Convert VR Headsets to Healthcare Devices
Exploring Innovations in Biosignal Monitoring

Yi Qu*
yi.qu@temple.edu
Temple University
Philadelphia, Pennsylvania, USA

Abstract

Photoplethysmography (PPG) utilizes optical signals to measure the blood flow, which can be viewed as an indicator of heart physiology. Electrocardiography (ECG) measures the bioelectric signals from the human heart. Previous studies have shown a strong correlation between PPG and ECG signals. Modern commercial electronic devices have adapted PPG sensors. However, for VR devices, due to their installation limitations, the adaptation of PPG sensors requires special engineering design. This project focuses on the reconstruction from VR accelerometer data to ECG. I borrowed the idea of self-attention blocks from the Transformer and applied it to Generative Adversarial Networks. This project serves as the feasibility study of my future research.

Keywords: PPG, ECG, Accelerometer, VR Headsets, Transformer, GANs

1 Introduction

Clinically, the electrocardiogram (ECG) is considered the preferred method for monitoring vital signs and for the diagnosis, management, and prevention of cardiovascular diseases (CVDs), which are a leading cause of death globally, accounting for approximately 32% of all deaths in 2017 according to Global Burden of Disease reports [1] [12]. Despite its accuracy, the measurement of ECG requires the proper position of ECG sensors (usually 3-12 leads). Misplacement of such sensors results in noisy outputs. The reading of ECG output needs expert knowledge, which prevents ECG from being integrated into mobile devices. Photoplethysmograms (PPG), on the other hand, consist of two components placed on the skin. First, a light source is utilized to reflect light to the skin surface. The red, infrared, or green light can be selected according to the application. Second, a photodetector collects the light reflection [11]. Thanks to their low energy consumption, PPG signals are widely used in commercial wearable smart devices. However, input PPG signals might be distorted due to noises caused by motion artifacts and other environmental sources, which are ubiquitous and unavoidable in everyday life settings [6]. Body hair, skin color, moisture level, or misplacement of the PPG sensors may cause the inaccurate measurement of PPG. Modern VR headsets have encapsulated many sensors, including inner camera, outer camera, accelerometers, gyroscopes, LiDAR, etc. However, the installation of PPG sensors requires additional engineering design since the way people wear VR headsets provides limited space for optical sensors.

In order to solve such problems, people have come up to many innovative solutions. Previous studies have shown a high correlation between PPG and ECG signals [5]. The peaks of the PPG signal in Fig. 1 are in alignment with the ECG signals since the rise of blood volume co-occur with the start of the extraction of human hearts. On the other hand, the sensitivity of the accelerometers in VR headsets guarantees the nominal gain of 0.0015 m/s^2, whereas, the sensitivity of human skin is 0.02 m/s^2. Theoretically, if human hands can feel the pulse, VR accelerometer can do better. Fig. 2 shows the correspondence between the accelerometer and PPG. In this project, I would like to study the deep neural network models that learn from VR accelerometer signals and reconstruct ECG signals.

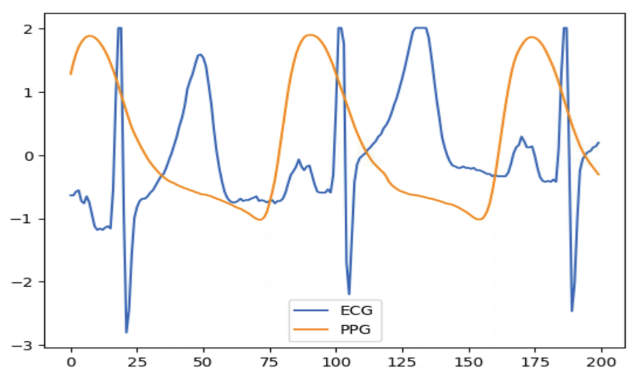

Figure 1. The Correspondence between PPG and ECG

2 Related Works

To leverage the convenience of PPG and the accuracy of ECG, many previous researchers have introduced various novel solutions. Dhanya et al [3]. showed the correlation between the accelerometer readings and the PPG signals. Santos et al. [8] studied the accelerometer-assisted PPG measurements using the LAVIMO sensor system. P. Sarkar et al. proposed CardioGAN [9], which implemented Generative Adversarial Networks on ECG measurements. It incorporated two discriminators targeting both time domain and

