CardioGAN: An Attention-based Generative Adversarial Network for Generation of Electrocardiograms

Subhrajyoti Dasgupta[1], Sudip Das[2], Ujjwal Bhattacharya[2]

[1] Amity University, Kolkata
[2] Indian Statistical Institute, Kolkata

Paper #965

Problem Statement

- Electrocardiogram(ECG) provides useful representation of the condition of human heart for diagnostic purposes
- Manual diagnosis of ailments from ECG is error-prone
- Automatic diagnostic systems can be learned using large volume of labelled ECG samples
- This gives rise to the data-imbalance problem
- Moreover, direct usage of medical records like ECG is highly prone to breaches of privacy

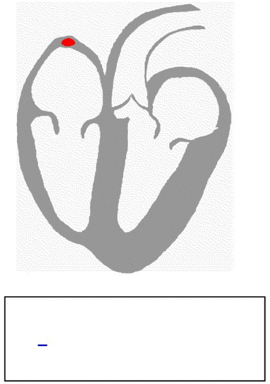

Fig ECG - a diagramatic record of the electrical activity of heart[1]

Can we generate synthetic data that closely replicates real human ECGs?

[1] https://en.wikipedia.org/wiki/Electrocardiography#/media/File:ECG_principle_slow.gif

Existing approaches

- A dynamical model based on three coupled ordinary differential equations [1]
- A nonlinear model based on Runge-Kutta integration of three ordinary differential equations to generate 24-hr long ECG [2]
- A model to generate artificial electrocardiogram signals based on certain reaction-diffusion system [3]
- Deep Learning based generative model to generate synthetic ECGs [4]

[1] McSharry et al., “A dynamical model for generating synthetic electrocardiogram signals,” IEEE Transactions on Biomedical Engineering, 2003.

[2] Clifford et al., “Generating 24-hour ECG, BP and respiratory signals with realistic linear and nonlinear clinical characteristics using a nonlinear model,” in Proceedings of Computers in Cardiology, 2004

[3] Quiroz-Juárez et al., “Generation of ECG signals from a reaction-diffusion model spatially discretized,” Scientific Reports, 2019.

[4] Zhu et al., “Zhu, F. Ye, Y. Fu, Q. Liu, and B. Shen, “Electrocardiogram generation with a bidirectional LSTM-CNN generative adversarial network,”Scientific Reports, vol. 9, no. 1, pp. 1–11, 2019.

Limitations of existing approaches

- Hardly possess human-like characteristics
- Dependent on predetermined initial parameters
- No well-defined PQRST curve
- Not consistent in terms of peak duration and characteristics

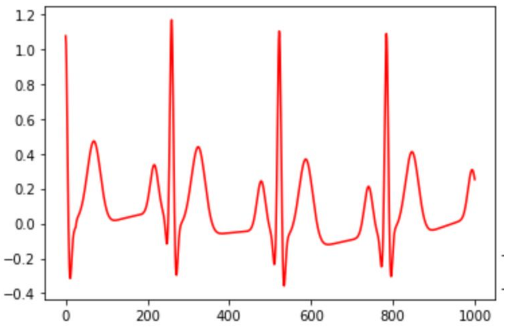

Fig - Results of ECGSYN [1]

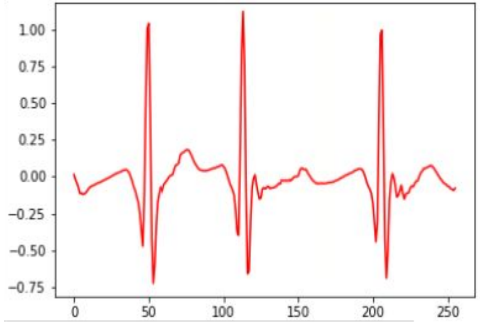

Fig - Results of Zhu et al. [2]

[1] McSharry et al., A dynamical model for generating synthetic electrocardiogram signals," IEEE Transactions on Biomedical Engineering, 2003.

