A Generative Methodology for PPG to ECG Reconstruction Based on Dual-Critic Approach

Rashmi Kumari[*,▲], Prateek Agrawal[+,▲], Spandan More[*], Nikhil Praveen[*], Surita Sarkar[*,▲], Pabitra Das[*,▲], Amit Acharyya[*,▲]

[*] Indian Institute of Technology, Hyderabad
[▲] Sense Health Technologies Private Limited
[+] Indian Institute of Information Technology, Dharwad

Abstract

Cardiovascular diseases (CVDs) are the leading cause of mortality worldwide. While ECG provides valuable cardiac information, its multi-lead setup is inconvenient for long-term monitoring. Despite various trials of ECG measurement in wearable devices, photoplethysmography (PPG) is found to be a more suitable alternative for wearable devices. This study introduces a novel PPG-to-ECG reconstruction using a Dual-Critic GAN, leveraging Wasserstein-Loss and Discrete Wavelet Transform (DWT) for improved ECG reconstruction than other state-of-the-art methods. The model achieves lower errors compared to other methods, yields a root-mean-squared error (RMSE), mean absolute error of heart rate (MAE_{HR}), and percentage root mean square (PRD) difference between the desired and generated signals as 0.134, 3.75, and 6.064, respectively, reducing RMSE by 14.6%, PRD by 27.4%, and MAE_{HR} by 21.4%.

1. Introduction

Cardiovascular diseases (CVDs) are a major global health concern, responsible for around 17.9 million deaths in 2019[1], according to the WHO. Despite advances in healthcare, CVDs often go undiagnosed due to unawareness, lack of routine monitoring, and insufficient healthcare access [2]. Regular cardiac monitoring and early detection can help reduce this issue.

An electrocardiogram (ECG) is crucial for detecting cardiac activities, particularly subtle abnormalities that may require long-term monitoring [3]. However, traditional ECG setups with multiple lead attachments restrict patient mobility. Although Holter monitors offer continuous recording, they are bulky and inconvenient. New wearable devices minimizing size and lead attachments, but many still require a chest patch that can irritate the skin or necessitate finger contact, making them impractical for prolonged ECG monitoring.

On the contrary, photoplethysmography (PPG) is widely used in wearable devices and smartwatches due to its portability and ease of integration. It is an optical way of sensing cardiac conditions, requires only a skin sensor, eliminating the need for leads or patches. Previous studies showed a high correlation between time and frequency domains of heart rate variations (HRV) derived from PPG and ECG [4]. This can be especially useful for athletes or patients with serious cardiac conditions for continuous monitoring. Although PPG is cost-effective and more convenient for continuous and ambulatory monitoring of cardiac activities, it retains some helpful information that may be insufficient to diagnose some specific cardiac diseases. It also alters skin tone and skin type and is sensitive to motion artifacts, which would prevent us from getting better information about cardiac activities.

Recently, A few techniques have been developed for extracting ECG from PPG signals. A study by Ho et al. used a banded kernel ensemble learning technique to generate ECG from PPG and use it to predict heart diseases [5]. However, implementing an ensemble makes the model output hard to predict and explain. Furthermore, any incorrect selection causes reduced predictive accuracy than an individual model [6]. Attention block was employed in this study to extract salient ECG features, including QRS complex from ECG. However, the architecture failed to obtain P and T waves, which are equally essential to identify CVD detection [7] accurately. Zhu et al. [8] proposed a bidirectional long short-term memory (LSTM)- convolutional neural network (CNN) architecture to generate ECG from Gaussian noise. They employed a novel loss function in their generative adversarial network (GAN) model to generate synthetic patient-specific ECG data from input noise. In their paper, Sarkar et al. [9] have reconstructed ECG waveform from PPG using the CycleGAN-based method, combined with attention units which incorporated dual generator and dual discriminator to analyze

 ISSN: 2325-887X DOI: 10.22489/CinC.2024.486

the signal in both time and frequency domains. However, the technique failed to generate an ECG waveform correctly for noisy PPG inputs. Many of these studies have suffered from vanishing gradient problems, affecting the reconstructed ECG's morphology.

