

[文章编号] 1003-4684(2020)01-0025-04

下采样迭代和超分辨率重建的图像风格迁移

周浩^{1,2}, 周先军^{1,2}, 邱书畅²

(1 湖北工业大学湖北省电网智能控制与装备工程技术研究中心, 湖北 武汉 430068;

2 湖北工业大学电气与电子工程学院, 湖北 武汉, 430068)

[摘要] 提出一种基于下采样迭代和超分辨率重建的图像风格转移算法,即在输入端对输入图像进行下采样,以加速整个图像风格迁移网络的迭代速度,在输出端进行超分辨率的重建,最终输出高分辨率的图像。实验结果表明,该方法减少了整个网络的迭代时间,输出的超分辨率图像也有较好效果。

[关键词] 深度学习; 图像风格迁移; 图像下采样; 超分辨率重建

[中图分类号] TP183 **[文献标识码]** A

图像风格迁移即将一张图片的语义内容用不同的风格表现出来的技术。传统图像风格迁移算法主要思路便是运用数学或者统计模型,去表征图片的纹理,从而近似达到再现图像风格的目的。深度学习算法在图像处理方面得到广泛运用之后,基于深度学习的图像风格迁移算法得到了快速的发展。Gatys 等^[1]开创性地提出了一种基于卷积神经网络的图像风格迁移算法,能很好地提取图像风格特征与内容特征,生成较好的图像风格迁移图像。Risser 等^[2]在 Gatys 的基础上添加梯度直方图损失进行约束,纹理合成和风格迁移的效果均优于 Gatys 的算法。Li 等^[3]使用马尔科夫随机场损失函数替换 Gram 损失函数,利用 MRF 的空间布局限制,提高生成图像的质量。Johnson 等^[4]在 Gatys 的算法前添加一个前置图像转移网络,通过训练前置网络参数实现快速风格转移。Ulyanov 等^[5]提出 Instance Normalization 方法代替卷积神经网络中常用的 Batch Normalization 方法,显著提升了生成网络的效果。图像风格迁移技术已经取得了较好的成果,基于深度学习的图像风格迁移算法也分为两类:基于图像迭代的图像风格迁移和基于模型迭代的图像风格迁移。

本文在前人基于图像迭代的图像风格迁移算法上做出了改进,提出了一种基于下采样迭代超分辨率重建的图像风格迁移,大量减少了网络迭代的时间,优化了整个网络的迭代效率,生成的图片效果较好。

1 基于下采样迭代超分辨率重建的图像风格迁移算法

传统基于图像迭代的风格迁移算法需要大量的计算量和时间去训练生成的图片。本文利用深度学习算法在图像超分辨率重建方面的高效性,提出一种使用下采样之后的低分辨率图像进行图像风格迁移的迭代,然后在输出端进行超分辨率重建的风格迁移算法。算法流程见图 1。

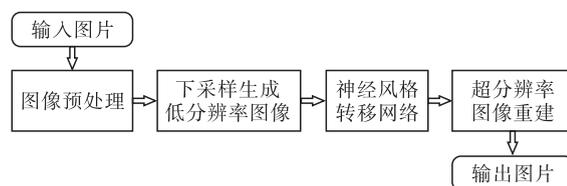


图 1 本文算法示意图

输入目标内容图片和风格图片后,首先经过图像预处理网络对内容图片和风格图片进行锐化,使图像特征更加明显;接着对预处理过后的内容图片和风格图片进行下采样,得到低分辨率图像,以降低数据量;将下采样后的内容图片和风格图片送入图像风格迁移网络进行训练,产生风格迁移图片;最后将得到的风格迁移图片进行超分辨率重建,得到同输入图像分辨率相同的风格迁移图片。

2 图像预处理和下采样

2.1 图像预处理

为了使图像的细节更突出,特征更加明显,需要

对图像进行预处理,选择常用的拉普拉斯算子对图像进行锐化,使得图像效果更好(图2)。



图2 图像预处理结果

2.2 图像下采样

为了减少图像风格迁移网络的运算量,将预处理后产生的图片进行下采样处理,选择的下采样倍数不宜过大,否则最后重建得到的风格迁移图片视觉效果会很差。



图3 图像下采样结果

3 基于图像迭代的图像风格迁移网络

基于图像迭代的风格迁移网络,是利用已经预训练好的 VGG19 网络作为图像特征提取器来分别提取内容图像的内容特征和风格图像的风格特征,并加权相加得到总的特征矩阵,并以此来设计总的损失函数。在网络训练过程中,输入训练的是白噪声图片,最后白噪声图像会变成既具有内容图像内容特征又具有风格图像风格特征的风格迁移图像。算法流程见图4。

