

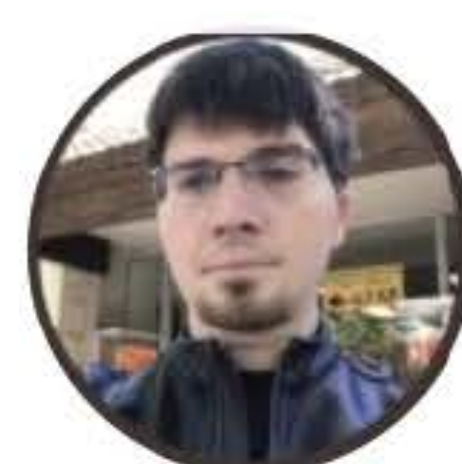
# 生成对抗网络

## An Introduction to Generative Adversarial Networks

- 生成对抗网络的背景
- 生成对抗网络的原理
- 生成对抗网络的变种


# 生成对抗网络的提出

- Ian Goodfellow 在 *Generative Adversarial Networks* 中首次提出了生成对抗网络 (GAN), 他也因此被誉为“GANs之父”



**Ian Goodfellow**

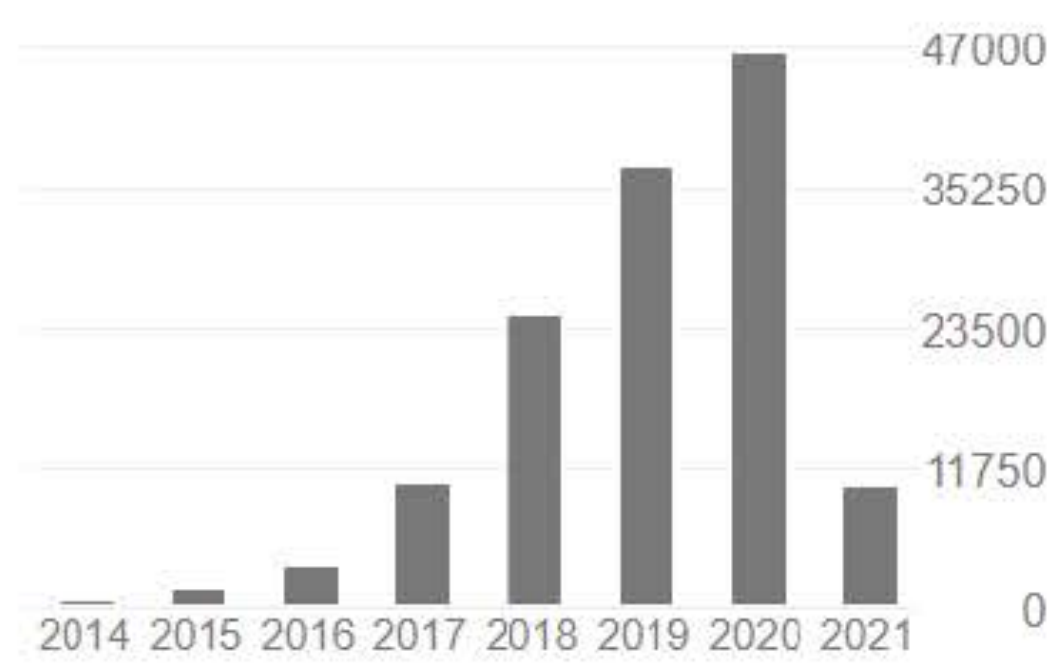
Unknown affiliation  
Verified email at cs.stanford.edu - [Homepage](#)  
[Deep Learning](#)



TITLE	CITED BY	YEAR
<a href="#">Generative adversarial networks</a> IJ Goodfellow, J Pouget-Abadie, M Mirza, B Xu, D Warde-Farley, S Ozair, ... arXiv preprint arXiv:1406.2661	29329	2014
<a href="#">Deep learning</a> I Goodfellow, Y Bengio, A Courville, Y Bengio MIT press 1 (2)	26673	2016

Cited by

	All	Since 2016
Citations	135513	132294
h-index	70	69
i10-index	110	107



# 酒后脑洞

- 2014年的一个晚上，Ian Goodfellow在酒吧给师兄庆祝博士毕业。一群工程师聚在一起探讨如何让自动生成照片。
- Goodfellow的朋友们提出：对构成照片的像素进行统计分析，使机器自己生成图像。
- Goodfellow一听觉得这个想法行不通，马上否决了它，然后边喝酒边思考。
- 突然他灵光一现：如果让两个神经网络互相对抗呢？（一个负责生成，一个负责判断生成的到底像不像）
- Goodfellow连夜写代码实现并测试，他本人都没想到，第一次就测试成功了，生成了质量不错的照片。那晚发明的东西就叫做GAN。



# GAN能做什么?

- 自然图像生成



请问上面四幅图像哪些是真实拍摄的照片？哪些是模型生成的假图？



# GAN能做什么?

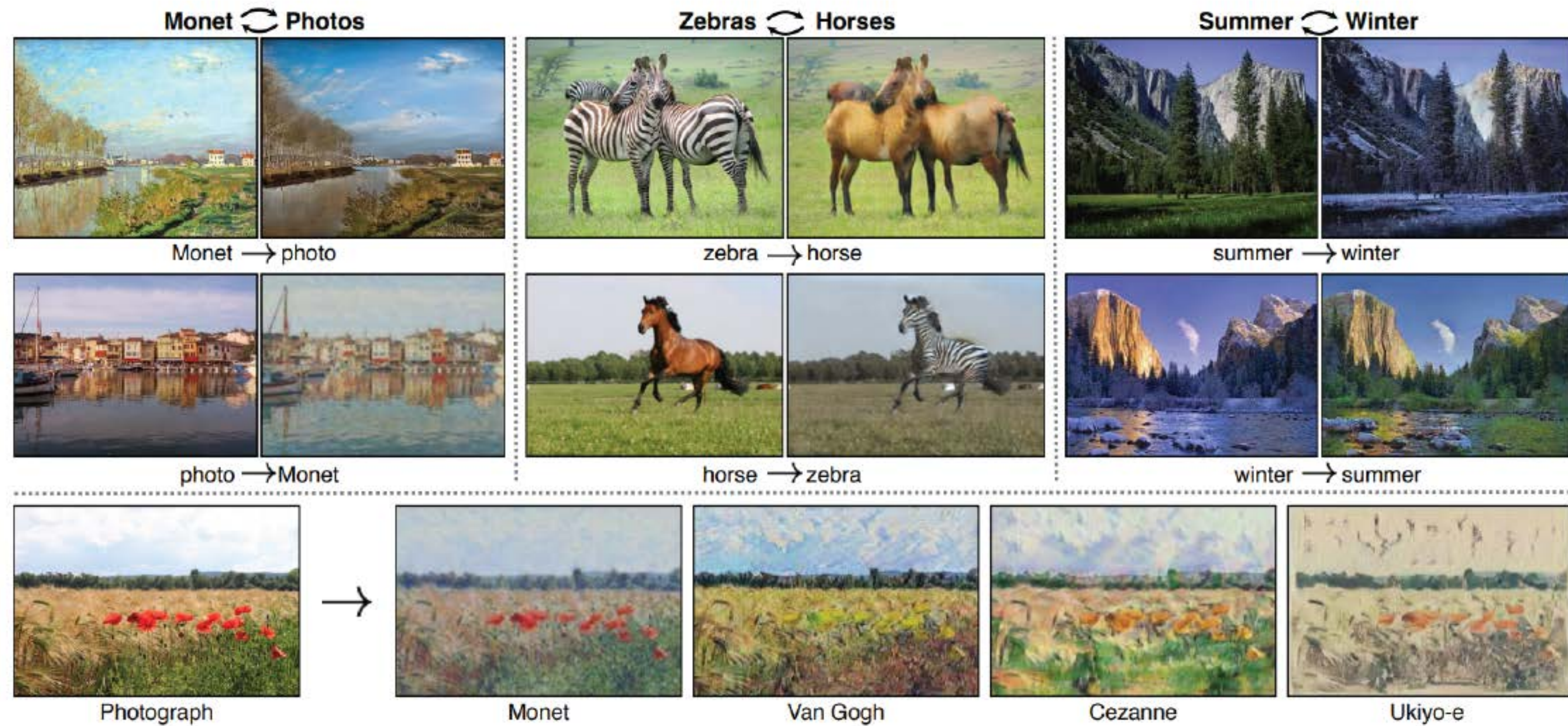
- 动漫头像生成





# GAN能做什么?

- 图像风格迁移





# GAN能做什么?

- 文字转图像

A family standing in front of a sign while wearing skis and holding ski poles.



A train being operated on a train track.



Three boys playing a soccer game on a green soccer field.



Two people in a speed boat on a body of water.



A bird with a brown and black wings, red crown and throat and the bill is short and pointed.



This is a white and grey bird with black wings and a black stripe by its eyes.



This bird has a yellow throat, belly, abdomen and sides with lots of brown streaks on them.



This bird has a white belly and breast, with a blue crown and nape.





# GAN能做什么?

- 图像编辑 (<https://www.artbreeder.com>)





