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

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

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]

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

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

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

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

Fig - The generator of the architecture
## Proposed Architecture - The Critic

### Configuration of the Critic with the Number of Parameters

<table>
<thead>
<tr>
<th>LAYER</th>
<th>Kernel Size/Stride</th>
<th>Output Size</th>
<th>Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>convolution(1D)</td>
<td>1×40 / 2</td>
<td>237×20</td>
<td>820</td>
</tr>
<tr>
<td>max-pooling(1D)</td>
<td>1×10 / 2</td>
<td>114×20</td>
<td>0</td>
</tr>
<tr>
<td>convolution(1D)</td>
<td>1×12 / 2</td>
<td>52×10</td>
<td>2410</td>
</tr>
<tr>
<td>max-pooling(1D)</td>
<td>1×5 / 2</td>
<td>24×10</td>
<td>0</td>
</tr>
<tr>
<td>convolution(1D)</td>
<td>1×6 / 2</td>
<td>10×5</td>
<td>305</td>
</tr>
<tr>
<td>max-pooling(1D)</td>
<td>1×2 / 1</td>
<td>9×5</td>
<td>0</td>
</tr>
<tr>
<td>linear</td>
<td>–</td>
<td>1×45</td>
<td>0</td>
</tr>
<tr>
<td>linear</td>
<td>–</td>
<td>1×1</td>
<td>46</td>
</tr>
</tbody>
</table>
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** -
  \[
  L = E_{x' \sim P_g}[D(x')] - E_{x \sim P_r}[D(x)] + \lambda E_{\hat{x} \sim P_x}[\left(\| \Delta \hat{x} D(\hat{x}) \|_2 - 1 \right)^2]
  \]

---

**Performance Comparison of CardioGAN with Different Loss Functions**

<table>
<thead>
<tr>
<th>Architecture</th>
<th>PRD</th>
<th>RMSE</th>
<th>FD</th>
<th>DTW</th>
</tr>
</thead>
<tbody>
<tr>
<td>CardioGAN (with Wasserstein loss)</td>
<td>38.566</td>
<td>0.157</td>
<td>0.606</td>
<td>4.146</td>
</tr>
<tr>
<td>CardioGAN (with minimax loss)</td>
<td>46.784</td>
<td>0.186</td>
<td>0.652</td>
<td>5.732</td>
</tr>
</tbody>
</table>

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.

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]

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.

<table>
<thead>
<tr>
<th>ARCHITECTURE</th>
<th>PRD</th>
<th>RMSE</th>
<th>FD</th>
<th>DTW</th>
</tr>
</thead>
<tbody>
<tr>
<td>CardioGAN [ours]</td>
<td>38.566</td>
<td>0.157</td>
<td>0.606</td>
<td>4.146</td>
</tr>
<tr>
<td>BiLSTM-GAN</td>
<td>66.408</td>
<td>0.276</td>
<td>0.756</td>
<td>9.375</td>
</tr>
<tr>
<td>ECGSYN</td>
<td>78.331</td>
<td>0.363</td>
<td>0.784</td>
<td>15.737</td>
</tr>
<tr>
<td>RNN-AE</td>
<td>121.877</td>
<td>0.506</td>
<td>0.969</td>
<td>–</td>
</tr>
<tr>
<td>LSTM-AE</td>
<td>148.650</td>
<td>0.618</td>
<td>0.996</td>
<td>–</td>
</tr>
<tr>
<td>RNN-VAE</td>
<td>146.656</td>
<td>0.609</td>
<td>0.982</td>
<td>–</td>
</tr>
<tr>
<td>LSTM-VAE</td>
<td>145.978</td>
<td>0.607</td>
<td>0.975</td>
<td>–</td>
</tr>
</tbody>
</table>
### Performance of CardioGAN with Different Critic Architectures

<table>
<thead>
<tr>
<th>Architecture</th>
<th>PRD</th>
<th>RMSE</th>
<th>FD</th>
<th>DTW</th>
</tr>
</thead>
<tbody>
<tr>
<td>CardioGAN (with 1D-CNN critic)</td>
<td>38.566</td>
<td>0.157</td>
<td>0.606</td>
<td>4.146</td>
</tr>
<tr>
<td>CardioGAN (with GRU critic)</td>
<td>59.674</td>
<td>0.294</td>
<td>0.762</td>
<td>9.604</td>
</tr>
<tr>
<td>CardioGAN (with LSTM critic)</td>
<td>65.101</td>
<td>0.327</td>
<td>0.781</td>
<td>11.871</td>
</tr>
<tr>
<td>CardioGAN (with RNN critic)</td>
<td>78.142</td>
<td>0.375</td>
<td>0.824</td>
<td>10.848</td>
</tr>
</tbody>
</table>
Experiments - Visual Results

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.

Thank You!

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