In this paper, we introduced a dual-critic GAN-based architecture utilizing Wasserstein loss to address the limitations of existing works in ECG reconstruction from PPG. The overall flow of our model is illustrated in Fig. 1. Furthermore, incorporating LSTM in the generator section helped to eliminate the vanishing gradient problem that occurs during the training of the GAN network [10]. The enhanced learning of the generator leads to better-reconstructed ECG signals. The primary contributions of this proposed work are as follows:

- We proposed a novel dual-critic GAN-based deep neural network that employs discrete wavelet transform in the frequency domain for one critic, enabling ECG inference from a single-channel PPG input.
- The implementation of a Multi-Head Attention Block in the generator allows for the adjustment of weights and focuses on salient signal features separately.
- We utilized Wasserstein loss and gradient penalty alongside dual discriminators as critics of the generator to address the vanishing gradient problem and minimize the generator's loss.

2. METHODOLOGY

2.1. Discrete Wavelet Transform

The Discrete Wavelet Transform (DWT) has grown in popularity in biomedical signal processing. A window or wavelet function is first chosen to perform a DWT on a signal. The signal is decomposed into different frequency bands and time intervals by convolving the wavelet function with the signal at different scales and translations. The convolution operation produces a set of coefficients representing the wavelet function's contribution at each scale and translation. The coefficients obtained from the convolution operation are typically too numerous to be practical for analysis, so they are downsampled to reduce their number. This decomposition process is repeated recursively on the approximation coefficients obtained in the previous step. The process is repeated until a desired decomposition level is reached or the coefficients become insignificant. Finally, the signal can be reconstructed by combining the coefficients obtained at different scales and translations. The reconstruction is performed by applying an inverse wavelet transform to the coefficients, which produces an approximation of the original signal.

The Discrete Wavelet Transform (DWT) is an effective signal processing technique for analyzing biomedical signals [11]. It enables multiresolution analysis by decomposing signals into various frequency bands, providing excellent time-frequency localization. This makes DWT ideal for tasks like peak detection in PPG/ECG signals and preserving ECG signal complexes [12]. In our model architecture, we used the WaveTF Python library to compute the DWT of TensorFlow tensors, employing Daubechies and Haar wavelets. During the DWT process, the resulting coefficients were downsampled by a factor of 2.

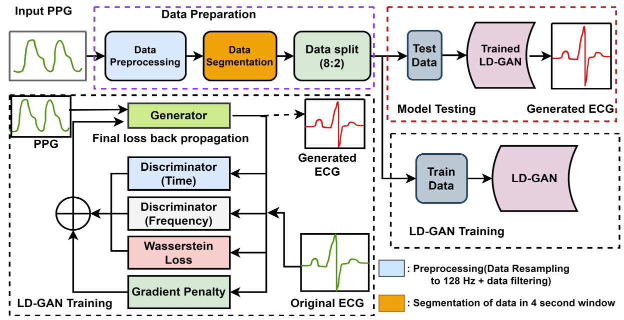

Figure 1. The proposed LDGAN process flow.

2.2. Wasserstein GAN

One common issue with GANs is mode collapse, where the generator produces only a few variations of data due to overfitting, making the output independent of the input. This often results from vanishing gradients when using KL or JS losses. To address this, Khuong et al. [13] introduced the Earth Mover Distance (or the Wasserstein-1 distance) as an alternative to KL Divergence. This changes GAN training by replacing the discriminator with a critic, which grades how close a generated image is to a real one rather than just classifying it as real or fake. With Wasserstein Loss, the model continues learning regardless of its current performance, overcoming issues with KL or JS Divergence.

$$JS(Divrgn) = \underset{D}{min}\,\underset{D\in\mathcal{D}}{max} \underset{x\sim\mathbb{P}_r}{\mathbb{E}}\big[D(x)\big] - \underset{x\sim\mathbb{P}_g}{\mathbb{E}}\big[D(\widetilde{x})\big]$$

2.3. Gradient Penalty

A gradient penalty was added to the loss function to enforce the Lipschitz constraint, improving training stability. By calculating the gradient between real and generated samples, this penalty addressed issues where the original GAN cost function struggled, ensuring better results. An interpolation between the real and fake instances was made, and the gradient was calculated. The aim was to prevent exploding or vanishing gradients, a huge problem for GANs.

$$L = \underset{x\sim\mathbb{P}_r}{\mathbb{E}}[f(x)] - \underset{x\sim\mathbb{P}_g}{\mathbb{E}}[D(\widetilde{x})] + \lambda\mathbb{E}[||\nabla_{\widehat{x}} f(\widehat{x}||_x)^2] \quad (1)$$

Here, λ is the penalty coefficient used to weigh the gradient penalty term, and we use ten here.

