

# 控制与决策

Control and Decision

## 基于深度残差生成式对抗网络的样本生成方法

王星, 杜伟, 陈吉, 陈海涛

引用本文:

王星, 杜伟, 陈吉, 等. 基于深度残差生成式对抗网络的样本生成方法[J]. *控制与决策*, 2020, 35(8): 1887–1894.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2018.1700>

---

## 您可能感兴趣的其他文章

Articles you may be interested in

### 基于深度去噪核映射的长期预测模型

Deep denoising kernel mapping-based long-term prediction model

控制与决策. 2019, 34(5): 989–996 <https://doi.org/10.13195/j.kzyjc.2017.1451>

### 一种利用知识迁移的卷积神经网络训练策略

Convolutional neural network training strategy using knowledge transfer

控制与决策. 2019, 34(3): 511–518 <https://doi.org/10.13195/j.kzyjc.2017.1183>

### 面向原油总氢物性预测的数据扩增预处理方法

Data pretreatment approach for crude oil hydrogen properties prediction

控制与决策. 2018, 33(12): 2153–2160 <https://doi.org/10.13195/j.kzyjc.2017.0937>

### 基于LM算法的在线自适应RBF网结构优化算法

Online self-adaptive optimal algorithm for RBF network based on Levenberg–Marquardt algorithm

控制与决策. 2017, 32(7): 1247–1252 <https://doi.org/10.13195/j.kzyjc.2016.0887>

### DBN网络的深度确定方法

Calculation for depth of deep belief network

控制与决策. 2015(2): 256–260 <https://doi.org/10.13195/j.kzyjc.2013.1390>

# 基于深度残差生成式对抗网络的样本生成方法

王星<sup>†</sup>, 杜伟, 陈吉, 陈海涛

(辽宁工程技术大学 电子与信息工程学院, 辽宁 葫芦岛 125105)

**摘要:** 作为样本生成的重要方法之一,生成式对抗网络(GAN)可以根据任意给定数据集中的数据分布生成样本,但它在实际的训练过程中存在生成样本纹理模糊、训练过程不稳定以及模式坍塌等问题.针对以上问题,在深度卷积生成式对抗网络(DCGAN)的基础上,结合残差网络,设计一种基于深度残差生成式对抗网络的样本生成方法RGAN.该样本生成方法利用残差网络和卷积网络分别构建生成模型和判别模型,并结合正负样本融合训练的学习优化策略进行优化训练.其中:深度残差网络可以恢复出丰富的图像纹理;正负样本融合训练的方式可以增加对抗网络的鲁棒性,有效缓解对抗网络训练不稳定和模式坍塌现象的发生.在102 Category Flower Dataset数据集上设计多个仿真实验,实验结果表明RGAN能有效提高生成样本的质量.

**关键词:** 生成式对抗网络; 残差网络; 深度学习; 对抗训练; RGAN; FID

**中图分类号:** TP183

**文献标志码:** A

## Sample generation based on residual generative adversarial network

WANG Xing<sup>†</sup>, DU Wei, CHEN Ji, CHEN Hai-tao

(School of Electronic and Information Engineering, Liaoning Technical University, Huludao 125105, China)

**Abstract:** As one of the important methods of sample generation, the generative adversarial network (GAN) can generate samples based on the data distribution in any given data set, but it has generated sample texture blur, unstable training process and mode collapse in the actual training process. In order to solve the above problems, this paper proposes a residual generative adversarial network, (RGAN) based on the deep convolutional generative adversarial network (DCGAN) and the residual network. The sample generation method uses the residual network and the convolution network to construct the generator and the discriminator respectively, and combines the learning optimization strategy of the positive and negative sample fusion training to optimize the training process. Among them, the depth residual network can recover the rich image texture and, the positive and negative sample fusion training can increase the robustness against the network and effectively alleviate the instability of network training and the collapse of the model. This paper designs several simulation experiments on the 102 category flower dataset. The experimental results show that the RGAN can effectively improve the quality of generated samples.

**Keywords:** generative adversarial network; residual network; deep learning; adversarial training; residual generative adversarial network; FID

## 0 引言

生成式对抗网络(generative adversarial networks, GAN)是一种生成式模型<sup>[1]</sup>.与传统的生成式模型不同,GAN网络结构除了包含生成模型外,还拥有判别模型.其中,生成模型根据真实数据的分布规律生成新的样本数据,而判别模型用于鉴定数据是来自真实数据还是来自新生成的样本数据.GAN的训练是一个极小极大博弈的过程,训练的最终目标是让生成模型完全捕获真实数据的样本分布规律,从而生

成样本<sup>[1-4]</sup>.因为GAN不需要对生成分布进行显示表达,避免了传统生成模型的复杂推断运算,故与传统生成模型相比优势显著.GAN在图像和文本的生成任务中取得了较大的成功,包括图像到图像的转换,图像超分辨率以及文本到图像的合成<sup>[3-5]</sup>.

虽然GAN解决了生成模型普遍存在的一些问题,但它仍然有许多亟待解决的问题,例如:生成样本质量较低、模式坍塌<sup>[3-6]</sup>、超参数敏感<sup>[7]</sup>等,这些问题导致GAN难以生成较高质量的样本<sup>[5]</sup>.目前,人

收稿日期: 2018-12-12; 修回日期: 2019-04-08.

基金项目: 国家自然科学基金项目(61402212); 辽宁省高等学校杰出青年学者成长计划项目(LJQ2015045); 中国博士后基金面上项目(2016M591452); 辽宁省自然科学基金面上项目(2015020098).

