

Article

UNet-BiLSTM: A Deep Learning Method for Reconstructing Electrocardiography from Photoplethysmography

Yanke Guo ¹, Qunfeng Tang ², Zhencheng Chen ^{2,*} and Shiyong Li ^{1,*}

¹ School of Electronic Engineering and Automation, Guilin University of Electronic Technology, Guilin 541004, China; guoyk1219@foxmail.com

² School of Life and Environmental Sciences, Guilin University of Electronic Technology, Guilin 541004, China; tangqunfeng@guet.edu.cn

* Correspondence: chenzhcheng@guet.edu.cn (Z.C.); lishiyong@guet.edu.cn (S.L.)

Abstract: Electrocardiography (ECG) is generally used in clinical practice for cardiovascular diagnosis and for monitoring cardiovascular status. It is considered to be the gold standard for diagnosing cardiovascular diseases and assessing cardiovascular status. However, it is not always easy to obtain. Unlike ECG devices, photoplethysmography (PPG) devices can be placed on body parts such as the earlobes, fingertips, and wrists, making them more comfortable and easier to obtain. Several methods for reconstructing ECG signals using PPG signals have been proposed, but some of these methods are subject-specific models. These models cannot be applied to multiple subjects and have limitations. This study proposes a neural network model based on UNet and bidirectional long short-term memory (BiLSTM) networks as a group model for reconstructing ECG from PPG. The model was verified using 125 records from the MIMIC III matched subset. The experimental results demonstrated that the proposed model was, on average, able to achieve a Pearson's correlation coefficient, root mean square error, percentage root mean square difference, and Fréchet distance of 0.861, 0.077, 5.302, and 0.278, respectively. This research can use the correlation between PPG and ECG to reconstruct a better ECG signal from PPG, which is crucial for diagnosing cardiovascular diseases.

Keywords: ECG reconstruction; electrocardiography (ECG); photoplethysmography (PPG); bidirectional long short-term memory network (BiLSTM); UNet



Citation: Guo, Y.; Tang, Q.; Chen, Z.; Li, S. UNet-BiLSTM: A Deep Learning Method for Reconstructing Electrocardiography from Photoplethysmography. *Electronics* **2024**, *13*, 1869. <https://doi.org/10.3390/electronics13101869>

Academic Editors: Luca Mesin and Maria Filomena Santarelli

Received: 9 March 2024

Revised: 20 April 2024

Accepted: 8 May 2024

Published: 10 May 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

According to the World Health Organization (WHO), cardiovascular diseases (CVDs) are the leading cause of death worldwide [1]. They have become an important problem that seriously threatens global public health. The report states that an estimated 17.9 million people died from cardiovascular diseases in 2019, accounting for 32% of deaths globally. Electrocardiography (ECG) is considered to be the gold standard for diagnosing cardiovascular diseases [2]. An ECG is typically conducted by placing electrodes on the skin to measure the electrical activity of the heart. However, an ECG device is inconvenient because it requires placing multiple electrodes at different locations on the body. This may cause skin irritation and discomfort during recording. Electrodes also may fall off the patient during recording, resulting in incomplete data acquisition. Photoplethysmography (PPG) is a noninvasive method for detecting variations in blood volume to reflect the amount of blood pulsation in tissues [3]. This method can be used to evaluate some cardiac information, such as oxygen saturation [4], blood pressure [5], and cardiac output [6]. Compared with ECG, PPG is easier to set up, more convenient, and more economical. Recently, PPG has been widely used in wearable devices because of its continuous, long-term monitoring capabilities. Although PPG has been widely used for health monitoring [7], ECG remains the standard fundamental measurement for medical diagnosis, with extensive supporting documents and research. PPG and ECG are intrinsically related because the heart's electrical activity influences changes in blood volume. The peak-to-peak interval of

PPG is known to be highly correlated with the RR interval, suggesting the possibility of deriving other ECG parameters from PPG [8]. Therefore, the correlation between ECG and PPG can be used to improve the effectiveness of a method for reconstructing ECG from PPG waveforms. This would enable cost-effective and user-friendly ECG screening for continuous and long-term monitoring, provided that it is possible to successfully reconstruct ECG from PPG obtained from modern wearable devices.

Several studies have used PPG signals to reconstruct ECG signals using various techniques. Several studies have utilized the discrete cosine transform (DCT) method [9], cross-domain joint dictionary learning (XDJDL) method [10], lightweight neural networks [11], bidirectional long short-term memory (BiLSTM) models [12], and the PPG2ECGps model [13] to reconstruct ECG signals for subject-specific models. The first two studies proposed reconstructing ECG from PPG using a mathematical model. The signal preprocessing process required peak detection, data alignment, and beat segmentation. However, it could contain errors due to the extensive preprocessing of the original data. After the signal was divided into beats, the signal lengths could vary. For optimal model training, it was necessary to linearly interpolate different signal lengths to ensure that all signal lengths were the same. At the same time, data alignment was performed after peak detection. During this process, the accuracy of the peak detection algorithm was crucial. However, there could still be some errors in peak detection. The last three studies proposed models that aimed to reconstruct ECG from PPG using deep learning models. The former study proposed using a lightweight neural network to reconstruct ECG signals. This study used the same preprocessing as that in the previous two studies, except that the models proposed in the first two studies were mathematically based models, and this study was based on deep learning models. The last two studies proposed using a BiLSTM model and an end-to-end deep learning neural network model with a W-Net architecture to reconstruct ECG signals, respectively. In the signal preprocessing phase of the former two studies, the signal was not divided based on beats but rather into segments. This avoided the errors associated with beat segmentation. However, these models were all proposed for specific subjects and were unsuitable for multi-subject situations; they have certain limitations.

Several studies utilized the discrete cosine transform (DCT) method [9], P2E-WGAN model [14], CardioGAN model [15], scattering wavelet transform (SWT) method [16], and PPG2ECG model [17] to reconstruct ECG signals for group models. The DCT [9] and SWT [16] model used a beat-to-beat method to reconstruct ECG signals. These methods required that the starting point in the PPG be aligned with the R peak in the ECG signal. Then, the aligned ECG and PPG signals were segmented into beats. The model performed data alignment and beat segmentation on the signal during signal preprocessing. However, this algorithm's accuracy depended on the extraction algorithms' accuracy for the R-wave in the ECG signal and the peak (or onset) of contraction in the PPG signal. If the accuracy of the extraction algorithm was not high, the accuracy of the ECG reconstruction would also decrease. The P2E-WGAN [14] and CardioGAN [15] model proposed using deep neural networks to reconstruct ECG from PPG. They did not require signal alignment during preprocessing. They focused on the destination heart rate without emphasizing the quality of the ECG waveform. The last study reconstructed ECG from PPG in a population model, but the dataset segmentation method differed from the above studies, and beat segmentation was used. The datasets used in these studies were divided as follows: 80% for the training set and 20% for the test set. Their performance on smaller training datasets has not been validated.

