生成式对抗网络

Generative Adversarial Nets
（ GAN ）

生成式对抗网络

- 生成式对抗网络是非监督式学习的一种方法，通过让两个神经网络相互博弈的方式进行学习。该方法由Ian Goodfellow等人于2014年提出。

[PDF] Generative Adversarial Nets - NIPS Proceedings
https://papers.nips.cc/paper/5423-generative-adversarial-nets.pdf ▾ 翻译此页
作者：I Goodfellow - 2014 - 被引用次数：5577 - 相关文章
Like generative adversarial networks, variational autoencoders pair a differentiable generator network with a second neural network. Unlike generative adversarial networks, the sec- ond network in a VAE is a recognition model that performs approximate inference.

Content

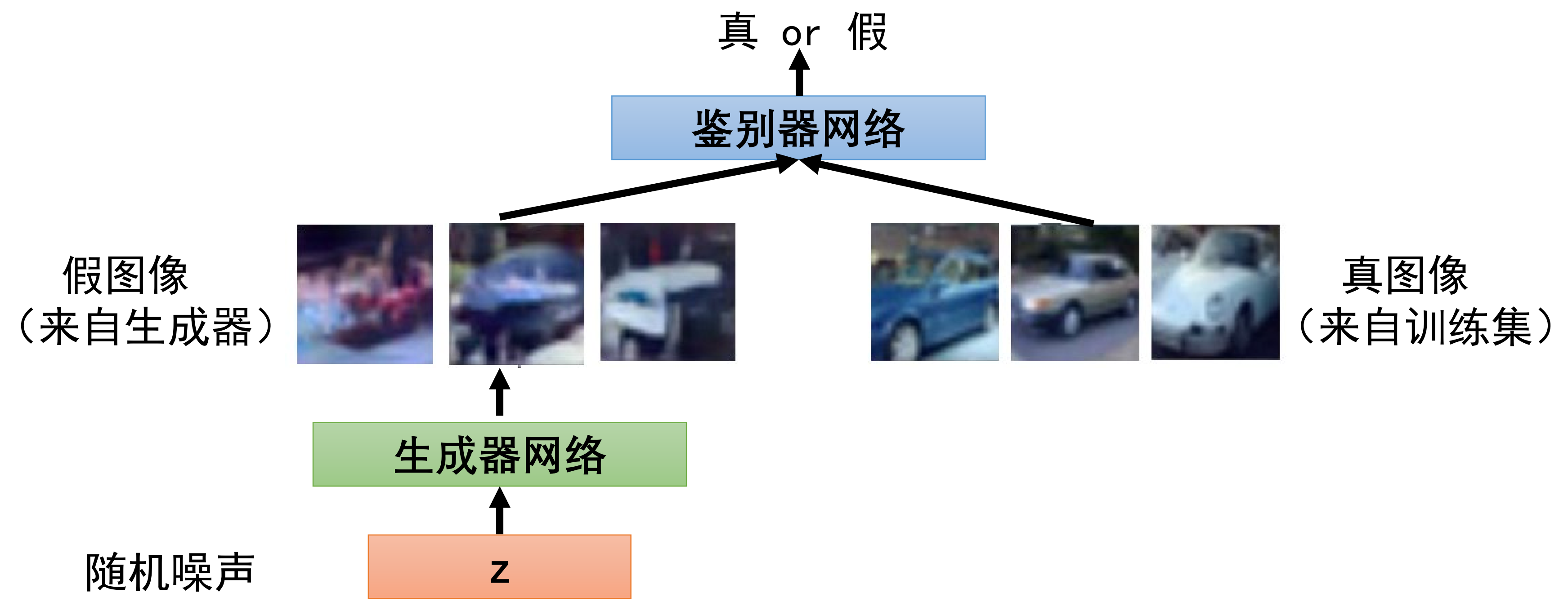
GAN网络的基本思想

GAN网络的算法

GAN网络的应用

训练GAN：双人游戏

生成器网络(Generator)：尝试通过生成逼真的图像来欺骗鉴别器
鉴别器/判别器(Discriminator)网络：尝试区分真实和虚假的图像



GAN-Generator

图像生成

$$\begin{bmatrix} 0.3 \\ -0.1 \\ \vdots \\ -0.7 \end{bmatrix} \begin{bmatrix} 0.1 \\ -0.1 \\ \vdots \\ 0.7 \end{bmatrix} \begin{bmatrix} -0.3 \\ 0.1 \\ \vdots \\ 0.9 \end{bmatrix} \rightarrow \boxed{\text{Generator 生成器}} \rightarrow$$

文本生成

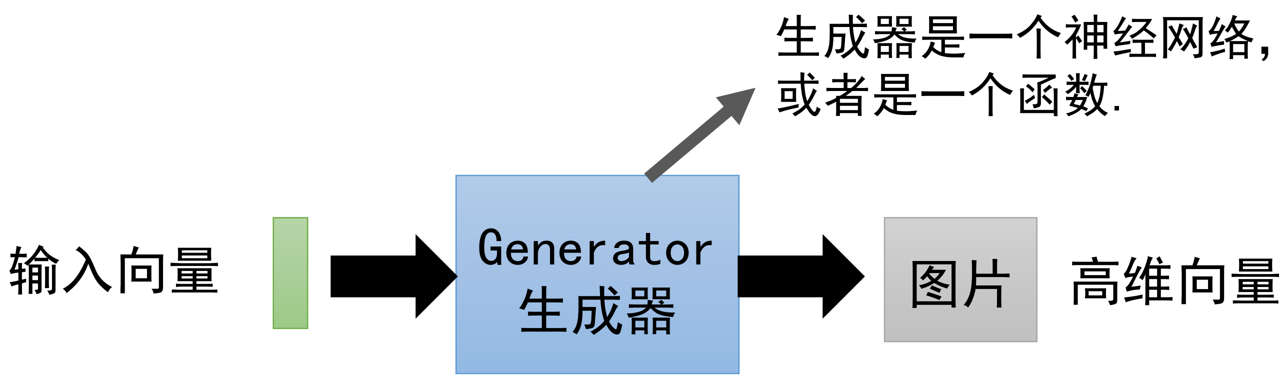
$$\begin{bmatrix} 0.3 \\ -0.1 \\ \vdots \\ -0.7 \end{bmatrix} \begin{bmatrix} 0.1 \\ -0.1 \\ \vdots \\ 0.2 \end{bmatrix} \begin{bmatrix} -0.3 \\ 0.1 \\ \vdots \\ 0.5 \end{bmatrix} \rightarrow \boxed{\text{Generator 生成器}} \rightarrow$$

