Pixel-wise Conditioning of Generative Adversarial Networks

Cyprien Ruffino[1] , Romain Hérault[1] , Eric Laloy[2] and Gilles Gasso[1]

December 6, 2019

[1]Normandie Univ, UNIROUEN, UNIHAVRE, INSA Rouen, LITIS, 76 000 Rouen, France
[2]Belgian Nuclear Research, Institute Environment, Health and Safety, Boeretang 200 - BE-2400 Mol, Belgium

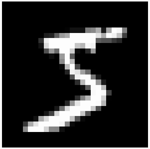
Problem

Objective

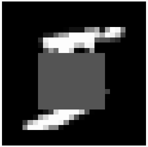
- Image reconstruction task
- Generation under pixel constraints
- Motivation: applications in geosciences

Differences with inpainting

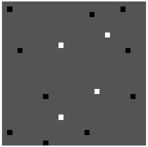
- Very few information ($\sim 0.5\%$)
- Full-size image generation
- Unstructured information



(a) Original Image



(b) Regular Inpainting



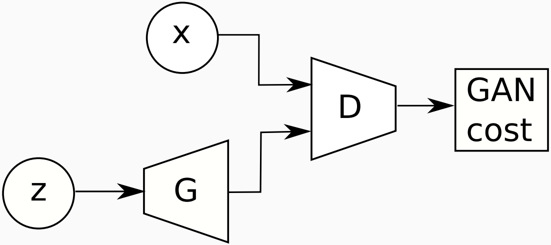
(c) Pixel Constraints

Generative Adversarial Networks [GPAM+14]

Two networks, a generator G and a discriminator D :

- **Generator**: produces synthetic data from a random $z \sim P_z$, where P_z is a known distribution
- **Discriminator**: binary classifier, tries to distinguish real samples from fake ones

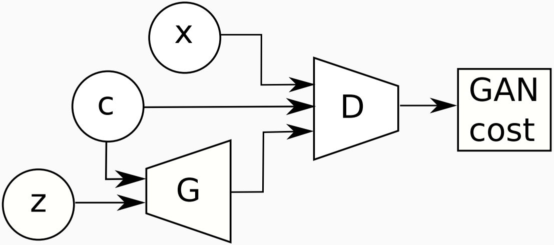
$$\min_G \max_D L(D, G) = \mathop{\mathbb{E}}_{X \sim P_r} \left[\log(D(X)) \right] + \mathop{\mathbb{E}}_{z \sim P_z} \left[\log(1 - D(G(z))) \right]$$



Conditional GAN [MO14]

- Conditional variant of the GANs
- A constraint/label c is simply given as an input to both G and D
- Works well for generating image with a class constraints

$$\min_G \max_D L(D, G) = \mathop{\mathbb{E}}_{\substack{X \sim P_r \\ C \sim P_{C|X}}} \Big[\log(D(X, C))\Big] + \mathop{\mathbb{E}}_{\substack{z \sim P_z \\ C' \sim P_C}} \Big[\log(1 - D(G(z, C'), C'))\Big]$$



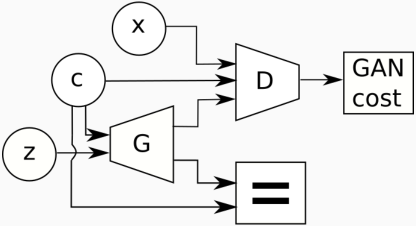
Constrained CGAN

Theoretical objective

Explicit verification, hard-constraint on the respect of C

$$\min_G \max_D L(D,G) = \mathop{\mathbb{E}}_{\substack{X\sim P_r \\ C\sim P_{C|X}}} \Big[\log(D(X,C))\Big] + \mathop{\mathbb{E}}_{\substack{z\sim P_z \\ C'\sim P_C}} \Big[\log(1-D(G(z,C'),C'))\Big]$$

s.c. $C = M(C) \odot G(z,C)$



where $M(C)$ gives the binary mask of the constraints

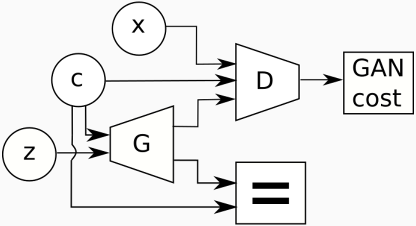
Constrained CGAN

Theoretical objective

Explicit verification, hard-constraint on the respect of C

$$\min_G \max_D L(D,G) = \mathop{\mathbb{E}}_{\substack{X\sim P_r \\ C\sim P_{C|X}}}\Big[\log(D(X,C))\Big] + \mathop{\mathbb{E}}_{\substack{z\sim P_z \\ C'\sim P_C}}\Big[\log(1-D(G(z,C'),C'))\Big]$$

s.c. $C = M(C) \odot G(z,C)$



where $M(C)$ gives the binary mask of the constraints

Problem

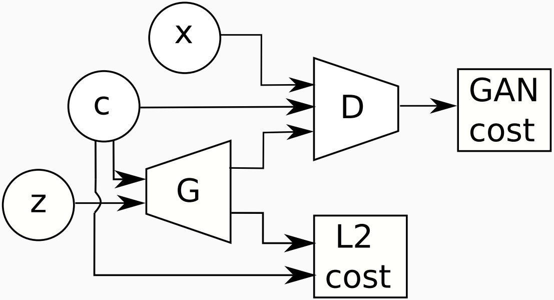
Strictly constrained objective is non-differentiable

Relaxation of the constrained CGAN

Our approach

Relaxation of the strict constraints by a regularization term

$$\min_G \max_D L_{reg}(D,G) = L(D,G) + \lambda \mathop{\mathbb{E}}_{\substack{z\sim P_z \\ C\sim P_C}} \left[\|C - M(C) \odot G(z,C)\|_2^2 \right]$$



Experiments i

Task: Hyperparameter search on λ

- Objective: find evidence of a controllable trade-off between quality and respect of the constraints
- Experiments repeated 10 times each

Metrics

- Respect of the constraints: Mean Square Error on constrained pixels
- Visual quality: Fréchet Inception Distance [HRU+17]:
  distance between the distributions of the features of real and generated samples at the output of a deep classifier.