This study proposes a new deep neural network model for reconstructing ECG signals from PPG signals. This model is based on bidirectional long short-term memory (BiLSTM) and UNet networks. The method has the advantage of targeting a group model rather than subject-specific models. It also does not require beat segmentation during signal preprocessing. This study divided the dataset into 60% as a training set, 20% as a validation set, and 20% as a test set. In comparison with previous work, this study used a smaller training set to reconstruct an ECG signal that was highly similar to an actual ECG signal.

2. Materials and Methods

This section introduces datasets, data preprocessing, UNet-BiLSTM model structure, and the model performance evaluation. Figure 1 shows a flowchart of the method. Figure 1a shows the training and validation process. Figure 1b shows the testing process.

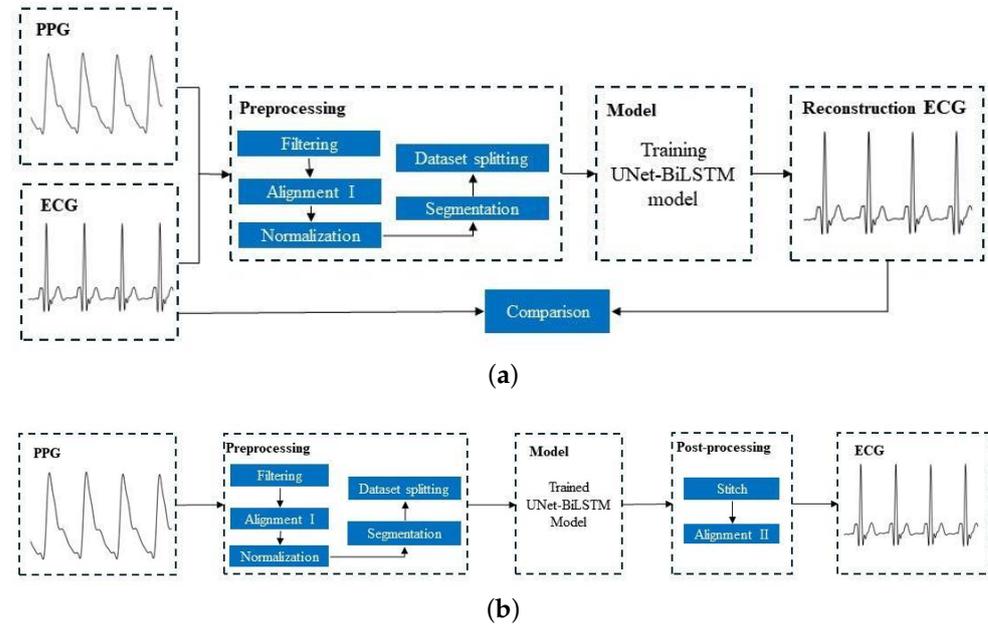


Figure 1. Flowchart of the reconstruction of ECG signals from PPG signals. (a) Training and validation process. (b) Testing process.

All codes in the experiment were implemented in Python 3.9.16, and the UNet-BiLSTM network was implemented using Pytorch 2.0.0, an end-to-end open-source machine learning platform. The UNet-BiLSTM model was trained on a server with the following configuration: CPU 11th Gen Intel(R) Core(TM) i7-11700 @ 2.50 GHz and GPU NVIDIA GeForce RTX 3060 Ti.

2.1. Dataset

The data used to test the model in this study came from the MIMIC III matched subset [18]. The MIMIC III database contains multiple physiological signals from patients in an intensive care unit, and there are numerous records in this subset. This study utilized 125 recordings from different subjects, which included lead II ECG and PPG signals. The sampling rate of both signals was 125 Hz. The length of each record was 5 min.

2.2. Preprocessing

The data preprocessing included filtering, alignment I, normalization, segmentation, and dataset splitting.

- **Filtering:** The ECG signal and PPG signal were filtered. We applied a fourth-order Chebyshev bandpass filter to the ECG signal with a passband frequency of 0.5–20 Hz. Similarly, a fourth-order Chebyshev bandpass filter was applied to the PPG signal with a passband frequency of 0.5–10 Hz.
- **Alignment I:** The Pan–Tompkins method [19] was used to detect the R-wave peak in the ECG signal. A block-based method [20] was used to detect the systolic peak in the PPG signal. Then, the third systolic peak in the PPG signal was aligned with the corresponding R peak in the ECG signal. Figure 2 shows the signals before and after alignment I. Figure 2a shows the ECG and PPG before Alignment I. Figure 2b shows the ECG and PPG after alignment I. This step produced a pair of aligned ECG and PPG signals.

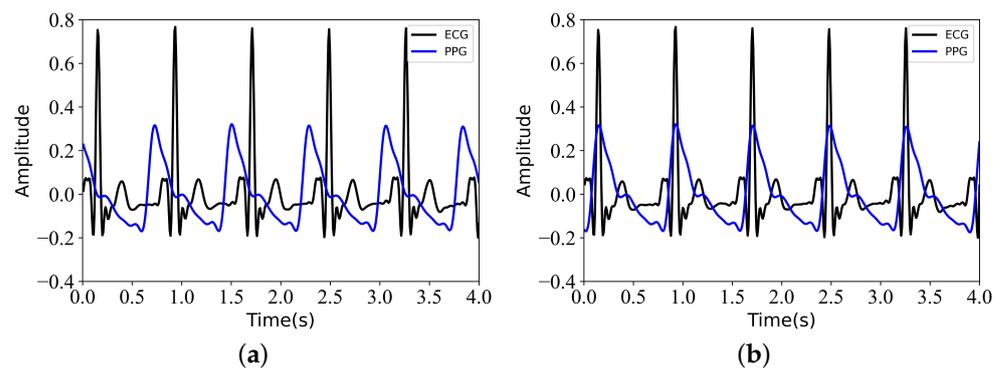


Figure 2. Signals before and after Alignment I. (a) ECG and PPG before Alignment I. (b) ECG and PPG after Alignment I.

- **Normalization:** Since the ECG signal needed to be compared with the reconstructed ECG signal, the PPG signal only needed to be scaled to the range of [0, 1] after aligning the data.
- **Segmentation:** The ECG and PPG signals obtained in the previous step were divided into segments of 3 s. Since the signal alignment would result in a signal length of less than 300 s, it was necessary to ensure that the length of each record was consistent. To maintain consistency in the length of the training data, we only considered the first 294 s of data and disregarded any data beyond that. Specifically, each record was divided into 3 s.
- **Dataset splitting:** In particular, the first 60% of each recording was used for training, the next 20% of each recording was used for validation, and the remaining 20% of each recording was used for testing.

2.3. Model Architecture

The model structure of the proposed combination of UNet and BiLSTM is shown in Figure 3. In Figure 3, the terms ‘Conv’, ‘ConvTrans’, and ‘Upsample’ represent a one-dimensional convolution layer, a one-dimensional transposed convolution layer, and an upsampling layer, respectively. ‘ReLU’ and ‘Tanh’ refer to the activation functions of the corresponding convolution layers. ‘BN’ represents a one-dimensional batch normalization layer. ‘Dropout’ represents a dropout layer. ‘BiLSTM’ represents a bidirectional long short-term memory layer. The slope of the ‘Dropout’ activation was set to 0.5.