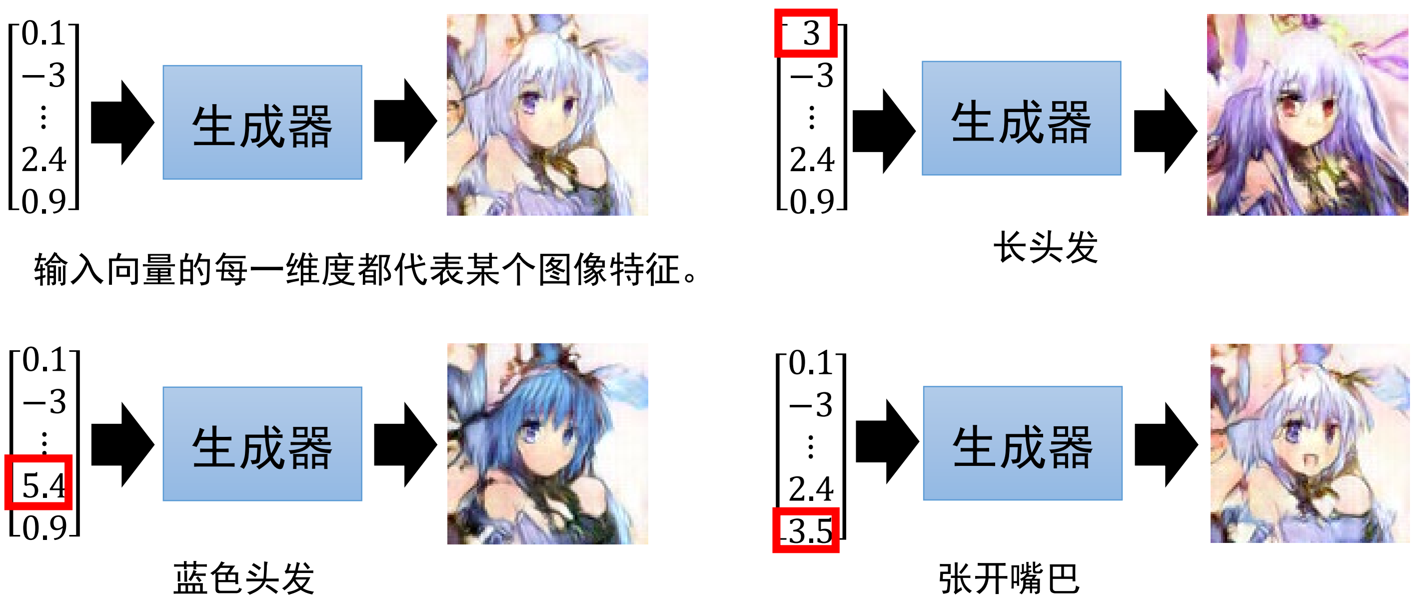
How are you?

Good morning.

Good afternoon.

GAN-Generator



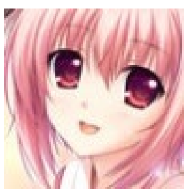


GAN-Discriminator

判别器是个神经网络，
或者是一个函数。

图像 → 判别器 → 标量

值越大代表越真实，
值越小代表图像越假。

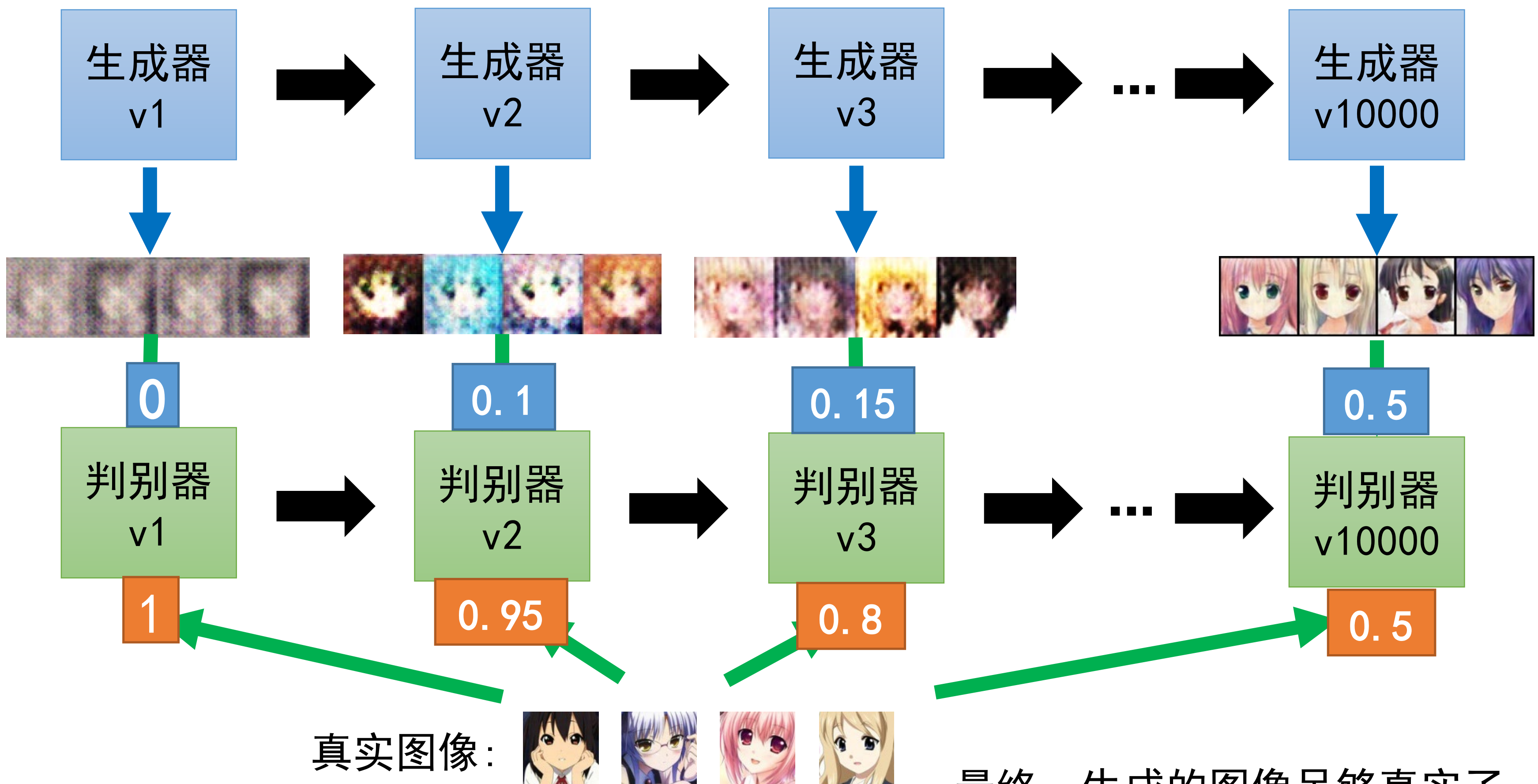
 → 判别器 → 1.0

 → 判别器 → 1.0

 → 判别器 → 0.1

 → 判别器 → 0.1

GAN-Adversarial



最终，生成的图像足够真实了，判别器只能输出1/2.

