, $\mathbf{a} \in \mathbb{R}^{1 \times |N' - N|}$ is a vector of scalar penalty terms of value a , and N' and N are the length of $\hat{r}\hat{r}$ and rr . If the number of peaks found in

$\hat{y}$ matches the number of true peaks in y , we simply compute the MSE between $\hat{r}\hat{r}$ and rr . On the other hand, if $\hat{y}$ contains more (i.e., spurious) peaks, we replace the exceeding $N' - N$ terms in $\hat{r}\hat{r}$ with zeros, and we extend rr to the length of $\hat{r}\hat{r}$ by appending a vector of penalty terms a to it. Similarly, if the number of peaks found in $\hat{y}$ is less than the expected number (i.e., the model has "missed" some peak), we extend the length of $\hat{r}\hat{r}$ appending $|N' - N|$ penalty terms to it. Accordingly, the error between the two IBI time series will be maximum (i.e., a) for those peaks that have been mistakenly reconstructed (spurious or missing peaks), whereas the true error will be computed for the distances between all the other detected peaks.

III. MATERIALS AND METHODS

In this section, we evaluate the performances of our denoising model on a publicly available dataset [18], [19]. First, we describe the dataset. We then illustrate the procedure followed to train the model, as well as the analyses conducted to evaluate its performances against a standard procedure based on bandpass filtering and power spectrum analysis for the estimation of mean HR.

A. Dataset

We adopted the publicly available dataset proposed for the IEEE Signal Processing Cup (IEEE SPC) 2015 [18], [19]. Twelve healthy participants (age: 18-36) were asked to perform the following motion activities: (1) 30s of resting-state, (2) running at the speed of 6km/h for 1min, (3) running at the speed of 12km/h for 1min, (4) decrease the speed to 6km/h for 1min, (5) increase again the running speed to 12km/h, and (6) 30s of final resting-state. Recordings included signals from a three-axis accelerometer, two equivalent PPG channels, and one ECG channel. Signals were simultaneously collected at the sampling rate of 125 Hz, and recordings lasted for about 5min. In this work, we used the first PPG channel and the ECG channel.

For each subject, we segmented PPG data into 8s-long windows (1000 samples) [14]. We adopted an overlap of 7s between consecutive windows to augment the data size. Finally, we normalized the amplitude within each window between 0 and 1. The same procedure was applied to the reference clean PPG data, which was obtained according with the procedure illustrated in Section II-B.

B. Model training

We trained the CNN-DAE model described in Section II-A using a batch size of 32. Accordingly, input data had a shape of $\{32, 1, 1000\}$, where the first dimension is the batch size, the second is the number of features per sample (i.e., the PPG amplitude), and the third is the window length in samples. We set the parameters of the loss function to $\lambda = 10^{-3}$ and $a = 0.250$. We adopted the Adam optimizer with an initial learning rate of 5×10^{-4} , and we gradually decreased it by a factor of 0.5 down to 1×10^{-6} whenever the validation loss did not improve for five consecutive epochs. A procedure based on model check-points and early stopping was implemented to prevent overfitting. Specifically, the model parameters were saved whenever an improvement on the validation set was observed, and the training was stopped when no improvements occurred for 10 consecutive epochs. We implemented the CNN-DAE model and the loss function computation in PyTorch, and the training was carried out using a server from the University of Pisa with four GPU NVIDIA A100 Tensor Core.

We trained the denoising model through a leave-one-subject-out (LOSO) 10-fold cross-validation (CV). Accordingly, for each subject, the model was trained on the data from the other 11 participants, and tested on the excluded subject. For each subject, we estimated the IBI from the model output. Then, we derived the PRV time series through cubic interpolation at the sampling frequency of 4Hz, and we estimated the mean HR within windows of 30s with no overlap. These steps were performed using the software Kubios HRV [20].

C. Model performance evaluation

To evaluate the performances of our approach, we repeated the estimation of mean HR from the choosen dataset using a standard procedure. Specifically, for each subject, we bandpass filtered PPG data in the (0.5-5)Hz, and we estimated the power spectral density within non-overlapping windows of length 30s using an autoregressive (AR) model of order 100 [14]. We estimated mean HR as $60 \cdot f_{HR}$, where f_{HR} is frequency associated with the maximum of the power spectrum in the (0.75-3) Hz range, corresponding to a beats-per-minute (bpm) range of (45-180) bpm.

