

改进生成对抗网络在场景图像转换中的应用^①



金 阳¹, 何利力²

¹(浙江理工大学 智能无人系统软件技术与应用重点实验室, 杭州 310018)

²(浙江理工大学 浙江省服装个性化定制 2011 协同创新中心, 杭州 310018)

通讯作者: 金 阳, E-mail: xinyangs.ai@gmail.com

摘 要: 本文针对不同场景图像之间的转换问题, 提出了一种改进的生成对抗网络模型, 能够生成高质量的目标场景图像. 在生成目标图像过程中存在因为向下采样而丢失原图像空间位置信息的现象, 因此本文设计了一个包含跳跃连接和残差块的生成网络, 通过在网络中加入多个跳跃连接部分, 将图像的空间位置信息在网络中保持传递. 同时为提高训练过程中生成图像在结构上的稳定性, 引入 SSIM 图像结构相似指数, 作为结构重建损失, 以指导模型生成更优结构的目标图像. 此外, 为使得转换后的目标场景图像保留更多的色彩细节, 加入了身份保持损失, 明显增强了目标生成图像的色彩表现力. 实验结果表明, 本文所提的改进生成对抗网络模型能够在场景图像转换中得到有效地应用.

关键词: 生成对抗网络; 场景图像转换; 结构重建损失; 残差块

引用格式: 金阳, 何利力. 改进生成对抗网络在场景图像转换中的应用. 计算机系统应用, 2021, 30(3): 151-157. <http://www.c-s-a.org.cn/1003-3254/7816.html>

Application of Improved Generative Adversarial Network in Scene Image Translation

JIN Yang¹, HE Li-Li²

¹(Key Laboratory of Software Technology and Application of Intelligent Unmanned System, Zhejiang Sci-Tech University, Hangzhou 310018, China)

²(2011 Collaborative Innovation Center for Garment Personalized Customization of Zhejiang Province, Zhejiang Sci-Tech University, Hangzhou 310018, China)

Abstract: Aiming at the translation between different scene images, we propose an improved generative adversarial network model that can generate high-quality target scene images. In the process of generating a target image, the spatial position information of the original image will lose due to down sampling. Therefore, a generative network that includes jump connections and residual blocks is designed in this paper. By adding multiple jump connections to the network, we can keep the spatial position information of the image transmitting in the network. At the same time, to improve the stability of the generated image during the training, we introduce the Structural Similarity Index Measure (SSIM) as a structure reconstruction loss to guide the model to generate a target image with a better structure. In addition, in order to make the translated target scene image retain more color details, we add an identity preservation loss, obviously enhancing the color expressiveness of the target generated image. The experimental results show that the improved generative adversarial network model proposed in this study can be effectively applied in scene image translation.

Key words: Generative Adversarial Network (GAN); scene image translation; structure reconstruction loss; residual block

① 基金项目: 国家重点研发计划 (2018YFB1700702)

Foundation item: National Key Research and Development Program of China (2018YFB1700702)

收稿时间: 2020-07-18; 修改时间: 2020-08-13; 采用时间: 2020-08-17; csa 在线出版时间: 2021-03-03

1 引言

近几年来,生成对抗网络(Generative Adversarial Network, GAN)在计算机视觉、图像处理等多个领域得到越来越多的应用.其中包括应用最广泛的图像到图像的转换,它可以将来自不同域的图像进行相互之间转换.场景图像转换可以看成是图形处理和计算机视觉领域相结合的问题.传统的场景图像转换算法大多是基于模型的方法.孙毅刚等^[1]提出一种基于改进的snake模型,该算法利用模糊连接度准确找到雨雪轮廓点,并通过模糊相似度函数确定雨雪边界,实现雨雪场景图像转换.刘杰平等^[2]基于大气散射模型,提出了一种适应人眼视觉特征的HSI色彩空间估计大气光和透射率,实现雾霾场景图像转换.费子豪等^[3]建立了基于高频分量直方图均衡化,低频分量同态滤波处理的小波变换模型,完成晴天多云场景图像转换.

