Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

*Digital Object Identifier 10.1109/ACCESS.2023.0322000*

# Synthesis of Multimodal Cardiological Signals using a Conditional Wasserstein Generative Adversarial Network

**IOANA CRETU[1], ALEXANDER TINDALE[2], WAMADEVA BALACHANDRAN[1], MAYSAM ABBOD[1], ASHRAF WILLIAM KHIR[3], and HONGYING MENG[1]**

[1]Brunel University London, London, UK
[2]Royal Brompton and Harefield Hospitals, NHS Foundation Trust, London, UK
[3]Durham University, Durham, UK

Corresponding author: Ioana Cretu (e-mail: ioana.cretu@brunel.ac.uk).

This work was supported by British Heart Foundation under Grant No.FS/19/73/34690.

**ABSTRACT** Cardiovascular diseases (CVDs) are the leading cause of mortality worldwide. Recent advancements in machine learning have significantly enhanced early detection and treatment strategies for CVDs. While electrocardiogram (ECG) signals are commonly used for detection, additional signals like arterial blood pressure (ABP) and central venous pressure (CVP) provide a comprehensive view of the cardiovascular system. However, acquiring such extensive datasets is challenging due to resource constraints, privacy issues, and ethical considerations. This paper introduces a novel Multichannel Conditional Wasserstein Generative Adversarial Network (MC-WGAN) capable of simultaneously generating synthetic ECG, ABP, and CVP signals. The MC-WGAN model addresses the data scarcity issue by providing high-fidelity synthetic data that mirrors real physiological signals, facilitating better simulation, diagnosis, and treatment planning. Evaluation against the MIT-BIH Arrhythmia Database demonstrated the model's strong performance, with competitive metrics such as RMSE, PRD, and FD, particularly excelling in the generation of ECG and ABP signals. MC-WGAN surpasses other generative models by simultaneously replicating multiple physiological signals, offering a comprehensive view of cardiovascular health. This advancement enhances diagnostic accuracy and risk stratification, setting a new standard in synthetic biomedical signal generation, and paving the way for more personalized and effective clinical interventions.

**INDEX TERMS** Generative Adversarial Network, Electrocardiogram, Blood Pressure, Biosignals

## I. INTRODUCTION

CARDIOVASCULAR diseases (CVDs) are the leading cause of death worldwide, responsible for 32% of global deaths annually, according to the World Health Organization (WHO) in 2019. Arrhythmias such as atrial fibrillation (F), atrial tachycardia (T), left bundle branch block (L), and right bundle branch block (R) are early signs of CVDs. Fortunately, the rapid advancement of artificial intelligence (AI) models for early arrhythmia detection has significantly enhanced the treatment and prevention of CVDs.

Computer-aided diagnosis (CAD) systems leveraging state-of-the-art machine learning algorithms based on cardiovascular signal analysis, such as ECG, have significantly improved the timely detection of CVDs. However, despite these technological advancements, the number of CVD cases increases annually, and human supervision remains essential in analyzing cardiovascular signal data, especially in intensive care units (ICUs). In ICUs, patients often present with complex conditions, and patient movement frequently results in lead disconnections, posing a significant challenge for continuous monitoring.

While the ECG is the primary tool in CAD systems, blood pressure signals such as arterial blood pressure (ABP), central venous pressure (CVP), and photoplethysmography (PPG) offer insights into blood pressure dynamics and arterial health, reflecting volume status and right ventricular function. Classification models that utilize ECG, ABP, and CVP signals provide a multifaceted view of cardiovascular health, enhancing diagnostic accuracy and predictive power [1], [2]. By integrating data from these three distinct yet interrelated sources, such models can capture a broader range of physiological variations and anomalies that might be missed

when analyzing each signal in isolation. This comprehensive approach allows for the detection of subtle patterns and interactions between electrical activity, blood pressure dynamics, and venous return, leading to more precise identification of cardiovascular diseases, better risk stratification, and more personalized treatment plans. Moreover, leveraging all three signals can improve the model's robustness and reliability, ultimately leading to better patient outcomes. However, creating these AI models requires extensive datasets, and the collection of such databases is not only time-consuming and costly but also demands medical expertise for accurate labeling. Furthermore, signals of rare heart abnormalities are underrepresented in existing databases, primarily because these conditions require immediate medical intervention, thereby limiting opportunities for data collection and leading to class imbalance within the datasets.

In response to the challenges presented by the lack of data, there has been significant research toward generating synthetic data that accurately reflects actual data. This approach has become a viable solution, facilitating the development of sophisticated techniques for the creation of synthetic data. These techniques allow researchers to access enriched datasets that emulate the attributes and correlations present in real-world data, thereby eliminating the costs and ethical concerns associated with traditional data collection. However, existing methods in the literature primarily focus on generating single or multi-lead ECG signals [3], [4], [5], synthesizing ECG signals from PPG signals [6], [7], and, to a lesser extent, individual blood pressure signals [8], [9]. None of these methods adequately address the simultaneous generation of multimodal cardiovascular signals, which is essential for a more holistic understanding of cardiovascular health.

In this paper, we propose a Multichannel Conditional Wasserstein Generative Adversarial Network (MC-WGAN), a novel GAN-based model designed with the unique capability to simultaneously replicate multiple critical physiological signals: ECG, ABP, and CVP. These signals are crucial for monitoring patients' cardiovascular health, providing essential insights into their cardiac rhythm, blood pressure, and the correlation between these parameters. The capability to accurately replicate these signals has significant implications for medical research, education, and potentially even clinical practice, allowing for enhanced simulations, diagnostics, and treatment planning. Beyond its multisignal generation capacity, the MC-WGAN distinguishes itself by offering the flexibility to specify the class of signals it generates. This feature enables the production of synthetic multichannel biosignals tailored to various heart dysfunction types, paving the way for more comprehensive and nuanced cardiovascular diseases research and simulation. To the best of our knowledge, there is no existing work that successfully replicates ECG, ABP, and CVP signals simultaneously. This makes MC-WGAN a groundbreaking tool in cardiovascular research, offering unprecedented opportunities for advancements in diagnosis, treatment, and patient care.

The MC-WGAN model is particularly suited for this application because it addresses specific limitations of existing generative models. Unlike other GAN variants, which often focus on single-signal generation or lack the ability to handle the complexity of multimodal data, MC-WGAN offers the unique capability to generate multiple synchronized cardiovascular signals. This model not only captures the intricate relationships between these signals but also provides the flexibility to specify the class of signals it generates, enabling the production of synthetic multichannel biosignals tailored to various heart dysfunction types.

The rest of this paper is organized as follows: Section II reviews the relevant literature. Section III details the methodology. Section IV presents the experimental results. Section V discusses these results and offers conclusions. Finally, Section VI outlines the study's limitations and suggests directions for future work.

## II. RELATED WORK

Previous studies proposed different approaches for synthetic signal generation that can be categorized into two groups: mathematical and AI-based methods. McSharry *et al.* [10] proposed a dynamical model able to produce realistic ECG signals using three coupled ordinary differential equations. This approach generates synthetic ECG using the heart rate and a set of morphological parameters for the PQRST cycle inserted by the user. Later, Clifford *et al.* [11] proposed a nonlinear model that can generate ECG, blood pressure, and respiratory signals using both nonlinear and linear characteristics. Although mathematical models can generate qualitative data, they require expert domain knowledge and the generated signals do not match the patterns of real signals as shown in [12]. As a result, Deep Neural Networks (DNNs) models have been recently proposed for synthetic signal generation.

Generative Adversarial Network (GAN) is a DNN model that has had a major impact on synthetic image generation and, most recently, synthetic signal generation. Central to GANs' architecture is their innovative approach to training data privacy, employing two distinct networks (a generator and a discriminator) that engage in adversarial training. Introduced by Goodfellow *et al.* [13] in 2014, this model utilizes a multilayer perceptron for both networks, setting a foundation for subsequent variations like the conditional GAN (CGAN), which enhances image generation with conditionality on class labels [14]. A deep convolutional GAN (DCGAN) was subsequently developed by Radford *et al.* [15] who demonstrated that deeper architectures can elevate the capabilities of GANs for image generation. In the DCGAN, however, the discriminator experienced mode collapse, and the loss graphs did not provide any meaningful information. As a result, a new type of GAN called the Wasserstein GAN (WGAN) was proposed to alleviate this problem [16]. The main advantage of GANs is that they do not require input from experts and the generator is not directly connected to the real data, therefore, offering privacy of the training dataset.

One of the pioneering applications of GANs that gener-

ated medical signals incorporated Long Short-Term Memory (LSTM) layers in both the generator and discriminator to generate synthetic ECG and Electroencephalography (EEG) signals [17]. This methodology was further refined by employing bidirectional LSTM layers (Bi-LSTM) in the generator and pairing it with a one dimensional (1D) Convolutional Neural Network (CNN) in the discriminator [18], [19]. Hazra and Byun [19] demonstrated the effectiveness of the Bi-LSTM GAN on ECG, EEG, electromyogram (EMG), and PPG signals. Further advancements are demonstrated in [20], where the authors employed the least-squares loss function within their GAN architecture (LSGAN) to produce simultaneous multichannel ECG data for multivariate ECG generation. In [21], LSGAN and Cycle-GAN were separately employed to generate phonocardiogram (PCG) and ECG data. A different approach was taken by Golany *et al.* [22] who incorporated the ordinary differential equations developed by McSharry *et al.* in the GAN architecture to improve the morphology of the resulted synthetic signals. Similarly, Neifar *et al.* [23] proposed a WGAN that incorporates advanced prior knowledge modeling of ECG shape and dynamics to improve control over the generation process by using statistical shape modeling to capture the 2-D dynamics and shape variations of ECG signals, enabling a more comprehensive modeling of ECG dynamics. In [24], Tran *et al.* generated long ECG sequences by converting the 1D signals into images and feeding it into a WGAN model for training.

Recent works have proposed the use of attention layers within the GAN architecture to enhance the quality of generated samples [25], [26]. In [25], the BC-GAN model was introduced for single lead ECG signal generation. The BC-GAN model captures temporal correlations and essential features in the signal by integrating a Bi-LSTM architecture with a Convolutional Block Attention Mechanism (CBAM) in the generator, which consists of two branches, each preceded and followed by two layers of deconvolution. In [26], the authors improve the GAN by integrating attention layers, using a transformer with a convolutional neural network (CNN) in the generator and a transformer in the discriminator.

