Synthetic ECG Signal Generation Using Probabilistic Diffusion Models

Pakpoom Danjittisiri 6536114
M.Sc. Data Science for Healthcare and Clinical Informatics

I. Introduction

- Deep learning models are used for automated ECG-based diagnosis due to their ability to classify ECG signals based on their morphological patterns.
- ECG datasets used for training are often imbalanced, with normal beats being more abundant than abnormal ones.
- Various approaches are employed to address imbalances, including oversampling methods like SMOTE, new loss functions such as focal loss, and new training schemes like few-shot training.
- Deep generative algorithms like Generative Adversarial Networks (GAN) and Variational Auto-Encoders (VAE) are increasingly used to generate synthetic ECG signals.
- Due to being outstandingly successful in classification as well as in generation tasks of computer vision deep learning models (e.g. AlexNet, VGG-16, ResNet-18,..), utilizing 2D data space for ECG signal generation seems enabling more augmentation techniques, such as flipping, rotation, and mirroring, compared to 1D space and also improves classification performance.
- The Improved Denoising Diffusion Probabilistic Models (DDPM) is explored as a method for generating images (2D) with claims of superior quality compared to GAN models.

I. Introduction

CONTRIBUTION

- To present the first-used pipeline for generating 1D synthetic ECG signals using 2D Denoising Diffusion Probabilistic Models (DDPM)

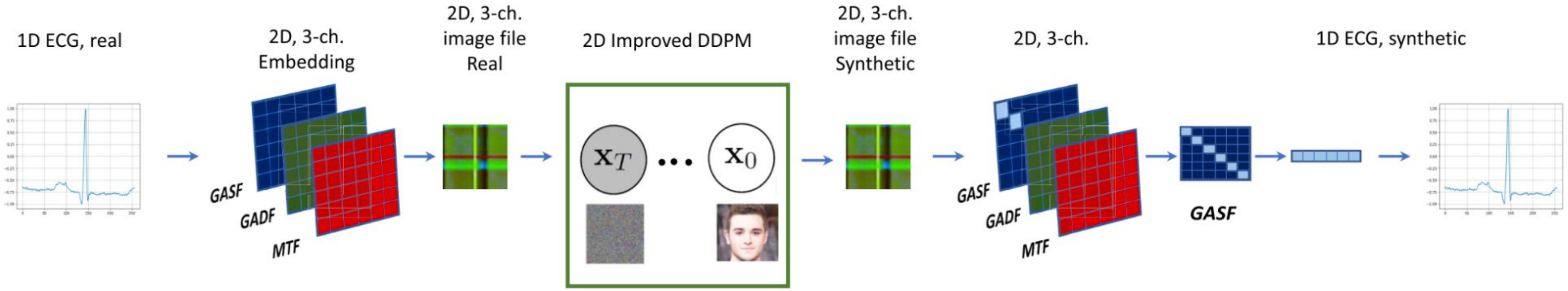
II. Methodology

A. DATASET AND SEGMENTATION

- MIT-BIH Arrhythmia dataset was used.
- It contained a set of 48 Holter recordings, of which each 30 minutes long with two channels.
- The modified limb lead II (MLII) was utilised in this study.
- ECG signals were segmented by the Adaptive Window method into 109,338 individual beats in 15 classes.
- All the segmented beats were resampled to 256 Hz.

II. Methodology

B. PIPELINE



Where
- *Gramian Angular Summation/Difference Fields (GASF/GADF)*
- *Markov Transition Fields (MTF)*

FIGURE 1. Block diagram of the proposed model. First, the normalized ECG time series are transformed into 2*D* space using GASF, GADF, and MTF separately, and 3-channel RGB-like data are generated. Then, the Improved DDPM model is trained and sampled to generate 2*D* ECG data. Finally, the 1*D* ECG time series are reconstructed using the diagonals of the GASF channel.

II. Methodology

C. Authenticity of Generated Beats

- The authenticity test was conducted for comparison between the generated beats and the real beats in a classification task to observe whether or not the generated beats can replace the real ones.
- The pre-trained ResNest-34 model was used with 1D implementation for the classification task.
- The training set was consisted of two classes, i.e. N (Normal Beat) and L (Left Bundle Branch Block Beat).
- The same test set was used in all cases with 1,000 samples of unseen real data in each class with no synthetic beats

TABLE 1. Authenticity test training set supports.