Content

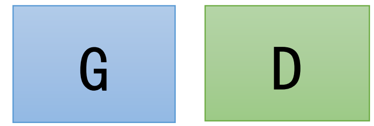
GAN网络的基本思想

GAN网络的算法

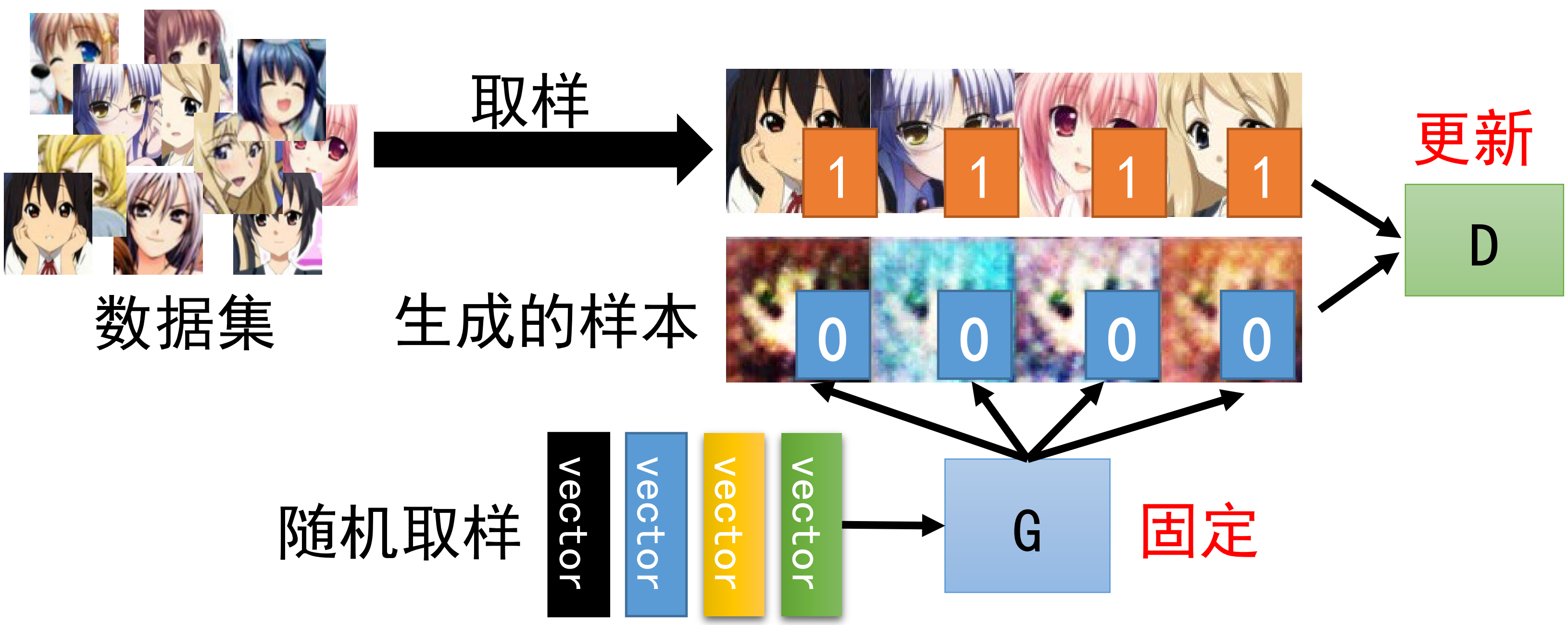
GAN网络的应用

算法

- 初始化生成器G和判别器D
- 在每个循环的迭代中：



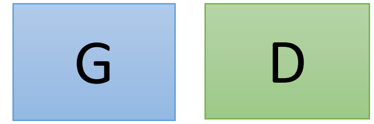
Step1：固定生成器G的参数，更新判别器D的参数



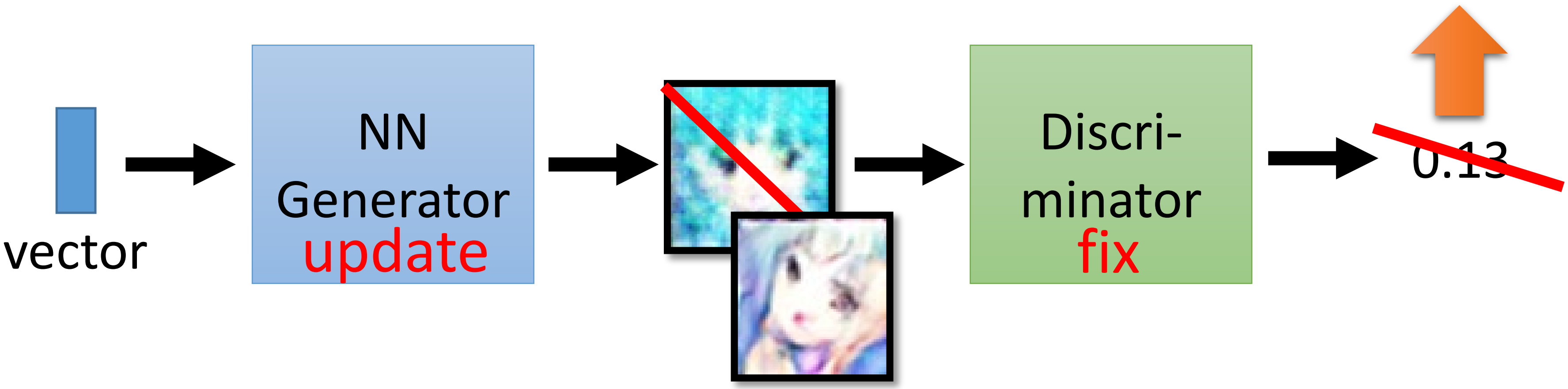
训练判别器赋予真实样本较高的输出值，赋予生成样本较低的输出值。

算法

- 初始化生成器G和判别器D
- 在每个循环的迭代中：



Step2：固定判别器D的参数，更新生成器G的参数：



训练生成器的生成样本能够骗过判别器，即希望判别器赋予生成的样本较高的输出值。

训练GAN：双人游戏

生成器网络：尝试通过生成逼真的图像来欺骗鉴别器
鉴别器网络：尝试区分真实和虚假的图像

结合Minimax游戏进行训练

对于真实图像，鉴别器输出的对数似然范围为（0,1）

Minimax目标函数：

$$\min_{\theta_g}\max_{\theta_d}\left[E_{x\sim p_{data}}\log \underbrace{D_{\theta_d}(x)}_{\substack{\text{对于真实数据}\\ \boldsymbol{x}\\ \text{鉴别器输出}}} + E_{z\sim p(z)}\log\left(1-\underbrace{D_{\theta_d}\left(G_{\theta_g}(z)\right)}_{\substack{\text{对于生成的假数据}\\ \boldsymbol{G(z)}\\ \text{鉴别器输出}}}\right)\right]$$

- 鉴别器(θ_d)想要最大化目标，因此$D(x)$接近1(真)并且$D\left(G(z)\right)$接近0(假)；
- 生成器(θ_g)想要最小化目标，因此$D\left(G(z)\right)$接近1(鉴别器被骗，以为$G(z)$为真)；

训练GAN：双人游戏

Minimax目标函数：

$$\min_{\theta_g}\max_{\theta_d}\left[E_{x\sim p_{data}}\log D_{\theta_d}(x)+E_{z\sim p(z)}\log\left(1-D_{\theta_d}\left(G_{\theta_g}(z)\right)\right)\right]$$

轮流：

1. 鉴别器梯度上升

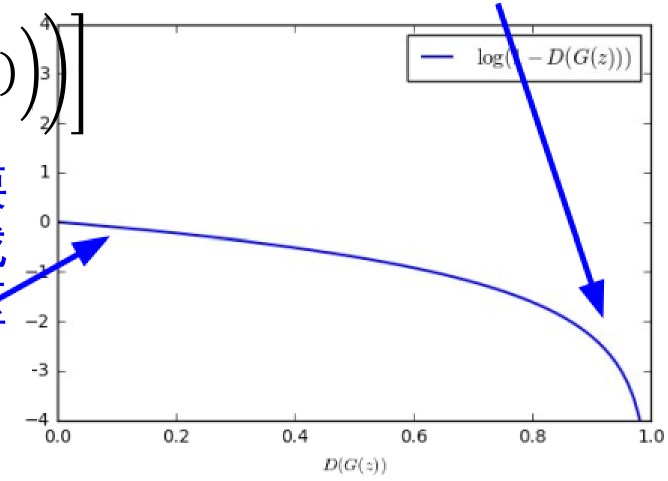
$$\max_{\theta_d}\left[E_{x\sim p_{data}}\log D_{\theta_d}(x)+E_{z\sim p(z)}\log\left(1-D_{\theta_d}\left(G_{\theta_g}(z)\right)\right)\right]$$

