Synthesizing Electrocardiograms with Generative Adversarial Networks

Xianglong Wang[1,2,3], Berkman Sahiner[1], Christopher Scully[1], Kenny Cha[1]

[1]Center for Devices and Radiological Health, U.S. Food and Drug Administration, Silver Spring, MD [2]University of Michigan, Ann Arbor, MI

[3]Tulane University, New Orleans, LA

Background

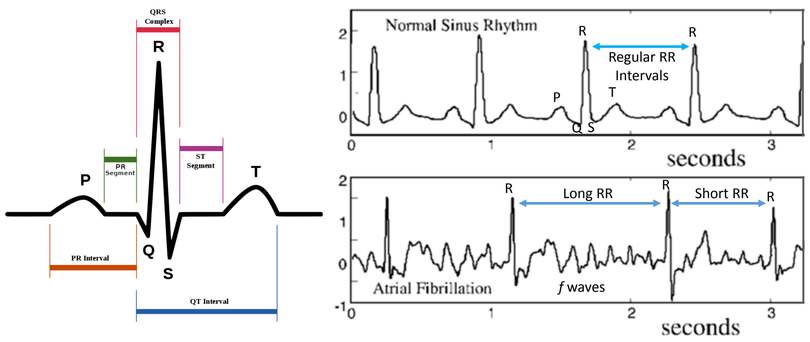
Atrial Fibrillation (AF) is:

- The most common type of arrhythmia
- Linked to increased risks of stroke and heart failure

Electrocardiogram (ECG) analysis is the primary tool for diagnosing AF.

The AF ECG signals are often characterized by:

- P-wave suppressed or absent
- Fibrillatory *f* waves in QT interval
- Higher variation in RR intervals



Wearable ECG recording devices have flourished in forms of dedicated devices or as part of wearable smart devices. These developments have sparked interest in running on-device AF prediction algorithms, including deep learning based algorithms. The performance of deep learning based predictions are often limited by availability of labeled training ECG signals that are from AF patients.

Generative adversarial networks (GANs) have been successfully used in synthesizing realistic-looking data from trained samples, especially in images and speech generation. To generate AF ECG signals, we applied a WarpGAN-based 1D GAN to convert real normal sinus rhythms into synthetic AF ECG signals.

Regulatory Relevance

Supplementing real training data with realistic synthetic data can increase the size of training set and reduce the burden associated with obtaining and annotating new data. Less real data for training would be required so that more real data can be applied towards testing the performance of the algorithm.

Objective

Our objective is to synthesize AF ECG data from normal ECG data with GAN-based deep learning methods. The synthesized AF data could be applied towards training AF detection algorithms to mitigate limited representation of AF ECG signals in training sets.

Dataset and Workflow

The 2017 Physionet Challenge aimed to encourage intelligent algorithms for ECG rhythms classification and provided a training database of single-lead ECG signals that included 5154 normal and 771 AF records.

Type	# Samples	Recording Range (s)	Recording Length (s)
Normal	5,154	[9.0, 61.0]	31.9 ± 10.0
AF	771	[10.0, 60.0]	31.6 ± 12.5

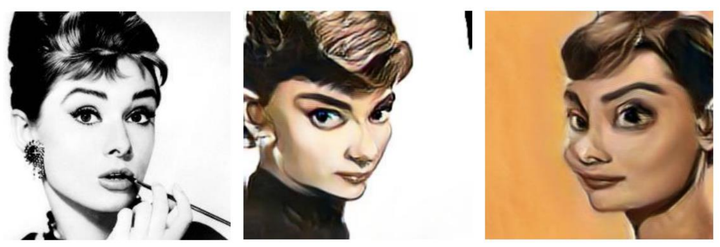
We assembled the dataset for GAN training according to the following flow diagram:

Load Data
- Load the database and extract normal and AF ECG records

Preprocess
- Bandpass filter the signal from 0.25-55 Hz
- Flip inverted signals

Extract
- Length > 10s: extract signal at 6-10s
- Length < 10s: extract last 5s

We based our GAN on the WarpGAN structure, originally used to generate caricatures from real photos. Examples are shown below.