Cases	Train Set Class N	Train Set Class L
Reference (Balanced, all real)	7000 r	7000 r
Compromised (Imbalanced)	350 r	7000 r
Augmented with case 00	350 r + 6650 s	7000 r
Augmented with case 01	350 r + 6650 s	7000 r
Augmented with case 02	350 r + 6650 s	7000 r
Augmented with case GAN	350 r + 6650 s	7000 r

r: Real Beat *s: Synthetically Generated Beat*

III. Experimental Setup

A. Study Cases

- Three study cases for the Improved DDPM included
    - Case 00
    - Case 01
    - Case 02
- The fourth study case was the data generated by the WGAN-GP model
    - GAN

III. Experimental Setup

B. The Improved DDPM

- Three different hyperparameter settings for the DM case studies were considered, while the rest of the parameters are the same for all cases.

TABLE 4. DM case studies.

Cases	Learn Sigma	Noise Schedule	Use KL [1]	Schedule Sampler
00	False	Linear	False	Uniform
01	True	Cosine	True	Uniform
02	True	Cosine	True	loss second moment

1: *Kullback-Leibeler (KL) Divergence*

III. Experimental Setup

C. WGAN-GP Model design

- The architectures of the generator and the critic in the Wasserstein GAN with Gradient Penalty (WGAN-GP) model are comprised of building blocks that are repeated multiple times.
- The processing are all in 1D space
- This model was constructed for comparison with the Improved DDPM in part generating realistic beats that resemble and function like real beats.

TABLE 2. WGAN-GP building blocks.

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

1: *kernel size* = *4*, *stride* = *2*, *padding* = *1*

TABLE 3. WGAN-GP architecture.

Layer	Generator	Critic
Input	$16 \times 100 \times 1$	$16 \times 1 \times 64$
1	Block	Vonv1d, LeakyReLU
2	Block	Block
3	Block	Block
4	Block	Block
5	ConvTranspose1d	Conv1d
6	FC	FC
7	tanh	-
Output	$16 \times 1 \times 64$	$16 \times 1 \times 1$

IV. Results

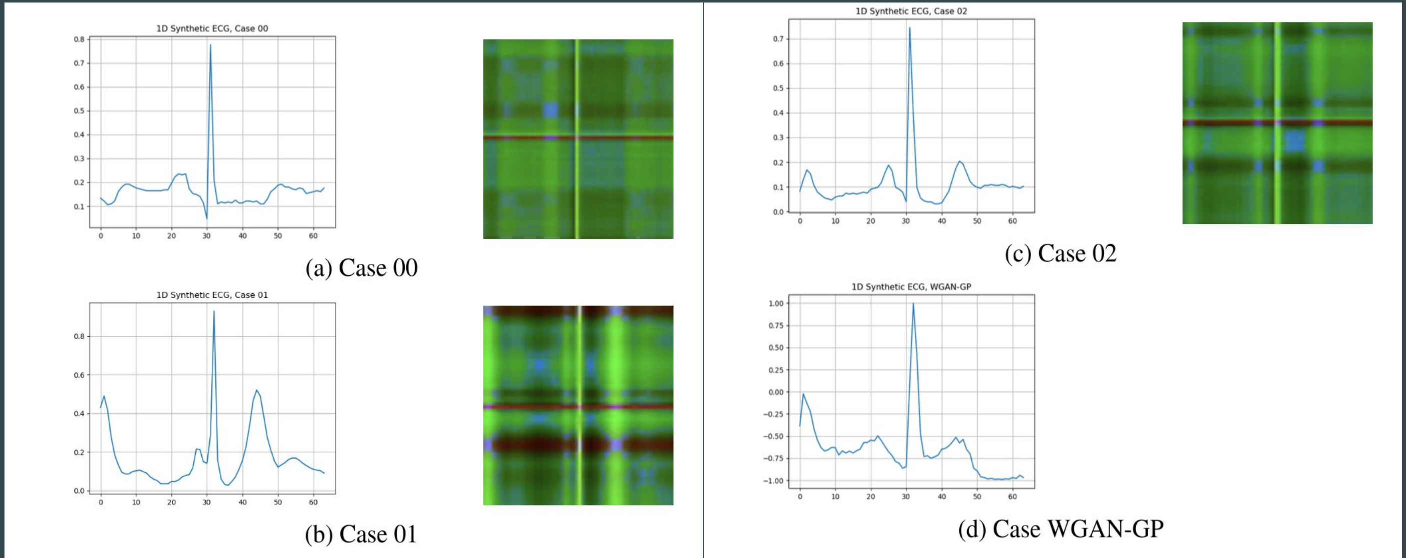


FIGURE 3. Samples of synthetically generated ECG signals.