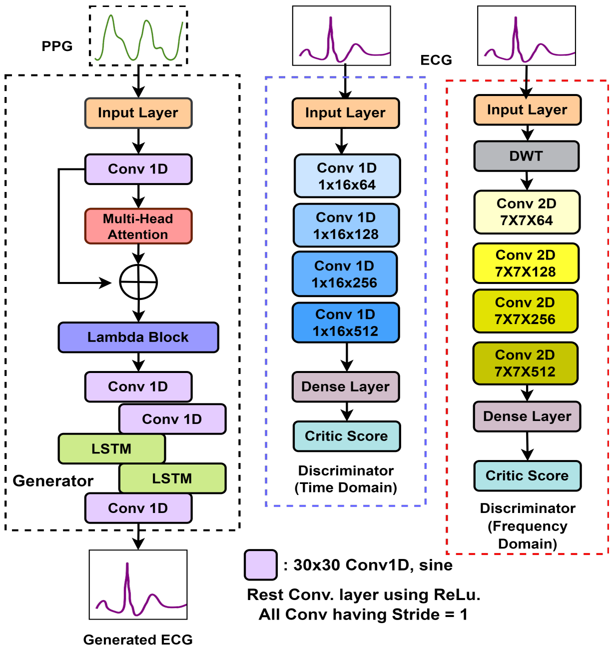

Figure 2. Left to Right: Generator, Discriminator(time domain), Discriminator (frequency domain).

The final loss combines the Wasserstein Loss, Gradient Penalty, and Double Discriminator Loss.

2.4. Wasserstein Loss

$$L = \mathop{\mathbb{E}}_{x\sim\mathbb{P}_r}[f(x)] - \mathop{\mathbb{E}}_{x\sim\mathbb{P}_g}[D(\widetilde{x})] + \lambda\mathbb{E}[||\nabla_{\widehat{x}} f(\widehat{x}||_x - 1)^2] \quad (2)$$

2.5. Final Loss

Final Loss, $L_{final} = \alpha * L_{time} + L_{DWT} - \lambda GP$. where,

$$L_{time} = \mathop{\mathbb{E}}_{x\sim\mathbb{P}_g}[f(x)] - \mathop{\mathbb{E}}_{x\sim\mathbb{P}_g}[D(\widetilde{x})] \quad (3)$$

$$L_{DWT} = \mathop{\mathbb{E}}_{x\sim\mathbb{P}_g}[f(DWT(x))] - \mathop{\mathbb{E}}_{x\sim\mathbb{P}_g}[D((DWT)\widetilde{x})] \quad (4)$$

$$\text{GP}= \mathop{\mathbb{E}}_{x\sim\mathbb{P}_x}[(||\nabla_{\widehat{x}} f(\widehat{x})||_x - 1)^2] + \mathop{\mathbb{E}}_{x\sim\mathbb{P}_x}[(||\nabla_{\widehat{x}} f(\widehat{x}_{DWT})||_x - 1)^2] \quad (5)$$

Here, $\widehat{x} = \epsilon x + (1-\epsilon)\widetilde{x}$, and $\widehat{x}_{DWT} = \epsilon DWT(x) + (1-\epsilon)DWT(\widetilde{x})$

2.6. Model Architecture

The model includes a generator designed to convert PPG to ECG. The generator includes attention and LSTM layers. Two discriminators evaluate the generated ECG: one in the time domain and the other using Discrete Wavelet Transform. The model architecture of the generator and discriminator sections are shown in Fig. 2.

Table 1. Error Results Comparison of Different Model.

