Journal of Information Systems and Telecommunication
Vol.14, No.1, January-March 2026, 79-98
http://jist.acecr.org
ISSN 2322-1437 / EISSN:2345-2773

ABC-GAN: An Attention-Based Conditional GAN with DTW Loss for ECG Signal Generation to Address Class Imbalance

Parmida Behain[1*] , Noushin Riahi[1]

[1].Department of Computer Engineering, Alzahra University, Tehran, Iran

Received: 13 Sep 2025/ Revised: 04 Jun 2026/ Accepted: 19 Jul 2026

Abstract

The imbalance between normal and pathological cases in Electrocardiogram (ECG) datasets significantly degrades the performance of automated deep learning-based diagnostic systems, particularly for minority classes. This paper introduces ABC-GAN, a novel Attention-Based Conditional Generative Adversarial Network designed to mitigate this critical data imbalance by generating high-fidelity, class-specific synthetic ECG signals. Our proposed model incorporates two key innovations: an attention mechanism within the generator to focus on critical morphological features of the ECG waveform, and the integration of Dynamic Time Warping (DTW) as a distance metric in the generator's loss function to better preserve essential temporal dynamics. Trained and thoroughly evaluated on the MIT-BIH Arrhythmia dataset across seven different heartbeat classes, ABC-GAN successfully generates highly realistic signals that closely match the original data's distribution. When used to augment the training set, these synthetic signals significantly enhance classifier performance and generalization. A downstream classifier achieved an accuracy of 95%, representing a substantial improvement over the baseline trained on the raw imbalanced data and clearly outperforming both standard oversampling techniques and a vanilla Conditional GAN (cGAN). The overall findings demonstrate that ABC-GAN is a powerful and effective tool for data augmentation, fully capable of improving diagnostic accuracy in cardiac research and practical clinical applications.

Keywords: ECG Signal Synthesis; Generative Adversarial Network (GAN); Attention Mechanism; Conditional GAN; Dynamic Time Warping (DTW); Class Imbalance; Data Augmentation.

1- Introduction

The electrocardiogram (ECG) is a fundamental diagnostic tool that records the heart's electrical activity, enabling the detection of cardiac abnormalities, functional disorders, and arrhythmias. The accurate and automated analysis of ECG signals is therefore critical for proactive cardiac care. However, the performance of deep learning models for this task is heavily reliant on large, balanced datasets. In practice, ECG datasets are notoriously imbalanced, as normal heartbeats vastly outnumber pathological ones, and some arrhythmias are exceptionally rare. This imbalance causes classifiers to be biased toward the majority class, leading to poor generalization and reduced sensitivity for detecting critical cardiac events [1][2].

This challenge of analyzing complex physiological signals is not unique to ECG data. The broader field of medical informatics has seen a significant rise in the application of deep learning to various types of biomedical data. Recent research has explored areas such as the remote monitoring of cardiac conditions [3], deep learning-based analysis of cardiac MRI [4], and stress detection from physiological signals [5]. These advancements motivate further exploration of sophisticated deep learning models, particularly generative models, to address inherent data limitations such as class imbalance in biomedical signal analysis.

Generative Adversarial Networks (GANs) have emerged as a promising solution to this problem, capable of synthesizing realistic synthetic data to augment minority classes [6]. While previous studies have successfully applied GANs for ECG generation [7]–[9], they often rely on standard loss functions like Mean Squared Error (MSE)

✉ **Parmida Behain**
behain.parmida@gmail.com

or adversarial loss alone. These losses can lead to overly smoothed outputs that fail to capture the precise morphological nuances, such as the shape and duration of the P-wave, QRS complex, and T-wave, that are essential for accurate diagnosis. Furthermore, while conditional GANs (cGANs) can generate class-specific signals [9], ensuring the morphological fidelity of these generated signals, especially for rare classes, remains a significant challenge.

In this paper, we propose the Attention-Based Conditional GAN (ABC-GAN), a novel framework designed to address these limitations. Our work makes the following key contributions:

1. We introduce an attention mechanism into the generator of a cGAN, enabling the model to dynamically focus on the most diagnostically relevant segments of the ECG waveform during synthesis.
2. We innovatively incorporate the Dynamic Time Warping (DTW) distance metric into the generator's loss function. DTW is particularly suited for ECG signals as it accounts for temporal misalignments and variations in phase, ensuring the generated signals are morphologically faithful to real heartbeats.
3. We demonstrate that signals generated by ABC-GAN effectively balance the MIT-BIH Arrhythmia dataset. Training a classifier on this augmented dataset leads to a significant improvement in accuracy, particularly for minority classes, outperforming classifiers trained on data augmented by other standard methods.

Unlike conventional ECG signal generation methods that rely solely on convolutional architectures, the proposed ABC-GAN integrates an attention mechanism within a conditional GAN framework to enhance the modeling of long-range temporal dependencies in ECG signals. Existing approaches such as CardioGAN and AC-WGAN-GP primarily focus on adversarial learning and class conditioning but lack explicit mechanisms for selectively emphasizing diagnostically relevant waveform regions.

The proposed method further incorporates Dynamic Time Warping (DTW) loss into the generator objective. While conventional GAN losses focus on distribution matching, DTW provides temporal alignment between generated and real ECG signals, preserving clinically significant morphological characteristics. This is particularly important for arrhythmia classes exhibiting temporal variability.

The main contributions of this study are:

1. Development of an Attention-Based Conditional GAN (ABC-GAN) for ECG signal generation.
2. Integration of an attention module to capture long-range temporal dependencies and emphasize diagnostically relevant waveform segments.
3. Incorporation of DTW loss to improve temporal alignment and waveform fidelity.
4. Generation of synthetic minority-class ECG samples to alleviate class imbalance.
5. Comprehensive evaluation against conventional GAN-based ECG augmentation methods on the MIT-BIH Arrhythmia Database.

The remainder of this paper is organized as follows. Section 2 reviews related work on ECG synthesis. Section 3 provides the necessary background on cGANs and attention mechanisms. Section 4 details the architecture of our proposed ABC-GAN and the experimental setup. Section 5 presents and discusses the results. Finally, Sections 6 and 7 conclude the paper and outline future research directions.

2- Related Works

The use of generative models, particularly Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs), for synthesizing ECG signals has been extensively explored to address data scarcity and imbalance. Previous research can be categorized based on architectural innovations and application goals.

A significant body of work has focused on adapting established GAN architectures for ECG data. Edmond Adib et al. [2] utilized WGAN-GP and AC-WGAN-GP for class-conditioned heartbeat generation, employing a post-generation filtering threshold. Similarly, Fei Zhu et al. [6] and Yu-He Zhang et al. [7] leveraged recurrent architectures (BiLSTM) in the generator to capture temporal dependencies in single and 12-lead ECG signals, respectively. Others, like Naren Wulana et al. [8], have explored architectures like WaveNet and SpectroGAN for generating specific heartbeat types. These approaches demonstrated the feasibility of GANs for ECG synthesis but often relied on standard loss functions that can blur critical features.

To enhance utility, several studies have integrated generation with downstream classification tasks. Pu Wang et al. [9] and Wenqiang Li et al. [10] proposed frameworks combining ACGAN-based data augmentation with CNN classifiers for anomaly and myocardial infarction detection, showing improved performance by leveraging synthetic data.

Addressing mode collapse and improving diversity has been another research focus. Haixu Yang et al. [11] proposed ProEGAN-MS to generate multi-scale ECG signals with high diversity. Subhrajyoti Dasgupta et al. [12] introduced CardioGAN, which incorporated an attention-based generator and a gradient penalty loss, marking a notable step towards focusing on salient waveform features. This is the closest to our use of attention; however, their work did not explore a specialized loss function like DTW to further enforce morphological accuracy.

Rohan Banerjee et al. [13] presented a GAN-based model for generating ECG signals of Atrial Fibrillation (AF), a type of cardiac arrhythmia. Their GAN architecture includes a combination of GAN pairs to simulate Heart Rate Variability (HRV) patterns and unique signal morphology. Results show significant improvement in AF classification performance with the addition of generated data to both datasets.

Recent advances have incorporated state-of-the-art architectures. Li et al. [14] and E. Brophy et al. [15] employed Transformer-based generators for multivariate time-series synthesis, showcasing their ability to model long-range dependencies. Shota Harada et al. [16] proposed a GAN-based method for time series data generation, with LSTM-based generator and discriminator networks. They demonstrated improved classification performance with the addition of generated data. Jintai Chen et al. [17] presented the ME-GAN model for generating ECG signals, incorporating Mixup normalization to inject disease information appropriately. Their evaluation using rFID metrics shows the model's effectiveness in generating Multi-view ECG signals.

Other efforts have tackled the challenge of generating variable-length signals [18] or using VAEs for feature-space generation [19].

A significant branch of research focuses on improving GAN training dynamics for ECG data. Munia et al. [21] leveraged the Wasserstein GAN (WGAN) to stabilize training on one-dimensional signals. Building on this, Nankani et al. [22] introduced the DCCGAN, which incorporated a 'Twist' parameter in the discriminator to mitigate vanishing gradients and accelerate convergence, specifically for generating minority class samples. Conversely, Golany et al. [20] addressed the issue of label scarcity directly with their Personalized GAN (PGAN), which generates patient-specific signals without relying on labeled arrhythmia data, enhancing performance for individualized models.

Hatamian et al. [23] demonstrated that for ventricular fibrillation classification, GAN-based augmentation (using DCGAN) significantly outperformed conventional oversampling techniques, a finding reinforced by qualitative assessments of morphological similarity.

Further innovations involve the structural decomposition of the generation task. Agarwal et al. [24] proposed a two-module framework (CardiacGen), where separate WGAN-GP models dedicated to generating heart rate variability and morphological features are combined to produce highly realistic signals. In one of the most complex approaches, Adib et al. [25] abandoned the 1D domain entirely, mapping ECG signals to 2D representations (using GASF/GADF/MTF) for processing with a Denoising Diffusion Probabilistic Model (DDPM), before mapping back to 1D; notably, their results indicated that a well-tuned WGAN-GP model could still outperform this sophisticated DDPM-based approach on certain metrics.

Recent advances in ECG synthesis have shifted from traditional GAN-based architectures toward transformer, diffusion, and hybrid self-supervised frameworks that better capture long-range temporal dependencies and improve physiological realism.
Class imbalance remains a critical issue in ECG analysis, particularly for rare arrhythmia classes. To address this, CECG-GAN introduces a conditional generative adversarial framework that augments minority ECG classes using class-conditioned synthesis. By improving the distribution of underrepresented rhythms, the model enhances downstream diagnostic robustness and classification sensitivity for rare cardiac conditions [26].

