

DOI:10.19651/j.cnki.emt.1802073

# 强化组合式生成对抗网络

孙靖文 王敏

(河海大学 计算机与信息学院 南京 211100)

**摘要:**近年来通过在生成对抗网络中加入条件标签,控制生成图像的类别或属性已经取得很大进展,但是生成图像的类别或属性的准确性有待提高。为此,在生成对抗网络的判别器中加入了强化学习,通过上一次的分类结果去指导当前的分类。另外为了让生成的细粒度图像更加逼真,使用注意力机制在只增加少量的计算损失下让图像有全局感受野,将多属性的星型生成对抗网络与自注意力生成对抗网络组合后的生成图像的质量较高。强化组合式生成对抗网络的最大均差达到0.036 93,最近邻指标效果较优,能自动化较准确地生成指定了某些属性的艺术图像,实验生成的图片也能用来解决缺乏数据的问题。

**关键词:**生成对抗网络;注意力机制;强化学习;图像生成;风格迁移

中图分类号: TP391; TN919.81 文献标识码: A 国家标准学科分类代码: 520.2040

## Reinforced-combined generative adversarial networks

Sun Jingwen Wang Min

(College of Computer and Information, Hohai University, Nanjing 211100, China)

**Abstract:** In recent years, great progress has been made in controlling the categories or attributes of generated images by adding condition tags to the generation adversarial networks. However, the accuracy of the category or attribute of generated image needs to be improved. In order to solve this problem, we add reinforcement learning to the generator of generation adversarial networks, which guides the current classification by previous. In addition, the attention mechanism is used which makes a global sensory field to the images with only a small amount of computational loss. We combines multi-attribute star generation adversarial networks with self-attention generation adversarial networks which improves the quality of generated. maximum mean discrepancy reaches to 0.036 93 and the 1-nearest neighbor classifier has a better effect by reinforced-combined generative adversarial networks, which can generate the art images that certain attributes are assigned automatically and accurately. The generated images can also be used to address the lack of data.

**Keywords:** generative adversarial networks; attention mechanism; reinforcement learning; image generation; style transfer

## 0 引言

生成对抗网络使计算机视觉中合成图像的分支得到了很大的发展。通过在生成对抗网络的输入中加入类别条件信息可以在一定程度上控制生成图像的类别条件。周期一致生成对抗网络(cycle-consistent adversarial networks, CycleGAN)把条件类别条件信息用某类风格的图片代替实现了不同物体间风格的转化。让生成对抗网络得到条件信息不仅可以在输入上加入标签,还可以添加类别损失,在分类取得了一定效果,但是判别图像类别的准确率有待提高<sup>[1-3]</sup>。星型生成对抗网络(star generate adversarial networks, SatrGAN)改进了CycleGAN单属性的标签输入,

并且加入了类别失,解决了如果需要其他属性信息就要再次训练的问题,但是提取特征受到卷积网络的局部感受野限制。为了生成较好的细粒度图片自注意力生成对抗网络(self-attention generative adversarial networks, SAGAN)采用了能够利用全局信息的注意力机制,同时不会增加大量的计算。在PASCAL VOC2012数据集上的实验表明了强化学习<sup>[4]</sup>用在多属性分类上性能有所提升<sup>[5]</sup>。

强化组合式生成对抗网络(reinforced-combined generative adversarial networks, RCGAN)在生成对抗网络的判别器中加入强化学习,即使没有得到很好的特征用来分类,也能通过奖励或者惩罚去指导分类结果。对于纹理比较细微的图像,通过注意力机制的全局感受野获得关

注区域,较忽略次要的部分。

## 1 相关工作

### 1.1 生成对抗网络

生成对抗网络在 2014 年被 Goodfellow 提出,由一个判别器  $D$  和一个生成器  $G$  组成<sup>[6]</sup>。 $G$  从  $P_z(z)$  分布的低维噪声中学习生成高维的图片  $G(z)$ , 其中噪声分布  $z \sim N(0,1)$ 。 $G$  的作用是生成误导  $D$  的图片,  $D$  的作用是区分生成图片  $G(z)$  和真实图片  $x$ ,  $D$  与  $G$  在博弈中学习到各自的参数。当  $D$  判别  $G(z)$  与  $x$  的概率近似相等训练完成<sup>[7]</sup>。如下:

$$V = E_{x \sim P_{data}(x)} [\log D(x) + E_{z \sim P_z} [\log(1 - D(G(z)))] \quad (1)$$

