Differentially Private GAN for Time Series

CSE3000 Research Project
Pepijn te Marvelde, P.H.temarvelde@student.tudelft.nl

Responsible Professor: Dr. Lydia Y. Chen
Supervisors: Aditya Kunar & Zilong Zhao

1. Background Information

Generative Adversarial Networks (GANs) [1]

- Deep learning model used **to generate synthetic data**.
- **Joint learning** of Generator and Discriminator.
- Generator creates fake data.
- Discriminator tries to distinguish real from fake.

Differential Privacy (DP) [2]

- Mathematical framework to quantify level of privacy.
- If including information of an individual does **not change**, e.g., the **result** of an aggregation, or query, on a dataset, then that person is likely not opposed to being included in the dataset.
- DP **limits** the **effect** of a **single sample has on a mechanism**.
  - Achieved by adding carefully constructed **noise**.
- **Quantitative** level of privacy, ε. Higher ε means less privacy.

Differential Privacy in GANs

- If the **training procedure** of a GAN is made DP, the generated dataset is guaranteed to adhere to a certain level of privacy, ε.
- Can be implemented in either the Generator, or the Discriminator.
  - Done respectively in GS-WGAN [3] and DP-CGAN [4].

2. Research questions

i. How does GS-WGAN's performance compare to DP-CGAN.

- **Verify** GS-WGAN performance through **quantitative** and **qualitative** analysis of generated images.

ii. Is GS-WGAN capable of generating high quality **synthetic time series**?

- Does performance of an image-based GAN **carry over** to time series?

3. Comparing GS-WGAN and DP-CGAN

Train both GANs on the MNIST datasets with same privacy budget: (ε,δ)=(10.0, 10e-5), and **compare** results.

- **Qualitative** comparison:

Model	MNIST	Fashion-MNIST
Real		
GS-WGAN		
DP-CGAN		

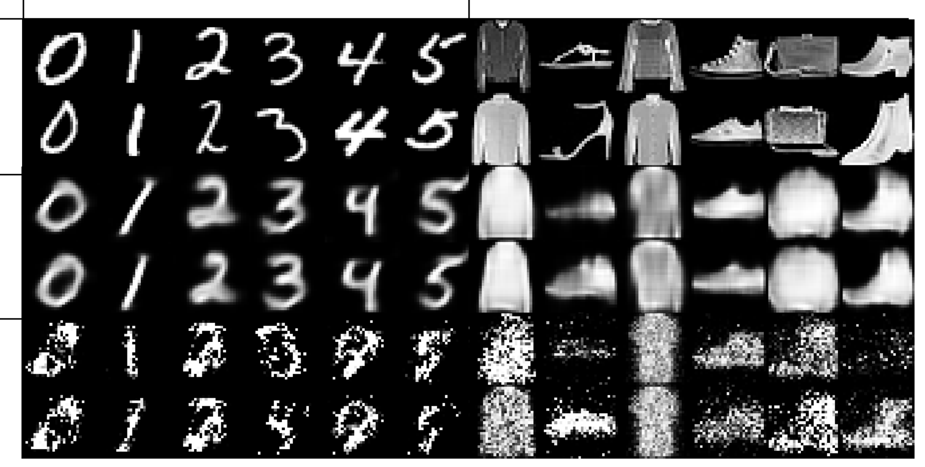


Table 1: Qualitative comparison of GS-WGAN and DP-CGAN generated samples for MNIST and Fashion-MNIST datasets with a privacy budget of (ε,δ)=(10.0, 10e-5).

- **Quantitative** comparison:
  - **Assess quality** of dataset through different **metrics**.
  - Most representative: accuracy of a classifier trained on synthetic data, **tested** on **real** samples. See Figure 1.

➢ GS-WGAN, on average, **performs 40% better** than DP-CGAN, consistently generating **clearer**, **less noisy** images for a privacy budget of ε = 10.0.

