

## 基于改进循环生成式对抗网络的图像风格迁移

张惊雷\* 厚雅伟

(天津理工大学电气电子工程学院 天津 300384)

(天津市复杂系统控制理论及应用重点实验室 天津 300384)

**摘 要:** 图像间的风格迁移是一类将图片在不同领域进行转换的方法。随着生成式对抗网络在深度学习中的快速发展,其在图像风格迁移领域中的应用被日益关注。但经典算法存在配对训练数据较难获取,生成图片效果差的缺点。该文提出一种改进循环生成式对抗网络(CycleGAN++),取消了环形网络,并在图像生成阶段将目标域与源域的先验信息与相应图片进行纵深级联;优化了损失函数,采用分类损失代替循环一致损失,实现了不依赖训练数据映射的图像风格迁移。采用CelebA和Cityscapes数据集进行实验评测,结果表明在亚马逊劳务平台感知研究(AMT perceptual studies)与全卷积网络得分(FCN score)两个经典测试指标中,该文算法比CycleGAN, IcGAN, CoGAN, DIAT等经典算法取得了更高的精度。

**关键词:** 图像风格迁移; 深度学习; 生成式对抗网络; 损失函数

中图分类号: TN911.73; TP317

文献标识码: A

文章编号: 1009-5896(2020)05-1216-07

DOI: 10.11999/JEIT190407

## Image-to-image Translation Based on Improved Cycle-consistent Generative Adversarial Network

ZHANG Jinglei HOU Yawei

(School of Electrical and Electronic Engineering, Tianjin University of Technology, Tianjin 300384, China)  
(Tianjin Key Laboratory of Complex System Control Theory and Application, Tianjin 300384, China)

**Abstract:** Image-to-image translation is a method to convert images in different domains. With the rapid development of the Generative Adversarial Network(GAN) in deep learning, GAN applications are increasingly concerned in the field of image-to-image translation. However, classical algorithms have disadvantages that the paired training data is difficult to obtain and the convert effect of generation image is poor. An improved Cycle-consistent Generative Adversarial Network(CycleGAN++) is proposed. New algorithm removes the loop network, and cascades the prior information of the target domain and the source domain in the image generation stage, The loss function is optimized as well, using classification loss instead of cycle consistency loss, realizing image-to-image translation without training data mapping. The evaluation of experiments on the CelebA and Cityscapes dataset show that new method can reach higher precision under the two classical criteria—Amazon Mechanical Turk perceptual studies(AMT perceptual studies) and Full-Convolutional Network score(FCN score), than the classical algorithms such as CycleGAN, IcGAN, CoGAN, and DIAT.

**Key words:** Image-to-image translation; Deep learning; Generative Adversarial Network (GAN); Loss function

### 1 引言

图像风格迁移,也称图像风格转换,是指将输入图像的风格,转变成指定一幅或多幅图像风格的方法。算法须保证原图像结构,只将图像风格进行转换,使得最终输出的合成图像呈现出与输入图像内容和风格的完美结合。图像风格迁移可以追溯到Hertzmann等人<sup>[1]</sup>发表的Image Analogies,使用非

参数纹理模型在单一输入输出配对训练图像中实现了风格迁移。由于非参数纹理模型只能提取底层特征而非高层抽象特征,且对不同类型图像处理的局限,此类方法多年来发展缓慢。近几年来,随着深度学习类算法的发展,基于卷积神经网络(Convolutional Neural Network, CNN)的监督学习类算法得到广泛应用,但人工创建的配对训练图像库规模有限,导致图片生成效果较差。直到Goodfellow等人<sup>[2]</sup>提出生成式对抗网络(Generative Adversarial Network, GAN),这一领域才焕发新的生机。典型