# GAN能做什么?

- 图像复原





# 生成对抗网络

## An Introduction to Generative Adversarial Networks

- 生成对抗网络的背景
- 生成对抗网络的原理
- 生成对抗网络的变种

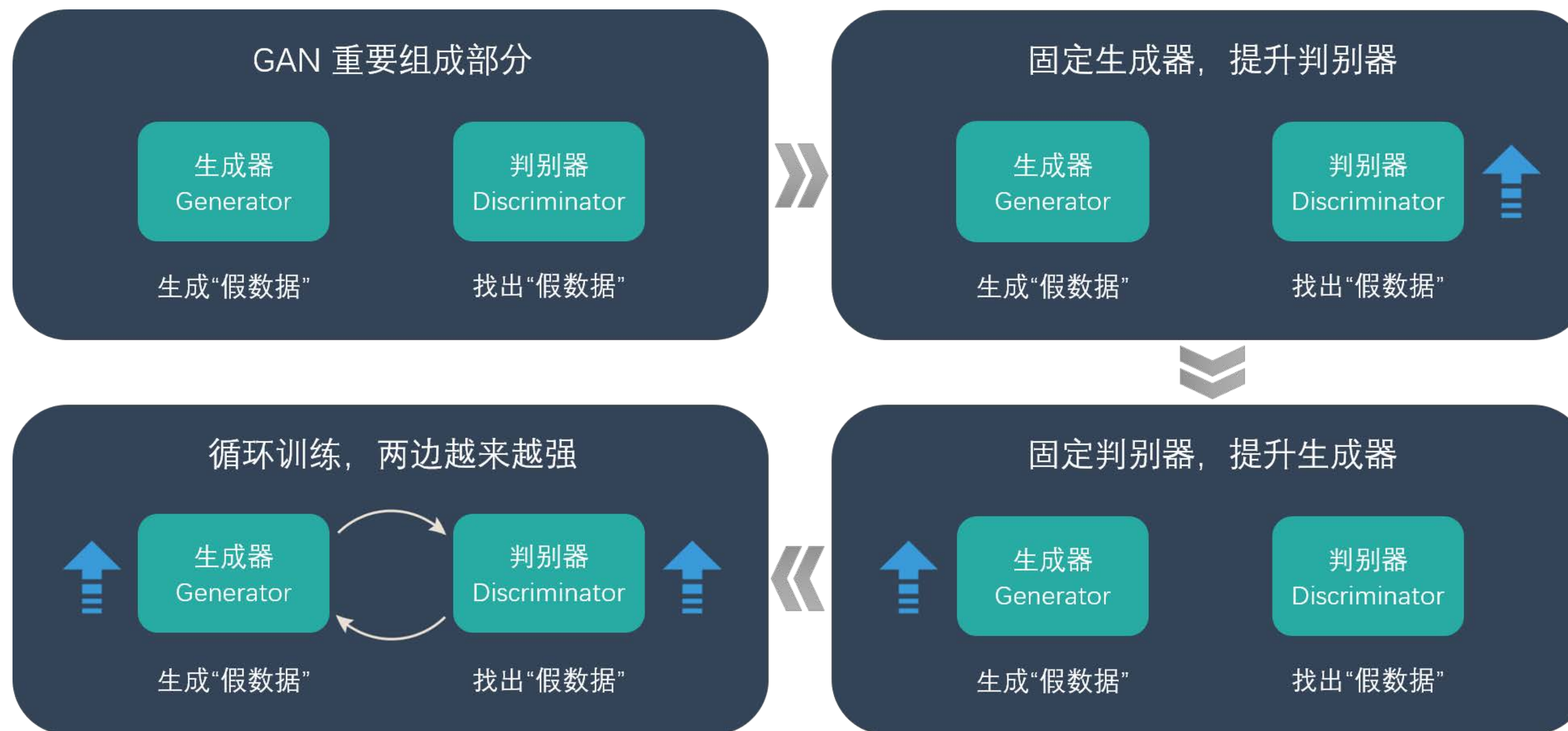


# 生成对抗？ 零和博弈？





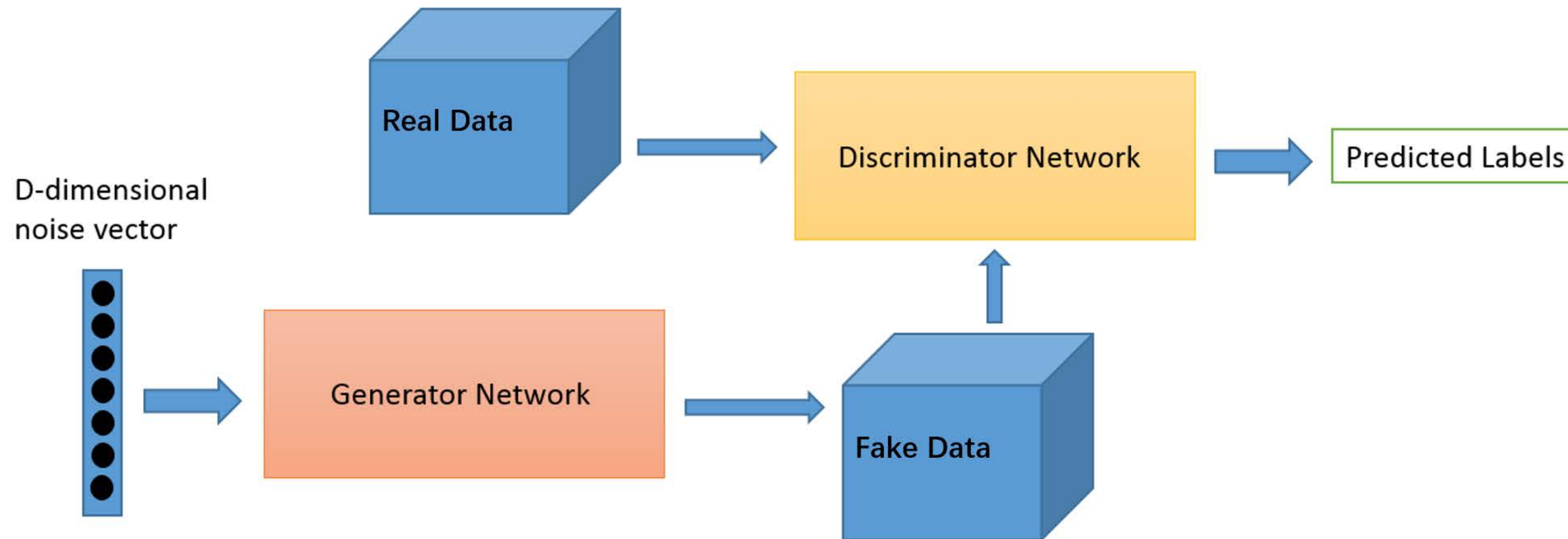
# 生成对抗？ 零和博弈？





# GAN的框架

- 生成网络接收一个随机的噪声，并生成“假数据”；判别网络接收“真数据”和生成的“假数据”，并预测不同数据的真实性概率。





# GAN的学习目标

- 假设 $G$ 为生成器， $D$ 为判别器，真实数据  $x$  服从 $p_{data}(x)$ 分布，噪声  $z$  服从 $p_z(z)$ 分布，目标函数 $V(D,G)$ 的优化可表示为：

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

- $D(\cdot)$ 计算数据的真实性概率（1为真，0为假）。
- 生成器的目标是生成尽可能真实的“假数据”，即 $D(G(z)) \rightarrow 1$ ；
- 判别器的目标是尽可能判断出数据的真假，即 $D(x) \rightarrow 1, D(G(z)) \rightarrow 0$ 。

目标函数可看作真实数据和生成数据的交叉熵之和。





# GAN的学习目标

实际操作分两步：

- 训练判别器 $D$ 时，固定生成器 $G$ 参数不变：

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

- 训练生成器 $G$ 时，固定判别器 $D$ 参数不变：

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

- 循环优化，直至收敛。





# GAN的优缺点

- 优点

1. 并未局限网络结构的具体形式，理论上，只要是可微分函数都可以用于构建生成器和判别器；
2. 生成器的参数更新不是直接来自数据样本，而是使用来自判别器的反向传播，因此网络设计更加灵活，降低了损失函数设计的困难。

- 缺点

1. 可解释性差，生成模型的分布没有显式的表达；
2. 比较难训练，生成器与判别器之间需要很好的同步，例如判别器更新  $k$  次而生成器只更新1次。





# 生成对抗网络

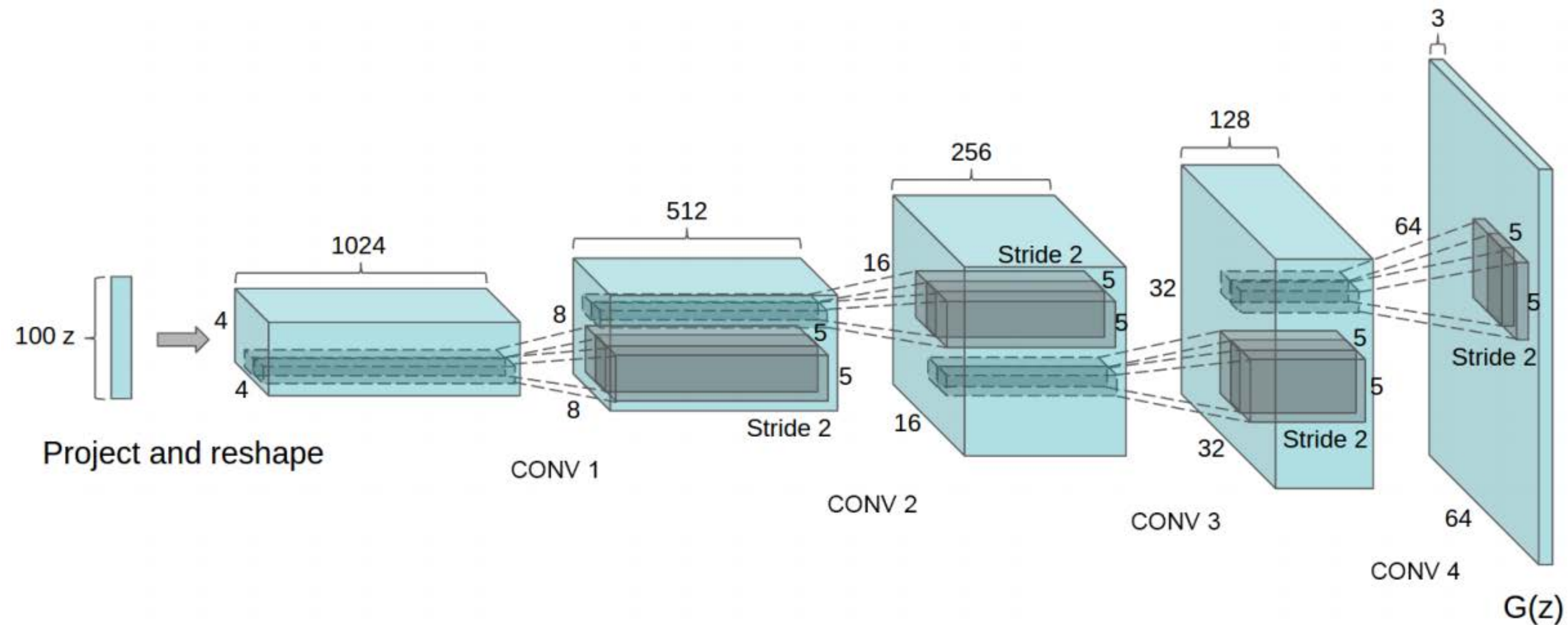
## An Introduction to Generative Adversarial Networks

- 生成对抗网络的背景
- 生成对抗网络的原理
- 生成对抗网络的变种



# DCGAN

- Deep Convolutional GAN, 将深度卷积神经网络CNN与生成对抗网络GAN结合, 设计了一种通用的图像生成网络结构。





# DCGAN

- DCGAN的改进包含以下几个方面:
- (1) 在生成器中使用转置卷积进行上采样, 判别器中使用步长卷积代替池化层
- (2) 在生成器和判别器中都添加了批量归一化操作 (batch normalization)
- (3) 判别器中去掉了全连接层, 使用全局池化层替代, 使网络变为全卷积网络
- (4) 生成器中使用ReLU作为激活函数, 最后输出层使用Tanh激活函数
- (5) 判别器中使用LeakyReLU激活函数

# DCGAN

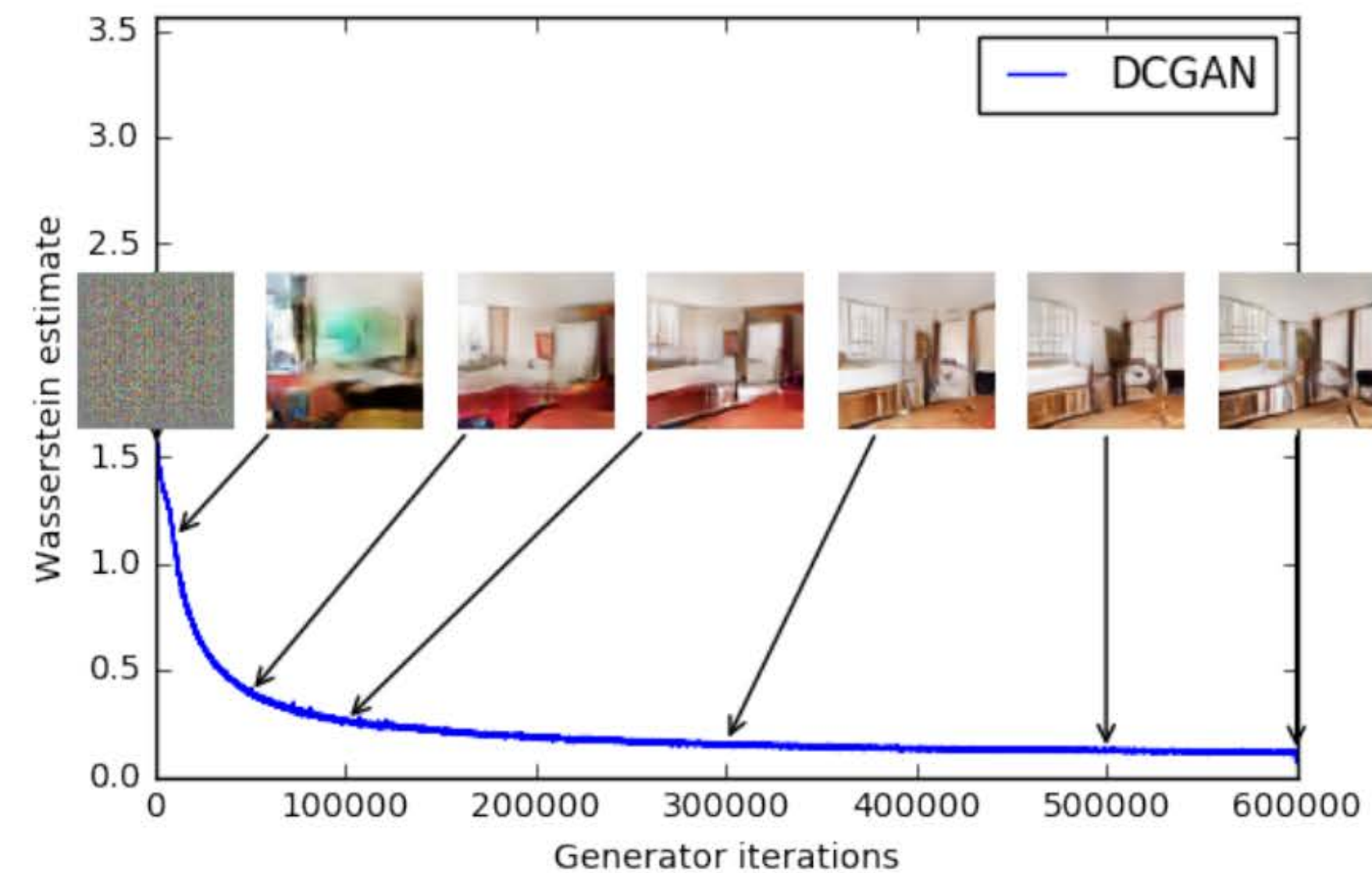
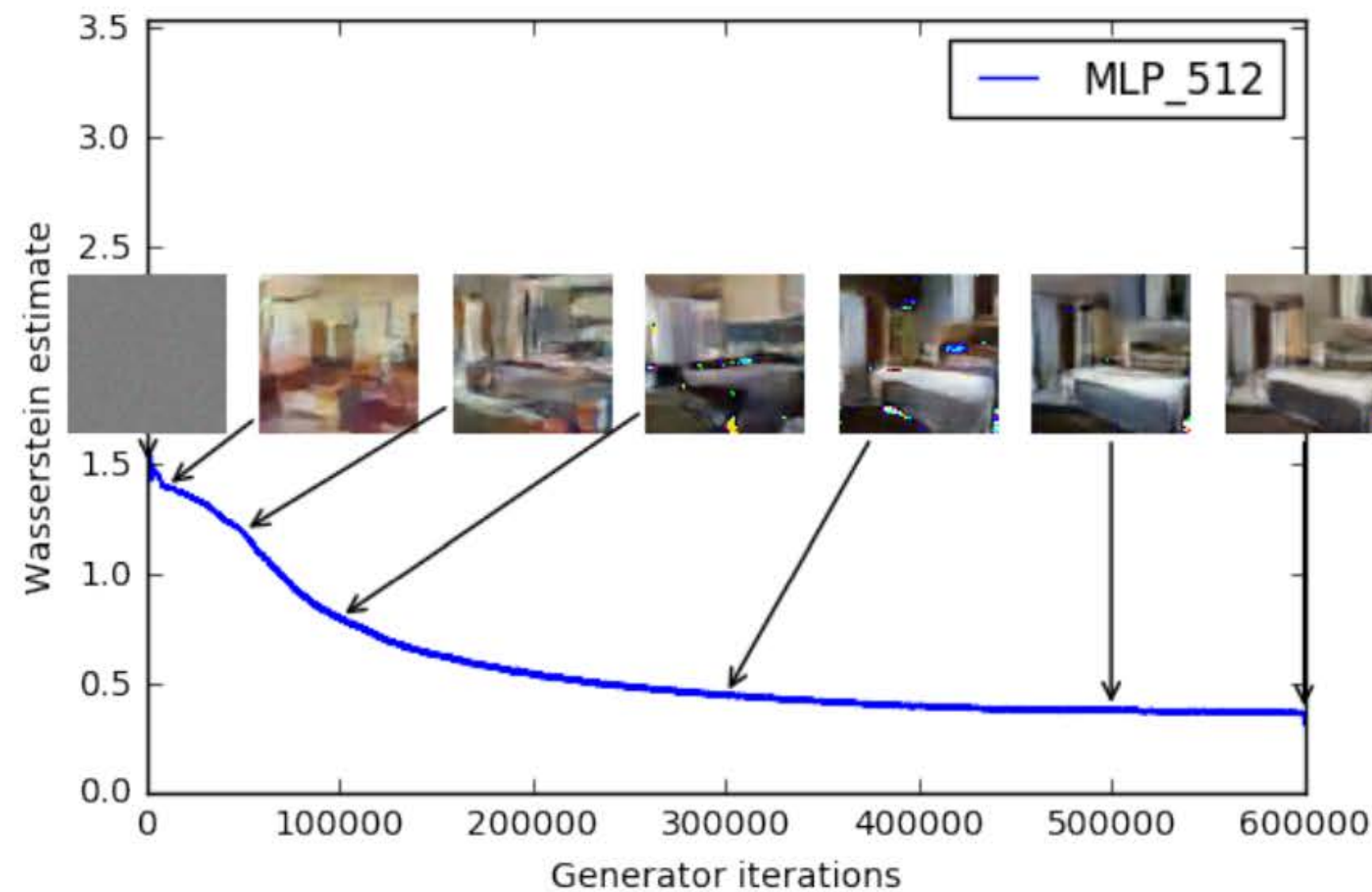
- **局限性**: DCGAN依靠实验，找到了一组比较好的网络架构设置，但治标不治本，网络的训练仍然存在以下问题：
- 生成器和判别器的损失函数无法指导训练进程
- 生成样本缺乏多样性，容易出现模式崩塌（mode collapse）