As shown in Figure 3, the proposed Unet-BiLSTM model consisted of a one-dimensional convolution-based “UNet” encoder–decoder architecture [21] and a BiLSTM network. We chose the BiLSTM model because it has been proven to effectively solve sequential and time-series problems [22,23]. A study on generating ECG signals also demonstrated that the BiLSTM model is robust when generating ECG signals [24]. Long short-term memory (LSTM) and BiLSTM are suitable for handling time-series problems. BiLSTM models take longer to reach equilibrium than LSTM models but provide better performance [25]. The U-block was inspired by the wave UNet [26]. The motivation for employing UNet in this study was its simple structure, which allows feature extraction and reconstruction from multiple dimensions through a symmetric cross-layer connection, even with a limited dataset. In general, UNet has a contracting path (the left side) and an expansive path (the right side), which are symmetric. There were four downsampling blocks and BiLSTM layers on the left side and four upsampling blocks on the right side. In this downsampling block, a convolutional layer was used instead of a pooling layer. The kernel size and stride of the convolutional layer were 4 and 2, respectively. In this upsampling block, a transposed convolutional layer was used. The kernel size and stride of the transposed convolutional layer were 4 and 2, respectively. A BiLSTM layer was added to the downsampling process to process the sequence data effectively. Dropout layers were added to improve the generalization ability of UNet–BiLSTM and to reduce overfitting.

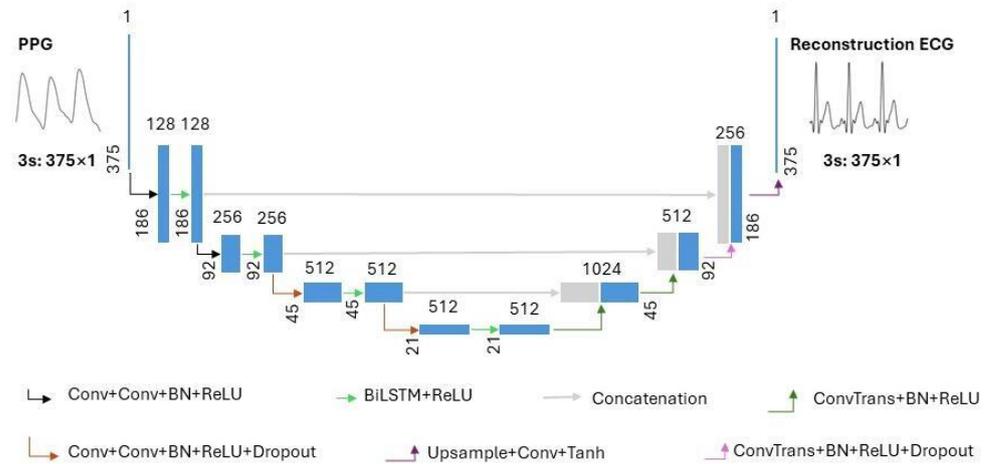


Figure 3. The architecture of the proposed Unet-BiLSTM model. ‘Conv’, ‘ConvTrans’, and ‘Upsample’ represent a one-dimensional convolution layer, a one-dimensional transposed convolution layer, and an upsampling layer, respectively. ‘ReLU’ and ‘Tanh’ refer to the activation functions used in the corresponding convolution layers. ‘BN’ represents a one-dimensional batch normalization layer. ‘Dropout’ represents a dropout layer. ‘BiLSTM’ represents bidirectional long short-term memory.

2.4. Training Options

The UNet–BiLSTM model proposed in this study used the Adam optimizer for training. Setting appropriate stopping criteria during the training of a neural network is of utmost importance to achieve optimal performance while avoiding overfitting. The neural network trained for 500 epochs while utilizing a batch size of 256 pairs of ECG and PPG segments for all recordings. The learning rate was set to 0.001 and decayed by a factor of 0.1 every 200 steps. The loss function used in this study is defined as follows:

$$Loss = \frac{1}{l} \sum_{i=1}^l (E(i) - E_r(i))^2 \quad (1)$$

The loss function used the mean square error. $E(i)$ and $E_r(i)$ represent the i th sample points of the reference and reconstructed ECG signals, respectively. The variable l represents the sample size of the reference ECG.

Regularization was implemented to address or prevent overfitting of ill-posed problems [27]. In this study, UNet–BiLSTM used Tikhonov regularization L_2 . The kernel regularizer parameter in UNet–BiLSTM was $L_2 = 1 \times 10^{-6}$.

2.5. Stitching the Reconstructed ECG Segments and Alignment II

- **Stitching the reconstructed ECG segments:** The neural network’s output consisted of 375 samples of reconstructed ECG segments, each of which was 3 s long. Therefore, they needed to be spliced together to form a continuous reconstructed ECG signal. The second ECG segment was placed after the first ECG segment when combining two ECG segments. The spliced signal was used as the first segment, and the subsequent segment was used as the second segment for further merging. This step was repeated until all test segments in the record were joined together.
- **Alignment II:** The result of splicing was an ECG signal that had already been reconstructed, and it was aligned using cross-correlation. After visualizing the reconstructed and reference ECGs, it was discovered that there was some offset between some of the recorded ECGs (some distance between the R-wave crests of the reference and reconstructed ECGs). Cross-correlation alignment is used to minimize the distance between the R-wave peaks of the reference and reconstructed ECGs. This alignment was primarily performed to improve the evaluation of the similarity between the reconstructed and reference signal.

2.6. Performance Evaluation

To evaluate the performance of the proposed model on both the reference and reconstructed ECG, we used several metrics for evaluation in the test set. These metrics included Pearson's correlation coefficient (r) [28], the root mean squared error (RMSE), the Fréchet distance (FD) [29], and the percentage root mean squared difference (PRD).

- **Pearson's correlation coefficient (r):** The r is a statistical measure that can be used to assess the strength and direction of the linear correlation between two variables. The absolute value of r is in the range of $[0, 1]$. A correlation coefficient approaching 1 indicates a strong correlation, whereas a coefficient approaching 0 indicates a weak correlation. r is given by the following equation:

$$r = \frac{\sum_{i=1}^l (E(i) - \bar{E}) \sum_{i=1}^l (E_r(i) - \bar{E}_r)}{\sum_{i=1}^l (E(i) - \bar{E})^2 \sum_{i=1}^l (E_r(i) - \bar{E}_r)^2} \quad (2)$$

In the given formula, $E(i)$ and $E_r(i)$ represent the individual sample points of the reference ECG signal and the reconstructed ECG signal, respectively, with both being indexed by i . The variable l represents the number of samples of the reference ECG. The symbols \bar{E} and \bar{E}_r denote the mean values of the ECG signal and the reconstructed ECG signal, respectively.

