

GANs

Group Ganglions

Constant Bourdrez, Mohamed Ali Srir, Sebastiano Scardera

Optimal Transport methods for GAN's improvement
Data Science Lab - IASD Master



November 2024

Table of contents

Discriminator Optimal Transport

- Presentation of OT method

- Results on latent space and on target space

Improving Vanilla GANs

- Effect of spectral normalization

- WGAN vs Vanilla GAN

Future works

- Embedded OT

- Metrics vs Geometry

- Further future works

Discriminator Optimal Transport

Main idea

Improve image generation using Discriminator after training.

If $y = G(z)$,

$$\tilde{y} = \operatorname{argmin}_x \left\{ \underbrace{\|x - y\|_2}_{\text{Local}} - \underbrace{\frac{1}{K} D(x)}_{\text{Improve}} \right\} \text{ s.t. } D(\tilde{y}) > D(y).$$

Target space version

$$\tilde{y} = \operatorname{argmin}_x \{ \|x - y\|_2 - D(x) \}$$

Latent space version

$$\tilde{z}_y = \operatorname{argmin}_z \{ \|z - z_y\|_2 - D(z) \}$$

Transport of images in the latent space

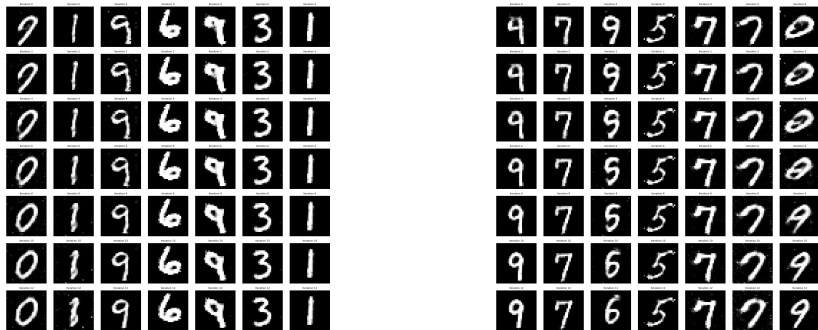
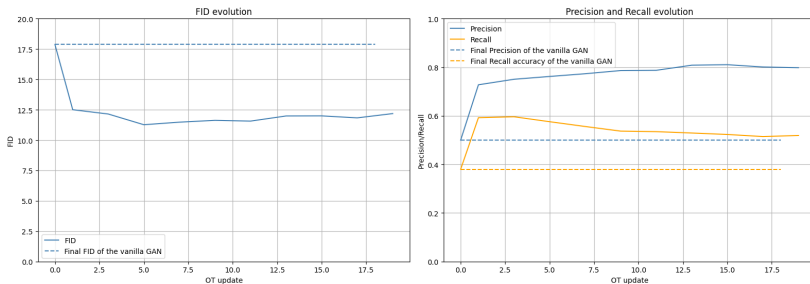


Figure: Both images show Optimal Transport processes during iterations.

Performance of OT method on latent space

Metrics evolution for the latent space OT



Transport of images in the target space

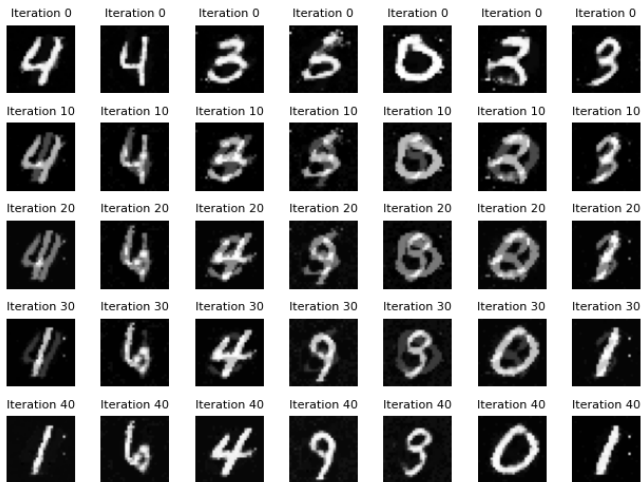
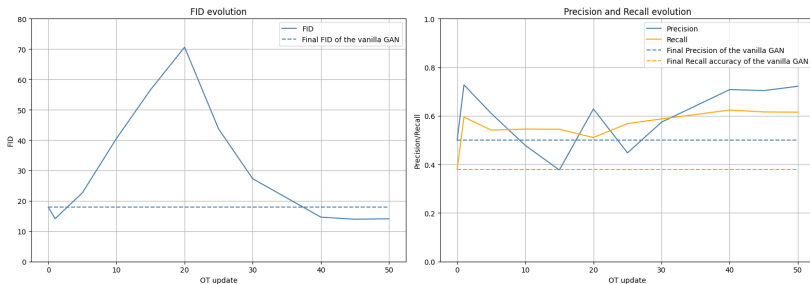