# WGAN

- Wasserstein GAN (Wasserstein 是一种经典的距离度量方式，也叫做推土机距离) 解释了GAN训练不稳定的原因，并给出了理论证明和解决方法。



# WGAN

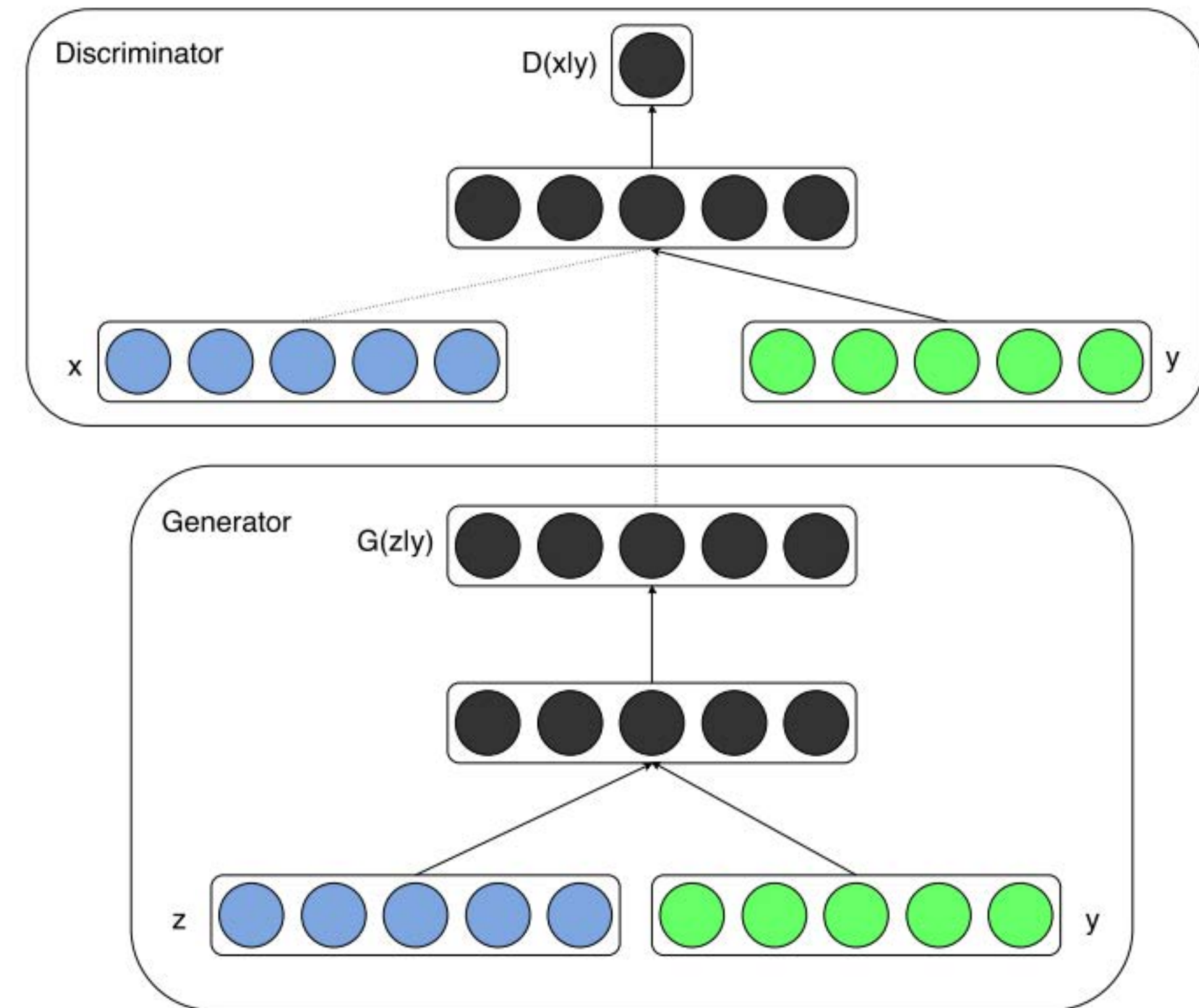
- WGAN 的具体改进如下：
  - 判别器最后一层去掉sigmoid
  - 生成器和判别器的损失函数不取log（Wasserstein 距离）
  - 每次更新判别器的参数后，把参数截断在某个范围内
  - 不使用基于动量的优化算法如momentum和Adam，推荐使用RMSProp，SGD也行



# CGAN

- Conditional GAN, 在生成器和判别器中同时加入条件约束来引导数据的生成过程，实现控制图像的生成类别。

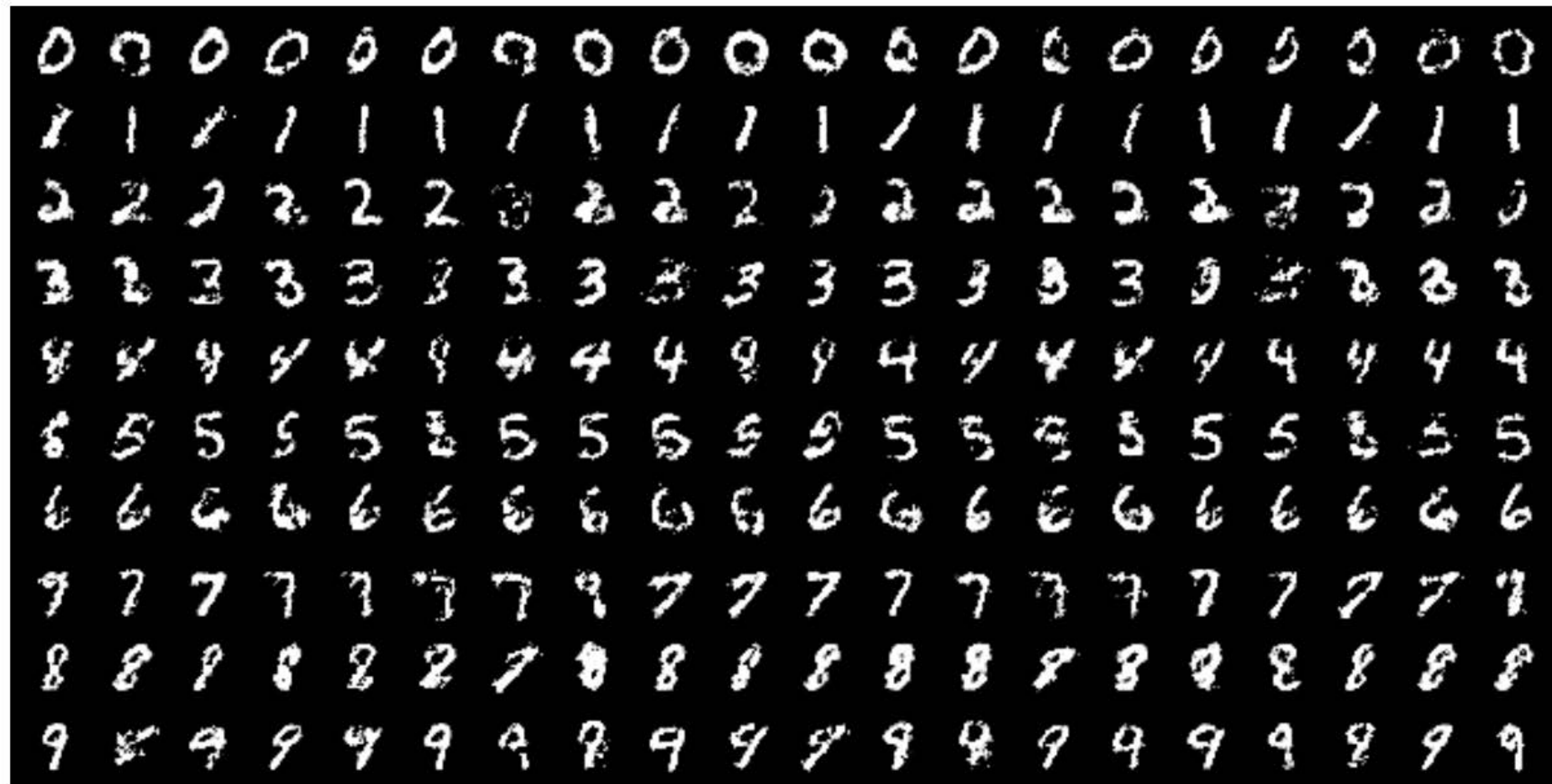
$$\begin{aligned} & \min_G \max_D V(D, G) \\ &= \mathbb{E}_{\mathbf{x} \sim p_{data}(\mathbf{x})} [\log D(\mathbf{x}|\mathbf{y})] + \\ & \quad \mathbb{E}_{\mathbf{z} \sim p_{\mathbf{Z}}(\mathbf{z})} [\log(1 - D(G(\mathbf{z}|\mathbf{y})))] \end{aligned}$$