2. 生成器梯度下降

$$\min_{\theta_g}E_{z\sim p(z)}\log\left(1-D_{\theta_d}\left(G_{\theta_g}(z)\right)\right)$$

梯度信号由样本已经很真实的区域主导。

当样本太假时，需要从中学习已提升生成器的性能。但是此区域的梯度比较平缓。



在实践中，优化该生成器目标效果并不是很好。

训练GAN：双人游戏

备注：联合培训两个网络具有挑战性，可能不稳定。选择具有更好的损失呈现的目标有助于训练，这是一个热门的研究领域。

Minimax目标函数：

$$\min_{\theta_g}\max_{\theta_d}\left[E_{x\sim p_{data}}\log D_{\theta_d}(x)+E_{z\sim p(z)}\log\left(1-D_{\theta_d}\left(G_{\theta_g}(z)\right)\right)\right]$$

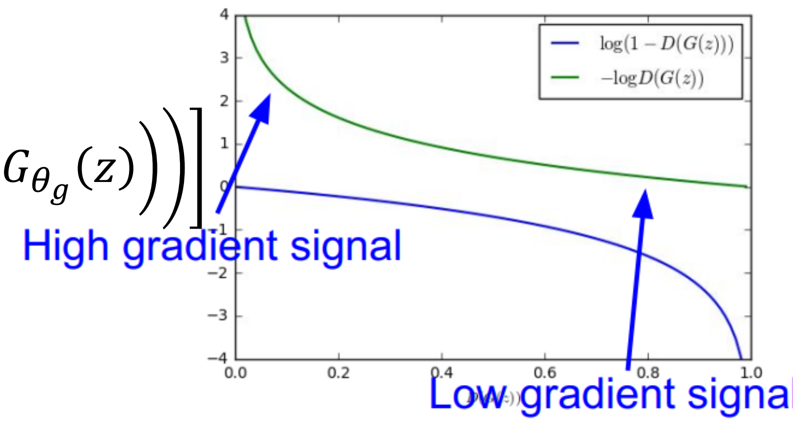
轮流：

1. 鉴别器梯度上升

$$\max_{\theta_d}\left[E_{x\sim p_{data}}\log D_{\theta_d}(x)+E_{z\sim p(z)}\log\left(1-D_{\theta_d}\left(G_{\theta_g}(z)\right)\right)\right]$$

2. 生成器梯度下降，不同的目标

$$\max_{\theta_g}E_{z\sim p(z)}\log\left(D_{\theta_d}\left(G_{\theta_g}(z)\right)\right)$$



现在最大化鉴别器判断错的可能性，来代替最小化鉴别器判断正确的可能性。
同样为了骗过鉴别器，但现在不好的样本对应更高的梯度信号=>效果更好！（实践标准）

算法

- 初始化判别器参数 θ_d 和生成器参数 θ_g
- 在每个训练的迭代中:

训练 D

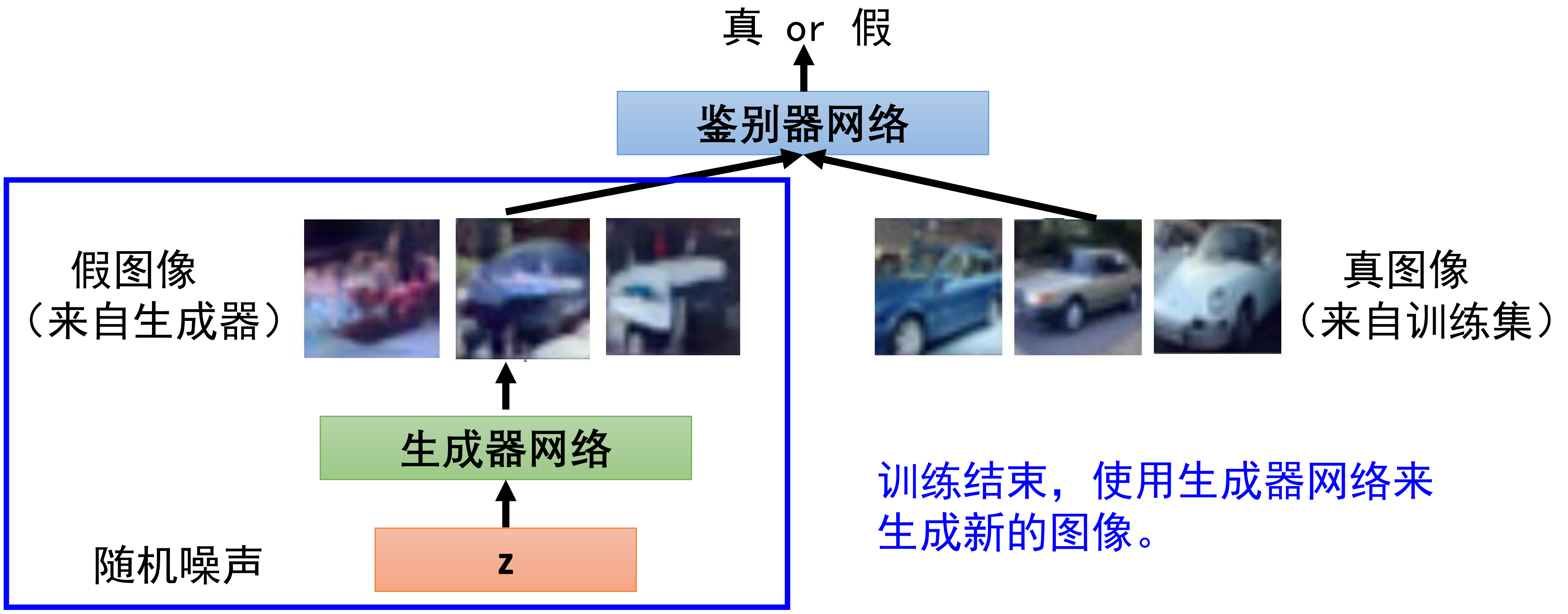
- 从数据集中取样m个样本$\{x^1, x^2, \dots, x^m\}$
- 从噪声先验分布中取样m个噪声样本$\{z^1, z^2, \dots, z^m\}$
- 得到生成的样本$\{\tilde{x}^1, \tilde{x}^2, \dots, \tilde{x}^m\}$, $\tilde{x}^i = G\left(z^i\right)$
- 更新判别器参数 θ_d 来最大化$\tilde{V}$
  - $\tilde{V} = \frac{1}{m}\sum_{i=1}^{m} log D_{\theta d}\left(x^i\right) + \frac{1}{m}\sum_{i=1}^{m} log\left(1 - D_{\theta d}\left(G_{\theta_g}\left(z^i\right)\right)\right)$
  - $\theta_d \leftarrow \theta_d + \eta \nabla \tilde{V}(\theta_d)$

训练 G

- 从噪声先验分布中取样m个噪声样本$\{z^1, z^2, \dots, z^m\}$
- 更新生成器参数 θ_g来最小化$\tilde{V}$
  - $\tilde{V} = \frac{1}{m}\sum_{i=1}^{m} log\left(D_{\theta d}\left(G_{\theta_g}\left(z^i\right)\right)\right)$
  - $\theta_g \leftarrow \theta_g + \eta \nabla \tilde{V}(\theta_g)$