收稿日期: 2019-06-05; 改回日期: 2019-12-23; 网络出版: 2019-12-31

\*通信作者: 张惊雷 zhangjinglei@tjut.edu.cn

GAN模型由生成器与判别器两个模块组成, 生成器产生假图并使其与真图难以区分, 判别器学习区分真假图像。

随着GAN的提出, 基于GAN思想的深度神经网络逐渐应用于图像风格迁移领域中<sup>[3-5]</sup>。典型如“pix2pix”<sup>[6]</sup>方法, 在cGANs<sup>[7]</sup>的基础上采用监督学习进行训练, 将对抗损失与L1损失相结合, 提出了解决图像迁移问题的统一框架。由于监督学习类算法存在需预先人工处理大量配对数据并进行训练的通病, 因此基于无监督学习的图像风格转换框架被提出。Rosales等人<sup>[8]</sup>设计出一种从源图像计算的先验马尔可夫随机场和从多类图像获得似然项的贝叶斯框架。UNIT<sup>[9]</sup>把CoGAN<sup>[10]</sup>与Variational AutoEncoders<sup>[11]</sup>(VAEs)相结合, 在相关域中用两个生成器联合训练, 共享权重。循环生成式对抗网络(Cycle-consistent Generative Adversarial Network, CycleGAN)<sup>[12]</sup>, DiscoGAN<sup>[13]</sup>和DualGAN<sup>[14]</sup>利用循环一致损失将输入与转化图像的关键信息保存, 实现了数据不配对情况下的图像风格迁移。为提高生成图像的质量, 学者们试图在输入和输出之间共享不同风格的图像特征信息<sup>[15-17]</sup>。这些方法使用生成对抗网络并结合其他研究使输出在预定义度量空间内逼近输入, 如类标签空间<sup>[15]</sup>、图像像素空间<sup>[16]</sup>和图像特征空间<sup>[17]</sup>。

CycleGAN作为GAN类的经典图像风格迁移算法, 使用两个镜像对称的GAN, 使图片关键信息得以保留, 无需配对数据即可进行训练。但是由于缺少目标域的先验信息, 生成图像仍未达到较为理想的效果。

本文在CycleGAN上进行改进, 取消环形网络设计, 使用单组GAN进行训练。在生成网络中添加目标域特征信息, 改善生成图像的逼真程度, 提

升了转换精度。改进原损失函数, 将对抗损失, 交叉熵类别损失与重建损失相结合, 使网络拟合能力更强。

## 2 CycleGAN算法

### 2.1 CycleGAN算法原理

CycleGAN是由Zhu等人<sup>[12]</sup>发表于ICCV2017的成果, 采用两个单向GAN构成环形GAN, 解决了图像风格迁移时需要配对训练数据的问题。算法的网络结构如图1所示。首先将领域A真实图片 $I_A$ 通过生成器 $G_X$ 转化为领域B虚假图片 $I_{B'}$ , 然后把 $I_{B'}$ 通过生成器 $G_Y$ 进行图片重构, 得到重建图片 $I_{RA}$ , 使原始图片信息得以保存。接着将 $I_{B'}$ 与真实图片 $I_B$ 送入判别器 $D_Y$ 以此判断真伪, 最后得到完整的单向GAN。CycleGAN为两个单向GAN的组合, 由循环一致损失<sup>[6]</sup>监督训练网络。相关公式为

$$\mathcal{L}_{cyc}(G, F) = \mathbb{E}_{x \sim p_{data}(x)} [\|F(G(x)) - x\|_1] + \mathbb{E}_{y \sim p_{data}(y)} [\|G(F(y)) - y\|_1] \quad (1)$$

对抗损失由两部分组成

$$\mathcal{L}_{GAN}(G, D_Y, X, Y) = \mathbb{E}_{y \sim p_{data}(y)} [\ln D_Y(y)] + \mathbb{E}_{x \sim p_{data}(x)} [\ln(1 - D_Y(G(x)))] \quad (2)$$

以及与式(2)相似的 $\mathcal{L}(F, D_X, Y, X)$ 。

两个GAN共享两个生成器, 各有一个判别器, 因此通过4个损失函数组成联合损失训练整个环形GAN网络, 如式(3)

$$\mathcal{L}(G, F, D_X, D_Y) = \mathcal{L}_{GAN}(G, D_Y, X, Y) + \mathcal{L}_{GAN}(F, D_X, Y, X) + \lambda \mathcal{L}_{cyc}(G, F) \quad (3)$$

