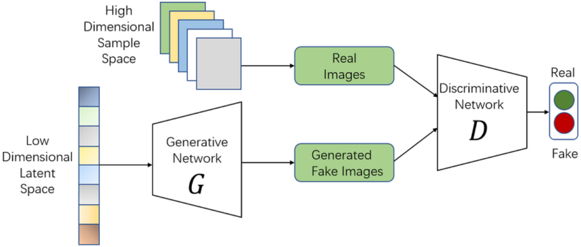
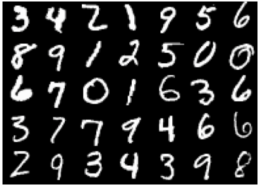
VeGANs project presentation

Mathilde Kretz, Amélie Leroy, Marius Roger

Problem statement

Goal : represent a latent (learning) space as a manner to implement image generation.

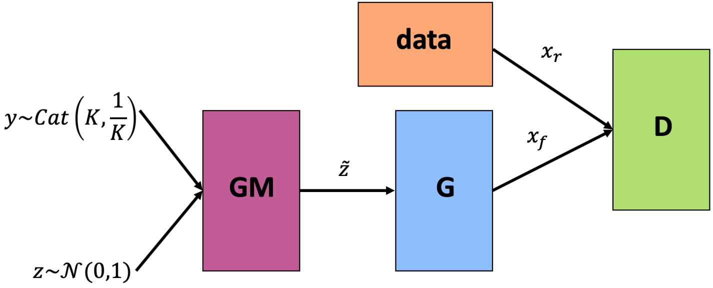




$$\min_G \max_D \mathbb{E}_{X_r \sim P}[\log D(x)] + \mathbb{E}_{X_g \sim \hat{P}_G}[\log(1 - D(x))]$$

Gaussian Mixture Model

K : number of gaussians inside the latent space



$\widetilde{z} = \mu(y) + \sigma(y) z$ follows $\mathcal{N}(\mu_k, \sigma_k)$, the k^{th} Gaussian distribution of the latent space.

Allows backpropagation through the GM module $\longrightarrow$ dynamic implementation

First implementation : Supervised GM-GAN

Supervised objective: separate real / fake samples **and** find class of real samples

Discriminator returns :

$$\begin{cases} (0, .., 1^{ith}, .., 0) & \text{if the input is of class } i \\ (0, ..., 0, 1) & \text{if the input is a fake sample} \end{cases}$$

Generator should map each Gaussian to a unique class.

CrossEntropy Loss :

- **G training**: between $D(G(\widetilde{z}))$ and label y used to sample $\widetilde{z}$
- **D training** : between $D(sample_{real})$ and label y_{real}
  \+ between $D(G(\widetilde{z}))$ and "label" $y_{fake} = 10$

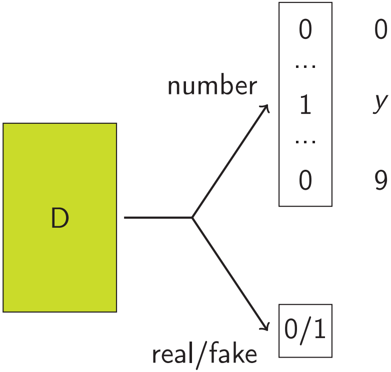
Second implementation : alternative objective

Change: We separate class prediction and reality prediction objectives.

Cross-Entropy between first 10 dims of output and the label y

Binary CE between last dim and fake/real

Training D, CE loss is used only on real samples
$G(GM(y,z))$ and y are not correlated enough



Approaches we tested

- Mean μ and variance σ are the representative parameters of our problem, we initialized them with different values : μ is uniformly sampled in $[-c, c]$ and σ is set to a constant σ
- Hyperparameter tuning : learning rate and batch size
- Data augmentation (future)
- Dropout
- Weighting the losses (approach 2)

Results for initial approach

Figure: Numbers generated with approach 1 (per-category mode collapse)

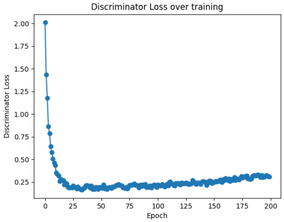


Figure: D loss

Figure: G loss

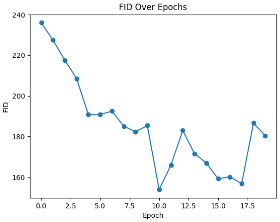


Figure: FID over epochs

Alternative results

σ	c	learning rate	FID
1.7	3	8e-5	100
0.7	3	2e-4	150
1	3	2e-4	134
1	1	2e-4	210

Table: Results for approach 1 @ 150 epochs

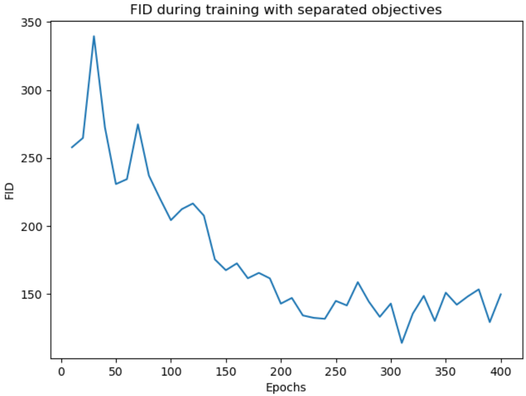


Figure: FID for approach 2

References

Matan Ben-Yosef and Daphna Weinshall : Gaussian Mixture Generative Adversarial Networks for Diverse Datasets, and the Unsupervised Clustering of Images, 2018

Teodora Pandeva and Matthias Schubert : MMGAN: Generative Adversarial Networks for Multi-Modal Distributions, 2019