责任编辑: 薛建儒.

<sup>†</sup>通讯作者. E-mail: xingwang\_lntu@163.com.

们提出了许多改进的方法. Mirza等<sup>[8]</sup>提出了条件生成式对抗网络(CGAN),它在传统生成模型和判别模型的基础上增加限制条件,约束了GAN的自由度,从而提高了生成样本的质量.为了找到纳什均衡状态,Salimans等<sup>[9]</sup>提出Improved-GAN,其使用的启发式动机是以鼓励融合的技巧代替传统生成对抗网络中梯度下降算法,并取得了不错的效果. Radford等<sup>[10]</sup>在GAN的基础上结合卷积神经网络提出了深度卷积生成式对抗网络(DCGAN),该方法利用卷积神经网络强大的特征提取能力,生成质量相对较高的样本,但DCGAN依然存在GAN具有的模式坍塌和超参数敏感问题.为了缓解以上问题,Mao等<sup>[7]</sup>将传统交叉熵损失函数改进为最小二乘法损失函数,提出了最小二乘法生成式对抗网络(LSGAN),该方法虽然有效缓解了模式坍塌和超参数敏感现象,但在实际的训练过程中容易出现训练过程不稳定性问题.

针对GAN及其衍生算法的诸多问题,本文结合正负样本融合训练<sup>[11-12]</sup>的模式以及两个时间尺度更新准则(two time-scale update rule, TTUR)<sup>[6,13]</sup>,设计并实现了基于深度残差生成式对抗网络的样本生成算法RGAN(residual generative adversarial network).该算法在训练过程中通过正负样本融合训练的模式增加了噪声,提高了对抗网络的鲁棒性,从而提高训练过程的稳定性,并结合TTUR缓解了模式坍塌现象的发生.此外,网络中使用的深度残差网络不但可以恢复出如真实照片般丰富的图像纹理,提高了样本生成质量,还能有效缓解深层网络容易出现的退化现象.

### 1 生成式对抗网络原理

生成式对抗网络由生成模型和判别模型两部分构成.生成模型主要根据输入的随机噪声生成服从真实数据概率分布的新样本,而判别模型则是判断出输入样本是真实样本还是由生成模型生成的新样本.生成模型和判别模型是竞争关系,均不断优化自身,当判别模型无法正确区分样本来源时,可以认为生成模型捕获到了真实样本的概率分布,具体流程如图1所示.

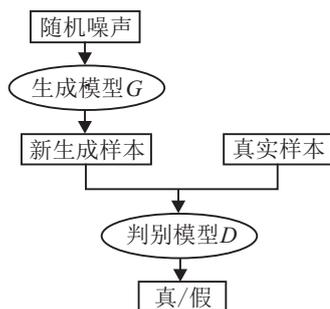


图1 GAN结构

生成模型和判别模型的具体博弈过程如下所示:

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim P_{\text{data}}(x)} [\log(D(x))] + \mathbb{E}_{z \sim P_z(z)} [\log(1 - D(G(z)))]. \quad (1)$$

其中:  $\mathbb{E}$ 表示数学期望,  $x$ 表示采样于真实数据分布  $P_{\text{data}}(x)$  的样本,  $z$ 表示采样于先验分布  $P_z(z)$  的噪声,  $D(x)$ 表示真实样本通过判别模型判别出其为真实样本的概率,  $D(G(z))$ 表示判别模型鉴定由生成模型通过随机噪声生成的样本的真实概率,  $P_g$ 是先验分布  $P_z(z)$  根据真实分布希望生成的理想分布,  $V(D, G)$ 表示判别模型和生成模型的值函数.理想状态下希望判别模型的值函数取得最大值,生成模型的值函数取得最小值.训练判别模型是为了最大化地判断出数据是来自真实样本,还是来自生成模型生成的新样本,训练生成模型是为了最小化  $\log(1 - D(G(z)))$  的值<sup>[1]</sup>.

当生成模型固定时,有

$$V(D, G) = \int_x P_{\text{data}}(x) \log(D(x)) + P_g(x) \log(1 - D(x)) dx. \quad (2)$$

令  $P_{\text{data}}(x) = a, P_g(x) = b$ ,将式(2)转化为

$$y \rightarrow a \log(y) + b \log(1 - y), \quad (3)$$

$$D_G^*(x) = \frac{P_{\text{data}}(x)}{P_{\text{data}}(x) + P_g(x)}. \quad (4)$$

其中对于任意的  $(a+b) \in \mathbb{R}^2 \setminus \{0, 0\}$ ,为了让方程(3)在  $[0, 1]$  区间上取得最大值,需对其进行求导运算.经运算,在点  $a/(a+b)$  处导数为零,故方程(3)取得最大值,必须满足方程(4),其中  $D_G^*(x)$ 表示当生成模型固定时,判断模型取得的最优值.

$$C(G) = \max_D V(D, G) = \mathbb{E}_{x \sim P_{\text{data}}(x)} [\log(D_G^*(x))] + \mathbb{E}_{z \sim P_z(z)} [\log(1 - D_G^*(x))] = \mathbb{E}_{x \sim P_{\text{data}}(x)} \left[ \log \frac{P_{\text{data}}(x)}{P_{\text{data}}(x) + P_g(x)} \right] + \mathbb{E}_{z \sim P_z(z)} \left[ \log \frac{P_{\text{data}}(x)}{P_{\text{data}}(x) + P_g(x)} \right]. \quad (5)$$

当判别模型值函数取得最大值时,为了使对抗网络达到纳什均衡,生成模型值函数需要在判别模型值函数取得最大值的基础上取得最小值,如下式所示:

$$C(G) = \mathbb{E}_{x \sim P_{\text{data}}(x)} [\log(D_G^*(x))] + \mathbb{E}_{z \sim P_z(z)} [\log(1 - D_G^*(x))] = -\log 4, \quad (6)$$

为了让真实分布与生成分布相等(即  $P_{data} = P_g$ ), 此时, 式(6)取得全局最小值.

