

文章编号: 2096-1618(2021)06-0629-05

# 基于生成对抗网络的图像风格转换算法

李权学, 牛梦晨, 陈睿麟, 陈伊婷, 杜 帅, 徐梓欣  
(成都信息工程大学, 四川 成都 610225)

**摘要:**针对图像样式转换产生的图像质量不高的问题,提出一种基于生成对抗网络的高质量图像样式转换方法。借鉴循环GAN网络结构上的发电网络相结合的方法采用跳层结构和U-Net网络中的残差网络,增强网络的多尺度不变性;其次,在判别网络方面,提出一种多尺度扩展卷积判别器,以改善图像样式的空间几何变换。实验证明,与Cycle-GAN算法相比,该算法在图像样式转换中的效果有很大提升,图像样式转换的质量也得到了提高。

**关键词:**图像风格转换;GAN;Cycle-GAN;U-Net 残差网络

**中图分类号:**TP302.4

**文献标志码:**A

**doi:**10.16836/j.cnki.jcuit.2021.06.008

## 0 引言

随着图像技术的发展,改变图像样式已成为图像处理领域的一个分支。图像迁移技术是将一幅图像的内容与一幅或多幅图像的内容进行融合,得到一幅新图像的技术。在神经网络出现前,为使一幅图像具有另一幅图像的特征,通过建立数学模型进行分析,不断调整偏移模型,然后将图像输入到模型输出结果中,完成图像风格的偏移。直到2015年,Gayts等提出利用深度学习提取图像纹理,通过梯度下降算法,不断优化模型参数,对迁移后的图像进行建模,最终生成另一幅符合预期的图像。其原理是在用深度神经网络(CNN)训练图像的过程中,靠近输入层的特征映射包含更多的图像纹理信息,而靠近输出层的特征映射包含更多的内容信息。这种CNN训练图像的性质已经成为最早的图像风格转换技术。

在基于深度学习的图像风格转换处理技术中,目前应用最广泛的是基于生成对抗网络的图像风格转换处理技术。生成性对抗网络GAN(Generative),是从数学中的博弈论中衍生的一种对抗性学习方法。一般的理论是通过不断对数据的训练和迭代,网络可以学习到图像的特征。根据这一特性,可以生成接近原始数据的新数据。该方法具有很强的性能,目前已广泛应用于各个领域。文献[1]提出一种基于生成对抗网络的固定物体步态识别方法;文献[2]提出一种基于条件生成对抗网络的焊接图像修复方法;文献[3]使用生成对抗网络对自动视觉检测进行建模和分析;文献[4]研究具有图像放大的超分辨率生成对抗网络的

降噪性能;文献[5]提出一种基于一维卷积神经网络和条件生成对抗网络的模型。该模型可以利用苗木的早期电信号快速有效地识别苗木的耐盐期。生成性对抗网络在图像上也有很好的效果。文献[6]提出一种使用生成对抗网络生成人类胚胎细胞合成图像的方法。文献[7]使用生成对抗网络对高光谱遥感图像进行分类,并提出一种多鉴别器生成对抗网络来研究鉴别器数目对分类结果的影响。

目前,应用最成熟的图像风格转换技术是商用产品Prisma,原理是通过深度神经网络学习某幅图像的风格,然后将这种风格应用到其他图像中。深度神经网络中应用最广泛的技术是循环GAN(cycleGAN)。循环GAN相当于两个传统GAN。每个GAN发生器通过最小化损耗学习相应的变换函数。测量生成的数据和目标数据的差异,以计算发电机损耗,差值越大,生成器受到的惩罚将越高。虽然图像样式转换技术在很多情况下都采用循环生成对抗网络,但在生成的图像中容易忽略一些细节,且存在噪声等问题<sup>[4]</sup>。

参照循环生成对抗网络的结构,提出一种高质量的图像样式转换网络。首先,在生成网络方面,将跳层结构与U网络中的残差网络结合,可以增加网络的多尺度不变性;其次,在判别网络方面,多尺度扩展卷积判别器可以提高图像样式转换的空间几何变换效率,并能生成高分辨率的图像。将目标域数据和源域数据的数据特征经过两次处理后视为一个双代对抗网络。