场景图像转换是图像转换的一类问题.解决图像转换问题大多使用基于学习的方法,利用卷积神经网络的强大特征表示能力,学习源域到目标域图像的映射函数.为了学习到这个映射函数,许多国内外学者展开了广泛的科学研究.得益于文献[4],其将生成对抗网络(GAN)引入到深度学习领域,给生成图像工作带来了更加新颖的方法.基于GAN网络模型,提出各种改进的图像转换模型. Mirza等^[5]提出了带条件约束的cGAN(conditional GAN)模型,该网络模型通过添加条件变量限制生成结果.基于cGAN的思想, Isola等^[6]提出Pix2Pix算法模型,该算法要求配对的图像数据集,能够无监督地学习到不同域图像之间的映射关系. Kim等^[7]提出DiscoGAN模型在跨域图像转换中发挥不错的效果. Yi等^[8]借鉴对偶学习的思想提出DualGAN. Radford等^[9]提出DCGAN,极大的提高了图像生成的实用性. Zhu等^[10]提出了CycleGAN网络,该方法突破了要求成对图像数据集的限制,无需配对数据集就能够实现不同域图像之间的转换工作.

本文正是基于CycleGAN网络,提出了一种场景图像转换的新方法.通过改进的生成网络结构、SSIM结构重建损失和身份保持损失,应用到场景图像转换上能够取得很好的效果.

2 相关工作

2.1 生成对抗网络(GAN)

生成对抗网络(GAN)是一个十分强大的生成模型,它能够在各种不同的计算机视觉任务中取得优秀

的结果. GAN基于连续数据,且假设有无限建模能力^[11].其灵感来源于零和博弈思想,通过双方的博弈过程,达到两者之间纳什平衡.其由两种网络结构模型组成:生成网络 G (Generator)和判别网络 D (Discriminator).判别网络 D 用来判断样本图像是原始图像还是由生成网络 G 生成的图像,目的是将他们区分开来.而生成网络 G 则尽可能的生成更加逼真的图像,以此来迷惑判别网络 D ,达到以假乱真的目的.通过 G 和 D 之间相互对抗,直到生成的图像判别网络无法分辨.

GAN的优化目标函数为:

$$\min_G \max_D V(G, D) = E_{x \sim p_{\text{data}}(x)} [\log D(x)] + E_{z \sim p_z(z)} [\log(1 - D(G(z)))] \quad (1)$$

其中, x 表示真实图片,采样于真实数据分布, z 表示输入生成网络 G 的随机噪声,采样于正态分布. $D(x)$ 表示输入的图片 x 接近真实图片的概率.训练是批量进行的, E 为数学期望. D 希望 $D(x)$ 变大, $D(G(z))$ 变小,所以 D 尽可能的最大化目标函数. G 希望 $D(G(z))$ 变大,这时 G 尽可能的最小化目标函数.

2.2 CycleGAN模型

CycleGAN模型工作在无配对的数据集上,实现跨域图像转换.其网络结构由两个生成网络 G_{XY} 、 G_{YX} 和两个判别网络 D_X 、 D_Y 组成. G 与 D 相互对抗过程类似于一种双向环结构.其整体结构如图1所示.

图1中,左边部分是 G_{XY} 与 D_Y 、 G_{YX} 与 D_X 之间的对抗.中间部分是将 X 域中的图像经过两次转换回原域后反馈调节网络参数,右边部分与之类似,作用于 Y 域中的图像.

该模型的目标函数由两部分损失函数组成,对抗损失和循环一致损失.对抗损失表示生成网络和判别网络之间的对抗,生成网络尽可能生成接近真实样本分布的图像,而判别网络尽可能将他们分辨出来.通过网络之间的对抗,模型可以学习到很好的映射函数 G_{XY} 和 G_{YX} ,然而单靠对抗损失还不能够保证学习函数能够将单个输入 x_i 映射到期望的输出 y_i ,因此需要循环一致损失缩小映射函数空间.循环一致损失类似于对偶学习,它通过前后两次转换将输入 x 映射回原始域,例如 $x \rightarrow G_{XY}(x) \rightarrow G_{YX}(G_{XY}(x)) \approx x$,然后根据反馈回来的信息就能指导模型优化改进.