### 1.2 自注意力生成对抗网络

取类内的谱正则化和双时间尺度更新规则能够促进达到纳什均衡,稳定性有所提高。将损失函数整合成一个分布去逼近真实标签沃瑟斯坦生成对抗网络(Wasserstein generate adversarial networks, WGAN),于此改进使用普希茨连续性的限制,解决了梯度爆炸使其更快收敛,不易模式崩溃。注意力机制曾被的文本上, SAGAN 首次使用在生成对抗网络的内部。注意力机制是全局依赖,能有效找到全局、长期具有内部代表性的图像区域。使用注意力机制之后的 SAGAN 生成图片的质量达到当前最好。

### 1.3 多属性的星型生成对抗网络

它使用一个生成器和一个判别器训练一次可以得到多属性的图片。通过设置数据集的标记向量,能够跨数据集

合并属性特征。将属性标签经过形变和扩充后和真实图片连接,通过生成器后产生假图片,再把该图和目标属性标签变换后连接,经过生成器后,计算当前产生的图片与上一步产生的假图片的重塑结构损失作为损失函数的一部分。图像的类别损失为判别器判断图像类别的辅助损失函数。真假图像的损失函数与 WGAN 的计算方式相同。图像翻译效果较好<sup>[8]</sup>。

### 1.4 多属性的强化学习分类

选择图像的一个标签作为动作,若有  $N$  个标签则有  $N$  个动作。决策过程就是运用神经网络来预测每个状态动作对应的奖励或者惩罚。状态是全连接层和预测标签向量拼接的结果。预测标签向量为选择值最大的标签,若预测正确第一个位置的值加 1,否则减 1,作为奖励或者惩罚。再次从更新的全连接层开始运行,选择除之前选择的标签的最大值作为第 2 个数,依次累加奖励或惩罚。在第 1 个全连接和之后的网络层进行迭代每个标签<sup>[9-11]</sup>。

## 2 提出的方法

### 2.1 网络总体结构

网络的总体解构如图 1 所示。网络训练过程如图 1 中间箭头及标号,判别器先训练 5 次(标号 1、2、4),生成器再训练 1 次(标号 3、5、6)。在判别器的判别类别的最后一层使用了强化学习。图形表示对应的描述在其正下方或正下方,相同的颜色表示相同的含义。“ $\times$ ”后的数字表示该网络重复的次数。除了生成器的最后一层使用双曲正切 Tanh 激活,其余的激活函数均为修正线性单元。

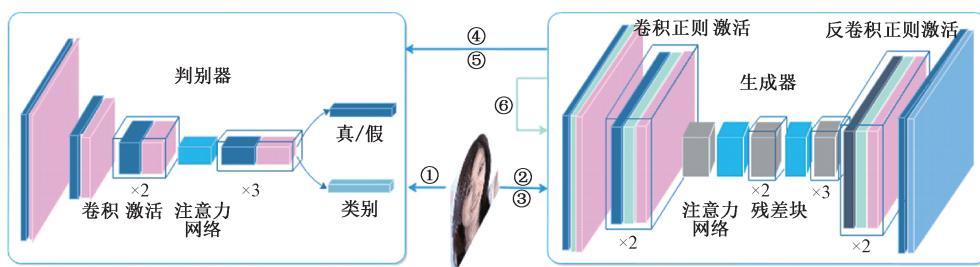


图 1 网络总体解构

网络的损失函数分成 3 部分:生成损失、类别损失、重构损失。此处与式(1)不同之处在于输入的条件信息不仅有图像种类  $c$ ,还有该图像本身  $x$ 。判断真假的判别器记为  $D_{src}$ ,如下:

$$\zeta_{adv} = E_x [\log D_{src}(x)] + E_{x,c} [\log(1 - D_{src}(G(x, c)))] \quad (2)$$