Figure: Optimal Transport processes during iterations.

Performance of OT method on target space

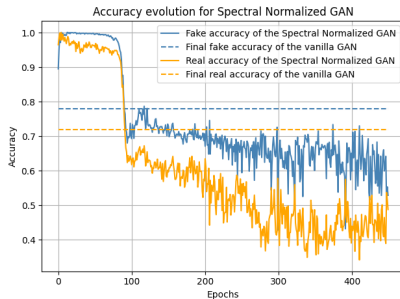
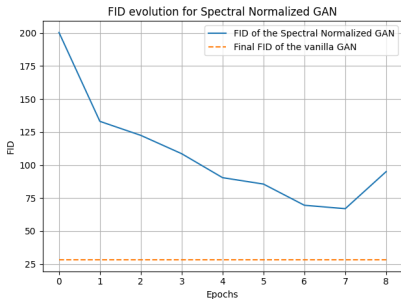
Metrics evolution for the target OT



Improving Vanilla GANs

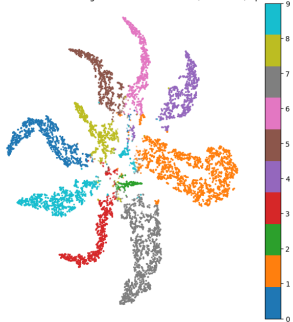
Spectral normalization formula

$$W_{SN} = W / \sigma(W) \text{ where } \sigma(W) = \max_{\|x\|=1} \|Wx\|$$

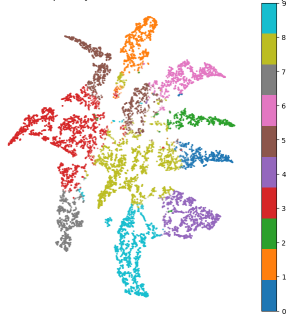


Effect of spectral normalization

MNIST Fake Dataset with Original Generator - VGG 5 (fine-tuned) Space



MNIST Fake Dataset with Spectrally Normalized Generator - VGG 5 (fine-tuned) Space



Traditional GAN

- ▶ **Loss Function:**
 - ▶ Uses Jensen-Shannon divergence
 - ▶ Binary cross-entropy loss
- ▶ **Architecture:**
 - ▶ Discriminator outputs $[0,1]$
 - ▶ Final sigmoid activation
 - ▶ No constraint on weights
- ▶ **Training Issues:**
 - ▶ Mode collapse common
 - ▶ Vanishing gradients
 - ▶ Training instability

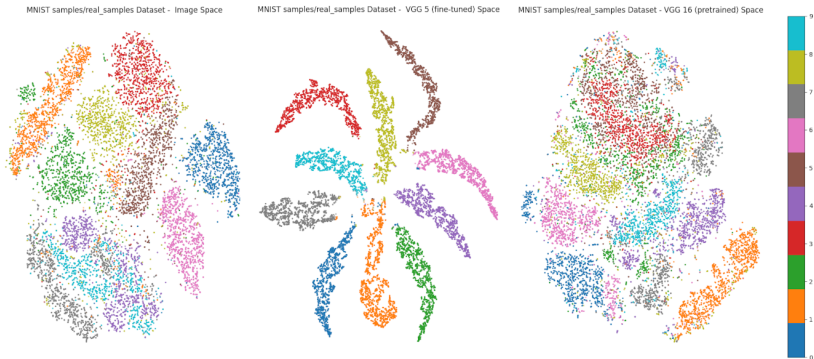
Wasserstein GAN

- ▶ **Loss Function:**
 - ▶ Uses Wasserstein distance
 - ▶ Earth mover's distance
- ▶ **Architecture:**
 - ▶ Critic outputs $(-\infty, \infty)$
 - ▶ No sigmoid activation
 - ▶ Weight clipping
- ▶ **Advantages:**
 - ▶ Better stability
 - ▶ Meaningful loss metric
 - ▶ Improved gradients