Incorporating task-specific constraints into generative modeling has also gained attention. The Task-Aware Conditional GAN introduces a multi-objective optimization strategy combining adversarial loss with temporal alignment constraints such as Dynamic Time Warping (DTW), reconstruction loss, and statistical distance measures. This enables the generator to preserve both global distributional characteristics and fine-grained temporal structure, which is particularly important for ECG signals where waveform alignment (e.g., R-peak consistency) is clinically significant [27].

More recently, diffusion-based generative models have demonstrated superior stability and sample quality compared to GAN-based approaches. BioDiffusion extends diffusion modeling to biomedical time-series synthesis, providing improved coverage of complex signal distributions and mitigating mode collapse issues commonly observed in GANs. Its probabilistic denoising formulation allows more faithful reconstruction of physiological variability in ECG signals [28].

Personalized ECG synthesis has also emerged as a key research direction. The Personalized 12-lead ECG generation framework conditions the generative process on patient-specific attributes, enabling individualized cardiac waveform synthesis. This improves clinical relevance by preserving subject-specific morphology while maintaining consistency across leads [29]. Hybrid self-supervised generative approaches further improve representation learning in low-label scenarios. ECG synthesis models combining self-supervised pretraining with adversarial generation improve feature extraction from limited ECG datasets, leading to more stable training and higher-quality pathological waveform generation [30].

Transformer-based diffusion models such as TransDiffECG integrate attention mechanisms with diffusion processes to enable semantically controllable ECG synthesis. This allows explicit conditioning on cardiac attributes while maintaining high fidelity in waveform generation, demonstrating strong performance in both generation quality and downstream diagnostic tasks [31]. Finally, RDiffGAN-ECG combines diffusion models, adversarial training, and state-space sequence modeling to enhance long-range dependency modeling in ECG signals. By explicitly conditioning on R-peak information and leveraging structured temporal modeling, it produces high-fidelity ECG waveforms with improved morphological consistency and heart-rate realism [32].

2-1- Limitations of Existing Methods

1. **Morphological Smoothing from MSE-based Losses.** Standard GANs and cGANs for ECG generation [6][9] predominantly rely on Mean Squared Error (MSE) or adversarial loss alone. MSE produces overly smooth outputs by averaging multiple plausible reconstructions, leading to loss of sharp morphological features such as precise QRS complex width and ST-segment elevation patterns that are clinically significant [2][12]. ABC-GAN overcomes this by replacing MSE with Dynamic Time Warping (DTW) loss, which preserves sharp transitions by matching temporal patterns rather than pointwise differences.

2. **Temporal Misalignment Insensitivity.**
   Conventional loss functions treat each time point independently, failing to penalize small temporal shifts that preserve overall shape but misalign critical landmarks (P-wave onset, R-peak, T-wave offset). This is particularly problematic for ECG signals where beat-to-beat variability is natural [13][41]. DTW loss explicitly optimizes for optimal temporal alignment, making the generator invariant to benign temporal variations while penalizing morphological distortions.

3. **Insufficient Focus on Diagnostically Relevant Regions.**
   While CardioGAN [12] introduced attention for ECG generation, their attention mechanism operates without temporal alignment guidance, potentially focusing on irrelevant segments. ABC-GAN combines self-attention with DTW loss, where the attention weights are indirectly optimized to focus on regions that minimize temporal alignment cost, directing attention to diagnostically critical waveform components.

4. **Mode Collapse in Minority Classes.**
   Standard GANs often fail to capture the full distribution of rare arrhythmia classes [2]. Our conditional architecture with class-specific attention maps and DTW loss for each class separately ensures that minority class distributions are modeled with higher fidelity.

Most models optimize using losses like MSE or adversarial loss, which are insufficient for capturing complex, non-linear temporal alignments and precise morphological patterns in ECG waveforms. The use of Dynamic Time Warping (DTW), a powerful similarity measure for time series that is invariant to non-linear warps in the time dimension, as a loss function for training ECG GANs remains largely unexplored.

Our work, ABC-GAN, bridges this gap. While building upon the foundation of conditional generation [2][6] and inspired by the use of attention [12], we uniquely integrate a DTW-based loss into the adversarial training regimen. This combination ensures that the generated signals are not only realistic and class-specific but also morphologically precise, addressing the core challenge of fidelity in ECG synthesis for imbalanced learning.

3- Background

This section outlines the fundamental concepts underlying our proposed model: Conditional Generative Adversarial Networks and the Attention Mechanism.

3-1- Conditional GAN

The Conditional Generative Adversarial Network (CGAN), introduced by Mirza and Osindero [33], is an extension of the original GAN framework [34]. While a standard GAN learns to generate data from a random noise vector z, a CGAN conditions the generation process on additional information y, such as a class label. This allows for the targeted generation of data from specific classes.

The generator G and discriminator D play a two-player minmax game with the following objective function:

$$\min_G \max_D V(D,G) = E_{x\sim p_{data}(x)}[\log D(x|y)] + E_{z\sim p_z(z)}\left[\log\left(1 - D\left(G(z|y)\right)\right)\right] \quad (1)$$

where x is real data, z is the noise vector, and y is the conditional label. The discriminator learns to distinguish between real pairs (x,y) and fake pairs $(G(z|y),y)$, while the generator learns to fool the discriminator by producing outputs that are realistic for the given label y.

In the context of ECG synthesis, CGANs enable the generation of heartbeat signals for a specific arrhythmia class (PVC, APC, LBBB). However, a known limitation of standard CGANs is that they often produce samples that lack fine-grained morphological details, as the standard adversarial and reconstruction losses (MSE) may not sufficiently penalize small temporal misalignments or shape inaccuracies crucial for ECG diagnosis.

Fig. 1 illustrates the process of conditioning the GAN model, which involves providing class labels (C) as auxiliary information along with input noise (Z) to the generator to produce ECG signals.

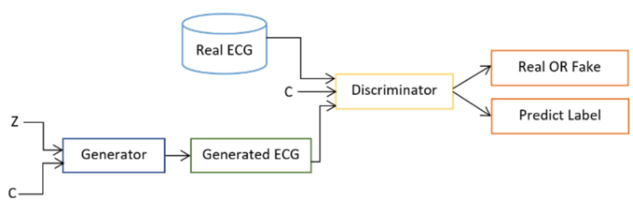


Fig. 1 Conditional GAN

3-2- Attention Mechanism

The attention mechanism, a cornerstone of the Transformer architecture [35] allows a model to dynamically focus on different parts of the input sequence when producing an output. It has since been successfully integrated into GANs (SAGAN [36])to capture long-range dependencies in images.

The core component is the Scaled Dot-Product Attention. It operates on learnable linear projections of the input: Query (Q), Key (K), and Value (V). The output is a weighted sum of the values, where the weight assigned to each value is determined by the compatibility of the query with the corresponding key:

$$\text{Attention}(Q,K,V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (2)$$

where dk is the dimensionality of the keys. The scaling factor $\frac{1}{\sqrt{d_k}}$ prevents the softmax function from having extremely small gradients.

For ECG generation, self-attention enables the model to learn intra-signal dependencies, focusing on critical waveform regions (P-wave onset, R-wave peak, ST segment) during generation. We hypothesize that this focused generation is crucial for morphological fidelity and works synergistically with DTW loss.

Fig. 2 visualizes the Waveform Attention Mechanism employed in the model architecture.

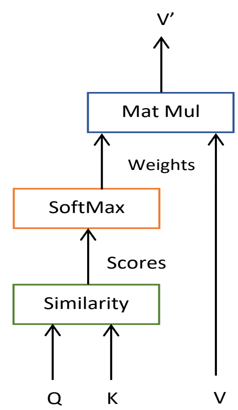


Fig. 2 Illustration of Waveform Attention Mechanism

In our work, we integrate a self-attention block [36] into the generator of the CGAN. For ECG generation, this mechanism allows the model to learn intra-signal dependencies, effectively focusing its "attention" on critical regions of the waveform (the onset of the P-wave, the peak of the R-wave, or the duration of the ST segment) while generating each point of the output signal. We hypothesize that this focused generation is crucial for achieving high morphological fidelity and works synergistically with our DTW-based loss function, which directly optimizes for temporal alignment.

4- Materials and Methods

The proposed method consists of four main steps: preprocessing, GAN training, post-processing, and data classification. Initially, real ECG signals from the dataset are preprocessed to prepare them for GAN training.

Following this, the GAN is trained on the preprocessed data to generate artificial ECG signals. Post-processing is then applied to the generated signals to refine them. Finally, the generated signals are validated to ensure high accuracy and similarity to the original signals, measured using appropriate quantitative metrics (DTW, MSE).
Fig. 3 illustrates the block diagram of the proposed method and its four main steps.

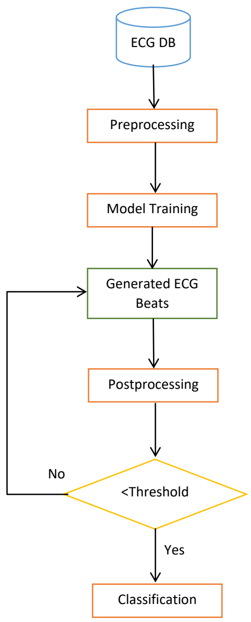


Fig. 3 Block diagram of the proposed method

4-1- Preprocessing

Preprocessing is a crucial step to prepare the dataset for GAN training. From the provided 15-class dataset, a subset of 7 classes is selected. Using the R-peak locations specified in the dataset, intervals to the preceding and succeeding R-peaks are determined. Segments corresponding to 75% of these intervals before and after each R-peak are extracted to define the boundaries of individual heartbeats. Additionally, the sampling frequency is reduced from 360 Hz to 256 Hz to decrease computational load and training time [2].

4-2- Train Attention-based Conditional GAN Model

The Wasserstein Generative Adversarial Network (WGAN) is an extension of the traditional GAN that replaces the discriminator with a critic to estimate the Wasserstein distance between real and generated samples. This modification stabilizes training and makes it less sensitive to model architecture and hyperparameter choices [37].
By conditioning the WGAN on class labels, the generator learns class-specific features of ECG signals, which are then used to generate realistic synthetic signals. The architecture of the proposed model and its loss function are described in the following sections.