记  $D_{cls}$  为判别类别损失;  $r$  为真实图片;  $f$  为生成图片;  $c'$  为原始图像类别标签;  $c$  为目标标签,即希望得到的生成图片的类别,真实图像与生成图像的类别损失分别如下:

$$\zeta_{cls}^r = E_{x,c'} [-\log D_{cls}(c' | x)] \quad (3)$$

$$\zeta_{cls}^f = E_{x,c} [-\log D_{cls}(c | G(x, c))] \quad (4)$$

为了只改变输入图像主要相关部分,采用了重构结构损失,保持次要部分的图像尽量变化小。

$$\zeta_{res} = E_{x,c,c'} [\|x - G(G(x, c), c')\|_1] \quad (5)$$

生成器和判别器总的损失函数如下,  $\lambda_{cls}$  和  $\lambda_{res}$  是控制损失重要程度的超参数,在实验里分别设置为 1 和 10。

$$\zeta_D = -\zeta_{adv} + \lambda_{cls} \zeta_{cls}^r \quad (6)$$

$$\zeta_G = \zeta_{adv} + \lambda_{cls} \zeta_{cls}^f + \lambda_{res} \zeta_{res} \quad (7)$$

### 2.2 注意力网络

注意力网络在中间或者偏后层的网络能发挥较好的效果。残差网络在增加深度的情况下能提取到有效特征,其

使用注意力网络在不同的标准数据集上识别达到了当时最好的水平。在 RCGAN 生成器的第 5 层插入一层注意力网络, 在判别器的第 1 个和第 3 个残差块后各插入一层。

$f(x)$ 、 $g(x)$  和  $h(x)$  分别是同一张特征图通过不同的  $1 \times 1$  的卷积核得到新的特征图。取两个特征图  $f(x)$  的转置和  $g(x)$  相乘后经过激活函数的值  $\beta_{j,i}$ , 即获得从  $i$  区域映射到合成图的  $j$  区域的关系, 公式如下:

$$\beta_{j,i} = \frac{\exp(s_{i,j})}{\sum_{i=1}^N \exp(s_{i,j})}, \text{ 其中 } S_{i,j} = f(x_i)^T g(x_i) \quad (8)$$

自注意力网络的输出为  $o_j$ , 公式如下:

$$o_j = \sum_{i=1}^N \beta_{j,i} h(x_i) \quad (9)$$

$\gamma$  初始化为 0,  $x$  为原始输入, 得到叠合的注意力网络的输出  $\gamma$  的公式(式(10))和图 2 所示, 因为矩阵相乘:

$$\gamma_i = \gamma o_i + x_i \quad (10)$$

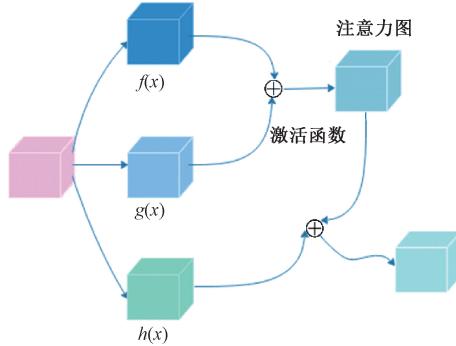


图 2 注意力机制式(8)~(10)的图形解析

### 2.3 强化学习

分类标签  $A$  相当于强化学习里的动作。状态  $S$  由两个元素组成, 第 1 个元素是判别器中的用作图像分类的最后一层, 第 2 个元素是之前的动作  $h$ , 代表之前预测, 每一个状态都属于整个状态空间。第 2 个元素的赋值过程如下图的粉红色部分, 每个方块代表一个标签, 按顺序依次自上而下排列, 将此纵向的向量复制 4 份连接到判别器判别图像类别的最后一层的特征图的每一个像素点上, 把拼接后的特征图放回到判别类别损失的卷积层中, 选取值最大的  $A$ , 若值最大的是第 2 类且判断正确, 将向量的第 2 个值改为该标签数(0 或 1)作为奖励, 迭代重复上述过程选出除了第 2 类值最大的类别, 假设是第 1 类, 若分类结果错误, 则把该处的值改成 -2 作为惩罚。依次循环标签个数的次数得到最后的结果。上述变化过程相当于强化学习中的状态转移。最后把分类结果的好坏累计在了  $h$  中<sup>[12-13]</sup>, 如图 3 所示。