联立式(5)和(6)可得

$$C(G) = -\log 4 + 2\text{JSD}(P_{data} \| P_g). \quad (7)$$

由于两个分布之间詹森香农散度(JSD)总是非负的, 只有当两个分布相等时(即同分布), JSD为0, 式(7)取得全局最小值, 即  $C(G) = -\log 4$ .

在实际训练的过程中, 为了让判别模型和生成模型更好地学习参数还需建立映射函数  $D(x, w)$  和  $G(z, \theta)$ . 其中  $D(x, w)$  为判别模型和真实数据所建立

的映射函数,  $w$  为判别模型所学参数. 同时, 为了使该网络达到纳什均衡, 生成模型需要在判别模型取得最大值的情况下取得最小值, 因此还需建立生成模型和随机噪声的映射函数  $G(z, \theta)$ , 其中  $\theta$  为生成模型所学参数.

## 2 深度残差生成式对抗网络样本生成方法

### 2.1 RGAN 模型结构

利用残差网络<sup>[14]</sup>思想构建本文所提出的RGAN模型.

1) 构建生成模型(如图2所示).

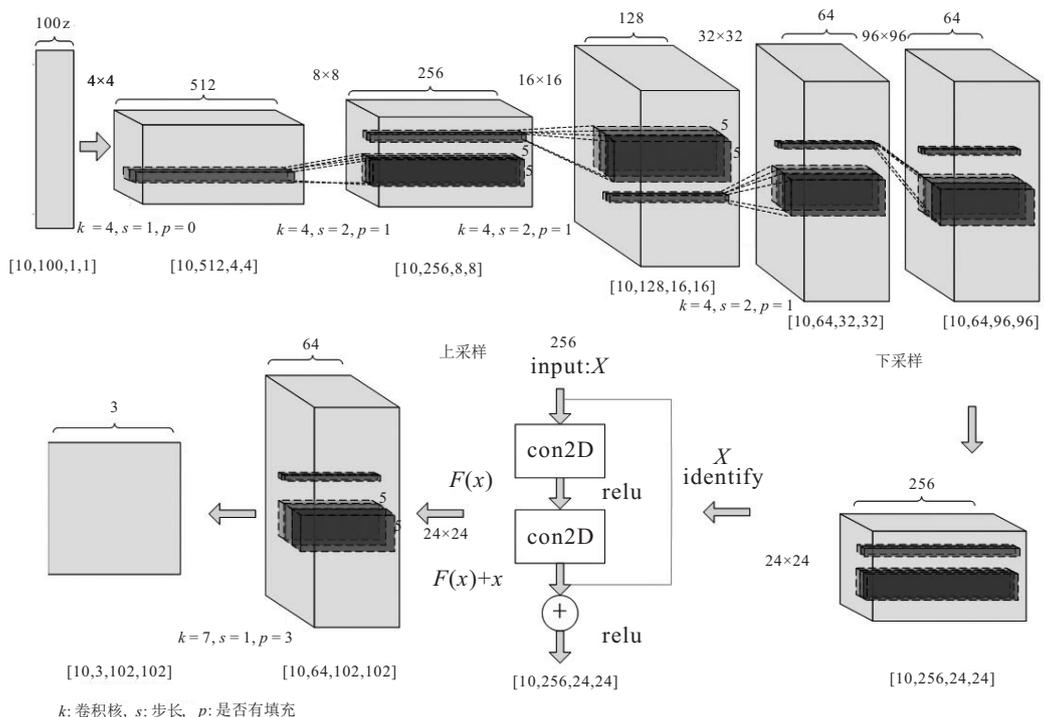


图2 生成模型示意图

step 1: 转置卷积运算. 首先, 给定3个超参数, 即批处理样本个数、噪声维度、初始噪声样本的像素, 并由这3个参数组成生成模型所需的四维张量, 默认四维张量(1, 100, 1, 1), 其中1表示批处理样本个数, 100表示输入噪声的维度,  $1 \times 1$ 表示初始噪声样本的像素; 其次, 经过1次卷积核为  $4 \times 4$ 、步长为1的二维转置卷积运算后将维度转换为(1, 512, 4, 4)的四维张量; 然后, 经过3次卷积核为  $4 \times 4$ 、步长为2的二维转置卷积运算后将维度转换为(1, 64, 32, 32)的四维张量; 最后, 经过1次卷积核为  $5 \times 5$ 、步长为3的二维转置卷积运算后输出张量为(1, 64, 96, 96).

step 2: 下采样运算<sup>[15]</sup>. 经过两次卷积核为  $4 \times 4$ 、步长为2的二维卷积运算后, 将维度转换为(1, 256, 24, 24)的四维张量. 使用下采样运算<sup>[15]</sup>的首要目的是减少参数, 加快运算速度. 此外, 采样后的感受也更

适合网络提取特征.

step 3: 残差网络运算. 通过由6个残差块组成的残差网络后输出的维度保持不变.

step 4: 上采样运算. 分别经过两次卷积核为  $4 \times 4$ 、步长为2和3次卷积核为  $4 \times 4$ 、步长为1的二维转置卷积运算后, 将维度转换为(1, 64, 102, 102)的四维张量. 最后经过一个卷积核为  $7 \times 7$ 、步长为1的二维卷积运算输出一张  $102 \times 102$  像素的三通道图片.

