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 (Tanaka (2019))

Idea: Improve image generation using Discriminator after training.
If $\mathrm{y} = G(\mathrm{z})$,

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

Target space version

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

Latent space version

$$\tilde{\mathrm{z}}_{\mathrm{y}} = \operatorname{argmin}_{\mathrm{z}} \{\|\mathrm{z}-\mathrm{z}_{\mathrm{y}}\|_2 - D(\mathrm{z})\}$$

Transport of images in the latent space

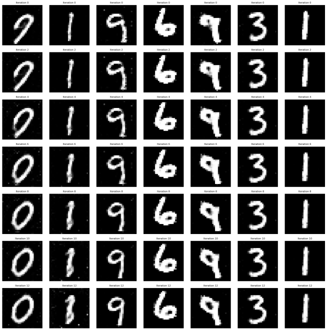
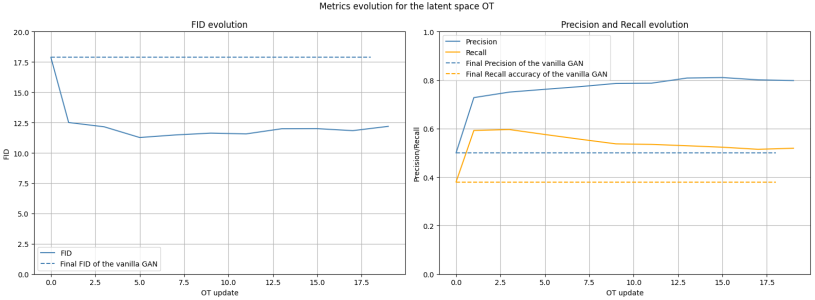


Figure: Both images show Optimal Transport processes during iterations.

Performance of OT method on latent space



Transport of images in the target space

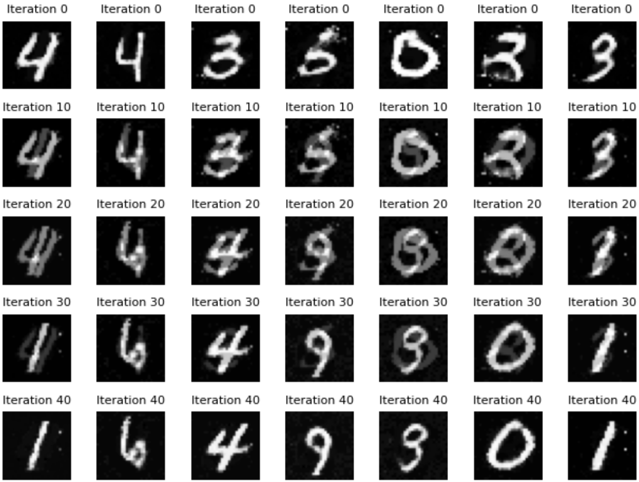
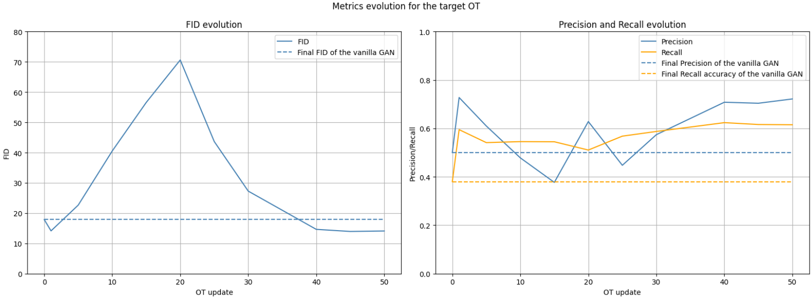


Figure: Optimal Transport processes during iterations.