1) Model Architecture

The proposed method aims to enhance the conditional WGAN model to generate ECG signals with higher fidelity, thereby improving classification accuracy. To achieve this, three main strategies are adopted:

1. **Attention Mechanism:** Incorporated into the generator to focus on critical temporal and morphological features of ECG signals, enabling the capture of class-specific patterns.
2. **Generator Loss Function Modification:** Different loss functions, including MSE, DTW, and Kullback-Leibler divergence, were evaluated. DTW was ultimately selected for its superior performance in generating signals that improve classification accuracy.
3. **WGAN Architecture:** The discriminator is replaced with a critic, which stabilizes training by estimating the Wasserstein distance between real and generated signals.

Below, the proposed architectures for the generator, critic, and the proposed loss function in the generator are provided.
The architecture of the generator and critic utilizes the blocks specified in Table 1.

Table 1: Blocks in the generator and critic

Layer	*Generator*	*Critic*
1	ConvTranspose1d[1]	ConvTranspose1d[2]
2	BatchNorm1d	InstanceNorm1d
3	ReLU	LeakyReLU

The input noise depicted in Table 2 has a length of 100 with dimensions of 100×1. Additionally, class labels are connected to it through a mapping to dimensions of 100×1. Therefore, the generator receives input of dimensions 200×1. An attention mechanism is incorporated into the architecture to enhance the generator's learning for better ECG signal generation. This mechanism assists the generator in learning more precise dependencies using attention mechanisms, allowing it to better capture the patterns and characteristics specific to each class of ECG signals.
In WGANs, the generator network utilizes CNN layers to transform input noise vectors into synthetic ECG waveforms that closely resemble real ECG signals. These CNN layers allow the generator to capture temporal and

[1] kernel size = 4, stride = 2, padding = 1
[2] kernel size = 4, stride = 2, padding = 1

spatial features inherent in ECG signals, producing realistic synthetic waveforms. The generator's ability to learn from the underlying distribution of ECG signals is crucial for generating accurate and clinically relevant synthetic data.
The architecture of the generator is shown in Table 2:

Table 2: Architecture of the generator

Layer	*Generator*
Input	16 × 200 × 1
1	block[1]
2	block
3	block
4	block
5	ConvTranspose1d[2]
6	FC
7	Tanh
8	Attention layer
Output	16 × 1 ×256

The critic network in WGANs, which replaces the discriminator in traditional GANs, also employs CNNs. The critic evaluates both generated and real ECG signals to estimate the Wasserstein distance between their distributions. By leveraging CNNs, the critic can effectively learn temporal and spatial patterns present in ECG signals, providing more precise assessments of the synthetic signal quality. In the critic, class labels are mapped to dimensions of 1×256. Subsequently, they are connected to either the generated or real heartbeat signal with dimensions of 1×256. The input to the critic is of dimensions 2×256.
The architecture of the critic is shown in Table 3:

Table 3: Architecture of critic

Layer	*Critic*
Input	16 × 2 ×256
1	Conv1D, Leaky ReLU[3]
2	block[4]
3	block
4	block
5	Conv1D[5]
6	FC
7	FC
Output	16× 1 × 1

[1] The architecture of the blocks is presented in **Error! Reference source not found.**
[2] kernel size=4, stride=2, padding=1
[3] kernel size=4, stride=2, padding=1
[4] The architecture of the blocks is presented in **Error! Reference source not found.**
[5] kernel size=4, stride=2, padding=0

2) Proposed Loss Function

The critic's loss function comprises both the Wasserstein loss and the gradient penalty, defined as follows [2]:

$$L_{critic} = - E_{(x,y) \sim p_{data}} [critic(x, y)] + E_{(\tilde{x},y) \sim p_{fake}} [critic(\tilde{x}, y)] + \lambda_{GP}\, gp \tag{3}$$

This loss function in Eq.(3) is used to update the critic's parameters. Both the generator and critic loss functions incorporate conditional information and class labels. The term mean(critic(real, labels)) refers to the average output of the critic when evaluating real data. On the other hand, mean(critic(fake, labels)) is the average output of the critic when given fake (generated) data. The difference between these two terms is central to the critic's goal of distinguishing between real and fake samples. The constant λ_{GP} represents the strength of the gradient penalty (GP), which is a regularization term denoted by gp that enforces smooth gradients to stabilize the training process. Together, these terms help the critic model maintain a balance between accurately identifying real and fake data while controlling the gradients for stable training.

The Wasserstein loss function for the critic typically relies on the Wasserstein distance, also known as Earth Mover's Distance (EMD). It measures the dissimilarity between two probability distributions and guides the generator toward producing samples closer to the distribution of real data [37].

$$L_{\text{wasserstein}} = -(E[\text{critic(real,labels)}] - E[\text{critic(fake,labels)}]) \tag{4}$$

In Eq.(4), critic(real, labels) represents the scores assigned by the critic to real samples, and critic(fake, labels) represents the scores assigned to generated samples. The goal is to maximize the difference between the average scores of real and generated samples, encouraging the critic to assign higher scores to real samples and lower scores to generated samples, guiding the generator toward producing more realistic samples.

The gradient penalty component is a crucial part of the WGAN with Gradient Penalty (WGAN-GP) algorithm. This function helps enforce Lipschitz continuity in the discriminator, which is essential for stable training. Lipschitz continuity is a mathematical property often applied to functions, particularly in the context of WGANs, to ensure training stability. The Lipschitz constraint restricts the rate of change of a function [25].

In the case of WGANs, this constraint is applied to the critic function. Lipschitz constraint ensures that the output of the function changes smoothly, preventing abrupt changes. Applying a Lipschitz constraint to the critic helps stabilize the training process by preventing the critic from assigning very high gradients to the generator, thus avoiding erratic behavior during training. This constraint also aids in reducing mode collapse by softening the critic's response, guiding the generator to cover a broader range of data distribution, and reducing the risk of collapsing into a single mode.

The gradient penalty term is given by:

$$\mathrm{gp} = \lambda\,(|\nabla_x \mathrm{critic}(\hat{x},\mathrm{labels})|_2 - 1) \tag{5}$$

In Eq.(5), $\hat{x}$ denotes an *interpolated sample*, which is a convex combination of a real data sample x and a generated sample $\tilde{x}$. It is typically computed as $\hat{x} = \alpha\, x + (1-\alpha)\, \tilde{x}$, where $\alpha \in [0,1]$. The coefficient λ, usually set to 1/2, controls the strength of the gradient penalty. $\nabla_{\hat{x}}$ represents the gradient of the critic's output with respect to the interpolated sample, indicating how the critic's score changes with small variations in the input. Finally, $\|\cdot\|_2$ denotes the L_2 norm, which measures the magnitude of this gradient vector.

The proposed generator's loss function is defined as the negative average score assigned by the critic to the generated samples, augmented with the Dynamic Time Warping (DTW) loss to capture the characteristics and waveforms of the generated signals:

$$L_{\mathrm{gen}} = -E[\mathrm{critic}(\mathrm{fake},\mathrm{labels})] + \lambda_{\mathrm{dtw}} \cdot \mathrm{dtw\_loss} \tag{6}$$

Eq.(6), λdtw is the DTW coefficient. This loss function is used to update the generator's parameters.
Minimizing this function aligns with the generator's objective of receiving high scores from the critic, indicating a higher similarity to the distribution of real data. Minimizing this negative average is equivalent to maximizing the positive average, aligning to generate realistic samples. The ultimate aim is to minimize the generator's loss throughout training.
The DTW function measures the dynamic time-warping distance between real and generated samples, quantifying the temporal alignment and dissimilarity between two sequences:

$$\mathrm{dtw\_loss} = \sum_{i=1}^{\mathrm{cur\_batch\_size}} \mathrm{DTW}(\mathrm{real}[i],\mathrm{fake}[i]) \tag{7}$$

This loss, as shown in Eq.(7) is computed for each pair of real and generated samples within the batch.

4-3- Postprocessing

In this step, generated ECG signals that are highly dissimilar to real signals are filtered out, retaining only morphologically similar signals. The similarity is assessed using the Dynamic Time Warping (DTW) distance between each generated signal and the nearest confirmed real signal from the corresponding class. Signals with a DTW distance exceeding a predefined threshold are removed, ensuring that the retained signals closely resemble the real ECG morphology.

4-4- Classification

The filtered generated ECG signals are then incorporated into the training dataset for classification. A one-dimensional implementation of ResNet34, adapted for ECG signal analysis, is employed for classification [38]. ResNet34 consists of 34 layers with residual blocks, each containing two 3×3 convolutional layers with skip connections [39]. The model is trained on a combined dataset of real and generated signals to evaluate the effectiveness of the generated signals in improving classification accuracy.

5- Results and Discussion

In this section, we first describe the dataset used and then present two approaches to evaluate the generated ECG signals by comparing them with real signals. The first approach involves determining cluster centers for each class, and the second involves examining average signals and standard deviations.

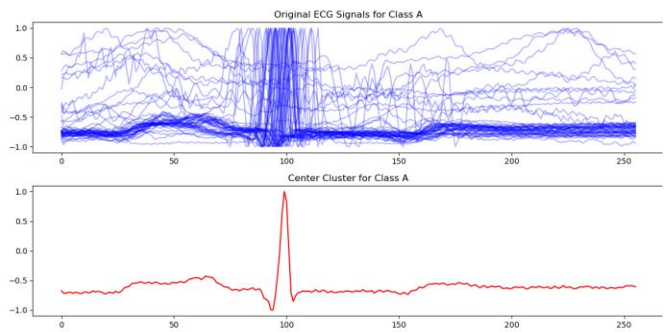
5-1- Dataset

The dataset used for this research is the MIT-BIH Arrhythmia Database [40]. This database, collected by the Massachusetts Institute of Technology (MIT), comprises a set of electrocardiogram (ECG) records widely used for research in arrhythmia detection and analysis. It stands as one of the most recognized and authoritative databases in the field of biomedical signal processing and cardiovascular research.
The dataset includes 48 half-hour two-channel ECG records sampled at a frequency of 360 Hz. These records encompass both normal sinus rhythm and various arrhythmias. The MIT-BIH dataset consists of a total of 15 classes, with only 7 classes selected for this research: P (premature atrial contraction), A (atrial premature contraction), L (left bundle branch block beat), N (normal

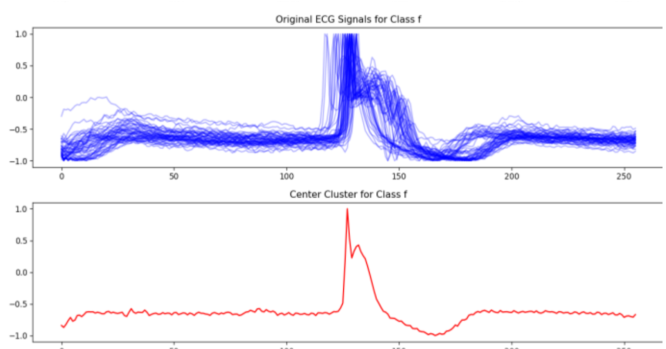
beat), R (right bundle branch block beat), f (fusion of paced and normal beat), j (junctional beat).

