基于残差式神经网络的局部风格迁移方法

孙劲光，刘鑫松*
辽宁工程技术大学电子与信息工程学院，辽宁 葫芦岛 125105

摘要 风格迁移技术迅速发展的今天，全局风格迁移技术已基本成型，但它在实际的应用过程中存在不能对图片的目标区域进行局部风格迁移等问题。针对以上问题，本文在卷积神经网络的基础上结合残差网络，提出了一种基于残差式神经网络的局部风格迁移方法。首先，利用掩模技术对内容图进行分割，提取目标区域；其次，卷积神经网络提取图片特征并进行特征融合；然后，使用残差网络加快生成图的形成速度；最后，通过反卷积生成一张只对目标区域完成风格迁移的图片。在 Microsoft Coco2014 数据集上设计了多个实验，实验结果表明，所提出的基于残差式神经网络的局部风格迁移网络模型具有较好的局部风格转换能力，并且具有较高的执行效率。

关键词 图像处理；风格迁移；局部分割；特征融合；残差网络；反卷积
中图分类号 TP389.1 **文献标志码** A **doi**：10.3788/LOP57.081012

Local Style Migration Method Based on Residual Neural Network

Sun Jinguang, Liu Xinsong*
School of Electronic and Information Engineering, Liaoning Technical University, Huludao, Liaoning 125105, *China*

Abstract With the rapid development of style migration technology, the global style migration technology has basically taken shape, but in the actual application process, there are problems such as the local style migration of the target area of the picture. Aiming at the above problems, this paper combines the residual network based on the convolutional neural network, and proposes a local style migration method based on residual neural network. Firstly, the mask is used to segment the content map to extract the target region. Secondly, the convolutional neural network extracts the image features and performs feature fusion. Then, the residual network is used to speed up the formation of the graph. Finally, the deconvolution is generated. A picture that only completes the style transition for the target area. In this paper, the several experiments are designed on the Microsoft Coco2014 dataset. The experimental results show that the local style migration network model based on residual neural network has better local style conversion ability and higher execution efficiency.

Key words image processing; style migration; local segmentation; feature fusion; residual network; deconvolution
OCIS codes 100.3008; 100.4994; 100.4996

1 引 言

随着计算机视觉在电影、动画、游戏创作、时装设计等方面的快速发展，风格迁移已经成为创作特定风格艺术作品的重要手段。风格迁移就是把一个图片的风格变成另一种风格，使得创作出的作品更具有艺术魅力，既可以降低创作成本，又可以节省制作时间。

正因为风格迁移方法在计算机视觉方面所表现出的强大优势，所以它深受研究学者的青睐，进而一系列研究成果被相继提出。Gatys 等[1]提出了一种基于神经网络的图片风格迁移方法，这也是最先被提出的方法。该方法一经提出便风靡一时，它解决了传统方法手动建模所存在的复杂过程问题。其采用深度学习技术对纹理特征进行建模，进而生成一张具有给定风格的内容图片，但同时它也存在着一定的弊端，需要不断迭代来生成图片，耗时过长。为了解决以上的问题，Johnson 等[2]在 2016 年提出采

收稿日期：2019-07-30；**修回日期**：2019-09-09；**录用日期**：2019-09-12
基金项目：国家自然科学基金青年基金(61702241)
* **E-mail**：song_0501qq.com@foxmail.com

用感知损失函数训练一个前馈网络来完成图片风格转换的任务，这种方法把转换的速度提高了三个数量级，但可转换的风格仅限于经过训练的风格。Yanai 等[3]于 2017 年提出一种有条件的快速转换网络，这种方法通过添加条件输入，使网络可以同时训练多个风格。Li 等[4]在 2017 年提出了一种简单而有效的通用风格转换方法，该方法不需要学习每种风格，通过训练一个用于图像重建的自动编码器来展开图像生成过程，将增白和着色转换集成到前馈过程中，以匹配内容和风格的中间特征之间的统计分布和相关性。

虽然以上研究都取得了不错的效果，但在某些情况下，创作者需要在特定的目标区域创作完成特定风格的艺术作品。例如，在室内设计过程中，用户只希望将渲染图中电视背景墙区域的风格由田园风改为现代简约风等。而现有技术只能完成整张渲染图所有区域的风格迁移，因此，局部风格迁移方法的研究具有十分重要的实际意义。为了解决图片目标区域的风格迁移这一问题，本文提出了一种基于残差式神经网络的局部风格迁移方法，该方法可以完成图片目标区域的局部风格迁移。