2) 构建判别模型(如图3所示). 与生成模型网络结构的构建过程正好相反, 深度残差生成式对抗网络的判别模型网络结构是将一个四维张量转化为一个标量. 首先, 输入一张样本数据, 经数字化后转换为(1, 3, 102, 102)的四维张量; 其次, 经过一个  $1 \times 1$  卷积运算后, 将维度转换为(1, 64, 102, 102)的四维张量; 再次, 经过4个卷积核为  $4 \times 4$ 、步长为2的卷积运

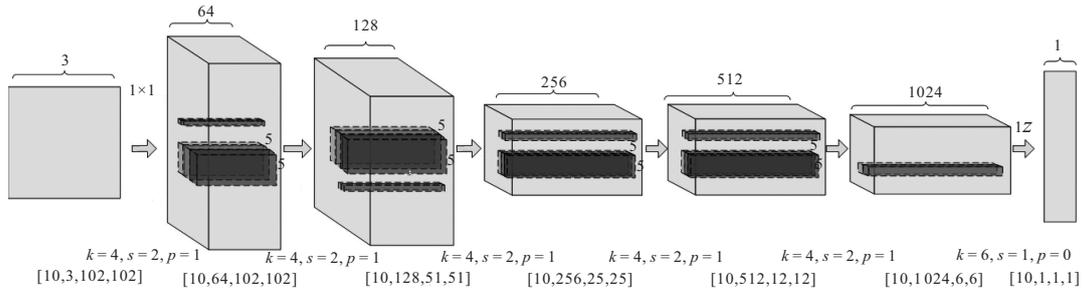


图3 判别模型示意图

算后,将维度转换为(1, 1024, 1, 1)的四维张量;最后,通过一个卷积核为 $6 \times 6$ 、步长为1的卷积运算转换为标量并输出。

## 2.2 RGAN算法

为了进一步提高对抗网络在训练过程的稳定性,缓解模式坍塌<sup>[3-6]</sup>现象,本文在DCGAN<sup>[10]</sup>的基础上,结合正负样本融合训练模式<sup>[11]</sup>设计并实现RGAN算法。

### 算法1 RGAN算法.

实验使用默认参数 $a_g=0.0001$ ,  $a_d=0.0003$ ,  $b=64$ ,  $\beta_1=0.5$ ,  $\beta_2=0.99$ 。

**注1**  $a_g, a_d$ 分别代表生成模型和判别模型的学习率;  $w$ 为判别模型学习到的参数;  $\theta$ 为生成模型学习到的参数。

**注2** 损失函数默认为BCEWithLogitsLoss (BCEWLoss). 优化器默认为Adam, 其参数为 $a_g, a_d, \beta_1, \beta_2$ 。

input: 随机噪声;

output: 是否为真实样本.

```

1 for 训练总次数 do
2   训练判别模型
3    $L_r = \text{BCEWLoss}(\text{RGAN-D}(\text{real-img}), \text{true-}$ 
labels)
4   real-img = RGAN-G(随机噪声)
5   gap1 = RGAN-D(fake-img) - RGAN-D
(real-img)

6    $L_f = \text{BCEWLoss}(\text{gap1}, \text{fake-img})$ 
7    $g_w \leftarrow \nabla_w \left[ \frac{1}{b} \sum_{n=1}^b \log(f_w(L_r)) - \right.$ 
 $\left. \frac{1}{b} \sum_{n=1}^b \log(f_w(g_\theta(L_f))) \right]$ 
8    $w \leftarrow \text{Adam}(g_w, w, a_d, \beta_1, \beta_2)$ 
9   判别模型的损失函数  $L_D = L_r + L_f$ 
10  训练生成模型
11  fake-img = RGAN-G(随机噪声)

```

```

12  gap2 = RGAN-D(fake-img) - RGAN-D(real-
img)

```

```

13   $L_G = \text{BCEWLoss}(\text{gap2}, \text{true-labels})$ 

```

```

14   $g_\theta \leftarrow \nabla_\theta \frac{1}{b} \sum_{n=1}^b (1 - D(G_\theta(L_G)))$ 

```

```

15   $\theta \leftarrow \text{Adam}(g_\theta, \theta, a_g, \beta_1, \beta_2)$ 

```

```

16  生成模型的损失函数  $L_G$ 

```

```

17  计算FID距离  $\frac{1}{b} \sum_{n=1}^b \text{FID}(\text{real-img}, \text{fake-img})$ 

```

```

18 end

```

文献[11]在逐点损失函数(pairwise objective functions)中的一对一正负实例样本训练方法的基础上改变了正负样本的比例,从而提高了隐式反馈中的分类效果. 本文在其工作的基础上进行改进,并运用到生成式对抗网络中,具体做法如下。

1) 当生成模型固定,判别模型的值函数求取最大值时,为了让生成分布更加接近原始数据集的真实分布,本文在传统GAN求取交叉熵损失函数之前,先计算出生成分布与真实分布之间的差值gap1(如算法1中第5行所示),然后与假标签作交叉熵运算(如算法1中第6行所示). 具体做法是:首先,将经判别模型的真实样本与真标签作交叉熵运算来构造真实样本的损失函数,同时将分别经由判别模型的融合样本与假标签作交叉熵运算来构造生成样本的损失函数,通过构造损失函数的方式建立 $x \sim P_{\text{data}}$ 和 $z \sim P_z$ 两组映射函数;然后,使用优化函数对所求取梯度的损失函数进行优化,并建立 $G(z, \theta)$ 之间的函数映射;最后,将二者损失函数之和作为判别模型总损失函数(如算法1中第9行所示). 判别模型目标函数公式如下所示:

$$\text{gap1} = D(G(z)) - D(x),$$

$$\max_D V(D, G) =$$

$$\mathbb{E}_{x \sim P_{\text{data}}} [\log(D(x))] +$$

$$\mathbb{E}_{(x,z) \sim (P_{\text{data}}, P_z)} [\log(1 - \text{gap1})]. \quad (8)$$