Other works combine the advantages of the encoder-decoder architectures with GAN architectures to create synthetic signals, demonstrating the versatility and advantages of these models [27], [28]. However, these models often introduce increased complexity, making it difficult to troubleshoot issues within the models. Consequently, more recent studies have proposed using U-net based architectures for the generator within GAN models [29]–[33]. In [32], the authors used a WGAN to generate 12-lead ECG signals using a U-Net generator and a CNN discriminator. Employing a similar architecture, Seo *et al.* [32] demonstrated the capabilities of the U-Net generator in complex data modeling by using it within a GAN to generate multiple synthetic ECG leads from a single lead ECG. This approach leverages the benefits of the encoder-decoder architecture while reducing complexity by integrating it directly into the GAN model, rather than maintaining separate models.

The above works demonstrate the potential of GANs for synthetic biosignals generation, with a notable focus on generating single-channel ECG signals or standalone biosignals.

## III. METHODOLOGY

This section provides a detailed description of our proposed method and its components. It begins with an overview of the system, followed by an in-depth discussion of its elements, including MC-WGAN, GAN networks, generator, discriminator, and loss functions.

### A. OVERVIEW

The first step in the proposed methodology is the preprocessing our input signals, which includes the segmentation of the signals into heartbeats. This process includes the detection of R-peaks and careful segmentation of the signals. Each type of signal—ECG, ABP, and CVP—possesses distinct characteristics; therefore, segmentation is uniquely tailored to capture the delays between the heart's electrical activity and the subsequent responses observed in blood pressure signals. This precision in preprocessing helps refine the input data, ensuring it more accurately reflects true physiological interactions. Before segmenting the signals, we excluded some that were heavily contaminated by noise due to lead dislocation or sudden patient movement, as these predominantly contained noise rather than useful data. However, we did not add any noise filtering steps because GANs are inherently robust to noise, and the presence of noise in the real signals allows the GAN to produce synthetic data that more accurately reflects real-world conditions. This approach helps replicate normal variations in signal quality.

The architecture of the MC-WGAN is designed to manage and synthesize this complex data effectively. The generator component of the network begins with an input layer that receives three-channel data from the ECG, ABP, and CVP signals, along with their respective labels. These inputs are first passed through an embedding layer, which transforms them into a suitable format for subsequent processing. The generator's primary objective is to produce realistic multichannel signals that mimic the natural distribution and interactions observed in real physiological data, thereby enhancing the model's ability to generate credible synthetic outputs.

Parallel to the generator, the discriminator is configured as a 1D convolutional network. Its main function is to evaluate the realism of the signals generated by the generator, determining whether they are real or fabricated. This assessment is critical as it directly influences the training dynamics of the network. The discriminator scores the authenticity of the data, effectively learning to distinguish between genuine and generated signals. During the training process, the discriminator's evaluations lead to the computation of losses for both the generator and discriminator. These calculated losses are crucial for the iterative refinement and optimization of the model's performance, ensuring that the generated outputs are both realistic and physiologically plausible.

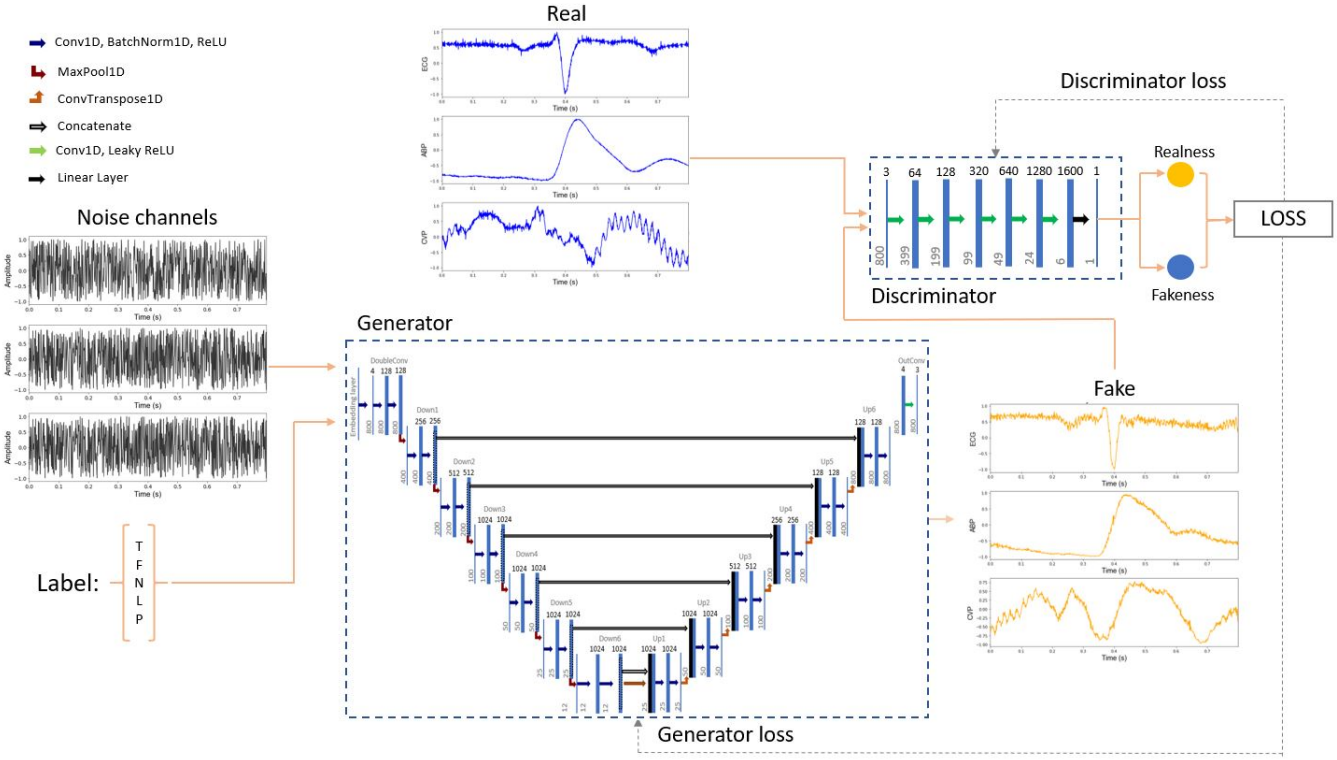



**FIGURE 1. Architecture of the Proposed MC-WGAN Model: This model incorporates a conditional generator and discriminator designed for processing multimodal signals. The input layer of the generator accepts three-channel data from multimodal signals (ECG, ABP, CVP), along with their associated labels. These inputs are initially processed through an embedding layer, transforming them into a format suitable for further processing. The generator's architecture, inspired by the U-net model, includes a double convolution layer, followed by six downsampling layers, then six upsampling layers, concluding with a final output convolution layer. The generator is designed to create realistic multichannel signals that mimic the distribution and interaction observed in the real signals.Structured as a 1D convolutional network, the discriminator evaluates whether the signals produced by the generator are real or synthetic. It features six 1D convolution layers, each followed by a Leaky ReLU activation function, and ends with a linear layer. This configuration is used to assess the authenticity of the data by scoring its realness or fakeness. During training, the discriminator assesses the generator's outputs, which leads to the calculation of losses for both the generator and discriminator. These losses are then utilized to iteratively refine the performance of the model.**

The detailed configurations of the generator and the discriminator, which are central to the MC-WGAN's functionality, are further elaborated in subsequent sections and their architectures can be seen in Figure 1.

### B. GENERATIVE ADVERSARIAL NETWORKS

GANs are composed of two separate models that are trained together: the generator model $G$ and the discriminator model $D$, as shown in Figure 1. The generator's role is to produce samples that resemble the distribution of real data by starting with a random noise input. The discriminator, on the other hand, acts as a classifier, receiving both real data and the synthetic samples produced by the generator as input, with the goal of distinguishing real samples from the generated ones. In the traditional GAN, the interaction between the generator and discriminator is a minmax problem. Each time the discriminator reviews a batch of data (containing both real and generated samples), it makes its classifications, which are then used as feedback for the generator. The goal of the generator is to adjust its parameters based on this feedback to improve its ability to produce data that the discriminator will misclassify as real. Meanwhile, the discriminator continuously attempts to enhance its ability to distinguish real from fake, minimizing its classification error. This training process continues until the Nash equilibrium is achieved. The loss function of a GAN is defined as follows:

$$\min_G \max_D (D, G) = \mathbb{E}_{x \sim P_r} [\log (D(x))] + \mathbb{E}_{z \sim P_z} [\log (1 - D(G(z)))] \quad (1)$$

where D(x) represents the loss of the discriminator, $P_r$ represents the distribution of the real data $x$ and $P_z$ represents the distribution of the generated data $G(z)$.