5-2- Determining Cluster Centers for Each Class

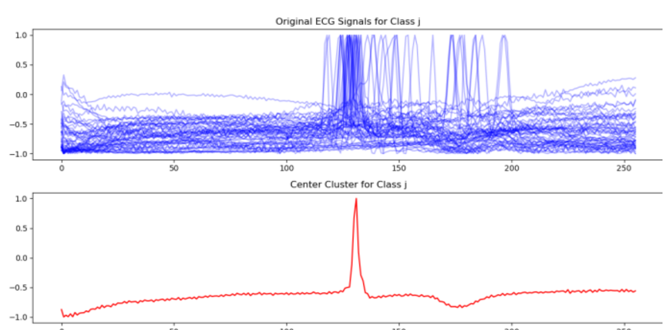
For each class, 50 real signals were selected as representative samples. The cluster center was defined as the signal whose maximum DTW distance to all other signals in the class was minimal. This ensures that the selected signal is the most representative of its class.
Fig. 4 shows the cluster centers and the corresponding 50 sample templates:



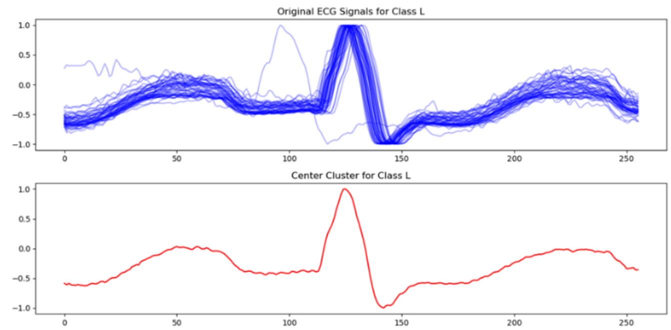
(a)



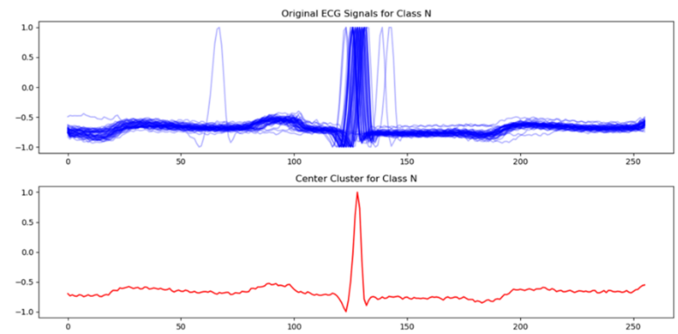
(b)



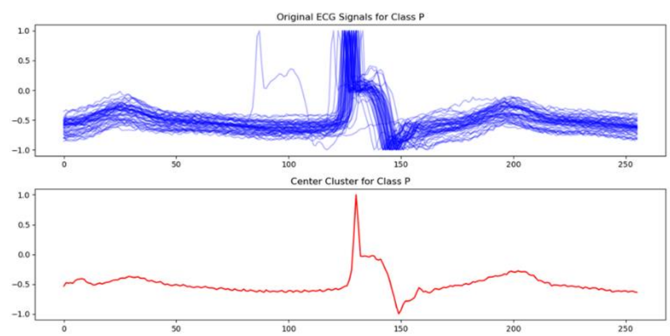
(c)



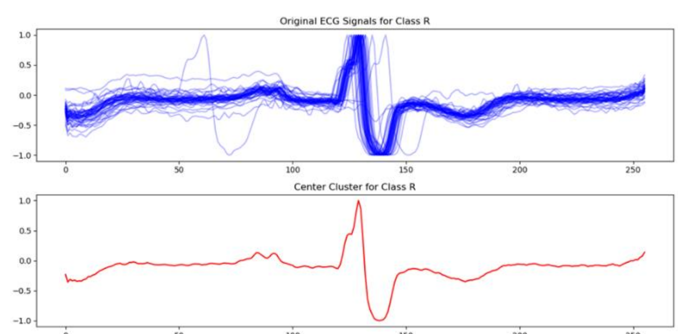
(d)



(e)



(f)



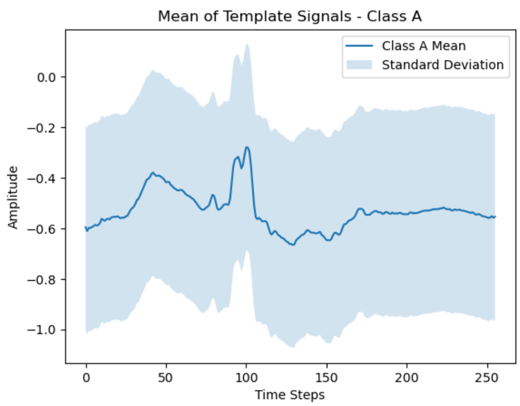
(g)

Fig. 4 Cluster centers of each class: (a) Class A (b) Class F (c) Class J (d) Class L (e) Class N (f) Class P (g) Class R. The red-colored signals represent the cluster center obtained for each class. The blue-colored signals are the same 50 real samples selected from the dataset.

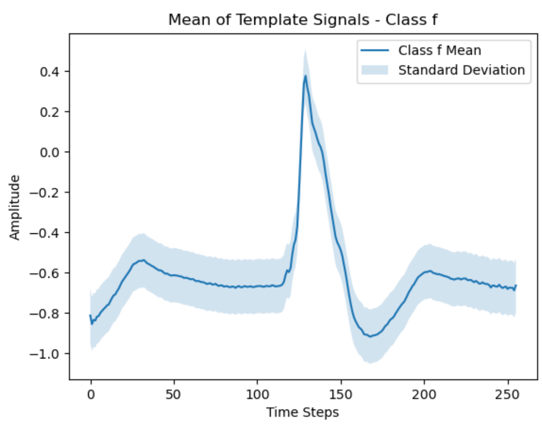
Notably, classes A and j exhibit higher intra-class variability compared to other classes. These cluster centers serve as references for evaluating the similarity of generated signals to real ECG signals.

5-3- Determining Mean Signals and Standard Deviation

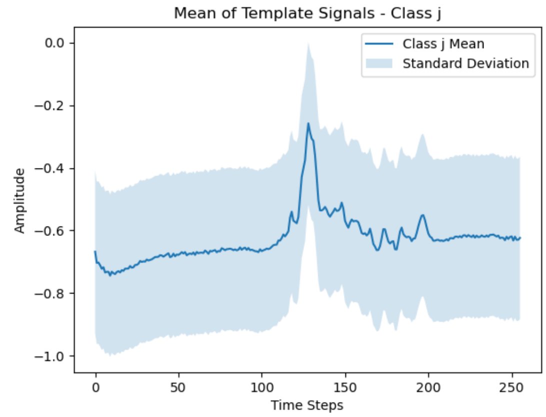
To further characterize the statistical properties of the signals, the mean and standard deviation of the 50 samples in each class were computed (Fig. 5).



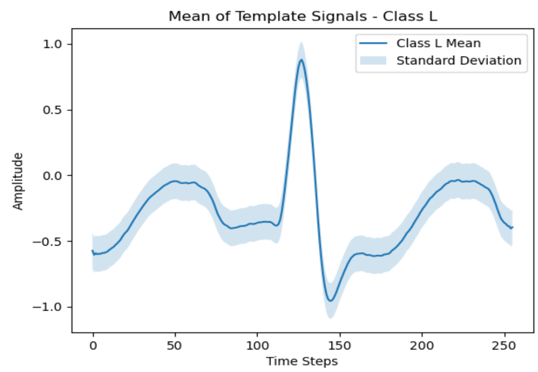
(a)



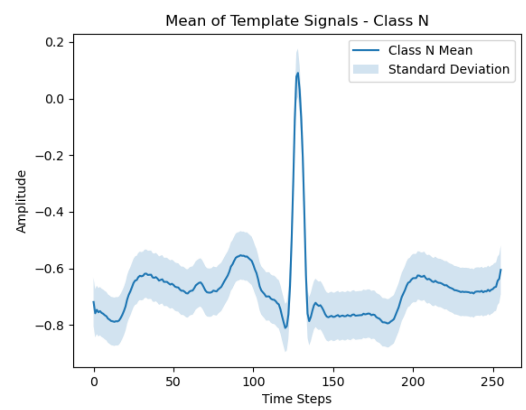
(b)



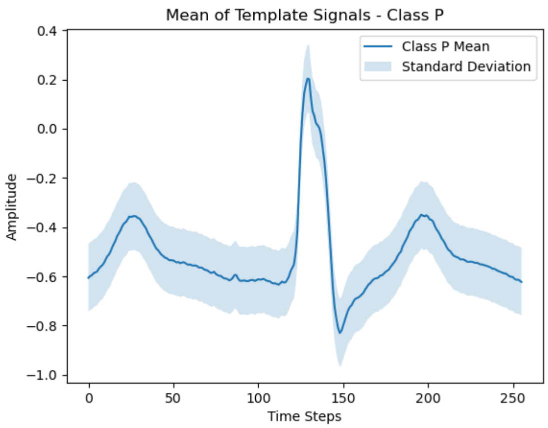
(c)



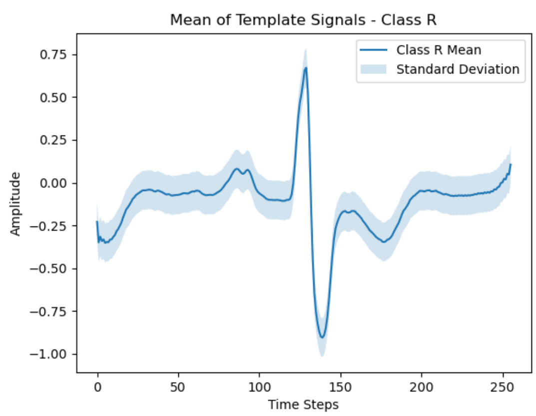
(d)



(e)



(f)



(g)

Fig. 5 Mean signals for each class: (a) Class A (b) Class f (c) Class j (d) Class L (e) Class N (f) Class P (g) Class R. It can also be observed that in classes A and j, a well-defined average signal has not been obtained so the average signal cannot be considered as a template for each class, and there is also a higher standard deviation in these two classes compared to other classes.