Experiments ii

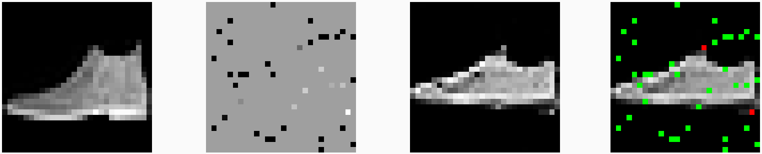
Datasets

- MNIST and FashionMNIST
- Split in train, validation and test sets
- 10% of each set used to sample constraints, then discarded

Networks architecture

- DCGAN [RMC15]-like, with only 2 convolutional/transposed convolutional layers in D and G

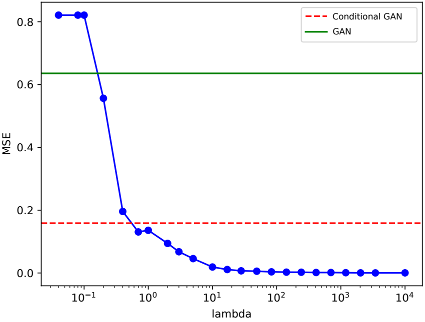
Results on FashionMNIST

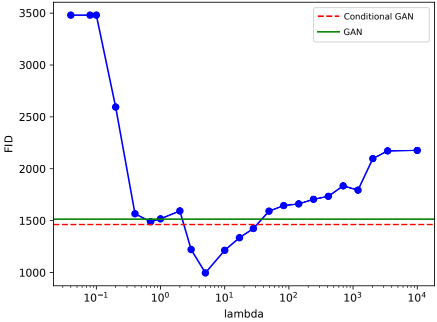


(a) Original Image (b) Constraints (c) Generated Image (d) Satisfied Consts.

This method can generate samples that respects pixel precise constraints

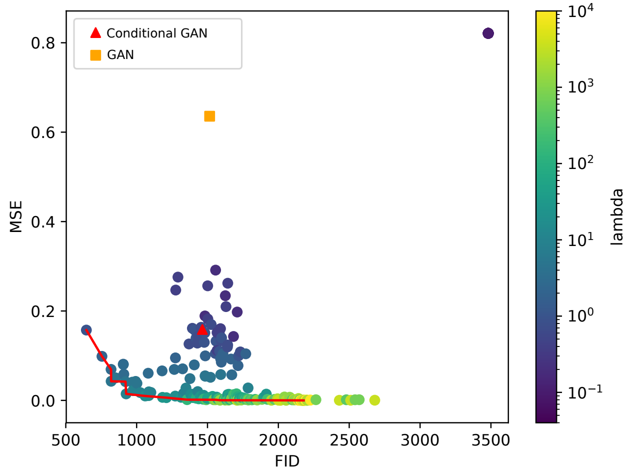
MSE / FID relative to λ



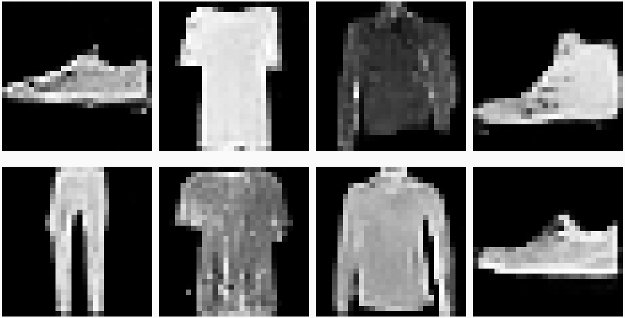


- The constraints seem able to enhance quality
- From $\lambda = 0.1$ upwards, there seem to be a trade-off between MSE and FID

Results on FashionMNIST



- Trade-off clearly visible
- Adding constraints can enhance visual quality
- Regularization enhance both quality and respect of constraints



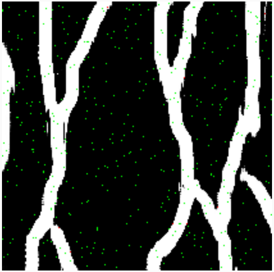
Some generated samples at $\lambda = 1$ (best ratio between quality and respect of the constraints)

Conclusion

- Conditional GANs can learn pixel-wise constraints
- The $L2$ regularization term allows to control a trade-off between visual quality and respect of the constraints

Extensions

- Applications on real-world datasets
- Extension to other kind of constraints (moments on zones, ...)



Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio.
Generative adversarial nets.
In *Advances in neural information processing systems*, pages 2672–2680, 2014.

Martin Heusel, Hubert Ramsauer, Thomas Unterthiner, Bernhard Nessler, and Sepp Hochreiter.
Gans trained by a two time-scale update rule converge to a local nash equilibrium.
In *Advances in Neural Information Processing Systems*, pages 6626–6637, 2017.

Mehdi Mirza and Simon Osindero.
Conditional generative adversarial nets.
arXiv preprint arXiv:1411.1784, 2014.

Alec Radford, Luke Metz, and Soumith Chintala.
Unsupervised representation learning with deep convolutional generative adversarial networks.
arXiv preprint arXiv:1511.06434, 2015.