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

# CGAN

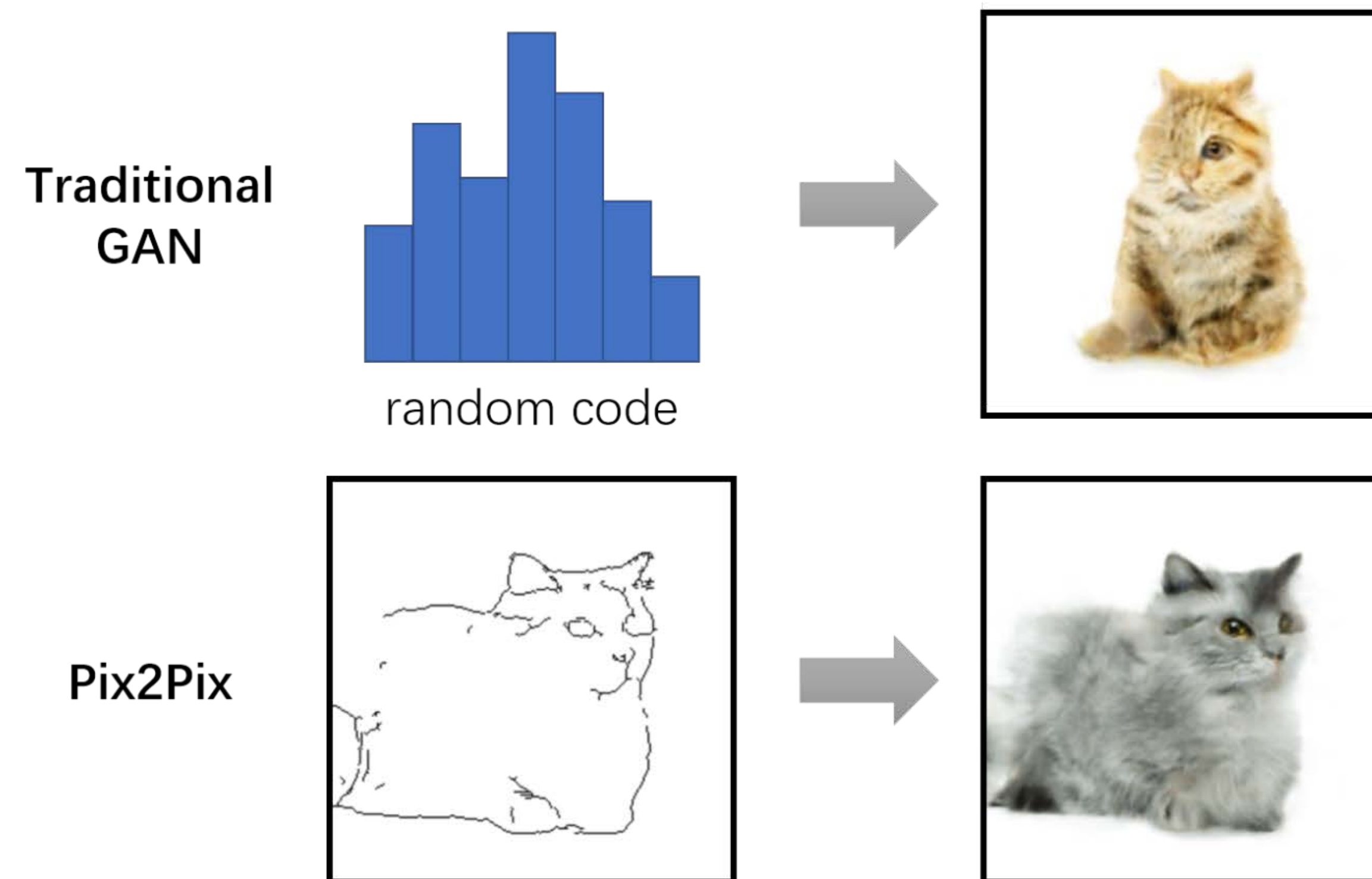
- 在MNIST数据集上训练好的CGAN能够自由控制生成数字的类别。





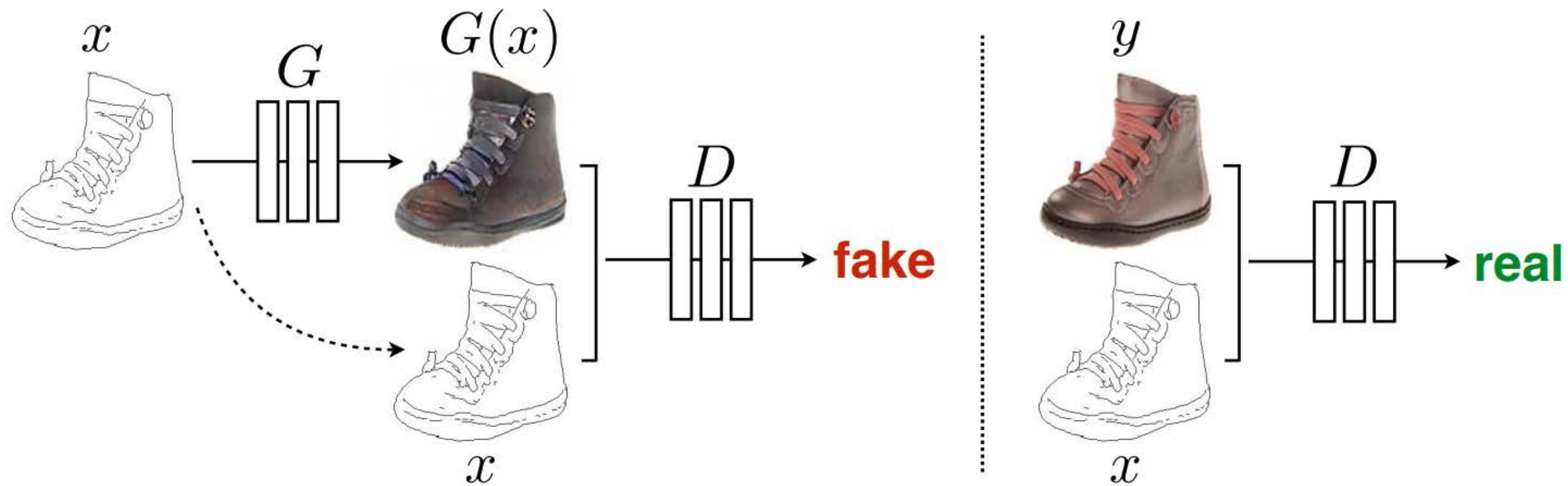
# Pix2Pix

- 传统GAN输入随机噪声，输出随机图像。但有时候用户希望输出图像和输入是对应的、有关联的。



# Pix2Pix

- Pix2Pix是将CGAN应用于有监督的图像到图像翻译任务，有监督表示Pix2Pix使用成对的数据进行训练。

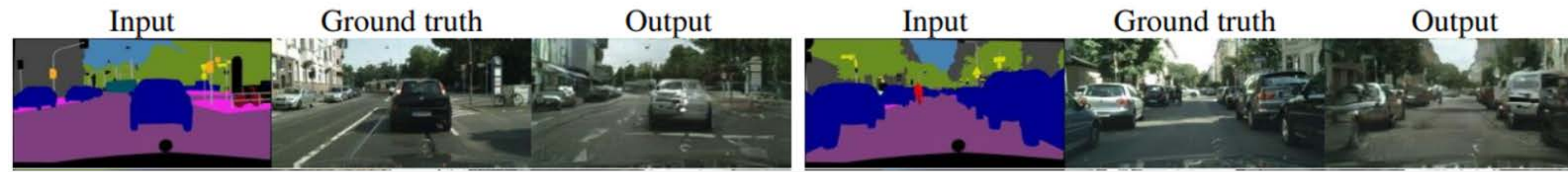


$$\mathcal{L}(G, D) = \mathcal{L}_{CGAN}(G, D) + \mathcal{L}_{\ell_1}(G)$$

$$\mathcal{L}_{\ell_1}(G) = \mathbb{E}_{\mathbf{x}, \mathbf{y}, \mathbf{z}} [\|\mathbf{y} - G(\mathbf{x}, \mathbf{z})\|_1]$$