While the average signal provides a summary of the class characteristics, it may be insufficient for assessing the similarity of generated signals, particularly in classes with high variability, such as A and j.

5-4- Eliminating outlier signals

As the mean signal may not provide a reliable reference for classes with high variability, we employed Dynamic Time Warping (DTW) to measure the similarity between generated signals and the 50 template signals of each class. For each generated signal, the nearest template based on the minimum DTW distance was identified.

The nearest real signal serves as a robust criterion for evaluating the generated signals and determining outliers. To remove outlier signals, the DTW distance between each generated signal and its closest template was computed. Signals with a DTW distance exceeding a threshold of 3 were considered outliers and excluded from further analysis. This threshold ensures that only generated signals with sufficient morphological similarity to real signals are retained for subsequent classification.

5-5- Evaluating Metrics

We employ two categories of metrics: (1) classification performance metrics to assess downstream task improvement, and (2) similarity metrics to quantify how closely generated signals resemble real ECG signals.

For classification, we use standard metrics including Precision, Recall, Accuracy, and F1-score as defined in standard textbooks [48]. These are computed from the confusion matrix of the ResNet34 classifier on held-out test data.

For similarity assessment, we employ three complementary metrics:

1) **DTW (Dynamic Time Warping):**
A classic approach for measuring the similarity (or dissimilarity) of time series curves is by considering the Dynamic Time Warping (DTW) similarity measure. It calculates the optimal alignment of two given time series with the minimum cost [12]. DTW aligns time series along the time axis to optimally align them and computes the distance between them [41].

$$D_{i,j} = f(x_i, y_j) + \min\{D_{i-1,j},\ D_{i,j-1},\ D_{i-1,j-1}\} \tag{8}$$

It's particularly useful for comparing time series data where the sequences might have different lengths or where there's some temporal distortion between them.
Eq.(8) is the recursive formula used to calculate the cost of aligning two elements, (*xi*,*yj*), from two sequences, *X* and *Y*, respectively. *Di*,*j* represents the cumulative cost of aligning the *i*th element of sequence *X* with the *j*th element of sequence *Y*. *f*(*x*,*y*) represents the local cost (or distance) between elements *xi* and *yj*. The idea is to find the optimal alignment path between the two sequences by minimizing the cumulative cost, considering all possible alignments at each step[12][41].

2) **Mean Squared Error (MSE):**
The MSE as shown in Eq.(9) measures the average squared difference between corresponding elements of

two signals. For ECG signals X and Y of length N, the MSE is calculated as:

$$MSE(X,Y) = \frac{1}{N}\sum_{i=1}^{N}(x_i - y_i)^2 \tag{9}$$

Where xi is the ith sample of the generated ECG signal. yi is the ith sample of the real ECG signal. N is the length of the signals (number of samples).
The MSE provides a measure of the average squared difference between corresponding samples of the two signals.

3) Kullback-Leibler (KL) Divergence:
The KL divergence in Eq.(10) measures the discrepancy between two probability distributions. For discrete probability distributions P and Q, the KL divergence from Q to P is computed as:

$$D_{KL}(P|Q) = \sum_{i=1}^{N} p_i \, log \frac{p_i}{q_i} \tag{10}$$

Where pi and qi are the probabilities of the ith bin in the histograms corresponding to the generated and real ECG signals, respectively. N is the number of bins in the histograms.
These metrics collectively provide a comprehensive evaluation of both classification performance and signal fidelity, ensuring generated ECG signals are accurate and morphologically consistent with real data.

5-6- Training model

To train the model, initially, we need to create two sets: a training set and a testing set, which will be used for both training the generative model and the classification model. 90% of the data is allocated to the training set, and the remaining 10% is assigned to the testing set. Table 4 illustrates the distribution of the number of samples in each dataset.

Table 4: Train and test sets

Class	*Total*	*Training set*	*Test set*
P	7028	6325	702
A	25046	2291	254
L	8075	7267	807
N	75052	67546	7505
R	70259	6533	752
f	982	883	98
j	229	206	22

Only 50% of the samples in each class were used in training because we want the classifier to experience weak training on a highly imbalanced dataset. Due to the insufficient number of samples in small classes, the model is intentionally weakened. Therefore, only half of the training samples, as shown in Table 4 were utilized [2].

The proposed generative model has been trained for 30 epochs. In each epoch, the generator takes a randomly sampled noise from a Gaussian distribution along with class labels as input. The random noise acts as a source of variability, while class labels guide the generator to produce ECG signals belonging to specific classes. Thus, using the defined architecture for the generator, it transforms input noise and labels into artificial ECG signals.

The main task of the critic is to distinguish between real ECG signals from the datasets and artificial signals generated by the generator. The number of critic iterations is a hyperparameter that determines how many updates the critic undergoes for each update of the generator during the training process. The goal of multiple critic iterations is to ensure that the critic is effective in accurately distinguishing between real and fake samples. In this iterative loop, the critic is trained to maximize the Wasserstein distance between the distributions of real and fake samples. By repeating the training of the critic several times before updating the generator, the training process stabilizes, assisting in achieving the Wasserstein GAN objective. Then, the generator is updated once for each batch to minimize the generator loss. This alternating training process continues throughout the epochs.
Therefore, the generator and critic are trained iteratively. The training process involves finding a balance where the generator is neither too weak to be easily detected by the critic nor too strong.

The Loss plot for the generator and critic over 30 epochs is presented in Fig. 6.

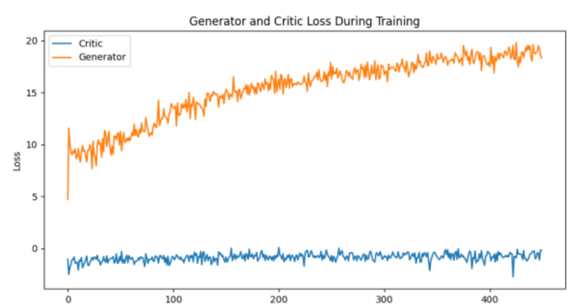


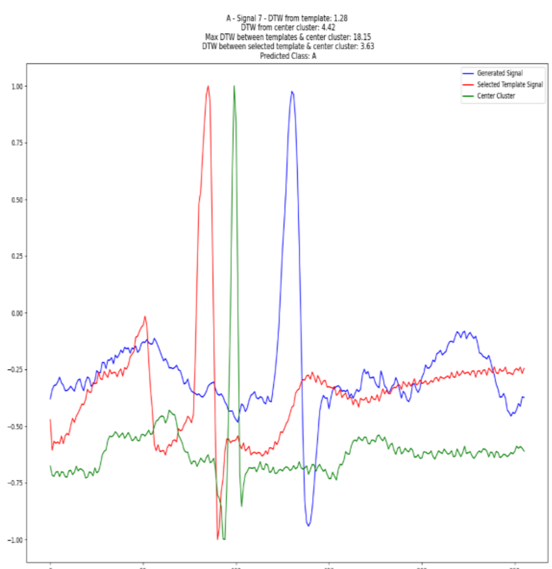
Fig. 6 Loss Plot for Generator and Critic

As indicated in Fig. 6, both the generator and critic have reached the desired stability during the training process, showing no significant fluctuations throughout learning. It can be concluded that the model has converged, and learning has taken place. It should be noted that the generator initially requires some time to adjust its parameters to achieve both the standard GAN objective simultaneously (deceiving the critic) and satisfying the DTW loss. The initial increase in generator loss indicates that the generator is learning to align its generated sequences with the desired patterns and gradually reaches a stable and consistent state.

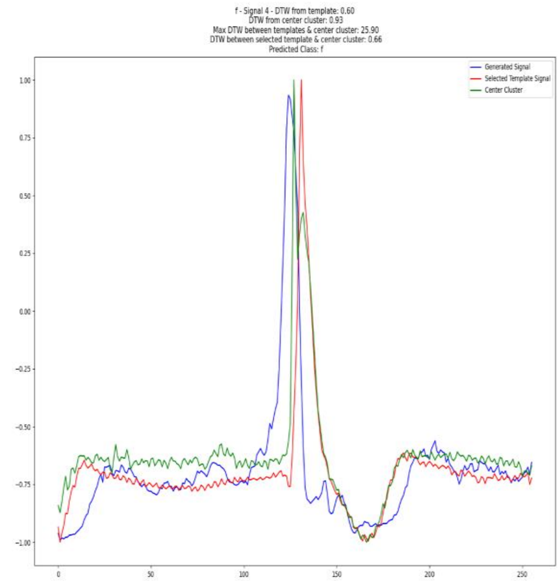
A one-dimensional implementation of ResNet34 is first trained using only real ECG signals. Subsequently, generated signals are added to the training set, and the model is retrained. Performance is finally evaluated on the test set, which contains only real ECG signals, to measure the contribution of the generated data to classification accuracy.

5-7- Evaluate Generated ECG Signals

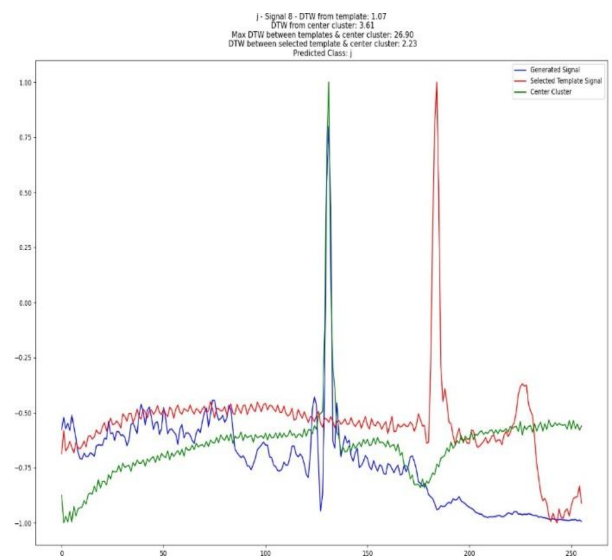
After training the generative model, it was used to synthesize ECG signals for each class. A total of 10,000 samples were generated per class, without applying any threshold at this stage. This initial set of generated signals serves as the basis for further evaluation, including comparison with real signals and outlier removal. Representative generated signals for each class are presented in Fig. 7.



