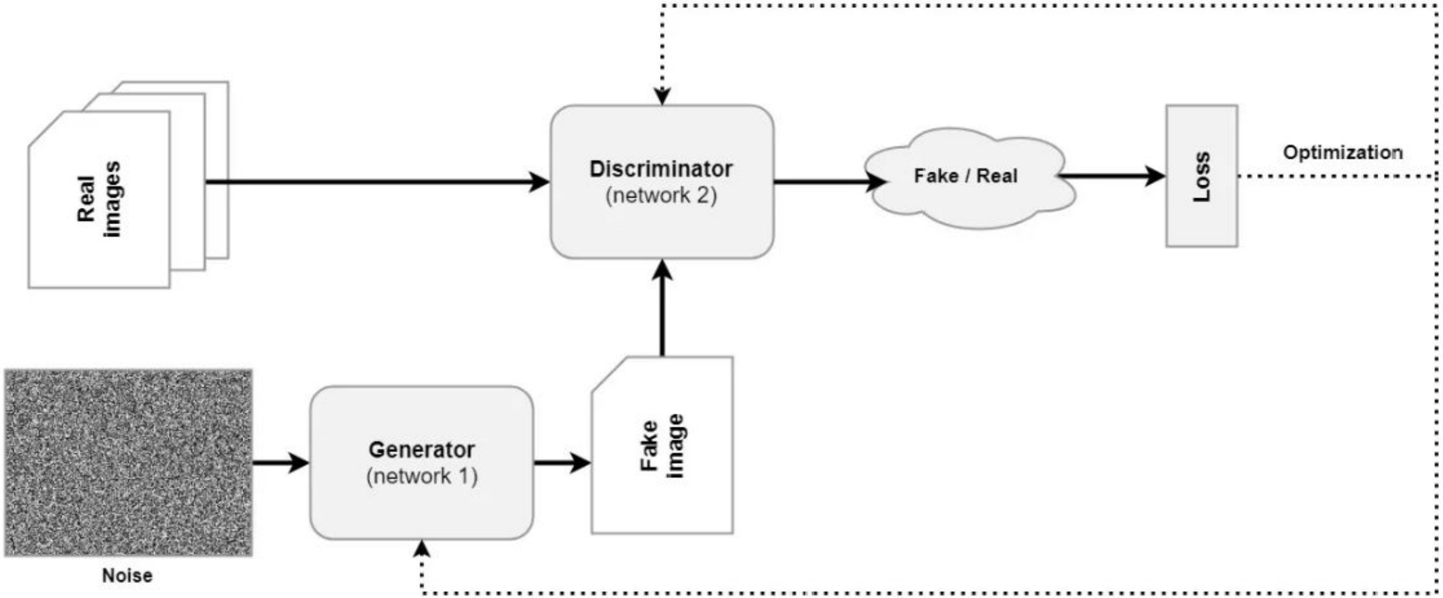
StyleGAN

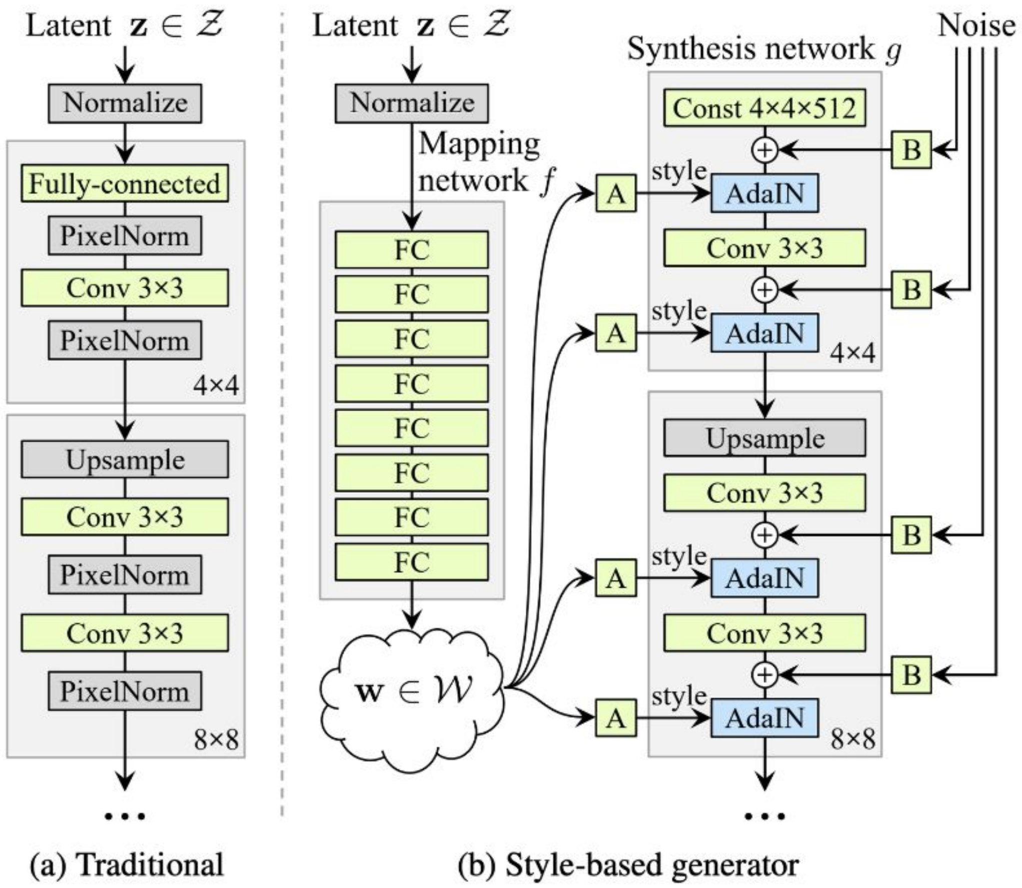
GAN



Problems with classic Generative Adversarial Networks

- Generators operate as black boxes
- GANs must be heavily regularized
- There is little control over image synthesis.

StyleGAN



Чем отличается StyleGan от традиционного Gan на примере ProGan и что это вообще за стили?

1. Mapper network

Рандомно сгенерированный вектор z[512] (standard normal distribution), пропускается через сеть и получается W[512]