[2] Zhu et al., "Zhu, F. Ye, Y. Fu, Q. Liu, and B. Shen, "Electrocardiogram generation with a bidirectional LSTM-CNN generative adversarial network,"Scientific Reports, vol. 9, no. 1, pp. 1–11, 2019.

Our Contributions

- Proposed an attention-based generator that improves the performance of the GAN to generate synthetic ECG data, by learning the intricate inter-dependencies of real samples
- An efficient loss function (Wasserstein GAN loss [1]) based on gradient penalty was used that improves training stability and, consequently, the synthetic ECG generation performance
- Several experiments for comparisons with different generative models are performed that demonstrate the improved performance of the proposed framework to generate synthetic samples that are morphologically similar to real-life ECGs

[1] Arjovsky et al., “Wasserstein Generative Adversarial Networks,” in ICML, 2017

Proposed Architecture - The Generator

Pre-Attention Layers:
- Stack of 2 bidirectional LSTM layers

Post-Attention Layers:
- Single LSTM layer
- Helps to generate the ECG signals based on the latent information and contexts

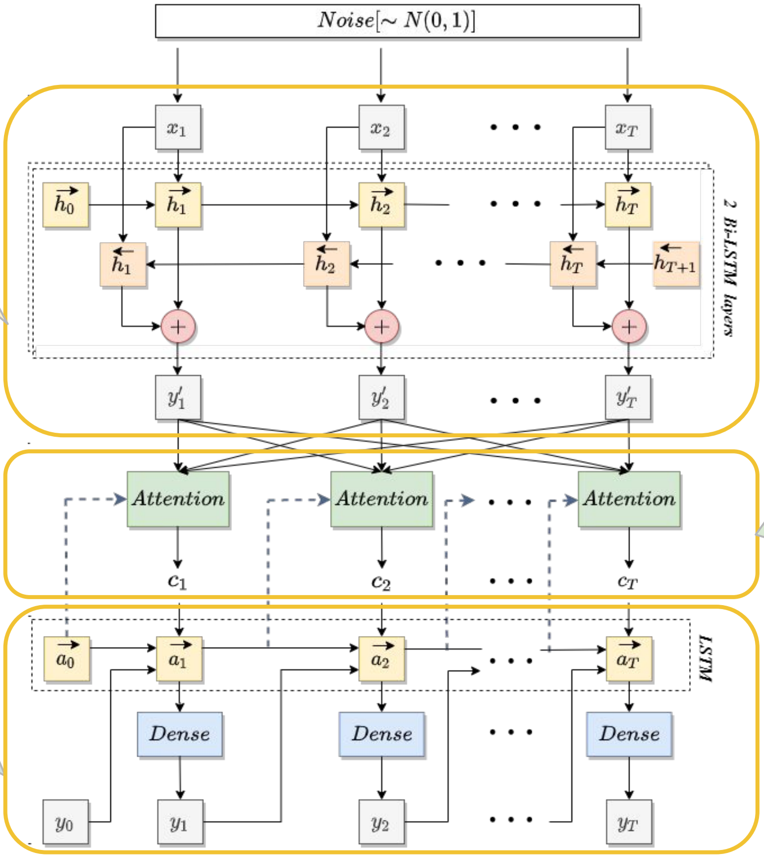

Fig - The generator of the architecture

Attention Layers:
- Dense feed-forward NNs
- Helps capture the intricate inter-dependencies

Proposed Architecture - The Critic

CONFIGURATION OF THE CRITIC WITH THE NUMBER OF PARAMETERS

LAYER	KERNEL SIZE/ STRIDE	OUTPUT SIZE	PARAMETERS
convolution(1D)	1×40 / 2	237×20	820
max-pooling(1D)	1×10 / 2	114×20	0
convolution(1D)	1×12 / 2	52×10	2410
max-pooling(1D)	1×5 / 2	24×10	0
convolution(1D)	1×6 / 2	10×5	305
max-pooling(1D)	1×2 / 1	9×5	0
linear	–	1×45	0
linear	–	1×1	46