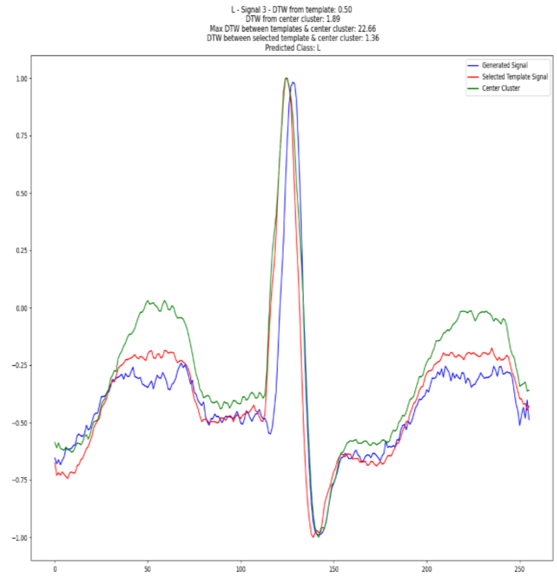
(a)



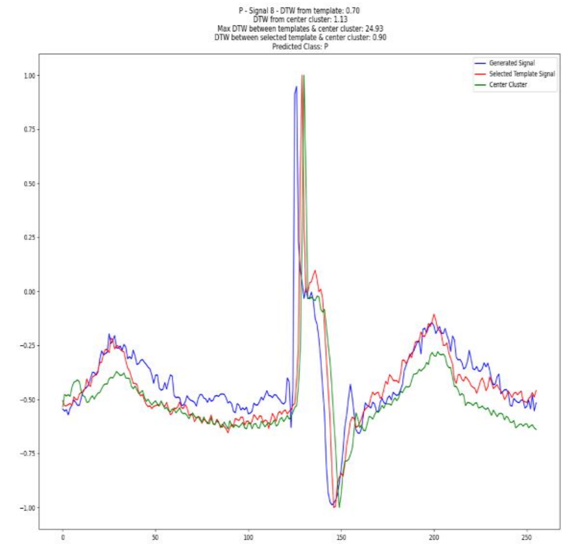
(b)



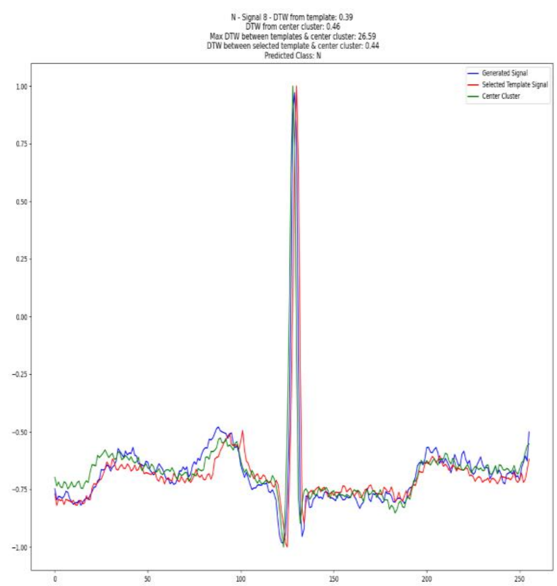
(c)



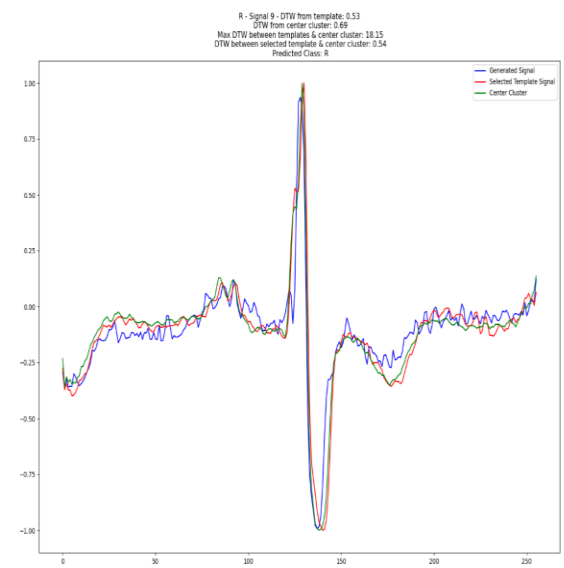
(d)



(f)



(e)

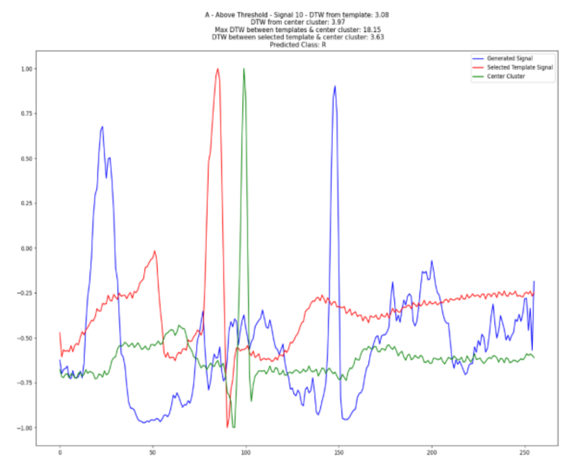


(g)

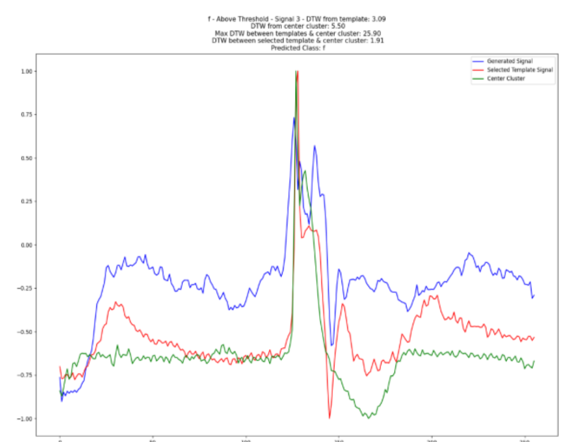
Fig. 7 Generated ECG signals for each class: (a) Class A (b) Class f (c) Class j (d) Class L (e) Class N (f) Class P (g) Class R. The generated signals are indicated in blue, and the nearest real and template sample to each generated signal, based on DTW distance, is shown in red color. Additionally, for a better understanding of the generated signals, the cluster center for each class is also indicated in green along with them. In the figures, the distances of the generated signals from the nearest sample, from the cluster center, the maximum distance between all template samples and the cluster center, and the distance from the cluster center to the nearest real sample are provided. Furthermore, the predicted class by the trained classification model is also given for each generated signal in each class.

As shown in Fig. 7, the generated signals closely resemble the real ECG samples and exhibit the characteristic patterns of their respective classes.

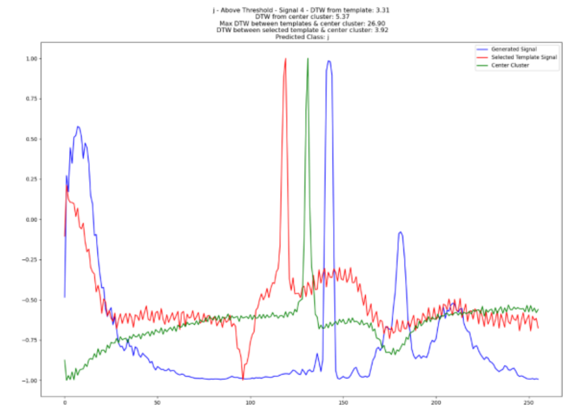
To improve classification accuracy, generated signals with a DTW distance exceeding the defined threshold of 3 from the nearest real sample are considered outliers and removed. Due to computational limitations, only the first 1,000 generated signals per class were evaluated in this step. Representative outlier signals identified for each class are presented in Fig. 8.



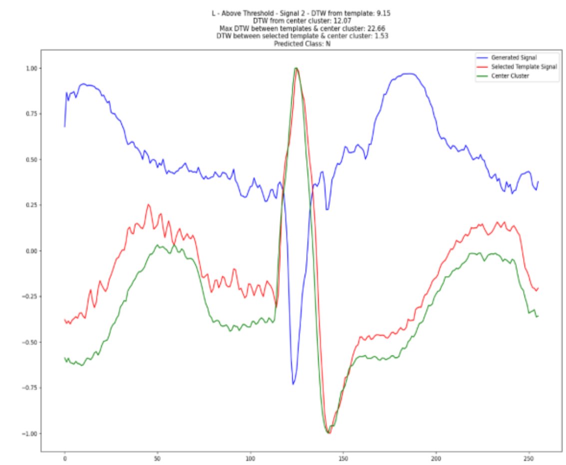
(a)



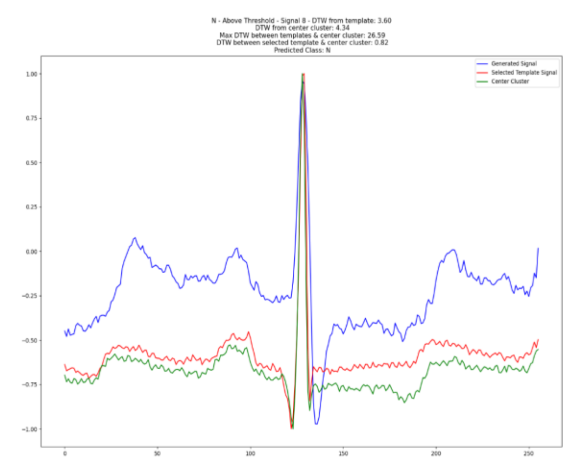
(b)



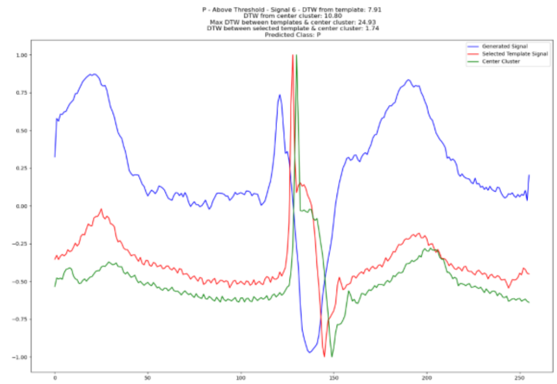
(c)



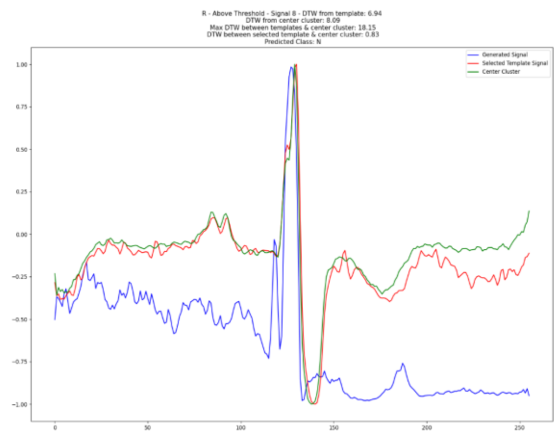
(d)



(e)



(f)



(g)

Fig. 8 Outlier ECG signals for each class: (a) Class A (b) Class f (c) Class j (d) Class L (e) Class N (f) Class P (g) Class R. The generated signals are indicated in blue, and the nearest real and template sample to each generated signal, based on DTW distance, is shown in red color.