- **Root mean square error (RMSE):** The RMSE is a metric used to quantify the discrepancy—commonly referred to as the error—between the measured value of an ECG signal and its corresponding reconstructed value. The RMSE is a quantitative measure used to assess the level of deviation between predicted and actual values. The value in question is a non-negative value that ranges from zero to positive infinity. The closer the value of the RMSE is to zero, the more optimal the reconstruction outcomes become. The RMSE was calculated with the following equation:

$$\text{RMSE} = \sqrt{\frac{1}{l} \sum_{i=1}^l (E(i) - E_r(i))^2} \quad (3)$$

- **Percentage root mean squared difference (PRD):** The PRD was calculated to quantify the distortion between the reference signal E and the reconstructed signal E_r . The value of the PRD was defined within the interval $[0, +\infty]$. The quality of the reconstruction results was enhanced, with a decrease in the PRD value. The following equation was used to calculate the PRD:

$$\text{PRD} = \sqrt{\frac{\sum_{i=1}^l (E(i) - E_r(i))^2}{\sum_{i=1}^N E(i)^2}} \times 100 \quad (4)$$

- **Fréchet distance (FD):** The FD is a metric that was utilized to assess the similarity of signals by analyzing the position and order of points on the ECG signal waveform and synthesizing them into a curve. The Fréchet distance quantified the minimum Euclidean distance between corresponding points in the reference and reconstructed ECG signal curve. When calculating the distance between two curves, the distance metric considered the spatial arrangement and sequence of the data points, allowing for a more accurate evaluation of the similarity between the two time-series signals. The value of the FD was defined within the interval $[0, +\infty]$. The closer the FD was to 0, the higher the degree of similarity observed between the reference and reconstructed ECG. The following equation was used to determine the value of the FD:

$$\text{FD} = \min_{i \in Q} (\max(d(E(i), E_r(i))), Q = [1, m]) \quad (5)$$

The function $d(*)$ represents the Euclidean distance between two corresponding points on the reference ECG signal curve and the reconstructed ECG signal curve. The variable m represents the number of sampling points. The maximum distance under this sampling is denoted as $\max_{i \in Q}(d(E(i), E_r(i)))$. The Fréchet distance is the value in the sampling method that minimizes the maximum distance.

3. Results

We tested the accuracy of the model when dividing the data into 3 s segments. After checking the fitting effects for various records, it was found that some records had time delays. Mutual correlation can be used to quantify the displacement between two similar time series. After calculating the cross-correlation between two time series, the maximum value of the cross-correlation function represents the point at which the signals are optimally aligned. The presence of a time delay has the potential to decrease the r -value. To better evaluate the model's performance, we used cross-correlation to align the reconstructed ECG signals with the reference ECG signals.

Two records were selected in this study—one without a time delay and the other with a time delay. Figures 4 and 5 show the experimental results for two selected recordings in 3 s segments. Figure 4 shows the results of the experiment conducted without any time delay. Figure 4a,b show the PPG when using data alignment and when not using data alignment in the preprocessing stage, respectively. Figure 4c,d show the reference and reconstructed ECG when using data alignment and when not using data alignment during preprocessing, respectively. The results of the reference and reconstructed ECG obtained after using cross-correlation in Figure 4c,d are shown in Figure 4e,f, respectively. As shown in Figure 4d,f, r between the reconstructed and reference ECG signal was found to be 0.875. It is important to note that this correlation was obtained when the data were not aligned during preprocessing. r between the reconstructed and reference ECG increased to 0.913 following the implementation of data alignment. r remained unchanged following the application of cross-correlation, indicating an absence of any time delays in the data.

Figure 5 presents the results obtained when a time delay was present. Figure 5c,d show the time delay in the data. The results obtained using cross-correlation are shown in Figure 5e,f. As can be seen in Figure 5, r between the reference and reconstructed ECG signal was 0.819 when using data alignment during data preprocessing. r between the reconstructed and reference ECG signal increased to 0.901 after using the cross-correlation alignment. r between the reference and reconstructed ECG signal was 0.912 when data alignment was not used during data preprocessing. r between the reconstructed ECG signal and the reference ECG signal increased to 0.924 after using the cross-correlation alignment.

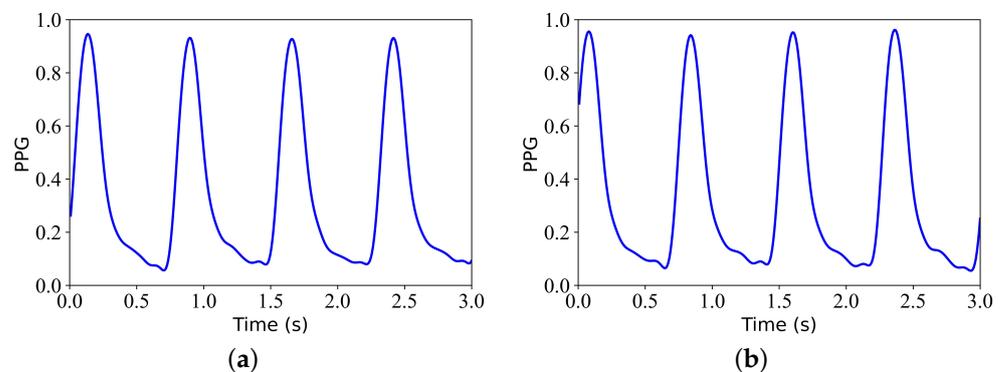


Figure 4. Cont.

