Advancing GANs: Optimizing Latent Representations and Loss Functions

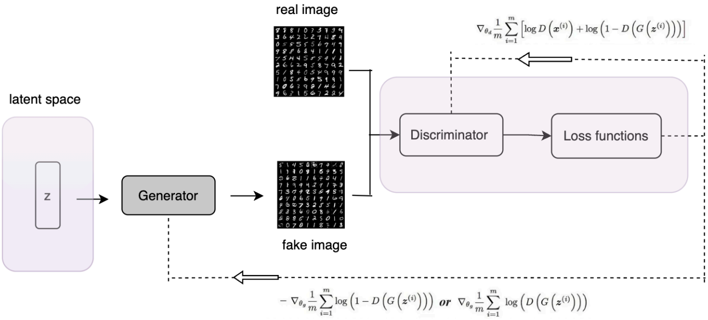
Data Science Lab A2

Team: Yesterday
Xichen ZHANG
Mariem AALABOU
Hangyue ZHAO

November 6, 2024

Table of Contents

Fundamentals of GANs



A2 GAN's Playground

The G architecture is fixed, there exist 4 playgrounds (1.discriminator, 2.losses functions, 3.latent space, 4.Optimization) and our focus is on two main areas for experimentation and improvement:

- Loss function (W-GAN, W-GAN-GP)
- Latent space(DRS, DDLS, Dgflow)

Challenges with Strong Discriminators

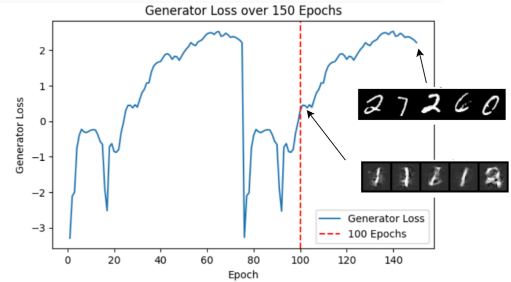


Figure: Generator Loss with a Discriminator Enhanced by Convolutional Layers

Strong discriminators can overpower the generator. This leads to:

- Mode collapse: The generator produces limited variety.
- Training instability: The generator struggles to learn effectively.

Wasserstein GAN

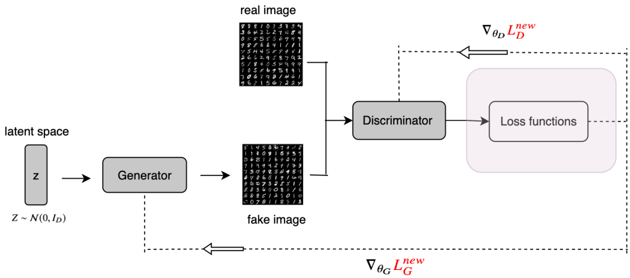


Figure: W-GAN

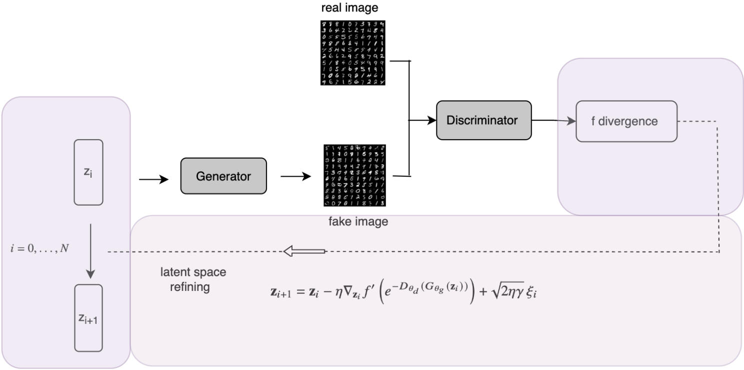
	Discriminator/Critic	**Generator**
GAN	$\nabla_{\theta_d} \frac{1}{m} \sum_{i=1}^{m} \left[\log D\left(x^{(i)}\right) + \log\left(1 - D\left(G\left(z^{(i)}\right)\right)\right)\right]$	$\nabla_{\theta_g} \frac{1}{m} \sum_{i=1}^{m} \log\left(D\left(G\left(z^{(i)}\right)\right)\right)$
WGAN	$\nabla_w \frac{1}{m} \sum_{i=1}^{m} \left[f\left(x^{(i)}\right) - f\left(G\left(z^{(i)}\right)\right)\right]$	$\nabla_\theta \frac{1}{m} \sum_{i=1}^{m} f\left(G\left(z^{(i)}\right)\right)$