*Artificial Intelligence Fall 2024 Project Report

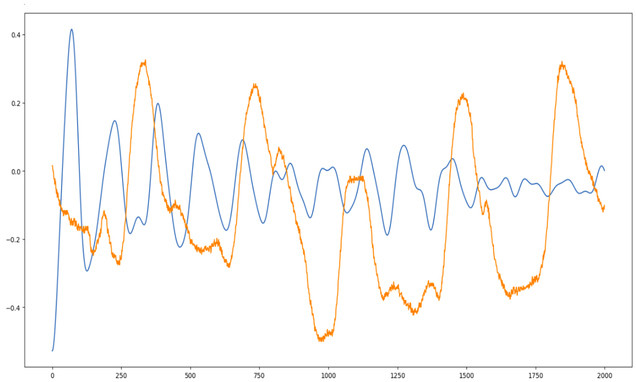

Figure 2. The Correspondence between Accelerometer and PPG

frequency domain information. The forward mapping learns features from PPG and generates ECG. The inverse mapping summarizes ECG and recovers PPG. Knowing that training GANs takes more computational power than training other deep neural networks, incorporating two discriminators would inevitably slow down the training time. D. Shome et al. applied the diffusion model to the PPG-ECG translation field. They brought solutions to address region-specific intricate details of ECG signals and introduced the Region-Disentangled Diffusion Model (RDDM) [10]. By including an additional ROI noise module in the forward process and an additional denoise module in the reverse process, RDDM successfully captures regional features in ECG data. T. Zhang et al. [13, 14] used bidirectional LSTM with attention layers in order to generate PPG signals from other sensors with waveform. Lan et al. [4] take advantage of the Transformer model to reconstruct ECG from PPG.

3 Methodology

Lan et al. applied the attention mechanism directly from the Transformer model. However, this mechanism requires both PPG and ECG to be present at the same time. During the testing phase, the presence of the ECG signal as the decoder input may suffer the critique of data leakage. My intentional idea was to use self-attention instead of cross-attention. In order to enhance the features in the frequency domain, the Fourier Transform was applied to the raw PPG signal and concatenated as the second channel alongside the PPG signal. Fig. 3 demonstrates the Transformer block. The input for both the encoder and decoder is constructed by the vertical stack of the raw PPG signal and FFT of the PPG signal, of which both the real part and the imaginary part are included.

Inspired by the previous results from my college, Ye et al. [15] was the first time introducing the generative model

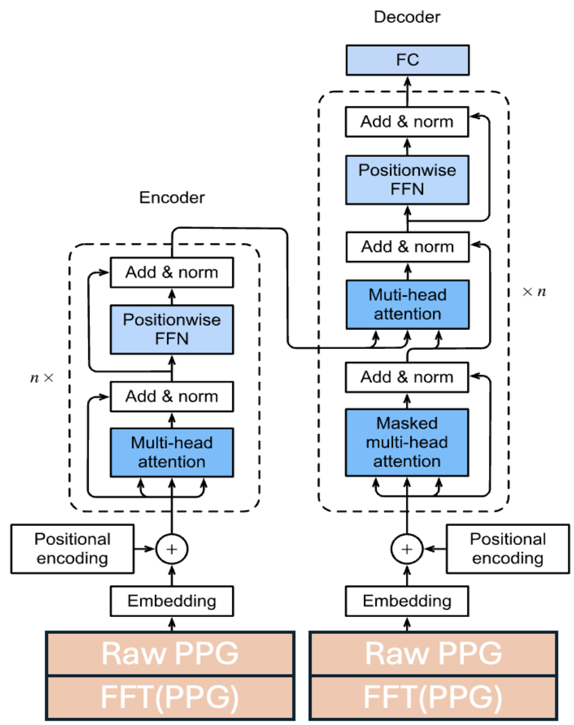

Figure 3. The Transformer Block with Raw PPG and Transformed PPG

in the recovery of vital signs. In their approach, they implemented the Variational Auto Encoder and used the accelerometer data from VR headsets as the input. My improved idea was to use another generative model, which was GANs. Chen et al. [2] used UNet in their GANs generator. The objective of this research is to reconstruct cardiac auscultation, an audio signal generated from heartbeats with a frequency of less than 500 MHz. To further improve the performance, I also took advantage of the self-attention mechanism and applied it at the end of each down-sample and up-sample block in the traditional UNet. The discriminator used the traditional CNN. Fig. 4 illustrates the generator structure.