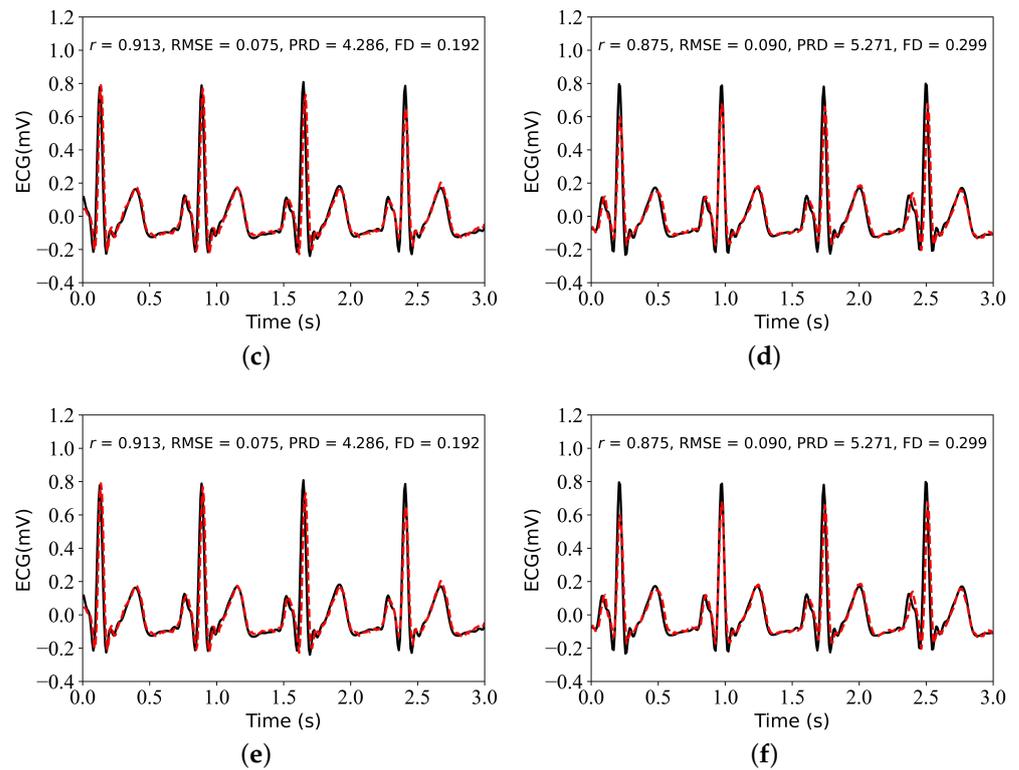


Figure 4. ECG signal reconstruction results without a time delay. The abbreviations r , RMSE, PRD, and FD refer to Pearson’s correlation coefficient, the root mean square error, the percentage root mean squared difference, and the Fréchet distance, respectively. The black line represents the reference ECG. The red line represents the reconstructed ECG. The blue represents the PPG. (a) PPG signal using alignment I. (b) PPG signal without alignment I. (c) Comparison of the reconstructed and reference ECG signal with alignment I. (d) Comparison of the reconstructed and reference ECG signal without alignment I. (e) Comparison of the reconstructed and reference ECG signal with alignment I and alignment II. (f) Comparison of the reconstructed and reference ECG signal without alignment I and using alignment II.

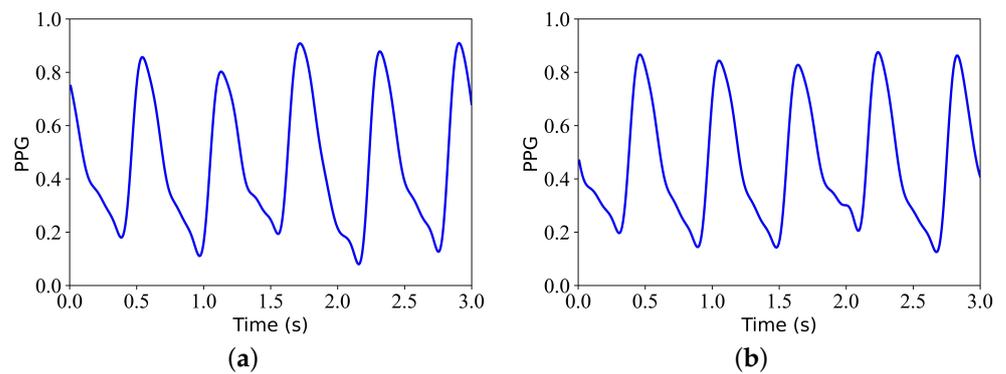


Figure 5. Cont.

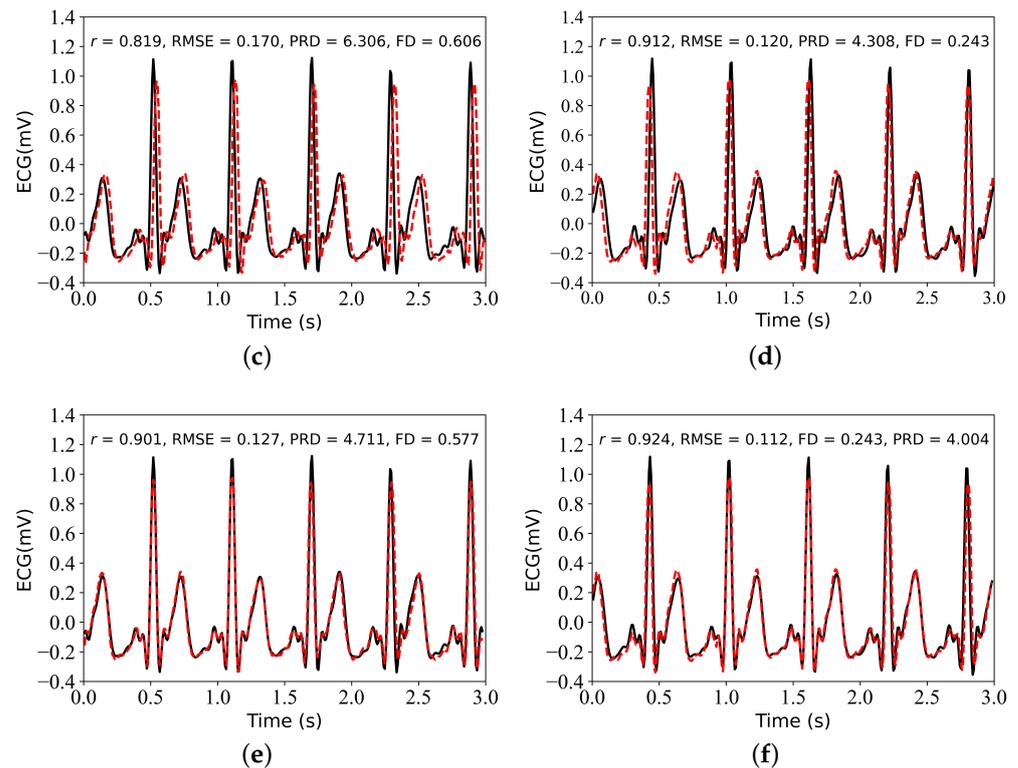


Figure 5. ECG signal reconstruction results with a time delay. The abbreviations r , RMSE, PRD, and FD refer to Pearson's correlation coefficient, the root mean square error, the percentage root mean squared difference, and the Fréchet distance, respectively. The black line represents the reference ECG. The red line represents the reconstructed ECG. The blue represents the PPG. (a) PPG signal using alignment I. (b) PPG signal without alignment I. (c) Comparison of the reconstructed and reference ECG signal with alignment I. (d) Comparison of the reconstructed and reference ECG signal without alignment I. (e) Comparison of the reconstructed and reference ECG signal with alignment I and alignment II. (f) Comparison of the reconstructed and reference ECG signal without alignment I and using alignment II.

This study introduced four experiments aimed at examining the impacts of data alignment on the model performance during the data preprocessing stage, as well as the impacts of cross-correlation alignment on model performance. In Experiments I and II, the data were aligned during preprocessing, but the reconstructed and the reference ECG were not and were aligned using cross-correlation, respectively. In Experiments III and IV, the data were not aligned during preprocessing, but the reconstructed and the reference ECG were and were not aligned using cross-correlation, respectively. Here, the data were not aligned during the preprocessing process, and the length of the divided data was consistent. Therefore, the data length was selected to be 300 s. When data alignment was used, the data length was 294 s.

Figure 6 presents a box plot comparison of the values of r , the RMSE, the PRD, and the FD for the ECG reconstruction in the four experiments. This visualization allows for a comprehensive understanding of the overall distribution of results across these four metrics. We observed that the median and mean values of r , the RMSE, the PRD, and the FD in the four experiments had some discrepancies in the proposed model. In Experiment II, the variable P exhibited the highest median and mean values, whereas the RMSE, FD, and PRD demonstrated the lowest median and mean values. The model's performance was evaluated based on these four experiments, and the corresponding results are shown in Table 1. It can be seen in Table 1 that the model using cross-correlation alignment for the

reconstructed and reference ECG signal performed better. However, the performance of models with and without data alignment in the data preprocessing stage was similar.