## 3 实验

使用了比较大型的 CelebA 人脸数据库生成多属性人脸, 与 SAGAN 和 StarGAN 进行了性能比较。另外使用了

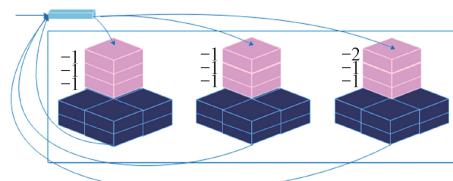


图 3 判别器最后一层判断类别的强化学习的图形表示

较小的数据, 实验了对于小数据集的效果和不同风格的图片的生成结果。

### 3.1 数据集

CelebA 图片张数 202 599, 具有 40 个不同标注的二分类, 图像分辨率为  $178 \times 218$ 。输入时网络时把图像裁剪成  $178 \times 178$  个像素, 之后缩小成  $128 \times 128$  的大小。随机的取出 2 000 张作为测试, 其余的用来训练。采用了标注类别中的 5 类, 包括是否黑发、金发、棕发和性别、年龄。

第 1 个网络图片的小型数据集包括夕阳和云雾的风景图各 1 000 张, 第 2 个有 4 类包括人脸石雕、人脸彩陶、人脸素描、国画的人脸各 700 张, 采用同样的裁剪方式, 分别还加入从 CelebA 中抽取的 2 000 张图像。

### 3.2 实现细节

用大量数据集训练时设置学习速率为 0.000 1, 步长设置为 64, 训练使用随机梯度优化算法, 一半的概率让图片水平翻转, 判别器更新 5 次, 生成器更新 1 次。

合并小型的数据集的属性时, 步长设置为 16, 训练到 10 个回合后逐渐降低学习速率。训练时从 CelebA 中随机挑选了 2 000 张图像, 分别加入网络风景数据集和网络艺术人脸数据集

构成两组训练的图片, 把每组的两个数据集的类别标签拼接, 并且将两个数据集的标号设置成独热标签, 分别加到已拼接标签的后边, 两部分的数据集标签的维度一致。交替训练两部分数据。

### 3.3 模型评估

将 RCGAN、StarGAN 和 SAGAN 生成图像的质量进行了对比, 同时比较了 StarGAN 和 RCGAN 判别器划分图像类别的性能。

如图 4 所示, 左侧为 StarGAN 生成的图片, 第 2 行的肤色受到输入标签的影响, 右侧为 RCGAN 生成的图片, 对标签之外的其他特征影响较小。

通过带附加角度损失的人脸识别测试了三者生成人脸的准确程度, 反映了生成人脸的真实度<sup>[14]</sup>。如表 1 所示, RCGAN 生成的人脸真实度更高。

表 1 人脸识别准确率

测试方法	人脸识别
StarGAN	0.995 99
SAGAN	0.997 13
RCGAN	0.999 00



图 4 StarGAN 与 RCGAN 的效果对比

基于核的最大均方差(maximum mean discrepancy, MMD)计算方法简单、计算量小,固定的核方法衡量真实图像与生成图像的分布,测其真实度与多样性,值小的较好。若真实图片与生成图片分布相同,最近邻分类器(1-nearest neighbor classifier, 1-NN)的准确率在 50%。当生成图片过拟合该值小于 50%,两者分布相同时准确率为 0。如果生成对抗网络的模型崩塌,该值会很高。两者都采用 Inception-v3 训练好的模型提取真实图像的特征和生成图像的特征<sup>[15]</sup>。

MMD 和 1-NN 方法是判断生成对抗网络生成图像质量和多样性标准比较好的衡量标准。基于 CelebA 人脸测试的 MMD 与 1-NN 分类器的值如表 2 所示。RCGAN 生成图片的质量得到提升。判别器或生成器都有真实标签的参与,无法用来衡量无标签的图像分类的准确率。所以在测试图像中随机采样 2 000 张,人工统计其分类的准确率,把 3 种发色的判别效果合成了 1 个参数,得到的分类结果如表 3 所示。加入强化学习后各个类别的分类效果明显提升,尤其在性别方面。

表 2 不同的生成对抗网络生成图像的效果对比

测试方法	MMD	1-NN								
		真实图像	生成图像	总精度	准确率					
StarGAN	0.150	78	0.799	00	0.752	00	0.846	00	0.830	00
SAGAN	0.053	86	0.611	04	0.577	78	0.664	17	0.623	60
RCGAN	0.036	93	0.577	39	0.551	28	0.563	50	0.557	97

表 3 各标签的准确率 (%)