2.3 跳跃连接和残差块

跳跃连接(skip connections)^[12]常用在生成网

络结构中, 以保证图像空间位置信息在网络传递中的连续性. 通常生成网络包含下采样部分和上采样部分, 在下采样中, 输入是原始图像, 每经过一次下采样, 图像都会在空间上被压缩, 同时扩张图像的通

道数. 相反, 上采样中, 图像在空间上拓展, 相应的通道数减少. 这样的网络是线性的, 图像数据从输入层流入, 通过网络一层接一层的向下流动, 最后到达输出层.

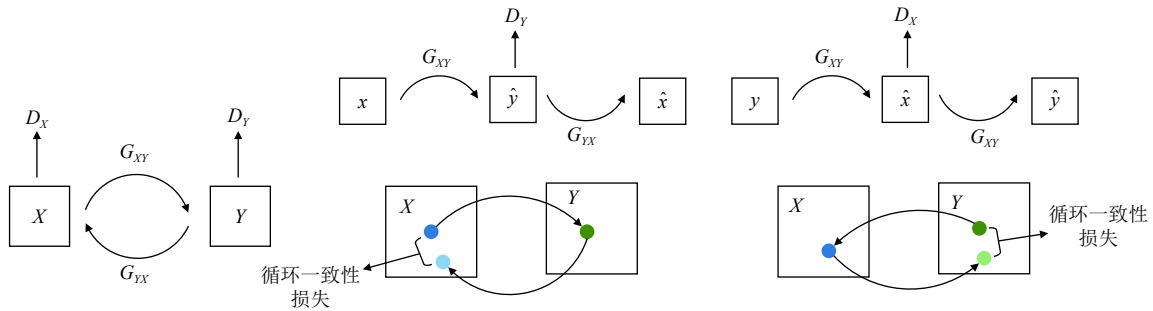


图1 CycleGAN 对抗过程整体结构图

在下采样部分的网络中, 经过每一个连续的层, 模型都会提取到图像特征内容, 但却丢失了图像空间位置信息. 在下采样的最后一层, 特征图已经有足够强的能力学习到图像的内容, 但几乎不理解空间位置. 对于预测分类问题, 这已经足够了, 所以我们可以把它连接到一个稠密层上, 来输出一个特定类出现在图像中的概率. 然而, 对于图像生成或者图像转换问题, 需要在上采样回到原始图像大小过程中将下采样丢失的空间位置信息传递回每一层. 这就是需要跳跃连接, 它允许网络将在下采样过程中捕获的高级抽象信息与从网络的先前层反馈来的特定空间信息进行混合.

为了让网络提取到的特征更加丰富, 避免在非采样层中, 过深的网络进行卷积和池化操作时导致的图像信息丢失, 我们也引入了残差块 (residual block)^[13]. 其结构如图2所示.

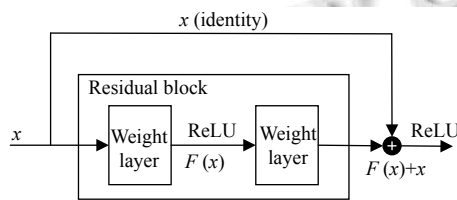


图2 残差块结构

残差块通过跳跃连接将浅层特征直接传递到深层特征, 其工作方式是将输入与残差块的网络输出叠加后作为最后的输出. 通过跳跃连接将图像特征向神经网络的更深层传递, 即使网络结构更深, 也不会出现特征丢失的现象. 同时还有有效的解决了梯度消失和梯度爆炸问题, 提高了网络的稳定性.

3 改进的网络模型

本文从网络结构和损失函数两个方面改进网络模型.

3.1 判别网络结构

本文设计了一个全卷积多层下采样网络结构, 通过提取场景图像特征, 将真假图像辨别出来. 所提的判别网络包含两类 D_X 和 D_Y , D_X 用来判别输入的图像是否来自 X 域中的真实图像, D_Y 将 Y 域中的真实图像 y 与生成的图像区分开来. 它们共用同一网络结构, 但模型参数不共享. 判别网络结构如图3所示.