## 1 相关理论

### 1.1 生成对抗网络

生成对抗网络由生成网络和判别网络两个子网络

组成。采用数学博弈的思想。生成网络用于生成真实的图像,而判别网络则试图将生成的图像与真实的图像区分开。各网络不断对抗和优化,最终达到预期效果。

生成对抗网络的结构如图 1 所示。它由生成网络  $G$  和判别网络  $D$  组成。将随机变量  $z$  输入生成网络  $G$ ,然后判别网络  $D$  判别输入目标,并确定其是生成图像还是真实图像,生成网络的结果是  $G(z)$ 。生成网络  $G$  使  $G(z)$  与真实图像保持接近,直到判别网络不能区分生成的内容与真实内容为止。此时,生成的网络和判别网络之间形成纳什均衡。

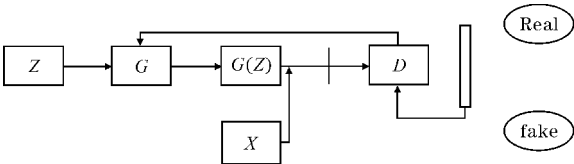


图 1 生成对抗网络的结构图

在生成对抗网络的训练过程中,生成网络和判别网络不断地进行迭代博弈,最终对网络进行优化。从图 2 可以看出, $z$  代表输入数据,生成网络  $G$  生成的结果为  $G(z)$ ,然后判别网络  $D$  识别  $G(z)$  和真实图像  $X$ ,真实图像的概率表示为  $D(X)$ ,当训练达到纳什均衡时, $D(X) = 1/2$ 。

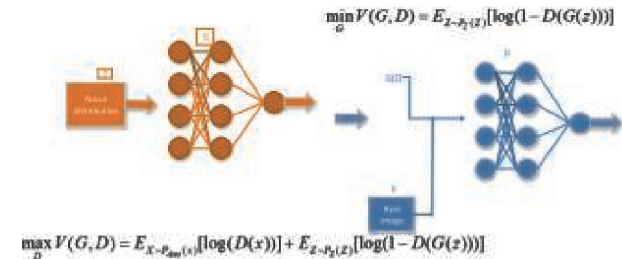


图 2 生成对抗网络的训练过程

在网络优化过程中,首先是对判别网络进行优化,目的是更好地区分生成的图像和真实图像,即使判别网络更加敏感,即使  $D(x)$  最大化;其次,对于生成的图像,需要最小化判别网络处理后的  $D(G(z))$ 。因此,第一步的目标定义为

$$\max_D V(G, D) = E_{X \sim P_{data}(x)} [\lg(D(x))] + E_{Z \sim P_Z(z)} [\lg(1 - D(G(z)))] \quad (1)$$

在优化生成网络的过程中,生成的图像可以尽可能接近真实图像,因此  $D(G(z))$  必须最大化,因此第二步的目标是:

$$\min_G V(G, D) = E_{Z \sim P_Z(z)} [\lg(1 - D(G(z)))] \quad (2)$$

形成对抗网络的总体目标是:

$$\min_{G, D} \max \{ \text{Obj}(D, G) = E_{X \sim P_{data}(x)} [\lg(D(X))] + E_{Z \sim P_Z(z)} [\lg(1 - D(G(z)))] \} \quad (3)$$

其中,  $P_{data}$  是真实数据的概率分布,  $P_Z$  是生成数据的概率分布。GAN 的训练过程如图 2 所示。它实际上是一个生成网络和判别网络相互对抗以实现最优解的过程。在这个过程中,需要对两个网络进行优化,直到达到纳什均衡。

1.2 循环生成对抗网络

循环生成对抗网络 (Cycle-Generative antiparial Network, Cycle-GAN) 是基于双重学习思想,在源域和目标域之间没有对应关系的情况下进行学习,然后变换图像样式。循环 GAN 可以先从源域转换到目标域,再从目标域转换到源域,利用这种方法实现图像配对<sup>[8]</sup>。与 GAN 模型不同的是,生成网络和判别网络在对抗中相互博弈,优化模型,使生成的图像更接近目标图像。循环生成对抗网络的结构如图 3 所示。

