

# 不同纹理复杂度图像生成中 CycleGAN 循环一致损失系数优化选择方法

徐 强 钟尚平 陈开志 张春阳

(福州大学数学与计算机科学学院 福州 350116)

(福州大学网络系统信息安全福建省高校重点实验室 福州 350116)

**摘 要** 高质量的图像生成一直是计算机视觉等领域探索的难点和热点。通过使用循环一致损失, CycleGAN 在无监督图像生成任务中取得了良好效果。但是面对不同纹理复杂度的图像生成任务, CycleGAN 的循环一致损失系数是默认不变的, 使得生成图像存在纹理变形甚至消失等弱点, 不能很好地保证生成图像的质量。文中融合图像的空间维度和时间维度来度量图像的纹理复杂性, 阐明循环一致损失函数在优化目标函数中的重要性, 发现并解释循环一致损失系数的大小与不同纹理复杂度图像生成质量的关联性: 纹理复杂度越高, 应选择越大的循环一致损失系数; 反之, 应取越小的循环一致损失系数。文中使用基准和自采集的图像数据集, 引入了基于迁移学习的分类准确性等生成图像质量评估指标。实验结果表明, 优化选择大小合适的循环一致损失系数, 可有效提高生成图像的质量。

**关键词** 图像生成, CycleGAN, 优化选择系数, 循环一致损失, 纹理复杂度

中图分类号 TP183 文献标识码 A DOI 10.11896/j.issn.1002-137X.2019.01.015

## Optimized Selection Method of Cycle-consistent Loss Coefficient of CycleGAN in Image Generation with Different Texture Complexity

XU Qiang ZHONG Shang-ping CHEN Kai-zhi ZHANG Chun-yang

(College of Mathematics and Computer Science, Fuzhou University, Fuzhou 350116, China)

(Network System Information Security Fujian Provincial University Key Laboratory, Fuzhou University, Fuzhou 350116, China)

**Abstract** High-quality image generation has always been a difficult and hot topic in the field of computer vision and other exploration. CycleGAN achieves good results in unsupervised image generation tasks by using cycle-consistent losses. However, in face of image generation tasks with different texture complexity, CycleGAN's cycle-consistent loss coefficient is unchanged by default, and its generated images have weak points such as texture distortion or even disappear, which can not guarantee the quality of generated images. In this paper, the complexity of image texture was measured by integrating the spatial dimension and time dimension of images, the importance of cycle-consistent loss function in optimizing objective function was clarified, the correlation between the size of the cycle-consistent loss coefficient and the quality of image with different texture complexity was discovered and explained. The higher the texture complexity, the larger the cycle-consistent loss coefficient should be selected. Otherwise, the smaller coefficient should be taken. Using benchmarks and self-acquired image data sets, the classification accuracy based on migration learning was introduced to generate image quality assessment indicators. The experimental results show that the optimal choice of the appropriate cycle-consistent loss factor can effectively improve the quality of generated images.

**Keywords** Image generation, CycleGAN, Optimization of selection coefficient, Cycle-consistent loss, Texture complexity

## 1 引言

许多图像处理等计算机视觉任务, 例如图像分割<sup>[1]</sup>、图像风格迁移<sup>[2-4]</sup>等, 都可以归结成图像生成问题。建立在非合作博弈<sup>[5]</sup>基础上的生成对抗网络(Generative Adversarial Networks, GANs)<sup>[6]</sup>为这类任务提供了强有力的支持, 被广泛研

究并取得了令人印象深刻的良好结果<sup>[7]</sup>。鉴于经典 GANs 存在模型崩溃、难以平衡训练等明显弱点, 目前业界提出了众多的 GANs 改进算法, 致力于解决不同场景的任务<sup>[8-10]</sup>。这些 GANs 改进算法目前主要基于有监督和无监督两种学习模式。

基于有监督模式构建的 GANs 包含如下典型算法: Mirza 等<sup>[11]</sup>提出的 cGAN 通过在生成器和鉴别器中加入条件变量,

到稿日期: 2018-05-11 返修日期: 2018-07-25

徐 强(1993—), 男, 硕士生, 主要研究方向为深度学习、图像处理, E-mail: happyxuwork@163.com; 钟尚平(1969—), 男, 教授, 硕士生导师, 主要研究方向为机器学习、模式识别、大规模核学习, E-mail: spzhong@fzu.edu.cn(通信作者); 陈开志(1983—), 男, 讲师, 硕士生导师, 主要研究方向为生物特征识别、图像安全; 张春阳(1987—), 男, 副教授, 硕士生导师, 主要研究方向为深度学习、大数据与云计算。

对生成器的数据生成发挥指导作用;Zhu 等<sup>[12]</sup>提出的 iGAN 为了使得生成图像纹理更加清晰,引入了两种指导信息,即用户的输入信息和差值空间中的光场信息;Isola 等<sup>[13]</sup>提出的 pix2pix 将鉴别器的判别任务变为判断输出的一对图像是否为一个真实的变换。上述的 cGAN, iGAN 以及 pip2pix 都存在一个问题:标记数据的获取困难。基于无监督模式构建的 GANs 包含如下典型算法:Radford 等<sup>[14]</sup>提出的 DCGAN 将 CNN<sup>[15]</sup>引入 GANs 的结构中,利用 CNN 强大的特征提取能力来提高模型的学习能力;Yi 等<sup>[16]</sup>提出的 DualGAN 引入 Arjovsky 等<sup>[17]</sup>提出的 Wasserstein 距离以及图像的 L1 范数<sup>[18]</sup>来构建损失函数;Kim 等<sup>[19]</sup>提出的 DiscoGAN 用于学习发现跨域关系;Zhu 等<sup>[20]</sup>提出的 CycleGAN 引入循环一致损失学习域  $X$  和域  $Y$  之间的映射函数。DualGAN, DiscoGAN 以及 CycleGAN 都借鉴了对偶学习<sup>[21]</sup>的想法,用于控制模型的学习。

