

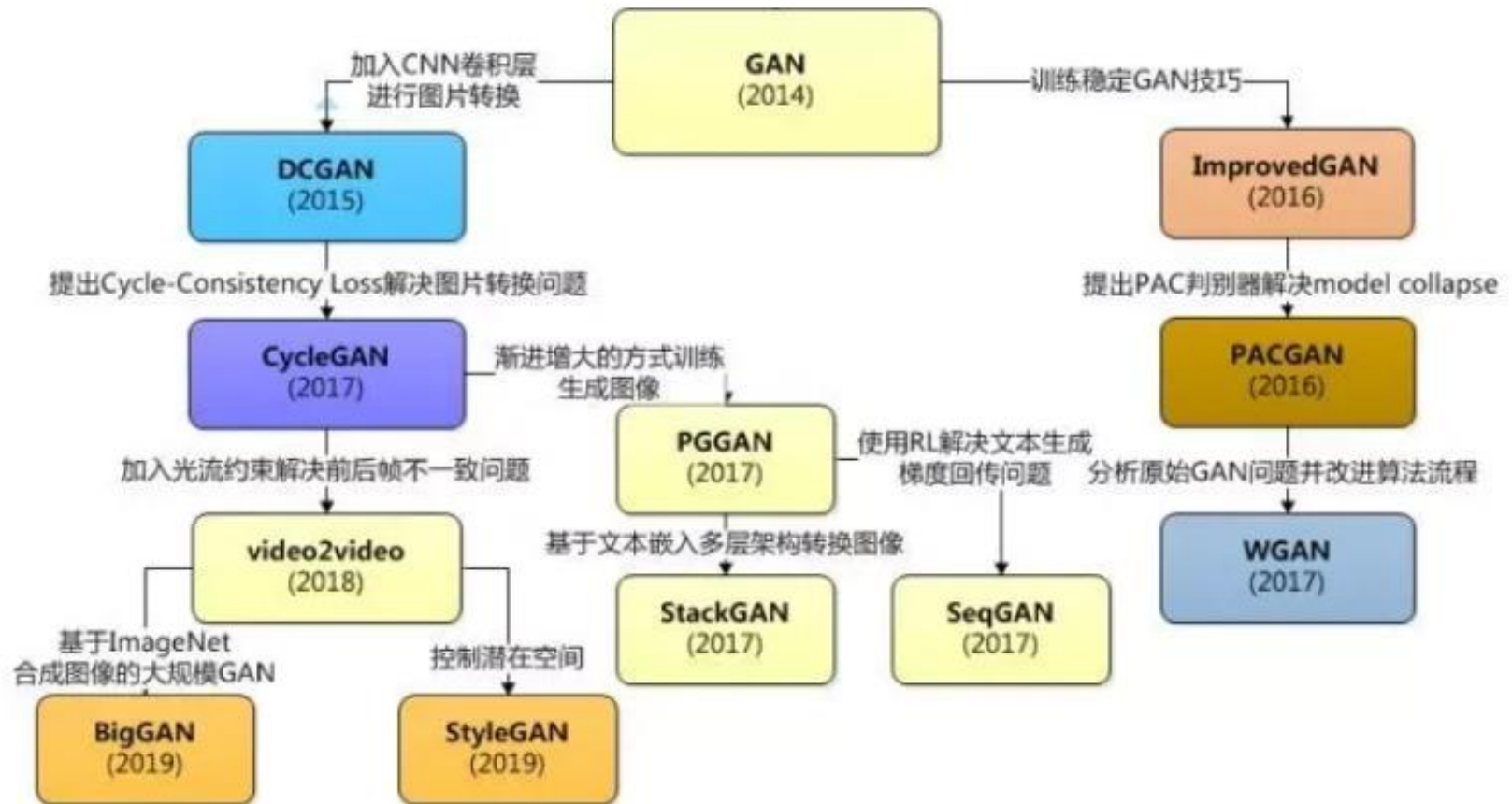


GAN 2

赵洲

浙江大学计算机学院

GAN模型家族



Generative Adversarial Network

1. StackGAN
2. Pix2PixGAN
3. CycleGAN
4. StyleGAN

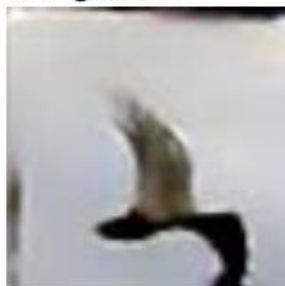
StackGAN模型动机

This bird is white with some black on its head and wings, and has a long orange beak

This bird has a yellow belly and tarsus, grey back, wings, and brown throat, nape with a black face

This flower has overlapping pink pointed petals surrounding a ring of short yellow filaments