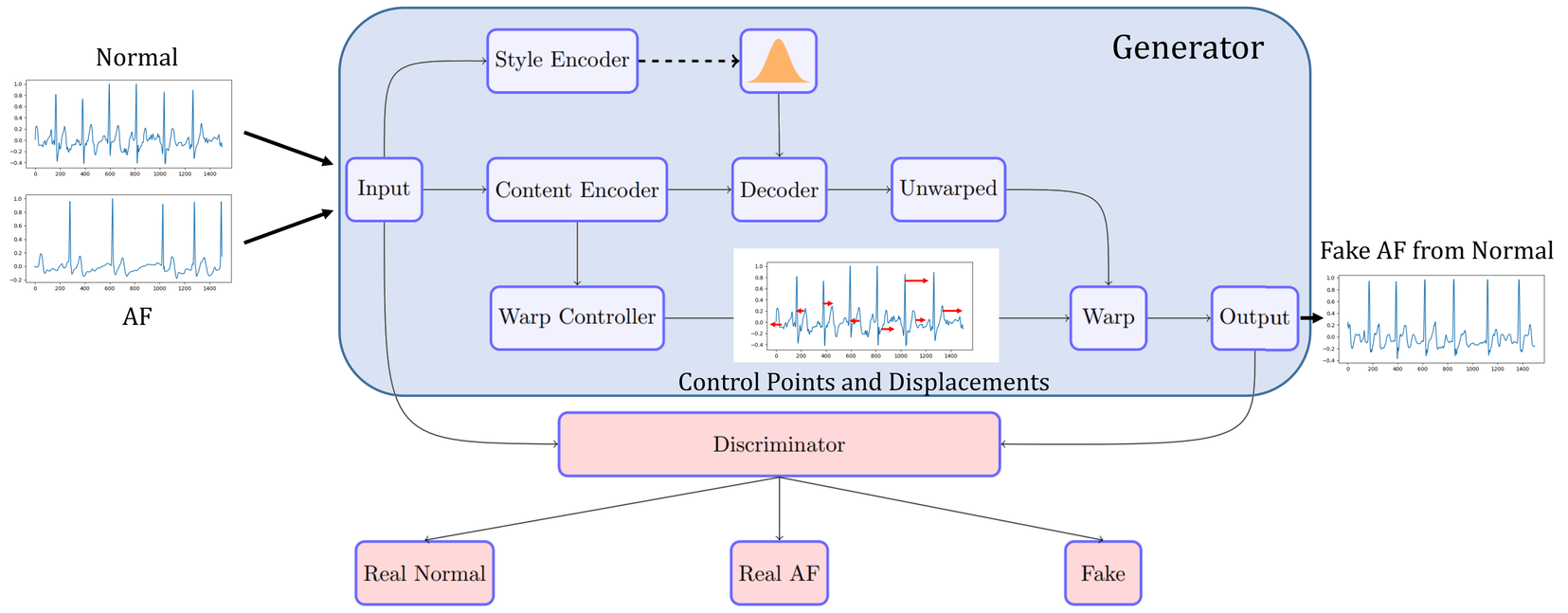


(a) Photo (b) WarpGAN (c) WarpGAN

The generator of our WarpGAN contains an autoencoder-based style transfer GAN that changes the morphology of normal ECG signals into AF ECG signals. The warp controller learns the control points of the signal and then warps the signal to obtain varied RR intervals.

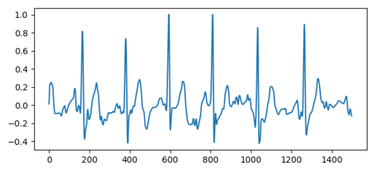
The discriminator differentiates between three classes: real normal, real AF, and fake. The discriminator sees the signal globally (whole signal) and locally (as small patches) to make the classification.

Network Structure

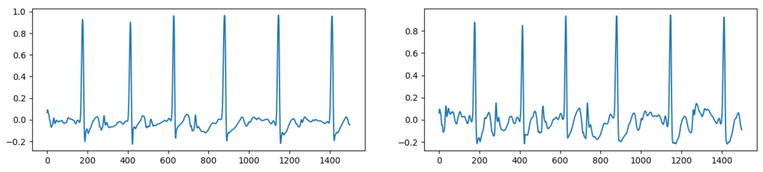


Results

The signals that our GAN generates exhibit characteristics found in AF. The P-waves are suppressed and fibrillatory waves similar to *f* waves appear during QT interval. The RR-interval are more irregular compared to the source signals. Our GAN is capable of generating multiple signals from a single normal ECG signal. For example:

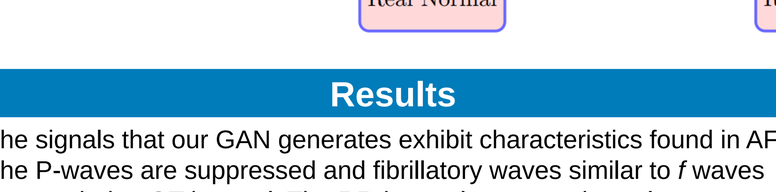


(a) Source Normal ECG



(b) Generated AF ECG (c) Generated AF ECG

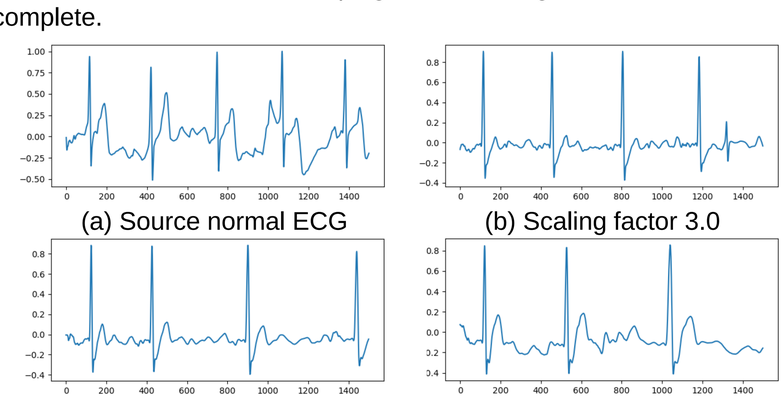
The synthesized signals exhibit different behaviors during different stages of training. In early and late stages, changes are mostly restricted to morphology changes and variations in RR intervals. In mid stages of training, new beats are sometimes inserted and skipped beats could be observed. Late stages partially revert the generated beats. The synthesized signals from early, mid, and late stages could be combined together to make a more comprehensive set of generated AF signals.