For both the denoising model and the spectral approach, we computed the mean absolute error (MAE) between the estimated mean HR and the gold standard reference values derived from the HRV of ECG. We compared the MAE distributions through a Wilcoxon sign-rank test ($\alpha=0.05$) to assess whether the model showed a significant improvement in the estimation error with respect to the standard approach.

Moreover, we conducted a Bland-Altman analysis to investigate for estimation biases and correlations with the gold standard measurements.

IV. RESULTS

A. Model training results

In Fig.2 we report the training and validation loss curves for an exemplary subject. Across subjects, the model required an average of 80 epochs to train. Training was stopped when the performances on the validation set did not improve for 10 consecutive epochs, and the model weights were restored to those associated with the last epoch with an improvement.

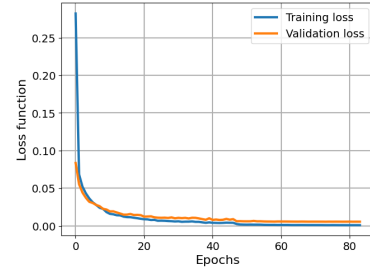


Fig. 2. Model's training loss (blue) and validation loss (red) trends across epochs for an exemplary subject during LOSO procedure.

In Fig.3c we also report an example of the predicted output provided by the proposed denoising algorithm, together with the corresponding noisy input (Fig.3a) and reference clean (Fig.3b) PPGs, for an interval of 8s. It can be observed that the model successfully removed motion artifacts and reconstructed original pulse peaks.

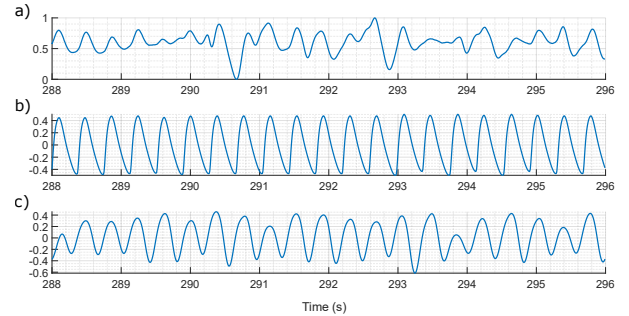


Fig. 3. PPG signals for an exemplary subject: a) original time series; b) clean reference; c) denoised and reconstructed model output. Time series are depicted over a window of 8s. The amplitude is normalized in the (0-1) interval and then mean-centered around 0.

B. Model performance evaluation results

Our model showed a significant improvement of the mean HR estimation error with respect to the standard spectral method. Specifically, as depicted in Fig.4, the standard method estimated mean HR with a median MAE of 12.31bpm (interquartile range (IQR): 5.83-18.86 bpm). Conversely, the median MAE committed by our model was 2.49bpm (IQR: 0.89-4.44 bpm).

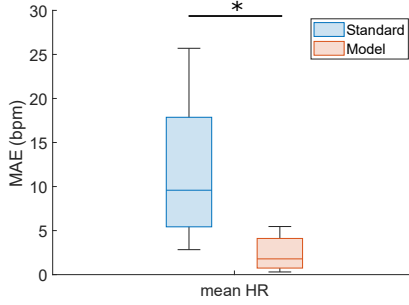


Fig. 4. Comparison of the mean absolute error (MAE) between the standard spectral approach (blue) and the denoising model (red). The solid horizontal line inside each box indicates the median, whereas the top and bottom of the boxes indicate the interquartile range (IQR). The denoising model demonstrated a significant reduction in estimation error compared to the standard approach (*: $p < 0.05$).

Table I shows the MAE of the denoising model and the standard method for each subject in the dataset. There is a consistent improvement of the estimation error across the entire dataset, with an overall average MAE of 3.68bpm for the model, compared to the average MAE of 13.49bpm committed using the standard approach.