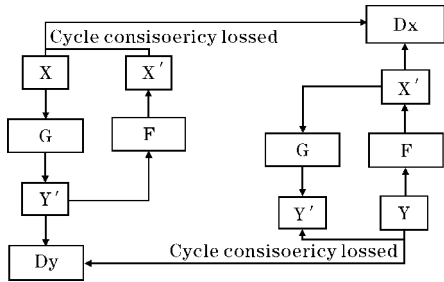


图 3 循环生成对抗网络结构图

在图像样式转换中,利用循环一致性解决图像转换问题。在模型的判断网络部分,输入原始图像和生成的图像,判断网络对两幅图像进行判断。同时,生成的图像也会随着目标图像的不断逼近,而不断进行优化迭代。

基于 CycleGAN 的图像样式转换可以将一幅图像转换为另一幅图像,但生成的图像质量不高。因此,提出一种基于循环生成对抗网络生成高质量图像的新型图像。图像样式转换方法。

2 高质量图像样式转换算法

2.1 建立网络

循环 GAN 结构仅利用残差网络进行单尺度学习,提高训练性能,限制网络的学习功能。在 U-Net 网络中,采用跳层结构和残差网络相结合的方法。增加网络的多尺度不变性。

U-Net 网络模型采用完全卷积网络结构,增加一个跳跃层结构,它可以将解码前的特征映射和解码后的特征映射按信道进行组合,也可以与传统编码相结

合,有不同的结构。传统的编码结构是先下采样,再上采样到原始分辨率。U-Net 以不同的分辨率保存信息,与传统的编解码结构相比,U-Net 融合了网络的高层和低层结构信息,模型的分割精度大大提高。U-Net 分割模型中没有使用全连接层,并且网络模型的输入与输出都是图像,通常在卷积神经网络的浅层的特征图中包含图像很多的位置信息、纹理信息等,在网络深层的卷积特征图包含的是分割的抽象特征有利于像素分类。具有对称结构的 U-Net 分割模型很好的继承全卷积神经网络模型的优势,网络模型的输入图像大小没有局限性,并且具有语义分割模型的端对端特点<sup>[9]</sup>。因此 U-Net 可以生成更高质量的图像,改进效果非常明显。图 4 是传统的编码结构,图 5 是 U-Net 结构。

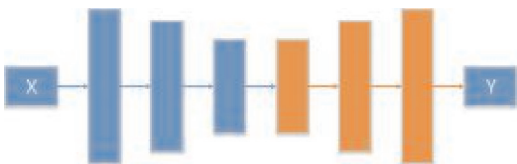


图 4 传统编码结构



图 5 U-Net 结构

目前,神经网络的层次越来越深,模型越来越复杂,训练和优化的难度越来越大。一般来说,深度模型越深,训练效果越好。但是,也会带来很多问题,如训练难度增加、模型梯度爆炸、梯度消失等。在这些问题上,向网络中添加剩余块可以防止这些问题的发生。剩余块是具有跳转结构的网络,如图 6 所示。

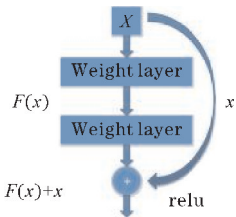


图 6 残差网络结构

从图 6 可以看出, $x$  是残差块的第一层的输入值。在第一层线性变化和激活函数  $\text{relu}$  后,输出  $F(x)$ ,然后在线性变化而不是激活函数后, $F(x)$  将该层的输入值  $x$  相加,然后激活并输出。添加  $x$  的路径称为快捷连接。残差网络是一种跳转结构,可以将前一个网络的信息跳到下一个网络,然后传递给下一个网络。确实,后面的网络可以学习前面的信息,这种结构可以保

护信息的完整性。

在 U-Net 网络和残差网络的基础上,提出一种综合两者优点的生成结构网络。与 GAN 标准的编解码结构相比,这种生成网络能够保留更多的信息,生成的图像分辨率更高。在 U-Net 网络中,加入残差块,使网络的信息容量更大,学习能力更强,网络可以在较低和较高的范围内学习。改进的网络结构如图 7 所示。