"Распутываем" латентное пространство

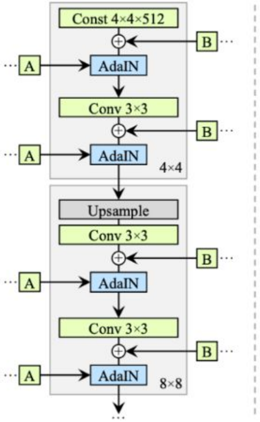
цвет	волосы	глаза	• • •	• • •

2. Synthesis network

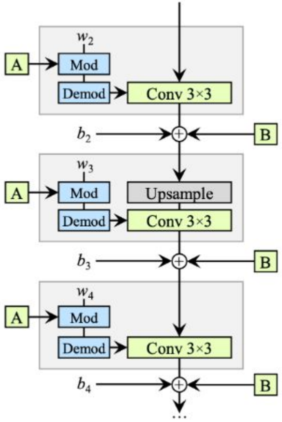
Основные компонентами каждого блока StyleGAN:
- Constant starting point
- Upsampling (except for the first synthesis block)
- Convolution layer
- Adaptive Instance Normalization (AdaIN)
- Style vectors (A) and noise vectors (B).

$$\mathrm{AdaIN}(\mathbf{x}_i, \mathbf{y}) = \mathbf{y}_{s,i} \frac{\mathbf{x}_i - \mu(\mathbf{x}_i)}{\sigma(\mathbf{x}_i)} + \mathbf{y}_{b,i},$$