CGANs have been introduced to gain control over the type of generated samples by adding an additional layer of information about each sample's label. The loss function then becomes:

$$\min_G \max_D (D, G) = \mathbb{E}_{x \sim P_r} [\log (D(x|c))] + \mathbb{E}_{z \sim P_z} [\log (1 - D(G(z|c)))] \quad (2)$$

where c stands for condition and represents the vector corresponding to the signal's labels.

| Block | Layer | Input Size | Feature Maps | Kernel Size | Stride | Padding | Output Size |
|---|---|---|---|---|---|---|---|
| DoubleConv | Conv1D | 4 x 800 | 128 | 3 | 1 | 1 | 4 x 128 x 800 |
| | BatchNorm1D | 4 x 128 x 800 | - | - | - | - | 4 x 128 x 800 |
| | ReLU | 4 x 128 x 800 | - | - | - | - | 4 x 128 x 800 |
| | Conv1D | 4 x 128 x 800 | 128 | 3 | 1 | 1 | 4 x 128 x 800 |
| | BatchNorm1D | 4 x 128 x 800 | - | - | - | - | 4 x 128 x 800 |
| | ReLU | 4 x 128 x 800 | - | - | - | - | 4 x 128 x 800 |
| Down1 | MaxPool1D | 4 x 128 x 800 | 128 | 2 | 2 | 0 | 4 x 128 x 400 |
| | DoubleConv | 4 x 128 x 400 | 256 | 3 | 1 | 1 | 4 x 256 x 400 |
| Down2 | MaxPool1D | 4 x 256 x 400 | 256 | 2 | 2 | 0 | 4 x 256 x 200 |
| | DoubleConv | 4 x 256 x 200 | 512 | 3 | 1 | 1 | 4 x 512 x 200 |
| Down3 | MaxPool1D | 4 x 512 x 200 | 512 | 2 | 2 | 0 | 4 x 512 x 100 |
| | DoubleConv | 4 x 512 x 100 | 1024 | 3 | 1 | 1 | 4 x 1024 x 100 |
| Down4 | MaxPool1D | 4 x 1024 x 100 | 1024 | 2 | 2 | 0 | 4 x 1024 x 50 |
| | DoubleConv | 4 x 1024 x 50 | 1024 | 3 | 1 | 1 | 4 x 1024 x 50 |
| Down5 | MaxPool1D | 4 x 1024 x 50 | 1024 | 2 | 2 | 0 | 4 x 1024 x 25 |
| | DoubleConv | 4 x 1024 x 25 | 1024 | 3 | 1 | 1 | 4 x 1024 x 25 |
| Down6 | MaxPool1D | 4 x 1024 x 25 | 1024 | 2 | 2 | 0 | 4 x 1024 x 12 |
| | DoubleConv | 4 x 1024 x 12 | 1024 | 3 | 1 | 1 | 4 x 1024 x 12 |
| Up1 | ConvTranspose1D | 4 x 1024 x 12 | 1024 | 2 | 2 | - | 4 x 1024 x 25 |
| | DoubleConv | 4 x 1024 x 25 | 1024 | 3 | 1 | 1 | 4 x 1024 x 25 |
| Up2 | ConvTranspose1D | 4 x 1024 x 25 | 1024 | 2 | 2 | - | 4 x 1024 x 50 |
| | DoubleConv | 4 x 1024 x 50 | 1024 | 3 | 1 | 1 | 4 x 1024 x 50 |
| Up3 | ConvTranspose1D | 4 x 1024 x 50 | 1024 | 2 | 2 | - | 4 x 1024 x 100 |
| | DoubleConv | 4 x 1024 x 100 | 512 | 3 | 1 | 1 | 4 x 512 x 100 |
| Up4 | ConvTranspose1D | 4 x 512 x 100 | 512 | 2 | 2 | - | 4 x 512 x 200 |
| | DoubleConv | 4 x 512 x 200 | 256 | 3 | 1 | 1 | 4 x 256 x 200 |
| Up5 | ConvTranspose1D | 4 x 256 x 200 | 256 | 2 | 2 | - | 4 x 256 x 400 |
| | DoubleConv | 4 x 256 x 400 | 128 | 3 | 1 | 1 | 4 x 128 x 400 |
| Up6 | ConvTranspose1D | 4 x 128 x 400 | 128 | 2 | 2 | - | 4 x 128 x 800 |
| | DoubleConv | 4 x 128 x 800 | 128 | 3 | 1 | 1 | 4 x 128 x 800 |
| OutConv | Conv1D | 4 x 128 x 800 | 3 | 1 | 1 | - | 3 x 800 |

**TABLE 1. The detailed architecture of the generator. It begins with a Double Convolution (DoubleConv) block, which comprises two convolutional layers, each followed by batch normalization and a ReLU activation function. Subsequently, the architecture includes six downsampling (Down) layers; each layer consists of one max pooling layer to reduce spatial dimensions, followed by another DoubleConv block for further feature processing. The data then passes through six upsampling (Up) layers, each comprising a convolutional transpose layer designed to gradually restore the data's dimensions, paired with a DoubleConv block to refine the features during this reconstruction phase. The architecture end with a final output convolution (OutConv) layer, which consolidates the processed features into the final output of the generator.**

Despite their potential, GANs have limitations that make their training difficult and sometimes unstable. Traditional GANs are notoriously hard to train due to training instability, mode collapse and the vanishing gradient problem. The training instability is caused by the models being trained simultaneously to find the Nash equilibrium in the above described minmax game. However, the generator and discriminator models update their cost functions independently, without considering the adjustments being made by the other model. This independent and concurrent gradient update mechanism does not guarantee convergence, making the training process challenging and often unstable. Mode collapse, on the other hand, happens when the generator produces the same output or a limited variety of outputs. This happens when the generator find a specific distribution that consistently tricks the corresponding discriminator, and as a result it has no incentive to generate a variety of samples. Moreover, for GANs to work effectively, the capabilities of the discriminator and generator need to be balanced. If the discriminator does not work well, it can not provide good feedback that can improve the generator. If it is too strong, the gradient of the loss function is close to zero and the learning becomes slow, leading to gradient vanishing problems.

WGAN networks were introduced to solve these issues by modifying the loss function of the traditional GANs. Arjovsky *et al.* [16] emphasised that using Jensen–Shannon (JS) Divergence to measure the similarity of the real and generated distributions is what causes the vanishing gradient problem [19]. WGAN significantly elevates the training stability by using a new type of loss function called Wasserstein loss, also known as Earth Mover's (EM) distance, to calculate the distance between the real and synthetic distribution of the samples. To ensure that the Wasserstein loss is valid, the 1-Lipschitz (1-L) continuity is enforced on the discriminator, known as a critic in WGAN. There are two methods that can be used to enforce the 1-L continuity: weight clipping or gradient penalty (GP), with GP proven to be more effective. Therefore, the loss of the WGAN with GP becomes:

$$\begin{aligned}\min_G \max_D (D,G) = \mathbb{E}_{z\sim P_z}[G(z)] - \\ \mathbb{E}_{x\sim P_r}[D(x)] + \\ \lambda \mathbb{E}_{\hat{x}\sim P_{\hat{x}}}\left[\left(\|\triangle_{\hat{x}} D(\hat{x})\|_2 - 1\right)^2\right]\end{aligned} \tag{3}$$

where $E_{z\sim P_z}[G(z)] - E_{x\sim P_r}[D(x)]$ represents the critic loss and $\lambda E_{\hat{x}\sim P_{\hat{x}}}[(||\triangle \hat{x} D(\hat{x})||_2 - 1)^2]$ represents the gradient

penalty.

### C. GENERATOR

The architecture of the generator within our proposed MC-WGAN model is inspired by the U-Net framework [30]. However, it has been specifically adapted for processing multi-channel time-series data. Unlike the original U-Net design, which utilizes 2D convolutional layers and upsampling layers suited for image data, our model employs 1D convolutional layers and 1D transpose convolution layers. This modification enables the effective handling of time-series signals by capturing the temporal dynamics inherent in multi-channel biomedical data.

The generator takes as input three channels of Gaussian noise and one vector containing the desired labels we want to obtain as output. At first, the labels are passed though an embedding layer which contains a lookup table where each integer index corresponds to a dense vector. The embedding layer transforms each label, represented by an integer, into its corresponding vector according to the lookup table. The output vector from the embedding layer is then concatenated with the multi-channel signals along the channel dimension. The new input, that carries both the noise channel and the embedded label information, is then passed through an input block, 6 down-sampling blocks, 6 up-sampling blocks and an output block as shown in Table 1. The down-sampling blocks consist of one max pooling layer and two 1D convolutional layers, each followed by batch normalisation and ReLU activation function. The up-sampling blocks consist of a 1D convolutional transpose layer and two 1D convolutional layers, each followed by batch normalisation and ReLU activation. The transpose convolutional layers not only upsample the input but also have learnable parameters (weights and biases) that are updated during the training process, which makes them more suitable for generating signals from noise. During up-sampling, the features from the down-sampling blocks are concatenated to the features of the up-sampling blocks, and padding is applied if the two do not have the same dimensions. In this manner, the the generator produces synthetic signals with the same number of channels and length as the input noise.

### D. DISCRIMINATOR

The discriminator model consists of six 1D convolutional layers, each of them being followed by a Leaky ReLU activation function, and a fully connected layer at the end of the network. In this work, we use the Wasserstein loss function for training, therefore, the critic outputs a score that denotes the realness or fakeness of a given signal. This alternative way of training the generator provides a more stable training and makes the GAN to be less sensitive to the choice of hyperparameters.

### E. LOSS FUNCTION

The implemented model combines the benefits of the WGAN architecture described in section III-B with the ability of the CGAN to control the labels of the generated samples. In this way we are able to generate only the desired types of samples, while having the training stability and performance of the WGAN. Using the label dependency from equation (2) and the WGAN-GP loss from equation (3) we then obtain the following loss function that is used in our training:

$$\begin{aligned}\min_G \max_D (D,G) = \mathbb{E}_{z\sim P_z}\left[G(z|c)\right] - \\ \mathbb{E}_{x\sim P_r}\left[D(x|c)\right] + \\ \lambda \mathbb{E}_{\hat{x}\sim P_{\hat{x}}}\left[\left(\|\triangle_{\hat{x}} D(\hat{x}|c)\|_2 - 1\right)^2\right]\end{aligned} \tag{4}$$

## IV. EXPERIMENTAL RESULTS

This section provides a detailed overview of the evaluation metrics, datasets, preprocessing stages, and training processes utilized to assess the quality of the signals generated by our model. Initially, we describe the specific metrics employed to evaluate the authenticity and accuracy of the synthetic signals. Following this, we outline the datasets used in our study, including their sources and characteristics. We then detail the preprocessing procedures and finally, we elaborate on the training methodology of the model, discussing the configurations, settings, and strategies implemented to optimize performance and ensure robust signal generation.

### A. EVALUATION METRICS

To evaluate the quality of the generated signals, we calculated a series of distance metrics. Starting with the Fréchet Distance (FD), this metric measures the morphological similarity between the real and generated signals by considering both the location and ordering of the values along the signal curve. This evaluation is particularly relevant in cardiovascular signal analysis, where the shape of the signals carries critical diagnostic information.

If we let $X = (x_1, x_2, \ldots, x_N)$ and $Y = (y_1, y_2, \ldots, y_N)$ be the real and generated signal curves, the Fréchet distance is defined as:

$$d_F(X,Y) = \inf_{\alpha,\beta} \max_{s\in[0,1]} \|X(\alpha(s)) - Y(\beta(s))\| \tag{5}$$

where $\alpha$ and $\beta$ range over all continuous, non-decreasing, and surjective mappings from $[0,1]$ to $([0,1])$ and $\|\cdot\|$ denotes the Euclidean distance.