During training, the discriminator loss function is the binary cross-entropy loss. The generator loss function contains the traditional MSE loss combined with peak alignment loss. The MSE loss can be defined as:

$$LOSS_{mse} = \frac{1}{N}\sum_{i=1}^{N}(y_{pred}^{(i)} - y^{(i)}) \tag{1}$$

Noticing that the model may misalign the peaks in the prediction. I also include a peak alignment loss, which is defined as:

$$LOSS_{peak} = \frac{1}{N}\sum_{i=1}^{N}(\frac{dy_{pred}^{(i)}}{di} - \frac{dy^{(i)}}{di}) \tag{2}$$

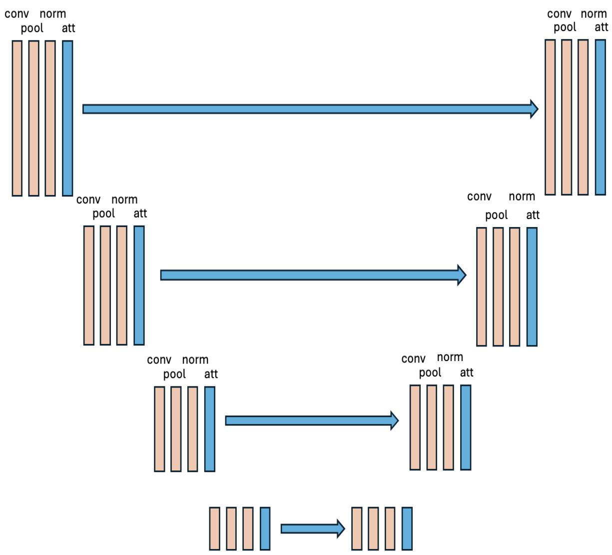

Figure 4. The Generator of GANs

The overall loss is the sum of the MSE loss and the alignment loss, where α is the weight to balance those two losses.

$$LOSS_{pred} = \alpha * LOSS_{mse} + (1 - \alpha) * LOSS_{peak} \quad (3)$$

4 Experiments

4.1 Dataset and EDA

The dataset I used for the accelerometer to PPG was collected in my research group, which consists of 5 people. Each of us wore a VR headset for 25 min and collected both accelerometer data and the PPG data. The PPG data was collected from a PPG sensor connected to an Arduino Uno developer board. Each experiment was repeated for 4 times. Fig. 5 demonstrates the PPG sensor connected to the Arduino. Before feeding into the model, the accelerometer data was preprocessed to filter out the high frequency noise using a Butterworth bandpass filter. The cutting frequency was set to be 40Hz. Then the accelerometer signal was smoothed by a convolution of a constant signal of 1.

The dataset I used for PPG to ECG is the BIDMC PPG and Respiration Dataset [7]. This dataset was collected from 53 patients from clinical records. Each patient collected 8 minutes of respiration, PPG, ECG lead V, II, and AVR. The sample rate was 125Hz. One of the data samples is illustrated in Fig. 6.

The overall dataset was formed by randomly selecting 60000 samples out off the entire dataset. The training set was built by 80% of 60000 samples, which was 48000 samples. The test set took the rest 20%. Before feeding into the model, I performed the min-max scale to make both PPG and ECG in the range [-1, 1]. The dataset was split into continuous windows of 160 samples since the heart rate was around 1Hz. Setting the window size to 160 guaranteed at least 1 beat in each window.

4.2 Hyperparameters

The Transformer model has an input dimension of 3, of which the first is the raw PPG, the second is the real part of the Fourier Transform of PPG, and the third is the imaginary part of the Fourier Transform of PPG. In order to capture the inner-related feature, I also add two CNN layers before feeding the signal into the Transformer. The embedding dimension is 160. The number of heads is 16. The number of encoder layers is 16.

The GANs model generator has 4 down-sample and up-sample blocks. Each block is then followed by a self-attention layer. The discriminator has 4 consecutive convolutional layers.

Both Transformer and GANs are trained for 100 epochs. The learning rate is set to 10e-4 for both discriminator and generator. The optimizer is Adam with weight decay to be 0.1.