作为典型的无监督域映射 GANs 模型, CycleGAN<sup>[20]</sup> 可以生成质量较高的图像,特别是针对目标输出无法明确定义的图像生成任务。CycleGAN 的成功主要源于在 GANs 的基础上引入了循环一致损失函数<sup>[22]</sup>。CycleGAN 优化模型中的损失函数由对抗损失和循环一致损失组成,其中对抗损失的作用是控制生成图像的风格近似目标图像<sup>[6]</sup>,循环一致损失的作用是保留输入的轮廓信息,其实质就是:更好地保留输入中的内容结构并捕获目标域的特征,如纹理、颜色或样式<sup>[16]</sup>。这样的损失组成能够兼顾到输入的轮廓和目标图像的风格。

但是,经典 CycleGAN 模型面对不同纹理复杂度的图像生成任务时,其循环一致损失系数是默认不变的,因此基于 CycleGAN 模型生成的图像存在纹理变形甚至消失等问题。理论分析表明:循环一致损失系数在图像生成中起到重要作用,通过调节系数大小可以控制生成图像的质量。

本文利用模糊方法<sup>[23]</sup>融合图像的空间维度和时间维度特征,度量图像集的纹理复杂度,得到循环一致损失系数的大小与不同纹理复杂度图像生成质量的关联性。

面对不同纹理复杂度的图像生成任务,本文将动态调整循环一致损失系数的取值。对于纹理复杂度较高的图像集,例如石材图像,其具有丰富的纹理细节,在进行图像的生成任务时,如果希望更多地保留输入图像的纹理信息,那么需要适当地调高循环一致损失系数的取值,使得优化过程适当倾向于轮廓的保留。对于复杂度较低的图像数据集,例如苹果和橙子图像集,其轮廓相对简单,适当调低循环一致损失系数也能很好地保留输入图像的轮廓信息,使得优化的过程适当倾向于生成图像的风格。

本文的主要贡献如下:

1) 针对不同纹理复杂度的图像生成任务,发现循环一致损失系数的大小与图像生成质量存在关联性,给出了循环一致损失系数的优化选择策略。

2) 融合图像的空间维度和时间维度特征,将模糊方法扩展到评估图像集复杂度。

3) 基于 ImageNet 数据集<sup>[24]</sup>训练好的 Inception-V3 模型<sup>[25]</sup>进行迁移学习,训练图像分类器,缩短训练时间并提高分类准确性,将分类准确性作为生成图像质量的一个度量指标<sup>[26]</sup>。

## 2 CycleGAN 模型概述

CycleGAN 模型的目标是在域  $X$  和域  $Y$  之间学习映射函数,训练样本  $\{x_i\}_{i=1}^N (x_i \in X)$  以及  $\{y_j\}_{j=1}^M (y_j \in Y)$ , 数据分布为  $x \sim p_{\text{data}}(x), y \sim p_{\text{data}}(y)$ , 如图 1(a) 所示。模型学习到的映射函数应该是循环一致的:对于域  $X$  中的每一张图像  $x$ , 图像的循环转换都能将  $x$  还原成原始图像。例如  $x \rightarrow G(x) \rightarrow F(G(x)) \approx x$ , 如图 1(b) 所示。同样,对于域  $Y$  中的每一张图像  $y$ , 图像的循环转换都能将  $y$  还原成原始图像。例如  $y \rightarrow F(y) \rightarrow G(F(y)) \approx y$ , 如图 1(c) 所示。

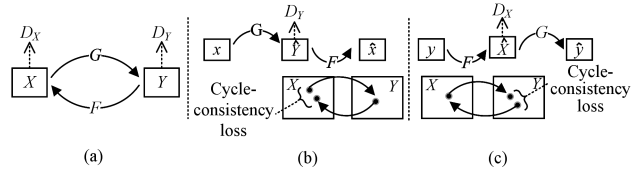


图 1 CycleGAN 结构示意图

Fig. 1 Structure schematic of CycleGAN

CycleGAN 模型的损失函数由 3 部分组成,具体如下:

$$\mathcal{L}(G, F, D_X, D_Y) = \mathcal{L}_{\text{GAN}}(G, D_Y, X, Y) + \mathcal{L}_{\text{GAN}}(F, D_X, Y, X) + \lambda \mathcal{L}_{\text{cyc}}(G, F) \quad (1)$$

$$\mathcal{L}_{\text{cyc}}(G, F) = \mathbb{E}_{x \sim p_{\text{data}}(x)} [\|F(G(x)) - x\|_1] + \mathbb{E}_{y \sim p_{\text{data}}(y)} [\|G(F(y)) - y\|_1] \quad (2)$$

其中,  $\mathcal{L}_{\text{GAN}}(G, D_Y, X, Y), \mathcal{L}_{\text{GAN}}(F, D_X, Y, X)$  分别是双边映射中的对抗损失;  $\lambda \mathcal{L}_{\text{cyc}}(G, F)$  为循环一致损失,其中  $\lambda$  称为循环一致损失系数,控制着循环一致损失相对于其他两个对抗损失的重要性<sup>[20]</sup>;  $\mathcal{L}_{\text{cyc}}(G, F)$  为输入图像与重构图像之间的 L1 范数。因此, CycleGAN 模型用于图像生成时可归结为如下优化问题:

$$G^*, F^* = \arg \min_{G, F} \max_{D_X, D_Y} \mathcal{L}(G, F, D_X, D_Y) \quad (3)$$