其中,  $X, Y$ 分别代表两个数据域,  $x, y$ 为两数据域

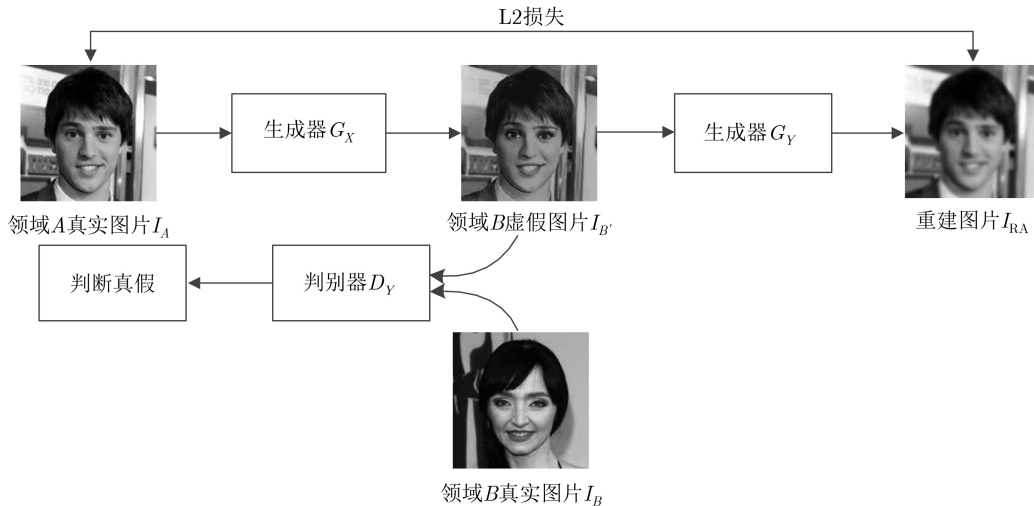


图1 CycleGAN中单向GAN网络结构图

中的样本数据,  $G$ 为从 $X$ 到 $Y$ 的映射函数,  $F$ 为从 $Y$ 到 $X$ 的映射函数,  $D_X, D_Y$ 为判别器,  $\lambda$ 用于控制循环一致损失函数的权重。

## 2.2 CycleGAN的缺点

CycleGAN因使用两个单向GAN组成网络, 网络结构的加深会使计算复杂度提高, 训练速率降低, 导致应用性与实时性较差。而且由于生成器没有得到目标变换域图像的先验信息, 生成图像的效果较差。在部分领域的图像变换中, 测试图片会出现模糊, 虚化, 变换效果差等问题。

## 3 本文算法

针对上述问题, 本文对网络结构重新进行设计, 取消环形结构, 移除循环一致损失并加入分类损失提高生成图像的效果, 最终得到改进算法CycleGAN++。

### 3.1 CycleGAN++的网络结构

本文重新设计了单向GAN, 采用纵深级联的方式导入域信息并加入交叉熵分类损失作为约束条件。纵深级联是一种将原始域图像与目标域信息使用级联(concat)操作进行融合的方法。例如, 图像与目标域信息均为3通道张量, 则生成器A的输入通过纵深级联可得到结果为6通道的张量, 实现了图像与域信息通道级别的联结, 在训练生成器的过程中提供了先验知识, 缩短了训练时间, 提升了生成图像的精确度。算法结构如图2所示。因算法采用无配对数据进行训练, 所以输入 $X$ 与输出 $Y$ 为各自领域的大量图像。取每个领域中一幅图像描述网络结构, 将原始域真实图像 $I_X$ 与目标域图像特征信息通过纵深级联送入生成器A中, 得到虚假图像 $I_{Y'}$ , 然后把生成的 $I_{Y'}$ 与原始域图像信息通过相同方法融合送入生成器B, 通过这一反向生成过程实现

图像重建, 得到重建后的图像 $I_{RX}$ , 实现了原图像内容的保存, 确保生成图像能够还原到原领域分类中, 达成网络局部的一个循环。另一方面把 $I_{Y'}$ 送入初步训练好的判别器中与真实图像 $I_Y$ 进行真伪判定, 若为真, 说明生成器优于判别器, 应对判别器进行再训练; 若为假, 则判别器优于生成器, 应对生成器进行再训练。在真伪判定过程中生成器与判别器不断进行博弈, 直到二者得到充分训练后达到拟合状态, 最终得出满意的生成图像。

### 3.2 生成器与判别器的神经网络设计

如图3所示, CycleGAN++的生成网络由下采样、残差块及上采样3部分组成。其中下采样部分包括3层步长为2的卷积, 残差块含有6层具有残差连接的卷积, 上采样由两层反卷积与一层卷积组成。在训练过程中对生成器使用实例正则化(instance normalization), 实例正则化对单张图像的数据进行处理, 起到整合全局信息的作用, 应用于图像转换任务中有较好效果。利用PatchGANs<sup>[18]</sup>来实现判别网络, 判断生成图片的真伪, 如图4所示。

### 3.3 相关损失函数的改进

CycleGAN++的总损失函数由3部分损失构成, 首先加入对抗损失函数来使生成图片与真实图片难以区分, 公式为

$$\mathcal{L}_{adv} = \mathbb{E}_x [\ln D_{src}(x)] + \mathbb{E}_{x,c} [\ln (1 - D_{src}(G(x, c)))] \quad (4)$$

其中,  $G$ 生成以输入图像 $x$ 和目标域标签 $c$ 为条件的图像 $G(x; c)$ , 而 $D$ 试图判别图像的真伪。 $D_{src}(x)$ 称为 $D$ 给出的源域概率分布。 $G$ 尝试最小化目标,  $D$ 的目的与之相反。