StackGAN
Stage-I
64x64
images



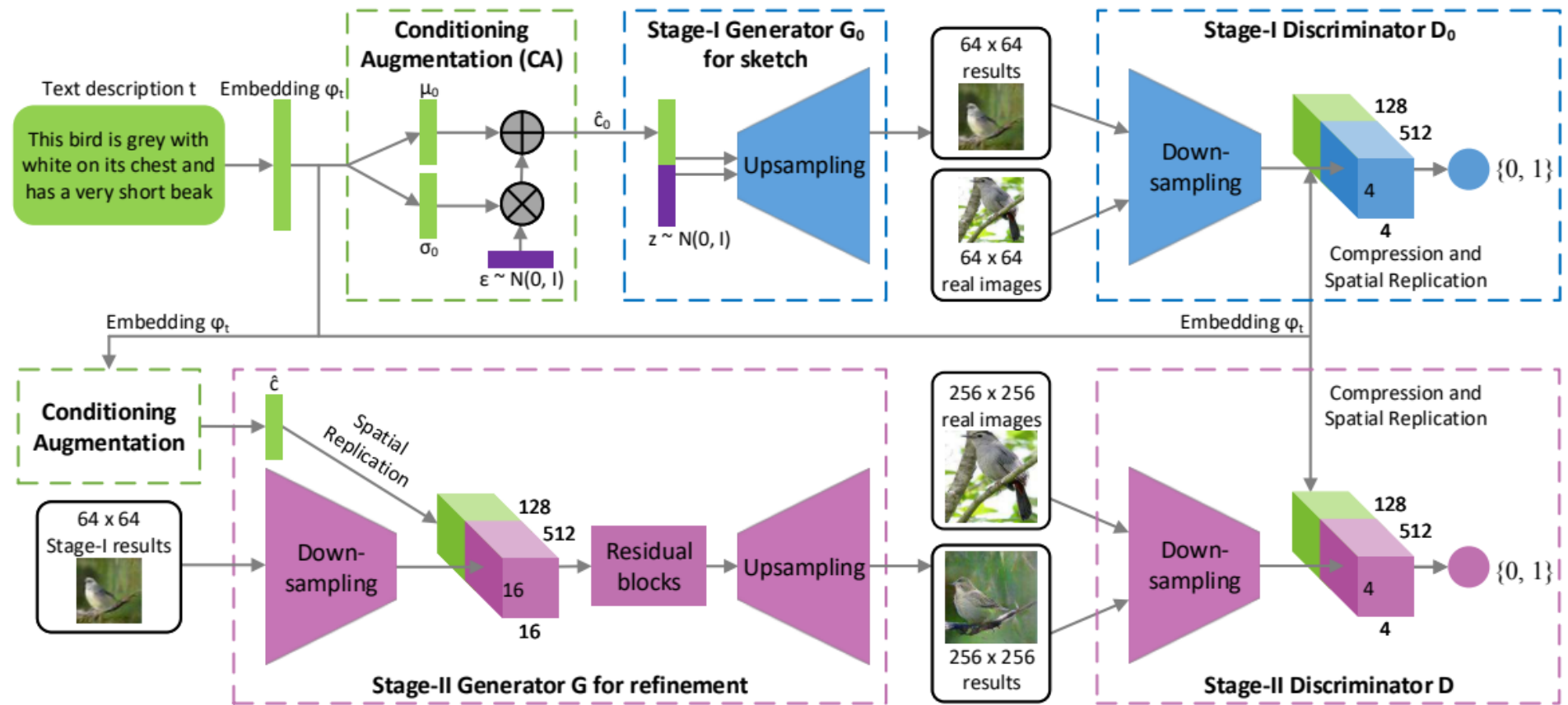
StackGAN
Stage-II
256x256
images



Vanilla GAN
256x256
images



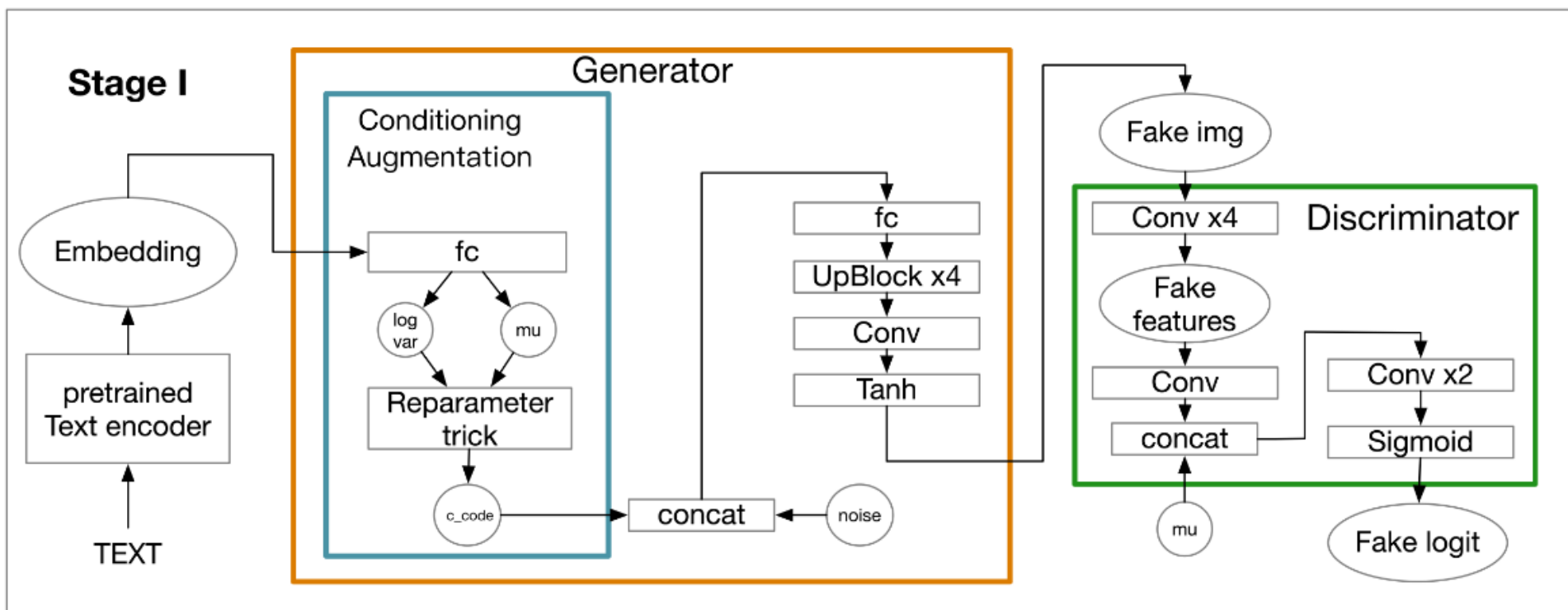
StackGAN模型结构



Stage 1结构

$$L_{D_0} = E_{(I_0, t) \sim p_{data}} [\log D_0(I_0, \psi_t)] + E_{z \sim p_z, t \sim p_{data}} [\log(1 - d_0(G_0(z, \hat{c}_0), \psi_t))]$$

$$L_{G_0} = E_{z \sim p_z, t \sim p_{data}} [\log(1 - D_0(G_0(z, \hat{c}_0), \psi_t))] + \lambda D_{KL}(N(\mu_0(\psi_t), \Sigma_0(\psi_t)) || N(0, I))$$