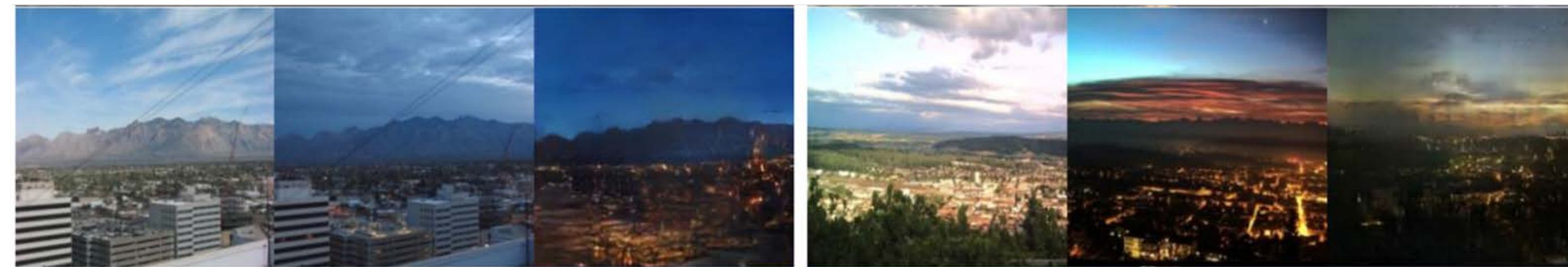
# Pix2Pix



Cityscapes labels → photo



facades labels → photo



day → night

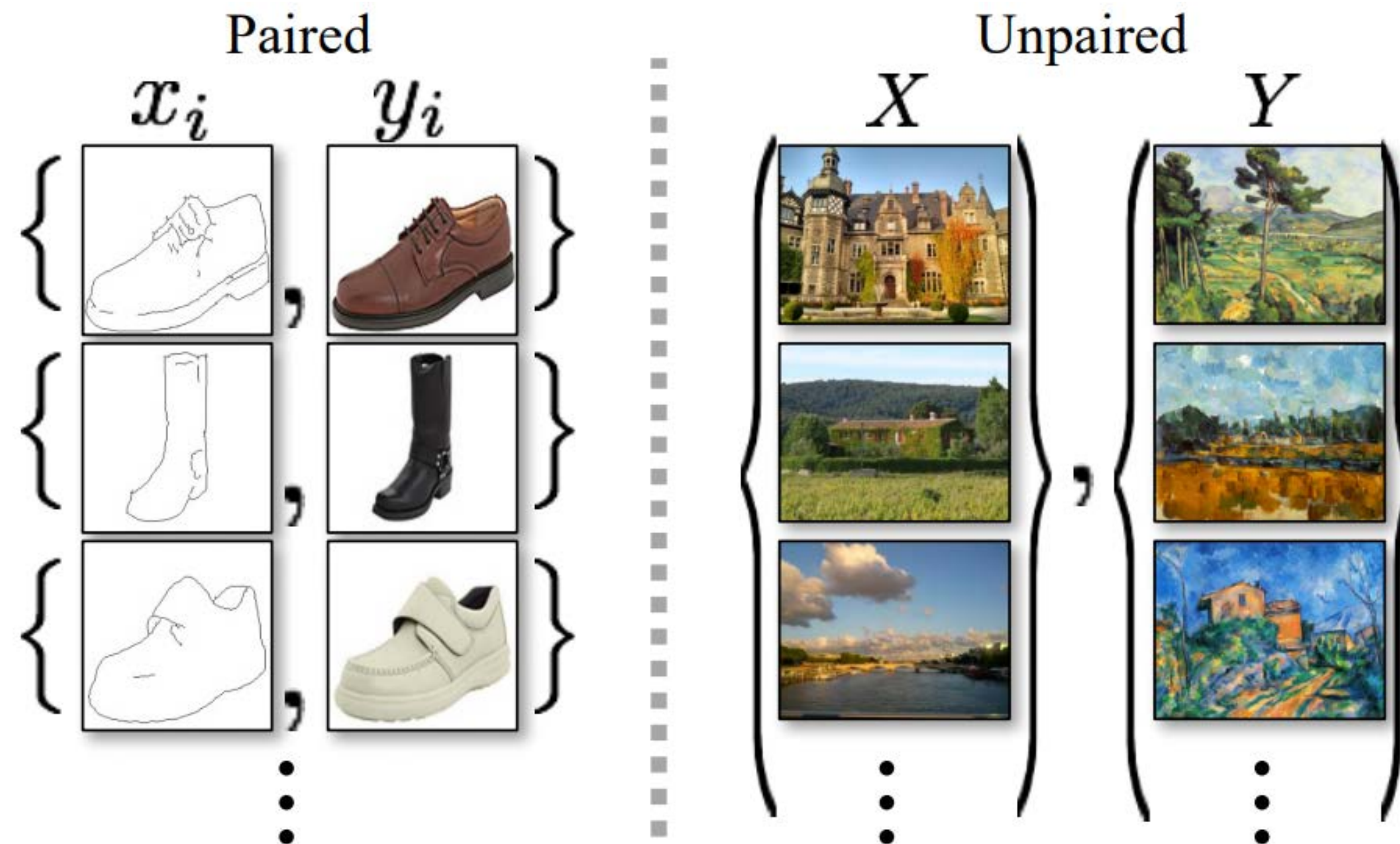


edges → handbags



# CycleGAN

- CycleGAN在Pix2Pix的基础上更进一步，使用不成对的数据即可训练网络，大大降低了数据收集的难度。

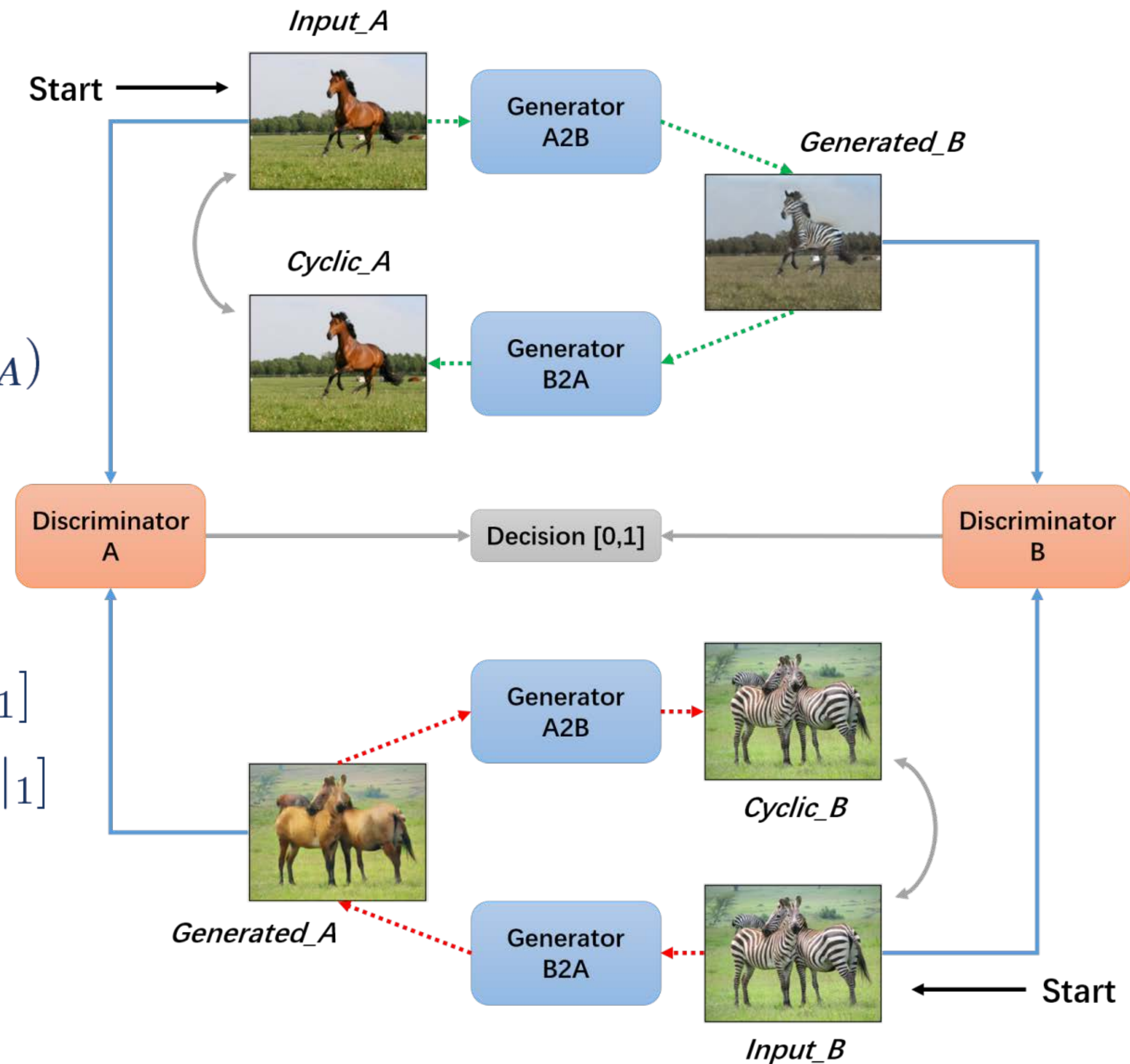




# CycleGAN

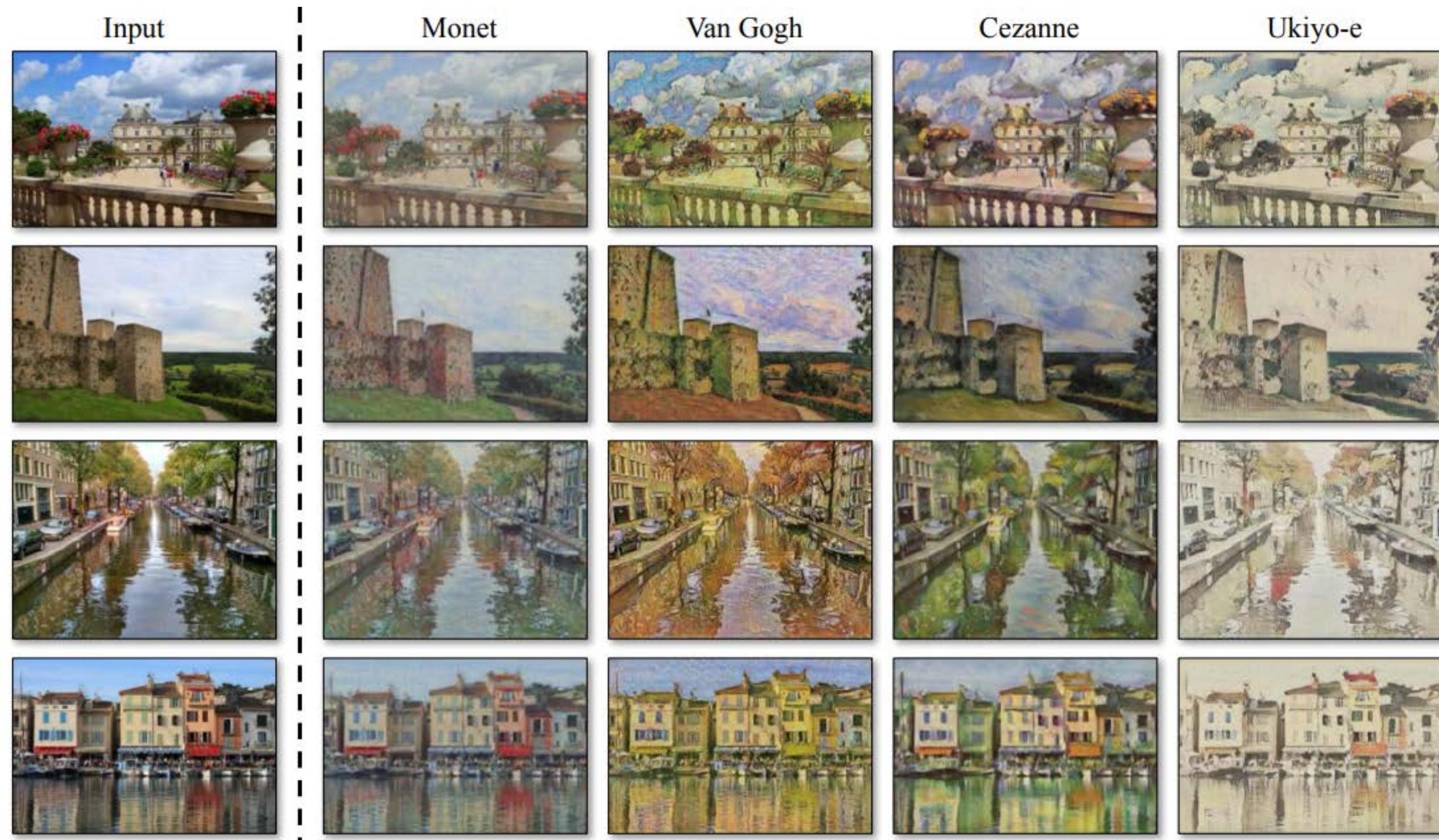
$$\begin{aligned} & \mathcal{L}(G_{A2B}, G_{B2A}, D_A, D_B) \\ &= \mathcal{L}_{GAN}(G_{A2B}, D_B) + \mathcal{L}_{GAN}(G_{B2A}, D_A) \\ &+ \lambda \mathcal{L}_{cyc}(G_{A2B}, G_{B2A}) \end{aligned}$$

$$\begin{aligned} & \mathcal{L}_{cyc}(G_{A2B}, G_{B2A}) \\ &= \mathbb{E}_{\mathbf{x} \sim p_{data}(\mathbf{x})} [\|G_{A2B}(G_{B2A}(\mathbf{x})) - \mathbf{x}\|_1] \\ &+ \mathbb{E}_{\mathbf{y} \sim p_{data}(\mathbf{y})} [\|G_{B2A}(G_{A2B}(\mathbf{y})) - \mathbf{y}\|_1] \end{aligned}$$



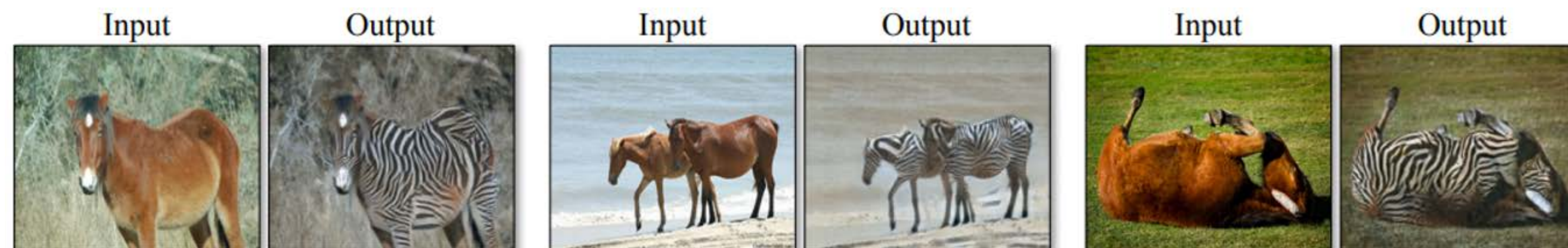


# CycleGAN





# CycleGAN



horse → zebra



zebra → horse



apple → orange

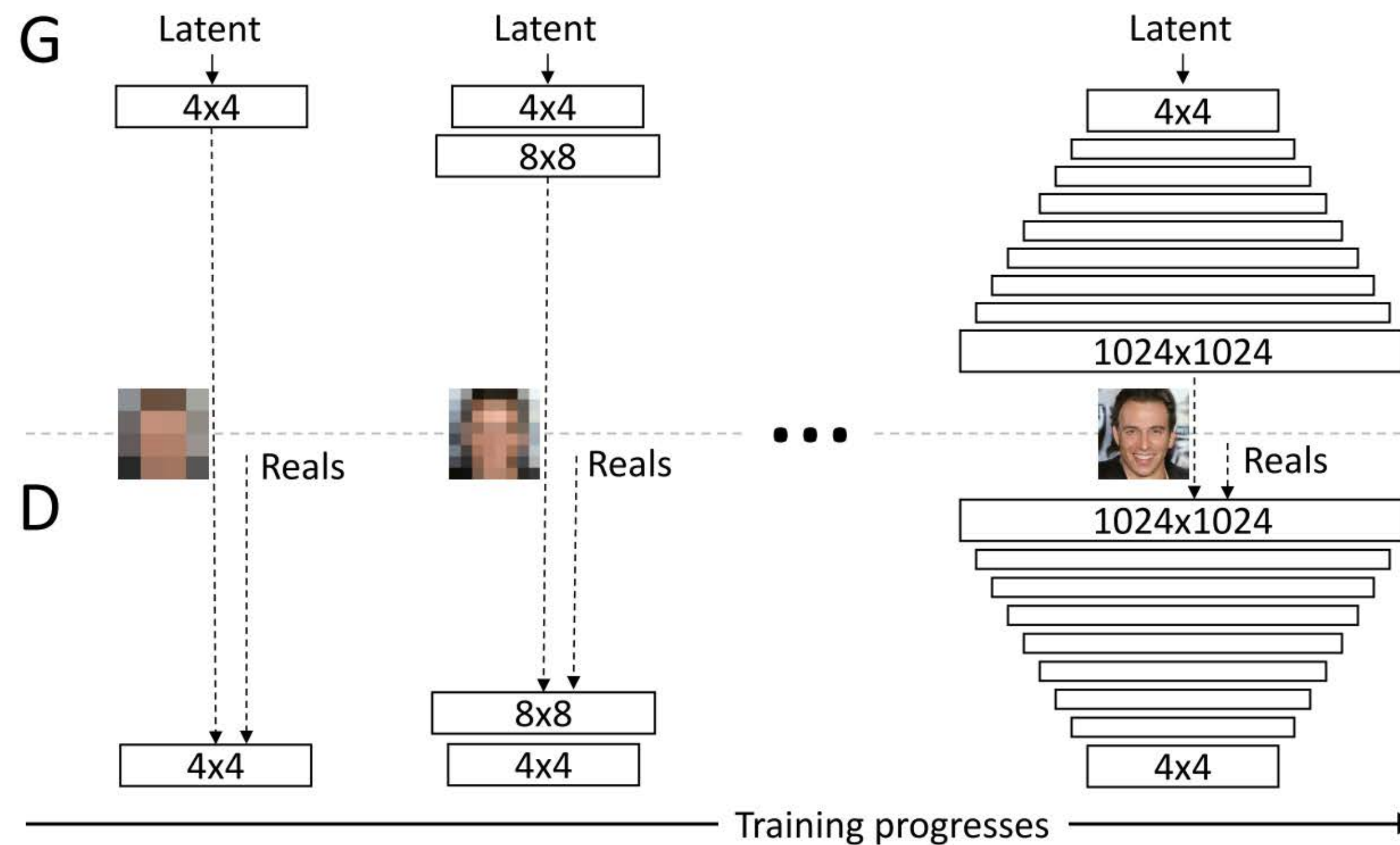


orange → apple



# PGGAN

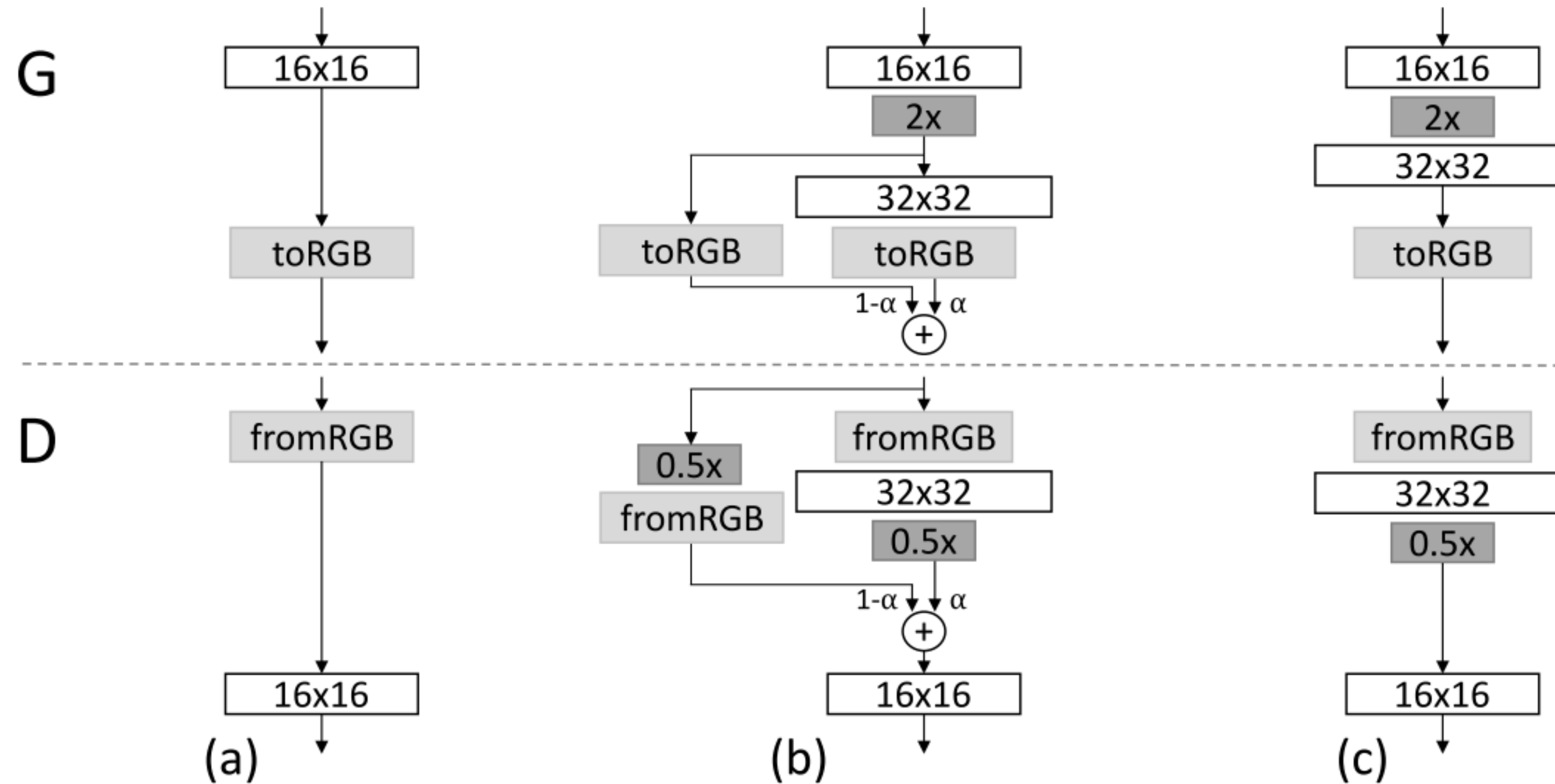
- 用GAN直接训练生成高分辨率图像时不稳定情况会加剧，Progressive GAN对此提出了一种渐进训练方式。