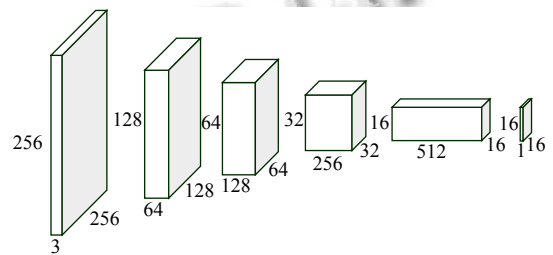


图3 判别网络结构

网络的输入为3通道大小为的图片, 随后通过边下采样边拓展通道的方式, 提取图像内容特征. 采用步幅为2, 卷积核大小为 3×3 的卷积操作, 每次卷积后图像宽高都减小一半, 除了输入层和输出层外, 其他层还需要连接 Instance Normalization 实例归一化层以及紧跟着 Leaky ReLU 激活层. 最后通过一个卷积核为1的卷积操作输出形状为 $1 \times 16 \times 16$ 的特征图, 对其求取平均值作为判别网络的输出, 也就是输入图片判别为真实图像的概率.

3.2 生成网络结构

仅用卷积层提取特征存在缺陷, 没有考虑输入图像与生成图像在纹理特征和图像结构上的联系, 造成生成的图像细节模糊. 为了弥补该缺陷, 本文改进生成网络结构, 通过引入跳跃连接, 将下采样过程中丢失的图像空间结构信息传递回上采样中, 以及在加深的网络层中加入残差块进一步抑制图像纹理特征信息在传递过程中的丢失现象.

所提的生成网络同样包含两类 $G_{XY}: X \rightarrow Y$ 和 G_{YX} 生成网络将 X 域中图像 x 作为输入, 尽可能去生成接近 Y 域数据样本分布的图像 \hat{y} . G_{YX} 则是与之相反, 将 Y 域中的图像尽可能 $G_{YX}: Y \rightarrow X$ 能转化为 X 域中图像, 它们的网络结构如图4所示.

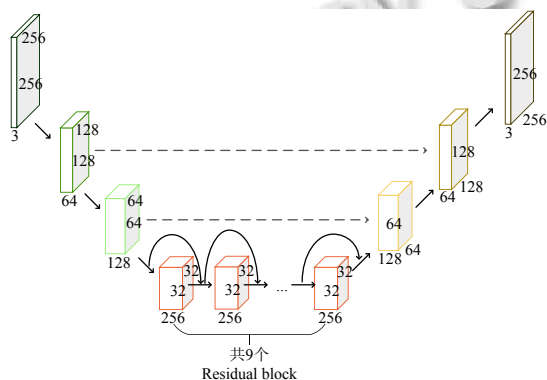


图4 生成网络结构

该网络输入输出图像均为3通道大小为 256×256 的图片. 模型左侧为下采样结构, 采用步幅为2卷积核为 5×5 的卷积层, 同时紧跟其后的是实例归一化层以及Leaky ReLU层. 在这里我们参考Ulyanov等^[14]的建议, 使用实例归一化层(instance normalization), 在生成网络中使用实例归一化层比批量归一化层^[15]更能生成令人满意的结果. 总共进行3次下采样, 每次图像宽高减半, 通道数加倍, 最终到达 $256 \times 32 \times 32$ 尺寸为止. 随后保持该尺寸经过9个残差块传递图像特征, 对应着模型底部的结构. 之后, 进行向上采样回原图像尺寸, 这部分属于模型右侧的结构, 其采用相同的步幅2、卷积核 5×5 的反卷积层, 以及Instance Normalization层和Leaky ReLU层, 每次上采样图像大小加倍, 通道数减半, 不同的是使用跳跃连接将下采样过程中对应大小尺寸的特征图连接过来. 最后, 将上采样的结果通过

通道数为3的反卷积层连接, 输出相同尺寸的图像.

3.3 损失函数

损失函数即本文所提模型的目标函数, 它是由对抗损失、SSIM重建损失和身份保持损失3个部分组成.