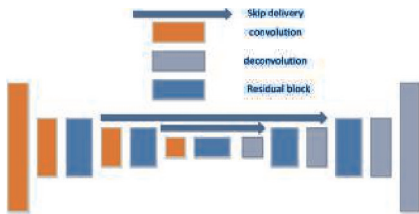


图 7 改进的网络结构

2.2 判别网络构造

扩展卷积是让卷积核的数目保持不变,感受野的大小随着扩展速率的增大而增大。扩张卷积图如图 8 所示。

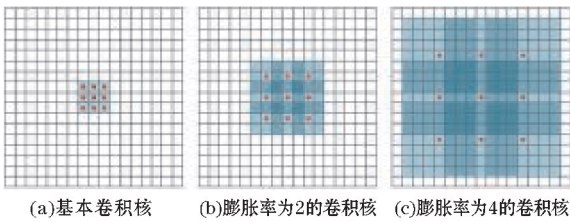


图 8 扩张卷积图

展开卷积运算仍然是很好的理解。图 8(a) 是基本卷积核,而扩展卷积就是在这个基本卷积核上加一个区间。图 8(b) 是图 8(a)  $3 \times 3$  膨胀率为 2 的卷积,但间隔为 1,相当于  $7$  对应的图像块乘  $7$ 。可以理解,内核大小已经变成  $7 \times 7$ ,但只有 9 个点有参数,其余位置参数为 0,通过卷积计算输入特征图对应位置的像素,跳过其余位置。图 8(c) 和图 8(b) 相似,只是扩张率为 4,相当于变成  $15 \times 15$  卷积核。当卷积核变大时,感受也自然会变大:

$$RF_i = RF_{i-1} + (k-1) \times s \tag{4}$$

其中, $RF_i$  是上层的感受野, $k$  是卷积核的大小, $s$  是步长。

具有大感受野的判别网络具有较高的分辨率,能够准确地区分输入数据是生成数据还是真实数据。为提高分辨率,一种方法是使用较深的网络,另一种方法是使用较大的卷积核,但这两种方法容易过拟合,占用较多的内存。为解决这一问题,采用不同的尺度训练多个鉴别器来判断图像的真实性。多尺度原理图如图 9 所示。



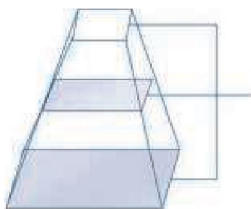


图 9 多尺度原理图

提出的判别网络在多尺度上实现了扩张卷积,即在不同尺度上使用不同的扩张系数来获取更多的信息。这种网络可以预测更大范围的数据,生成高分辨率的图像,具有更好的效果。改进的鉴别器网络结构如图 10 所示。

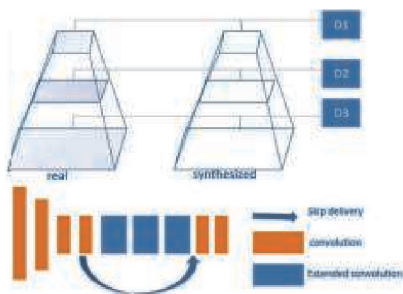


图 10 改进的鉴别器网络结构图

在本文提出的对抗性别网络生成实验中,选取 1800 张图片作为训练集,200 张图片作为测试集,图片大小为 1024×768,设置了 100 个历元,并对图像进行训练。

3 结果

3.1 实验结果

图 11(a)是原始图像,图 11(b)是性别转换后的图像。从图中可以看出,使用本文设计的生成对抗网络后,生成的图像具有明显的性别转换。



图 11 实验结果图

3.2 实验比较

为对提出的模型方法进行评价,采用初始得分

(IS)、Frechet 初始距离(FID)和 AMT 知觉研究(Ama-  
zon-Mechanical-Turk 知觉研究)测试进行定量评价。  
初始分数是一个评价指标,它与生成图片的质量和多样性更相关,但分数较高的图片实际上是锐利多样的  
图片,而不是生成图片与真实图片的关系。分数越高,  
生成图像的细节越多,质量越好。FID 值是生成数据  
和实际数据的更接近分布。AMT perception research  
是在 Amazon 劳务平台上发布“辨别图片真实性”任  
务,对输出图片的真实性进行评估。任务参与者将查  
看配对的真实图片,然后单击以选择认为是真实图  
片的图片。最后,对大家进行总结。实验结果表明,这  
是评价图像真实性的金科玉律。数据评价结果见表 1。