Future works

Geometry of the different embeddings

Different embeddings leads to different manifolds



Embedded OT

Previous version

$$\tilde{y} = \operatorname{argmin}_x \{ \|x - y\|_2 - D(x) \}$$



Embedded OT

Let E be a (differentiable) embedding map and $y = G(z)$ then

$$\tilde{y} = \operatorname{argmin}_x \{ \|E(x) - E(y)\|_2 - D(x) \}$$

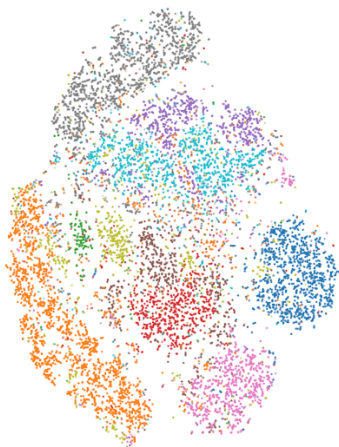


Metrics depend on the embedding

MNIST samples/real_samples Dataset - Image Space



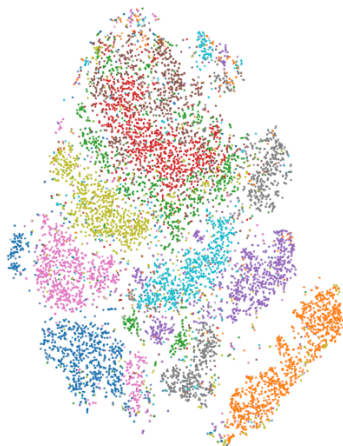
MNIST samples/fake_samples Dataset - Image Space



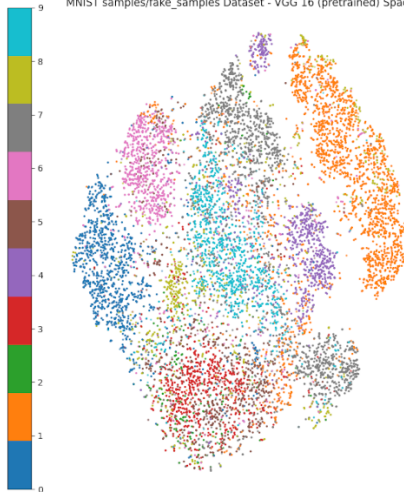
Precision = 0.78, Recall = 0.28

Metrics depend on the embedding

MNIST samples/real_samples Dataset - VGG 16 (pretrained) Space



MNIST samples/fake_samples Dataset - VGG 16 (pretrained) Space



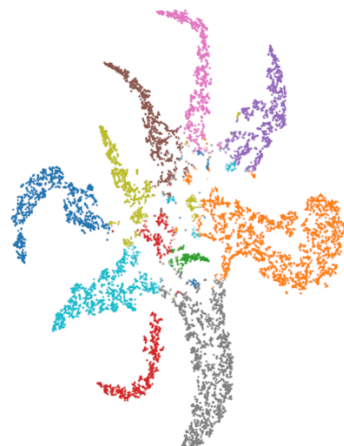
Precision = 0.50, Recall = 0.38

Metrics depend on the embedding

MNIST samples/real_samples Dataset - VGG 5 (fine-tuned) Space



MNIST samples/fake_samples Dataset - VGG 5 (fine-tuned) Space



Precision = 0.84, Recall = 0.77

- ▶ Better understanding of the behavior of the OT model on the target space
- ▶ Try other distributions on latent space
- ▶ Implement the methods with the normalized Discriminator:

$$\tilde{y} = \operatorname{argmin}_x \left\{ \|x - y\|_2 - \frac{1}{K} D(x) \right\}$$

Thank you!

Tanaka, A. (2019). Discriminator optimal transport. *Advances in Neural Information Processing Systems*, 32.