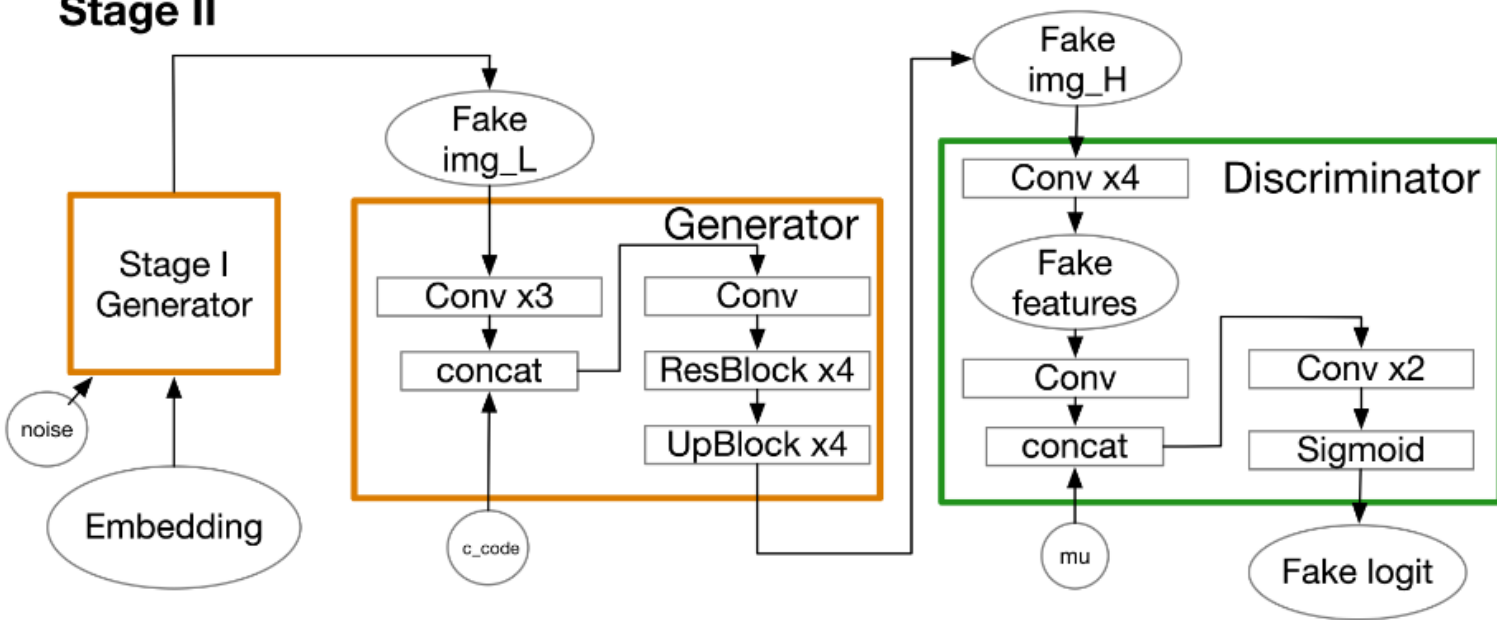


Stage 2结构

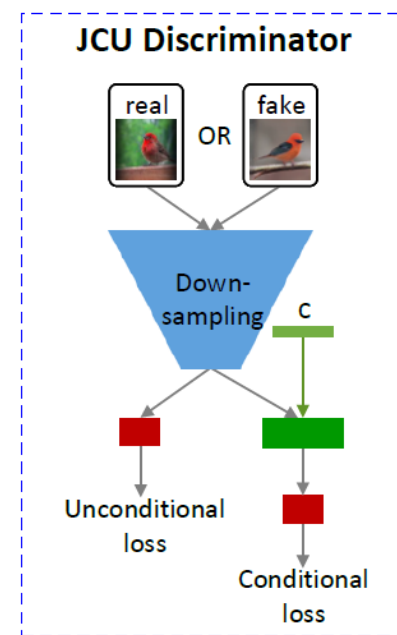
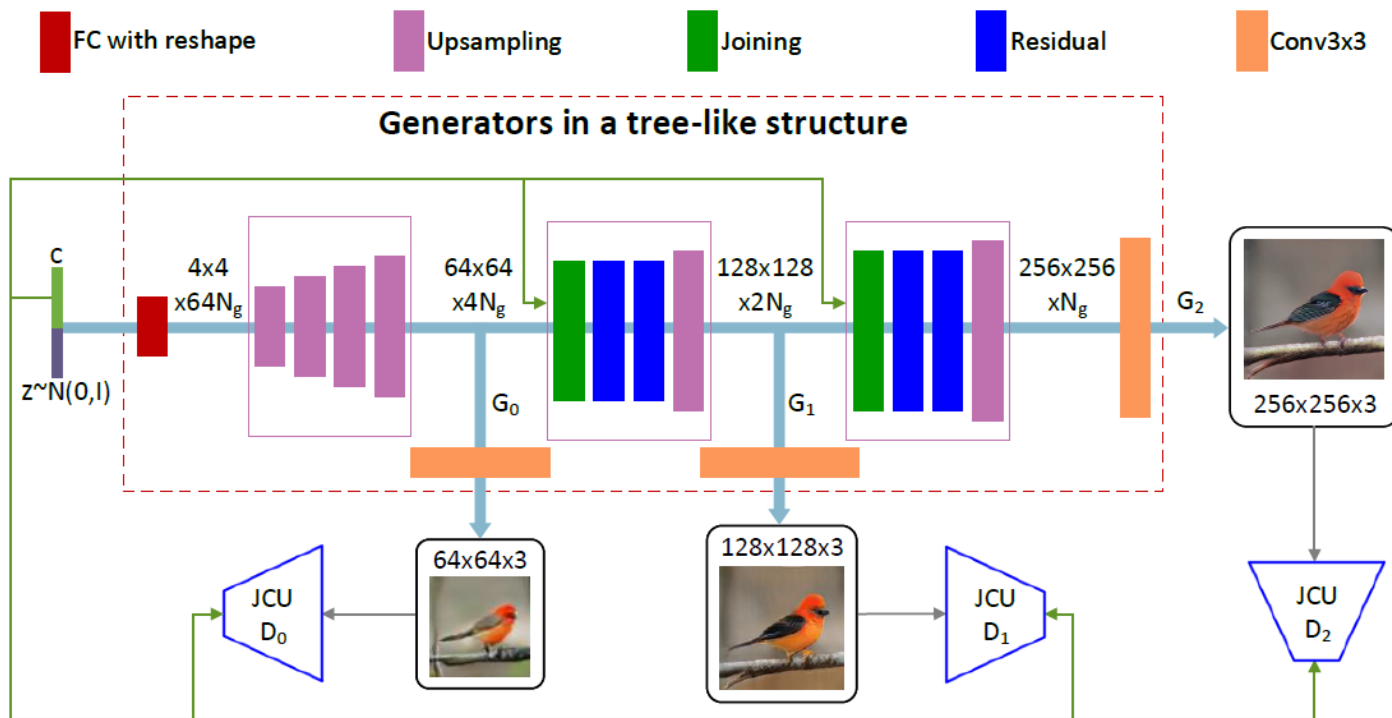
$$L_D = E_{(I,t) \sim p_{data}} [\log D(I, \psi_t)] + E_{s_0 \sim p_{G_0}, t \sim p_{data}} [\log(1 - D(G(s_0, \hat{c}_0), \psi_t))]$$

$$L_G = E_{s_0 \sim p_{G_0}, t \sim p_{data}} [\log(1 - D(G(s_0, \hat{c}_0), \psi_t))] + \lambda D_{KL}(N(\mu(\psi_t), \Sigma(\psi_t)) || N(0, I))$$

Stage II



StackGAN2 模型结构



StackGAN2 数学原理

$$\mathcal{L}_{D_i} = \underbrace{-\frac{1}{2}\mathbb{E}_{x_i \sim p_{data_i}} [\log D_i(x_i)] - \frac{1}{2}\mathbb{E}_{s_i \sim p_{G_i}} [\log(1 - D_i(s_i))]}_{\text{unconditional loss}} + \underbrace{-\frac{1}{2}\mathbb{E}_{x_i \sim p_{data_i}} [\log D_i(x_i, c)] - \frac{1}{2}\mathbb{E}_{s_i \sim p_{G_i}} [\log(1 - D_i(s_i, c))]}_{\text{conditional loss}}.$$

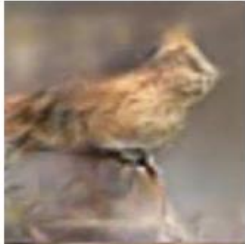





$$\mathcal{L}_{G_i} = \underbrace{\frac{1}{2}\mathbb{E}_{s_i \sim p_{G_i}} [\log(1 - D_i(s_i))]}_{\text{unconditional loss}} + \underbrace{\frac{1}{2}\mathbb{E}_{s_i \sim p_{G_i}} [\log(1 - D_i(s_i, c))]}_{\text{conditional loss}}.$$

$$\mathcal{L}_{C_i} = \frac{1}{n} \sum_{j=1}^n \left(\lambda_1 \|\mu_{s_i^j} - \mu_{s_{i-1}^j}\|_2^2 + \lambda_2 \|\Sigma_{s_i^j} - \Sigma_{s_{i-1}^j}\|_F^2 \right)$$

实验结果

Text description	This bird is red and brown in color, with a stubby beak	The bird is short and stubby with yellow on its body	A bird with a medium orange bill white body gray wings and webbed feet	This small black bird has a short, slightly curved bill and long legs	A small bird with varying shades of brown with white under the eyes	A small yellow bird with a black crown and a short black pointed beak	This small bird has a white breast, light grey head, and black wings and tail
64x64 GAN-INT-CLS							
128x128 GAWWN							
256x256 StackGAN							

实验结果

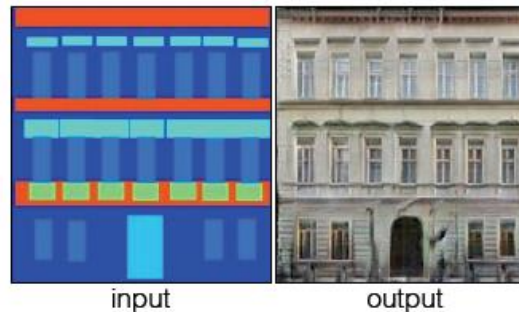
Text description	This bird is blue with white and has a very short beak	This bird has wings that are brown and has a yellow belly	A white bird with a black crown and yellow beak	This bird is white, black, and brown in color, with a brown beak	The bird has small beak, with reddish brown crown and gray belly	This is a small, black bird with a white breast and white on the wingbars.	This bird is white black and yellow in color, with a short black beak
Stage-I images							
Stage-II images							

