Assignment 2

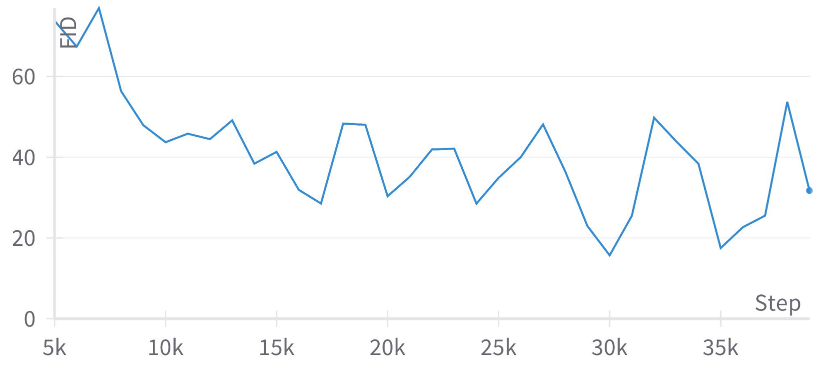
happyhippogan, Nikita Andreev

GANs' weaknesses

GANs $\min_G \max_D V(D,G) = \mathbb{E}_{\boldsymbol{x}\sim p_{\text{data}}(\boldsymbol{x})}[\log D(\boldsymbol{x})] + \mathbb{E}_{\boldsymbol{z}\sim p_{\boldsymbol{z}}(\boldsymbol{z})}[\log(1 - D(G(\boldsymbol{z})))]$

Main problems:

- Unstability
- image artifacts (poor quality with non-optimal parameters)
- Convergence speed
- Higher resolution (requires larger models and much more training time)



Unbalanced Optimal Transport [1]

On the border between GANs and OT. So, authors use some tricks like gradient penalty and advanced models for image generation.

$$C_{ub}(\mu, \nu) := \inf_{\pi \in \mathcal{M}_+(\mathcal{X} \times \mathcal{Y})} \left[\int_{\mathcal{X} \times \mathcal{Y}} c(x, y) d\pi(x, y) + D_{\Psi_1}(\pi_0 | \mu) + D_{\Psi_2}(\pi_1 | \nu) \right]$$

Dual formulation:

$$C_{ub}(\mu, \nu) = \sup_{u(x)+v(y) \leq c(x,y)} \left[\int_{\mathcal{X}} -\Psi_1^*(-u(x)) d\mu(x) + \int_{\mathcal{Y}} -\Psi_2^*(-v(y)) d\nu(y) \right]$$

Semi-dual formulation (psi functions usually are softplus or shifted exponent) leads to such optimization:

$$\mathcal{L}_{v_\phi, T_\theta} = \inf_{v_\phi} \left[\int_{\mathcal{X}} \Psi_1^* \left(-\inf_{T_\theta} \left[c\left(x, T_\theta(x)\right) - v_\phi\left(T_\theta(x)\right) \right] \right) d\mu(x) + \int_{\mathcal{Y}} \Psi_2^*\left(-v_\phi(y)\right) d\nu(y) \right]$$

But the optimization procedure is similar to OT and GANs.

I **implemented** UOT for image generation task, adopted losses, conducted experimental comparison to GANs and the comparison of Psi functions.

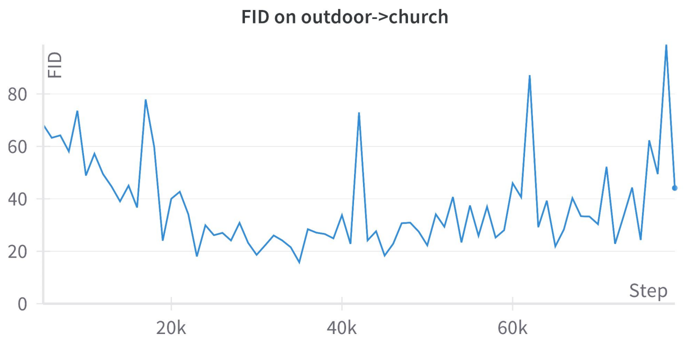
[1] - Choi et al. Generative Modeling through the Semi-dual Formulation of Unbalanced Optimal Transport, 2023. https://arxiv.org/abs/2305.14777

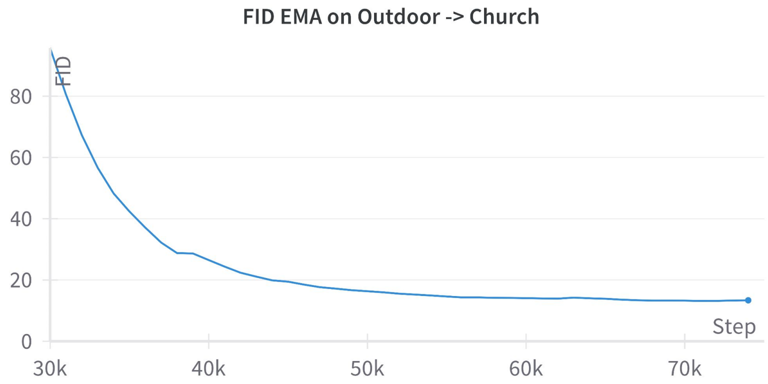
Exponential Moving Average (EMA)

$$\theta_{EMA}^{(t)} = \beta\theta_{EMA}^{(t-1)} + (1-\beta)\theta^{(t)}$$

I used EMA which is a common method of weight averaging for GANs [2].

Usually, EMA gives extreme stability and some quality boost, especially for more complex models like UNet or diffusions, but not in my case due to the weak architectures of the generator and discriminator models.





[2] - Y. Yazici et al. The Unusual Effectiveness of Averaging in GAN Training, 2018. https://arxiv.org/abs/1806.04498

Main results

- UOT works but not as good as GANs
- Psi=softplus is usually better than Psi=exp
- EMA did not give any improvement which is sad as it's a common practice for most models

Method	FID score
GAN baseline	29.78
UOT baseline (60 epochs)	35.6
UOT baseline (100 epochs)	46.8
UOT(100 epochs) + Ψ^*=*exp*	37.1
UOT + Ψ^*=*exp* + *EMA (50 epochs)*	143
UOT + Ψ^*=*exp* + *EMA (90 epochs)*	136

Conclusion

- I have tried to adapt Unbalanced Optimal Transport to MNIST generation.
- Additionally, I used several common techniques such as EMA
- Basically, UOT works but not as good as GAN for this task
- Techniques do not work, mainly because of the limitations of architectures of Generator (and discriminator), so they are not effective (but they are for the traditional architectures).