FIGURE 3. Samples of synthetically generated ECG signals.

- The four case studies were then compared by the quality, distribution, and authenticity of the generated beats in each case

IV. Results

A. Quality

- How much the generated synthetic beats resemble the real ones in appearance and morphology.
- The average distance of the generated beats from a randomly selected template was measured using two distance functions, i.e.
  - Dynamic Time Warping (DTW) function
  - Fre'chet distance function

TABLE 5. Quality of generated beats.

Cases	Ave. DTW Distance	Ave. Fréchet Distance
00	6.67	1.074
01	6.95	1.117
02	6.36	1.042
GAN	2.12	0.723
Real (rl)	2.09	0.718

IV. Results

B. Distribution

- Maximum Mean Discrepancy (MMD) is a kernel based statistical tool to measure the distance between two distributions.

TABLE 6. MMD value of synthetic and real beats.

Cases	00-rl	01-rl	02-rl	GAN-rl	rl-rl
MMD	39.8	44	35.9	1.00	0.0

IV. Results

C. Authenticity

- To measure how much the generated beats can replace the real ones in a classification test
- Measurement metrics included
  - Average Precision Scores
  - The Area Under the Curve (AUC) of the Precision-Recall Curves
  - The AUC of the Receiver Operating Characteristic curves (AUC ROC score)

TABLE 7. Authenticity of generated beats.

Cases	Ave Precision	PRC [1] AUC [2] Score	ROC [3] AUC Score
00	0.90	0.95	0.96
01	0.55	0.68	0.63
02	0.76	0.76	0.81
GAN	0.96	0.99	0.99
Real (rl)	0.98	1.00	1.00

1: *Precision-Recall Curve*
2: *Area Under Curve*
3: *Receiver Operating Characteristic Curve*

V. Discussion

In classification tasks, the mapping of data into an embedded space only requires an injective mapping, where each data point in the original space has a unique corresponding point in the embedded space. Thus, invertibility is not necessary. Spectrograms are useful in classification tasks as they fulfill this requirement. However, in generation tasks, the data must be de-embedded and mapped back into the original space, necessitating a bijective mapping. This is why spectrograms may not be suitable for the generation tasks.

In the context of ECG signal generation, using the WGAN-GP model generates 64 pieces of information per heartbeat, all of which are considered useful. On the other hand, with DDPM, 3 x 64 x 64 = 12,288 pieces of information are generated per beat, but only 64 (0.52%) are considered useful, and the rest are discarded. During training, the optimization process aims to find the optimal values of parameters such that the generated information is close to real beats. The error in the loss function accounts for the deviation between generated and real beats, with WGAN-GP typically having a tighter neighborhood (closer to real beats) and DDPM having a more relaxed one.

VI. Conclusion

The study found that beats generated by the WGAN-GP model consistently resemble real beats more closely than those generated by DDPM across all metrics used.

VII. Limitations and Future Works

- The study employs the Improved DDPM, a 2D model developed by OpenAI, with minimal alterations.
- The proposed pipeline involves mapping 1D ECG data into 2D space, converting them into image files, and feeding them to the DDPM model. Processing occurs in 2D space, and generated 2D data are de-embedded back into 1D space to reconstruct ECG data. In contrast, the WGAN-GP model used in the study is inherently 1D, requiring no embedding.
- The conclusions drawn from this research are specific to the employed setting and pipeline. The results may vary in different settings.
- While probabilistic diffusion models are commonly applied to images and 2D data, they can also be adapted for 1D data. In such cases, the model operates directly on 1D data, performing noise/denoising processes without embedding. This approach could provide a more realistic comparison in terms of the diffusion concept.

Thank You!