测试方法	发色	性别	年龄
StarGAN	66.2	31.9	70.6
RCGAN	81.3	67.8	75.7

RCGAN 在另外两个小数据集上生成的效果如图 5 所示。可以看出生成的彩陶风格的人脸图像颜色鲜艳,素描风格的是黑白轮廓,国画风格的轮廓连接较平滑,石雕风格的人脸面色暗沉。在数量较少的数据集上 RCGAN 也能够学习其代表性的特征。测试小型数据集上输入指定标签后 RCGAN 的多属性表现,如图 6 所示。可以看出混



图 5 融合两个小数据集特征的效果



图 6 自定义特征组合

合数据集的多属性图像较准确,训练一次就可以使用不同数据集所拥有的多重属性。对于色调变换不明显、纹理较大的图像如夕阳、云雾(右图)会改变整个图像色调,对人脸影响较少,而对于棱角分明、色彩鲜艳的图像如彩陶、国画等(左图),背景着色和人脸轮廓线条会都有较大的变化。

## 4 结 论

RCGAN 将多属性转化的生成对抗网络和自注意力的生成对抗网络组合,使多属性的生成对抗网络有注意力层,在不增加大量计算的情况下,有全局感受野。另外为了提高图像类别划分的准确性,在判别器区分图像类别的这个辅助分类器上加入强化学习,用前一次标签的分类去指导下次其它标签的分类。使用不同风格的图片探索了该网络对图片的约束性,寻找适合属性迁移的特征。本文的工作主要是在人脸方面,之后会尝试在自然场景下图像的多标签合成。

## 参考文献

- [1] 汪旗,贾兆红. 基于多示例学习的图像分类算法[J]. 计算机技术与发展, 2014, 24(4):88-91.
- [2] 柯善武,金聪. 基于优化空间金字塔匹配模型的图像分类[J]. 电子测量技术, 2017, 40(7):95-99.
- [3] 韩辰希,刘惠义,商国中.视觉显著性纹理——色彩特征融合的图像目标分类[J]. 电子测量技术, 2017, 40(11):94-98.
- [4] KAEELBLING L P, LITTMAN M L, MOORE A W. Reinforcement learning: A survey [J]. Journal of Artificial Intelligence Research, 1996, 4(1): 237-285.
- [5] GUO L L, DING S F. Research progress on deep learning[J]. Computer Science, 2015, 42(5): 28-33.
- [6] 张营营.生成对抗网络模型综述[J]. 电子设计工程, 2018, 26(5):34-37.
- [7] WANG K F, GOU C, DUAN Y J. Generative adversarial networks: the state of the art and beyond [J]. Acta Automatica Sinica, 2017, 43(3): 321-332.
- [8] 曹志义,牛少彰,张继威. 基于半监督学习生成对抗网络的人脸还原算法研究[J]. 电子与信息学报, 2018, 40(2):323-330.
- [9] 孙玉红,张元科,孟静,等. 基于纹理和草图的图像铅笔画绘制[J]. 计算机应用, 2016, 36(7):1976-1980.
- [10] 王坤峰,鲁越,王雨桐,等.平行图像:图像生成的一个新型理论框架[J]. 模式识别与人工智能, 2017, 30(7):577-587.
- [11] 谢志峰,杜胜,郭雨辰,等.基于字典学习的 HDR 照片风格转移方法[J]. 图学学报, 2017, 38(5): 706-714.
- [12] SALLAB A E, ABDOU M, PEROT E, et al. Deep reinforcement learning framework for autonomous driving[J]. Electronic Imaging, 2017, 19(8): 70-76.
- [13] 夏曼,宋稳柱,施必成,等.基于加权密集连接卷积网络的深度强化学习方法[J]. 计算机应用, 2018, 38(8): 2141-2147.
- [14] 王楠楠,李洁,高新波. 人脸画像合成研的综述与对比分析[J]. 模式识别与人工智能, 2018, 31(1): 37-48.
- [15] 林志洁,丰明坤. 深度视觉特征与策略互补融合的图像质量评价[J]. 模式识别与人工智能, 2017, 30(8): 682-691.

## 作者简介

王敏,工学博士、副教授,主要研究方向为计算机视觉。

孙靖文,工学硕士,主要研究方向为计算机视觉。

E-mail:1902686538@qq.com