# PGGAN

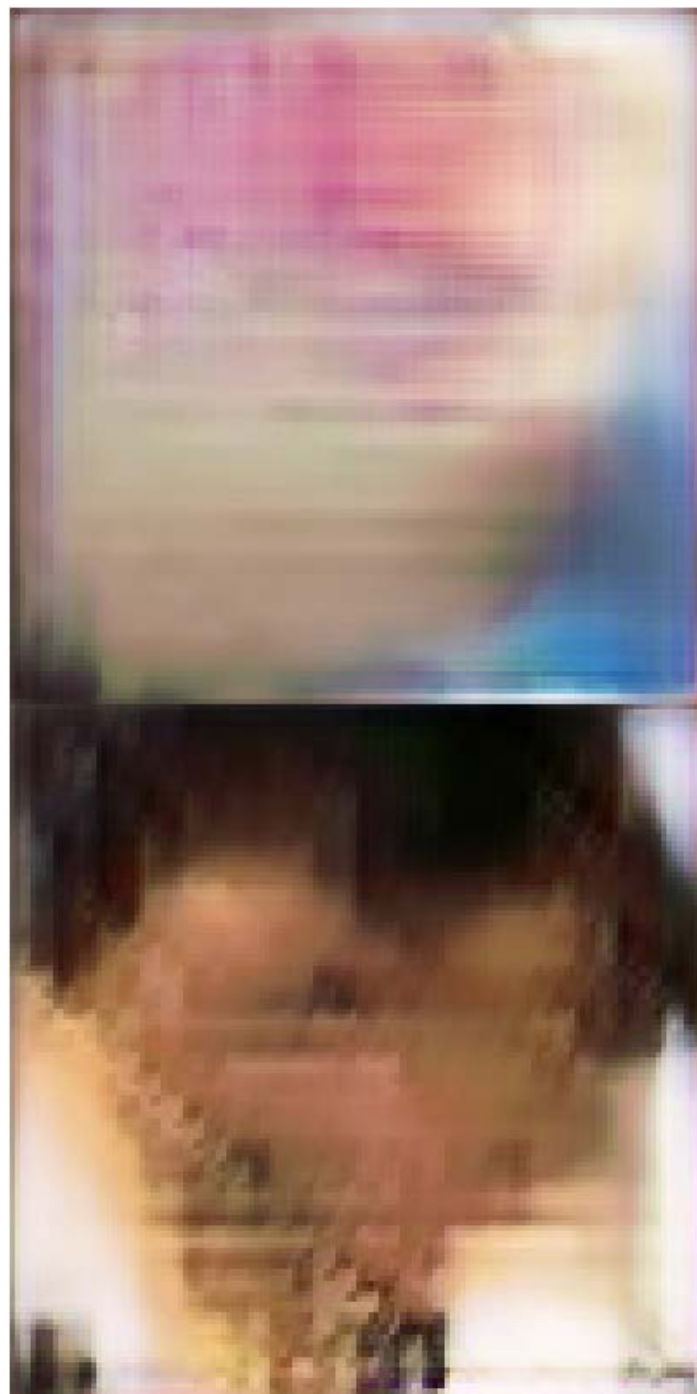
- PGGAN递增分辨率过程：增加新的网络层时利用  $\alpha$  进行平滑过渡