表 1 评价结果

Model	Inception Score	FID	AMT
Cycle GAN	12.5	32.5	28.4
Improved model	16.7	28.3	31.2

3.3 训练过程

图 12 是本文提出的生成对抗网络在训练过程中的  
发生器和判别损失变化曲线。生成器和鉴别器的目  
的是相反的,即两个生成器网络和鉴别器网络是相  
对的,一个互相交互。可见鉴别器的损耗曲线在下降,发  
电机的损耗曲线在上升,两人一直在博弈。

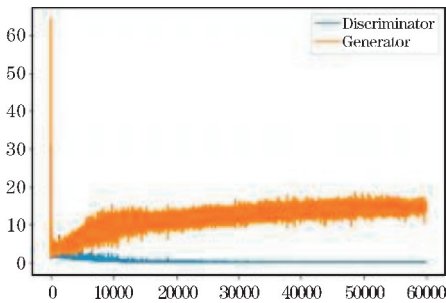


图 12 损失变化曲线图

4 结束语

参照循环生成对抗网络的结构,设计一个高质量  
的图像样式转换网络。基于 U 网的跳层结构和残差  
分块原理,发电网络利用跳层和残差网络增加多尺度  
差。判别网络采用多尺度展开卷积,提高空间几何变  
换和高分辨率图像生成效果。实验结果表明,该算法  
在图像风格转换方面有较理想的效果,并能保证图像  
风格转换的质量。

## 参考文献:

- [1] Xiang Li, Yasushi Makihara, Chi Xu, et al. Gait recognition invariant to carried objects using alpha blending generative adversarial networks[J]. Pattern Recognition, 2020, 105.
- [2] Yanbiao Zou, Xianzhong Wei, Jiaxin Chen. Conditional generative adversarial network-based training image inpainting for laser vision seam tracking [J]. Optics and Lasers in Engineering, 2020, 134.
- [3] Robert Skilton, Yang Gao. Combining object detection with generative adversarial networks for in-component anomaly detection [J]. Fusion Engineering and Design, 2020, 159.
- [4] Gaffari Çelik, Muhammed Fatih Talu. Resizing and cleaning of histopathological images using generative adversarial networks[J]. Physica A: Statistical Mechanics and its Applications, 2020, 554.
- [5] Xiaohuang Qin, Ziyang Wang, Jiepeng Yao, et al. Using a one-dimensional convolutional neural network with a conditional generative adversarial network to classify plant electrical signals[J]. Computers and Electronics in Agriculture, 2020, 174.
- [6] Robertas, Scherer Rafal. HEMIGEN: Human Embryo Image Generator Based on Generative Adversarial Networks[J]. Sensors (Basel, Switzerland), 2019, 19(16).
- [7] Gao Hongmin, Yao Dan, Wang Mingxia, et al. A Hyperspectral Image Classification Method Based on Multi-Discriminator Generative Adversarial Networks[J]. Sensors (Basel, Switzerland), 2019, 19(15).
- [8] Yao Shanjian. Research on pedestrian recognition technology based on cycle Gan and twin network [D]. Jinan: Shandong University of science and technology, 2020.
- [9] Deng Yongchang. Digital breast cancer recognition method based on convolution neural network [D]. Quanzhou: Huaqiao University, 2020.

## Image Style Conversion Algorithm based on Generative Confrontation Network

LI Quanyue, NIU Mengchen, CHEN Ruilin, CHEN Yiting, DU Shuai, XU Zixin  
(Chengdu University of Information Technology, Chengdu 610225, China)

**Abstract:** Aiming at the problem of low image quality generated by image style conversion, a high-quality image style conversion method based on generative confrontation network is proposed. This paper draws on the Cycle GAN network structure on the generation network, the method of combining the jump layer structure and the residual network in the U-Net network is used to increase the multi-scale invariance of the network; Secondly, in terms of discriminating network, a multi-scale dilated convolution discriminator is proposed to improve the spatial geometric transformation in image style conversion and generate high-resolution images. It has been verified that compared with Cycle GAN, the effect of this algorithm in image style conversion has been greatly improved, and the quality of image style conversion has been improved.

**Keywords:** image style conversion; GAN; Cycle-GAN; U-Net residual network