然后加入类别损失, 该损失被分为两类, 训练

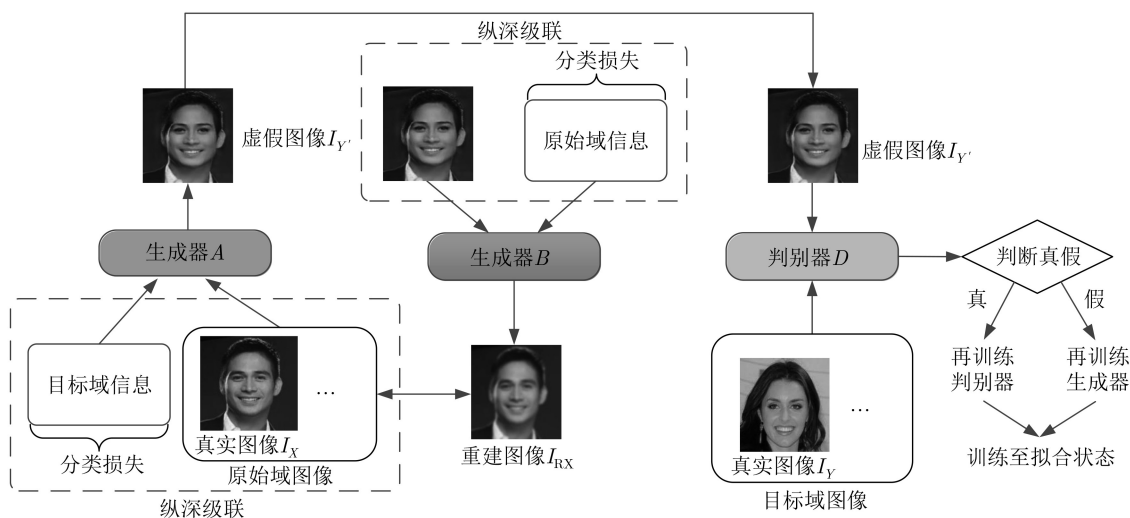


图2 CycleGAN++的网络结构

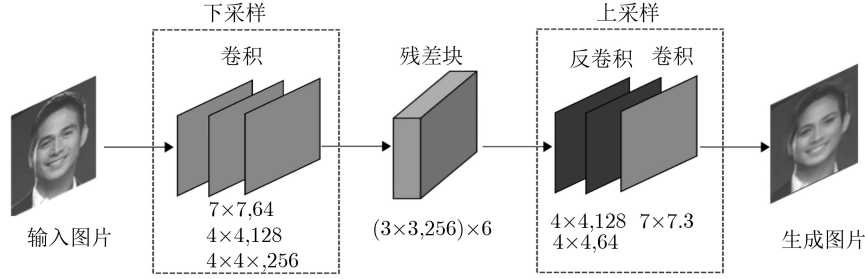


图3 CycleGAN++的生成网络

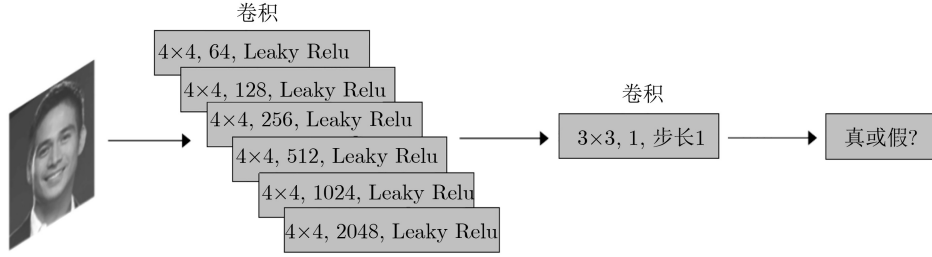


图4 CycleGAN++的判别网络

判别网络 $D$ 时，使用真实图像在源域监督信号下进行训练；而训练生成网络 $G$ 时，使用生成图像在目标域监督信号下训练。训练 $D$ 的损失函数为

$$\mathcal{L}_{\text{cls}}^r = \mathbb{E}_{x, c'} [-\ln D_{\text{cls}}(c' | x)] \quad (5)$$

对于给定的输入图像 $x$ 和目标域标签 $c$ ，目标是将 $x$ 转换为输出图像 $y$ 并被正确分类到目标域 $c$ 。选择在 $D$ 上添加一个辅助分类器，并在优化 $D$ 和 $G$ 时加入分类损失。训练 $G$ 的损失函数为

$$\mathcal{L}_{\text{cls}}^f = \mathbb{E}_{x, c} [-\ln D_{\text{cls}}(c | G(x, c))] \quad (6)$$

其中， $D_{\text{cls}}$ 表示由 $D$ 计算的域标签上的概率分布， $c'$ 表示源域。

最后一部分损失函数为重建损失，主要目的是保证图片风格迁移过程中将图片内容信息保存下来，实现了使用无配对数据训练网络并改善了生成图片的效果。采用L-1范数函数作为重建函数，公式为

$$\mathcal{L}_{\text{rec}} = \mathbb{E}_{x, c, c'} [\|x - G(G(x, c), c')\|_1] \quad (7)$$