In Fig.5-6 we report the BA results on the mean HR estimates obtained from the spectral approach (Fig.5a-6a) and our denoising model (Fig.5b-6b), with respect to the gold standard estimates obtained from ECG. Both approaches were able to estimate mean HR with no significant bias with respect to ECG. Nevertheless, the denoising model showed lower limits of agreement (LOA) compared to the spectral method, indicating that it is capable of providing more consistent and closer estimates to those of the gold standard. Regarding the correlation between estimates, the spectral method did not exhibit a significant relationship with ECG (Fig.6a). On the other hand, the estimates obtained through the denoising model showed a significant correlation with those provided by ECG, with a Pearson correlation coefficient ρ of 0.92 and an r^2 index of 0.84 (Fig.6b).

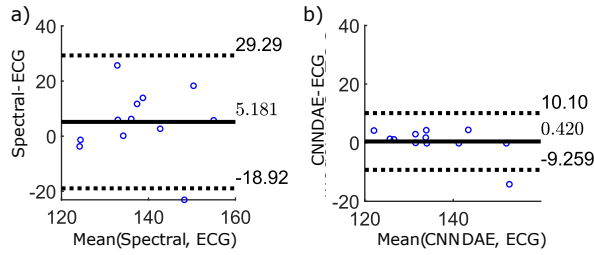


Fig. 5. Bland-Altman (BA) bias analysis of the mean HR provided by the spectral method (a) and CNN-DAE model (b), against the estimates obtained from ECG analysis. For each of the approaches, we report the bias, computed as the average difference between the method and gold standard estimates, and the limits of agreement (LOA), indicating the ± 1.96 standard deviation interval around the bias. Both approaches showed no significant estimation bias. However, the LOA were lower for the denoising model estimates, compared to those obtained from the spectral method.

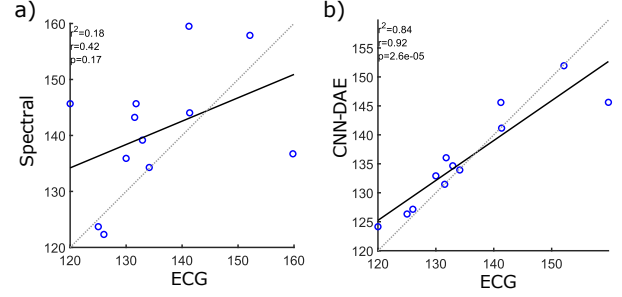


Fig. 6. Bland-Altman (BA) correlation analysis of the mean HR provided by the spectral method (a) and CNN-DAE model (b), against the estimates obtained from ECG analysis. For each of the estimation methods, we report the Pearson correlation coefficient r , the coefficient of determination r^2 , and the pvalue associated with the significance of the linear regression relationship between measurements. There is no significant linear correlation between spectral method's estimates and those obtained from ECG. Conversely, the CNN-DAE model estimates showed a significant correlation with ECG ones, reporting a correlation coefficient of 0.92.

V. DISCUSSION

We proposed a preliminary approach to denoise PPG data acquired during high-intensity activities from a wearable wrist device, aiming to improve the reliability of mean HR estimates from PRV. To this aim, we designed an ad-hoc CNN-DAE model that was trained using a physiologically-constrained custom loss function, and we evaluated the estimation error using ECG's HRV as the gold standard. We compared the performances of our approach against those obtained using a procedure which combines standard bandpass filtering and spectral analysis. Our preliminary results indicate that the proposed methodology could effectively denoise PPG time series and provide significant PRV estimates in everyday settings.

The performances' analysis showed that the proposed denoising algorithm was able to estimate mean HR with an average MAE of 3.68bpm, outperforming some previously proposed DL approaches tested on the same dataset [13]. We observed an improvement of the estimation error for all the subjects in the dataset, with errors as low as 0.30 bpm. Still, one subject showed a high residual error, suggesting that our methodology could be further improved against outliers. Furthermore, BA analysis showed a high agreement of mean HR estimates between the proposed model and ECG, without significant bias. However, we observed a tendency for both the standard approach and the model to overestimate mean HR with respect to ECG. This could be due to either a higher number of peaks detected by the peak detection algorithm or the intrinsic differences between ECG and PPG [1].