3.3.1 对抗损失

在原始的GAN模型中, 网络的训练过程存在着不稳定和生成图像质量不高等问题, 针对该问题, 我们将交叉熵损失替换为最小二乘损失^[16]. 对应的生成网络 G_{XY} 和判别网络 D_Y 的对抗损失定义如下:

$$L_{\text{gan}}(G_{XY}, D_Y) = \frac{1}{2} E_{y \sim p_{\text{data}}(y)} [(D_Y(y) - 1)^2] + \frac{1}{2} E_{x \sim p_{\text{data}}(x)} [(D_Y(G_{XY}(x)))^2] \quad (2)$$

式中, $y \sim p_{\text{data}}(y)$ 表示变量 y 采样于 Y 域空间服从的概率分布, $x \sim p_{\text{data}}(x)$ 表示变量 x 采样于 X 域空间服从的概率分布, $E[\cdot]$ 表示计算数学期望值. 对于生成网络 G_{XY} , 其目标是尽可能最大化 $L_{\text{gan}}(G_{XY}, D_Y)$, 即生成更接近目标域 Y 数据分布的图像来混淆判别网络 D_Y 的决策. 而判别网络 D_Y 则尽可能最小化目标, 将真实样本 y 和生成样本 \hat{y} 区分开.

$$G_{XY}^*, D_Y^* = \arg \min_{D_Y} \max_{G_{XY}} L_{\text{gan}}(G_{XY}, D_Y) \quad (3)$$

如式(3)所示, 通过两者对抗最终达到最优解 G_{XY}^* 和 D_Y^* .

同样的, 对于生成网络 G_{YX} 和判别网络 D_X 的对抗损失定义如式(4), 通过两者之间的对抗过程寻找到最优解 G_{YX}^*, D_X^* , 如式(5)所示.

$$L_{\text{gan}}(G_{YX}, D_X) = \frac{1}{2} E_{x \sim p_{\text{data}}(x)} [(D_X(x) - 1)^2] + \frac{1}{2} E_{y \sim p_{\text{data}}(y)} [(D_X(G_{YX}(y)))^2] \quad (4)$$

$$G_{YX}^*, D_X^* = \arg \min_{D_X} \max_{G_{YX}} L_{\text{gan}}(G_{YX}, D_X) \quad (5)$$

3.3.2 SSIM重建损失

通过优化上述对抗损失函数可以学习到很好的映射函数 G_{XY}^* 和 G_{YX}^* , 产生接近目标域空间分布的样本. 如果有足够大的样本容量, 生成网络可以将同一组输入图像映射到目标域中任意随机排列的图像, 对于 G_{XY}^* 生成网络, 它完全可以将所有的 X 域图像都映射到目标 Y 域的另一图像上, 从而导致对抗损失失效. 因此, 本文引入了重建损失来帮助网络进一步缩小无效

的映射函数空间, 确保映射函数能够将单个输入映射到期望的输出。

同时, 为进一步保证生成图像过程中结构特性不丢失, 引入 Wang 等^[17]提出的结构相似性指数 (Structural Similarity Index Measurement, SSIM) 来衡量两幅图像的结构差异。

对于给定的图像 x 和 y , 其结构相似性指数定义如下:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (6)$$

$$c_1 = (k_1L)^2, c_2 = (k_2L)^2 \quad (7)$$

式中, μ_x 是图像 x 的均值, σ_x^2 是图像 x 的方差, σ_{xy} 是图像 x 和 y 的协方差。 c_1 与 c_2 为两个常数, 避免除零, L 为像素范围, $k_1 = 0.01, k_2 = 0.03$ 为默认值。 SSIM 结构相似性范围为-1 到 1, 值越大表示图像 x 和 y 越相似。

因此, 我们引入 SSIM 重建损失, 如式 (8) 所示, 模型尽可能的最小化该目标。

$$\begin{aligned} L_{ssim+rec}(G_{XY}, G_{YX}) &= E_{x \sim p_{data}(x)} [1 - SSIM(x, G_{YX}(G_{XY}(x)))] \\ &+ E_{y \sim p_{data}(y)} [1 - SSIM(y, G_{XY}(G_{YX}(y)))] \end{aligned} \quad (8)$$