所以可得生成与判别网络的总损失为

$$\mathcal{L}_G = \mathcal{L}_{\text{adv}} + \lambda_{\text{cls}} \mathcal{L}_{\text{cls}}^f + \lambda_{\text{rec}} \mathcal{L}_{\text{rec}} \quad (8)$$

$$\mathcal{L}_D = -\mathcal{L}_{\text{adv}} + \lambda_{\text{cls}} \mathcal{L}_{\text{cls}}^r \quad (9)$$

使用重新设计的损失函数和GAN网络对CycleGAN进行改进，解决了其生成图像效果较差，训练过程收敛慢的问题。图5为本文算法与经典CycleGAN训练过程的对比，可见CycleGAN++在16 h后损失值趋于稳定，达到收敛状态。而CycleGAN因网络结构复杂，经过约40 h才达到收敛。

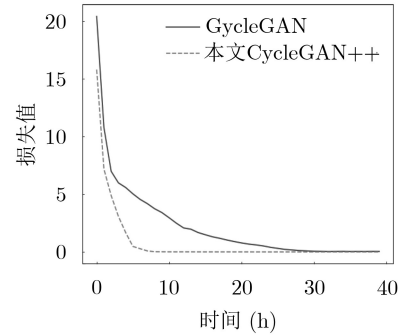


图5 CycleGAN与CycleGAN++的训练过程对比

## 4 实验

本文算法在CelebA<sup>[19]</sup>和Cityscapes数据集进行实验，CelebA包括20万张178×218的人物面部图像。将此数据集分为男女两类，实现男女性别间的风格转换。先将原始图像通过人脸中心裁剪的方式裁剪为178×178，然后缩放到128×128。随机选择2000张图片作为测试集，用剩余的所有图片作为训练集。Cityscapes数据集为图像分割数据集，包含50个城市不同场景、背景、季节的街景，共5000张图像、30类物体，分为2975张训练图，500张验证图和1525张测试图。实验方法采用AMT感知研究(Amazon Mechanical Turk perceptual studies)测试与FCN得分(Full-Convolutional Network score)进行算法性能的评估。

AMT感知研究与FCN得分为两种在GAN方向广泛使用的性能指标，在CycleGAN<sup>[12]</sup>，DualGAN<sup>[14]</sup>等论文中均使用这两种指标衡量算法的优劣。其中AMT感知研究为在亚马逊劳务平台发布“辨别图



片真伪”任务来评估输出图片的真实性,将成对的真伪图片供任务参与者观看并点击选择他们认为是真实图片的图像,最后汇总每个人的结果,得出实验数据。

虽然AMT感知研究可能是评估图像真实性的黄金准则,但是采用无需人类经验的自动质量检测方法也尤为重要。为此,采用FCN得分来进行Cityscapes数据集上的风格迁移任务评估,全卷积网络根据标准语义分割模型评估算法生成图片与输入图片的语义标签。为了评估风格迁移任务的性能,本文使用了Cityscapes数据集的标准评测指标,包括了每像素精度、每类精度、IoU分类平均值。

本文在CelebA和Cityscapes数据集上进行的实验均采用AMT感知研究进行测评,FCN得分只用于评估在Cityscapes数据集上进行的实验。

#### 4.1 实验环境与准备

实验采用浪潮NF5280M4服务器,Ubuntu 16.04系统,Nvidia GTX 1080TI 11 GB单显卡。采用 $\beta_1=0.5$ ,  $\beta_2=0.999$ 的Adam<sup>[20]</sup>优化器,批处理参数(batch size)设置为16,训练轮次(epoch)为20,在前10轮训练中的学习率为0.0001,在后10轮训练中的学习率从0.0001线性衰减为0。实验对照算法为DIAT<sup>[21]</sup>, CoGAN, IcGAN<sup>[22]</sup>, CycleGAN等4种经典图像风格迁移算法。

#### 4.2 CycleGAN++算法的实验过程

##### 4.2.1 网络结构的改进

相关结构图见图2。只改进网络结构的CycleGAN记为CycleGAN+。AMT测试,FCN得分如表1、表2所示。

##### 4.2.2 改进损失函数

因网络结构改变,取消原算法的循环一致函数,加入分类损失函数,并引入目标域信息,提高转换精度。加入改进损失函数的CycleGAN+记为CycleGAN++。测试结果如表3、表4所示。