WGAN with Gradient Penalty

- Wasserstein GAN (WGAN):
  - a critic(same architecture with G) output a scalar
  - Uses Wasserstein distance (1-Lipschitz constraint on the discrimintator) instead of JS divergence
  - Provides more stable gradients during training
- Gradient Penalty (GP) Addition:
  - introduces a gradient penalty term to ensure that the discriminator's gradients remain bounded
  - Better training instability
  - Better convergence
- Benefits for Our Project:
  - More stable training process
  - Better quality samples (FID: 40.78 vs 63.99)
  - Improved recall (0.294 vs 0.205)

https://arxiv.org/abs/1701.07875.

Latent Space Refinement with DGflow



The purpose of this refinement is to use information from the discriminator's feedback to iteratively adjust the latent vector z , improving the quality of generated samples.

Discriminator Driven Latent Sampling (DDLS)

Your GAN is Secretly an Energy-based Model and You Should Use Discriminator Driven Latent Sampling

- Sample from the energy-based model.
- Higher quality sample.
- FID: 51.66

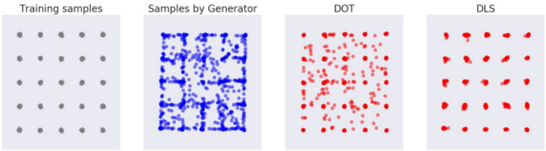


Figure: DDLS, generator alone, and generator + DOT, on 2d mixture of Gaussians distribution

https://arxiv.org/abs/2003.06060.

Discriminator Rejection Sampling (DRS)

- **Overview of DRS:**
  - Enhances GANs by using a pre-trained discriminator to filter generated samples.
  - Samples are retained based on a rejection threshold that aligns with discriminator confidence.
- **Advantages of DRS:**
  - Improves sample quality by rejecting low-confidence samples.
  - Provides a mechanism for balancing precision and recall through threshold tuning.
- **Performance at Current Threshold (0.7):**
  - FID: 38.02
  - Precision: 0.51, Recall: 0.25

https://arxiv.org/abs/1810.06758.

Model Metrics vs Threshold

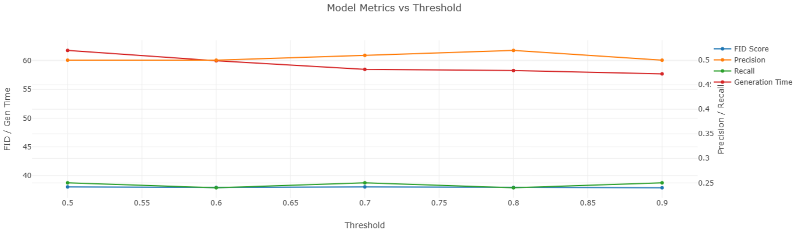


Figure: Impact of DRS Thresholds on Model Performance Metrics

DRS Threshold Effects:

- Lower thresholds (0.5-0.6): lower precision.
- Higher thresholds (0.8-0.9): higher precision.
- 0.7 provides balanced trade-off between metrics.

Results

Model	FID	P	R
Baseline VanillaGAN	63.99	0.39	0.21
VanillaGAN with DRS	38.49	0.34	0.28
W-GAN	40.78	0.34	0.29
W-GAN with DRS	38.02	0.51	0.25
W-GAN with Dgflow	40.02	0.47	0.25
VanillaGAN Hard $\psi = 2.0$	27.65	0.42	0.16
WGAN Hard $\psi = 0.9$	28.35	0.49	0.18

Table: Comparison of FID, Precision (P), Recall (R) for various models on MNIST.

- **Method Trade-offs:**
  - Hard truncation: Best FID, precision-focused
  - WGAN + DRS: Balanced metrics
  - DRS threshold allows flexible optimization
- **Complementary Strengths:**
  - Truncation improves precision
  - DRS helps with distribution coverage
  - Threshold tuning enables metric prioritization

4. Next Steps

NEXT Ups:

- Dive into f-divergence and PR-divergence to see the trade off beteween recall and precision
- Search and compare other method on latent space