A U-Net Based Discriminator for Generative Adversarial Networks

Schonfeld, Edgar and Schiele, Bernt and Khoreva, Anna
Presented by: Andrey Yashkin

Big Data AI
Novosibirsk State University

April 13, 2021

Generative Adversarial Networks

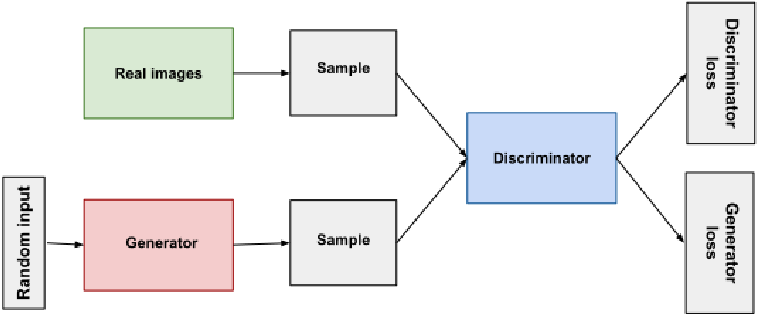

Figure: GAN structure

Optimisation problem

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

Ways to solve the minimax problems:

- Find $F(G) = \max_D V(D,G)$ and apply gradient descent
- Do several gradient descent steps for G and several gradient ascent steps for D
- Gradient descent ascent

Global and local realism

Discriminator is a classification network and learns only a representation that allows to efficiently penalize the generator based on the most discriminative difference between real and synthetic images. Thus, it often focuses either on the global structure or local details.

(a) Global realism

(b) Local realism

Figure: Global and local realism examples

U-NET based discriminator

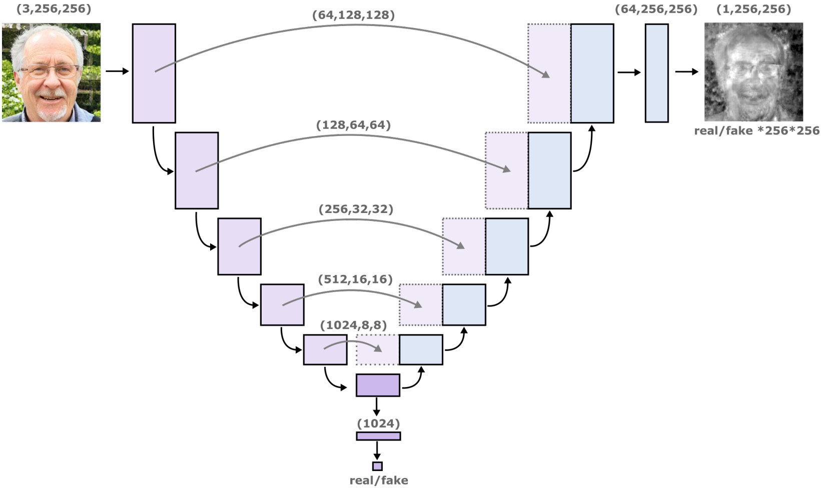

Figure: GAN structure

U-NET gan loss

Usual losses

$$L_D = -\mathbb{E}_x[logD(x)] - \mathbb{E}_z[1 - log(D(G(z)))]$$

$$L_G = \mathbb{E}_z[logD(G(z)))]$$

New losses

$$L_{D^U_{enc}} = -\mathbb{E}_x[logD^U_{enc}(x)] - \mathbb{E}_z[1 - log(D^U_{enc}(G(z)))]$$

$$L_{D^U_{dec}} = -\mathbb{E}_x \sum_{i,j}[logD^U_{dec}(x)_{i,j}] - \mathbb{E}_z \sum_{i,j}[1 - log(D^U_{dec}(G(z)))_{i,j}]$$

Here $D^U_{enc}(x)_{i,j}$ and $D^U_{dec}(G(z)))_{i,j}$ refer to the discriminator decision at pixel (i, j)

$$L_{D^U} = L_{D^U_{enc}} + L_{D^U_{dec}}$$

$$L_{G^U} = \mathbb{E}_z[logD(G(z))) + \sum_{i,j} logD(G(z)))_{i,j}]$$

Consistency Regularization

Using cutmix ($mix(x, y, M)$) with mask M

$$L^{cons}_{D^U_{dec}} = ||L_{D^U_{dec}}(mix(x, G(z), M)) - mix(L_{D^U_{dec}}(x), L_{D^U_{dec}}(G(z)), M)||^2$$

$$L_{D^U} = L_{D^U_{enc}} + L_{D^U_{dec}} + \lambda L^{cons}_{D^U_{dec}}$$

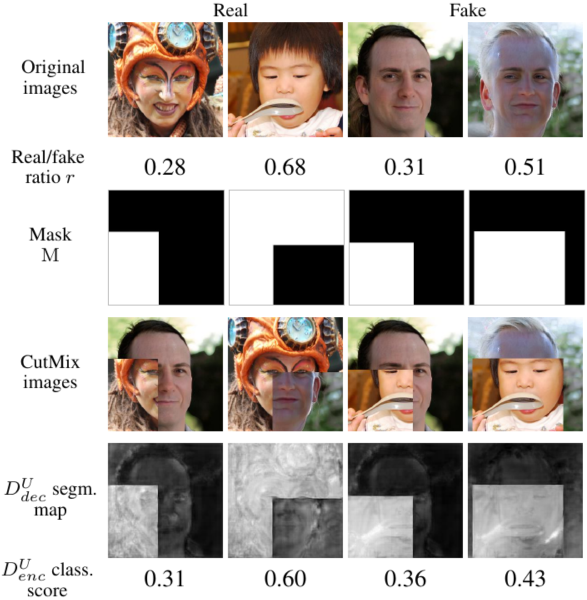

Figure: Cutmix

Results

Method	COCO-Animals	FFHQ
BigGAN [5]	16.55	12.42
U-Net based discriminator	15.86	10.86
+ CutMix augmentation	14.95	10.30
+ Consistency regularization	**13.87**	**7.63**

Figure: FID scores

Results

Method	FID ↓	IS ↑
PG-GAN [19]	7.30	–
COCO-GAN [27]	5.74	–
BigGAN [5]	4.54	3.23
U-Net GAN	**2.95**	**3.43**

Figure: Comparison with the state-of-the-art models on CelebA (128 $\times$ 128)

Results

Figure: Samples generated by U-Net GAN and the corresponding real-fake predictions of the U-Net decoder

Schonfeld, Edgar and Schiele, Bernt and Khoreva, Anna, Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, p. 8207–8216, 2020

The end