Pix2PixGAN

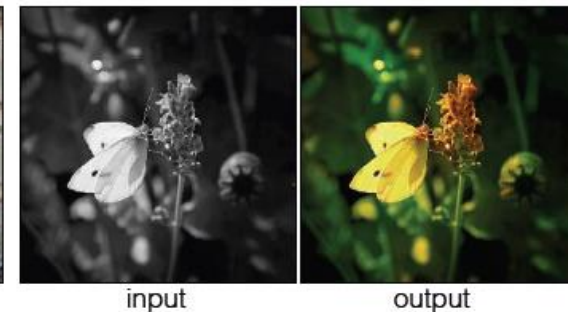
Labels to Street Scene



Labels to Facade



BW to Color



Aerial to Map



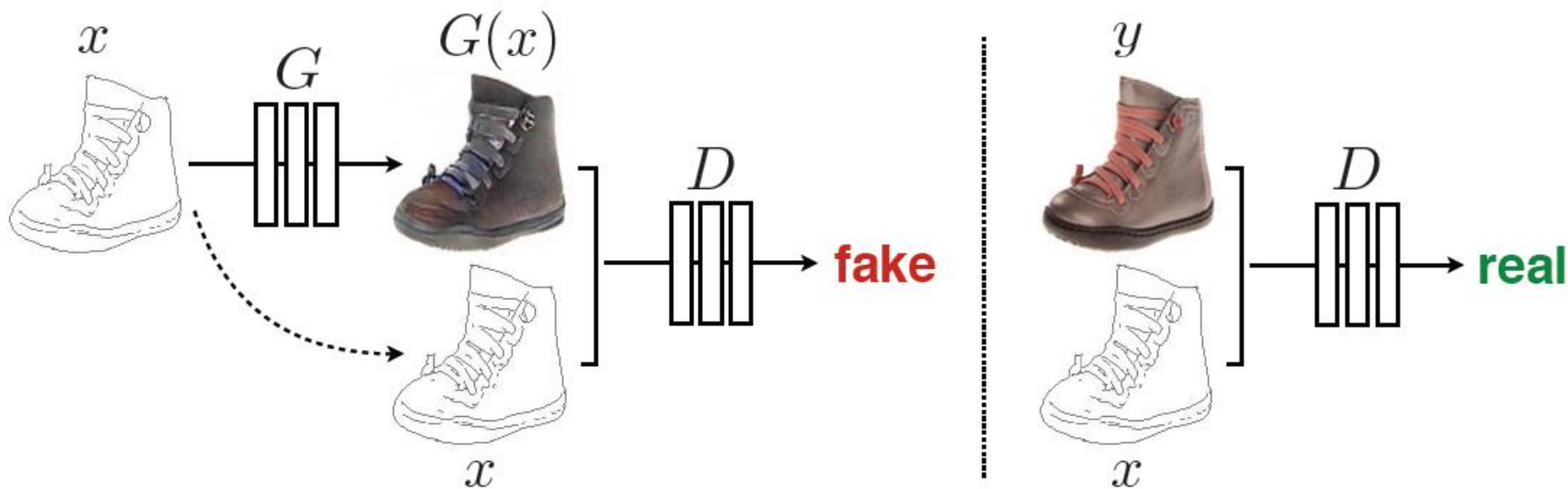
Day to Night



Edges to Photo

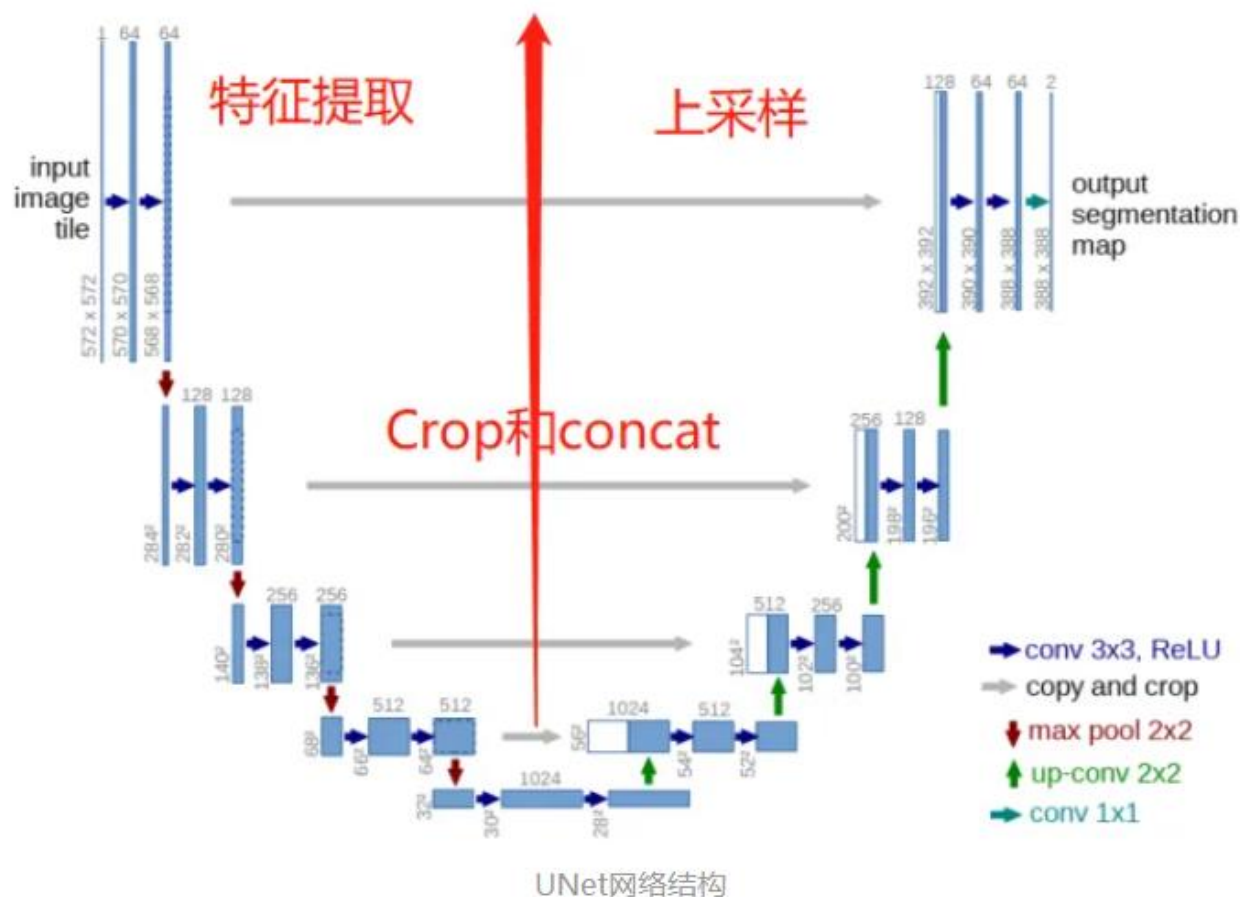


Pix2PixGAN模型结构



U-Net

- U-Net是编码解码结构，编码负责特征提取，解码负责恢复原始分辨率（上采样和拼接操作）。



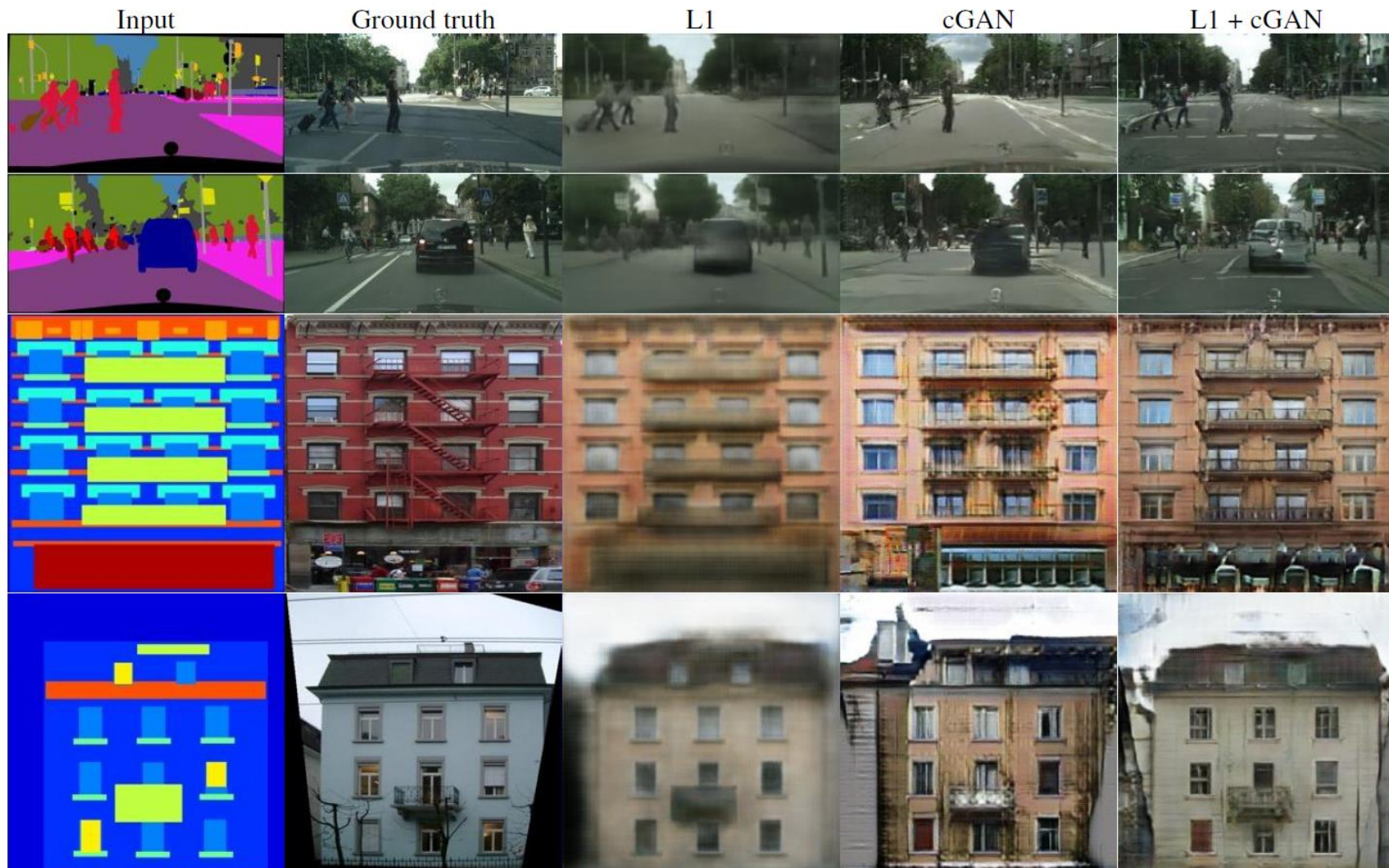
数学原理