2) 当判别模型的值函数取得最大值时,要求模型

的值函数取得最小值. 首先计算出生成分布与真实分布之间的差值 gap 2 (如算法 1 中第 12 行所示), 然后与真标签作交叉熵运算 (如算法 1 中第 13 行所示). 具体做法是: 首先, 将经由判别模型的融合样本与真标签作交叉熵运算建立生成样本的损失函数, 通过该损失函数建立  $z \sim P_z$  映射函数; 然后, 使用优化函数对求取梯度的损失函数进行优化, 并建立  $G(z, \theta)$  之间的函数映射; 最后, 将生成样本的损失函数作为生成模型损失函数 (如算法 1 中第 16 行所示). 生成模型目标函数如下所示:

$$\text{gap 2} = D(G(z)) - D(x),$$

$$\min_G V(G) = E_{(x,z) \sim (P_{\text{data}}, P_z)} [\log(1 - \text{gap 2})]. \quad (9)$$

3) 结合 TTUR<sup>[6,13,16]</sup> 学习策略进行网络的优化. TTUR 是由 Martin 等<sup>[6]</sup> 提出的一种学习策略, 目的是找到合适的学习率, 使得网络能够快速收敛. 一般情况下, 在训练的过程中需要保持慢速的生成模型学习率和快速的判别模型学习率, 故本文将判别模型的学习率设置为 0.000 3 (如算法 1 中第 8 行所示), 生成模型的学习率设置为 0.000 1 (如算法 1 中第 15 行所示).

4) 使用 FID 评价指标计算生成样本与真实数据之间的相似距离 (如算法 1 中第 17 行所示).

### 3 实现设计与分析

为了验证本文所提出算法的有效性, 在配置为 Inter(R)Corei7-6800 K@3.40 GHz 处理器, 操作系统 Ubuntu16.04 LTS, 运行内存 16 GB (RAM), 显卡为 GIGABYTE GeForce GTX1080Ti, Pytorch0.4.1 的计算机中, 利用 102 Category Flower 数据集<sup>[17]</sup> 进行实验, 并与 GAN<sup>[1]</sup>、CGAN<sup>[6]</sup>、DCGAN<sup>[8]</sup>、LSGAN<sup>[7]</sup>、BEGAN<sup>[18]</sup> 进行实验对比.

102 Category Flower 数据集<sup>[17]</sup> 是由牛津大学工程科学系视觉几何组建立的高分辨率彩色数据集, 单张图片的大小可达  $800 \times 800$  像素. 与传统数据集不同, 该数据集有着较大幅度的种类姿势和光照变化, 拥有丰富的颜色特征和纹理特征. 实验过程中超参数设置如下: 判别模型学习率为 0.000 3, 生成模型学习率为 0.000 1, 图像批处理大小为 64, 优化器 Adam 参数  $\beta_1$  为 0.5、 $\beta_2$  为 0.99, 训练次数为 4 000.

1) 为了解决梯度消失和梯度爆炸的问题, 生成模型中使用 ReLU 代替传统的激活函数 Sigmoid, 判别模型则采用 LeakyReLU 激活函数.

2) 为了解决网络在训练过程出现的过拟合现象, 使用归一化层来代替 Dropout, 相对于传统的归一化

BatchNormalization (BN)<sup>[19]</sup>, 本文采用的 Instance Normalization<sup>[20]</sup> 相当于在 BN 的基础上对批处理数据又作了一次归一化处理, 以便能够根据不同样本的属性生成与之对应的样本.

3) 为了解决卷积运算之间通道转换问题, 本文借鉴了 Network in Network<sup>[21]</sup> 的思想, 充分利用  $1 \times 1$  卷积运算构建了 RGAN 模型.

4) 为了生成更高质量的样本, 本文借鉴了 Wasserstein-GAN<sup>[22]</sup> 构建网络时的思想, 在判别模型的最后一层中去除了单独的 Sigmoid 层和与其对应的传统损失函数 BCELoss, 转而使用比其更加稳定的损失函数 BCEWithLogitsLoss.

图 4 为生成模型的损失函数变化曲线, 训练初期, 曲线的变化幅度相对缓慢, 后期则呈现大幅度震荡状态; 图 5 为判别模型的损失函数的变化曲线, 呈现先增后降的趋势, 最终降至 0.1 左右. 通过对比两条损失曲线可以得出判别模型在对抗过程中仍占据优势.