As illustrated in Fig. 8, the generated signals show clear dissimilarity from the nearest real signals in each class, which may lead to misclassification. To mitigate this, a DTW threshold of 3 was applied to remove outliers.

Overall, the number of generated outlier samples and the number of good ECG signals, based on only 1000 generated signals from each class, are presented in Table 5.

Table 5: Number of outliers Vs. good signals

Class	***Number of Correct Samples***	***Number of Outlier Samples***
P	463	537
A	985	15
L	821	179
N	670	330
R	792	208
f	895	105
j	993	7

Here, we only examined the number of outlier signals and good signals for the first 1000 generated signals. Therefore, the number of signals in Table 5 is only for the first 1000 signals for each class, not for all generated signals.

5-8- Evaluation using ResNet34 Classifier

In this study, two proposals were examined to investigate the generated ECG signals. One involves modifying the generator's loss function in a way that captures structural differences between the generated signals and real ones. The other proposal involves incorporating an attention mechanism alongside a CNN model in the generator to ensure that the system architecture focuses more on crucial parts of the ECG signals.

A comparison is made regarding the use of other loss functions in the generator loss function. The accuracy and performance of the ResNet34 classifier in classifying generated signals from each of these methods are examined, and the results are presented in Table 6.

Table 6: Comparison of models with different loss functions. In each row, the calculated Precision for each class is displayed, and the last row shows the Accuracy obtained from each method.

Class	***WGAN-GP***	***MSE***	***DTW (Our model)***	***Kullback -Leibler***
P	0.95	0.90	**0.97**	**0.97**
A	0.61	0.53	**0.82**	0.49
L	0.74	0.85	**0.92**	0.65
N	0.98	0.96	**0.99**	0.98
R	0.93	**0.94**	0.88	0.86
f	0.11	0.22	0.34	**0.36**
j	0.12	**0.19**	0.10	0.10
ACC	0.89	0.92	**0.95**	0.89

In total, four evaluation methods were considered. The first method used only the WGAN-GP loss (Eq.(3)) without any additional terms in the generator loss. The second method adds Mean Squared Error (MSE) to encourage similarity to real signals (Eq.(11)).

$$L_{gen} = -E_{x \sim p_{fake}}[D(x,y)] + \lambda_{MSE} \cdot MSE_{loss} \quad (11)$$

The third method, our proposed approach, incorporates DTW to better align generated and real ECG waveforms (Eq.(6)), while the fourth method employs Kullback-Leibler (KL) divergence to promote diversity in the generated signals (Eq.(12)).

$$L_{gen} = -E_{x \sim p_{fake}}[D(x,y)] + \lambda_{KL} \cdot KL_{divergence} \quad (12)$$

As indicated in Table 6, Using MSE increased precision in classes R and j, whereas our DTW-based method achieved higher precision in the remaining classes. The goal of both these approaches is to generate signals as similar as possible to real signals. However, the fourth method, which involves using Kullback-Leibler in the generator's loss function, has a different objective. Its primary goal is to create diversity in the generated signals. As is evident, the proposed method generally has a higher accuracy compared to other methods after adding generated data to the dataset.

The classification results of the proposed model are summarized in Table 7. Compared with the reference model [2] the proposed approach achieved higher overall accuracy (95% vs. 92%) and showed better F1-scores in the main classes (P, L, N, and R). Although the reference model obtained higher recall for the minority classes (f and j), it suffered from lower precision and reduced overall accuracy. These results indicate that the proposed model provides a more balanced and reliable performance, particularly in the clinically important classes related to cardiac signal analysis.

Table 7: Classification Report

Class	*Precision*	*Recall*	*F1-score*	*Support*
P	0.97	0.99	0.98	702
A	0.82	0.91	0.86	254
L	0.92	0.97	0.95	807
N	0.99	0.95	0.97	7504
R	0.88	0.98	0.92	725
f	0.34	0.62	0.44	98
j	0.10	0.36	0.16	22
Accuracy		0.95		10112
Macro avg	0.72	0.83	0.75	10112
Weighted avg	0.96	0.95	0.95	10112

In the next comparison, four generative models were evaluated, as summarized in Table 8. The first model is our proposed DTW-based attention model. The second model is an attention-based version without incorporating DTW in the generator loss. The third model is the baseline WGAN-GP model [2], and the fourth model replaces the critic with a discriminator to assess the impact of this architectural change.

For the baseline model with a discriminator, the generator loss was modified by replacing the Cross-Entropy term with WGAN-GP, ensuring a fair comparison with models using a critic.

All models were evaluated using 10,000 generated signals, without applying an outlier removal threshold, to provide a consistent basis for comparison.

Table 8: Comparison of models in terms of classification accuracy

Class	*Proposed Model*	*Proposed model without DTW*	*Baseline [2]*	*Baseline with Disc*	*No generated signal*
P	**0.97**	0.95	0.95	**0.97**	0.76
A	**0.82**	0.61	0.62	0.49	0.37
L	**0.92**	0.74	0.86	0.63	0.74
N	0.99	0.98	1.00	0.97	0.94
R	0.88	**0.93**	0.76	0.86	0.78
f	0.34	0.11	**0.35**	0.18	0.00
j	0.10	**0.12**	**0.12**	0.02	0.00
ACC	**0.95**	0.89	0.92	0.89	0.89

Based on the results in Table 8, the proposed model outperformed all other evaluated models on the test data.

Table 9 presents a comparison of the proposed generative method with other recently utilized generative approaches in terms of classification accuracy after incorporating generated signals into the dataset using ResNet34.

Table 9: Comparison of the proposed model with related models in terms of classification accuracy

Research	*Year*	*Model*	*Accuracy*
Our	2024	ABC-GAN	**95%**
[2]	2022	ACWGAN-GP	92%
[17]	2022	ME-GAN	90.2%
[25]	2023	Improved DDPM	90%
[26]	2024	CECG-GAN	94%

The proposed method achieves higher accuracy, which can be attributed to its ability to effectively capture class-specific characteristics through the attention mechanism in the generator architecture and the inclusion of the DTW function in the generator loss. By generating signals that more closely resemble real ECG signals within each class, the method improves classification performance when added to the dataset.

5-9- Computational Complexity Analysis

In addition to generation quality, computational complexity is an important consideration for practical deployment. Table 10 compares the computational complexity of ABC-GAN against baseline methods.

Table 10: Comparison of the proposed model with related models in terms of complexity

Method	*Core Mechanism*	*Relative Complexity*	*Training Stability*
CardioGAN [12]	Attention-based GAN with dual discriminators	Low	Medium
AC-WGAN-GP [2]	Wasserstein GAN with gradient penalty	Medium	High
ME-GAN [17]	Multi-view conditional GAN with disease-aware representations	Medium–High	Medium
TTS-GAN [14]	Transformer-based GAN using self-attention	High	Medium
CECG-GAN [26]	Conditional GAN for class-imbalanced ECG synthesis	Medium–High	Medium
Task-Aware CGAN (DTW) [27]	Conditional GAN with DTW + multi-objective temporal loss	Medium–High	Medium
TransDiffECG [31]	Transformer-based diffusion model	High	High
RDiffGAN-ECG [32]	Hybrid diffusion + GAN + state-space modeling for R-peak conditioned ECG	Very High	High
ABC-GAN (Proposed)	**Attention-based conditional GAN with explicit DTW loss in generator**	**Medium–High**	**Medium–High**

5-10- Limitations

In this study, a fixed DTW threshold of 3 was selected for outlier removal across all classes. This choice was primarily motivated by computational constraints and to ensure a uniform filtering criterion. While class-specific thresholds have been proposed in the baseline work, where variability differs significantly between majority and minority classes, our simplified approach applies the same cutoff to all classes. Consequently, certain signals from high-variability classes (A and j) may have been incorrectly discarded, while signals in lower-variability classes might have been retained despite subtle mismatches.

Another limitation arises from evaluating only the first 1,000 generated samples per class when applying the DTW threshold. Although this was necessary due to computational costs, restricting the analysis to a subset of generated signals could bias the assessment of outlier prevalence and overall model fidelity. Future studies should investigate adaptive, class-specific thresholds and conduct a large-scale evaluation across all generated signals to provide a more comprehensive and statistically reliable assessment of synthetic ECG quality.

6- Future Works

Based on the results obtained, combining the DTW and MSE functions in the generator's loss function may enhance classification accuracy across multiple ECG classes. Specifically, in some classes, adding the MSE term to the DTW-based generator loss showed measurable improvements compared to using DTW alone. Future studies could systematically explore this combination and quantify the improvements for each class.

Additionally, ensuring the privacy of generated ECG signals remains an important consideration. Future work could investigate methods to rigorously guarantee patient data privacy, such as differential privacy or data anonymization techniques, when generating synthetic ECG signals.

Finally, alternative evaluation metrics to DTW should be considered. While DTW effectively captures temporal alignment, it is computationally intensive. Exploring other similarity measures, such as Pearson correlation, cosine similarity, or feature-based metrics, could reduce computational time while maintaining evaluation accuracy. Combining multiple evaluation criteria could further enhance both efficiency and reliability in assessing generated signals.