StyleGAN2



(a) StyleGAN



(d) Weight demodulation

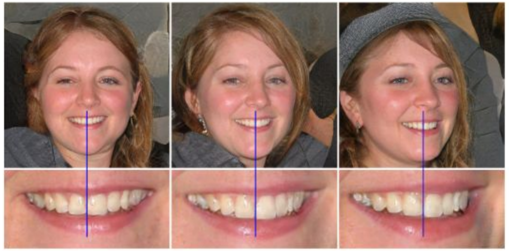
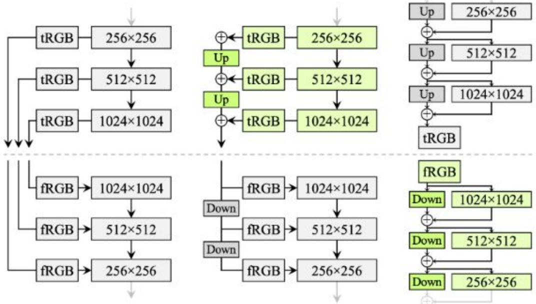


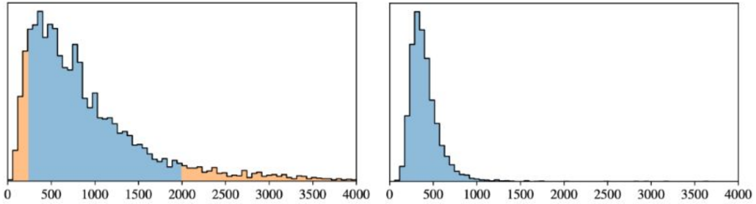
Figure 6. Progressive growing leads to "phase" artifacts. In this example the teeth do not follow the pose but stay aligned to the camera, as indicated by the blue line.



(a) MSG-GAN (b) Input/output skips (c) Residual nets

Perceptual path length (PPL)

(a) Low PPL scores (b) High PPL scores



(a) StyleGAN (config A) (b) StyleGAN2 (config F)

$$\mathbf{J_w} = \partial g(\mathbf{w})/\partial \mathbf{w}.$$

$$\mathbb{E}_{\mathbf{w},\mathbf{y}\sim\mathcal{N}(0,\mathbf{I})}\left(\left\|\mathbf{J}_{\mathbf{w}}^T\mathbf{y}\right\|_2 - a\right)^2,$$

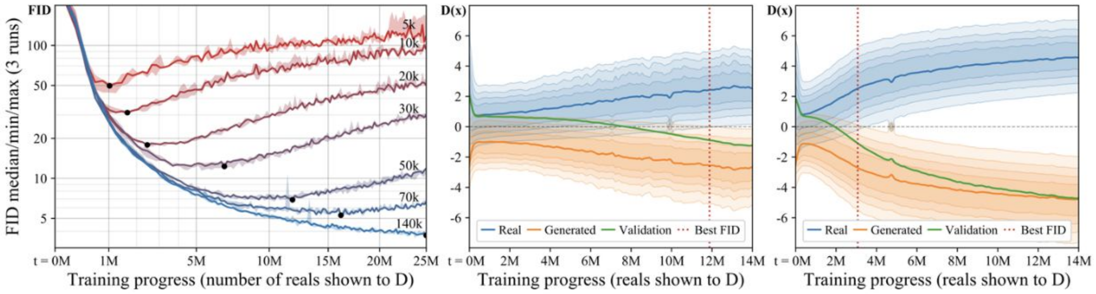
$$w'_{ijk} = s_i \cdot w_{ijk}, \quad (1)$$

$$w''_{ijk} = w'_{ijk} \Bigg/ \sqrt{\sum_{i,k} {w'_{ijk}}^2 + \epsilon}, \quad (3)$$

StyleGAN2-ADA

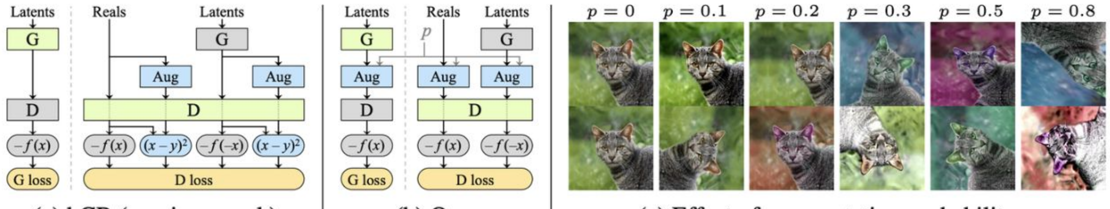
1. Overfitting in GANs (Как количество данных влияет на FID)
Figure 1a shows our baseline results for different subsets of FFHQ. Training starts the same way in each case, but eventually the progress stops and FID starts to rise.