More sophisticated spectral approaches compared to the baseline method adopted in this study showed an even higher accuracy with respect to our denoising model [9], [14]. It is worth noting, however, that these methodologies are focused on providing a direct estimation of the mean HR, which can not be adopted to estimate PRV time series. Our approach focuses instead on the accurate denoising and reconstruction

TABLE I

COMPARISON OF THE MEAN HR ESTIMATION ERROR BETWEEN THE PROPOSED CNN-DAE MODEL AND THE STANDARD SPECTRAL ANALYSIS. FOR EACH SUBJECT OF THE DATASET, IS REPORTED THE MEAN ABSOLUTE ERROR (MAE) OF THE MEAN HR ESTIMATES AGAINST THOSE OBTAINED FROM ECG'S HRV. THE LAST COLUMN OF THE TABLE REPORTS THE OVERALL MAE ACROSS SUBJECTS.

MAE(bpm)	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	Avg
CNN-DAE	5.46	4.51	1.37	3.34	0.60	3.19	0.30	1.18	1.79	17.51	0.56	4.37	3.68
Spectral	16.62	25.70	15.05	8.26	3.27	9.58	2.84	19.45	5.02	31.12	6.64	18.28	13.49

of PPG pulses, so that the distance between consecutive pulses and, hence, PRV can be derived.

In the methodology proposed by [14], they combine the outcome of several denoising CNNs to reconstruct PPG time series. While their DL approach shows on average a lower estimation error compared to ours, the performances drop when ensemble learning is not used. On the one hand, this suggests that our CNN-DAE architecture is capable of providing a better reconstruction of the PPG signal. On the other hand, their results highlight that combining the outcome from several models may yield to better denoising performances. Indeed, neural networks are prone to fall into different local minima due to weights initialization. In this context, combining the output of multiple models has been proved to reduce estimation variance and improve generalization [21]. Accordingly, future studies will aim to integrate ensemble learning techniques into the proposed denoising algorithm.

One of the greatest advantages of PRV time series analysis regards the investigation of autonomic balance through its power spectrum components. Nevertheless, reaching a high estimation accuracy of such spectral parameters is a more challenging task compared to mean HR estimation, as it requires highly-reliable PRV time series. In this light, future works will focus on improving our denoising algorithm to raise the accuracy of relevant PRV frequency-domain parameters and on the comparison with other previously validate deep-learning approaches.