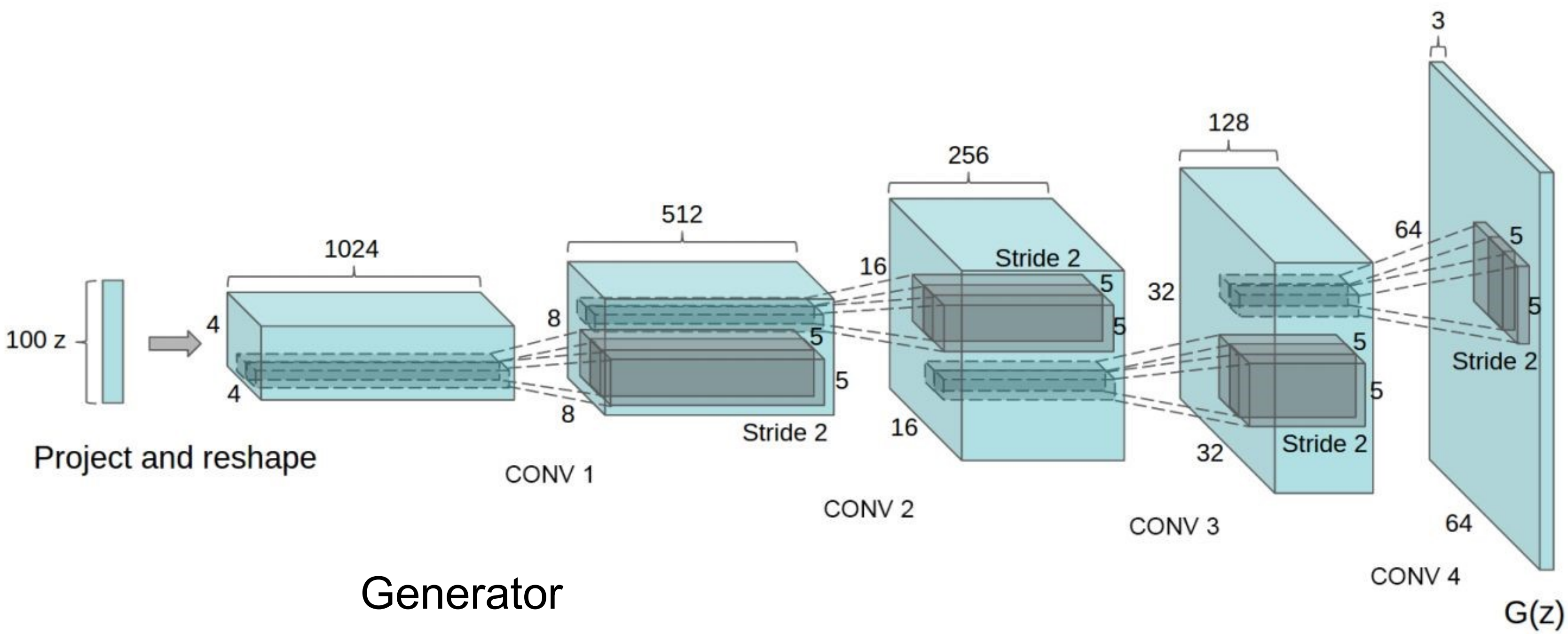
训练GAN：双人游戏

生成器网络：尝试通过生成逼真的图像来欺骗鉴别器
鉴别器网络：尝试区分真实和虚假的图像



训练结束，使用生成器网络来生成新的图像。

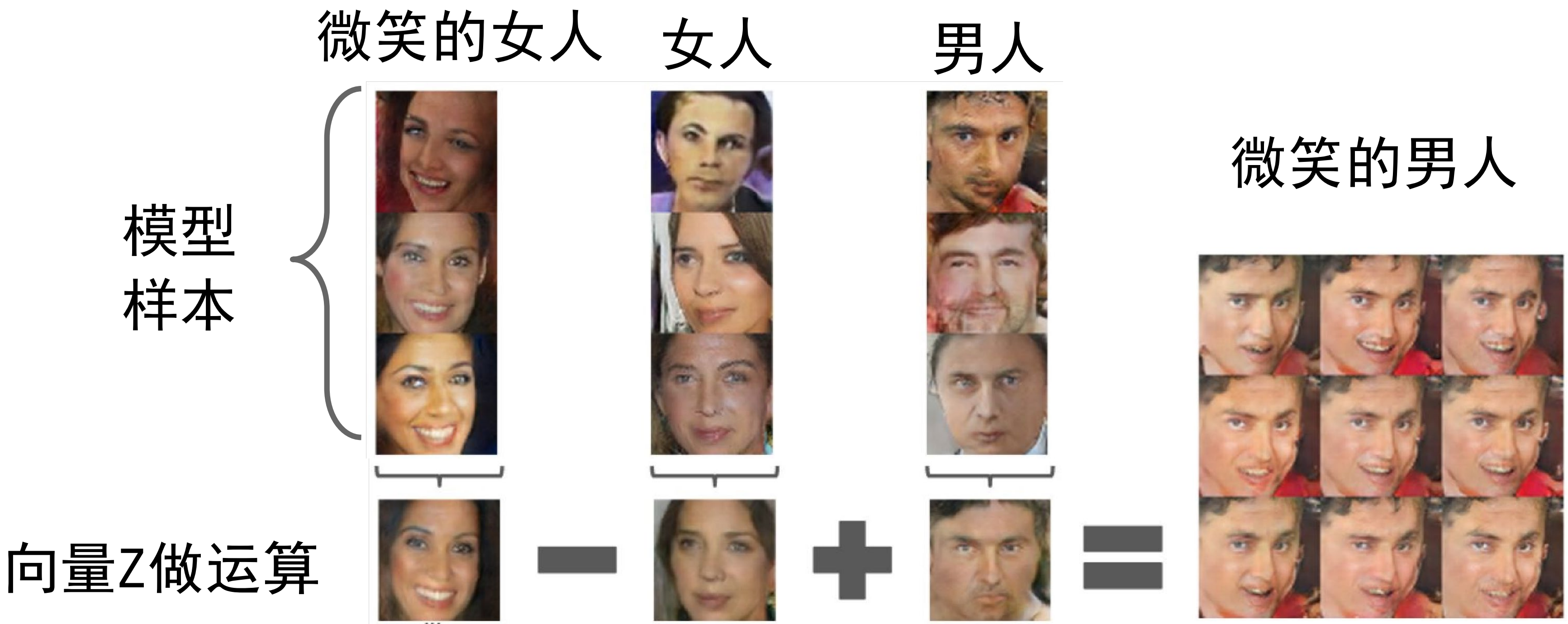
GAN：卷积结构


100 z
Project and reshape
1024
4
4
512
8
8
Stride 2
5
5
CONV 1
256
16
16
Stride 2
5
5
CONV 2
128
32
32
Stride 2
5
5
CONV 3
3
64
64
5
5
Stride 2
CONV 4
G(z)
Generator