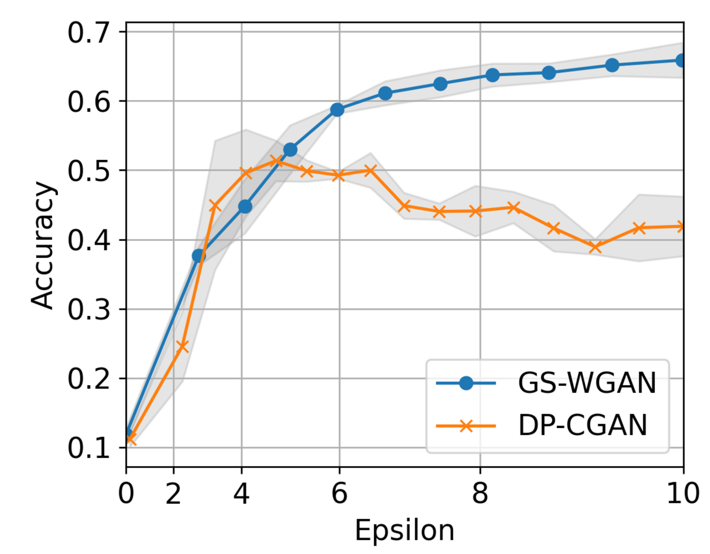
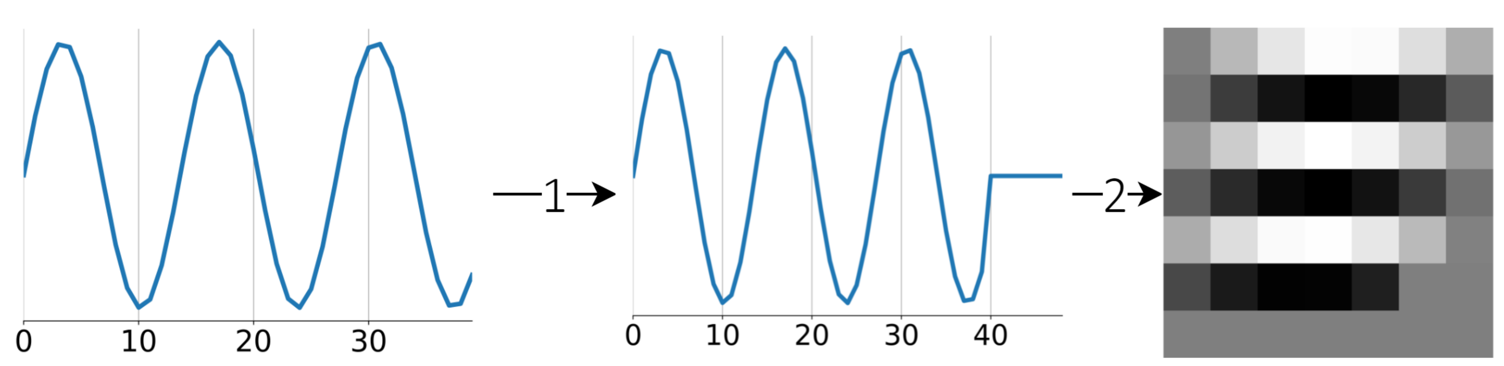


Figure 1: Downstream classifier accuracy on MNIST dataset divided by baseline classifier accuracy versus spent privacy budget ε. Higher accuracy is better. 3

4. DP Generation of Time Series

GS-WGAN not directly applicable to time series:

- **Convert** time series into **images**:
  1. **Pad** data to a square.
  2. Normalize and **reshape** to convert to an image.



Data shape: (40, 1) (49, 1) (7, 7)

Figure 2: Schematic showing the process of converting a time series sample into an image. Example shows a sine wave of length 40 converted into an 7x7 grayscale image.

- **Compare** performance against state-of-the-art [5]. ECG dataset (PTB): each sample is one heartbeat, captured over 187 timesteps.
  - Task: classify if normal or abnormal hearth rhythm.