(a) Source normal ECG (b) Early stage generated

(c) Mid stage generated (d) Late stage generated

We can also control the amount of variations in RR intervals by introducing a scaling factor during the synthesis stage. We can increase or decrease the amount of warping to make our generated AF set more complete.



(a) Source normal ECG (b) Scaling factor 3.0

(c) Scaling factor -6.0 (d) Scaling factor 6.0

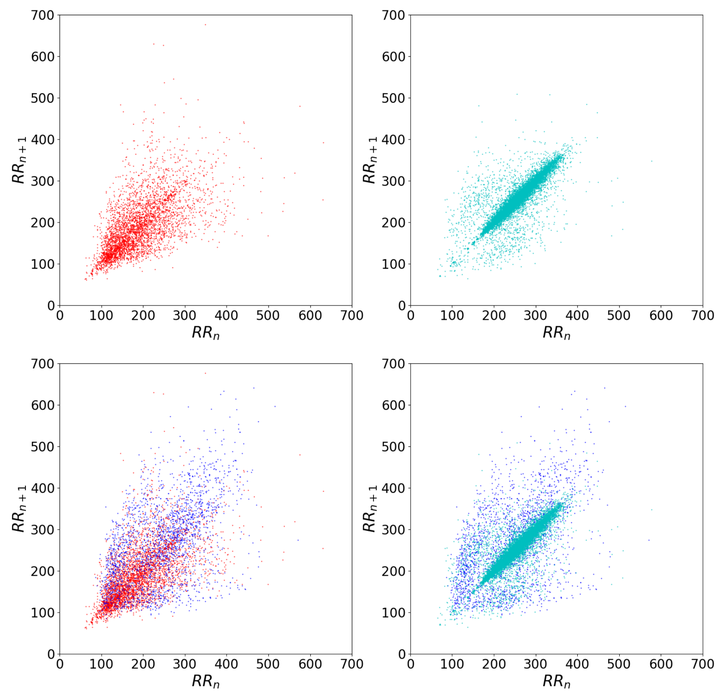
Evaluation – MSE, CC

Individual beats are extracted from each record and resampled to the same length in time. A median beat is obtained for each record. The mean square error (MSE) and correlation coefficient (CC) of each median beat in the record is computed against overall median beat of real normal and real AF records. The MSE and CC for the generated beats suggest the generated beats resemble more like real AF beats.

Source Record	Median Beat	MSE	CC
Normal Real	Normal Real	0.028 ± 0.039	0.736 ± 0.229
AF Real	AF Real	0.031 ± 0.042	0.629 ± 0.279
AF Generated	Normal Real	0.042 ± 0.045	0.541 ± 0.238
AF Generated	AF Real	0.036 ± 0.045	0.619 ± 0.262

Evaluation – Poincare Plots

The Poincare plot shows the correlation between two consecutive RR intervals. If the points are centered around the diagonal, low heart rate variability is indicated as in normal ECG. Scattered, dispersed points indicate higher variability in heart rates as in AF ECG. The generated AF ECG has a dispersed profile on the Poincare plots, indicating high RR variability. The following plot is from mid-stage training with scaling 6.0.



Legends: Red – Real AF, Teal – Real Normal, Blue- Generated AF

As in the sample ECG shown in the results, the Poincare plot also varies with training stages and scaling factors.

Conclusions

We were able to generate realistic looking AF ECG signals from normal ECG records that are similar to real AF in morphology with a WarpGAN based GAN structure. The generated signals have higher variability in RR-intervals compared to their input signals.

The generated signals are highly customizable using our GAN structure. According to our need, we could increase or decrease the amount of variability in RR-interval by introducing a scaling factor, or selecting results from early, mid, and late stages of training to have a complete AF dataset.

Our future work would be to:

- Apply new methods to evaluate our generated results.
- Apply the generated dataset towards training of an existing AF detection algorithm and see if the generated dataset improves the performance.

Acknowledgements

This work was supported by CDRH Critical Path funding and in party by an appointment to the Research Participation Program at the CDRH administrated by ORISE through an interagency agreement between the US Department of Energy and the US Food and Drug Administration.

References

G. Clifford et al. AF Classification from a Short Single Lead ECG Recording: the PhysioNet Computing in Cardiology Challenge 2017.

F. Castells et al. Multidimensional ICA for the Separation of Atrial and Ventricular Activities from Single Lead ECGs in Paroxysmal Atrial Fibrillation Episodes.

Shi, Yichun, Debayan Deb, and Anil K. Jain. WarpGAN: Automatic Caricature Generation.