算法

20000次迭代

GAN：向量解释



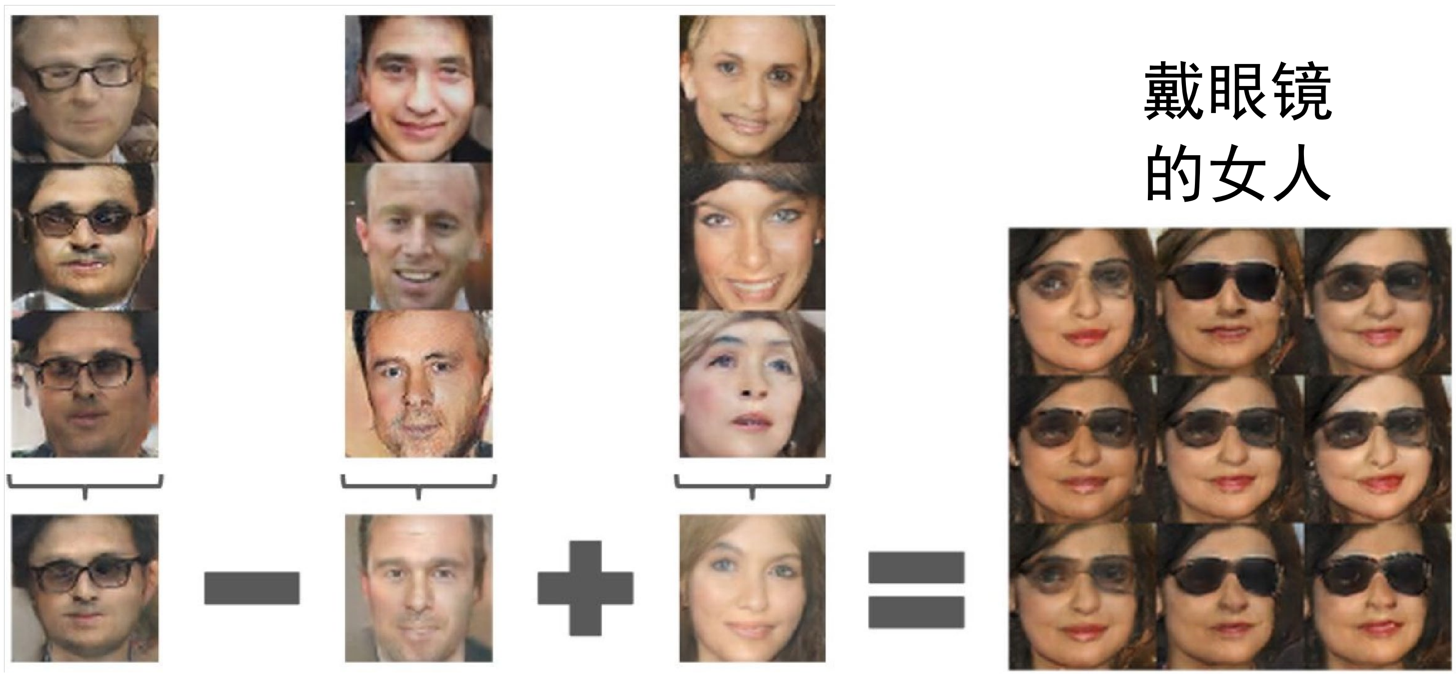
GAN：向量解释

戴眼镜的男人

不戴眼镜的男人

不戴眼镜的女人

戴眼镜的女人



GANs

不使用显式密度函数
采取博弈论的方法：通过双人游戏，从训练分配中生成

- **优点**:

  形式优雅，时下流行的样本生成方法
- **缺点**:

  训练不稳定

- **研究的主要领域**:

  更好的损失函数，更稳定的训练(Wasserstein GAN, LSGAN)