$$\mathcal{L}_{cGAN}(G, D) = \mathbb{E}_{x,y}[\log D(x, y)] + \mathbb{E}_{x,z}[\log(1 - D(x, G(x, z)))]$$

$$\mathcal{L}_{L1}(G) = \mathbb{E}_{x,y,z}[\|y - G(x, z)\|_1].$$

$$G^* = \arg \min_G \max_D \mathcal{L}_{cGAN}(G, D) + \lambda \mathcal{L}_{L1}(G)$$

实验结果



CycleGAN

Monet \leftrightarrow Photos



Monet \rightarrow photo

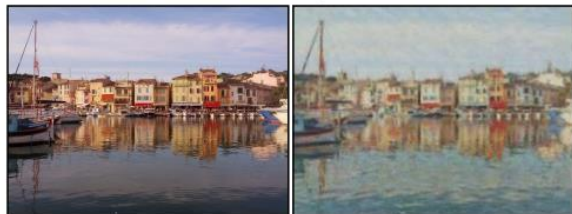
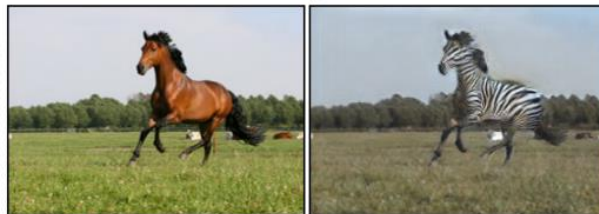