5 Results Analysis

Fig. 7 shows one of the reconstructed PPG from the accelerometer signals using GANs. Fig. 8 shows one of the reconstructed ECG from PPG data using Transformer. Fig. 9 shows one of the reconstructed ECG from PPG using GANs. The MSE of Transformer model was 0.0462. The MSE of GANs in ACC2PPG was 0.0358. The MSE of GANs in PPG2ECG was 0.0284. From the images, we could tell the difference between those two models. The Transformer model gave more smooth but inaccurate predictions. Whereas, the GANs model generated more accurate but fuzzy predictions. By changing the weight between $LOSS_{mse}$ and $LOSS_{peak}$, the GANs model would generate more smooth signals with high $LOSS_{mse}$ weight. Table 1 shows the relationship between the number of heads in Transformer, the number of encoder layers in Transformer, and the MSE values.

Table 1. The impact of the number of heads and number of encoder layers in Transformer

number of heads	number of encoder layers	MSE
4	4	0.0528
8	8	0.0499
16	16	0.0462

6 Conclusion

As a feasibility study, using accelerometer signals to reconstruct ECG signal is achivable. Both the Transformer and the GANs models can work as the PPG to ECG translator. However, the Transformer model suffers from the reconstructive accuracy. Whereas, the GANs model performs better in terms

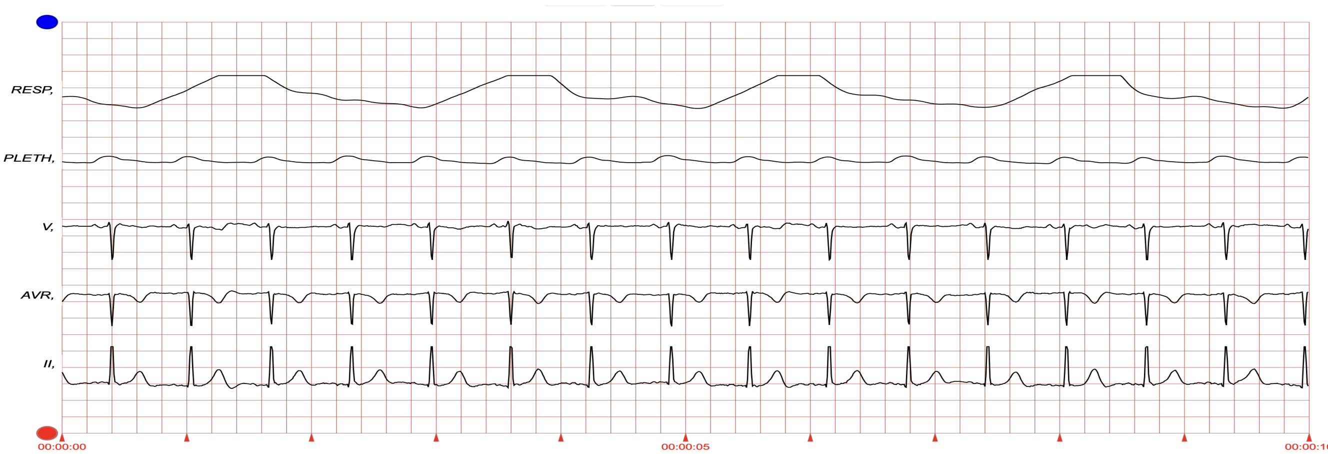

Figure 5. The BIDMC Data

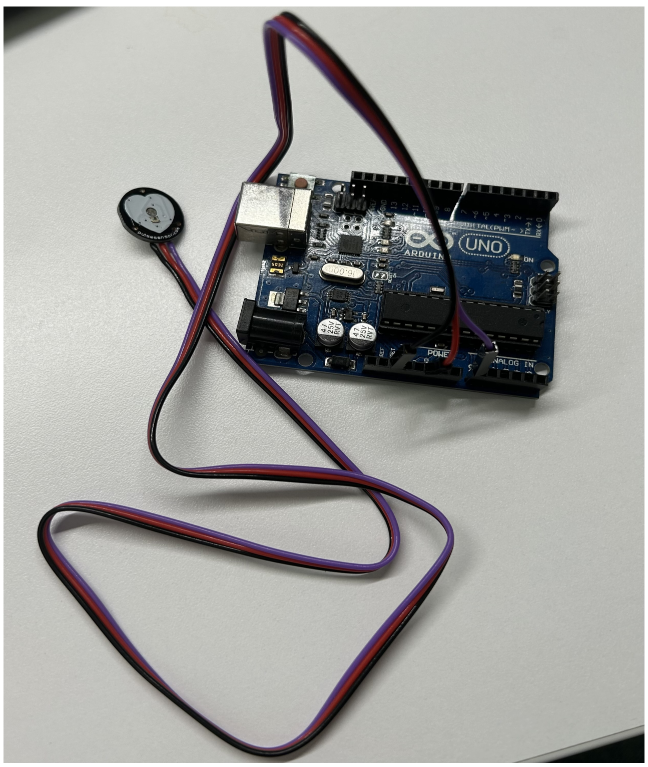

Figure 6. The PPG sensor connected with an Arduino Uno board

of accuracy, but the fuzziness of the output cannot be used in clinical conditions. Instead, the time interval between the reconstructive peaks in the GANs model could be used as a clinical indicator.