In the discrete case, the Fréchet distance can be represented as:

$$d_F(X,Y) = \min_{\sigma,\tau} \max_{k\in[1,N]} \|x_{\sigma(k)} - y_{\tau(k)}\| \tag{6}$$

where $\sigma$ and $\tau$ are order-preserving mappings of indices between the sequences $X$ and $Y$, such that $\sigma : [1,N] \to [1,N]$ and $\tau : [1,N] \to [1,N]$.

This definition ensures that the Fréchet distance $d_F(X,Y)$ is the minimum possible value of the maximum Euclidean distance between corresponding points on the real and generated signal curves, taken over all possible parameterizations $\alpha$ and $\beta$. In other words, it finds the best possible alignment of the two signals such that the greatest distance between

matched points is as small as possible, measuring the similarity between the two signals in terms of both their shapes and the sequence of their values.

The root mean square error (RMSE) measures the difference between the values of the generated and the original signals. RMSE is sensitive to large errors, making it particularly useful for detecting significant deviations that could affect the quality and usability of the generated signals. RMSE is calculated using the following equation:

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - y_i)^2} \tag{7}$$

The Mean Absolute Error (MAE) stands out as one of the simplest yet most effective ways of measuring the average of the absolute differences between generated and original signals. Compared to RMSE, MAE is less sensitive to outliers. This makes it particularly suitable for analyzing biosignals with types of rhythm disturbances where small, frequent errors can be more detrimental to diagnosis than rare large deviations. MAE is calculated by averaging the absolute differences between the generated values and the real values as follows:

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N}|x_i - y_i| \tag{8}$$

where $x_i$ represents the i-th value in the real signal, $y_i$ represents the i-th value in the generated signal, and N is the length of the signal.

The Percent Root Mean Square Difference (PRD) metric is widely used to assess the distortion between synthetic signals generated by the model and the original signals. It provides a percentage-based performance metric that reflects the relative difference to the real signal's scale, allowing for an intuitive understanding of the level of distortion. The PRD is calculated using the following formula:

$$\text{PRD} = \left(\sqrt{\frac{\sum_{i=1}^{N}(x_i - y_i)^2}{\sum_{i=1}^{N}x_i^2}}\right) \times 100 \tag{9}$$

Dynamic Time Warping (DTW) is used to measure the similarity between two time series curves that can vary in time or speed by finding the optimal alignment between them while considering shifts and distortions. The DTW distance can be formally defined as:

$$\text{DTW}(X, Y) = \sqrt{\sum_{k=1}^{K}(x_{\sigma(k)} - y_{\tau(k)})^2} \tag{10}$$

where $K$ is the length of the optimal warping path, and $\sigma$ and $\tau$ are mappings that define the alignment between the indices of X and Y. These mappings minimize the cumulative distance:

$$\text{DTW}(X, Y) = \min_{\sigma,\tau}\sum_{k=1}^{K}|x_{\sigma(k)} - y_{\tau(k)}| \tag{11}$$

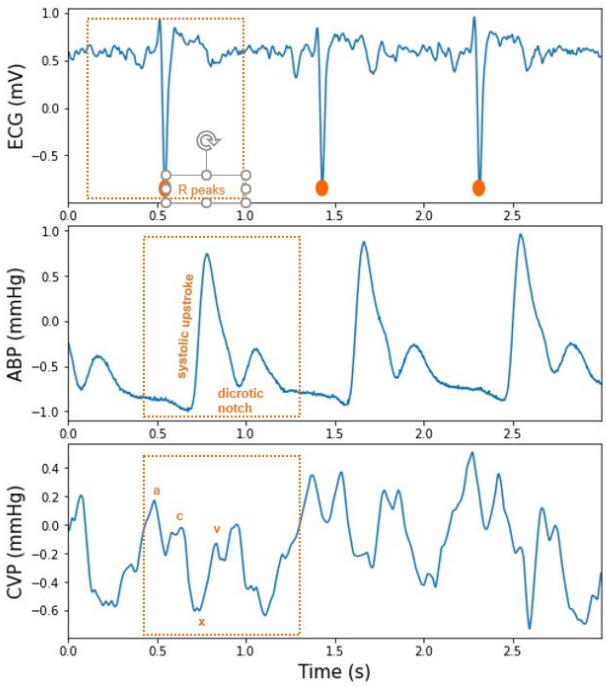



**FIGURE 2. Segmentation process of the signals. First we identify ECG R peaks and select heartbeats with 800ms centered around each R peak. then, to accommodate the timing disparity between ECG R peaks and the systolic events, we segment the blood pressure signals with 200ms before and 600ms after the R peak locations.**

Maximum Mean Discrepancy (MMD) is calculated to quantify the dissimilarity between the generated and real signals, aiming to identify differences in the distribution characteristics. MMD can effectively detect variations in the overall distribution characteristics, not just individual data points and is calculated as follow:

$$\text{MMD}^2(X, Y) = \frac{1}{m^2}\sum_{i=1}^{m}\sum_{j=1}^{m}k(x_i, x_j) + \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}k(y_i, y_j) - \frac{2}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}k(x_i, y_j) \tag{12}$$

where the function $k$ is a Gaussian kernel function that measures the similarity between data points, and $m$ and $n$ represent the number of samples from the real dataset and the generated dataset, respectively.

The generated signals are evaluated by calculating each of the above distance metrics between each generated signal and all real signals, and then averaging the obtained values. The smaller these values are, the closer the generated signals are to the real ones, indicating higher quality and fidelity.

### *B. DATA*

#### 1) Collected Dataset

The dataset integrates ABP, CVP, and ECG signals to provide a comprehensive view of a patient's hemodynamic status, cru-

cial for diagnosing and managing CVDs. This combination allows for a detailed correlation and analysis of cardiac and vascular interactions, essential for understanding complex cardiovascular conditions, optimizing treatment strategies, and improving patient outcomes.

Simultaneous lead V1 ECG, ABP, and CVP signals were collected from patients in the intensive care unit (ICU) at Harefield Hospital, London. ABP measurements are critical for assessing the pressure exerted by the blood on arterial walls during different phases of the heart cycle. The invasive ABP was transduced from the right radial artery using Edwards Lifesciences TruWave pressure transducers. CVP measurements provide insights into the pressure within the thoracic vena cava near the right atrium, reflecting right ventricular function and venous blood return to the heart. CVP was transduced from the right superior vena cava using the same type of pressure transducers. The electrical activity of the heart was recorded using a Boston Scientific Labsystem Pro electrophysiology recording system.

The signals underwent digital-to-analog conversion using a National Instruments DAQ card and LabVIEW software (National Instruments, TX, USA). Data were sampled at a frequency of 1000 Hz. The dataset includes signals from 29 patients, each with different baseline heart rhythms: sinus rhythm (N), atrial fibrillation (F), atrial tachycardia (T), and left bundle branch block (L).

N represents the normal regular rhythm of the heart, where the electrical impulses originate from the sinoatrial node. F is characterized by fast and irregular beating of the atria, leading to inefficient blood flow. T involves an abnormally fast heartbeat originating from the atria, distinct from normal sinus rhythm. L occurs when there is a delay or blockage along the pathway that electrical impulses travel to make the heart beat, specifically in the left bundle branch.

Additional data were collected from the same set of patients while they were paced under different settings. The pacing mode used was dual-chamber pacing (DDD mode), which involves pacing, sensing, inhibition, and stimulation. The pacing rate was set at 90 beats per minute (bpm) or 10 bpm above the patient's sinus rhythm. Patients began pacing at a reference AV delay of 120 ms, transitioning rapidly to a tested AV delay (ranging from 40 ms to 280 ms in 40 ms increments) for 20 beats before returning to the reference AV delay. This cycle repeated 8 times for each tested AV delay. Testing stopped when intrinsic conduction occurred or the tested AV delay reached 280 ms. These signals were labelled as paced beats (P).

This study has been approved by the South West - Cornwall and Plymouth Research Ethics Committee as part of the PACESIM trial (ISRCTN15383573). All patients provided informed consent prior to participation. They were thoroughly briefed about the study's aims, procedures, and potential risks. Patient data were anonymized to protect privacy. This process involved removing all personally identifiable information (PII) and replacing it with unique, non-identifiable codes. The anonymized data were then securely stored and accessed only by authorized personnel.

#### 2) Public Dataset

To evaluate our model against existing literature, we utilized lead II from the MIT-BIH Arrhythmia Dataset, a public database provided by the Massachusetts Institute of Technology. This dataset comprises annotated long-term ECG signals recorded at a sampling frequency of 360 Hz and includes a variety of 17 different rhythm classes.

For this work, we grouped the signals according to the AAMI standard into the following categories: N, atrial premature beat (A), premature ventricular contractions (V), right bundle branch block (R), and L. A refers to an early contraction of the atria, which can disrupt the regular heart rhythm. V are early heartbeats originating from the ventricles, often causing a sensation of skipped heartbeats. R occurs when there is a delay or blockage in the electrical impulses traveling to the right ventricle, leading to an uncoordinated heartbeat.

### C. DATA PREPROCESSING

The R peak detection of our collected ECG signals has been done using Pan-Tompkins algorithm [34]. Then, the segmentation of our signals has been tailored to accommodate the physiological delays observed between the ECG and blood pressure signals. As shown in Fig. 2, there is a brief delay between the ventricular contraction (as indicated by the R peak in the ECG signal) and the subsequent response in the blood pressure, which must be accounted for in our segmentation approach. To address this, we adopted distinct segmentation strategies for each type of signal. For the ECG, we used a window of 800 milliseconds (ms) centered around the R peak to capture the full cardiac cycle. Conversely, for the ABP and CVP signals, the segmentation was adjusted to include 200 ms prior to and 600 ms following the R peak. This methodological adjustment ensures that the segmented ABP and CVP accurately reflect the hemodynamic response associated with each ECG heartbeat, facilitating precise alignment and analysis of these crucial cardiovascular signals in our research. In the public dataset, on the other hand, we used the annotated R peak locations to perform the signal segmentation and selected 88 datapoints before and 168 datapoints after each R peak. After the segmentation all the signals were normalised in the [-1;1] range.