7- Conclusions

This study presents a novel approach for ECG signal generation and classification by integrating an attention mechanism into the generator architecture, employing Dynamic Time Warping (DTW) in the generator loss function, and utilizing a Critic block instead of a traditional discriminator. The proposed method successfully generates realistic ECG signals that, when added to the training dataset, improve classification performance using a ResNet34-based model. The results demonstrate that the proposed approach outperforms recent methods in terms of overall classification accuracy, achieving 95% compared to 92% for the baseline [2]. Key improvements include enhanced precision and F1-scores in major ECG classes (P, L, N, and R), while maintaining balanced performance across minority classes. Although the model incurs a higher computational cost due to DTW calculations, it achieves superior signal fidelity and class-specific learning. For future research, we recommend focusing on optimizing the thresholding process for generated signals, further preserving patient privacy, and exploring alternative or combined evaluation metrics to reduce computational complexity. Overall, the proposed method provides a robust framework for generating and classifying ECG signals, with significant potential for further enhancements in synthetic biomedical data generation.

References

[1] G. Chen et al., "EmotionalGAN: Generating ECG to Enhance Emotion State Classification," 2019.

[2] E. Adib, F. Afghah, and J. J. Prevost, "Arrhythmia Classification using CGAN-augmented ECG Signals," arXiv preprint arXiv:2201.12345, 2022.

[3] N. P. Singh et al., "Remote Monitoring System of Heart Conditions for Elderly," Journal of Information Systems and Telecommunication (JIST), (accessed online).

[4] A. Sandooghdar and F. Yaghmaee, "Deep Learning Approach for Cardiac MRI Images," Journal of Information Systems and Telecommunication (JIST), vol. 10, no. 37, pp. xx–yy, 2022.

[5] A. Halder Roy, "Integrating Questionnaires and Physiological Signals for Stress Detection," Journal of Information Systems and Telecommunication (JIST), Jul. 2025.

[6] F. Zhu et al., "Electrocardiogram generation with a bidirectional LSTM-CNN generative adversarial network," Scientific Reports, vol. 9, no. 1, p. 107, 2019.

[7] Y.-H. Zhang and S. Babaeizadeh, "Synthesis of standard 12-lead electrocardiograms using two-dimensional generative adversarial networks," Journal of Electrocardiology, vol. 69, pp. 1–9, 2021.

[8] N. Wulana et al., "Generating electrocardiogram signals by deep learning," Neurocomputing, vol. 404, pp. 1–11, 2020.

[9] P. Wang et al., "ECG Arrhythmias Detection Using Auxiliary Classifier Generative Adversarial Network and Residual Network," IEEE Access, vol. 7, pp. 1–10, 2019.

[10] W. Li et al., "SLC-GAN: An automated myocardial infarction detection model based on generative adversarial networks and convolutional neural networks with single-lead electrocardiogram synthesis," Information Sciences, vol. 589, pp. 1–12, 2021.

[11] H. Yang et al., "ProEGAN-MS: A Progressive Growing Generative Adversarial Networks for Electrocardiogram Generation," IEEE Access, vol. 9, pp. 1–10, 2021.

[12] S. Dasgupta, S. Das, and U. Bhattacharya, "CardioGAN: An Attention-based Generative Adversarial Network for Generation of Electrocardiograms," in Proc. 25th ICPR, 2021.

[13] R. Banerjee and A. Ghose, "Synthesis of Realistic ECG Waveforms Using a Composite Generative Adversarial Network for Classification of Atrial Fibrillation," in Proc. 29th EUSIPCO, 2021.

[14] X. Li et al., "TTS-GAN: A Transformer-Based Time-Series Generative Adversarial Network," in Artificial Intelligence in Medicine. AIME 2022. LNCS, vol. 13263, Springer, 2022.

[15] E. Brophy, "Synthesis of Dependent Multichannel ECG using Generative Adversarial Networks," in Proc. 29th ACM CIKM, 2020.

[16] S. Harada, H. Hayashi, and S. Uchida, "Biosignal Data Augmentation Based on Generative Adversarial Networks," in EMBC 2018, 2018.

[17] J. Chen et al., "ME-GAN: Learning Panoptic Electrocardio Representations for Multi-view ECG Synthesis Conditioned on Heart Diseases," in ICML 2022, 2022.

[18] F. Ye et al., "ECG Generation With Sequence Generative Adversarial Nets Optimized by Policy Gradient," IEEE Access, vol. 7, pp. 1–10, 2019.

[19] V. V. Kuznetsov et al., "Interpretable Feature Generation in ECG Using a Variational Autoencoder," Frontiers in Genetics, vol. 12, p. 634, 2021.

[20] T. Golany and K. Radinsky, "PGANs: Personalized Generative Adversarial Networks for ECG Synthesis to Improve Patient-Specific Deep ECG Classification," in AAAI 2019, 2019.

[21] M. S. Munia, M. Nourani, and S. Houari, "Biosignal Oversampling Using Wasserstein Generative Adversarial Network," in ICHI 2020, 2020.

[22] D. Nankani and R. D. Baruah, "Investigating Deep Convolution Conditional GANs," in IJCNN 2020, 2020.

[23] F. N. Hatamian et al., "The Effect of Data Augmentation on Classification of Atrial Fibrillation in Short Single-Lead," in ICASSP 2020, 2020.

[24] T. Agarwal and E. Ertin, "CardiacGen: A Hierarchical Deep Generative Model for ECG Synthesis," 2022.

[25] E. Adib et al., "Synthetic ECG Signal Generation Using Probabilistic Diffusion Models," IEEE Access, vol. 11, pp. 75818–75828, 2023.

[26] X. Li, E. Adib, and A. M. Rahmani, "CECG-GAN: Conditional ECG Generation for Class-Imbalanced Cardiac Diagnosis Using GANs," Scientific Reports, vol. 14, no. 1, p. 14767, 2024, doi: 10.1038/s41598-024-65619-8.

[27] K. Lang and Y. Li, "Task-Aware Conditional GAN with Multi-Objective Loss for Realistic and Efficient Time-Series Generation," Journal of Big Data, vol. 12, 2025, Art. no. 206, doi: 10.1186/s40537-025-01266-8.

[28] N. Neifar, A. Ben-Hamadou, A. Mdhaffar, and M. Jmaiel, "DiffECG: A Versatile Probabilistic Diffusion Model for ECG Signal Synthesis," in Proc. IEEE/ACIS Int. Conf. Software Engineering Research, Management and Applications (SERA), 2024, pp. 182–188, doi: 10.1109/SERA61261.2024.10685651.

[29] Y. Lai, B. Liu, X. Guan, Q. Zhao, and S. Hong, "ECGTwin: Personalized ECG Generation Using Controllable Diffusion Model," arXiv:2508.02720, 2025. [Online]. Available: https://arxiv.org/abs/2508.02720

[30] L. Simone, D. Bacciu, and V. Gervasi, "ECG Synthesis for Cardiac Arrhythmias: Integrating Self-Supervised Learning and Generative Adversarial Networks," Artificial Intelligence in Medicine, vol. 167, p. 103162, 2025, doi: 10.1016/j.artmed.2025.103162.

[31] Y. Lin, J. Ma, S. Dong, C. Sun, W. Cong, and K. Wang, "TransDiffECG: Semantically Controllable ECG Synthesis via Transformer-Based Diffusion Modeling," Journal of Biomedical Informatics, vol. 172, p. 104948, 2025, doi: 10.1016/j.jbi.2025.104948.

[32] P. Li, Y. Zhou, J. Min, Y. Wang, W. Liang, and W. Li, "RDiffGAN-ECG: High-Fidelity R-Peak Conditional ECG Generation Using Diffusion GAN with State Space Models," Biomedical Signal Processing and Control, vol. 108, p. 108544, 2025, doi: 10.1016/j.bspc.2025.108544.

[33] M. Mirza and S. Osindero, "Conditional Generative Adversarial Nets," arXiv preprint arXiv:1411.1784, 2014.

[34] I. Goodfellow et al., "Generative Adversarial Nets," in Advances in Neural Information Processing Systems (NIPS), 2014, pp. 2672–2680.

[35] A. Vaswani et al., "Attention Is All You Need," in Advances in Neural Information Processing Systems (NIPS), 2017, pp. 5998–6008.

[36] H. Zhang, I. Goodfellow, D. Metaxas, and A. Odena, "Self-Attention Generative Adversarial Networks," in Proc. ICML, 2019, pp. 7354–7363.

[37] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein Generative Adversarial Networks," in Proc. 34th ICML, 2017.

[38] A. Lyashuk, "ECG Classification," [Online]. Available: https://github.com/lxdv/ecgclassification/blob/master/README.md, 2021.

[39] K. He et al., "Deep residual learning for image recognition," in Proc. IEEE Conf. CVPR, 2016, pp. 770–778.

[40] G. B. Moody and R. G. Mark, "The impact of the MIT-BIH Arrhythmia Database," IEEE Engineering in Medicine and Biology Magazine, vol. 20, no. 3, pp. 45–50, 2001.

[41] A. M. Delaney, E. Brophy, and T. E. Ward, "Synthesis of realistic ECG using generative adversarial networks," arXiv preprint arXiv:1903.09633, 2019.

[42] S. Harada, H. Hayashi, and S. Uchida, "Biosignal Generation and Latent Variable Analysis With Recurrent Generative Adversarial Networks," IEEE Access, vol. 7, pp. 1–10, 2019.

[43] K. F. Hossain et al., "ECG-Adv-GAN: Detecting ECG Adversarial Examples with Conditional Generative Adversarial Networks," in Proc. 20th IEEE ICMLA, 2021.

[44] A. M. Shaker et al., "Generalization of Convolutional Neural Networks for ECG Classification Using Generative Adversarial Networks," IEEE Access, vol. 8, pp. 1–10, 2020.

[45] T. Golany et al., "Improving ECG Classification Using Generative Adversarial Networks," in IAAI-20, 2020.

[46] A. Furdui et al., "AC-WGAN-GP: Augmenting ECG and GSR Signals using Conditional Generative Models for Arousal Classification," 2021.

[47] Y.-Y. Jo et al., "ECGT2T: Towards Synthesizing Twelve-Lead Electrocardiograms from Two Asynchronous Leads," in ICASSP 2023, pp. 1–5.

[48] T. Lan et al., "Arrhythmias Classification Using Short-Time Fourier Transform," in EMBC 2020, 2020.

[49] A. Goldberger et al., "Components of a new research resource for complex physiologic signals: PhysioBank, PhysioToolkit, and PhysioNet," Circulation, vol. 101, no. 23, pp. E215–E220, 2000.