Methods	RMSE	MAE (HR)	PRD
CGAN+CNN [15]	0.668	–	–
CardioGAN [9]	0.364	4.77	8.356
BiLSTM-CNN GAN [8]	0.215	–	51.799
CardioGAN(Wasserstein loss) [14]	0.157	–	38.566
Proposed method	**0.134**	**3.75**	**6.064**

3. RESULT AND DISCUSSION

The primary focus of the dataset was the BIDMC dataset, having 53 adult ICU patients. We selected ECG data with lead II(among II, V, AVR) configuration for our study. Both ECG and PPG signals had been split, with 80% of the total dataset being used for training, while the remaining 20% of the dataset is kept as testing. In data pre-processing, data was resampled to 128Hz, passed through Bandpass Butterworth filters to remove noise, and windowed to 4 seconds segments. All training for the models was executed on an Nvidia Tesla V100-PCIE-16GB GPU using Tensorflow 2.0 for 30 epochs and alpha = 10, with the discriminator being trained thrice every time the generator was trained at a learning rate of 1e-5, which linearly decays after epoch 20.

The generated ECG from PPG using the proposed model and ground truth ECG is shown in Fig. 3. The graphical representation shows that our proposed method generates ECGs resembling the original ECG waveform. The model was fed 4-second PPG signal snapshots, and the output was a corresponding 4-second generated ECG signal. A graphical comparison between generated and original ECG measured using our proposed method and CardioGAN [9] work has been shown in Fig. 3.

We used the NVIDIA Jetson AGX Orin board to test our model output. The Orin board, a low-powered, high-performance GPU, significantly accelerates AI processing. As shown in the middle section of Fig. 3, the output from our model and the NVIDIA Jetson AGX Orin board were similar, but comparable in much less processing time. Training took 0.27 seconds per epoch, which involved processing data from 53 patients with 60,000 data points each. For one window of 512 data points, the processing time was calculated to be approximately $\frac{0.27\times512}{53\times60000}$ seconds. Performance was evaluated using metrics like RMSE, MAE for heart rate, and PRD between original and generated ECG signals. Obtained metrics have been compared with existing methods. The RMSE value had been calculated by normalizing the data [14]. The performance evaluation in Table I shows that the proposed work outperforms other existing works.

4. CONCLUSION

In this work, we proposed a dual-critic GAN model using DWT in the critic architecture for precise ECG reconstruction from PPG. Including Wasserstein loss, gra-

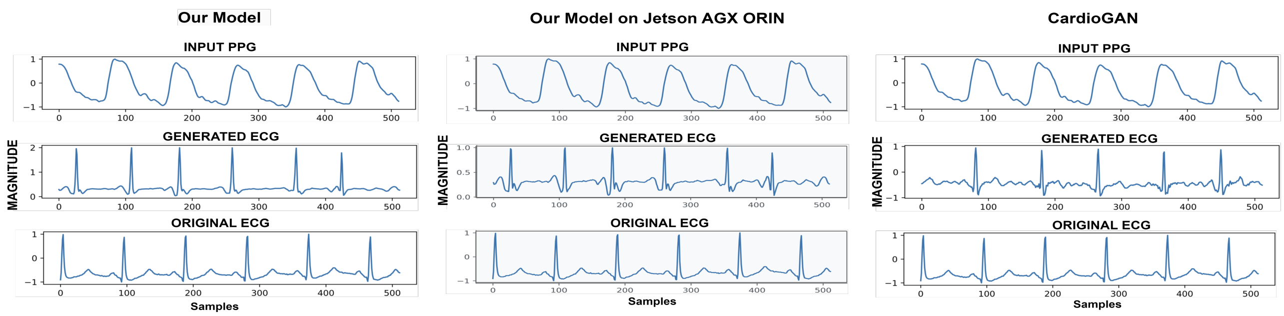

Figure 3. Generated ECG Graphs, Left: Our model, Middle: Our model tested on NVIDIA Jetson Orin board, Right: CardioGAN.

dient penalty, and dual discriminators as critics help prevent the vanishing gradient problem, reduce generator loss, and improve ECG reconstruction accuracy, closely resembling the ground truth. The low values of calculated performance evaluation metrics indicate the effectiveness of the proposed model architecture for accurate ECG reconstruction. Also, we reduces the complexity of the model by eliminating one generator compared to the dual generator-dual discriminator model. Moreover, the reliable conversion process provides unobtrusive and continuous monitoring of ECG from a single PPG input.