### D. TRAINING

Hyperparameter tuning was a critical aspect of developing the MC-WGAN, given the model's complexity and the need for stable training. The Adam optimizer was chosen for its ability to handle sparse gradients on noisy problems, with an initial learning rate set to $1e-5$. The choice of $\beta_1 = 0$ and $\beta_2 = 0.9$ was based on preliminary experiments that showed these values provided a good balance between convergence speed and training stability. We used a training ratio of 5:1 for the discriminator to the generator, meaning the discriminator was updated five times for each update of the generator. This ratio was chosen after experimenting with different values

and observing that a higher ratio improved training stability, reducing mode collapse and helping maintain a balance between the generator and discriminator. The batch size was set to 16, which was determined through grid search experiments. We found that smaller batch sizes led to more stable training, likely due to the increased noise in gradient estimates, which helped the generator escape local minima. However, too small of a batch size caused the training to be overly noisy, so 16 was chosen as a middle ground. In addition to these hyperparameters, the choice of weight initialization and the use of gradient penalty (instead of weight clipping) were crucial in ensuring that the model trained effectively without suffering from the common pitfalls associated with GANs, such as vanishing gradients and mode collapse. We experimented with different levels of gradient penalty $\lambda$ and found that setting $\lambda = 10$ provided the best balance between training stability and convergence speed.

By carefully tuning these hyperparameters and considering the specific requirements of the MC-WGAN architecture, we were able to develop a model that generates high-quality synthetic signals while maintaining stable and efficient training dynamics.

### *E. RESULTS*

This section presents the evaluation results for the signals generated by our MC-WGAN model, specifically focusing on the simultaneous generation of ECG, ABP, and CVP signals. All the signals were generated simultaneously for each class. However, for clarity and thoroughness, we discuss the results for each individual signal type and each generated class separately. Furthermore, to contextualize our findings within the broader field, we compare our results with those reported in existing literature. This comparison includes a detailed analysis of our model's performance using the MIT-BIH Arrhythmia Database, which serves as a benchmark for assessing our model against established standards. This approach allows us to present a clear and direct comparison between our results and those documented in previous studies.

The results obtained using our dataset are presented in Table 2, and a visual comparison between the signals is illustrated in Figure 3. The real signals are displayed on the left-hand side, while the corresponding synthetic signals are shown on the right-hand side. Each set of signals—ECG, ABP, and CVP—is captured within a single heartbeat cycle, where the real signals are shown in blue and the synthetic ones in orange.

The proposed MC-WGAN, which generates simultaneous multimodal data for three types of signals and different types of heart rhythms, was evaluated across all classes and signals, for each class group and for each signal type (Table 2). The average results across all signals and all classes show an RMSE value of 0.319, a PRD value of 33.159, an FD value of 0.490, a DTW value of 3.266, and an MMD value of 0.46899. The FD value of 0.490, which measures the similarity between two distributions, indicates that the distribution of the generated signals is fairly close to that of the original signals, implying that the synthetic data preserves the overall structure of the original data. The DTW value of 3.266, which measures the similarity between two temporal sequences, suggests that the synthetic signals are generally well-aligned with the original signals, though some minimal temporal discrepancies exist. The MMD of 0.46899, which measures the difference between two distributions in terms of their means, indicates that the distributions of the signals are not significantly different. However, the performance of the model shows slight variations across different signal types. Class P consistently exhibited the highest accuracy across all signal types, followed closely by class F. For the P rhythm class, the generated ECG signals showed an RMSE of 0.159, PRD of 17.049, FD of 0.386, DTW of 2.012, and an MMD of 0.006. In comparison, class F demonstrated an RMSE of 0.173, PRD of 17.343, FD of 0.238, DTW of 1.766, and MMD of 4.172. Conversely, class T consistently displayed the highest errors, with the generated ECG signals presenting an RMSE of 0.383, PRD of 38.341, FD of 0.688, DTW of 4.72, and MMD of 0.019. These variations may be attributed to differences in pattern complexity or the volume of training data available for each class.

A similar behavior is observed among different signal types. Among the three, the ECG and ABP signals show higher variability in performance, while CVP generally produces the highest errors, indicating that CVP might be the most challenging signal for the model to synthesize accurately. For the ECG signals, the MC-WGAN model achieved an RMSE of 0.248, a PRD of 25.094, an FD of 0.387, a DTW of 2.407, and an MMD of 1.402, averaged across all labels. The PRD value indicates that the generated signals closely follow the general trends of the real ECG data, though some errors in amplitude or waveform shape are present. The FD value reflects moderate variability in the signal's derivative, suggesting some discrepancies in waveform features. For ABP signals, the results show slightly better accuracy and fidelity, with an FD of 0.349, a DTW of 1.414, and a significantly lower MMD of 0.00065. These lower error metrics suggest a closer approximation to the true ABP signals than to the ECG signals, highlighting the model's capacity to handle ABP signals with slightly greater precision. The FD of 0.349, the lowest among the three, indicates high consistency in capturing the dynamics of ABP fluctuations, which are critical for accurate blood pressure monitoring. Conversely, the model's performance on CVP signals indicates room for improvement. The RMSE of 0.451, PRD of 48.334, FD of 0.735, and DTW of 5.978 are significantly higher than those for ECG and ABP, suggesting a poorer fit between the generated and actual CVP signals. This higher level of error may be due to the inherent complexity or variability of CVP signals, making them more challenging for the model to synthesize accurately.

In comparing our MC-WGAN model with existing generative models, it's crucial to highlight its unique capability to simultaneously generate three different types of cardiovascu-

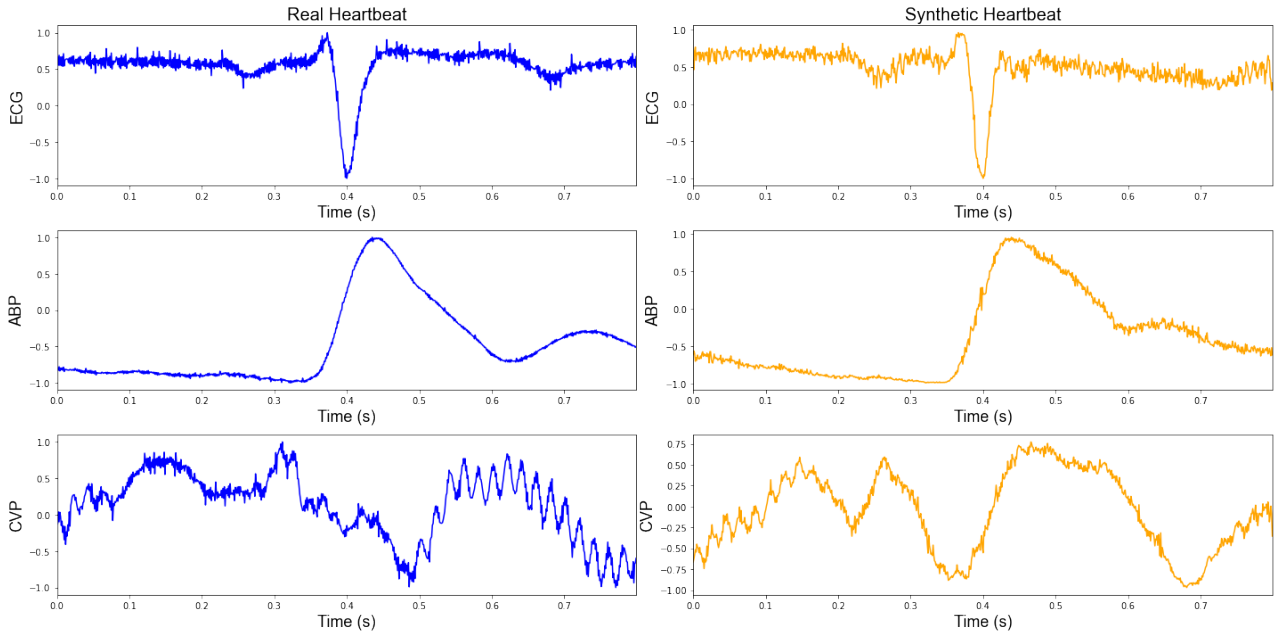



**FIGURE 3.** This figure displays real ECG, ABP, and CVP signals from a patient with N rhythm (shown in blue) alongside synthetic signals of the same types generated by the MC-WGAN model for N rhythm (shown in orange), illustrating the model's ability to replicate complex biomedical signals.

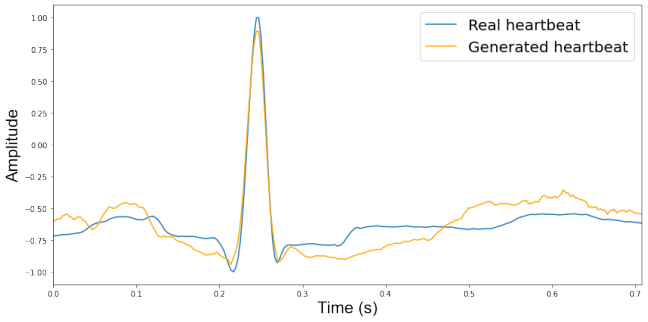



**FIGURE 4.** Comparison between a real ECG heartbeat (blue) from the MIT-BIH arrhythmia database and a synthetic ECG heartbeat (orange) generated by the MC-WGAN model for a N rhythm signal.

lar signals: ECG, ABP, and CVP. This stands in contrast to the other models listed in Table 4 and Figure 5, which either focus on a single signal type or produce simultaneous multichannel ECG signals. Our model was trained on both our proprietary dataset and the publicly available MIT-BIH dataset, allowing us to benchmark its performance against existing models. The results demonstrate a comprehensive approach, not only producing multimodal signals across different heart rhythms but also achieving competitive performance metrics on the public dataset.

The MC-WGAN achieved strong results on the MIT-BIH database, with an RMSE of 0.205, an MAE of 0.146, a PRD of 20.535, an FD of 0.225, a DTW of 1.355, and an MMD of 0.0051. An example comparison between real and synthetic samples generated by our model is provided in Figure 4, and performance metrics for each individual class are detailed in Table 3.