在训练的过程中,  $\mathcal{L}_{\text{GAN}}(G, D_Y, X, Y)$  被替换成  $\mathcal{L}_{\text{LSGAN}}(G, D_Y, X, Y)$ , 即将式(4)的负对数似然替换成式(5)的最小平方损失 LSGAN<sup>[27]</sup>。这样做的目的是稳定训练过程,并产生高质量的图像输出<sup>[20]</sup>。

$$\mathcal{L}_{\text{GAN}}(G, D_Y, X, Y) = \mathbb{E}_{y \sim p_{\text{data}}(y)} [\log D_Y(y)] + \mathbb{E}_{x \sim p_{\text{data}}(x)} [\log(1 - D_Y(G(x)))] \quad (4)$$

$$\mathcal{L}_{\text{LSGAN}}(G, D_Y, X, Y) = \mathbb{E}_{y \sim p_{\text{data}}(y)} [(D_Y(y) - 1)^2] + \mathbb{E}_{x \sim p_{\text{data}}(x)} [D_Y(G(x))^2] \quad (5)$$

循环一致的想法主要受对偶学习<sup>[21]</sup>的启发,其初衷是减少培训英语-法语和法语-英语翻译人员对标注数据的要求。对偶学习的思路为:充分利用大量无标注语义数据以及较少的标注数据,来对翻译模型进行训练。其极大地减少了对成对双语数据的依赖性。

单独的对抗损失函数(如式(4)所示),并不能保证对于一个输入  $x_i$ , 学习到的函数能输出一个预期的  $y_i$ 。CycleGAN 模型中,为了进一步减少可能映射函数的空间,认为学习映射函数应该是循环一致的<sup>[20]</sup>。DualGAN 中对循环一致的解释是:通过双发生器,使得输入能重建输出,并加强编码目标分布的反馈信号<sup>[16]</sup>。得到循环一致损失函数的作用在于:在像

素级别控制图像生成,缩小映射空间,更好地保留输入图像的轮廓。CycleGAN 模型损失函数(即式(1))中的循环一致损失函数项,通过引入 L1 范数使得优化的过程兼顾了图像像素级别的变化, $\lambda$  的大小决定了循环一致损失对整个损失函数的重要程度<sup>[20]</sup>。CycleGAN 原始模型中  $\lambda$  的取值默认为 10。

### 3 CycleGAN 模型生成图像时纹理的变形和消失

由第 2 节可知,循环一致损失(式(6)和式(7)的第一部分)的作用是保留输入的轮廓信息,而对抗损失(式(6)和式(7)的第二部分)的作用是保证生成图像近似于目标图像的风格<sup>[6]</sup>。根据 Sabour 等<sup>[28]</sup>提出的胶囊网络(Dynamic Routing Between Capsules),对应到式(6)和式(7),如果  $\lambda$  取值较大,则训练的过程由  $\mathcal{L}_G$  函数的第一部分主导,生成图像能较好地保留输入图像的轮廓,而不能较好地生成具有目标图像集分布的图像;如果  $\lambda$  取值较小,训练的过程由  $\mathcal{L}_G$  函数的第二部分主导,生成图像能较好地近似目标图像集分布,而不能较好地保留输入图像的轮廓信息。因此,需要平衡输入图像轮廓和目标图像集分布之间的关系。

CycleGAN 中引入了循环一致损失函数,目的是增加更多的限制,保留输入图像的轮廓特征,减小映射空间,同时也使得优化 GANs 的过程更加稳定。

$$\mathcal{L}_G(G, D_X, X, Y) = \lambda \mathcal{L}_{\text{cyc}}(G, F) + \mathbb{E}_{x \sim p_{\text{ann}}(x)} [(D_Y(G(x)) - 1)^2] \quad (6)$$

$$\mathcal{L}_F(F, D_X, X, Y) = \lambda \mathcal{L}_{\text{cyc}}(G, F) + \mathbb{E}_{y \sim p_{\text{ann}}(y)} [(D_X(F(y)) - 1)^2] \quad (7)$$

研究发现,在诸如大理石等高纹理复杂性图像的生成任务中,使用默认的循环一致损失系数( $\lambda = 10$ ),生成图像的纹理存在变形甚至消失的现象,如图 2 所示。

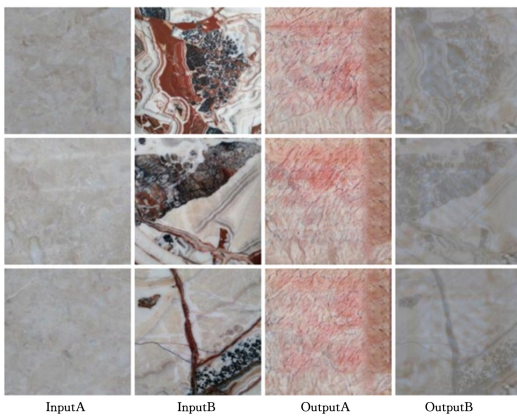


图 2 大理石图像生成过程中的纹理消失

Fig. 2 Disappearance of texture in marble image generation process

图 2 中,InputA 和 InputB 为输入的训练样本,OutputA 由 InputA 生成,OutputB 由 inputB 生成,OutputA 中出现了明显的图像纹理变形或消失的现象。针对低纹理复杂性图像(苹果图像集的)生成任务,如果选择较高的  $\lambda$  值,则生成的图像质量较差,不能很好地近似目标图像集风格,具体验证见第 5 节实验部分。原始 CycleGAN 中,循环一致损失系数的大