Loss Function

- **Vanilla GAN [1] loss -**

$$L = E_{x \sim P_r}[log(D(x))] + E_{x' \sim P_g}[1 - log(D(x'))]$$

- **Wasserstein GAN [2] loss -**

$$L = E_{x' \sim P_g}[\,D(x')\,] - E_{x \sim P_r}[\,D(x)\,] \; + \\ \lambda E_{\hat{x} \sim P_{\hat{x}}}[\,(\,||\,\Delta \hat{x} D(\hat{x})\,||_2 - 1\,)^2\,]$$

PERFORMANCE COMPARISON OF CARDIOGAN WITH DIFFERENT LOSS FUNCTIONS

ARCHITECTURE	PRD	RMSE	FD	DTW
CardioGAN(with Wasserstein loss)	**38.566**	**0.157**	**0.606**	**4.146**
CardioGAN(with minimax loss)	46.784	0.186	0.652	5.732

[1] Goodfellow et al., “Generative adversarial nets,” in NIPS, 2014

[2] Arjovsky et al., “Wasserstein Generative Adversarial Networks,” in ICML, 2017

Experiments - Datasets

- **MIT-BIH Arrhythmia Database [1]**
  - Contains 48 half-hour excerpts of two-channel ambulatory ECG recordings, obtained from 47 subjects studied by the BIH Arrhythmia Laboratory. The recordings are sampled at 360Hz amounting to about 650,000 time-steps.

- **MIT-BIH Normal Sinus Rhythm Database [2]**
  - 18 long-term ECG recordings of subjects at Boston's Beth Israel Hospital. Subjects included in this database were found to have had no significant arrhythmias; they include 5 men aged 26 to 45, and 13 women aged 20 to 50.

[1] Moody et al., “The impact of the MIT-BIH arrhythmia database,” IEEE Engineering in Medicine and Biology Magazine, 2001
[2] Goldberger et al., “The MIT-BIH normal sinus rhythm database,” Circulation, vol. 101, no. 23, pp. e215–e220

Experiments - Evaluation Metrics

- Percentage Root Mean Square Difference (PRD) [1]
- Root Mean Square Error (RMSE) [1]
- Frechét distance (FD) [2]
- Dynamic Time Warping (DTW) [3]

[1] AlMahamdy et al., “Performance study of different denoising methods for ECG signals,” Procedia Computer Science, 2014.
[2] Aronov et al., “Fréchet distance for curves, revisited,” in European Symposium on Algorithms. Springer, 2006
[3] Cuturi et al., “Soft-DTW: a differentiable loss function for time-series,” in ICML. PMLR, 2017

Experiments - Performance Comparison

- **For the experiments:**
  - One lead signal was considered for both the datasets
  - Signals were sub-sampled by a rate of 2 time-steps and for each iteration of training, a window size of 512 time-steps for each batch was considered.
  - To avoid unnecessary biasing, the batch size was fixed to the number of patients in each of the datasets.

QUANTITATIVE COMPARISON OF THE PERFORMANCE OF DIFFERENT ARCHITECTURES BASED ON DIFFERENT EVALUATION METRICS

ARCHITECTURE	PRD	RMSE	FD	DTW
CardioGAN [ours]	**38.566**	**0.157**	**0.606**	**4.146**
BiLSTM-GAN	66.408	0.276	0.756	9.375
ECGSYN	78.331	0.363	0.784	15.737
RNN-AE	121.877	0.506	0.969	–
LSTM-AE	148.650	0.618	0.996	–
RNN-VAE	146.656	0.609	0.982	–
LSTM-VAE	145.978	0.607	0.975	–

Experiments - Performance Comparison

PERFORMANCE OF CARDIOGAN WITH DIFFERENT CRITIC ARCHITECTURES

ARCHITECTURE	PRD	RMSE	FD	DTW
CardioGAN(with 1D-CNN critic)	**38.566**	**0.157**	**0.606**	**4.146**
CardioGAN(with GRU critic)	59.674	0.294	0.762	9.604
CardioGAN(with LSTM critic)	65.101	0.327	0.781	11.871
CardioGAN(with RNN critic)	78.142	0.375	0.824	10.848