##### 4.2.3 与各算法的对比

本文得到改进算法CycleGAN++,相比于原

表1 CycleGAN+与原算法的AMT测试结果对比(%)

| 方法        | 男性→女性    | 女性→男性    | 照片→标签    | 标签→照片    |
|-----------|----------|----------|----------|----------|
| CycleGAN  | 24.6±2.3 | 21.1±1.8 | 26.8±2.8 | 23.2±3.4 |
| CycleGAN+ | 29.5±3.2 | 29.2±4.1 | 27.8±2.2 | 28.2±2.4 |

表2 CycleGAN+与原算法的FCN得分结果对比

| 方法        | 每像素精度 | 每类精度 | IoU分类 |
|-----------|-------|------|-------|
| CycleGAN  | 0.52  | 0.17 | 0.11  |
| CycleGAN+ | 0.60  | 0.21 | 0.16  |

始CycleGAN,在AMT感知研究与FCN得分中取得更高的精度。为验证本算法的有效性,算法与其他4种经典算法进行了对比实验,如表5、表6所示。可以发现CycleGAN++相比于其他4种算法达到了较高的转换精度。可视化结果与实例对比见图6—图8。

在CelebA数据集中,图7(a)—图7(c)为男性转换为女性,图7(d)—图7(f)为女性转换为男性。

在Cityscapes数据集中,图8(a)—图8(c)为照片转换为语义标签,图8(d)—图8(f)为语义标签转换为照片。从图7,图8中可见,本文CycleGAN++在面部细节处理、人物脸型优化、背景色彩处理、语义分割准确度与清晰度等方面均优于原算法。

## 5 结束语

本文在CycleGAN基础上,提出改进算法——CycleGAN++,通过将目标域信息加入单向GAN中进行训练,提升了生成图片的转换精度。优化分类与重建损失函数,使预测值更加逼近真实值,图像生成效果更逼真。相比于经典GAN算法,如CycleGAN,

表3 CycleGAN++与CycleGAN+的AMT感知研究结果对比(%)

| 方法           | 男性→女性    | 女性→男性    | 照片→标签    | 标签→照片    |
|--------------|----------|----------|----------|----------|
| CycleGAN+    | 29.5±3.2 | 29.2±4.1 | 27.8±2.2 | 28.2±2.4 |
| 本文CycleGAN++ | 31.4±3.8 | 32.6±4.7 | 30.1±2.6 | 30.9±2.7 |

表4 CycleGAN++与CycleGAN+的FCN得分结果对比

| 方法           | 每像素精度 | 每类精度 | IoU分类 |
|--------------|-------|------|-------|
| CycleGAN+    | 0.60  | 0.21 | 0.16  |
| 本文CycleGAN++ | 0.69  | 0.27 | 0.23  |

表5 各算法的AMT感知研究结果对比(%)

| 方法                       | 男性→女性    | 女性→男性    | 照片→标签    | 标签→照片    |
|--------------------------|----------|----------|----------|----------|
| CycleGAN <sup>[12]</sup> | 24.6±2.3 | 21.1±1.8 | 26.8±2.8 | 23.2±3.4 |
| IcGAN <sup>[22]</sup>    | 23.2±2.5 | 22.4±2.9 | 22.8±2.6 | 19.8±1.9 |
| CoGAN <sup>[10]</sup>    | 6.8±1.1  | 5.1±0.9  | 0.6±0.5  | 0.9±0.5  |
| DIAT <sup>[21]</sup>     | 31.1±3.9 | 30.2±3.6 | 28.4±2.9 | 27.2±2.5 |
| 本文CycleGAN++             | 31.4±3.8 | 32.6±4.7 | 30.1±2.6 | 30.9±2.7 |

表6 各算法的FCN得分结果对比

| 方法                       | 每像素精度 | 每类精度 | IoU分类 |
|--------------------------|-------|------|-------|
| CycleGAN <sup>[12]</sup> | 0.52  | 0.17 | 0.11  |
| IcGAN <sup>[22]</sup>    | 0.43  | 0.11 | 0.07  |
| CoGAN <sup>[10]</sup>    | 0.40  | 0.10 | 0.06  |
| DIAT <sup>[21]</sup>     | 0.68  | 0.24 | 0.21  |
| 本文CycleGAN++             | 0.69  | 0.27 | 0.23  |