photo \rightarrow Monet

Zebras \leftrightarrow Horses



zebra \rightarrow horse



horse \rightarrow zebra

Summer \leftrightarrow Winter



summer \rightarrow winter



winter \rightarrow summer



Photograph



Monet



Van Gogh

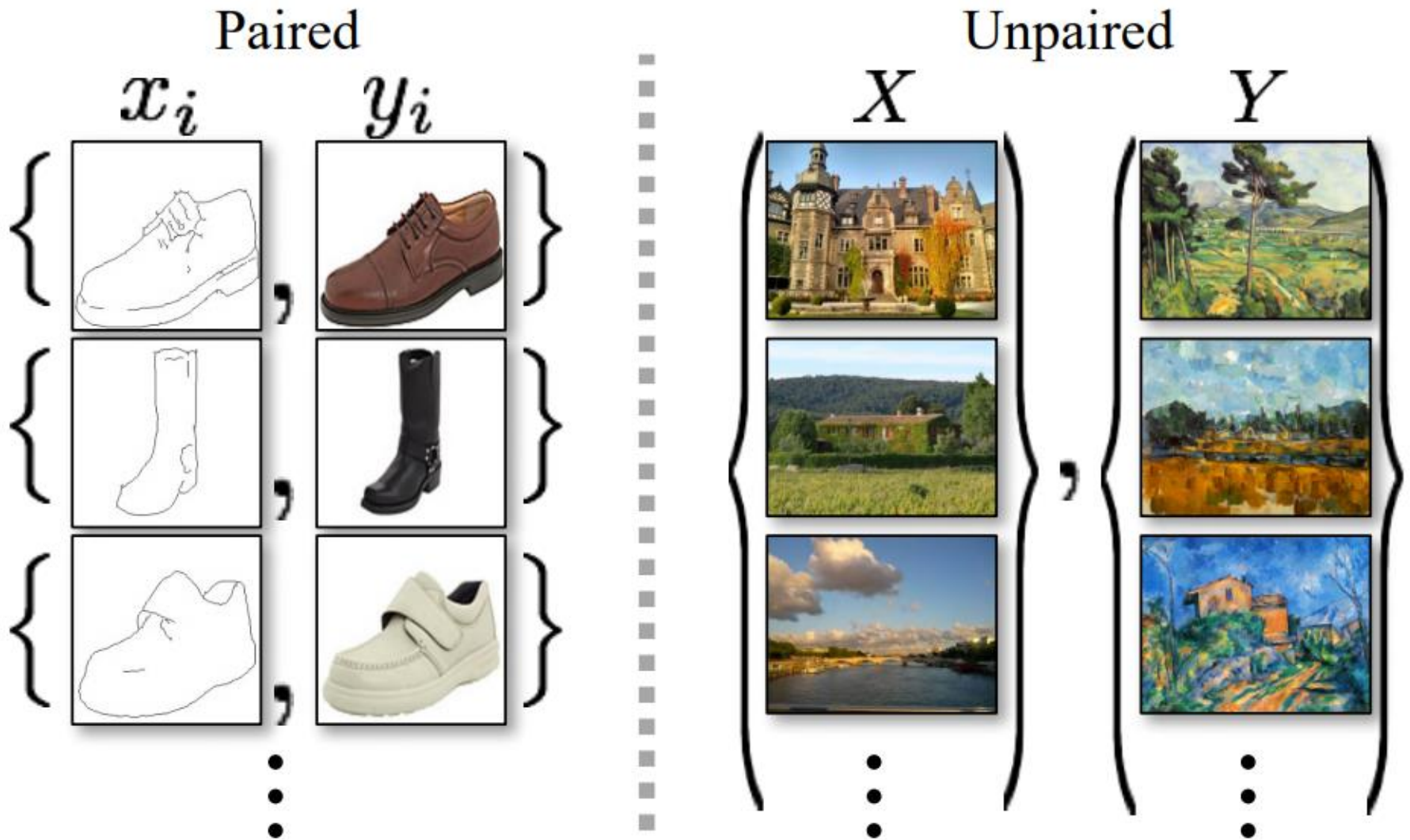


Cezanne

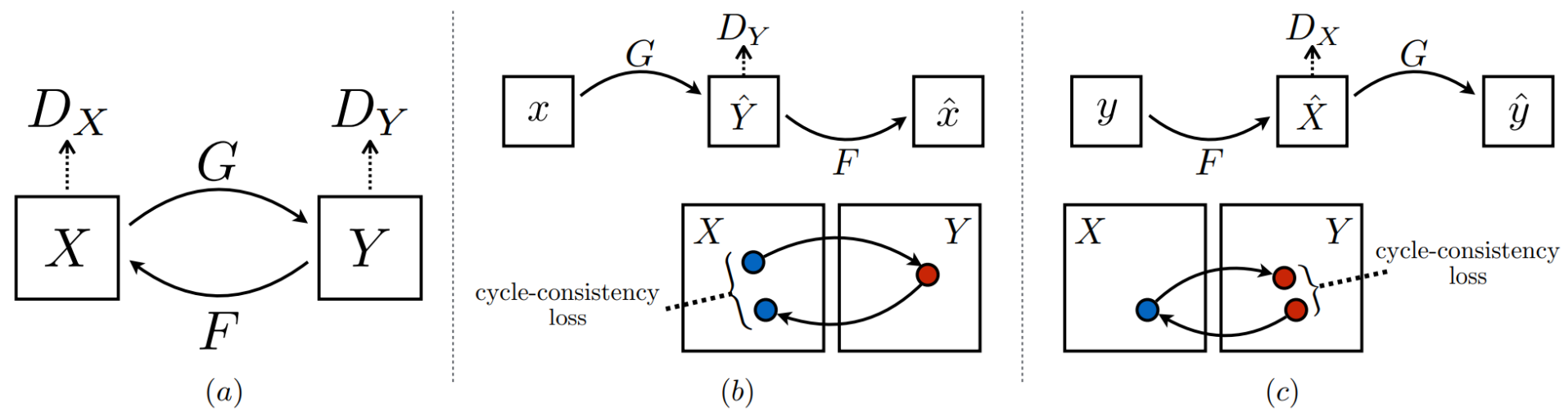


Ukiyo-e

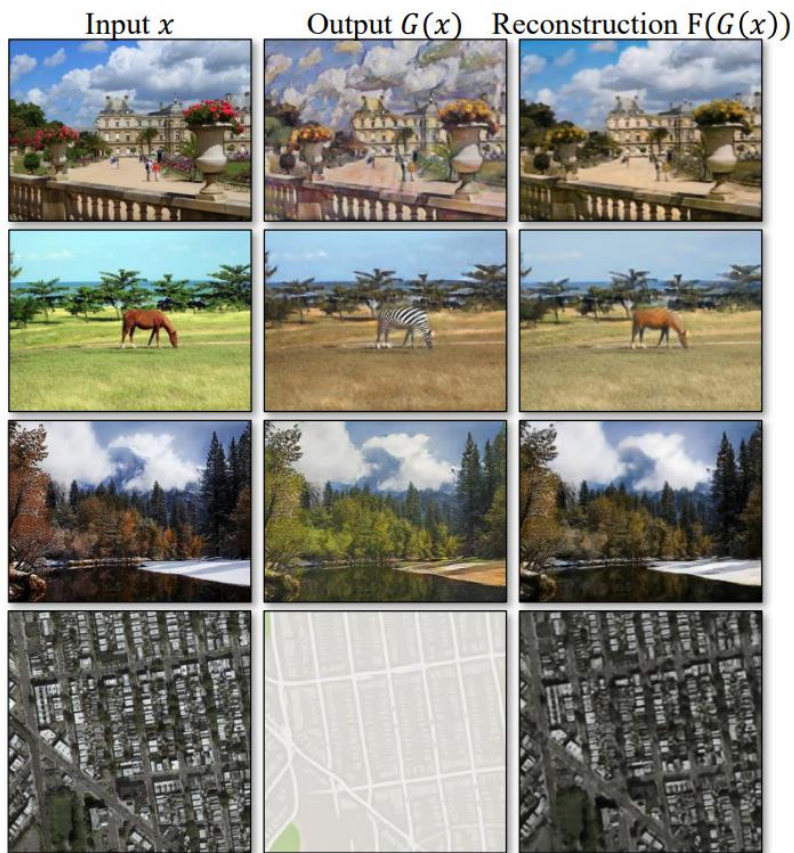
成对 v.s. 非成对



Cycle原理



数学原理

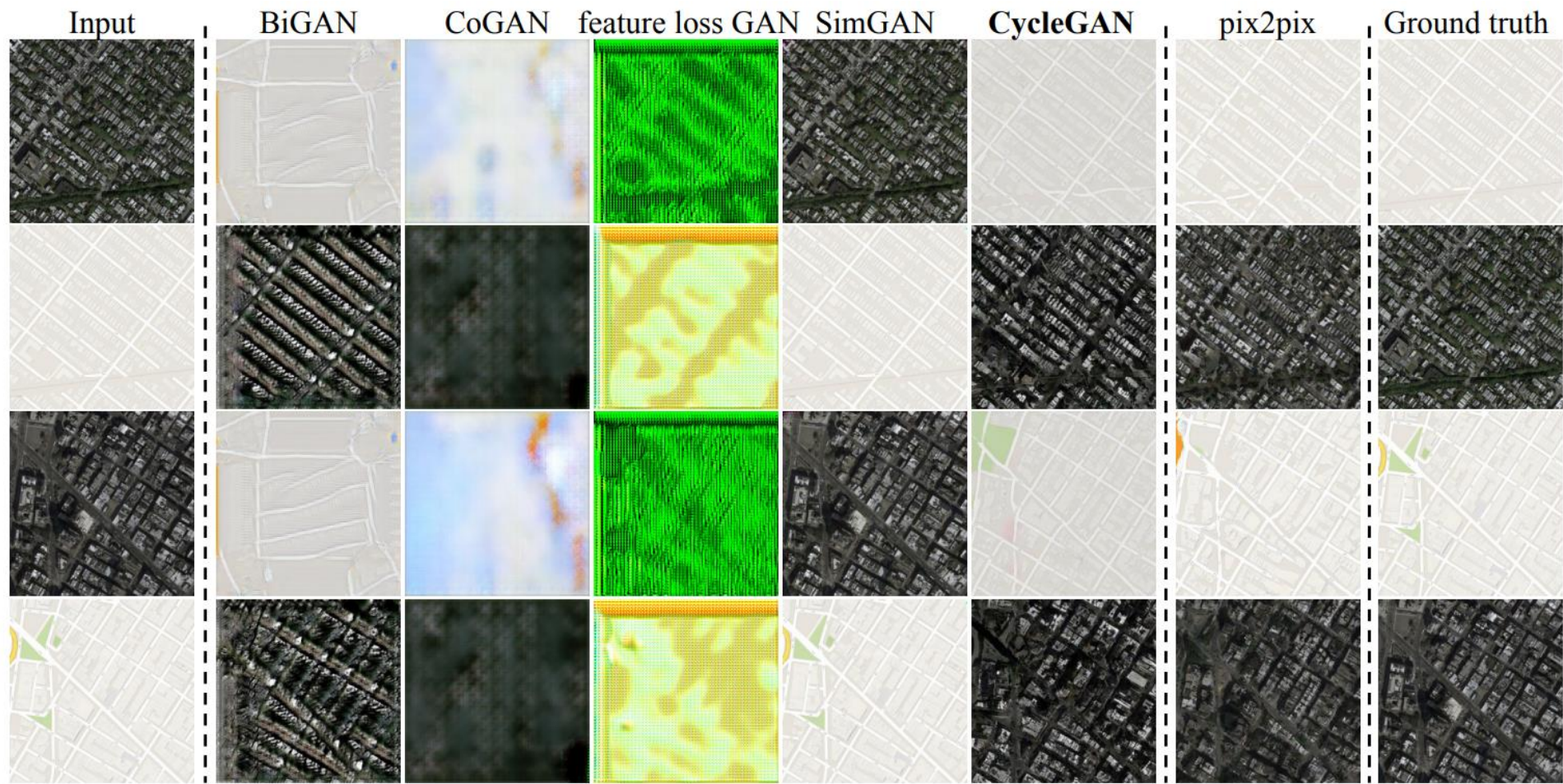