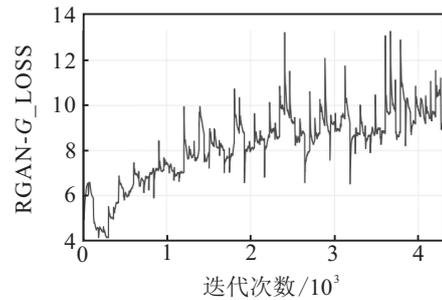


图 4 生成模型损失函数变化曲线

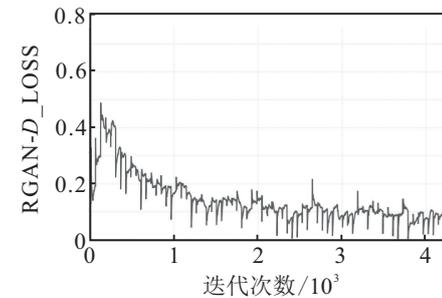


图 5 判别模型损失函数变化曲线

#### 3.1 客观评价指标

本实验采用弗雷歇距离 (fréchet inception distance, FID)<sup>[6,13,23]</sup> 作为客观评价指标. FID 是一个更具原则性和综合性的指标, 它可以更好地捕捉生成图像与真实图像的相似性, 并且在区分生成样本和真实数据上被证明符合人类的区分准则<sup>[23]</sup>, FID 主要用于计算生成样本与 Inception-v3<sup>[24]</sup> 网络的特征空间中的真实数据之间的相似距离. 具体做法为: 首先, 将生成样本和真实数据嵌入到由 Inception-v3<sup>[24]</sup> 网络所给出的特征空间中; 然后, 将特征空间所组成的嵌入层视为多个连续的多元高斯方程来估计生成

样本和实际数据的均值和协方差;最后,使用生成样本和实际数据的均值和协方差来量化生成样本的质量. 通常较低的FID值意味着生成样本与实际数据分布之间的距离更近. 计算公式如下:

$$FID(x, g) = \|u_x - u_g\|_2^2 + \text{Tr}(\Sigma_x + \Sigma_g - 2(\Sigma_x \Sigma_g)^{1/2}), \quad (10)$$

其中  $(u_x, \Sigma_x)$  和  $(u_g, \Sigma_g)$  分别是真实数据分布和样本数据分布的均值和协方差<sup>[6,23]</sup>.