小默认为  $\lambda = 10$ ,且没有给出任何合理性的解释。

### 4 循环一致损失系数的优化选择方法

本文融合空间维度和时间维度来度量图像的纹理复杂性,据此选择大小合适的  $\lambda$  值,以平衡生成图像的轮廓和风格,提高生成图像的质量。

#### 4.1 图像纹理的复杂性度量

传统的图像度量更多的是关注图像空间维度的信息,例如特征提取。主要的测量方法有:根据 Leeuwenberg 等<sup>[29]</sup>提出的结构信息理论,或是基于边界粗糙度和对称性<sup>[30]</sup>进行图像复杂性的度量,往往忽略了人眼对于图像认知的的时间维度。因此,本文使用 Cardaci 等<sup>[23]</sup>提出的基于模糊方法的图像复杂性度量方法,融合图像的空间维度和时间维度来计算图像的复杂性,其计算结果值越大,纹理复杂性越高。但是,Cardaci 等度量的是一张图像的复杂度,对此我们进行相应的扩展,用该方法计算图像集中每张图像的复杂度,并求平均值  $C(S)$ ,最后将该平均值作为图像集的复杂度。图 3 列举了一些不同  $C(S)$  复杂度的图像集。其中,图 3(a)大理石的复杂度为 0.917;图 3(b)花岗岩的复杂度为 0.887;图 3(c)马的复杂度为 0.776;图 3(d)苹果的复杂度为 0.561。



图 3 不同纹理复杂度图像集的展示

Fig. 3 Display of different texture complexity image sets

$$G_0(X) = -\frac{1}{N \log(2)} \times \sum_{i=1}^N (h_i \log(h_i) + (1 - h_i) \log(1 - h_i))$$

$$G_1(X) = \frac{2\sqrt{e}}{N(e-1)} \sum_{i=1}^N \eta_i$$

where

$$\eta_i = \begin{cases} h_i \times (e^{1-h_i} - e^{h_i-1}), & \text{if } h_i \leq 0.5 \\ (1-h_i) \times (e^{h_i} - e^{-h_i}), & \text{if } h_i > 0.5 \end{cases}$$

$$G_2(X) = \frac{4}{N} \sum_{i=1}^N h_i (1 - h_i)$$

$$G(X) = (G_0 + G_1 + G_2) / 3$$

$$C(S) = \frac{1}{K} \times \sum_{j=1}^K G(j)$$

$G_j (j=0,1,2)$  使用标准的距离函数,它遵循由熵函数增强的一致性、对称性和三角不等式的通常性质函数,满足模糊熵函数的性质:它们是凸的、对称的,取值区间为 $[0,1]$ 。图像的模糊化可以通过简单的归一化函数来执行,该归一化函数赋予与图像强度成比例的图像的归属感,其中  $h_i \in [0,1]$  是像素  $i$  的缩放灰度值,它被视为像素  $i$  对大小为  $N$  的矢量化图像  $O$  (即全白图像) 的归属程度<sup>[23]</sup>。 $k$  为一个图像集中图像的张数。

Cardaci 等<sup>[23]</sup> 根据单张图像的纹理复杂度,将图像分为 3 个级别:简单、中等、难。在此,本文根据图像集合的  $C(S)$  值将图像集的纹理复杂度分为 3 个层次,即 L, M, H。具体划分如图 4 所示。

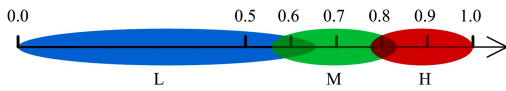


图 4 图像集的复杂度层次划分

Fig. 4 Complexity partition of image set

采用上述图像集纹理复杂度度量方法,我们度量了 CycleGAN<sup>[20]</sup> 实验中图像数据集的纹理复杂度,度量结果如表 1 所列。从中可以发现:绝大部分图像集的纹理复杂度为 M 层次。

表 1 CycleGAN<sup>[20]</sup> 实验中图像集的纹理复杂度

Table 1 Texture complexity of image sets in CycleGAN<sup>[20]</sup> experiments

Image sets	Image subsets	Texture complexity	Complexity level
Apple2orange	Apple	0.561	L
	Orange	0.587	L
Cezanne2photo	Cezanne	0.747	M
	Photot	0.706	M
Cityspace	Labels	0.743	M
	Photot	0.730	M
Facades	Lables	0.360	L
	Photo	0.780	M
Horse2zebra	Horse	0.776	M
	Zebra	0.743	M
Iphone2dslr_Flower	Iphone	0.682	M
	Flower	0.672	M
Maps	Map	0.424	L
	Photo	0.665	M
Monet2photo	Monet	0.811	M
	Photo	0.706	M
Summer2winter_Yosemite	Summer	0.696	M
	Winter	0.696	M
Ukiyoe2photo	Ukiyoe	0.718	M
	Photo	0.706	M
Vangoph2photo	Vango	0.761	M
	Photo	0.706	M