Acknowledgments

We acknowledge Sense Health Technologies Private Limited and STARS, Ministry of Human Resource Development (MHRD), Government of India, for their support.

References

[1] WHO, "Cardiovascular diseases (cvds)," vol. https://www.who.int/ news-room/ fact-sheets/ detail/ cardiovascular-diseases-(cvds), 2021.

[2] V. Naresh, P. Patra, P. K. Jha, K. B. Chivukula, C. K. Vala, A. Jagirdar, V. Y. Gudur, A. Amit, and A. Dutta, "Low power personalized ecg based system design methodology for remote cardiac health monitoring," *IEEE Access*, vol. 4, pp. 8407–8417, 2016.

[3] C.-C. Hsiao, F.-W. Hsu, R.-G. Lee, and R. Lin, "Correlation analysis of heart rate variability between ppg and ecg for wearable devices in different postures," in *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pp. 2957–2962, 2017.

[4] G. Lu, F. Yang, J. A. Taylor, and J. F. Stein, "A comparison of photoplethysmography and ecg recording to analyse heart rate variability in healthy subjects," *Journal of Medical Engineering & Technology*, vol. 33, no. 8, pp. 634–641, 2009. PMID: 19848857.

[5] W.-H. Ho, C.-T. Liao, Y. J. Chen, K.-S. Hwang, and Y. Tao, "Quickly convert photoplethysmography to electrocardiogram signals by a banded kernel ensemble learning method for heart diseases detection," *IEEE Access*, vol. 10, pp. 51079–51092, 2022.

[6] Y. Liu and Q. Zhao, "Ensemble learning." handbook on computer learning and intelligence: Volume 2: Deep learning," *Intelligent Control and Evolutionary Computation*, vol. 9, pp. 635–660, 2022.

[7] J. Liu, Y. Jin, Y. Liu, Z. Li, C. Qin, X. Chen, L. Zhao, and C. Liu, "A novel p-qrs-t wave localization method in ecg signals based on hybrid neural networks," *Computers in Biology and Medicine*, vol. 150, p. 106110, 2022.

[8] Y. F. F. Y. e. a. Zhu, F., "Electrocardiogram generation with a bidirectional lstm-cnn generative adversarial network," *scientific reports*, vol. 9, 2019.

[9] P. Sarkar and A. Etemad, "Cardiogan: Attentive generative adversarial network with dual discriminators for synthesis of ECG from PPG," *CoRR*, vol. abs/2010.00104, 2020.

[10] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," *Commun. ACM*, vol. 63, p. 139–144, oct 2020.

[11] L. Domenico, F. F. La, O. Gianluigi, and M. Francesco, "Empirical mode decomposition vs. wavelet decomposition for the extraction of respiratory signal from single-channel ecg: A comparison," *IEEE Sensors Journal*, vol. 13, no. 7, pp. 2666–2674, 2013.

[12] S. Maheshwari and A. Kumar, "Empirical mode decomposition : Theory & applications," 2014.

[13] K. Vo, E. K. Naeini, and A. Naderi, "P2e-wgan: Ecg waveform synthesis from ppg with conditional wasserstein generative adversarial networks," in *Proceedings of the 36th Annual ACM Symposium on Applied Computing*, p. 1030–1036, 2021.

[14] S. Dasgupta, S. Das, and U. Bhattacharya, "Cardiogan: An attention-based generative adversarial network for generation of electrocardiograms," in *2020 25th International Conference on Pattern Recognition (ICPR)*, pp. 3193–3200, 2021.

[15] B. Bollepalli, L. Juvela, and P. Alku, "Generative adversarial network-based glottal waveform model for statistical parametric speech synthesis," *arXiv preprint arXiv:1903.05955*, 2019.

Address for correspondence:
Name: Amit Acharyya
Full postal address: Indian Institute of Technology, Hyderabad, Telangana-502284, India
E-mail address: amit_acharyya@ee.iith.ac.in