Figure 1b,c shows the discriminator output distributions for real and generated images during training. The distributions overlap initially but keep drifting apart as the discriminator becomes more and more confident, and the point where FID starts to deteriorate is consistent with the loss of sufficient overlap between distributions.



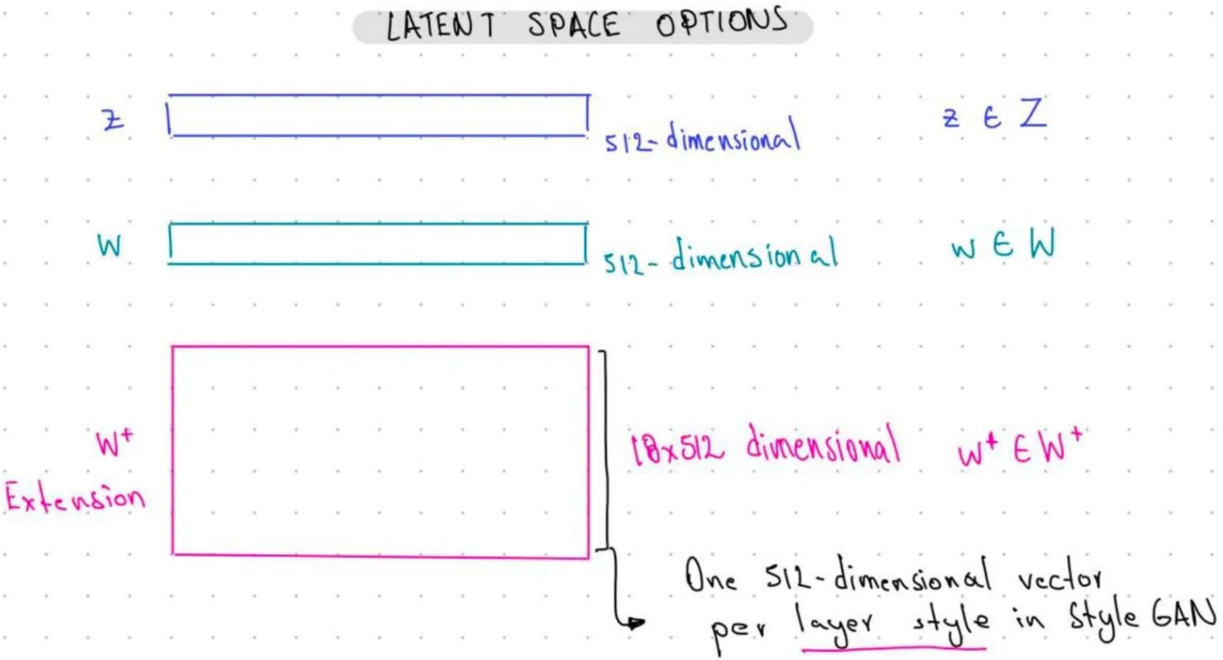
(a) Convergence of FFHQ (256 × 256) (b) Discriminator outputs, 50k (c) Discriminator outputs, 20k

2.1 Stochastic discriminator augmentation
By definition, any augmentation that is applied to the training dataset will get inherited to the generated images.
Our solution is similar to bCR in that we also apply a set of augmentations to all images shown to the discriminator