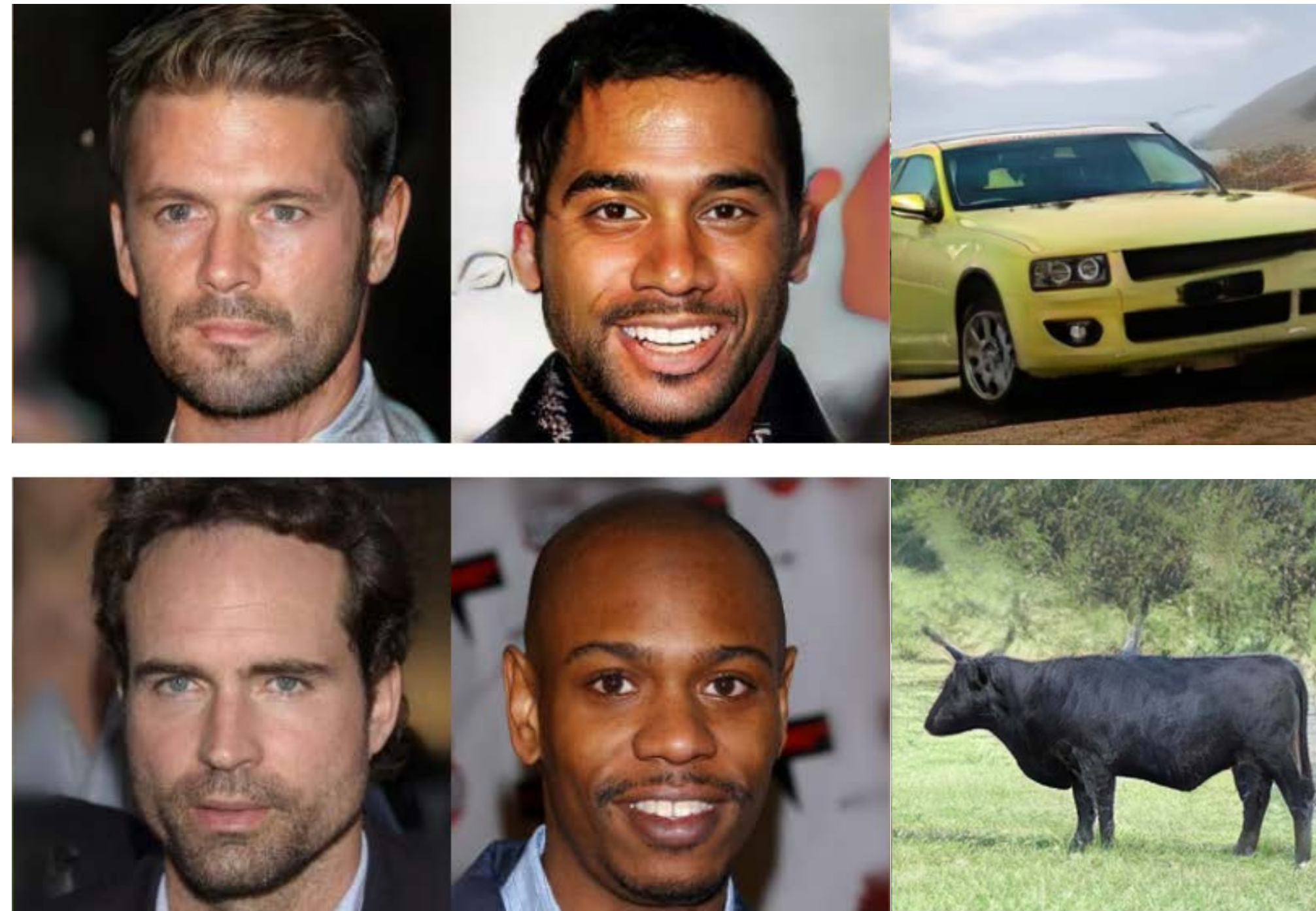
# PGGAN

- PGGAN在人脸与自然场景下生成的高分辨率图像

直接生成结果



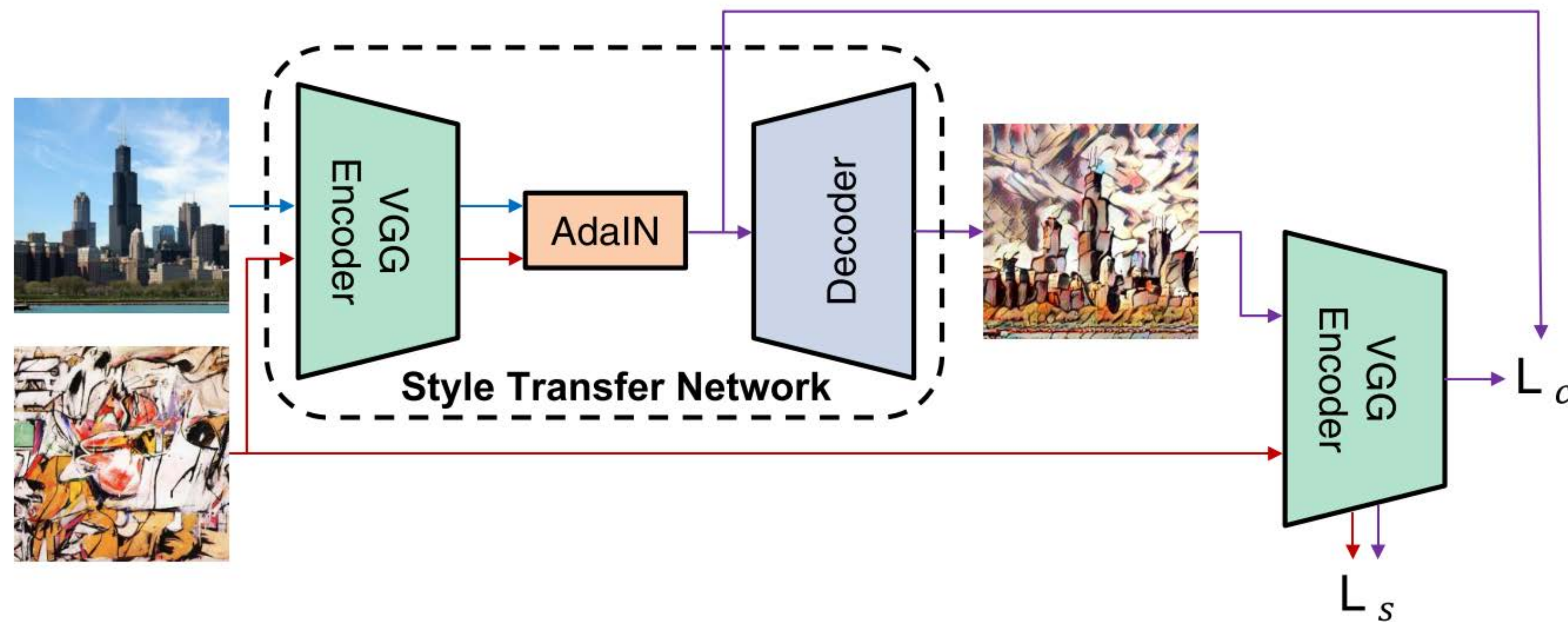
Progressive生成结果





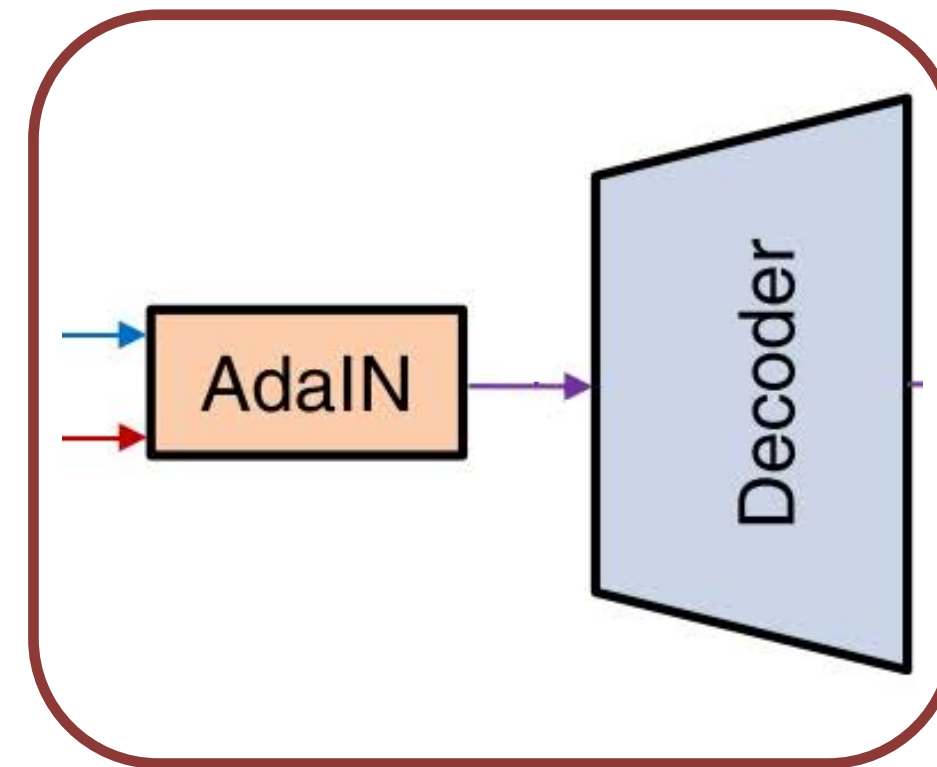
# StyleGAN

- StyleGAN在PGGAN的基础上，受风格迁移中AdaIN操作的启发，对传统生成器网络结构进行了修改



# StyleGAN

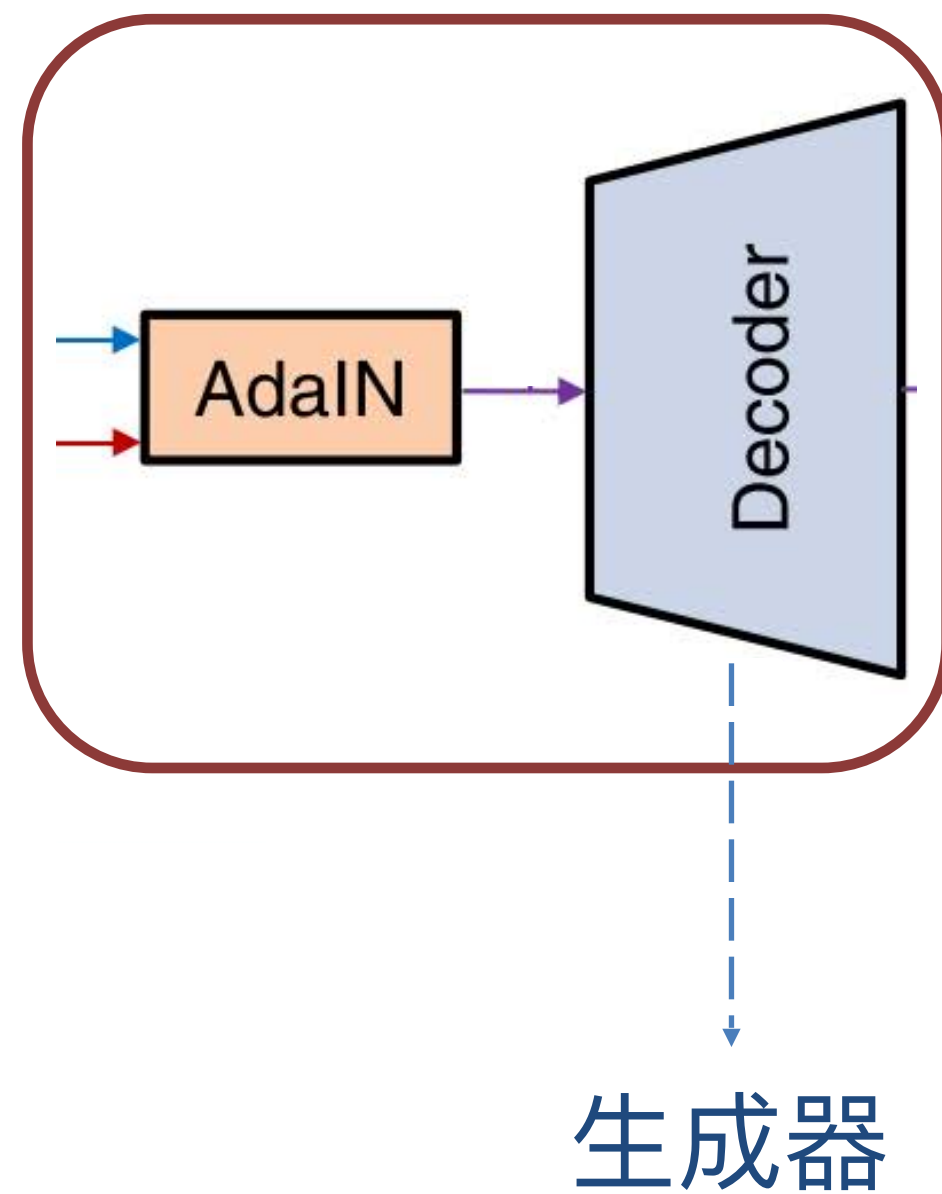
- StyleGAN在PGGAN的基础上，受启发于风格迁移中的AdaIN操作，对传统生成器网络结构进行修改





# StyleGAN

- StyleGAN在PGGAN的基础上，受启发于风格迁移中的AdaIN操作，对传统生成器网络结构进行修改

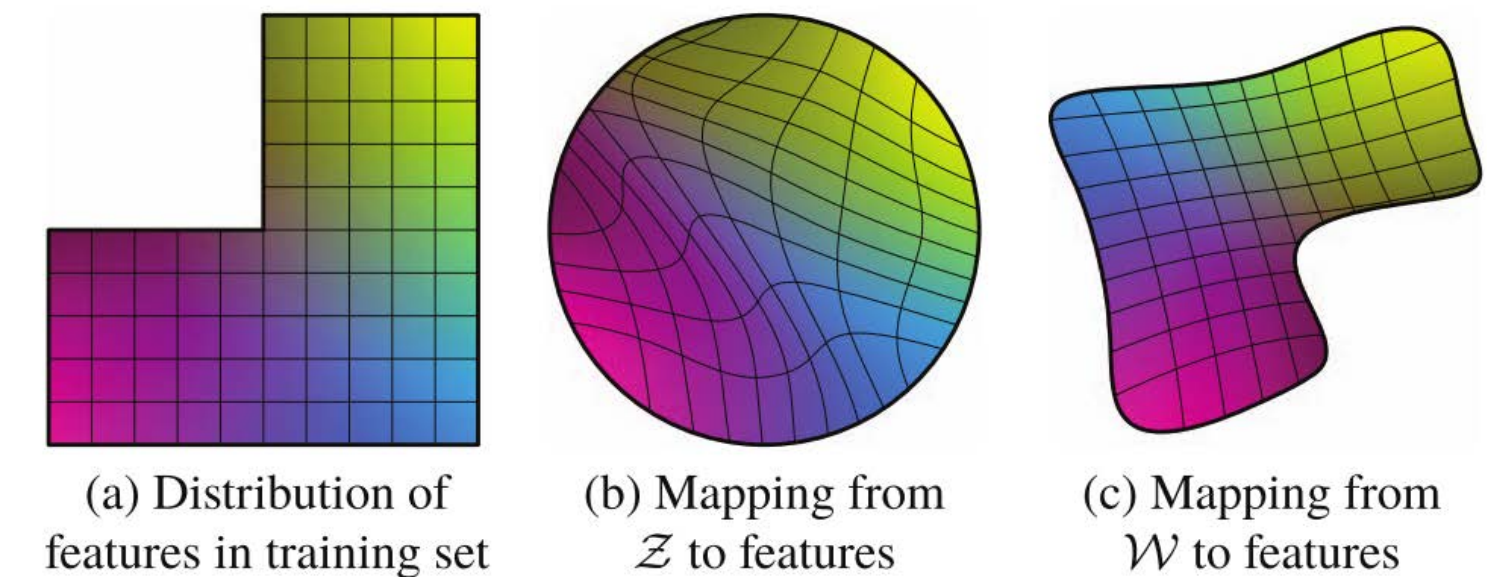
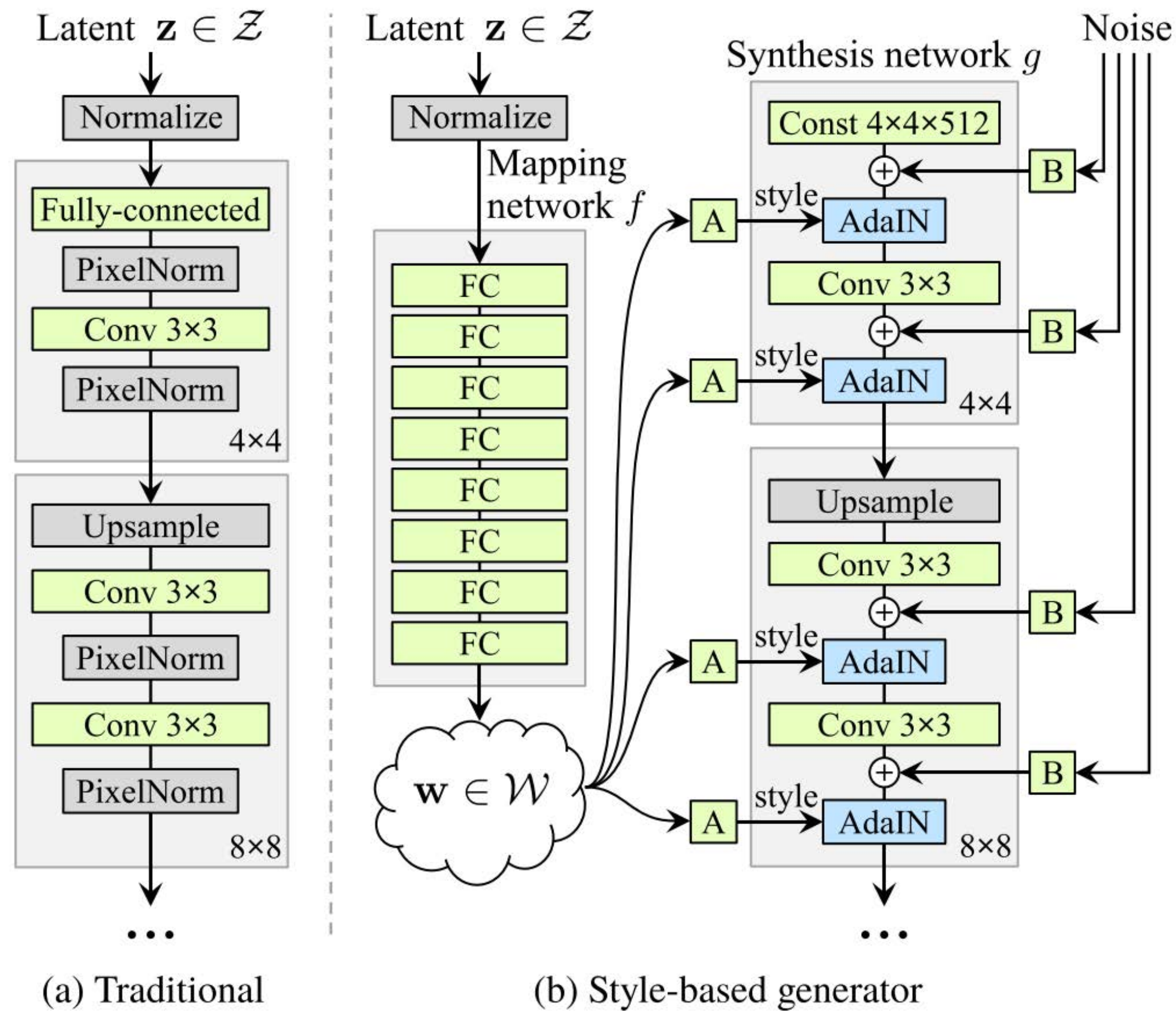


$$\text{IN}(x) = \gamma \left( \frac{x - \mu(x)}{\sigma(x)} \right) + \beta$$

$$\text{AdaIN}(x, y) = \sigma(y) \left( \frac{x - \mu(x)}{\sigma(x)} \right) + \mu(y)$$

如果不是将随机噪声直接输入生成器中，而是由随机噪声产生 $\sigma(y)$ ,  $\mu(y)$ ，使用AdaIN操作间接控制生成器前传过程会怎样？

# StyleGAN



## Mixing Regularization

Train with new  $w$

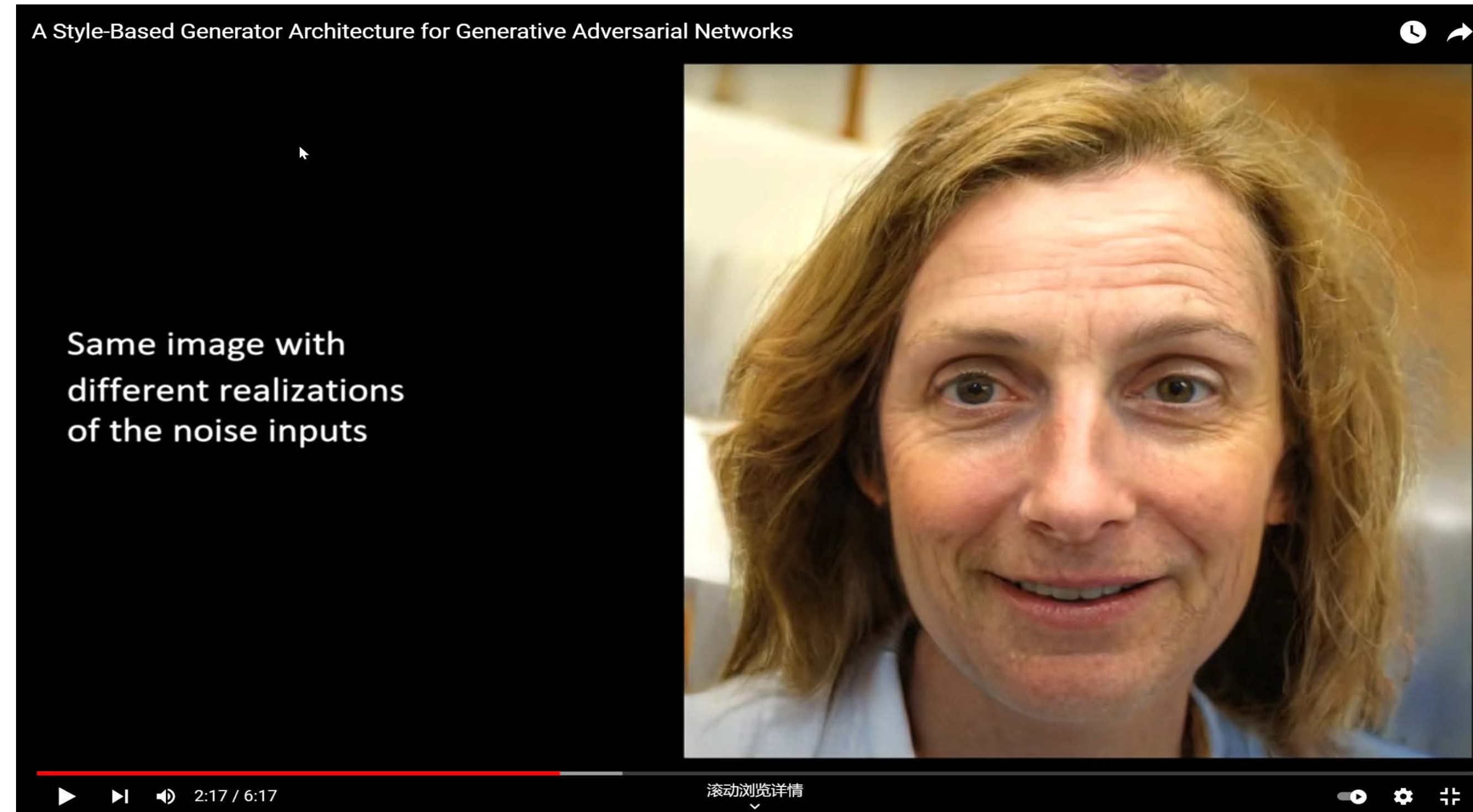
$$w_{new} = \text{concat}(w_1[0:n], w_2[n:0])$$