References

[1] John Allen. 2007. Photoplethysmography and its application in clinical physiological measurement. *Physiological measurement* 28, 3 (2007), R1.

[2] Tao Chen, Yongjie Yang, Xiaoran Fan, Xiuzhen Guo, Jie Xiong, and Longfei Shangguan. 2024. Exploring the Feasibility of Remote Cardiac Auscultation Using Earphones. In *Proceedings of the 30th Annual International Conference on Mobile Computing and Networking*. 357–372.

[3] V Dhanya and N Kayalvizhi. 2020. Correlation of PPG and Accelerometer Signals. In *2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA)*. 53–57. https://doi.org/10.1109/ICECA49313.2020.9297504

[4] Ella Lan. 2023. Performer: A novel ppg-to-ecg reconstruction transformer for a digital biomarker of cardiovascular disease detection. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. 1991–1999.

[5] Jaewon Lee, Jinwoo Kim, and Miyoung Shin. 2017. Correlation analysis between electrocardiography (ECG) and photoplethysmogram (PPG) data for driver's drowsiness detection using noise replacement method. *Procedia computer science* 116 (2017), 421–426.

[6] Mengmeng Ma, Jian Ren, Long Zhao, Sergey Tulyakov, Cathy Wu, and Xi Peng. 2021. Smil: Multimodal learning with severely missing modality. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 35. 2302–2310.

[7] Marco AF Pimentel, Alistair EW Johnson, Peter H Charlton, Drew Birrenkott, Peter J Watkinson, Lionel Tarassenko, and David A Clifton. 2016. Toward a robust estimation of respiratory rate from pulse oximeters. *IEEE Transactions on Biomedical Engineering* 64, 8 (2016), 1914–1923.

[8] Susana Aguiar Santos, Boudewijn Venema, and Steffen Leonhardt. 2012. Accelerometer-assisted PPG measurement during physical exercise using the LAVIMO sensor system. *Acta Polytechnica* 52, 5 (2012).

[9] Pritam Sarkar and Ali Etemad. 2021. Cardiogan: Attentive generative adversarial network with dual discriminators for synthesis of ecg from ppg. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 35. 488–496.

[10] Debaditya Shome, Pritam Sarkar, and Ali Etemad. 2024. Region-Disentangled Diffusion Model for High-Fidelity PPG-to-ECG Translation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 38. 15009–15019.

[11] Toshiyo Tamura, Yuka Maeda, Masaki Sekine, and Masaki Yoshida. 2014. Wearable photoplethysmographic sensors—past and present. *Electronics* 3, 2 (2014), 282–302.

[12] Khuong Vo, Mostafa El-Khamy, and Yoojin Choi. 2023. Ppg to ecg signal translation for continuous atrial fibrillation detection via attention-based deep state-space modeling. *arXiv preprint arXiv:2309.15375*