Performance of OT method on target space

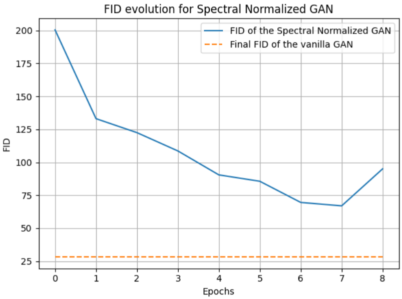


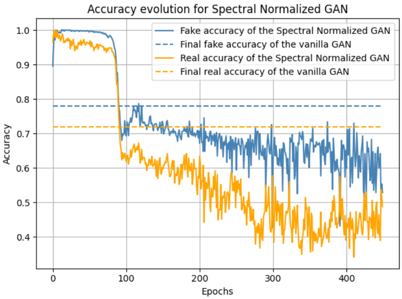
Improving GANs

Effect of spectral normalization

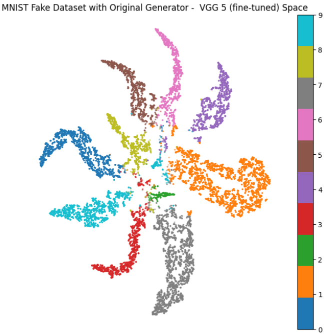
Spectral normalization formula

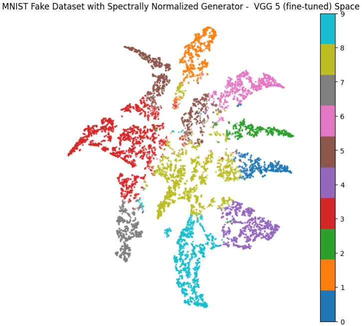
$$W_{SN} = W/\sigma(W) \text{ where } \sigma(W) = \max_{\|x\|=1} \|Wx\|$$





Effect of spectral normalization





WGAN vs Vanilla GAN

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

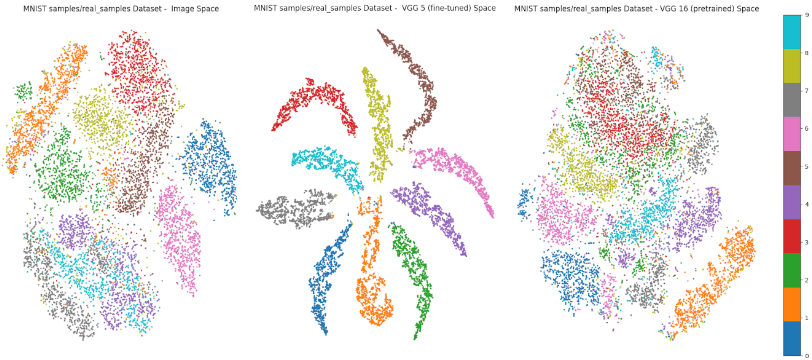
Wasserstein GAN

- **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

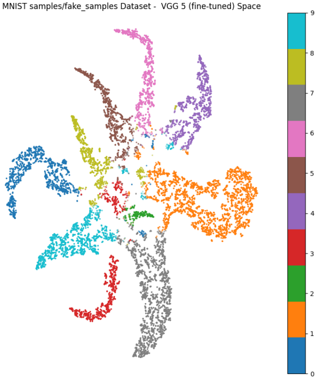
Different embeddings leads to different manifolds