Experiments - Visual Results

ICPR 2020

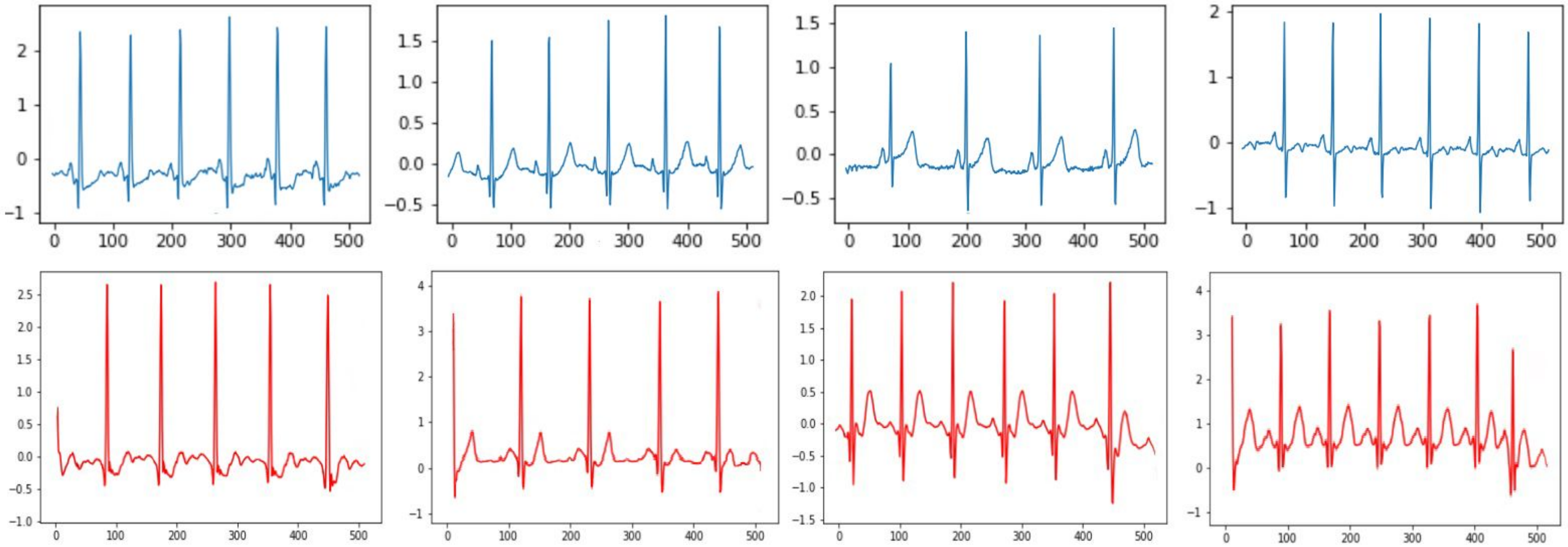

Fig - Plots representing the ECG signals in 521 timesteps.

Blue - Ground Truth Red - ECGs generated by CardioGAN

Conclusions

- A novel attention-based GAN is proposed for generation of synthetic ECGs
- Synthetic ECGs generated by CardioGAN possess human-like characteristics such as well-defined PQRST waves, consistent wave patterns, etc.
- Further development in generating synthetic human-like medical data can help preserve privacy
- We also plan to extend our research using Federated Learning [1] to further foster the privacy aspect in medical research.

[1] Kairouz et al., “Advances and open problems in federated learning,” arXiv preprint, arXiv:1912.04977, 2019.

Thank You!

Subhrajyoti Dasgupta - subhrajyotidg@gmail.com
Sudip Das - d.sudip47@gmail.com
Ujjwal Bhattacharya - ujjwal@isical.ac.in