REFERENCES

- [1] E. Mejía-Mejía, J. M. May, R. Torres, and P. A. Kyriacou, "Pulse rate variability in cardiovascular health: A review on its applications and relationship with heart rate variability," *Physiological Measurement*, vol. 41, no. 7, p. 07TR01, 2020.
- [2] N. Milstein and I. Gordon, "Validating measures of electrodermal activity and heart rate variability derived from the empatica e4 utilized in research settings that involve interactive dyadic states," *Frontiers in Behavioral Neuroscience*, vol. 14, p. 148, 2020.
- [3] G. Rho, F. Di Rienzo, C. Marinai, F. Giannetti, L. Arcarisi, P. Bufano, M. Zanoletti, F. Righetti, C. Vallati, M. Laurino, *et al.*, "Preliminary assessment of the samsung galaxy watch 5 accuracy for the monitoring of heart rate and heart rate variability parameters," in *Mediterranean Conference on Medical and Biological Engineering and Computing*, pp. 22–30, Springer, 2023.
- [4] M. R. Ram, K. V. Madhav, E. H. Krishna, N. R. Komalla, and K. A. Reddy, "A novel approach for motion artifact reduction in ppg signals based on as-lms adaptive filter," *IEEE Transactions on Instrumentation and Measurement*, vol. 61, no. 5, pp. 1445–1457, 2011.
- [5] A. Galli, C. Narduzzi, and G. Giorgi, "Measuring heart rate during physical exercise by subspace decomposition and kalman smoothing," *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 5, pp. 1102–1110, 2017.
- [6] R. Jaafar and M. A. A. Rozali, "Estimation of breathing rate and heart rate from photoplethysmogram," in *2017 6th International Conference on Electrical Engineering and Informatics (ICEEI)*, pp. 1–4, IEEE, 2017.
- [7] D. Dao, S. M. Salehizadeh, Y. Noh, J. W. Chong, C. H. Cho, D. McManus, C. E. Darling, Y. Mendelson, and K. H. Chon, "A robust motion artifact detection algorithm for accurate detection of heart rates from photoplethysmographic signals using time–frequency spectral features," *IEEE journal of biomedical and health informatics*, vol. 21, no. 5, pp. 1242–1253, 2016.
- [8] E. Khan, F. Al Hossain, S. Z. Uddin, S. K. Alam, and M. K. Hasan, "A robust heart rate monitoring scheme using photoplethysmographic signals corrupted by intense motion artifacts," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 3, pp. 550–562, 2015.
- [9] Z. Zhang, Z. Pi, and B. Liu, "Troika: A general framework for heart rate monitoring using wrist-type photoplethysmographic signals during intensive physical exercise," *IEEE Transactions on biomedical engineering*, vol. 62, no. 2, pp. 522–531, 2014.
- [10] S.-H. Park, S.-J. Choi, and K.-S. Park, "Advance continuous monitoring of blood pressure and respiration rate using denoising auto encoder and lstm," *Microsystem Technologies*, vol. 28, no. 10, pp. 2181–2190, 2022.
- [11] J. Lee, S. Sun, S. M. Yang, J. J. Sohn, J. Park, S. Lee, and H. C. Kim, "Bidirectional recurrent auto-encoder for photoplethysmogram denoising," *IEEE journal of biomedical and health informatics*, vol. 23, no. 6, pp. 2375–2385, 2018.
- [12] D. Biswas, L. Everson, M. Liu, M. Panwar, B.-E. Verhoef, S. Patki, C. H. Kim, A. Acharyya, C. Van Hoof, M. Konijnenburg, *et al.*, "Cornet: Deep learning framework for ppg-based heart rate estimation and biometric identification in ambulant environment," *IEEE transactions on biomedical circuits and systems*, vol. 13, no. 2, pp. 282–291, 2019.
- [13] A. Reiss, I. Indlekofer, P. Schmidt, and K. Van Laerhoven, "Deep ppg: Large-scale heart rate estimation with convolutional neural networks," *Sensors*, vol. 19, no. 14, p. 3079, 2019.
- [14] X. Chang, G. Li, G. Xing, K. Zhu, and L. Tu, "Deepheart: a deep learning approach for accurate heart rate estimation from ppg signals," *ACM Transactions on Sensor Networks (TOSN)*, vol. 17, no. 2, pp. 1–18, 2021.
- [15] A. Sološenko, A. Petrėnas, V. Marozas, and L. Sörnmo, "Modeling of the photoplethysmogram during atrial fibrillation," *Computers in biology and medicine*, vol. 81, pp. 130–138, 2017.
- [16] B. Paliakaitė, A. Petrėnas, A. Sološenko, and V. Marozas, "Photoplethysmogram modeling of extreme bradycardia and ventricular tachycardia," in *XV Mediterranean Conference on Medical and Biological Engineering and Computing–MEDICON 2019: Proceedings of MEDICON 2019, September 26-28, 2019, Coimbra, Portugal*, pp. 1165–1174, Springer, 2020.
- [17] S. M. Bishop and A. Ercole, "Multi-scale peak and trough detection optimised for periodic and quasi-periodic neuroscience data," in *Intracranial Pressure & Neuromonitoring XVI*, pp. 189–195, Springer, 2018.
- [18] A. K. Ahmadi, P. Moradi, M. Malihi, S. Karimi, and M. B. Shamsollahi, "Heart rate monitoring during physical exercise using wrist-type photoplethysmographic (ppg) signals," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 6166–6169, IEEE, 2015.
- [19] Z. Zhang, "Photoplethysmography-based heart rate monitoring in physical activities via joint sparse spectrum reconstruction," *IEEE transactions on biomedical engineering*, vol. 62, no. 8, pp. 1902–1910, 2015.
- [20] M. P. Tarvainen, J.-P. Niskanen, J. A. Lipponen, P. O. Ranta-Aho, and P. A. Karjalainen, "Kubios hrv–heart rate variability analysis software," *Computer methods and programs in biomedicine*, vol. 113, no. 1, pp. 210–220, 2014.
- [21] K. Kawaguchi, "Deep learning without poor local minima," *Advances in neural information processing systems*, vol. 29, 2016.