## 4.2 $\lambda$ 的优化选择策略

Sabour 等<sup>[28]</sup> 介绍了模型训练中的主导部分的作用;Zhu 等<sup>[20]</sup> 阐述了循环一致损失相对其他损失部分的重要性;Yi 等<sup>[16]</sup> 介绍了从生成器重建输出的有效性。上述 3 个研究团队从不同的层面和角度说明了  $\lambda$  的取值会影响图像的生成,

即纹理复杂度越高,应选择越大的循环一致损失系数;反之,应取较小的循环一致损失系数。因此,利用 CycleGAN 进行  $X$  和  $Y$  两图像集之间的转换时,可以根据图像集的纹理复杂性,动态调整  $\lambda$  的取值,更好地平衡输入图像轮廓和目标图像集分布之间的关系,提高生成图像的质量。本文给出如表 2 所列的优化选择策略。若图像集  $X$  和图像集  $Y$  所处复杂度层次不一致,则按所处层次较大的准则取  $\lambda$  值。通过该优化策略,针对不同纹理复杂性的图像生成任务,可以寻找到一个大小合适的  $\lambda$  值。

表 2 图像集纹理复杂度对应  $\lambda$  取值的优化选择策略

Table 2 Optimized selection strategy for image set texture complexity corresponding to  $\lambda$

Complexity Level	Set X,Y Complexity Level		$\lambda$ Range of Values
	H	M	
H			15~20
M			10~15
L			5~10

## 5 实验

实验的主要任务是利用 CycleGAN 进行图像生成,并根据图像集的纹理复杂性,找到一个大小合适的循环一致损失系数,以提高生成图像的质量。本文通过对高、中、低不同层次纹理复杂性的图像集进行不同取值实验,来验证所提优化选择策略的正确性。

### 5.1 数据集与实验环境

本文数据集主要来自:1)CycleGAN<sup>[20]</sup> 中使用的数据集,即马 $\leftrightarrow$ 斑马、苹果 $\leftrightarrow$ 橙子,它们均来源于 ImageNet;2)自采集数据集,即大理石集合 1 $\leftrightarrow$ 大理石集合 2。每个模型训练的次数统一为 200。限于篇幅,本文仅展示代表性实验结果。

通过控制变量法,取不同的  $\lambda$  进行实验,其他部分与 CycleGAN 保持一致,通过生成图像度量指标发现合适的  $\lambda$  值,从而验证所提  $\lambda$  优化选择策略的正确性。本文中的转换任务包括大理石集合 1 $\leftrightarrow$ 大理石集合 2、马 $\leftrightarrow$ 斑马、苹果 $\leftrightarrow$ 橙子。所有实验均是在 Tensorflow 平台上完成,开发语言为 Python 3.5.2。使用 2 块 NVIDIA GTX 1080Ti 显卡加速。

### 5.2 生成图像的质量评估方法

本文通过生成图像的分类准确性、多样性、清晰度,来评估生成图像的质量。

生成图像的分类准确性<sup>[26]</sup>:通过迁移学习<sup>[31]</sup>,利用 Inception-V3 预训练模型对 CycleGAN 生成图像进行所属类别的判断。如果生成图像足够逼真,则分类器在分类过程中对某一类给予高分<sup>[26]</sup>,前提是使用 CycleGAN 的训练集去训练这个模型时得到一个效果较好的分类器。

多样性<sup>[32]</sup>:引入图像多样性的度量方法来度量生成图像的多样性,得到的值越大,图像集的多样性越高。

清晰度<sup>[33]</sup>:仅仅产生多样性的图像结果是不够的,生成的图像还得具有可识别的清晰度,这对于实际的应用具有重要的意义。引入图像清晰度来度量生成图像的清晰度,得到的值越大,图像集的清晰度就越高。

使用上述 3 种图像度量方法,对生成的图像集合进行度

量。分类准确性度量中,将生成的两个图像集(OutputA, OutputB,其中 OutputA 由输入 InputA 生成,OutputB 由输入 InputB 生成)输入到训练好的分类模型中,得到对应的分类概率,取目标分类(由马生成斑马,则斑马为目标分类)的概率,最后对整个图像集取平均值并将其作为生成图像的分类准确性指标。

### 5.3 实验结果与分析

利用本文图像纹理复杂性度量方法,对实验中用到的图像集(大理石集合 1↔大理石集合 2、马↔斑马、苹果↔橙子)进行图形纹理复杂性的度量,得到的复杂度如表 3 所列。取不同的  $\lambda$  值( $\lambda = 5, 10, 15, 20, 25, 30, 40$ )进行实验,其中 CycleGAN 中默认  $\lambda = 10$ 。实验中,分别对高纹理、中等纹理、低纹理复杂性生成图像进行分类准确性、多样性以及清晰度度量。

表 3 实验图像集的复杂性

Table 3 Complexity of experimental image set

Image Sets	Texture Complexity	Complexity Level
Marble Set 1	0.957	H
Marble Set 2	0.917	H
Horse	0.776	M
Zebra	0.743	M
Apple	0.561	L
Orange	0.587	L

#### 5.3.1 高纹理复杂性图像的生成

图 5 和图 6 具体给出高纹理复杂性图像集(大理石集合 1↔大理石集合 2)生成图像的分类准确性、多样性和清晰度。图 7 和图 8 为生成图像在不同  $\lambda$  取值下的效果展示。其中, OutputA, OutputB 分别为输入图像集大理石 1 和大理石 2 生成图像的输出。