3.3.3 身份保持损失

为了让生成网络将其目标域的图像作为输入后保持不变性, 我们引入了身份保持损失。 存在这样一种情况, 将 Y 域中的图像 y 输入到 G_{XY} 中, 满意的结果应该是图像 y 原封不动的从生成网络中输出。 身份保持损失能够帮助我们做到这一点, 该项损失帮助调节生成网络, 以确保它只调整完成转换所需的图像部分, 而不是更多的内容。 在保持图像内容前后一致性的同时, 也间接地维持了图像的色彩饱和度。 身份保持损失公式如式 (9) 所示。

$$\begin{aligned} L_{id}(G_{XY}, G_{YX}) &= E_{y \sim p_{data}(y)} [\|G_{XY}(y) - y\|_1] \\ &+ E_{x \sim p_{data}(x)} [\|G_{YX}(x) - x\|_1] \end{aligned} \quad (9)$$

模型通过最小化目标函数 $L_{id}(G_{XY}, G_{YX})$, 保证转换图像前后内容的一致性, 同时也避免转换后的图像出现色彩失真现象, 以保留更多的色彩细节。

综合以上三项损失函数, 得到最终的目标函数为 $L(G_{XY}, G_{YX}, D_X, D_Y)$ 。 如式 (10) 所示。

$$\begin{aligned} L(G_{XY}, G_{YX}, D_X, D_Y) &= L_{gan}(G_{XY}, D_Y) + L_{gan}(G_{YX}, D_X) \\ &+ \lambda_1 L_{ssim+rec}(G_{XY}, G_{YX}) + \lambda_2 L_{id}(G_{XY}, G_{YX}) \end{aligned} \quad (10)$$

其中, λ_1 和 λ_2 分别控制 3 项损失之间的相对重要性。

4 实验结果与分析

4.1 实验环境和数据集

本文算法进行实验的平台环境及配置如表 1 所示。 选用 Tensorflow 作为深度学习框架, 在 GPU 上使用 CUDNN 加速训练模型, 使用的编程语言为 Python。

表 1 实验环境配置

设备名称	设备信息
CPU	Intel(R) Xeon(R) Gold 6148
GPU	Tesla V100 SXM2 16 GB
操作系统	Ubuntu 14.04 LTS 64位
CUDA版本	10.1.0
CUDNN版本	7.6.5
TensorFlow版本	2.0.0
Python版本	3.7.0

数据集包括 4 个类别的场景图像, 两类季节场景 summer、winter 和两类天气场景 cloudy、sunny。 数据分别选自 MWD (Multi-class Weather Dataset) 数据集、Two-class Weather Classification 数据集, 组成我们需要的两类场景转换数据集, 数据集一为多云与晴天场景转换数据集 (cloudy2sunny), 数据集二为夏季与冬季场景转换数据集 (summer2winter)。 每个数据集中每类场景各自包含训练数据 5000 张图片, 测试数据 1000 张图片, 图片大小都为 256×256。

4.2 训练过程及参数设置

基于 TensorFlow 深度学习框架, 使用 GPU 加速, 采用迭代交替的方式对本文算法模型进行训练。 优化算法使用 Adam, 批量大小为 1, λ_1 、 λ_2 分别取 10 和 5, 所有的网络从学习率为 0.0002 开始训练, 前 100 个 epochs 保持 0.0002 的学习率不变, 后 100 个 epochs 中学习率线性衰减为 0, 交替迭代训练生成网络和判别网络。 模型参数如表 2 所示。

表 2 网络模型参数

参数	值
epoch	200
offset	100
batch_size	1
learning_rate(η^d, η^g)	0.0002
λ_1	10
λ_2	5

4.3 对比实验验证

为验证所提模型的有效性, 分别进行两类场景图像转换对比实验。 第一类场景为天气场景图像转换 (多

云与晴天互转), 首先分别选取不同域的图片, 输入到对应的生成网络中, 得到转换后的图像结果. 然后将实验结果与 GAN、Pix2Pix、CycleGAN 模型得到的结果进行定性和定量对比. 第二类场景为季节场景图像转换 (夏季与冬季互转), 验证过程与上述类似.

(1) 定性对比实验

使用数据集一 (cloudy2sunny) 训练本文所提模型.