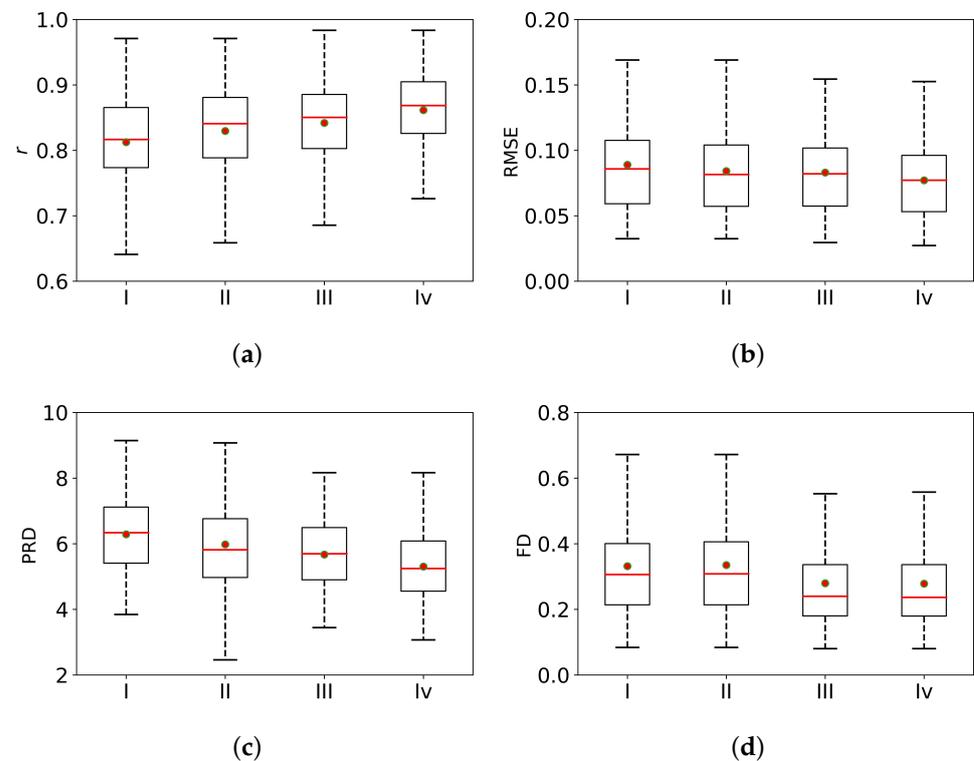


Figure 6. Comparison of the ECG signal reconstruction performance across Experiments I, II, III, and IV. The statistics of (a) the Pearson’s correlation coefficient r , (b) root mean squared error (RMSE), (c) percentage root mean squared difference (PRD), and (d) Fréchet distance (FD) are summarized using box plots.

Table 1. Comparison of the performance of the UNet–BiLSTM model with and without alignment of the reconstructed ECG signal with the reference ECG signal and with and without the alignment of the ECG signal with the PPG signal. Note: NR stands for “not reported”. r , RMSE, FD, and PRD represent Pearson’s correlation coefficient, the root mean square error, the Fréchet distance, and the percentage root mean squared difference, respectively. E , P , and E_r represent the ECG signal, PPG signal, and reconstructed ECG signal, respectively.

	Alignment I	Alignment II	r	RMSE	PRD	FD
Experiment I	Yes	No	0.842 ± 0.061	0.083 ± 0.035	5.672 ± 1.167	0.280 ± 0.149
Experiment II	Yes	Yes	0.861 ± 0.058	0.077 ± 0.030	5.302 ± 1.169	0.278 ± 0.149
Experiment III	No	No	0.812 ± 0.076	0.089 ± 0.036	6.287 ± 1.408	0.332 ± 0.157
Experiment IV	No	Yes	0.830 ± 0.076	0.084 ± 0.034	5.978 ± 1.447	0.335 ± 0.165

4. Discussion

This study proposed a novel model that combined the UNet architecture with a BiLSTM network to reconstruct ECG from PPG. Table 2 shows the results of this study and other studies regarding such group models. As seen in Table 2, the proposed model had some advantages over others. Unlike the DCT model [9] and SWT model [16], the model in this study did not use beat segmentation, but rather divided the signal into 3 s segments. When using the MIMIC III dataset, the DCT model [9] selected 103 records, while this study selected 125 records. The value of r for the DCT model was only 0.79. In contrast, the value of r in this study reached 0.842, and after using cross-correlation to align

the reconstructed ECG signal with the reference ECG signal, the correlation reached 0.861. The SWT model's [16] experimental results section only gives the RMSE and MAE (mean absolute error) values of the reference and reconstructed ECG. Its RMSE value is 0.1006, while the RMSE of this study is 0.077. This study did not calculate the value of MAE, but the RMSE of this study is smaller than the value of the SWT model [16]. Both the model in this study and the P2E-WGAN model [14] performed 3 s segmentation of the signal, and this study's model was better than the P2E-WGAN model in terms of r , the RMSE, and the Fréchet distance. In [14], the proposed model reached 6000 epochs. In [16], based on the loss function diagram in the experimental results section, we know that the number of iterations of the proposed model exceeded 3500. Unlike the P2E-WGAN model [14] and the SWT model [16], the number of epochs in this study was only 500.

Four datasets are utilized in the CardioGAN model [15] to validate the model's accuracy. However, due to the varying sampling frequencies, it was necessary to resample the data. The reconstructed signals were also required to undergo resampling to align with the sampling frequency of the original signal for analysis. Furthermore, the evaluation of the model's performance did not involve using r . The model in this study outperformed the CardioGAN model in terms of the root mean square error, Fréchet distance, and percentage root mean squared difference. Additionally, it did not necessitate resampling the dataset to mitigate errors that could arise during the data resampling process. The DCT model, P2E-WGAN model, and CardioGAN model employed similar approaches by partitioning the dataset into a training set comprising the first 80% of the data and a test set containing the remaining 20%. In this study, the training set consisted of the first 60% of each record, the test set comprised the next 20% of each record, and the validation set comprised the remaining 20% of each record. The division of the dataset in this manner offered the advantage of utilizing a smaller portion of the data to assess the model's accuracy.

Table 2. Evaluation of the UNet–BiLSTM algorithm against population models versus other existing algorithms in the literature for reconstructing ECG signals from PPG signals. Note: NR stands for “not reported”. r , RMSE, FD, and PRD represent Pearson's correlation coefficient, the root mean square error, the Fréchet distance, and the percentage root mean squared difference, respectively. Epoch represents the number of times that the model was run.