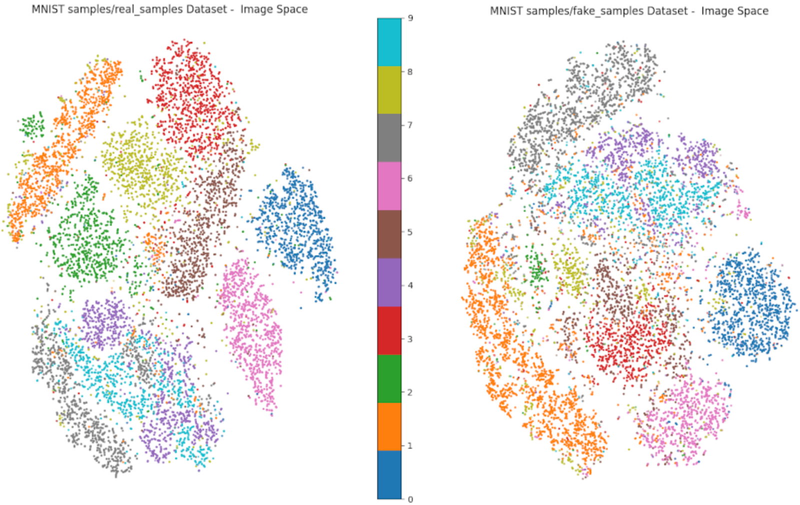
Embedded OT

Let E be the (differentiable) embedding map and $\mathrm{y} = G(\mathrm{z})$ then

$$\tilde{\mathrm{y}} = \mathrm{argmin}_{\mathrm{x}} \left\{ \|E(\mathrm{x}) - E(\mathrm{y})\|_2 - D(\mathrm{x}) \right\}$$

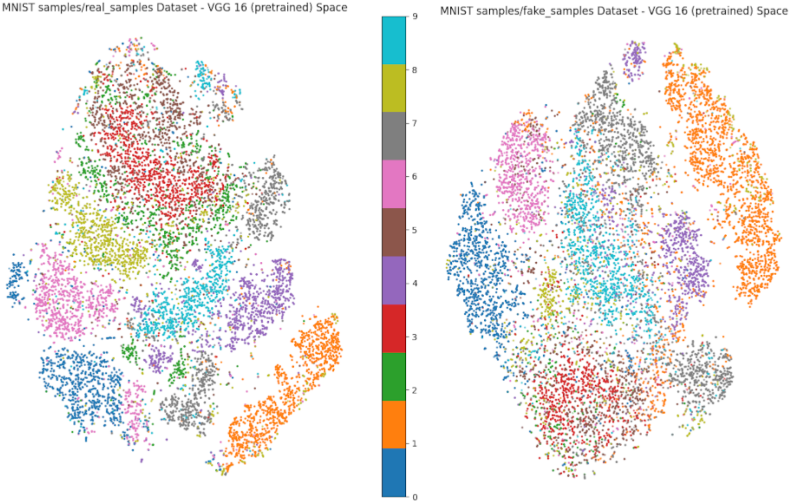


Metrics depend on the embedding



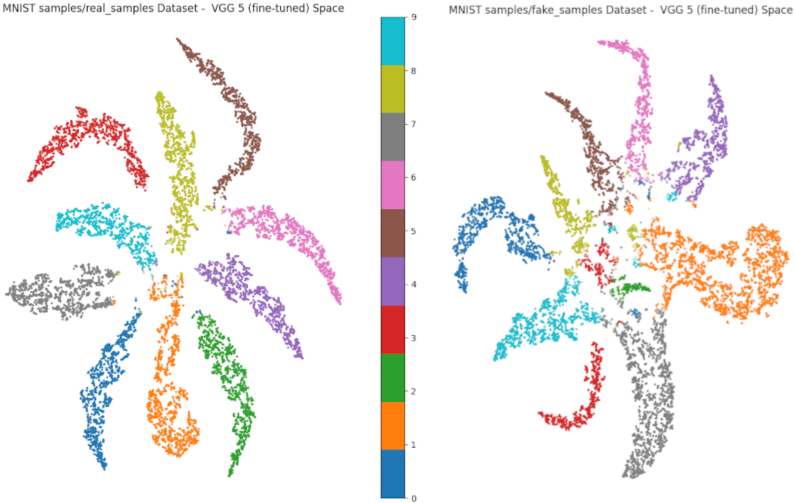
Precision = 0.78, Recall = 0.28

Metrics depend on the embedding



Precision $= 0.50$, Recall $= 0.38$

Metrics depend on the embedding



Precision $= 0.84$, Recall $= 0.77$

Further future works

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

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

Thank you!

References

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