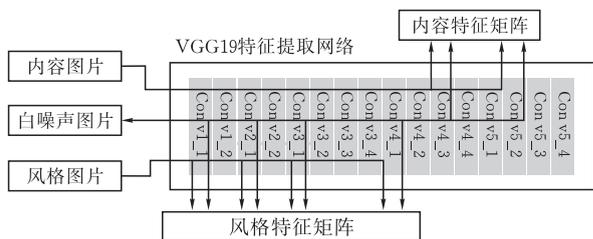


图4 图像风格迁移算法

3.1 特征提取网络建立

使用预训练好的 VGG19 网络模型来作为特征提取网络,这里提取出来的 VGG 参数全部是作为常量使用的,这些参数是不会再被训练的,在反向传播的过程中也不会改变,要训练的参数,是输入进来

的白噪声图像,使最初输入进来的白噪声图像同时具有指定的内容与风格。最开始输入一张噪音图片,然后不断地根据设定好的内容损失函数和风格损失函数对其进行调整,直到一定次数后,该图片兼具了风格图片的风格以及内容图片的内容。当训练结束时,输入层的参数就是图像风格迁移后的图片。

3.2 建立内容损失函数

在内容损失函数上,选择均方误差(MSE),该损失函数计算简单,效率高。在计算内容损失函数时,选择的内容特征模型是 conv4_2 和 conv5_2 所输出的矩阵,将这两个矩阵进行加权求和,权重各为 0.5,这样可以比较有效地反映内容图像不同层次的特征,使白噪声图像更好地学习到内容图像的特征。其中每一层的损失计算公式如下:

$$L_i = \frac{1}{2 \times M \times N} \sum_{ij} (\mathbf{X}_{ij} - \mathbf{P}_{ij})^2$$

其中, \mathbf{X} 是输入的噪声图片在对应层数所产生的特征矩阵, \mathbf{P} 是输入的内容图片在对应层所产生的特征矩阵。 M 是 \mathbf{P} 的长 \times 宽, N 是深度。最终的内容损失也就是各层内容损失之和,即

$$L_{\text{content}} = \sum L_i$$

3.3 建立风格损失函数

在风格损失函数上,选择 Gram 矩阵来表征图像风格特征,选取了 conv1_1, conv2_1, conv3_1, conv4_1 所输出的 4 个特征矩阵,然后求这些特征矩阵的 Gram 矩阵,最后加权相加得到最后的可以表征图片风格的 Gram 矩阵。同理,白噪声图像的风格特征矩阵,也是这样得出的。得到了两张图片的 Gram 矩阵,便可以用均方误差,来构建风格的损失函数:

$$L_i = \frac{1}{4 \times M^2 \times N^2} \sum_{ij} (G_{ij} - A_{ij})^2$$

其中: M 为特征矩阵的长 \times 宽, N 是特征矩阵的信道数。 G 为噪音图像特征的 Gram 矩阵, A 为风格图片特征的 GRAM 矩阵。总的风格损失如

$$L_{\text{style}} = \sum L_i$$

得到了内容损失与风格损失,要构建出总的损失函数,即将风格损失和内容损失加权相加得到总的损失函数:

$$L_{\text{total}} = \alpha L_{\text{content}} + \beta L_{\text{style}}$$

得到总的损失函数之后,将白噪声图片进行迭代,梯度下降反向传播去调整输入的白噪声的各个像素的值,从而使多次迭代之后产生的白噪声图片既具有内容图像的内容特征,又有风格图像的风格特征,从而实现风格迁移。

4 图像超分辨率重建

图像超分辨率重建是将低分辨率的图像转换成高分辨率图像的技术,传统算法中常用的方法是双线性插值法和双三次插值法,随着深度学习算法的发展,人们开始利用神经网络算法来更好地实现图像超分辨率重建。