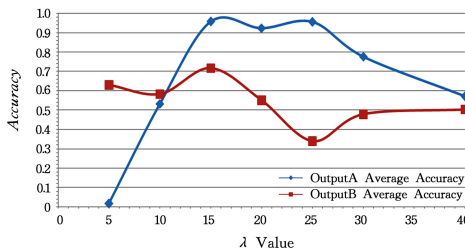


图 5 高纹理复杂性图像集生成图像的平均分类准确性

Fig. 5 Average classification accuracy of image generated by high-texture complexity image set

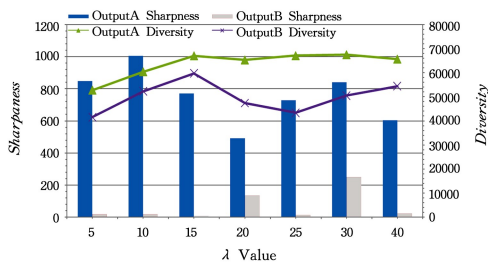


图 6 高纹理复杂性图像集生成图像的多样性和清晰度

Fig. 6 Diversity and sharpness of image generated by high-texture complexity image set

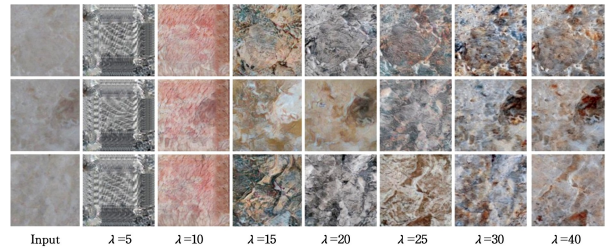


图 7 不同  $\lambda$  值的生成效果(1)

Fig. 7 Generate effect (1) for different  $\lambda$  values

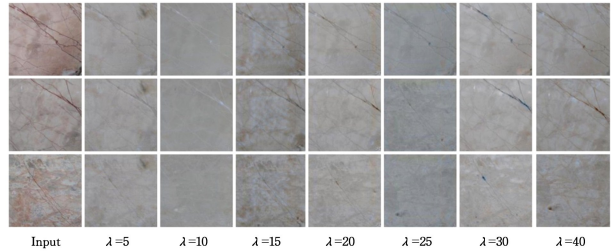


图 8 不同  $\lambda$  值的生成效果(2)

Fig. 8 Generate effect (2) for different  $\lambda$  values

进行高纹理复杂性图像生成时,在默认的  $\lambda = 10$  处生成图像的分类准确率并非最高,而在  $\lambda = 15$  处取得最高,如图 5 所示,并且随着  $\lambda$  的增大准确率逐渐降低,兼顾图 6 中生成图像的多样性和清晰度,发现  $\lambda$  的合理取值范围为 15~20。

#### 5.3.2 中等纹理复杂性图像的生成

图 9 和图 10 给出了中等纹理复杂性图像集(马↔斑马)生成图像的分类准确性、多样性和清晰度。图 11 和图 12 为生成图像在不同  $\lambda$  取值下的效果展示。其中, OutputA, OutputB 分别为输入图像集马和斑马生成图像的输出。

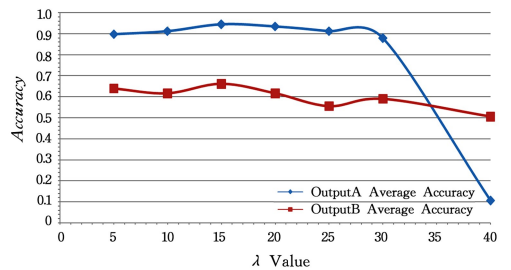


图 9 中等纹理复杂性图像集生成图像的平均分类准确性

Fig. 9 Average classification accuracy of image generated by medium-texture complexity image set

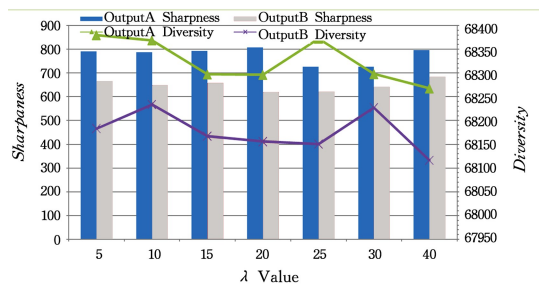


图 10 中等纹理复杂性图像集生成图像的多样性和清晰度

Fig. 10 Diversity and sharpness of image generated by medium-texture complexity image set

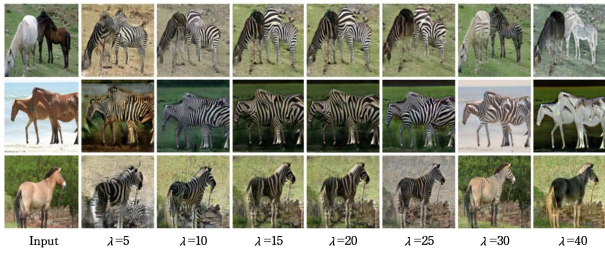


图 11 不同  $\lambda$  值的生成效果(3)

Fig. 11 Generate effect (3) for different  $\lambda$  values

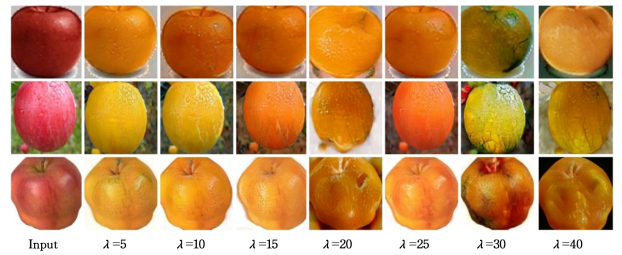


图 15 不同  $\lambda$  值的生成效果(5)