Method	Data	Segment Length	r	RMSE	PRD	FD	Epoch
DCT [9]	TBME-RR [30]: 42 Records	Beat	0.906	NR	NR	NR	NR
	MIMIC III [18]: 103 Records		0.790				
	Self-collected: 2 Records		0.895				
P2E-WGAN [14]	MIMIC II [31]: 276 Records	3 s	0.835	0.162	NR	0.375	6000
CardioGAN [15]	BIDMC [32]: 53 Records	4 s	NR	0.364	9.315	0.784	15
	CAPNO [30]: 42 Records						
	DALIA [33]: 15 Records						
	WESAD [34]: 15 Records						
SWT [16]	MIMIC II [31]		NR	0.1006	NR	NR	3500+
This study (UNet–BiLSTM)	MIMIC III [18]: 125 Records	3 s	0.861	0.077	5.302	0.278	500

This study validated the model's performance by verifying whether the data were aligned during preprocessing and whether cross-correlation alignment was used between the reconstructed and the reference ECG signal. When the preprocessing involved data alignment, on average, r between the reconstructed and reference ECG signal with cross-correlation alignment increased from 0.842 to 0.861, the RMSE decreased from 0.083 to 0.077, the Fréchet distance decreased from 0.280 to 0.278, and the percentage root mean squared difference decreased from 5.672 to 5.302. When data alignment was not used during preprocessing, on average, r between the reconstructed and reference ECG signal with cross-correlation alignment increased from 0.812 to 0.830, the RMSE decreased from

0.089 to 0.084, the Fréchet distance increased from 0.332 to 0.335, and the percentage root mean squared difference decreased from 6.287 to 5.978. It can be seen from the comparisons between Experiment I and II and between Experiment III and IV in Table 1 that the average value of r of the reconstructed and the reference ECG signal when using cross-correlation alignment only increased by 0.2. In contrast, the average RMSE, the average Fréchet distance, and the average percentage root mean squared difference slightly decreased. It can be seen from the comparisons between Experiment I and Experiment III and between Experiment II and Experiment IV in Table 1 that the average value of r of the reconstructed ECG and the reference ECG when using data alignment during the preprocessing only increased by 0.2. In contrast, the average RMSE and the average PRD slightly decreased; the average Fréchet distance remained relatively constant.

In this study, we proposed a UNet–BiLSTM model for reconstructing ECG. Although the UNet–BiLSTM model exhibited notable advantages over models used in previous studies, it has limitations.

- In the present study, the ECG signals were filtered using a frequency range below 20 Hz, while frequencies above 20 Hz were not considered. This approach has certain limitations. In subsequent research, we intend to evaluate the efficacy of this model across various frequency ranges.
- The dataset utilized for this research was obtained from the MIMIC III matched subset, which consisted of 125 records. Although the model proposed in this study was designed for group models, the dataset did not provide distinctions based on gender, age, disease, etc. In subsequent research, the dataset will be partitioned based on gender, age, disease, and other relevant factors to evaluate the efficacy of group models.
- This study exclusively focused on the attributes of a complete ECG waveform and did not examine additional features, such as QRS waves and ST segments. In subsequent research, it is imperative to conduct a more comprehensive evaluation of the disparities between reconstructed ECG features and reference ECG features.

5. Conclusions

This study presented a novel structural model that combined the UNet architecture with a BiLSTM network. The model was utilized to reconstruct ECG from PPG. Our proposed methodology involved using 3 s PPG segments to generate ECG segments of equal length. The proposed model verified the impact of the preprocessing process using Alignment I, reconstructed ECG, and referenced ECG using Alignment II on the model performance on the MIMIC III dataset. The experimental results show that using Alignment I and Alignment II improves the model's performance to a certain extent. The experimental results demonstrated this model's effectiveness in reconstructing ECG from PPG. The proposed model demonstrated a distinct impact on the group model. However, it is worth noting that the current average coefficient stands at 0.861, indicating the need for further enhancement. The dataset we selected has 125 records, and we need to verify the model performance on more datasets. Different deep learning techniques can be employed to enhance the performance of group models to obtain improved ECG reconstruction models in future research.

Author Contributions: Z.C. designed the study. Y.G., Q.T., S.L., and Z.C. conceived the study, provided directions and feedback, and revised the manuscript. Y.G. drafted the manuscript for submission with revisions and feedback from the contributing authors. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by a project supported by the Joint Funds of the National Natural Science Foundation of China (U22A2092), the National Major Scientific Research Instrument and Equipment Development Project (61627807), the Guangxi Science and Technology Major Special Project (2019AA12005), and the Innovation Project of GUET Graduate Education (Grant No. 2022YCXB08).

Data Availability Statement: Available upon request.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

BiLSTM	Bidirectional long short-term memory.
CVD	Cardiovascular disease.
DCT	Discrete cosine transform.
ECG	Electrocardiography.
FD	Fréchet distance
MIMIC	Multiparameter Intelligent Monitoring in Intensive Care.
r	Pearson's correlation coefficient.
PPG	Photoplethysmography.
PRD	Percentage root mean squared difference.
RMSE	Root mean square error.
WHO	World Health Organization.
XDJDL	Cross-domain joint dictionary learning.