While our model's MSE of 0.047 and RMSE of 0.205 are higher than the lowest values reported by models like the 2D WGAN [5] (0.002 and 0.024, respectively), these metrics still indicate solid performance, especially considering the simultaneous generation of multiple signal types. The MAE of 0.146 further underscores the model's ability to accurately replicate the characteristics of real signals. Compared to other models using the same database, our model achieves significantly lower DTW and MMD values. However, it shows a higher PRD compared to Transformer-GAN [26] and SynSigGAN [19], and slightly higher FD than BC-GAN [25], as illustrated in Figure 5 and Table 4. Notably, the PRD value of 20.535 is substantially lower than that of models like BiLSTM-CNN GAN (51.799) and BC-GAN (33.753), indicating better amplitude and waveform fidelity. The FD of 0.225 and DTW of 1.355 reflect strong distributional and temporal alignment between the generated and real signals, surpassing several other models. The MMD of 0.0051, while not directly comparable due to limited reporting, further validates the model's effectiveness.

It's important to highlight the complexity of our task compared to other models reported in the literature. Our model's ability to produce multimodal simultaneous signals involves not only replicating the patterns and morphologies of real signals but also capturing their interdependencies, unlike models that generate only one type of signal at a time or multiple leads of the same signal type. For example, while SynSigGAN excels in generating individual ECG and PPG signals with PRDs of 6.343 and 5.167 respectively, it generates these signals separately, potentially missing interdependencies. Overall, the MC-WGAN delivers competitive performance, particularly in terms of distributional similarity and amplitude fidelity,

| Signal | Class | MSE | RMSE | MAE | PRD | FD | DTW | MMD |
|---|---|---|---|---|---|---|---|---|
| ECG | TAH (T) | 0.147 | 0.383 | 0.265 | 38.341 | 0.688 | 4.72 | 0.019 |
| | AF (F) | 0.030 | 0.173 | 0.134 | 17.343 | 0.238 | 1.766 | 4.172 |
| | SR (N) | 0.042 | 0.206 | 0.144 | 20.683 | 0.248 | 2.043 | 2.801 |
| | LBBB (L) | 0.102 | 0.320 | 0.279 | 32.053 | 0.383 | 1.494 | 0.014 |
| | Paced (P) | 0.025 | 0.159 | 0.112 | 17.049 | 0.386 | 2.012 | 0.006 |
| | Average | 0.069 | 0.248 | 0.187 | 25.094 | 0.387 | 2.407 | 1.402 |
| ABP | TAH | 0.287 | 0.536 | 0.331 | 53.635 | 0.406 | 1.29 | 0.00079 |
| | AF | 0.029 | 0.171 | 0.128 | 17.058 | 0.355 | 1.376 | 0.00025 |
| | SR | 0.067 | 0.259 | 0.170 | 25.948 | 0.248 | 1.645 | 0.00067 |
| | LBBB | 0.051 | 0.227 | 0.174 | 22.713 | 0.557 | 1.299 | 0.00138 |
| | Paced (P) | 0.010 | 0.101 | 0.081 | 10.890 | 0.179 | 1.456 | 0.00015 |
| | Average | 0.089 | 0.259 | 0.177 | 26.049 | 0.349 | 1.413 | 0.00065 |
| CVP | TAH | 0.178 | 0.422 | 0.345 | 42.261 | 0.697 | 8.50 | 0.00036 |
| | AF | 0.144 | 0.379 | 0.304 | 37.973 | 0.821 | 6.743 | 0.00065 |
| | SR | 0.254 | 0.504 | 0.401 | 50.41 | 0.863 | 1.494 | 0.00222 |
| | LBBB | 0.332 | 0.576 | 0.465 | 57.669 | 0.866 | 9.648 | 0.01692 |
| | Paced (P) | 0.142 | 0.376 | 0.301 | 53.359 | 0.426 | 3.507 | 0.00144 |
| | Average | 0.210 | 0.451 | 0.363 | 48.334 | 0.735 | 5.978 | 0.00432 |
| **Total Average** | | **0.123** | **0.319** | **0.242** | **33.159** | **0.490** | **3.266** | **0.46899** |

**TABLE 2.** **The metrics obtained when training the MC-WGAN with our dataset. Our approach involved generating all signals—ECG, ABP, and CVP—simultaneously for each class using the MC-WGAN model. The table categorizes and presents the results for each signal type across the various classes generated.**

| Class | MSE | RMSE | MAE | PRD | FD | DTW | MMD |
|---|---|---|---|---|---|---|---|
| N | 0.016 | 0.128 | 0.105 | 12.854 | 0.162 | 0.668 | 0.0004 |
| A | 0.038 | 0.183 | 0.145 | 18.336 | 0.262 | 1.938 | 0.0030 |
| L | 0.039 | 0.198 | 0.125 | 19.842 | 0.164 | 1.221 | 0.0008 |
| R | 0.026 | 0.169 | 0.119 | 17.034 | 0.204 | 0.920 | 0.0006 |
| V | 0.119 | 0.345 | 0.234 | 34.609 | 0.333 | 2.028 | 0.0206 |
| Avg | 0.047 | 0.205 | 0.146 | 20.535 | 0.225 | 1.355 | 0.0051 |

**TABLE 3.** **The metrics obtained when training the MC-WGAN with the MIT-BIH arrhythmia database.**

though there is still room for improvement in reducing RMSE and PRD further.

## V. DISCUSSION AND CONCLUSION

The evaluation of our MC-WGAN method demonstrates its adeptness in generating synthetic ECG and ABP signals with an accuracy that aligns well with prevailing benchmarks in the literature. The measured metrics—MSE, RMSE, MAE, PRD, FD, DTW, MMD—indicate a close approximation to actual physiological signals, highlighting the model's capability in effectively mimicking real-world data. The performance varies with the type of signal being synthesized. ECG and ABP signals, characterized by their repetitive patterns, are the easiest to reproduce. However, generating CVP signals is more challenging, as evidenced by the elevated values in the same metrics. This challenge is largely due to the intrinsic complexity and variability of CVP signals, influenced by a range of dynamic physiological factors. Despite these challenges, the model exhibits a promising ability to generate simultaneous multimodal signals and learn the underlying patterns and dynamics essential for patient monitoring and diagnostic applications.

The capacity of our MC-WGAN to simultaneously process and generate multiple signal types marks a significant advancement for clinical applications, especially in scenarios requiring comprehensive monitoring. This capability not only enhances the relevance of the generated data for more holistic simulations and diagnostics but also represents a major leap in understanding the interdependencies and correlations between various biomedical signals. Access to different types of signals from a patient reduces the risk of misdiagnosis and improves the ability of diagnostic tools to detect patterns present in complex conditions, which often manifest through multiple physiological signals. Furthermore, models based on multimodal signals enable the establishment of patient-specific baselines, facilitating more accurate detection of deviations and leading to personalized diagnostic insights. These advancements are crucial in pushing the boundaries of current technologies and setting new standards in synthetic biomedical signal generation, thereby enhancing diagnostic and prediction tools.

Overall, our findings confirm the efficacy of the MC-WGAN architecture in generating realistic synthetic multimodal biomedical signals. This model's ability to produce ECG, ABP, and CVP signals concurrently sets it apart from existing models, marking a significant progression in the field of medical signal processing. While the model performs admirably with ECG and ABP signals, the generation of CVP signals, although less precise, still represents a critical area of learning and adaptation for the system. However, this study comes with some limitations.

**TABLE 4. Performance comparison of the MC-WGAN model on the MIT-BIH Arrhythmia Database (MIT-BIH) against various existing models documented in the literature. This table highlights the effectiveness of the MC-WGAN model in generating ECG signals, using metrics such as MSE, RMSE, MAE, PRD, FD, DTW, and MMD to demonstrate its performance.**

| Model | Year | Signal | Dataset | MSE | RMSE | MAE | PRD | FD | DTW | MMD |
|---|---|---|---|---|---|---|---|---|---|---|
| BiLSTM-CNN GAN [18] | 2019 | ECG | MIT-BIH | - | 0.215 | - | 51.799 | 0.803 | - | - |
| SynSigGAN [19] | 2020 | ECG | MIT-BIH | - | 0.126 | 0.218 | 6.343 | 0.936 | - | - |
| SynSigGAN [19] | 2020 | PPG | MIT-BIH | - | 0.596 | 0.063 | 5.167 | 0.783 | - | - |
| ECG-Adv-GAN [35] | 2021 | ECG | MIT-BIH | 0.002 | 0.029 | - | - | - | - | - |
| Multichannel LS GAN [20] | 2021 | ECG | NSR | - | - | - | - | - | 3.060 | 0.0057 |
| Multichannel LS GAN [20] | 2021 | ECG | ARR | - | - | - | - | - | 4.010 | 0.0548 |
| Multichannel GAN [32] | 2022 | ECG | PTB-XL; China | 0.024 | - | - | - | 7.237 | - | - |
| LSGAN [21] | 2023 | ECG | MIT-BIH | 0.070 | - | - | - | - | - | - |
| CycleGAN [21] | 2023 | ECG | MIT-BIH | 0.069 | - | - | - | - | - | - |
| Transformer-GAN [26] | 2023 | ECG | MIT-BIH | - | 0.180 | - | 7.10 | 0.660 | - | - |
| 2D WGAN [5] | 2024 | ECG | MIT-BIH | 0.002 | 0.024 | - | - | - | - | - |
| BC-GAN [25] | 2024 | ECG | MIT-BIH | - | 0.235 | - | 33.753 | 0.131 | 10.687 | - |
| Statistical WGAN [23] | 2024 | ECG | MIT-BIH | - | 0.003 | 0.002 | - | - | 17.410 | - |
| MC-WGAN (Ours) | 2024 | ECG | MIT-BIH | 0.047 | 0.205 | 0.146 | 20.535 | 0.225 | 1.355 | 0.0051 |