Fig. 15 Generate effect (5) for different  $\lambda$  values

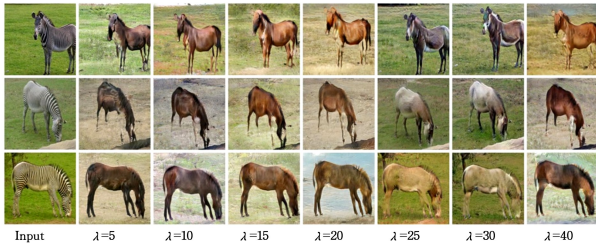


图 12 不同  $\lambda$  值的生成效果(4)

Fig. 12 Generate effect (4) for different  $\lambda$  values

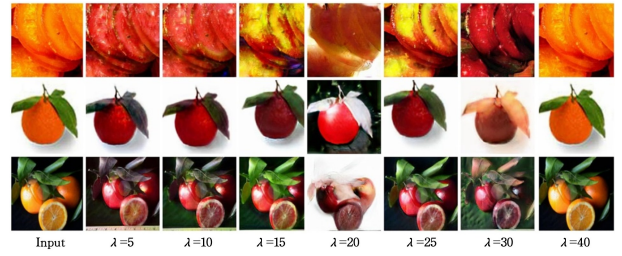


图 16 不同  $\lambda$  值的生成效果(6)

Fig. 16 Generate effect (6) for different  $\lambda$  values

进行中等纹理复杂性图像生成时发现,生成图像的分类准确率在  $\lambda=15$  处取得最高,如图 9 所示,兼顾图 10 中生成图像的多样性和清晰度,发现  $\lambda$  的合理取值范围为 10~15。

### 5.3.3 低纹理复杂性图像的生成

图 13 和图 14 为低纹理复杂性图像集(苹果橙子)生成图像的分类准确性、多样性和清晰度。

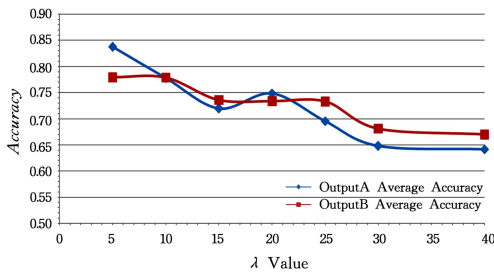


图 13 低纹理复杂性图像集生成图像的平均分类准确性

Fig. 13 Average classification accuracy of image generated by low-texture complexity image set

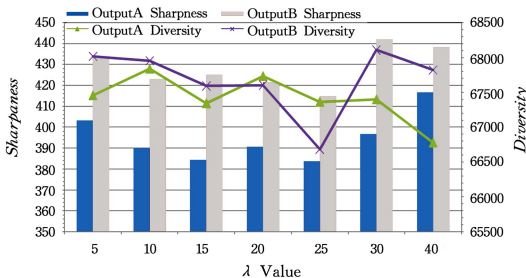


图 14 低纹理复杂性图像集生成图像的多样性和清晰度

Fig. 14 Diversity and sharpness of image generated by low-texture complexity image set

图 15 和图 16 为生成图像在不同  $\lambda$  取值下的效果展示。其中,OutputA,OutputB 分别为输入图像集苹果和橙子生成图像的输出。

进行低纹理复杂性图像生成时可以发现,生成图像的分类准确率由  $\lambda=5$  向  $\lambda=40$  呈逐渐下降趋势,如图 13 所示,同时兼顾图 14 中生成图像的多样性和清晰度,发现  $\lambda$  的合理取值范围为 5~10。

**结束语** 利用 CycleGAN 进行图像转换时,如何平衡输入图像轮廓和近似目标图像集风格之间的关系,取决于  $\lambda$  的大小。本文融合图像的空间维度和时间维度来度量图像的纹理复杂性,得到图像集纹理复杂性对应  $\lambda$  取值的优化选择策略;利用这个策略,在进行图像生成时,动态调整  $\lambda$  的取值,使得生成图像有更好的分类准确性、多样性和清晰度。

通过本文的策略,利用 CycleGAN 进行图像生成,找到合适的  $\lambda$  取值,提高了生成图像的质量。但是面对极端的情况,例如纹理复杂性非常高与纹理复杂性非常低的图像集进行相互转换时,并不能通过调节  $\lambda$  的取值来获得较好的生成效果。因此在未来工作中,我们希望通过增加一些有关图像复杂度的损失项,来调节图像集之间悬殊复杂度带来的图像生成问题。

## 参 考 文 献

- [1] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2015: 3431-3440.
- [2] GATYS L, ECKER A S, BETHGE M. Texture synthesis using convolutional neural networks[C] // Advances in Neural Information Processing Systems. Cambridge, Massachusetts: MIT Press, 2015: 262-270.
- [3] GATYS L A, ECKER A S, BETHGE M. Image style transfer using convolutional neural networks[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2016: 2414-2423.
- [4] JOHNSON J, ALAHI A, FEI-FEI L. Perceptual losses for real-