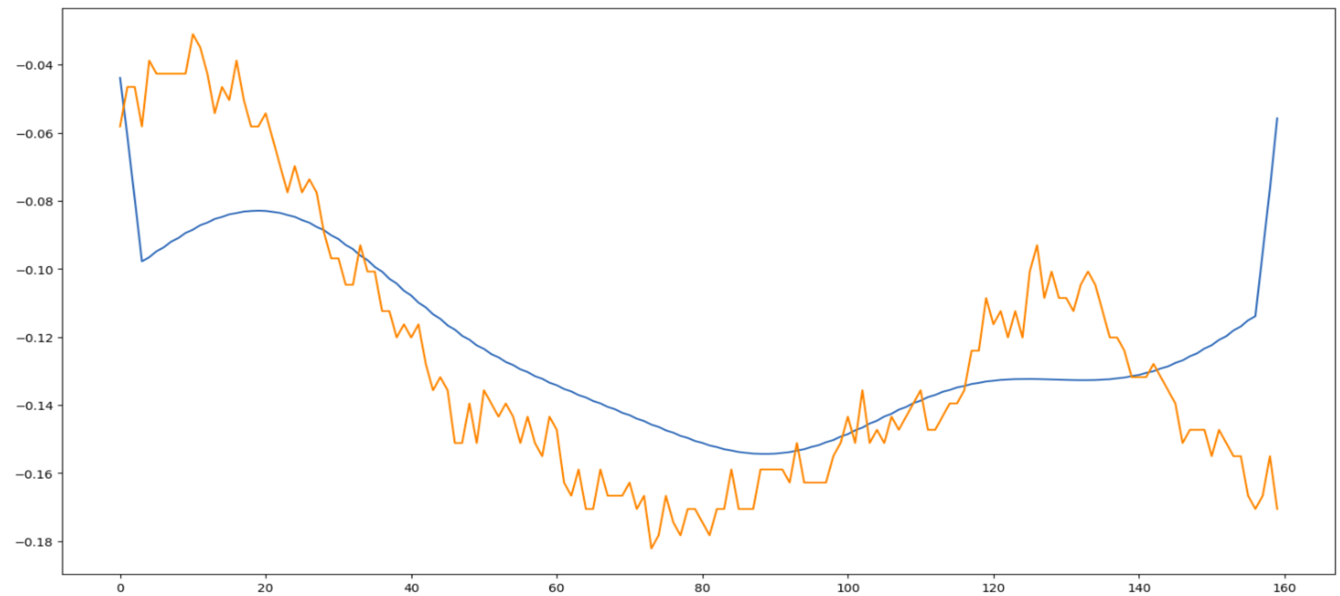

Figure 7. One of the Reconstructed PPG from accelerometer data using GNAs. Blue: prediction, Orange: ground truth

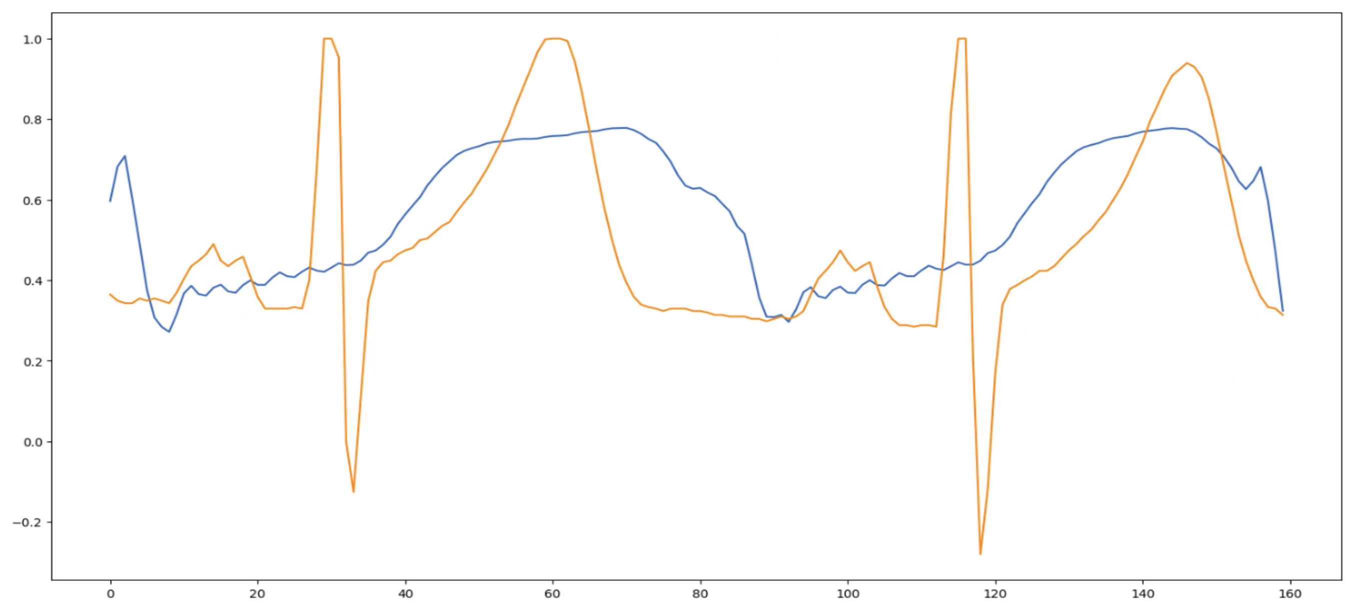

Figure 8. One of the Reconstructed ECG from PPG data using Transformer. Blue: prediction, Orange: ground truth

(2023).