$$\mathcal{L}_{\text{GAN}}(G, D_Y, X, Y) = \mathbb{E}_{y \sim p_{\text{data}}(y)} [\log D_Y(y)] \\ + \mathbb{E}_{x \sim p_{\text{data}}(x)} [\log(1 - D_Y(G(x)))]$$

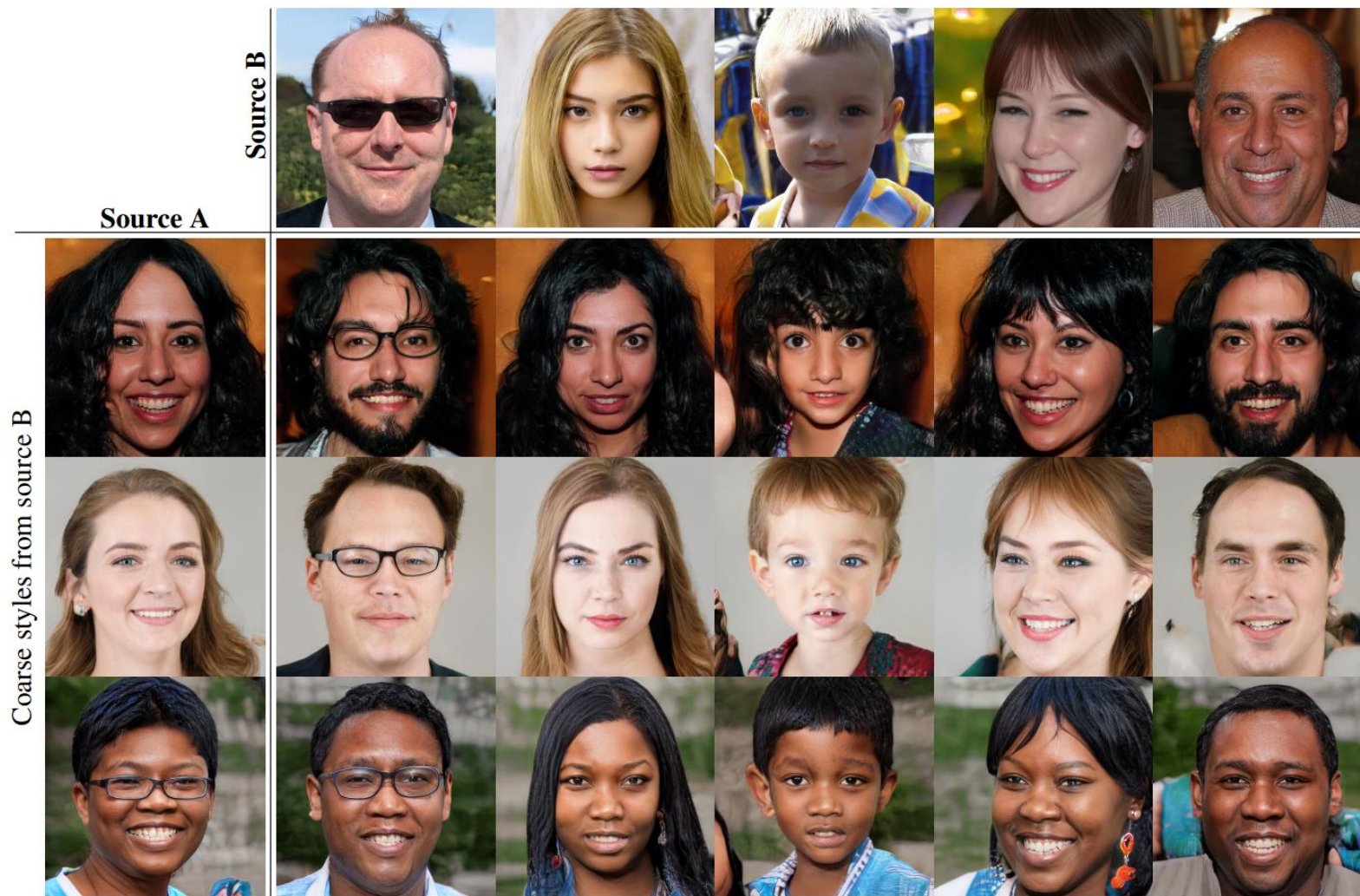
$$\mathcal{L}_{\text{cyc}}(G, F) = \mathbb{E}_{x \sim p_{\text{data}}(x)} [\|F(G(x)) - x\|_1] \\ + \mathbb{E}_{y \sim p_{\text{data}}(y)} [\|G(F(y)) - y\|_1]$$

$$\mathcal{L}(G, F, D_X, D_Y) = \mathcal{L}_{\text{GAN}}(G, D_Y, X, Y) \\ + \mathcal{L}_{\text{GAN}}(F, D_X, Y, X) \\ + \lambda \mathcal{L}_{\text{cyc}}(G, F)$$