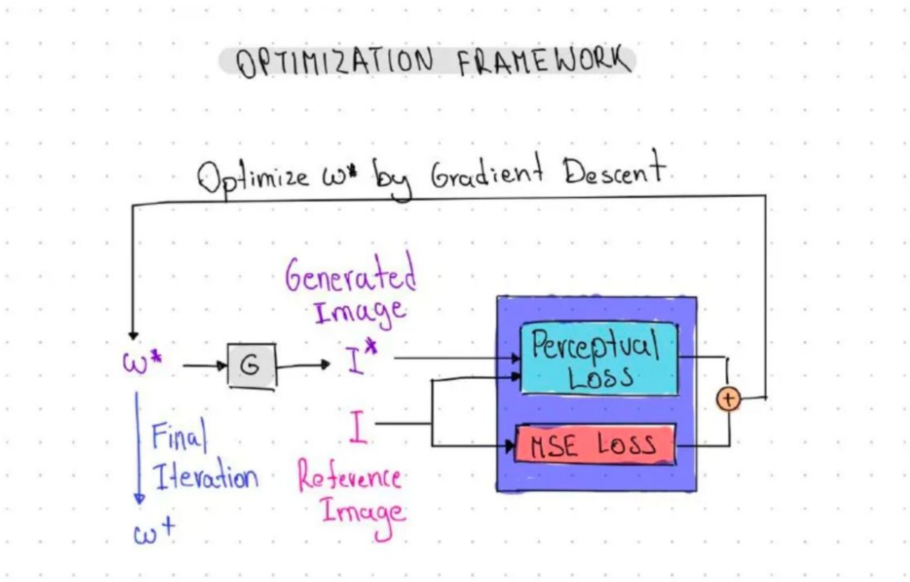


(a) bCR (previous work) (b) Ours (c) Effect of augmentation probability p

Latent space



Optimization



Parameter to control
the MSE influence

$$w^* = \min_{w} \underbrace{L_{\text{percept}}\left(G(w), I\right)}_{\text{perceptual loss}} + \overbrace{\frac{\lambda_{mse}}{N}\|G(w) - I\|_2^2}^{\text{MSE loss}}$$

Expression Transfer

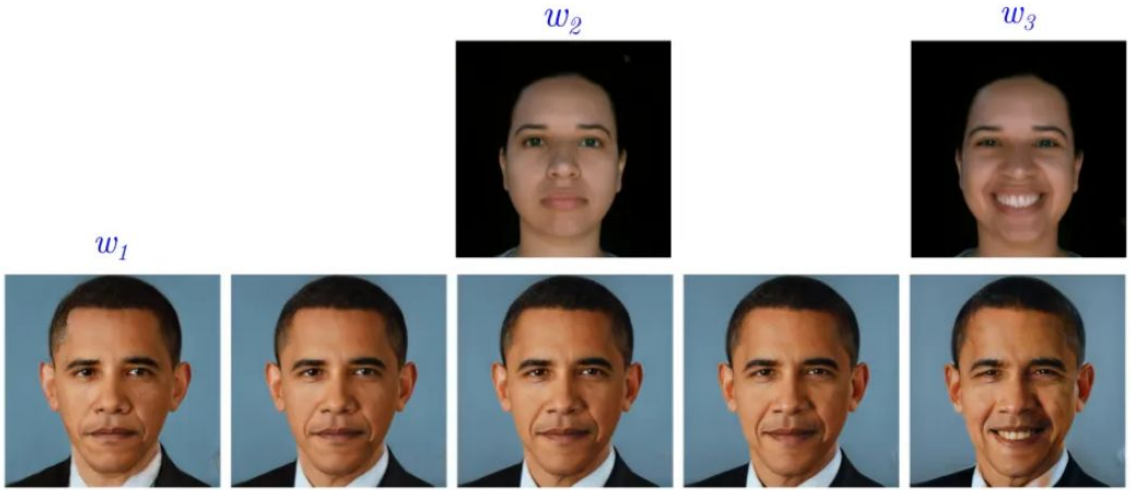
transfer expression parameter

latent code from neutral expression

$$w = w_1 + \lambda(w_3 - w_2)$$

latent code from target image

latent code from distinctive expression



Style Transfer