  条件GANs，各种应用

Content

GAN网络的基本思想

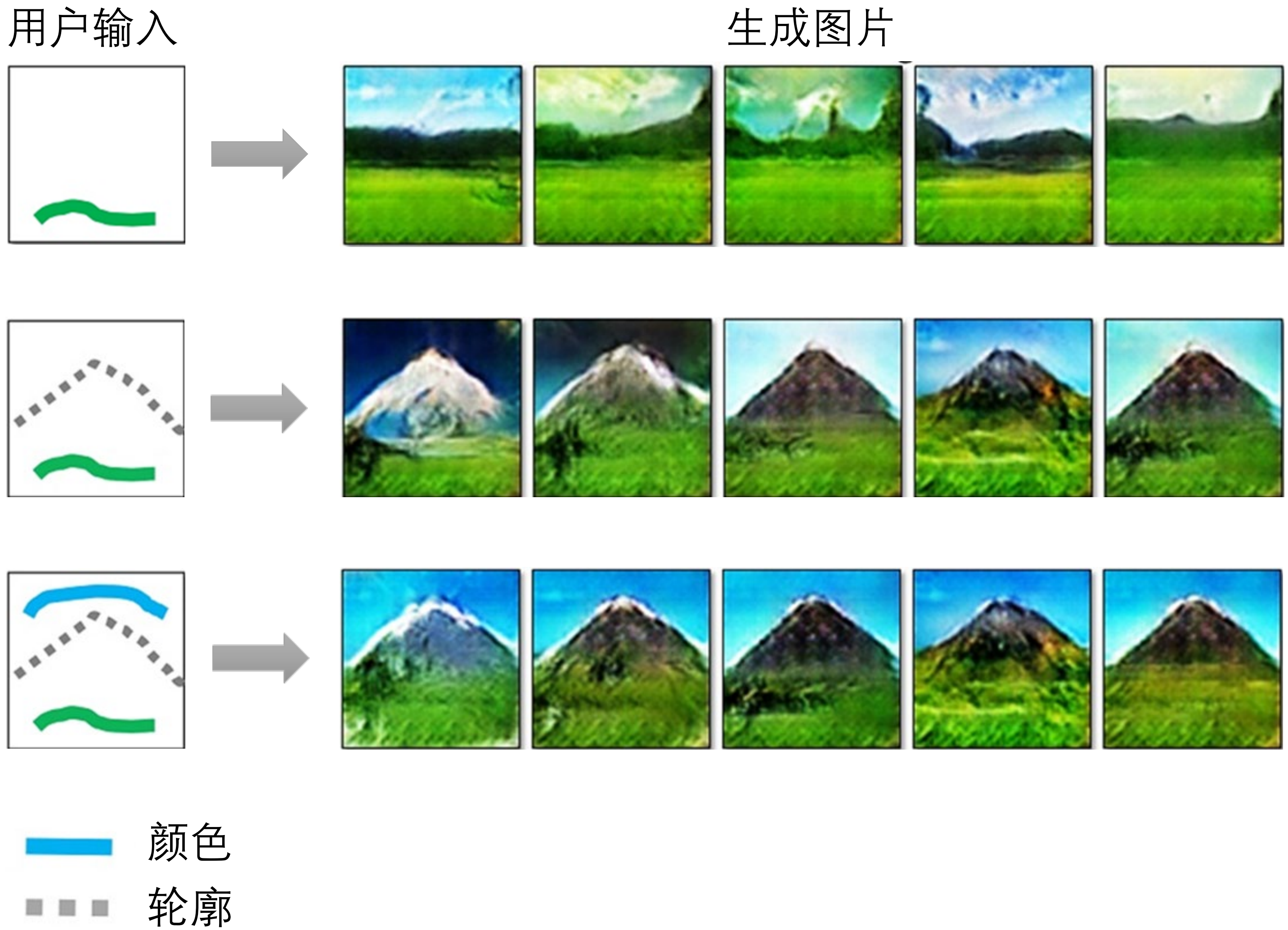
GAN网络的算法

GAN网络的应用

GAN网络的应用

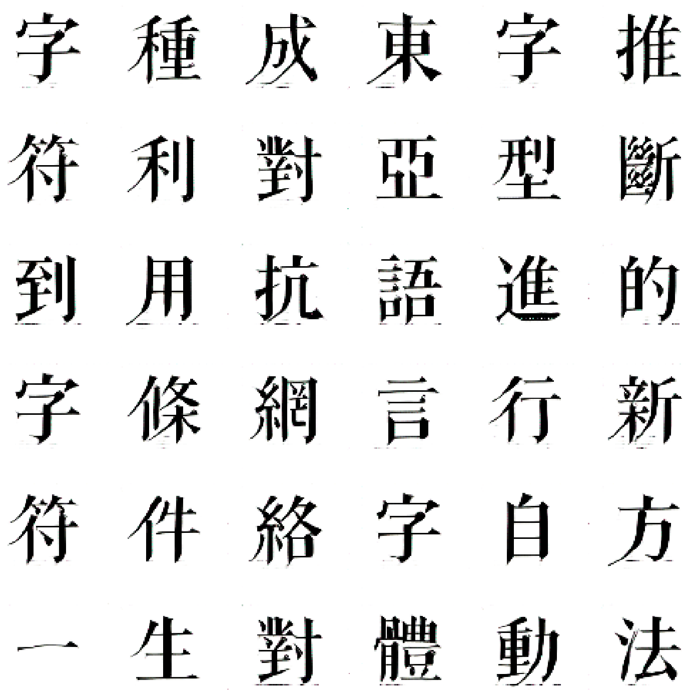
- 文本生成，语音、诗歌生成，文字描述生成图像
- 风格迁移、图像合成、图像修复
- 视频合成、目标跟踪、行人重识别

