2023 NYCU DLP Final Project – Group 11

Reconstructing ECG from Paired PPG using Dual Discriminator SEGAN with Attention Mechanism

國立陽明交通大學
NATIONAL YANG MING CHIAO TUNG UNIVERSITY

1.Abstract

Electrocardiogram (ECG) is the electrical measurement of cardiac activity, whereas Photoplethysmography (PPG) is the optical measurement of volumetric changes in blood circulation. While both are used for heart rate monitoring, ECG is more useful from a medical point of view as it carries additional cardiac information. Unfortunately, ECG sensors are not as readily available as those of PPG. To solve this problem, using GANs to generate ECG from PPG signals was proposed. In our result the person correlation coefficient of the generated signal from PPG is up to 0.855.
.

Keywords: ECG, PPG, GAN, Reconstruction

2.Introduction

Nowadays, healthcare has become a major focus. According to the World Health Organization (WHO), approximately 17.9 million people died from cardiovascular diseases (CVDs) in 2019, accounting for 32% of global deaths.[1] However, measuring ECG signals in a fast and simple way remains a challenge. The high cost and complexity of ECG devices, along with the need for trained professionals and multiple electrode leads, make remote healthcare difficult. On the other hand, devices like smartphones, smartwatches, and pulse oximeters offer easier access to data at lower prices, making them viable alternatives.

Generative Adversarial Networks (GAN) use a generator and a discriminator to produce realistic data samples. Previous studies using techniques like SEGAN and Transformer have certain limitations. For instance, some studies only focus on the QRS complex of the ECG waveform, overlooking other important components like PQRST waves associated with different diseases. Additionally, the training requirements of certain models, such as powerful servers, make them impractical for remote healthcare. To address these limitations, we propose a method that incorporates SEGAN as a bottleneck, dual discriminators, dilation, and attention mechanism to enhance the reconstruction process.

3.Dataset

The BIDMC dataset[4] is a collection of physiological signals and numbers. The dataset was collected from critically-ill patients during hospital care at the Beth Israel Deaconess Medical Centre (BIDMC) in Boston, MA, USA. The dataset includes 53 recordings of 8-minute duration, each containing:

1. Physiological signals: ECG, PPG, and impedance respiratory signal. These are sampled at 125 Hz.
2. Physiological parameters: heart rate (HR), respiratory rate (RR), and blood oxygen saturation level (SpO2). These are sampled at 1 Hz.
3. Fixed parameters: age and gender.
4. Manual annotations of breaths.

The BIDMC dataset is a valuable resource for researchers working in the field of biomedical engineering. In this experiment, the BIDMC dataset is primarily used for training and testing the effectiveness of reconstructing ECG signals from PPG data, and conducting overall analysis of the results.

4.Methodology

4.1 Data Preprocess

Firstly, we use a Chebyshev Type II bandpass filter with a frequency range from 0.05 to 15Hz to filter out noise from the PPG and ECG signals. This filtering process is necessary because signals below 15Hz result in the disappearance of PPG systolic peak features, and ECG signals lose the complete waveform of the PQRST waves.

4.2 GAN and Dual Discriminator

In this GAN model, we adopt the SEGAN architecture to enhance speech signals [2]. Training the generator produces clearer and more intelligible speech signals, benefiting applications like speech recognition and synthesis. We extend this model to generate high-quality ECG signals, introducing a latent vector between the encoder and decoder, obtained by downsampling the original signal. The encoder incorporates dilation mechanisms to capture temporal relationships in PPG or ECG signals. To improve signal quality, we employ dual discriminators. The first compares the ground truth ECG with the reconstructed ECG from paired PPG, while the second uses FFT to align frequencies with the ground truth [2]. Results demonstrate improved reconstruction accuracy. This approach has potential for camera-based rPPG applications, but challenges remain, such as individual fine-tuning and waveform alignment [2].