References

- Cardiovascular Diseases (CVDs). Available online: [https://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)) (accessed on 27 September 2023).
- Sulaiman, S.; Adam, J.R.; Felix, R.; Shan, Z.; Akhil, V.; Fayzan, C.; Jessica, K.D.F.; Nidhi, N.; Riccardo, M.; Girish, N.N.; et al. Deep learning and the electrocardiogram: Review of the current state-of-the-art. *EP Eur.* **2021**, *23*, 1179–1191.
- Reisner, A.; Shaltis, P.A.; McCombie, D.; Asada, H.H. Utility of the photoplethysmogram in circulatory monitoring. *Anesthesiol. J. Am. Soc. Anesthesiol.* **2008**, *108*, 950–958. [[CrossRef](#)] [[PubMed](#)]
- Shelley, K.H. Photoplethysmography: Beyond the calculation of arterial oxygen saturation and heart rate. *Anesth. Analg.* **2007**, *105*, S31–S36. [[CrossRef](#)] [[PubMed](#)]
- Elgendi, M.; Fletcher, R.; Liang, Y.; Howard, N.; Lovell, N.H.; Abbott, D.; Lim, K.; Ward, R. The use of photoplethysmography for assessing hypertension. *NPJ Digit. Med.* **2019**, *2*, 1. [[CrossRef](#)] [[PubMed](#)]
- Wang, L.; Pickwell-Macpherson, E.; Liang, Y.P.; Zhang, Y.T. Noninvasive cardiac output estimation using a novel photoplethysmogram index. In Proceedings of the 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Minneapolis, MN, USA, 3–6 September 2009.
- Denisse, C. A review on wearable photoplethysmography sensors and their potential future applications in health care. *Int. J. Biosens. Bioelectron.* **2018**, *4*, 195.
- Allen, J. Photoplethysmography and its application in clinical physiological measurement. *Physiol. Meas.* **2007**, *28*, R1–R39. [[CrossRef](#)] [[PubMed](#)]
- Zhu, Q.; Tian, X.; Wong, C.W.; Wu, M. Learning your heart actions from pulse: ECG waveform reconstruction from PPG. *IEEE Internet Things J.* **2021**, *8*, 16734–16748. [[CrossRef](#)]
- Tian, X.; Zhu, Q.; Li, Y.; Wu, M. Cross-domain joint dictionary learning for ECG reconstruction from PPG. In Proceedings of the 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Barcelona, Spain, 4–8 May 2020.
- Li, Y.; Tian, X.; Zhu, Q.; Wu, M. Inferring ECG from PPG for Continuous Cardiac Monitoring Using Lightweight Neural Network. *arXiv* **2012**, arXiv:201204949.
- Tang, Q.; Chen, Z.; Guo, Y.; Liang, Y.; Ward, R.; Menon, C.; Elgendi, M. Robust reconstruction of electrocardiogram using photoplethysmography: A subject-based Model. *Front. Physiol.* **2022**, *13*, 859763. [[CrossRef](#)]
- Tang, Q.; Chen, Z.; Ward, R.; Menon, C.; Elfendi, M. PPG2ECGps: An End-to-End Subject-Specific Deep Neural Network Model for Electrocardiogram Reconstruction from Photoplethysmography Signals without Pulse Arrival Time Adjustments. *Bioengineering* **2023**, *10*, 630. [[CrossRef](#)]
- Vo, K.; Naeini, E.K.; Naderi, A.; Jilani, D.; Rahmani, A.M.; Dutt, N.; Cao, H. P2E-WGAN: ECG waveform synthesis from PPG with conditional wasserstein generative adversarial networks. In Proceedings of the 36th Annual ACM Symposium on Applied Computing, Virtual Event, 22–26 March 2021; pp. 1030–1036.
- Sarkar, P.; Etemad, A. CardioGAN: Attentive Generative Adversarial Network with Dual Discriminators for Synthesis of ECG from PPG. In Proceedings of the AAAI Conference on Artificial Intelligence, Delhi, India, 2–9 February 2021; Volume 35, pp. 488–496.
- Omer, O.A.; Salah, M.; Hassan, A.M.; Mubarak, A.S. Beat-by-Beat ECG Monitoring from Photoplethysmography Based on Scattering Wavelet Transform. *Trait. Signal* **2022**, *39*, 1483–1488. [[CrossRef](#)]
- Abdelgaber, K.M.; Salah, M.; Omer, O.A.; Farghal, A.E.A.; Mubarak, A.S. Subject-Independent per Beat PPG to Single-Lead ECG Mapping. *Information* **2023**, *14*, 377. [[CrossRef](#)]

18. Johnson, A.E.W.; Pollard, T.J.; Shen, L.; Lehman, L.H.; Feng, M.; Ghassemi, M.; Moody, B.; Szolovits, P.; Celi, L.A.; Mark, R.G. MIMIC-III, a freely accessible critical care database. *Sci. Data* **2016**, *3*, 160035. [[CrossRef](#)] [[PubMed](#)]
19. Pan, J.; Tompkins, W.J. A Real-Time QRS Detection Algorithm. *IEEE Trans. Biomed. Eng. BME* **1985**, *32*, 230–236. [[CrossRef](#)]
20. Elgendi, M.; Norton, I.; Brearley, M.; Abbott, D.; Schuurmans, D. Systolic Peak Detection in Acceleration Photoplethysmograms Measured from Emergency Responders in Tropical Conditions. *PLoS ONE* **2013**, *8*, e76585. [[CrossRef](#)]
21. Olaf, R.; Philipp, F.; Thomas, B. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*; Springer: Berlin/Heidelberg, Germany, 2015; pp. 234–241.
22. Sakib, M.A.M.; Sharif, O.; Hoque, M.M. Offline Bengali Handwritten Sentence Recognition Using BiLSTM and CTC Networks. In *Proceedings of the Internet of Things and Connected Technologies*, Patna, India, 3–5 July 2020.
23. Wang, Q.; Feng, C.; Xu, Y.; Zhong, H.; Sheng, V.S. A Novel PrivacyPreserving Speech Recognition Framework Using Bidirectional LSTM. *J. Cloud Comput.* **2020**, *9*, 36. [[CrossRef](#)]
24. Zhu, F.; Ye, F.; Fu, Y.; Liu, Q.; Shen, B. Electrocardiogram Generation with a Bidirectional LSTM-CNN Generative Adversarial Network. *Sci. Rep.* **2019**, *9*, 6734. [[CrossRef](#)]
25. Siami-Namini, S.; Tavakoli, N.; Namin, A.S. The Performance of LSTM and BiLSTM in Forecasting Time Series. In *Proceedings of the 2019 IEEE International Conference on Big Data (Big Data)*, Los Angeles, CA, USA, 9–12 December 2019. [[CrossRef](#)]
26. Stoller, D.; Ewert, S.; Dixon, S. Wave-u-net: A multi-scale neural network for end-to-end audio source separation. *arXiv* **2018**, arXiv:1806.03185.
27. Bühlmann, P.; Van de Geer, S. *Statistics for High-Dimensional Data: Methods, Theory and Applications*; Springer: Berlin/Heidelberg, Germany, 2011.
28. Liu, J.; Tang, W.; Chen, G.; Lu, Y.; Feng, C. Correlation and agreement: Overview and clarification of competing concepts and measures. *Shanghai Arch. Psychiatry* **2016**, *28*, 115–120.
29. Alt, H.; Godau, M. Computing the Fréchet distance between two polygonal curves. *Int. J. Comput. Geom. Appl.* **1995**, *5*, 7591. [[CrossRef](#)]
30. Karlen, W.; Raman, S.; Ansermino, J.M.; Dumont, G.A. Multiparameter respiratory rate estimation from the photoplethysmogram. *IEEE Trans. Biomed. Eng.* **2013**, *60*, 1946–1953. [[CrossRef](#)] [[PubMed](#)]
31. Saeed, M.; Villarroel, M.; Reisner, A.T.; Clifford, G.; Lehman, L.; Moody, G.; Heldt, T.; Kyaw, T.H.; Moody, B.; Mark, R.G. Multiparameter Intelligent Monitoring in Intensive Care II: A public-access intensive care unit database. *Crit. Care Med.* **2011**, *39*, 952–960. [[CrossRef](#)] [[PubMed](#)]
32. Pimentel, M.A.; Johnson, A.E.; Charlton, P.H.; Birrenkott, D.; Watkinson, P.J.; Tarassenko, L.; Clifton, D.A. Toward a robust estimation of respiratory rate from pulse oximeters. *IEEE Trans. Biomed. Eng.* **2016**, *64*, 1914–1923. [[CrossRef](#)]
33. Reiss, A.; Indlekofer, I.; Schmidt, P.; Van Laerhoven, K. Deep PPG: Large-scale heart rate estimation with convolutional neural networks. *Sensors* **2019**, *19*, 3079. [[CrossRef](#)]
34. Schmidt, P.; Reiss, A.; Duerichen, R.; Marberger, C.; Van Laerhoven, K. Introducing wesad, a multimodal dataset for wearable stress and affect detection. In *Proceedings of the International Conference on Multimodal Interaction*, Boulder, CO, USA, 16–20 October 2018; pp. 400–408.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.