### 3.2 主观评价指标

图6为各模型算法在 102 Category Flower Dataset 数据集上以相同超参数进行训练所产生的实验结

果. 在实验过程中GAN、CGAN、LSGAN均出现了FID值回升现象,导致FID值回升的主要原因是训练

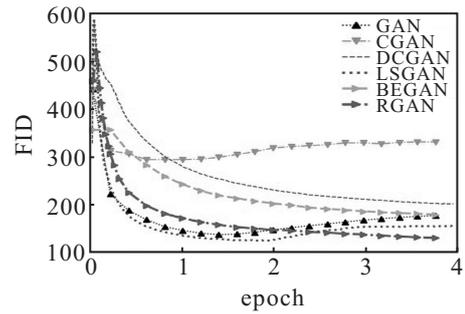


图6 102 Category Flower数据集实验结果分析

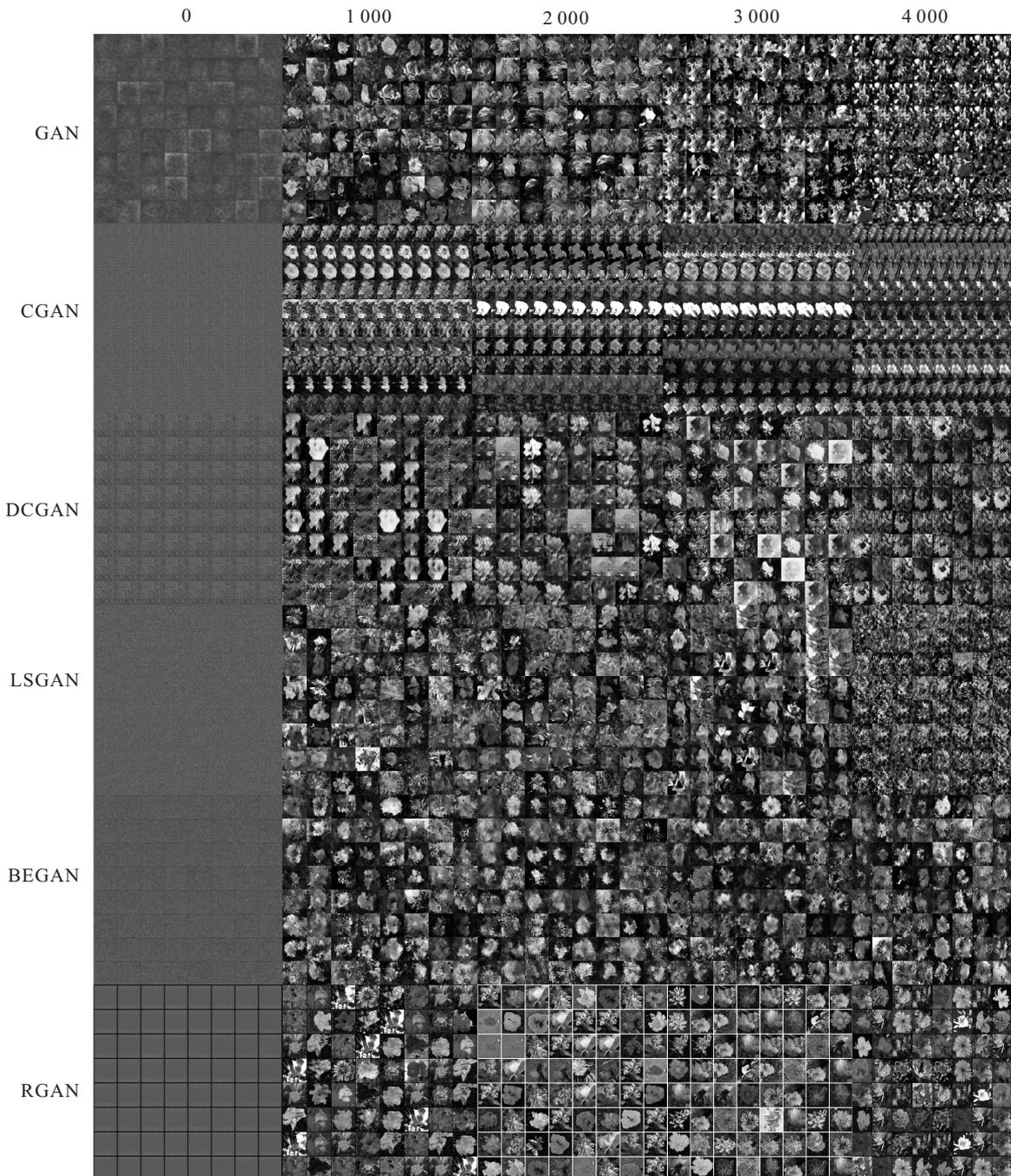


图7 样本生成算法结果对比图

过程不稳定. GAN的优化实质是找到真实分布与生成分布之间的詹森香农散度,当二者具有极少重叠或者没有重叠时,詹森香农散度为常数,从而导致优化梯度消失,造成训练过程不稳定现象的发生.通常伴随着训练过程不稳定现象一起出现的是模式坍塌<sup>[3-6]</sup>,模式坍塌<sup>[3-6]</sup>是一种大量真实数据概率被映射到几个小区域的现象,虽然生成的这些小区域也代表着有意义的样本,但却导致大量原型数据丢失.结合主观评价指标,不难发现训练过程不稳定现象往往伴随模式坍塌<sup>[3-6]</sup>一起出现,例如:图6中的CGAN曲线和图7中的CGAN生成样本.此外,为了验证图6中FID曲线与图7中生成样本之间的关系,本文以LSGAN为例,当LSGAN训练至2000次左右时,FID值为115.23,所生成的样本具有较高的图像纹理和色彩,而LSGAN当训练至4000次左右时,FID值已经回升至140.23,所生成的样本已经出现了模糊现象,完全不能分辨出生成样本的种类,故验证了FID评价指标的有效性.相对于以上算法出现FID值回升现象,DCGAN、BEGAN、RGAN均平稳下降,其中本文所采用算法RGAN在训练4000次时FID值平稳下降至120.37,远优于相同迭代次数下DCGAN的194.25和BEGAN的170.76.具体实验结果如表1所示.

表1 各模型在102 Category Flower数据集实验结果

模型	年度	FID	回升现象
GAN <sup>[1]</sup>	2014	129.18	有
CGAN <sup>[8]</sup>	2016	288.96	有
DCGAN <sup>[10]</sup>	2016	194.25	无
LSGAN <sup>[7]</sup>	2017	115.23	有
BEGAN <sup>[18]</sup>	2017	170.76	无
RGAN	2018	120.37	无

图7为各个生成模型在102 Category Flower数据集上5个训练阶段所生成的样本,纵向从上向下依次为GAN、CGAN、DCGAN、LSGAN、BEGAN、RGAN六个模型;横向从左向右依次为6个模型在0、1000、2000、3000、4000五个阶段所生成的样本.当训练次数为0时,各个模型生成的样本均模糊不清,无法辨识.当训练次数至1000时,GAN、LSGAN、RGAN生成的样本质量较高,具有较好种类辨识度,而其余模型算法生成的样本相对模糊,无法辨识生成样本的类别,且CGAN与DCGAN明显出现了模式坍塌<sup>[3-6]</sup>现象.当训练次数至2000时,除GAN模型外,其他模型算法生成的样本均较上一轮有所提高,而GAN生成的样本出现了图像模糊,样本扭曲现象,结合图6中客观评价指标曲线可以发现,GAN模型已

经发生了训练过程不稳定现象.当训练次数至3000时,GAN、CGAN、LSGAN均出现生成样本模糊现象,其中GAN模型最为明显,其生成的样本不仅出现了样本模糊现象,而且还发生了大范围模式坍塌<sup>[3-6]</sup>现象,而DCGAN、BEGAN、RGAN算法生成的样本在色泽以及清晰度方面均优于上一轮.最后,当训练次数至4000时,LSGAN和GAN生成的样本中充满大范围椒盐噪声,因而难以辨识生成样本的类别. CGAN在整个训练过程中都出现了模式坍塌<sup>[3-6]</sup>的现象.尽管DCGAN、BEGAN在整个训练过程中均不断优化生成的样本质量,但效果并不明显,且DCGAN生成的样本中常伴随模式坍塌<sup>[3-6]</sup>现象.虽然本文算法的FID值高于LSGAN的最低值115.23,但整个训练过程平稳,无FID回升现象,且生成样本具有丰富图像纹理和颜色特征.

## 4 结论

本文结合深度残差网络与生成式对抗网络,提出了一种基于深度残差的生成式对抗网络的样本生成方法,该生成方法主要通过构建深度残差网络和RGAN算法来提高样本生成质量.通过在102 Category Flower Dataset数据集中建立多组仿真对比实验,从主客观评价指标中证实了本文所提出算法和模型在绝大多数情况下效果优于其他算法及模型,从而验证了本文所提出的算法和模型的可行性.为了进一步提高训练过程速度较慢以及生成样本依然不够真实等问题,在未来的研究中,一方面将沿着如何让网络模型能够更快速达到全局收敛状态展开研究;另一方面将考虑沿着如何生成更高质量样本的方向继续探索,将生成式对抗网络应用于单张图像的超分辨率.