分别选取多云图片和晴天图片各一张进行对比实验, 得到如图 5 所示结果. 从实验结果中可以发现, GAN 模型转换后的晴天和多云场景图像模糊不清. Pix2Pix 模型生成的目标场景图像部分存在结构上的缺陷问题. CycleGAN 模型转换后的场景图像容易造成图像色彩失真, 与真实的图像存在差异. 本文算法生成的目标场景图像 (多云和晴天) 更加的接近真实场景图像.

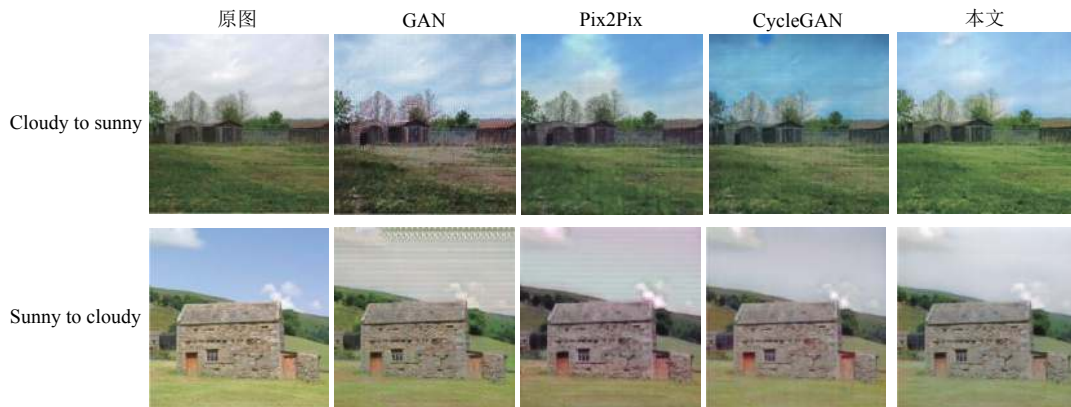


图 5 天气场景图像转换实验结果对比图

使用数据集二 (summer2winter) 时, 同样选取夏季和冬季图片各一张, 将实验结果图与 GAN、Pix2Pix、CycleGAN 模型得到的结果进行对比, 如图 6 所示. 通过对比可以看出, 所提模型转换后的结果更加接近目

标真实场景, 同时保持了完整的图像结构和纹理细节信息. 通过这两类场景图像转换的对比实验, 验证了本文所提模型应用到场景图像转换上的有效性和可行性.

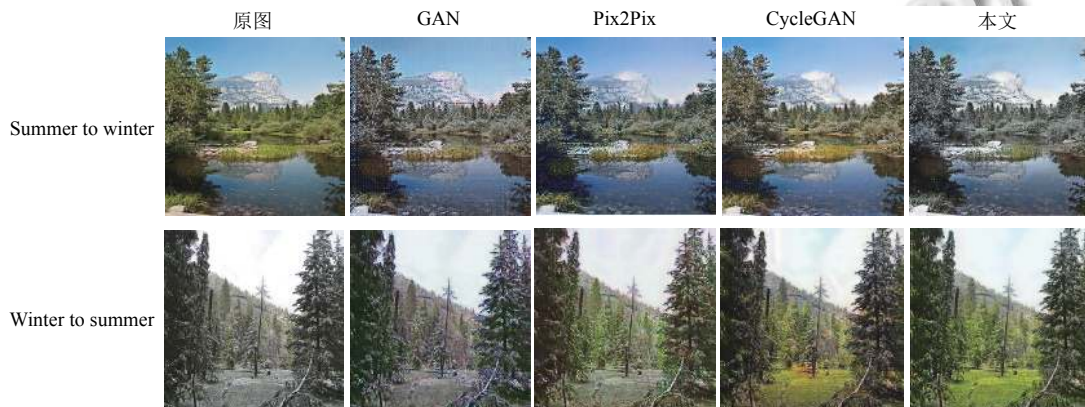


图 6 季节场景图像转换实验结果对比图

(2) 定量对比实验

定量对比部分, 我们采用广泛使用的 PSNR^[18] 和 SSIM 这两种全参考的图像质量评价指标, 来对生成的图像进行客观的定量评判. 在同一测试数据集上, 使用这两种评价指标, 将本文算法和其他算法生成的图像与真实图像进行定量对比. 各算法的评价指标结果如表 3 所示.