Dong 等人在 2015 年提出一种将卷积神经网络应用于单张图像超分辨率重建的算法 SRCNN^[6],网络结构为 (conv1 + relu) — (conv2 + relu2) — (conv3) 算法结构见图 5。

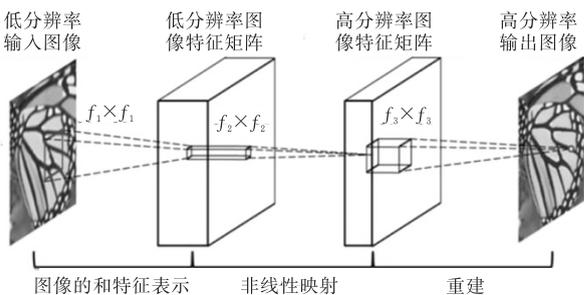


图 5 SRCNN 网络结构

图 5 算法分为三个步骤:

1) 图像特征的提取和表示,用 $f_1 \times f_1$ 的卷积核,去提取低分辨率图像的特征,并得到特征表示矩阵;

2) 低分辨率图像特征到高分辨率图像特征的非线性映射,用 1×1 的卷积核,将低分辨率图像特征非线性映射到高分辨率图像特征;

3) 高分辨率图像重建,用 $f_3 \times f_3 \times 1$ 的卷积核,得到一张单通道的高分辨率输出图像。

在该算法中,为了网络更快收敛,首先将 RGB 图像转移到 YCbCr 空间中,然后每个通道单独训练,最后再合成彩色图像。实验结果表明,训练好的 SRCNN 网络相较于传统算法,有更好的超分辨率图像重建能力。

5 实验结果与分析

5.1 实验结果

为验证本文提出算法的有效性,分别进行了两组实验。第一组 Gatys 等人的方法,仅仅使用图像风格迁移网络,输入图片为锐化之后的内容图片和风格图片,图像尺寸均为 $400 \times 400 \times 3$,图像风格迁移网络迭代次数为 3000 次;第二组,使用本文所提出的算法,下采样时将内容图像和风格图像均变成 $200 \times 200 \times 3$ 的大小,图像风格迁移网络迭代次数为 3000 次。实验结果见图 6。

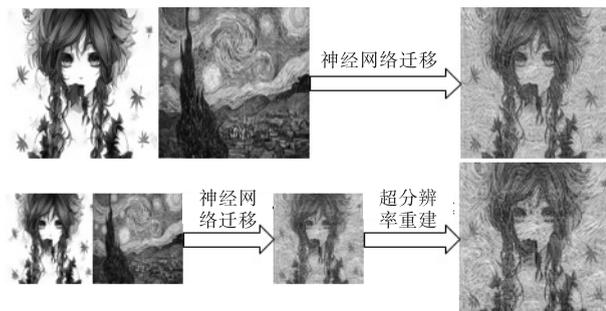


图 6 实验结果

5.2 结果分析

本次实验中,为了衡量生成的风格迁移图像的效果,分别将两组实验结果图片重新带入了损失函数,计算损失。损失函数表达式如

$$L_{\text{total}} = \alpha L_{\text{content}} + \beta L_{\text{style}}$$

其中第一组结果图片的损失值 (Current loss value) 为 1158316800.0;第二组结果图片的损失值 (Current loss value) 为 2087749200.0。

结果表明,在训练次数相同的情况下,虽然本文方法得到的图像风格迁移图片损失值大于 Gatys 等人的方法,但是本文所提出的方法相较于 Gatys 等人的算法有以下优点。

1) 节约时间成本。第二组要训练的数据只有 $200 \times 200 \times 3 = 120000$ 个,是第一组要训练的数据 $400 \times 400 \times 3 = 480000$ 个的 1/4,并且第二组中使用了已经预训练好的 SRCNN 网络,超分辨率重建过程时间很短可忽略不计,所以第二组算法所用的时间要远远低于第一组算法。在实验过程中,第一组算法运行时间约为 8 h,第二组算法运行时间约为 3 h,实验结果证明本算法可以大量减少迭代时间。