[13] Tianfang Zhang, Cong Shi, Payton Walker, Zhengkun Ye, Yan Wang, Nitesh Saxena, and Yingying Chen. 2023. Passive Vital Sign Monitoring via Facial Vibrations Leveraging AR/VR Headsets. In *Proceedings of the 21st Annual International Conference on Mobile Systems, Applications and Services*. 96–109.

[14] Tianfang Zhang, Zhengkun Ye, Ahmed Tanvir Mahdad, Md Mojibur Rahman Redoy Akanda, Cong Shi, Yan Wang, Nitesh Saxena, and Yingying Chen. 2023. FaceReader: Unobtrusively Mining Vital Signs and Vital Sign Embedded Sensitive Info via AR/VR Motion Sensors. In *Proceedings of the 2023 ACM SIGSAC Conference on Computer and Communications Security*. 446–459.

[15] Tianming Zhao, Zhengkun Ye, Tianfang Zhang, Cong Shi, Ahmed Tanvir Mahdad, Yan Wang, Yingying Chen, and Nitesh Saxena. 2022. Continuous blood pressure monitoring using low-cost motion sensors on AR/VR headsets. In *Proceedings of the 20th Annual International Conference on Mobile Systems, Applications and Services*. 589–590.

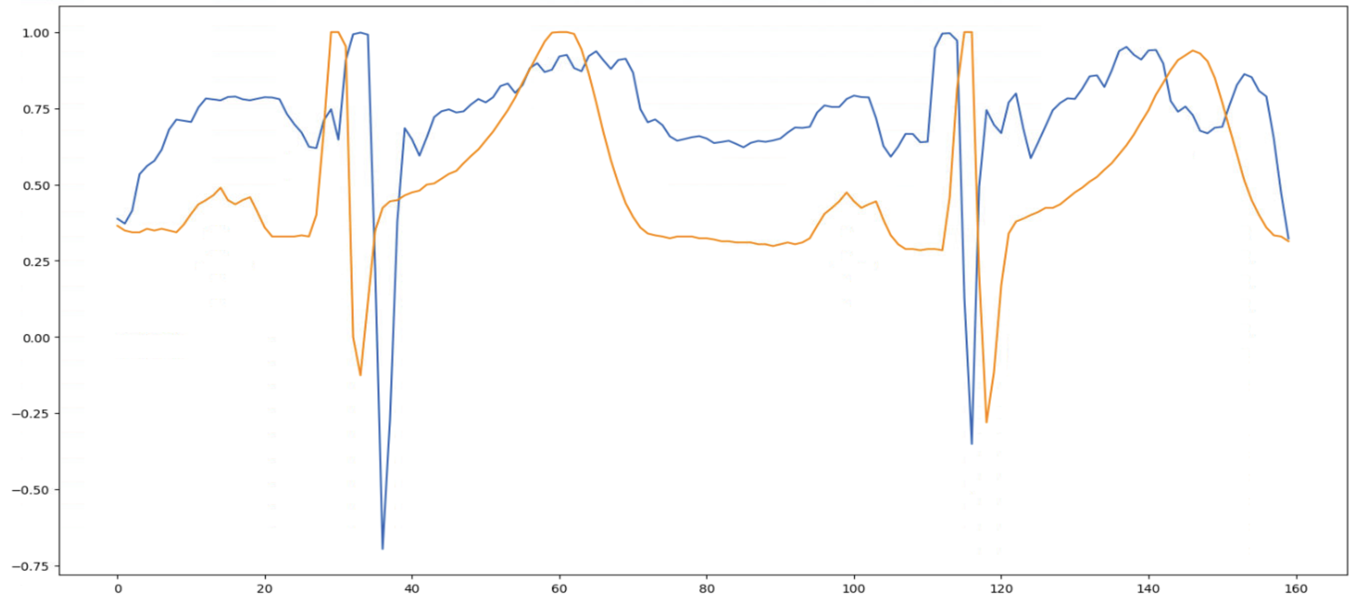

Figure 9. One of the Reconstructed ECG from PPG data using GNAs. Blue: prediction, Orange: ground truth