- Mapping网络一定程度上解耦了潜空间 (z-space  $\rightarrow$  w-space)
- 生成器架构中的分层AdaIN操作可以做到属性的分层控制



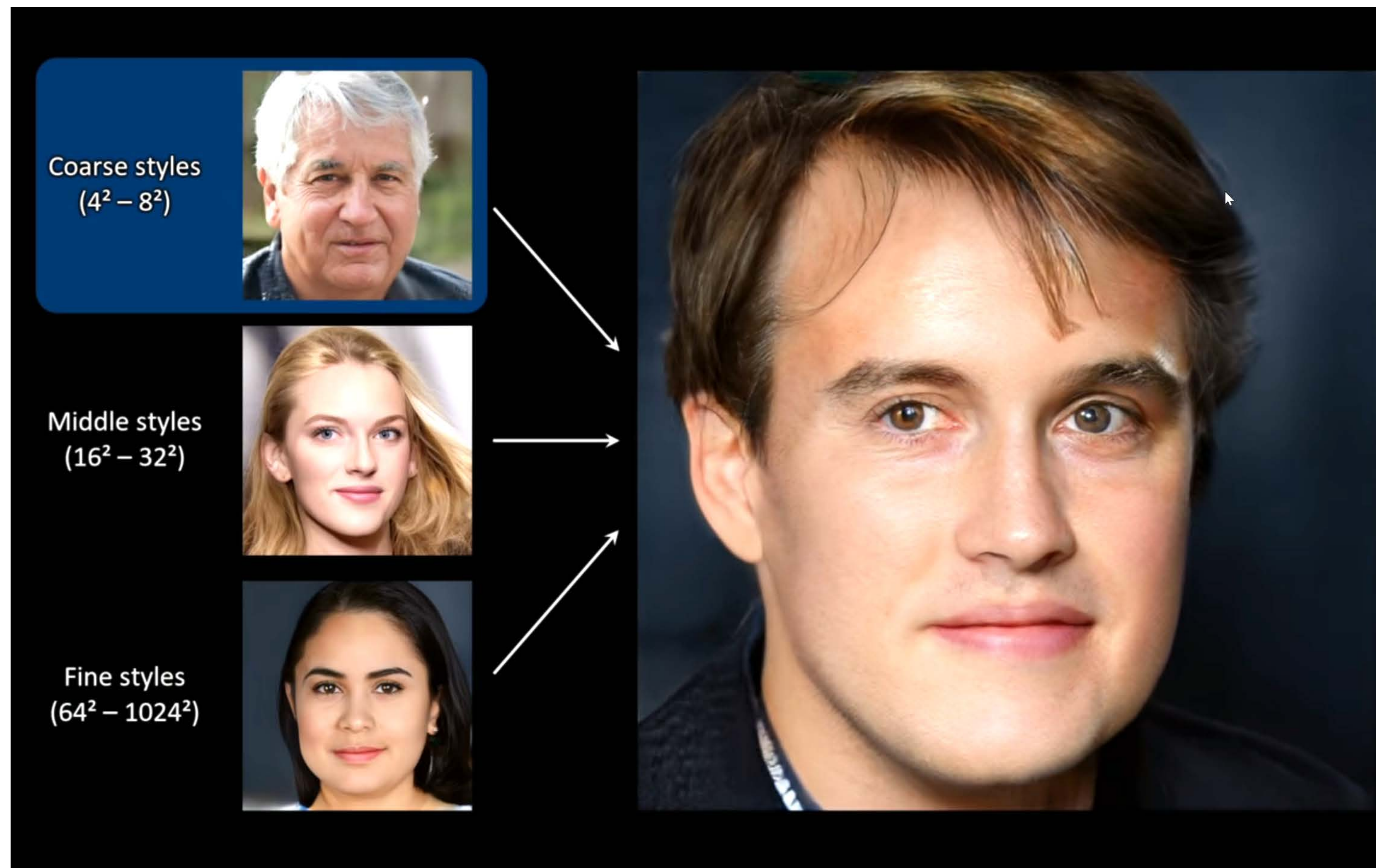
# StyleGAN

- 噪声B的影响





# StyleGAN





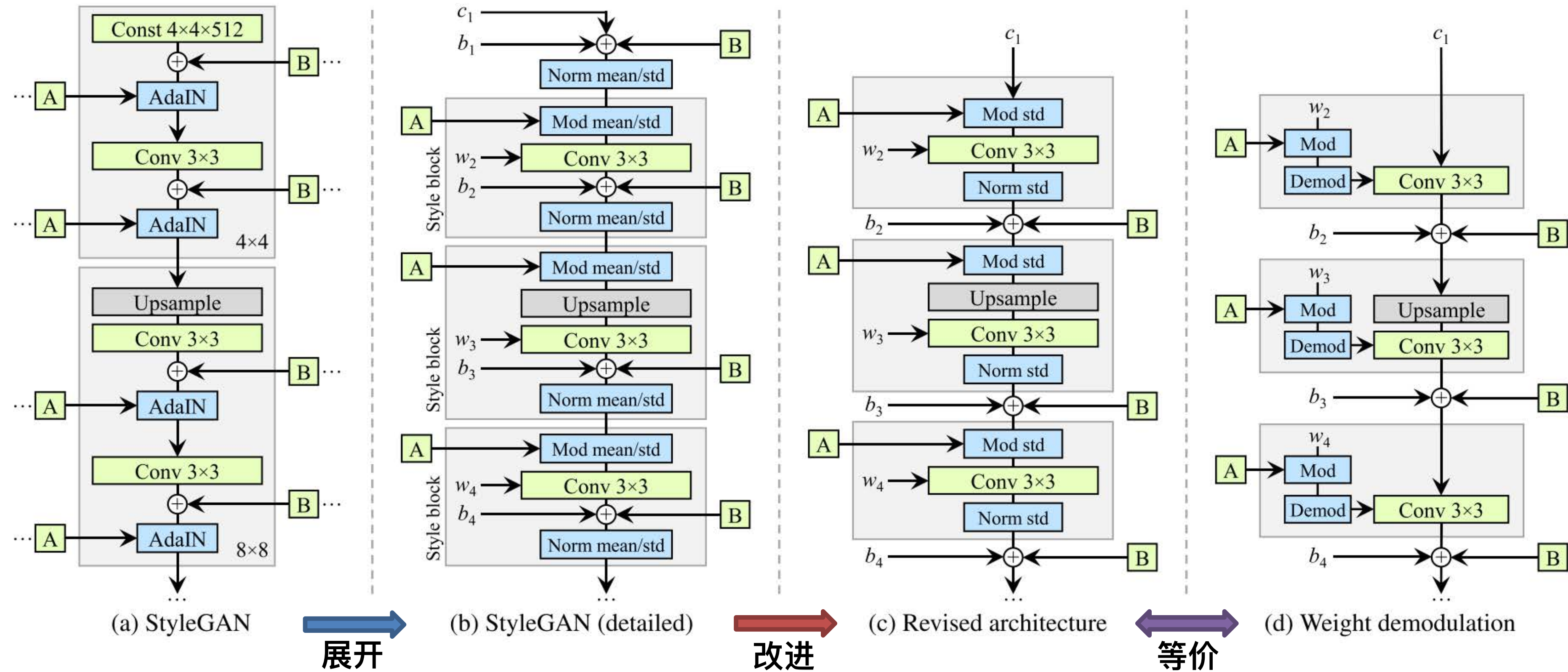
# StyleGAN V2

- StyleGAN2通过观察StyleGAN生成中的伪影瑕疵，改进了AdaIN操作与网络结构，进一步提升生成质量





# StyleGAN V2

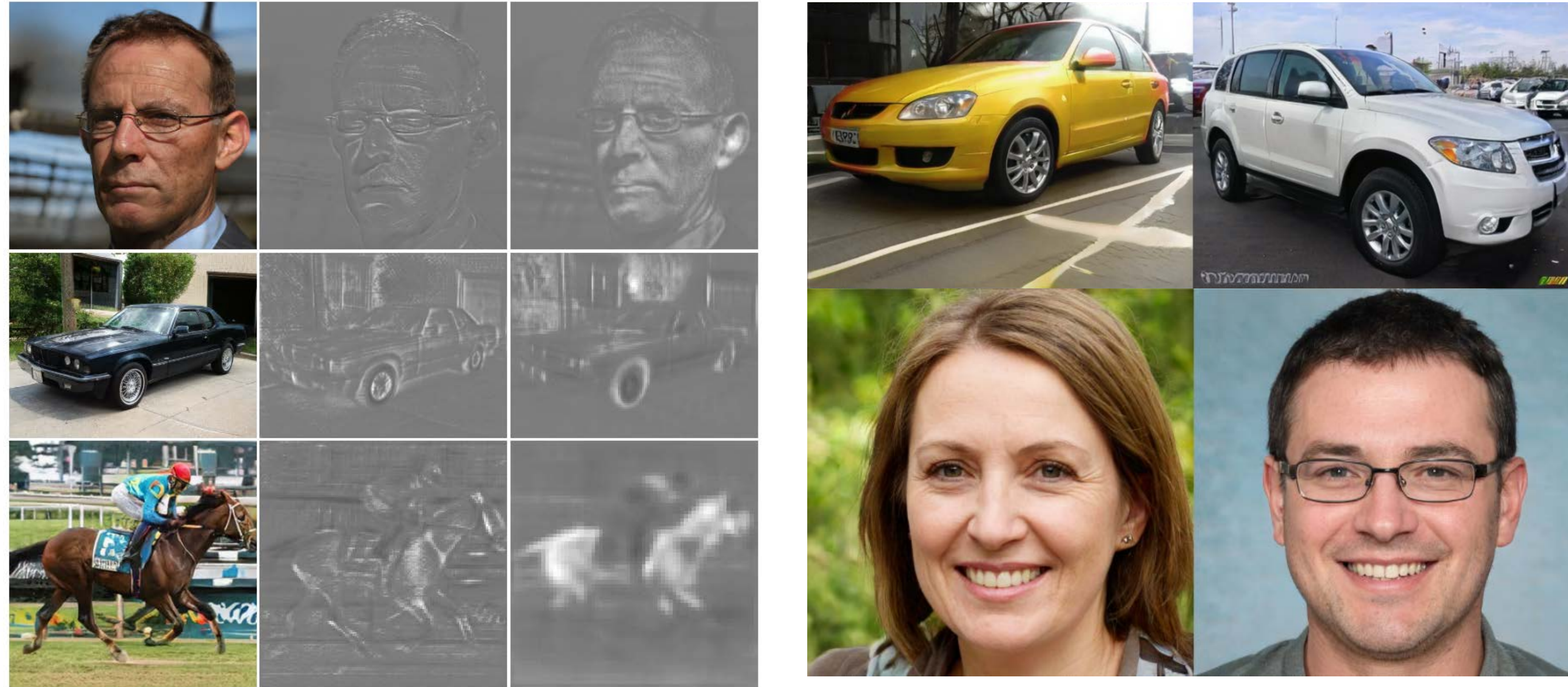


StyleGAN

StyleGAN2



# StyleGAN V2





# 参考文献

1. Goodfellow I J, Pouget-Abadie J, Mirza M, et al. "Generative Adversarial Nets." NIPS. 2014.
2. Radford, A., Metz, L., Chintala, S. "Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks." CoRR. 2016.
3. Arjovsky, M., Chintala, S., Bottou, L. "Wasserstein GAN." ICML. 2017.
4. Mirza, M., Osindero, S. "Conditional generative adversarial nets." arXiv preprint arXiv:1411.1784. 2014.
5. Isola P, Zhu J Y, Zhou T, et al. "Image-to-image translation with conditional adversarial networks." CVPR. 2017.
6. Zhu J Y, Park T, Isola P, et al. "Unpaired image-to-image translation using cycle-consistent adversarial networks." CVPR. 2017.



# 参考文献

7. Karras T, et al. “Progressive Growing of GANs for Improved Quality, Stability, and Variation.” ICLR. 2018.
8. Huang X, Belongie S. “Arbitrary Style Transfer in Real-time with Adaptive Instance Normalization.” ICCV. 2017.
9. Karras T, S. Laine, et al. “A Style-Based Generator Architecture for Generative Adversarial Networks.” CVPR. 2019.
10. Karras T, et al. “Analyzing and Improving the Image Quality of StyleGAN.” CVPR. 2020.