iGAN:灵魂画手



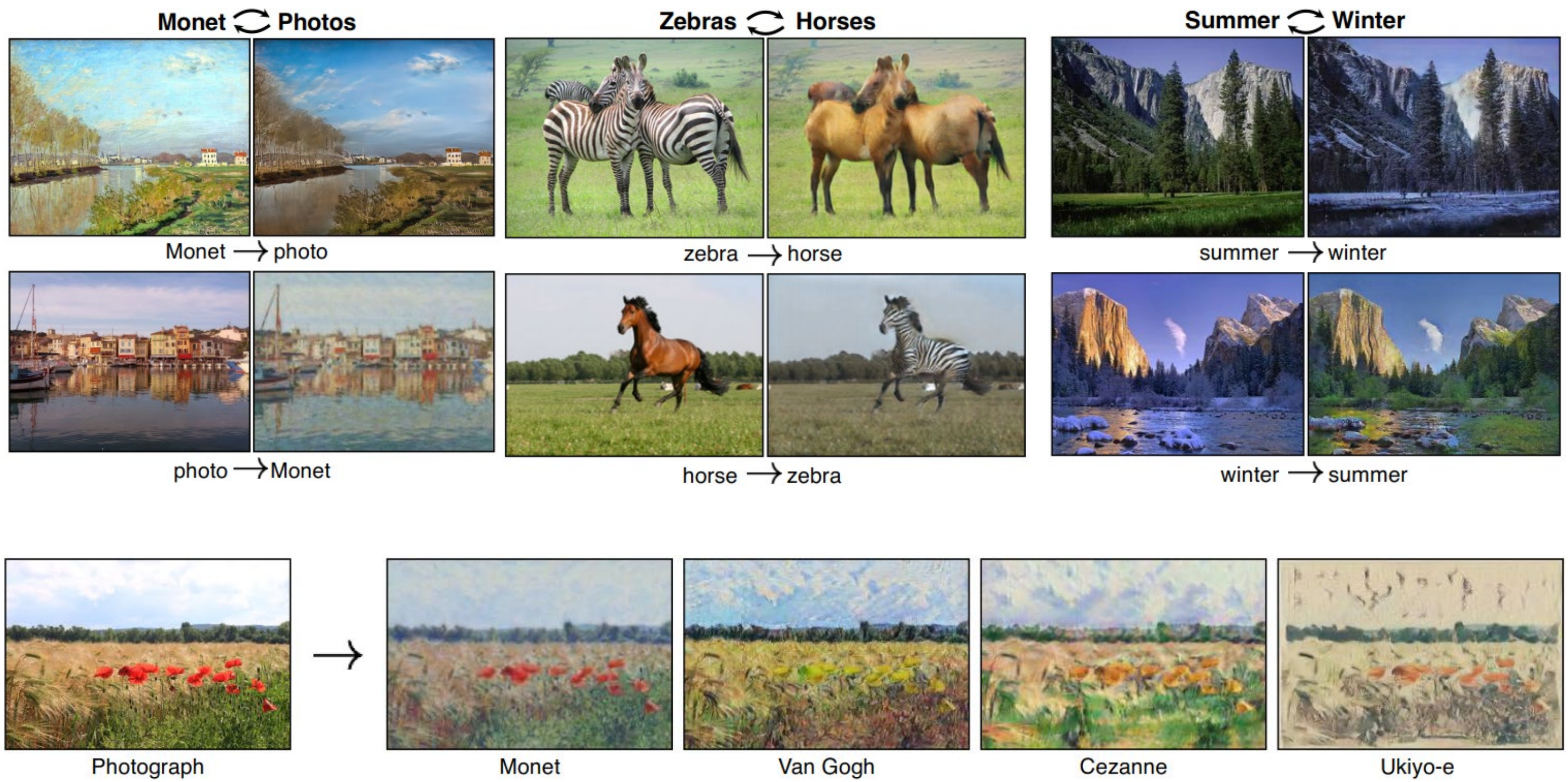
PairedCycleGAN:妆容迁移

Zi2zi-字体生成

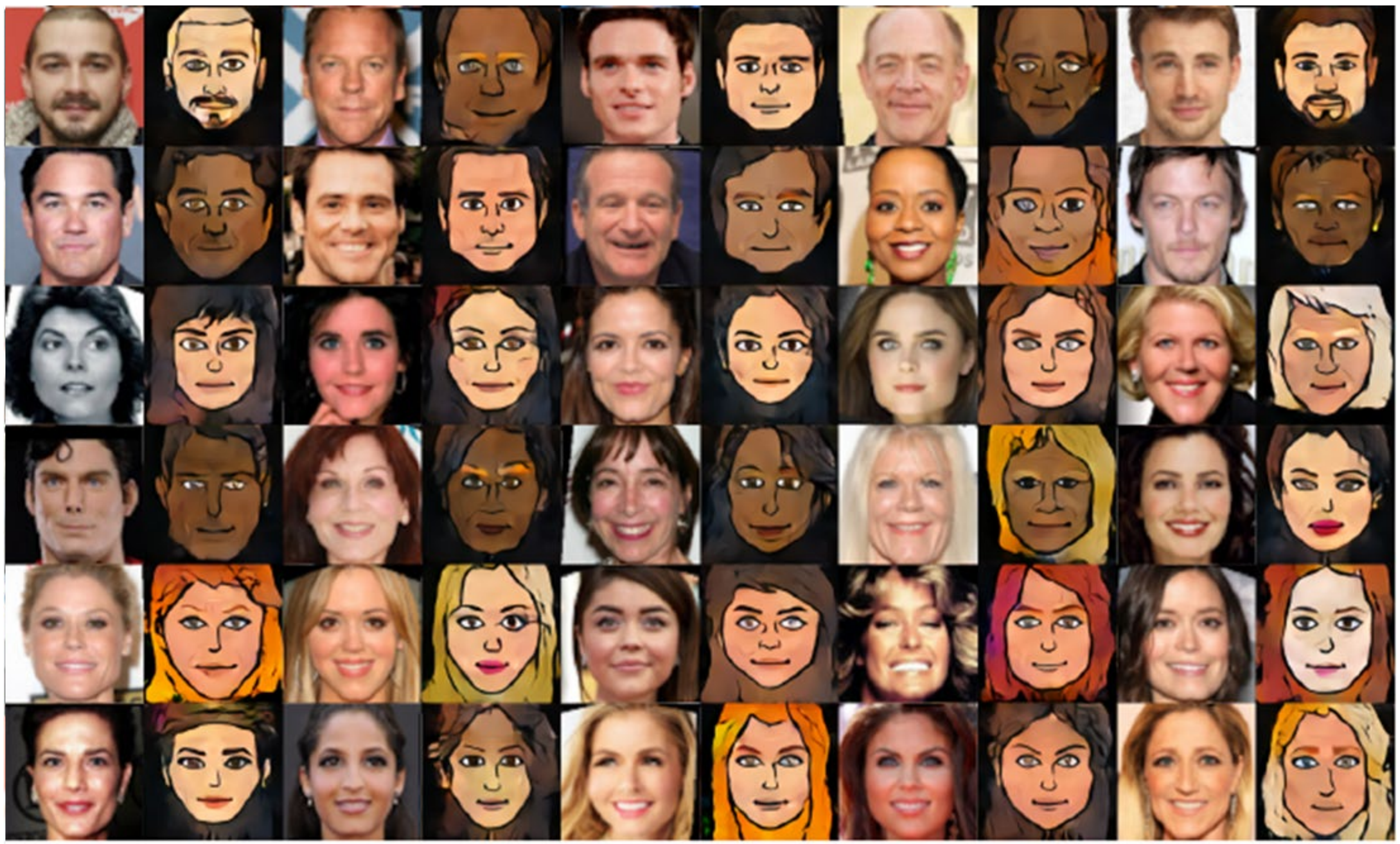

字 種 成 東 字 推
符 利 對 亞 型 斷
到 用 抗 語 進 的
字 條 網 言 行 新
符 件 絡 字 自 方
一 生 對 體 動 法


千 萬 孤 獨
山 徑 舟 釣
鳥 人 蓑 寒
飛 踪 笠 江
絕 滅 翁 雪

CycleGAN:风格迁移



DTN:emoji头像生成



Pix2pix:风格迁移

The GAN ZOO

- 3D-ED-GAN - Shape Inpainting using 3D Generative Adversarial Network and Recurrent Convolutional Networks
- 3D-GAN - Learning a Probabilistic Latent Space of Object Shapes via 3D Generative-Adversarial Modeling (github)
- 3D-IWGAN - Improved Adversarial Systems for 3D Object Generation and Reconstruction (github)
- 3D-PhysNet - 3D-PhysNet: Learning the Intuitive Physics of Non-Rigid Object Deformations
- 3D-RecGAN - 3D Object Reconstruction from a Single Depth View with Adversarial Learning (github)
- ABC-GAN - ABC-GAN: Adaptive Blur and Control for improved training stability of Generative Adversarial Networks (github)
- ABC-GAN - GANs for LIFE: Generative Adversarial Networks for Likelihood Free Inference
- AC-GAN - Conditional Image Synthesis With Auxiliary Classifier GANs
- acGAN - Face Aging With Conditional Generative Adversarial Networks
- ACGAN - Coverless Information Hiding Based on Generative adversarial networks
- acGAN - On-line Adaptative Curriculum Learning for GANs
- ACtuAL - ACtuAL: Actor-Critic Under Adversarial Learning
- AdaGAN - AdaGAN: Boosting Generative Models
- Adaptive GAN - Customizing an Adversarial Example Generator with Class-Conditional GANs
- AdvEntuRe - AdvEntuRe: Adversarial Training for Textual Entailment with Knowledge-Guided Examples
- AdvGAN - Generating adversarial examples with adversarial networks
- AE-GAN - AE-GAN: adversarial eliminating with GAN
- AE-OT - Latent Space Optimal Transport for Generative Models
- AEGAN - Learning Inverse Mapping by Autoencoder based Generative Adversarial Nets
- AF-DCGAN - AF-DCGAN: Amplitude Feature Deep Convolutional GAN for Fingerprint Construction in Indoor Localization System

- VEEGAN - VEEGAN: Reducing Mode Collapse in GANs using Implicit Variational Learning (github)
- VGAN - Generating Videos with Scene Dynamics (github)
- VGAN - Generative Adversarial Networks as Variational Training of Energy Based Models (github)
- VGAN - Text Generation Based on Generative Adversarial Nets with Latent Variable
- ViGAN - Image Generation and Editing with Variational Info Generative Adversarial Networks
- VIGAN - VIGAN: Missing View Imputation with Generative Adversarial Networks
- VoiceGAN - Voice Impersonation using Generative Adversarial Networks
- VOS-GAN - VOS-GAN: Adversarial Learning of Visual-Temporal Dynamics for Unsupervised Dense Prediction in Videos
- VRAL - Variance Regularizing Adversarial Learning
- WaterGAN - WaterGAN: Unsupervised Generative Network to Enable Real-time Color Correction of Monocular Underwater Images
- WaveGAN - Synthesizing Audio with Generative Adversarial Networks
- WaveletGLCA-GAN - Global and Local Consistent Wavelet-domain Age Synthesis
- weGAN - Generative Adversarial Nets for Multiple Text Corpora
- WGAN - Wasserstein GAN (github)
- WGAN-CLS - Text to Image Synthesis Using Generative Adversarial Networks
- WGAN-GP - Improved Training of Wasserstein GANs (github)
- WGAN-L1 - Subsampled Turbulence Removal Network
- WS-GAN - Weakly Supervised Generative Adversarial Networks for 3D Reconstruction
- X-GANs - X-GANs: Image Reconstruction Made Easy for Extreme Cases
- XGAN - XGAN: Unsupervised Image-to-Image Translation for many-to-many Mappings
- ZipNet-GAN - ZipNet-GAN: Inferring Fine-grained Mobile Traffic Patterns via a Generative Adversarial Neural Network
- α-GAN - Variational Approaches for Auto-Encoding Generative Adversarial Networks (github)
- β-GAN - Annealed Generative Adversarial Networks
- Δ-GAN - Triangle Generative Adversarial Networks

https://github.com/hindupuravinash/the-gan-zoo

How to Train a GAN? Tips and tricks to make GANs work

1. **Normalize the inputs**

2. **A modified loss function**

3. **Use a spherical Z**

4. **…….**

https://github.com/soumith/ganhacks

参考资料

- 1. Goodfellow, I., et al. Generative adversarial nets. in Advances in neural information processing systems. 2014.
- 2. 王坤峰, 苟超, 段艳杰, 林懿伦, 郑心湖, 王飞跃. 生成式对抗网络GAN的研究进展与展望. 自动化学报, 2017, 43(3): 321-332.
- 3. http://people.csail.mit.edu/junyanz/ 主页有：

  CycleGAN, pix2pix, pix2pixHD, video2video等论文
- 4. 其他论文（关键词搜索）：

  Zi2zi, iGAN, PairedCycleGAN, Domain Transfer Network(github),

  Exemplar GANs, DeblurGAN, AttentiveGAN；
- 5. Ian Goodfellow报告(NIPS2016)

  PDF: https://arxiv.org/pdf/1701.00160.pdf
- 6. 李宏毅线上课程：http://speech.ee.ntu.edu.tw/~tlkagk/courses_MLDS18.html