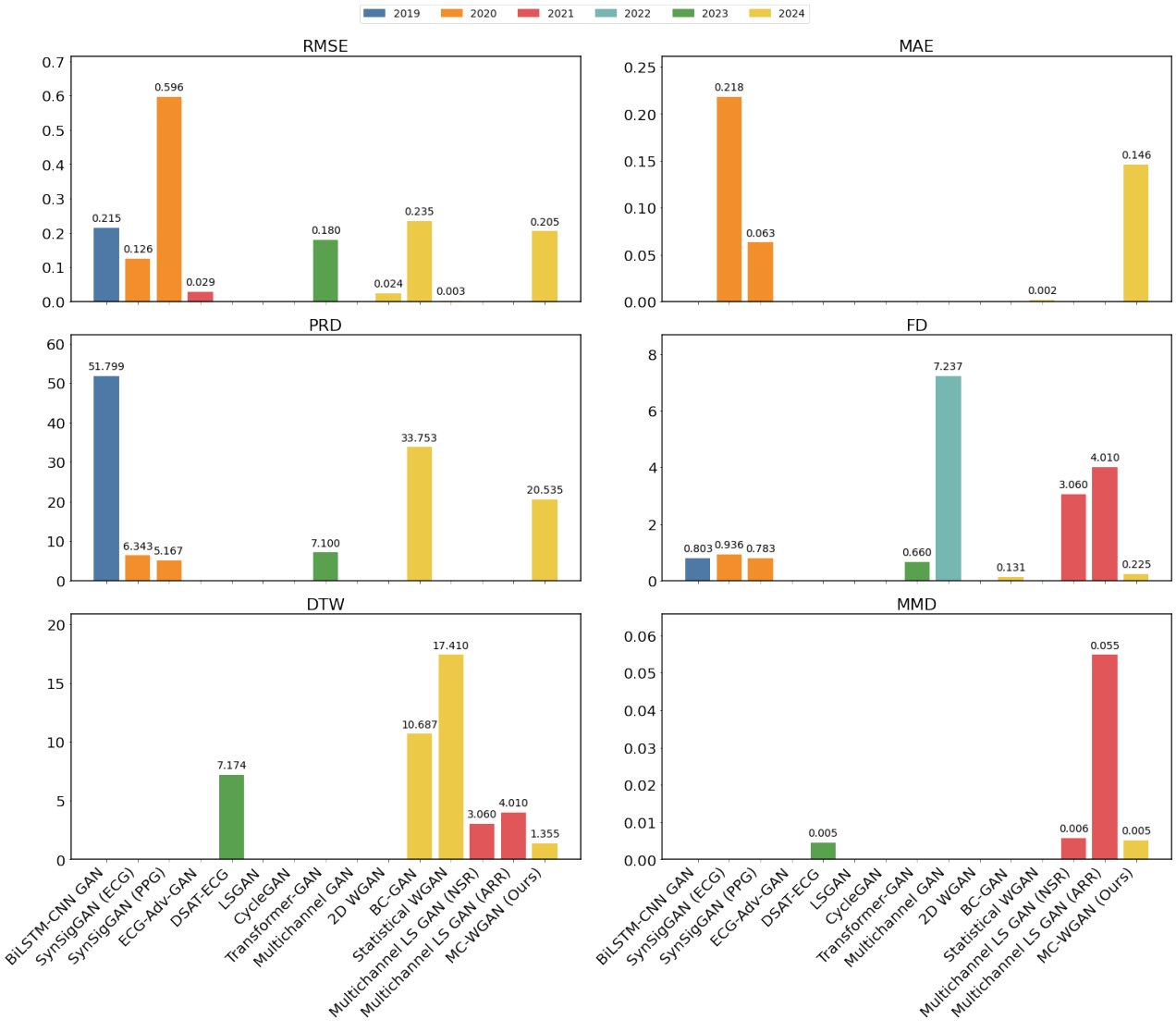



**FIGURE 5. Comparison of our model's performance with existing state-of-the-art models. Here, different colors represent the publication years of each study, while each subplot corresponds to a distinct evaluation metric. The results demonstrate that our model consistently outperforms others across various metrics. Gaps in the plots indicate metrics that were not reported in some studies.**

## VI. LIMITATIONS AND FUTURE WORK

Our study has several limitations. Firstly, it relies on a limited dataset that may not fully represent the diversity of physiological cardiological conditions encountered in clinical practice across different medical centers. This limitation could affect the generalizability of our findings. Additionally, the current focus of our study is on generating short-term signals, which might not be useful in the context of chronic conditions or long-term monitoring scenarios. Moreover, while our model excels with ABP and ECG signals, it struggles with the more complex CVP signals. This highlights the need for further refinement to handle highly variable and intricate physiological data. Furthermore, the practical integration of our model into existing clinical workflows has not been fully explored. Ensuring seamless adoption in real-world healthcare settings will require additional studies and collaborative efforts with clinical professionals.

Future efforts will concentrate on improving the model's capability to produce longer sequences of signals and to handle the generation of CVP signals more effectively. This may involve using enriched training datasets that encompass a broader spectrum of physiological scenarios or developing innovative model architectures to better capture the complex dynamics of such signals. Expanding our capability to produce long-term, continuous data that reflects chronic conditions and long-term monitoring scenarios is a key objective. Incorporating deeper clinical insights into the training process could further enhance the model's performance and reliability.

In conclusion, while our study provides a promising approach to generating physiological signals, addressing these limitations and exploring future work areas will be crucial for advancing the model's applicability and effectiveness in clinical settings.

## ACKNOWLEDGMENT

We would like to thank British Heart Foundation for sponsoring this project (Grant No.FS/19/73/34690).

## REFERENCES

[1] I. Cretu, A. Tindale, M. Abbod, A.W. Khir, W. Balachandran, and H. Meng, "Reliable Multimodal Heartbeat Classification using Deep Neural Networks," *Journal of Biomedical Engineering and Biosciences (JBEB)*, vol. 10, 2023. doi: 10.11159/jbeb.2023.007.

[2] I. Cretu, A. Tindale, M. Abbod, A.W. Khir, W. Balachandran, and H. Meng, "Multimodal Arrhythmia Classification Using Deep Neural Networks," in *Proceedings of the 9th World Congress on Electrical Engineering and Computer Systems and Sciences (EECSS'23)*, London, United Kingdom, 3-5 Aug. 2023, pp. ICBES 152-1 - ICBES 152-8. doi: 10.11159/icbes23.152.

[3] E. Adib, A.S. Fernandez, F. Afghah, and J.J. Prevost, "Synthetic ECG signal generation using probabilistic diffusion models," *IEEE Access*, vol. 11, pp. 75818-75828, 2023. doi: 10.1109/ACCESS.2023.3296542.

[4] J. Wu, L. Wang, H. Pan, and B. Wang, "MLCGAN: Multi-Lead ECG Synthesis with Multi-Label Conditional Generative Adversarial Network," in *ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Rhodes Island, Greece, 2023, pp. 1-5. doi: 10.1109/ICASSP49357.2023.10095035.

[5] T.D. Tran, T.T.K. Dang, and N.Q. Tran, "An Innovative Approach for Long ECG Synthesis with Wasserstein GAN Model," in *Computational Science and Its Applications – ICCSA 2024. Lecture Notes in Computer Science*, vol. 14814, Springer, Cham, 2024. doi: 10.1007/978-3-031-64608-9_22.

[6] K. Vo, E.K. Naeini, A. Naderi, D. Jilani, A.M. Rahmani, N. Dutt, and H. Cao, "P2E-WGAN: ECG waveform synthesis from PPG with conditional Wasserstein generative adversarial networks," in *Proceedings of the 36th Annual ACM Symposium on Applied Computing*, 2021, pp. 1030-1036.

[7] J. Sohn, H. Shin, J. Lee, and H.C. Kim, "Validation of Electrocardiogram Based Photoplethysmogram Generated Using U-Net Based Generative Adversarial Networks," *Journal of Healthcare Informatics Research*, vol. 8, no. 1, pp. 140-157, 2024.

[8] O. Mazumder, R. Banerjee, D. Roy, S. Bhattacharya, A. Ghose, and A. Sinha, "Synthetic PPG Signal Generation to Improve Coronary Artery Disease Classification: Study With Physical Model of Cardiovascular System," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 5, pp. 2136-2146, May 2022. doi: 10.1109/JBHI.2022.3147383.

[9] A. Mahdi, G.D. Clifford, and S.J. Payne, "A model for generating synthetic arterial blood pressure," *Physiological Measurement*, vol. 38, no. 3, pp. 477, 2017. doi: 10.1088/1361-6579/aa51b8.

[10] P.E. McSharry, G.D. Clifford, L. Tarassenko, and L.A. Smith, "A dynamical model for generating synthetic electrocardiogram signals," *IEEE Transactions on Biomedical Engineering*, vol. 50, no. 3, pp. 289-294, 2003. doi: 10.1109/TBME.2003.808805.

[11] G.D. Clifford and P.E. McSharry, "Generating 24-hour ECG, BP and respiratory signals with realistic linear and nonlinear clinical characteristics using a nonlinear model," in *Computers in Cardiology*, 2004, pp. 709-712. doi: 10.1109/CIC.2004.1443037.

[12] T. Golany and K. Radinsky, "PGANs: Personalized Generative Adversarial Networks for ECG Synthesis to Improve Patient-Specific Deep ECG Classification," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, no. 01, pp. 557-564, 2019. doi: 10.1609/aaai.v33i01.3301557.

[13] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative Adversarial Networks," in *Proceedings of Advances in Neural Information Processing Systems*, Z. Ghahramani, M. Welling, C. Cortes, N.D. Lawrence, and K.Q. Weinberger, Eds. Red Hook, NY, USA: Curran Associates, 2014, pp. 2672-2680.

[14] M. Mirza and S. Osindero, "Conditional generative adversarial nets," 2014, in *arXiv preprint arXiv:1411.1784*.

[15] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," *arXiv preprint arXiv:1511.06434*, 2015.

[16] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein generative adversarial networks," in *Proceedings of the International Conference on Machine Learning*, pp. 214-223, 2017.

[17] S. Haradal, H. Hayashi, and S. Uchida, "Biosignal Data Augmentation Based on Generative Adversarial Networks," in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2018, pp. 368-371. doi: 10.1109/EMBC.2018.8512396.

[18] F. Zhu, F. Ye, and Y. Fu, et al., "Electrocardiogram generation with a bidirectional LSTM-CNN generative adversarial network," *Scientific Reports*, vol. 9, p. 6734, 2019. doi: 10.1038/s41598-019-42516-z.