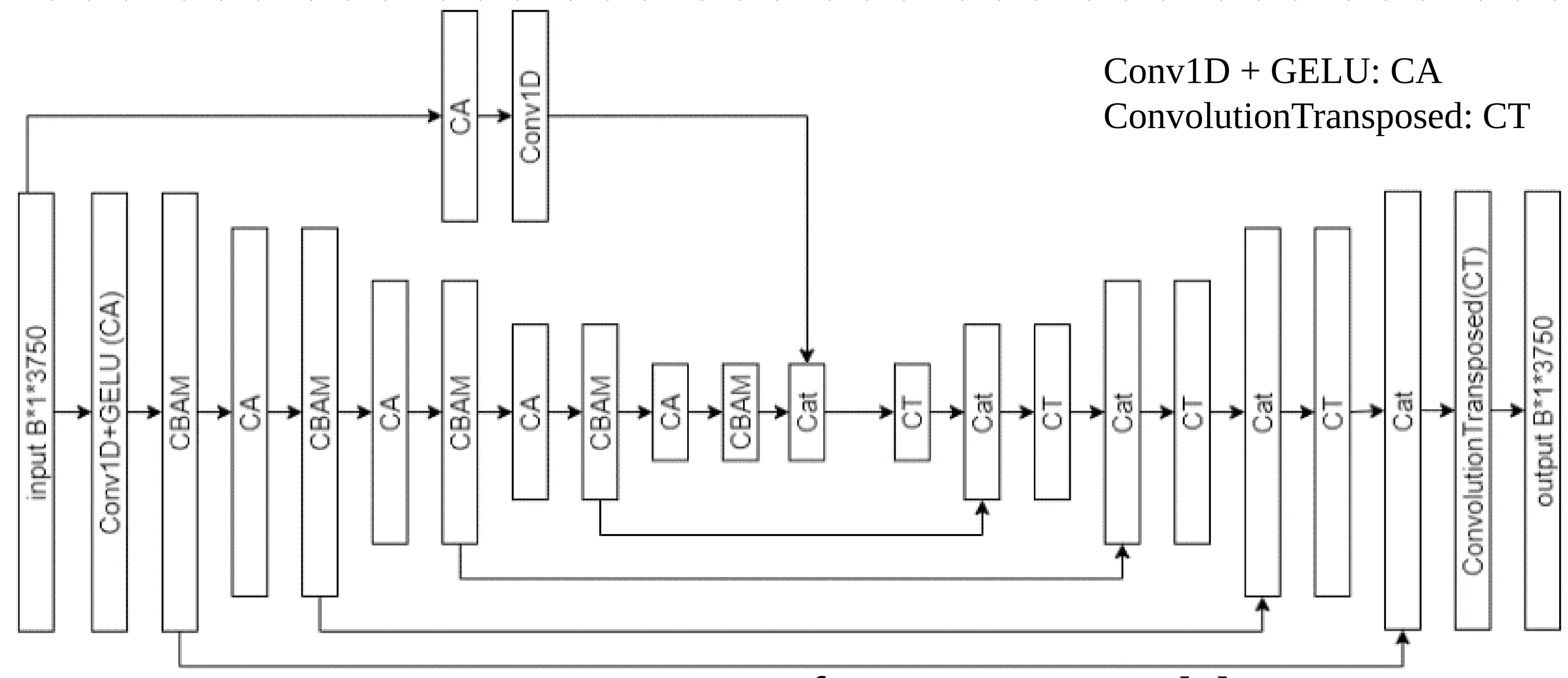


Figure 1. Structure of Generative Model

5.Result

5.1 Data Preprocess

Overlap	RMSE	MS	MAE	Correlation
/	0.3760	0.1551	0.3017	0.4584
step = 5s	0.2173	0.0529	0.1694	0.8257
step = 1s	**0.2046**	**0.0503**	**0.1605**	**0.8551**

Table 1. Data augmentation comparison

In this experiment, due to the limited number of subjects, the total number of samples available was approximately 700, which resulted in a lower overall accuracy. To address this limitation, we performed data augmentation using overlapped data with step lengths of 5 seconds and 1 second, while the testing was conducted using non-overlapping data.