- time style transfer and super-resolution[C] // European Conference on Computer Vision. Berlin, German; Springer, 2016: 694-711.
- [5] NASH J. Non-Cooperative Games[J]. *Annals of Mathematics*, 1951, 54(2):286-295.
- [6] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[C] // International Conference on Neural Information Processing Systems. Cambridge, Massachusetts; MIT Press, 2014:2672-2680.
- [7] DENTON E L, CHINTALA S, FERGUS R. Deep Generative Image Models using a Laplacian Pyramid of Adversarial Networks[C] // Advances in Neural Information Processing Systems. Cambridge, Massachusetts; MIT Press, 2015:1486-1494.
- [8] LI C, WAND M. Precomputed real-time texture synthesis with markovian generative adversarial networks[C] // European Conference on Computer Vision. Berlin, German; Springer, 2016: 702-716.
- [9] LI C, ZHAO X Y, XIAO L M, et al. Multi-layer perceptual de-fogging algorithm for image under generative adversarial networks[J]. *Journal of Computer-Aided Design & Computer Graphics*, 2017, 29(10):1835-1843. (in Chinese)  
李策, 赵新宇, 肖利梅, 等. 生成对抗映射网络下的图像多层感知去雾算法[J]. *计算机辅助设计与图形学学报*, 2017, 29(10): 1835-1843.
- [10] LEDIG C, WANG Z, SHI W, et al. Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network[C] // Computer Vision and Pattern Recognition. IEEE, 2017: 105-114.
- [11] MIRZA M, OSINDERO S. Conditional generative adversarial nets[J]. arXiv preprint arXiv:1411.1784, 2014.
- [12] ZHU J Y, KRÄHENBÜHL P, SHECHTMAN E, et al. Generative visual manipulation on the natural image manifold[C] // European Conference on Computer Vision. Berlin, German; Springer, 2016:597-613.
- [13] ISOLA P, ZHU J Y, ZHOU T, et al. Image-to-Image Translation with Conditional Adversarial Networks[C] // IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ; IEEE, 2017:5967-5976.
- [14] RADFORD A, METZ L, CHINTALA S. Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks[J]. arXiv:1511.06434, 2016.
- [15] FUKUSHIMA K. Neural network model for a mechanism of pattern recognition unaffected by shift in position-Neocognitron [J]. *IEICE Technical Report A*, 1979, 62(10):658-665.
- [16] YI Z, ZHANG H, TAN P, et al. DualGAN: Unsupervised Dual Learning for Image-to-Image Translation[C] // IEEE International Conference on Computer Vision. Piscataway, NJ; IEEE, 2017:2868-2876.
- [17] ARJOVSKY M, CHINTALA S, BOTTOU L. Wasserstein GAN [J]. arXiv preprint arXiv:1701.07875, 2017.
- [18] GOLDSTEIN T, OSHER S. The split Bregman method for L1-regularized problems[J]. *SIAM Journal on Imaging Sciences*, 2009, 2(2):323-343.
- [19] KIM T, CHA M, KIM H, et al. Learning to Discover Cross-Domain Relations with Generative Adversarial Networks[C] // Proceedings of the 34th International Conference on Machine Learning. New York; ACM, 2017:1857-1865.
- [20] ZHU J Y, PARK T, ISOLA P, et al. Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks[C] // IEEE International Conference on Computer Vision. Piscataway, NJ; IEEE, 2017:2242-2251.
- [21] HE D, XIA Y, QIN T, et al. Dual learning for machine translation[C] // Advances in Neural Information Processing Systems. Cambridge, Massachusetts; MIT Press, 2016:820-828.
- [22] ZHOU T, KRAHENBUHL P, AUBRY M, et al. Learning dense correspondence via 3d-guided cycle consistency[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ; IEEE, 2016:117-126.
- [23] CARDACI M, DI GESÙ V, PETROU M, et al. A fuzzy approach to the evaluation of image complexity[J]. *Fuzzy Sets and Systems*, 2009, 160(10):1474-1484.
- [24] RUSSAKOVSKY O, DENG J, SU H, et al. Imagenet large scale visual recognition challenge[J]. *International Journal of Computer Vision*, 2015, 115(3):211-252.
- [25] SZEGEDY C, VANHOUCKE V, IOFFE S, et al. Rethinking the inception architecture for computer vision[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ; IEEE, 2016:2818-2826.
- [26] WANG X, GUPTA A. Generative image modeling using style and structure adversarial networks[C] // European Conference on Computer Vision. Berlin, German; Springer, 2016:318-335.
- [27] MAO X, LI Q, XIE H, et al. Least squares generative adversarial networks[C] // 2017 IEEE International Conference on Computer Vision (ICCV). IEEE, 2017:2813-2821.
- [28] SABOUR S, FROSST N, HINTON G E. Dynamic routing between capsules[C] // Advances in Neural Information Processing Systems. Berlin, German; MIT Press, 2017:3859-3869.
- [29] LEEUWENBERG E, BUFFART H. An outline of coding theory [J]. *Advances in psychology*, 1983, 11:25-47.
- [30] SU H, BOURIDANE A, CROOKES D. Scale Adaptive Complexity Measure of 2D Shapes[C] // International Conference on Pattern Recognition. Piscataway, NJ; IEEE, 2006:134-137.
- [31] PAN S J, YANG Q. A survey on transfer learning[J]. *IEEE Transactions on knowledge and data engineering*, 2010, 22(10): 1345-1359.
- [32] ZHOU B, LAPEDRIZA A, XIAO J, et al. Learning deep features for scene recognition using places database[C] // Advances in neural information processing systems. Cambridge, Massachusetts; MIT Press, 2014:487-495.
- [33] GARCIA B, BRUNET P. 3D reconstruction with projective octrees and epipolar geometry[C] // International Conference on Computer Vision. Piscataway, NJ; IEEE, 2008:1067-1072.