图6 CycleGAN++在人物性别转换领域下的可视化结果

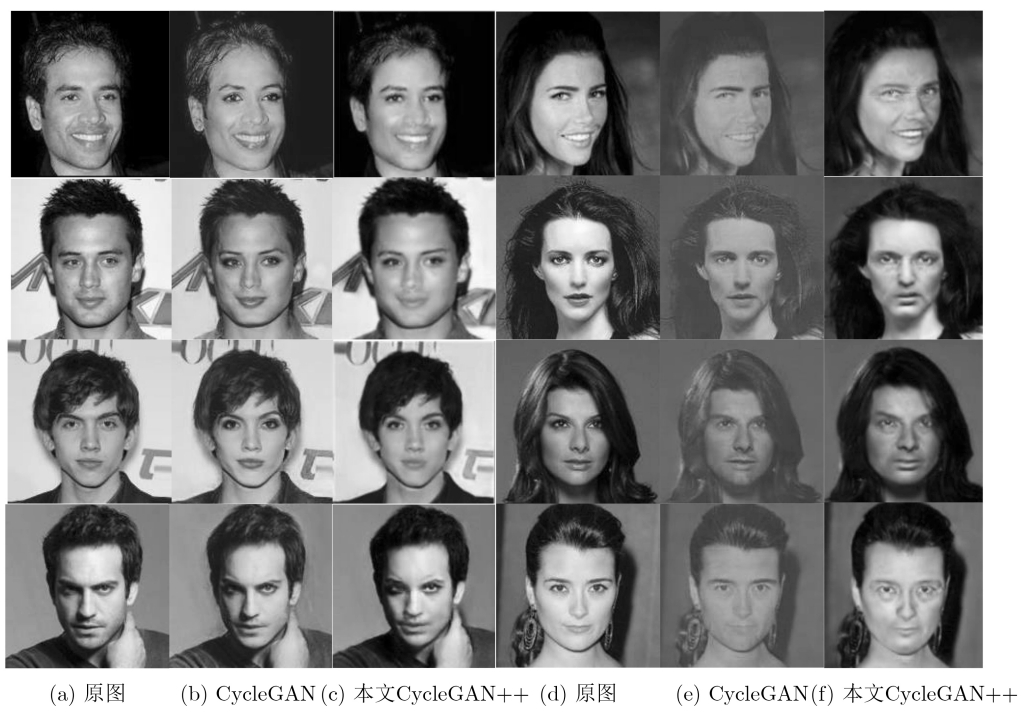


图7 CycleGAN++与原算法在CelebA测试集下的对比

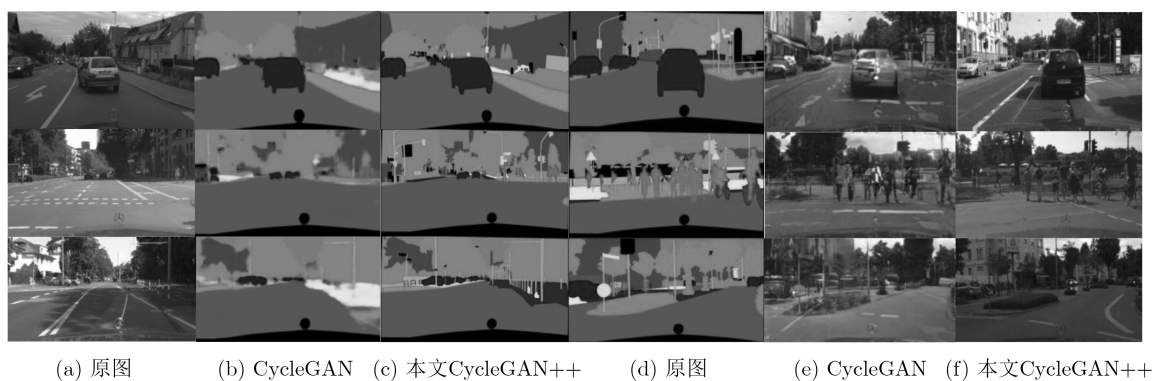


图8 CycleGAN++与原算法在Cityscapes测试集下的对比

CoGAN, DIAT等, CycleGAN++在AMT感知研究与FCN得分中取得了更逼真的生成结果与较高的转换精度, 但只能实现模型与领域一对一的风格迁移, 如何实现模型一对多领域的转换是今后需完善的工作。