表 3 测试数据集上各算法评价指标结果

算法	PSNR	SSIM
GAN	28.8914	0.7068
Pix2Pix	29.6726	0.7892
CycleGAN	29.8928	0.8132
本文	30.0114	0.8201

5 结束语

本文设计了基于改进生成对抗网络的图像转换算法,应用在天气和季节场景的图像转换上能够获得很好的效果.具体改进方法涉及以下几个方面:首先是改进生成网络的网络结构,通过在网络的下采样部分和上采样部分中间加入跳跃连接,保证下采样部分的图像空间位置信息传递的连续性,以及在非采样部分使用残差块,保证层数加深的网络不丢失图像内容信息.其次,选用 SSIM 相似性指数作为衡量图像结构差异的指标,组成 SSIM 结构重建损失,目的是使得模型生成的图像结构更完整.此外,还通过添加身份保持损失,保证转换前后图像色彩不失真和内容的一致性.在今后的研究工作中,我们将考虑进一步优化网络结构和损失函数,在纹理、细节等方面,增强原图和转换图之间的结构相似性和语义一致性.

参考文献

- 1 孙毅刚,段晓晔,张红颖,等.基于改进 snake 模型的图像中雨雪去除算法研究.计算机应用研究,2011,28(5):1991-1993.[doi:10.3969/j.issn.1001-3695.2011.05.113]
- 2 刘杰平,黄炳坤,韦岗.一种快速的单幅图像去雾算法.电子学报,2017,45(8):1896-1901.[doi:10.3969/j.issn.0372-2112.2017.08.013]
- 3 费子豪,刘勇,李虎,等.基于均衡化与同态滤波的光学图像去云模型.科技创新导报,2019,16(36):266-272.
- 4 Goodfellow IJ, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets. Proceedings of the 27th International Conference on Neural Information Processing Systems. Montreal, QB, Canada. 2014. 2672-2680.
- 5 Mirza M, Osindero S. Conditional generative adversarial nets. arXiv: 1411.1784, 2014.
- 6 Isola P, Zhu JY, Zhou TH, et al. Image-to-image translation with conditional adversarial networks. Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA. 2017. 5967-5976.
- 7 Kim T, Cha M, Kim H, et al. Learning to discover cross-domain relations with generative adversarial networks. Proceedings of the 34th International Conference on Machine Learning. Sydney, NSW, Australia. 2017. 1857-1865.
- 8 Yi ZL, Zhang H, Tan P, et al. Dualgan: Unsupervised dual learning for image-to-image translation. Proceedings of 2017 IEEE International Conference on Computer Vision. Venice, Italy. 2017. 2868-2876.
- 9 Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv: 1511.06434, 2015.
- 10 Zhu JY, Park T, Isola P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks. Proceedings of 2017 IEEE International Conference on Computer Vision. Venice, Italy. 2017. 2242-2251.
- 11 Zou XF, Zhu DJ. Review on generative adversarial network. Computer Systems & Applications, 2019, 28(11): 1-9.
- 12 Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. Proceedings of the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention. Munich, Germany. 2015. 234-241.
- 13 He KM, Zhang XY, Ren SQ, et al. Deep residual learning for image recognition. Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA. 2016. 770-778.
- 14 Ulyanov D, Vedaldi A, Lempitsky V. Instance normalization: The missing ingredient for fast stylization. arXiv: 1607.08022, 2016.
- 15 Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. Proceedings of the 32nd International Conference on Machine Learning. Lille, France. 2015. 448-456.
- 16 Mao XD, Li Q, Xie HR, et al. Multi-class generative adversarial networks with the L2 loss function. arXiv: 1611.04076, 2016.
- 17 Wang Z, Bovik AC, Sheikh HR, et al. Image quality assessment: From error visibility to structural similarity. IEEE Transactions on Image Processing, 2004, 13(4): 600-612. [doi:10.1109/TIP.2003.819861]
- 18 Huynh-Thu Q, Ghanbari M. Scope of validity of PSNR in image/video quality assessment. Electronics Letters, 2008, 44(13): 800-801.