5.2 Dilation

Method	RMSE	MSE	MAE	Correlation
w/o dilation	0.2470	0.0696	0.1964	0.7914
dilation = 2	**0.2046**	**0.0503**	**0.1605**	**0.8551**

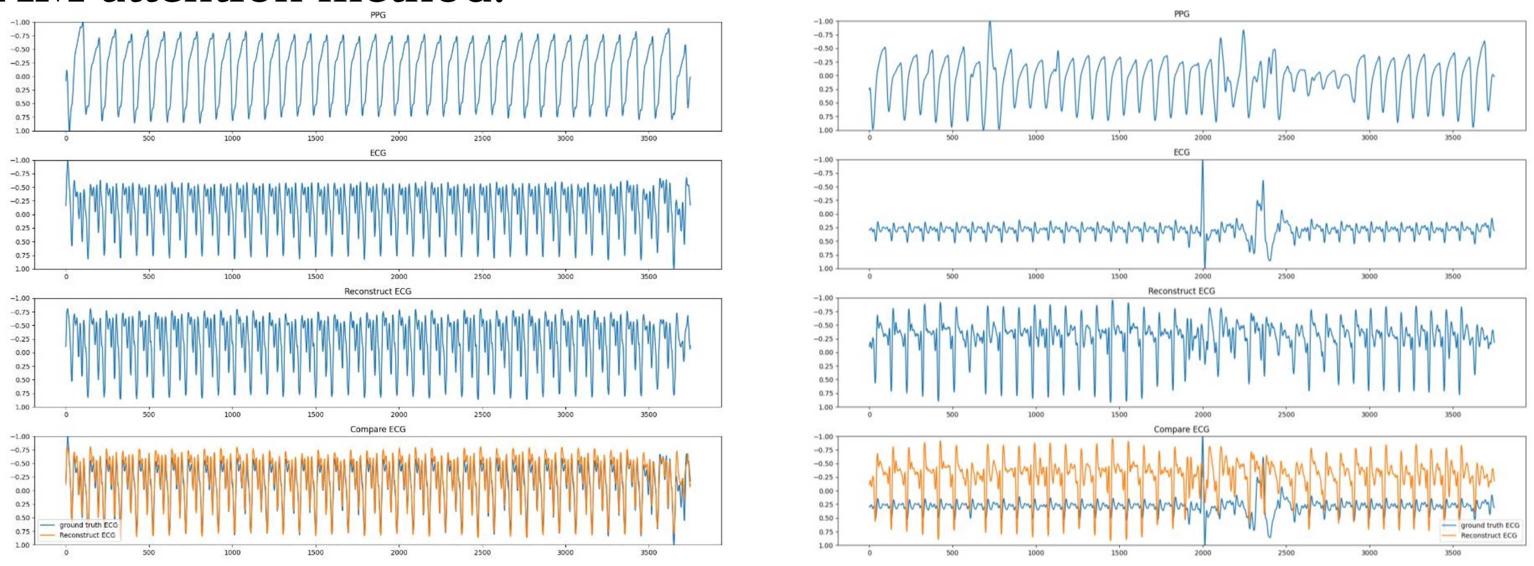
Table 2. Dilation comparison

5.3 Different Attention

Attention	RMSE	MS	MAE	Correlation
SE-Net	0.2896	0.0983	0.2297	0.6650
SK-Net	0.2838	0.0903	0.2256	0.6776
CBAM	**0.2046**	**0.0503**	**0.1605**	**0.8551**

Table 3. Attention comparison

We compare three attention methods, namely SE-Net, SK-Net, and CBAM, in our model. SE-Net captures information across different channels by assigning higher weights to important channels and lower weights to less informative ones. SK-Net uses different kernel sizes to focus on features at different scales, utilizing smaller kernels for finer details and larger kernels for global features. CBAM is an improved version of SE-Net that incorporates both channel and spatial attention. In Table 3, we observe significant performance improvement when using the CBAM attention method.



(a) (b)

Figure 2. Example of Reconstruction ECG

6.Conclusion

Figure 2(a) shows excellent overall reconstructive performance, with a PCC value of approximately 0.976, indicating a close resemblance to the original signal. This approach proves highly effective for signal reconstruction. In future applications, camera-based rPPG could potentially capture facial blood volume changes, eliminating the need for specialized devices. However, limitations include the need for patient-specific data for fine-tuning the model and challenges in aligning different baseline waveforms due to variations caused by symptoms or noise. Future research should address these limitations and seek solutions in this direction.