### 参 考 文 献

- [1] HERTZMANN A, JACOBS C E, OLIVER N, *et al.* Image analogies[C]. The 28th Annual Conference on Computer Graphics and Interactive Techniques, New York, USA, 2001: 327–340. doi: [10.1145/383259.383295](https://doi.org/10.1145/383259.383295).
- [2] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, *et al.* Generative adversarial nets[C]. The 27th International Conference on Neural Information Processing Systems, Montreal, Canada, 2014: 2672–2680.
- [3] RADFORD A, METZ L, and CHINTALA S. Unsupervised representation learning with deep convolutional generative adversarial networks[EB/OL]. <https://arxiv.org/abs/1511.06434>, 2015.
- [4] ARJOVSKY M, CHINTALA S, and BOTTOU L. Wasserstein GAN[EB/OL]. <https://arxiv.org/abs/1701.07875>, 2017.
- [5] GULRAJANI I, AHMED F, ARJOVSKY M, *et al.* Improved training of wasserstein GANs[C]. The 31st International Conference on Neural Information Processing Systems, Red Hook, USA, 2017: 5769–5779.
- [6] ISOLA P, ZHU Junyan, ZHOU Tinghui, *et al.* Image-to-image translation with conditional adversarial networks[C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, USA, 2017: 5967–5976. doi: [10.1109/CVPR.2017.632](https://doi.org/10.1109/CVPR.2017.632).
- [7] MIRZA M and OSINDERO S. Conditional generative adversarial nets[EB/OL]. <https://arxiv.org/abs/1411.1784>, 2014.
- [8] ROSALES R, ACHAN K, and FREY B. Unsupervised image translation[C]. The 9th IEEE International Conference on Computer Vision, Nice, France, 2003: 472–478. doi: [10.1109/ICCV.2003.1238384](https://doi.org/10.1109/ICCV.2003.1238384).
- [9] LIU Mingyu, BREUEL T, KAUTZ J, *et al.* Unsupervised image-to-image translation networks[C]. The 31st Conference on Neural Information Processing Systems, Long Beach, USA, 2017: 700–708.
- [10] LIU Mingyu and TUZEL O. Coupled generative adversarial networks[C]. The 30th Conference on Neural Information Processing Systems, Barcelona, Spain, 2016: 469–477.
- [11] KINGMA D P and WELLING M. Auto-encoding variational bayes[EB/OL]. <https://arxiv.org/abs/1312.6114>, 2013.
- [12] ZHU Junyan, PARK T, ISOLA P, *et al.* Unpaired image-to-image translation using cycle-consistent adversarial networks[C]. 2017 IEEE International Conference on Computer Vision, Venice, Italy, 2017: 2242–2251. doi: [10.1109/ICCV.2017.244](https://doi.org/10.1109/ICCV.2017.244).
- [13] KIM T, CHA M, KIM H, *et al.* Learning to discover cross-domain relations with generative adversarial networks[C]. The 34th International Conference on Machine Learning, Sydney, Australia, 2017: 1857–1865.
- [14] YI Zili, ZHANG Hao, TAN Ping, *et al.* DualGAN: Unsupervised dual learning for image-to-image translation[C]. 2017 IEEE International Conference on Computer Vision, Venice, Italy, 2017: 2868–2876. doi: [10.1109/ICCV.2017.310](https://doi.org/10.1109/ICCV.2017.310).
- [15] BOUSMALIS K, SILBERMAN N, DOHAN D, *et al.* Unsupervised pixel-level domain adaptation with generative adversarial networks[C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, USA, 2017: 95–104. doi: [10.1109/CVPR.2017.18](https://doi.org/10.1109/CVPR.2017.18).
- [16] SHRIVASTAVA A, PFISTER T, TUZEL O, *et al.* Learning from simulated and unsupervised images through adversarial training[C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, USA, 2017: 2242–2251. doi: [10.1109/CVPR.2017.241](https://doi.org/10.1109/CVPR.2017.241).
- [17] TAIGMAN Y, POLYAK A, and Wolf L. Unsupervised cross-domain image generation[EB/OL]. <https://arxiv.org/abs/1611.02200>, 2016.
- [18] LI Chuan and WAND M. Precomputed real-time texture synthesis with markovian generative adversarial networks[C]. The 14th European Conference on Computer Vision, Amsterdam, The Netherlands, 2016: 702–716. doi: [10.1007/978-3-319-46487-9\\_43](https://doi.org/10.1007/978-3-319-46487-9_43).
- [19] LIU Ziwei, LUO Ping, WANG Xiaogang, *et al.* Deep learning face attributes in the wild[C]. 2015 IEEE International Conference on Computer Vision, Santiago, Chile, 2015: 3730–3738. doi: [10.1109/ICCV.2015.425](https://doi.org/10.1109/ICCV.2015.425).
- [20] KINGMA D P and BA J. Adam: A method for stochastic optimization[EB/OL]. <https://arxiv.org/abs/1412.6980>, 2014.
- [21] LI Mu, ZUO Wangmeng, and ZHANG D. Deep identity-aware transfer of facial attributes[EB/OL]. <https://arxiv.org/abs/1610.05586>, 2016.
- [22] PERARNAU G, VAN DE WEIJER J, RADUCANU B, *et al.* Invertible conditional GANs for image editing[EB/OL]. <https://arxiv.org/abs/1611.06355>, 2016.

张惊雷: 男, 1969, 教授, 博士, 研究方向为模式识别、图像处理等。  
厚雅伟: 男, 1995, 硕士生, 研究方向为图像处理、目标检测等。