[19] D. Hazra and B. Yung-Cheol, "SynSigGAN: Generative Adversarial Networks for Synthetic Biomedical Signal Generation," *Biology*, vol. 9, no. 12, p. 441, 2020. doi: 10.3390/biology9120441.

[20] E. Brophy, M. De Vos, G. Boylan, and T. Ward, "Multivariate generative adversarial networks and their loss functions for synthesis of multichannel ECGs," *IEEE Access*, vol. 9, pp. 158936-158945, 2021.

[21] S.M. Rayavarapu, T.S. Prasanthi, G.S. Kumar, G.S. Rao, and G. Prashanti, "A generative model for deep fake augmentation of phonocardiogram and electrocardiogram signals using LSGAN and Cycle GAN," *IAPGOS*, vol. 13, no. 4, pp. 34-38, Dec. 2023.

[22] T. Golany, K. Radinsky, and D. Freedman, "SimGANs: Simulator-based generative adversarial networks for ECG synthesis to improve deep ECG classification," in *Proceedings of the International Conference on Machine Learning*, pp. 3597-3606, 2020.

[23] N. Neifar, A. Ben-Hamadou, A. Mdhaffar, M. Jmaiel, and B. Freisleben, "Leveraging Statistical Shape Priors in GAN-Based ECG Synthesis," *IEEE Access*, vol. 12, pp. 36002-36015, 2024. doi: 10.1109/ACCESS.2024.3373724.

[24] T.D. Tran, T.T.K. Dang, and N.Q. Tran, "An Innovative Approach for Long ECG Synthesis with Wasserstein GAN Model," in *Computational Science and Its Applications – ICCSA 2024. Lecture Notes in Computer Science*, vol. 14814, Springer, Cham, 2024. doi: 10.1007/978-3-031-64608-9_22.

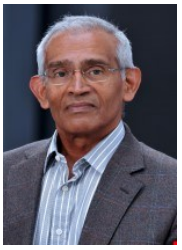
[25] F. Zhou and J. Li, "ECG data enhancement method using generative adversarial networks based on Bi-LSTM and CBAM," *Physiological Measurement*, vol. 45, no. 2, p. 025003, Feb. 2024. doi: 10.1088/1361-6579/ad2218.

[26] H.S. Kaleli and V. Dehalwar, "Generation of Synthetic ECG Signal Using Generative Adversarial Network With Transformers," in *2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, Delhi, India, 2023, pp. 1-6. doi: 10.1109/ICCCNT56998.2023.10307487.
[27] L. Simone and D. Bacciu, "ECGAN: Self-supervised Generative Adversarial Network for Electrocardiography," in *Artificial Intelligence in Medicine. AIME 2023. Lecture Notes in Computer Science*, vol. 13897, Springer, Cham, 2023.
[28] I. Cretu, A. Tindale, M. Abbod, W. Balachandran, M. Mason, A.W. Khir, and H. Meng, "Classification of arrhythmias using an LSTM- and GAN-based approach to ECG signal augmentation," *EP Europace*, vol. 25, no. Supplement 1, June 2023, euad122.622.
[29] J. Sohn, H. Shin, J. Lee, et al., "Validation of Electrocardiogram Based Photoplethysmogram Generated Using U-Net Based Generative Adversarial Networks," *Journal of Healthcare Informatics Research*, vol. 8, pp. 140-157, 2024. doi: 10.1007/s41666-023-00156-z.
[30] V. Thambawita, J.L. Isaksen, and S.A. Hicks, "DeepFake electrocardiograms using generative adversarial networks are the beginning of the end for privacy issues in medicine," *Scientific Reports*, vol. 11, no. 21896, 2021. doi: 10.1038/s41598-021-01295-2.
[31] S. Mahmud, N. Ibtehaz, A. Khandakar, A.M. Tahir, T. Rahman, K.R. Islam, M.S. Hossain, M.S. Rahman, F. Musharavati, M.A. Ayari, and M.T. Islam, "A shallow U-Net architecture for reliably predicting blood pressure (BP) from photoplethysmogram (PPG) and electrocardiogram (ECG) signals," *Sensors*, vol. 22, no. 3, p. 919, 2022.
[32] H.C. Seo, G.W. Yoon, S. Joo, and G.B. Nam, "Multiple electrocardiogram generator with single-lead electrocardiogram," *Computer Methods and Programs in Biomedicine*, vol. 221, 2022. doi: 10.1016/j.cmpb.2022.106858.
[33] Y. Xia, W. Wang, and K. Wang, "ECG signal generation based on conditional generative models," *Biomedical Signal Processing and Control*, vol. 82, 2023, p. 104587. doi: 10.1016/j.bspc.2023.104587.
[34] J. Pan and W.J. Tompkins, "A real-time QRS detection algorithm", *IEEE transactions on biomedical engineering*, 1985, pp.230-236.
[35] K. F. Hossain et al., "ECG-Adv-GAN: Detecting ECG Adversarial Examples with Conditional Generative Adversarial Networks," *20th IEEE International Conference on Machine Learning and Applications (ICMLA)*, Pasadena, CA, USA, 2021, pp. 50-56, doi: 10.1109/ICMLA52953.2021.00016.

**IOANA CRETU** received her B.Sc. degree in Biomedical Engineering from Politehnica University of Bucharest, Romania, in 2019, and her M.Sc. degree with distinction in Biomedical Engineering from Brunel University London, UK, in 2020. That same year, she was awarded a full-time PhD Studentship from the British Heart Foundation. Ioana is currently in the final stages of her PhD at Brunel University London, working in collaboration with Royal Brompton & Harefield Hospital NHS, UK. Her research is centered on the development of a temporary cardiac pacing simulator, designed as a training tool for clinicians.

Throughout her PhD journey, Ioana has published four papers as the first author and has made significant technical contributions to several other publications. She was recognized with the Vice-Chancellor's Conference Prize in both 2022 and 2023 and received the Best Paper Award at The 9th World Congress on Electrical Engineering and Computer Systems and Science in 2023. Additionally, she was recommended by the Alan Turing Institute for a six-month PhD internship at HSBC, where she worked as a Data Scientist on a Speech-to-Text project. Ioana is currently employed full-time at HSBC as a Senior Analyst. Her primary research interests include signal processing, deep learning, and cardiology.

**ALEXANDER TINDALE** is the Senior Cardiology Registrar at Harefield Hospital, London UK. He has Undergraduate degrees in Natural Sciences from Emmanuel College, University of Cambridge, as well as Business Management from the Judge Business School, Cambridge. His then completed his primary medical degree at Imperial College London with distinction. Since then, he has more than a decade's clinical experience as a general and interventional cardiologist. He is currently completing his PhD at Imperial College aimed at optimising haemodynamics in cardiac critical care



**WAMADEVA BALACHANDRAN** (M'91–SM'96-F'04-LF'20) received a B.Sc. degree from the University of Ceylon, Colombo, Sri Lanka, in 1970, and the M.Sc. and Ph.D. degrees from the University of Bradford, Bradford, U.K., in 1975 and 1979, respectively. Professor Balachandran is a Research Professor in the Electronic and Electrical Engineering Department and Director of the Electronic Systems Research Group. He was the Head of the Department of Systems Engineering from 1999 until 2005. Professor Balachandran has made significant contributions to Electronic Systems and Biomedical Engineering. His recent contributions include ECG signal detection using graphene sensors. In 2021, he led a team of three Universities in the UK and developed Vidiia 6, a COVID-19 diagnostic device. Vidiaa Ltd. is currently exploiting this commercially. Currently, he is engaged in developing AI-based early diagnosis of Diabetes type II, the development of a portable device for detecting Cardiovascular Biomarkers using GFET and the development of graphene EEG sensors for monitoring brain activity linked to Epileptic Seizures, Dementia and Alzheimer's diseases. He has published over 400 papers and filed 15 patent applications. He is a recipient of IEEE John Melcher Best Paper Award for Innovation and Creativity in 2005. Recipient of the Lifetime Achievement Award by the Electrostatic Society of America in 2017. His national and international esteem is evident from the accolades and awards he has received. Professor Balachandran's is consulted for his interdisciplinary engineering expertise by industries worldwide. His research has been featured in BBC World Service, TV broadcasts.

**MAYSAM ABBOD** (FIET, CEng, SMIEEE, SFHEA) received BSc degree in Electrical Engineering from University of Technology in 1987. PhD in Control Engineering from University of Sheffield in 1992. From 1993 to 2006 he was with the Department of Automatic Control and Systems Engineering at the University of Sheffield as a research associate and senior research fellow. In 2006, he has moved to Brunel University as a lecturer and progressed in his career. Currently, he is a Professor of Artificial Intelligence at the Brunel University London. His research interest is AI and its applications.

**ASHRAF KHIR** has carried out his PhD work (1996-2000) with Prof Kim H Parker in the Deartment of Bioengineering at Imperial College London and his Thesis dealt with the effect of aortic occlusion on the hemodynamics in the ascending aorta and left ventricular performance. The In Vivo experiments were carried out in the late Prof John V Tyberg's laboratory at Calgary Medical Centre and the human data were collected at the Derek Gibson Echocardiography Department at the Royal Brompton Hospital. Ashraf then spent 3 years (2000-2003) as a postdoctoral fellow at the National Heart and Lung Institute (NHLI), with Prof John R Pepper, studying the mechanics of the Intra Aortic Balloon Pump (IABP). Ashraf obtained his first tenure post as a lecturer at Brunel University London in 2003, and remained there for 18 years, during which he was promoted to Chair in Cardiovascular Mechanics in 2016 and was the Biomedical Engineering Theme Leader. In January 2022, Ashraf moved to Durham University as the Bioengineering Node Leader and his research is generally in the area of arterial and cardiac (cardiovascular) mechanics.

**HONGYING MENG** (M'10–SM'17) is a Professor with College of Engineering, Design and Physical Sciences, Brunel University London, UK. He received his Ph.D. degree in Communication and Electronic Systems from Xi'an Jiaotong University and has research interests in signal processing, machine learning, affective computing, computer vision and embedded systems. He is an associate editor for IEEE Transactions on Circuits and Systems for Video Technology (TCSVT) and IEEE Transactions on Cognitive and Developmental Systems (TCDS), and the associate Editors-in-Chief for Digital Twins and Applications (IET). He has published over 200 journal and conference papers with more than 6800 citations.