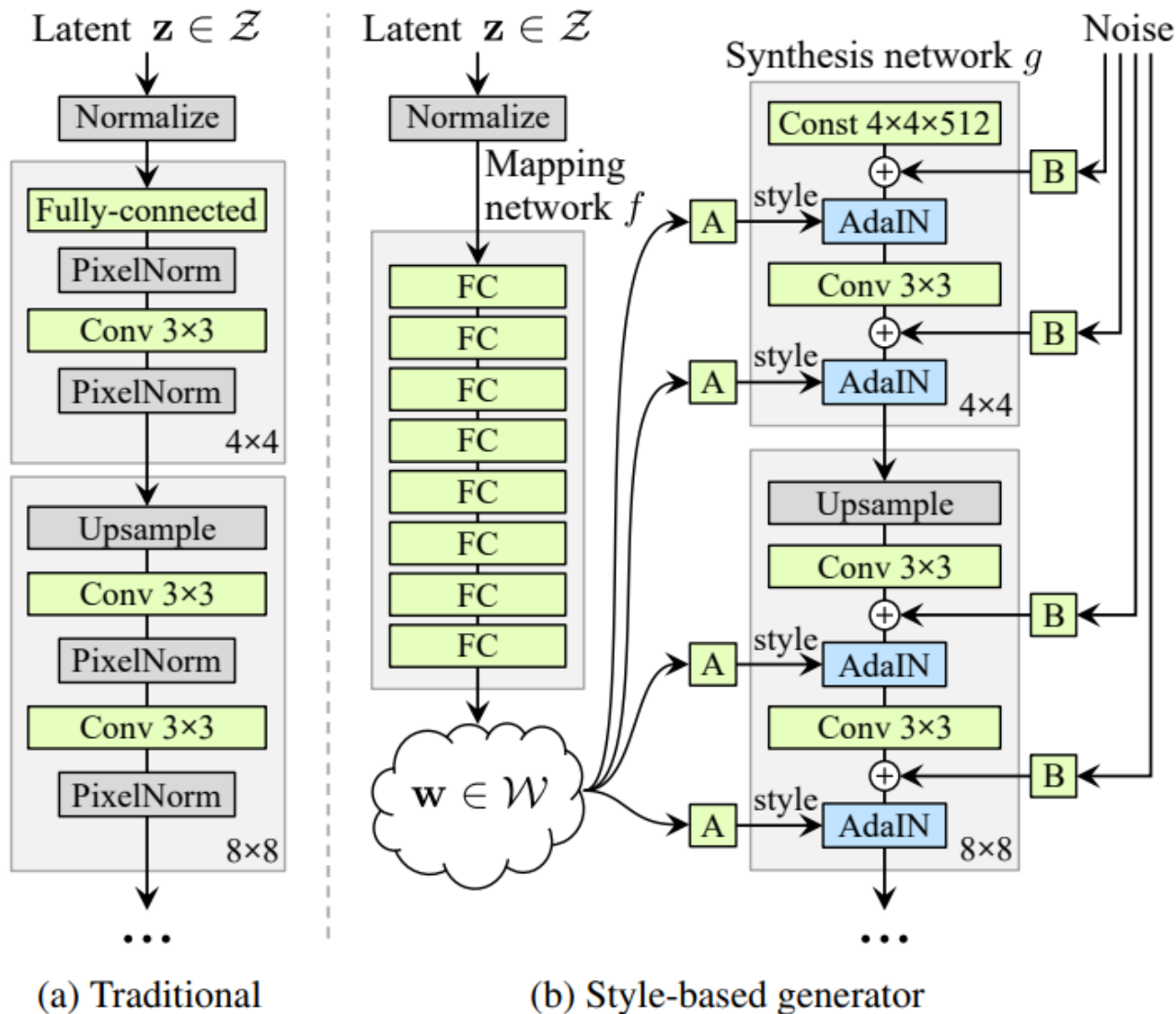
实验结果



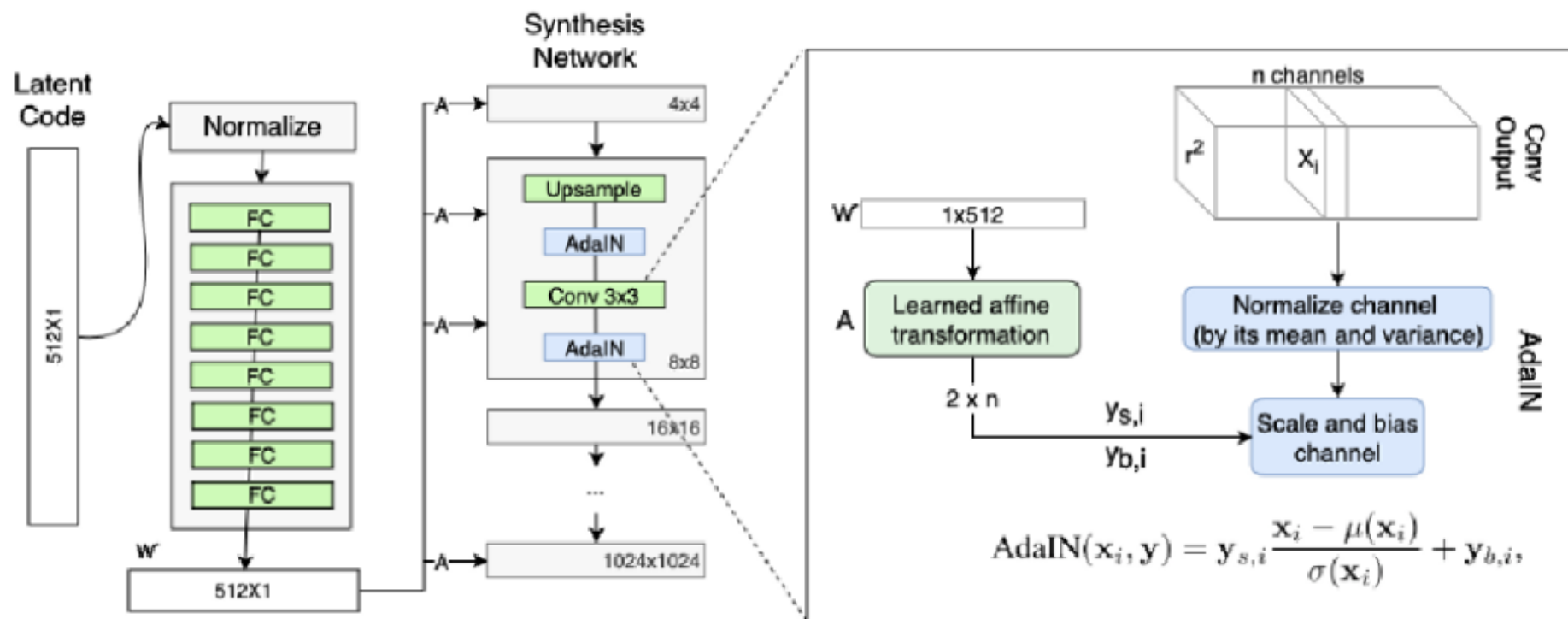
StyleGAN



StyleGAN 架构



AdaIN实现细节



AdaIN数学原理

- 在AdaIN公式中，特征图减去其均值除以方差，去掉自己的风格。再乘以新风格的方差加均值，实现风格转换。

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