	PATE	RDP-CGAN	DPGAN	GS-WGAN ResNet	GS-WGAN DCGAN
AUROC	0.75 ± 0.012	**0.79 ± 0.009**	0.71 ± 0.012	0.47 ± 0.051	0.50 ± 0.121
AUPRC	0.76 ± 0.011	0.80 ± 0.008	0.71 ± 0.018	**0.83 ± 0.069**	0.80 ± 0.083

Table 2: GS-WGAN comparison versus baselines for a low privacy setting with (ε,δ)=(1.0, 10e-5). AUROC is Area under Receiver Operating Characteristic Curve, AUPRC is Area Under Precision Recall curve. Best and second-best values are respectively bold and underlined. Higher values are better.

➢ Low privacy budget: Sub-optimal performance, not enough time to learn.

➢ High privacy budget: Good performance. Model is fit for timeseries generation.

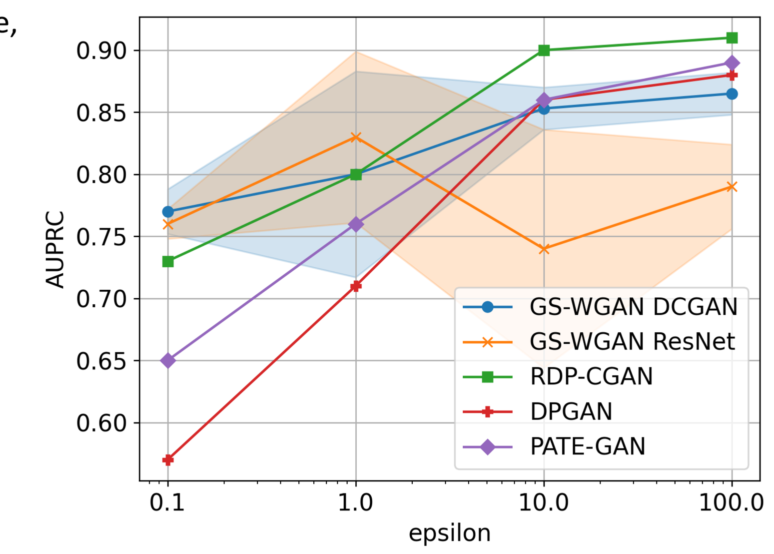


Figure 3: The effect of spent privacy budget on the quality of generated dataset measured by AUPRC.

5. Conclusions

➢ GS-WGAN **outperforms** DP-CGAN when tasked with generating MNIST images in a differentially private setting.

➢ Wrapping time series samples as images **allows image-based** GANs to be used on **time series**.

➢ GS-WGAN can generate high quality time series for the PTB dataset, **on par** with **best performing DP-GANs** for high privacy budget (ε > 10.0) setting.

References

1. I. Goodfellow, J. Pouget-Abadie, M. Mirza, B.Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative Adversarial Nets," in Advances in Neural Information Processing Systems, Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, and K.Q. Weinberger, Eds., vol. 27, Curran Associates, Inc.,2014. [Online]. Available: https://proceedings.neurips.cc/paper/2014/file/5ca3e9b122f61f8f06494c97b1afccf3-Paper.pdf
2. C. Dwork, "Differential Privacy: A Survey of Results," in Lecture Notes in Computer Science, Springer Berlin Heidelberg, pp. 1–19. DOI: 10.1007/978-3-540-79228-4 1. [Online]. Available: https://link.springer.com/chapter/10.1007%2F978-3-540-79228-4_1.
3. D. Chen, T. Orekondy, and M. Fritz, GSWGAN: A Gradient-Sanitized Approach for Learning Differentially Private Generators, 2021. arXiv: 2006.08265 [cs.LG]. Available: https://arxiv.org/abs/2006.08265.
4. R. Torkzadehmahani, P. Kairouz, and B. Paten, DP-CGAN: Differentially Private Synthetic Data and Label Generation, 2020. arXiv:2001.09700 https://arxiv.org/abs/2001.09700.
5. A. Torfi, E. A. Fox, and C. K. Reddy, "Differentially private synthetic medical data generation using convolutional gans," 2020.