2) 图片整体效果更好。在本次第一组实验和第二组实验中的风格迁移网络部分,图像风格和特征提取网络均为 VGG19 网络,卷积核大小均为 3×3 。而在第二组实验中,因为输入图像的分辨率更低,所以此时的卷积核可以感受到更多的图片信息,因而可以连贯地改变原图片的大部分区域,从而使内容图像和风格图像融合得更好,图像风格迁移整体效果更好。

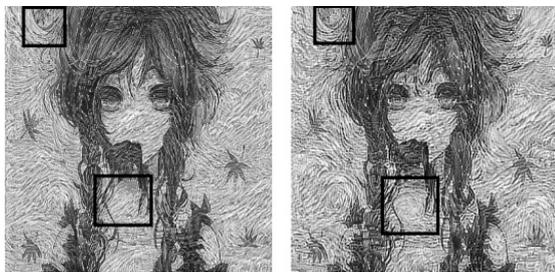


图 7 Gatys 等人实验结果 图 8 本文方法实验结果

如图 7—8,虽然本文方法得到的图片在图像细

节上有所欠缺,但是本文方法得到的图片更加贴合原风格图片(如图7—8中黑框所示),同时本文方法得到的图片在整个图像的风格和色调上也更加贴合风格图片,使内容图像和风格图像达到更好的融合。

3)更加贴合真实绘画风格。本文提出算法产生的结果图片,虽然牺牲了图像的细节,但是图像整体风格更加连贯,更加贴合平时所见到的艺术绘画作品,可以在艺术绘画辅助设计方面得到较好的应用。

6 结束语

在前人研究的基础上,针对基于图像迭代的图片风格迁移算法做出了改进,实验证明:本文所提出的方法大量减少了网络迭代所需要的时间,同时牺牲了一部分图像的细节使得内容图像和风格图像有一个更好的融合,达到了不错的风格迁移的效果。但是也有需要进一步改进的地方,例如图像下采样的倍数不宜过大,否则在输出端重建得到的图像质量较差;图像超分辨率重建网络有待优化,如果使用更多的数据,对应的训练集进行训练,在进行图像超分辨率重建时会得到效果更好的图片。

[参 考 文 献]

[1] Gatys L A, Ecker A S, Bethge M. Image style trans-

fer using convolutional neural networks [C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition(CVPR), IEEE, 2016. 2414-2423.[2] Risser E, Wilmot P, Barnes C. Stable and controllable neural texture synthesis and style transfer using histogram losses [EB/OL]. [2019-08-03], <https://arxiv.org/abs/1701.08893>.

[3] Li Chuan, Wand M. Combining markov random fields and convolutional neural networks for image synthesis [C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2016: 2479-2486.

[4] Johnson J, Alahi A, Li Fei-fei. Perceptual losses for real-time style transfer and super-resolution [C]. European Conference on Computer Vision Springer, 2016: 694-711.

[5] Ulyanov D, Vedaldi A, Lempitsky V. Instance normalization: the missing ingredient for fast stylization [EB/OL]. (2019-08-03), <https://arxiv.org/abs/1607.08022>.

[6] Dong C, Loy C C, He K, et al. Image super-resolution using deep convolutional networks[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2016, 38(2):295-307.

Image Style Transfer Algorithm Based on Downsampling Iterative and Super-Resolution Reconstruction

ZHOU Hao^{1,2}, ZHOU Xianjun^{1,2}, QIU Shuchang²

(1 Hubei Power Grid Intelligent Control and Equipment Engineering

Technology Research Center, Hubei Univ. of Tech., Wuhan 430068, China;

2 School of Electrical and Electronic Engineering, Hubei Univ. of Tech., Wuhan 430068, China)

Abstract: Based on image iteration, direct iteration takes much time for the image-style migration algorithm. In this paper, an image style transfer algorithm based on downsampling iterative super-resolution reconstruction is proposed, which is to downsample the input image at the input end to accelerate the iterative speed of the entire image style migration network, and to perform super-resolution reconstruction at the output end to output high resolution images. The experimental results show that the method reduces the iteration time of the whole network for content fusion at low resolution, and the reconstructed super-resolution image also has better effect.

Keywords: deep learning; image style transfer; image downsampling; super-resolution reconstruction

[责任编辑:张岩芳]