## 参考文献(References)

- [1] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets[C]. Proceedings of the International Conference on Neural Information Processing Systems. Montreal: Curran Associates, 2014: 2672-2680.
- [2] 唐贤伦, 杜一铭, 刘雨微, 等. 基于条件深度卷积生成对抗网络的图像识别方法[J]. 自动化学报, 2018, 44(5): 855-864.  
(Tang X L, Du Y M, Liu Y W, et al. Image recognition with conditional deep convolutional generative adversarial networks[J]. Acta Automatica Sinica, 2018, 44(5): 855-864.)
- [3] 张龙, 赵杰煜, 叶绪伦, 等. 协作式生成对抗网络[J]. 自动化学报, 2018, 44(5): 804-810.  
(Zhang L, Zhao J Y, Ye X L, et al. Co-operative generative adversarial nets[J]. Acta Automatica Sinica, 2018, 44(5): 804-810.)

- [4] 王坤峰, 苟超, 段艳杰, 等. 生成式对抗网络 GAN 的研究进展与展望[J]. 自动化学报, 2017, 43(3): 321-332. (Wang K F, Gou C, Duan Y J, et al. Generative adversarial networks: The state of the art and beyond[J]. Acta Automatica Sinica, 2017, 43(3): 321-332.)
- [5] 王万良, 李卓蓉. 生成式对抗网络研究进展[J]. 通信学报, 2018, 39(2): 135-148. (Wang W L, Li Z R. Research progress in generative confrontation network[J]. Journal of Communications, 2018, 39(2): 135-148.)
- [6] Heusel M, Ramsauer H, Unterthiner T, et al. Gans trained by a two time-scale update rule converge to a local nash equilibrium[C]. Proceedings of the Advances in Neural Information Processing Systems. Long Beach: Curran Associates, 2017: 6626-6637.
- [7] Mao X, Li Q, Xie H, et al. Least squares generative adversarial networks[C]. Proceedings of the Computer Vision 2017 IEEE International Conference. Venice: IEEE, 2017: 2813-2821.
- [8] Mirza M, Osindero S. Conditional generative adversarial nets[J]. arXiv:1411.1784, 2014.
- [9] Salimans T, Goodfellow I, Zaremba W, et al. Improved techniques for training gans[C]. Proceedings of the Advances in Neural Information Processing Systems. Barcelona: MIT Press, 2016: 2234-2242.
- [10] Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks[J]. arXiv:1511.06434, 2015.
- [11] He X, Liao L, Zhang H, et al. Neural collaborative filtering[C]. Proceedings of the 26th International Conference on World Wide Web. Perth, 2017: 173-182.
- [12] He X, Du X, Wang X, et al. Outer product-based neural collaborative filtering[C]. Proceedings of the International Joint Conferences on Artificial Intelligence. Stockholm: Morgan Kaufmann, 2018: 2227-2233.
- [13] Zhang H, Goodfellow I, Metaxas D, et al. Self-attention generative adversarial networks[C]. Proceedings of the 36th International Conference on Machine Learning. Long Beach, 2019: 7354-7363.
- [14] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 770-778.
- [15] Ledig C, Wang Z, Shi W, et al. Photo-realistic single image super-resolution using a generative adversarial network[C]. Proceedings of the Computer Vision and Pattern Recognition. Hawaii: IEEE, 2017: 105-114.
- [16] Konda V. Actor-critic algorithms[J]. Siam Journal on Control Optimization, 2003, 42(4): 1143-1166.
- [17] Nilsback M, Zisserman A. 102 category flower dataset[DB/OL]. [2008-12-12]. <http://www.robots.ox.ac.uk/~vgg/data/flowers/102/>.
- [18] Berthelot D, Schumm T, Metz L. BEGAN: Boundary equilibrium generative adversarial networks[J]. arXiv: 1703.10717, 2017.
- [19] Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift[C]. Proceedings of the 32nd International Conference on Machine Learning. Lille Grand Palais: ACM, 2015: 448-456.
- [20] Huang X, Belongie S J. Arbitrary style transfer in real-time with adaptive instance normalization[C]. Proceedings of the Computer Vision 2017 IEEE International Conference. Venice: IEEE, 2017: 1510-1519.
- [21] Lin M, Chen Q, Yan S. Network in network[J]. arXiv: 1312.4400, 2013.
- [22] Gulrajani I, Ahmed F, Arjovsky M, et al. Improved training of wasserstein gans[C]. Proceedings of the International Conference on Neural Information Processing Systems. Long Beach: MIT Press, 2017: 5767-5777.
- [23] Lucic M, Kurach K, Michalski M, et al. Are gans created equal? A large-scale study[C]. Proceedings of the Advances in Neural Information Processing Systems. Montreal: MIT Press, 2018: 697-706.
- [24] Xia X, Xu C, Nan B. Inception-v3 for flower classification[C]. Proceedings of the International Conference on Image, Vision and Computing. Chengdu: IEEE, 2017: 783-787.

### 作者简介

王星(1983-), 男, 教授, 博士生导师, 从事智能数据与知识工程等研究, E-mail: xingwang\_lntu@163.com;

杜伟(1992-), 男, 硕士生, 从事图像处理与机器学习的研究, E-mail: 18342918905@163.com;

陈吉(1981-), 女, 工程师, 硕士, 从事图像处理与机器学习的研究, E-mail: jichen\_lntu@163.com;

陈海涛(1995-), 男, 硕士生, 从事图像处理与机器学习的研究, E-mail: 15564952035@163.com.

(责任编辑: 孙艺红)