2 本文方法

为了解决图片目标区域的局部风格迁移问题，本文提出了一种基于残差式神经网络的局部风格迁移方法，网络模型如图 1 所示，算法流程如下所示。

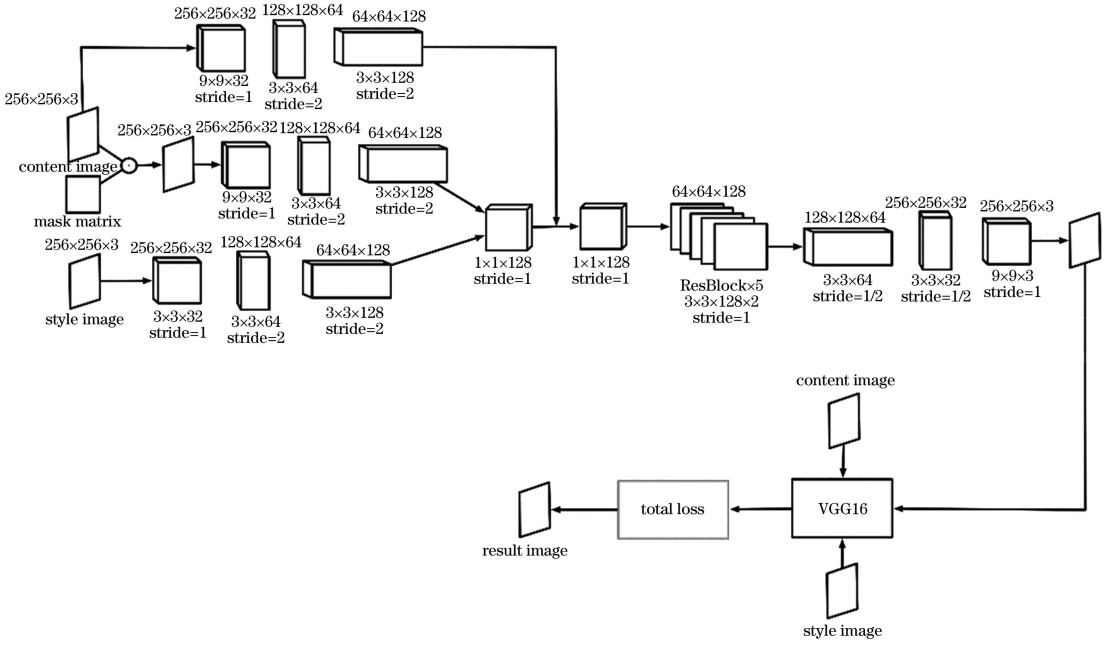

图 1 局部风格迁移网络模型

Fig. 1 Local style migration network model

Algorithm 1: Local style migration algorithm based on residual neural network

Input: mask map, content map, style map.

Output: a target area has style map style information, and the non-target area has a color picture with the original style information of the content map.

Model construction:

Step1, extracting the target area of the content map by combining the mask map of the target area with the content map;

Step 2, using two convolutions with a step size of 2, respectively down-sampling the target area, the content map, and the style map of the content map, extracting the image features, and obtaining the position information of the target area;

Step3, adding two 1×1 convolutional layers[5], the first one is used to realize the feature fusion of the target area feature and the style map feature of the content map[6-7], and the second one is used to implement the mask after the non-mask Feature fusion of target area features and content map features;

Step 4, is 5 residual blocks[8], and each residual block contains two 3×3 convolution layers, which effectively reduces the computational cost of the network;

Step 5, is completed by two convolutions of 1/2 step size[9];
Parameter training:
Step6, calculating content loss and style loss, adjusting parameters according to the loss value, until the end of the iteration, the loss value is minimized;
Step7, output a 256 × 256 color picture, the target area of the picture has style map style information, and the non-target area has the original style information of the content map.

2.1 图片分割

图片分割[10-12]的本质就是找到目标区域的边界，随后将图片按照纹理空间、形状结构等特征分割成若干个互不相容的小区域，使这些特征在目标区域呈现出一定的相似性，但与非目标区域有明显差异，从而将目标区域从背景中分离出来。

本文采用掩模技术实现图片的局部分割，提取目标区域。该方法可以屏蔽非目标区域，使网络在处理过程中不用计算该区域数据，减少了计算量，加快了运行速度。其原理是用目标区域掩模图中的每个像素和内容图中的每个对应像素进行与运算[13]，得到目标区域图片。例如，一个 3×3 的掩模图与 3×3 的内容图进行与运算，得到的结果图为

$$\begin{pmatrix}0 & 0 & 1\\ 1 & 0 & 1\\ 1 & 1 & 1\end{pmatrix}\&\begin{pmatrix}57 & 22 & 59\\ 0 & 0 & 220\\ 150 & 0 & 8\end{pmatrix}=\begin{pmatrix}0 & 0 & 59\\ 0 & 0 & 220\\ 150 & 0 & 8\end{pmatrix}。$$

此时，目标区域内图像值保持不变，而非目标区域图像值全为 0，如图 2 所示。

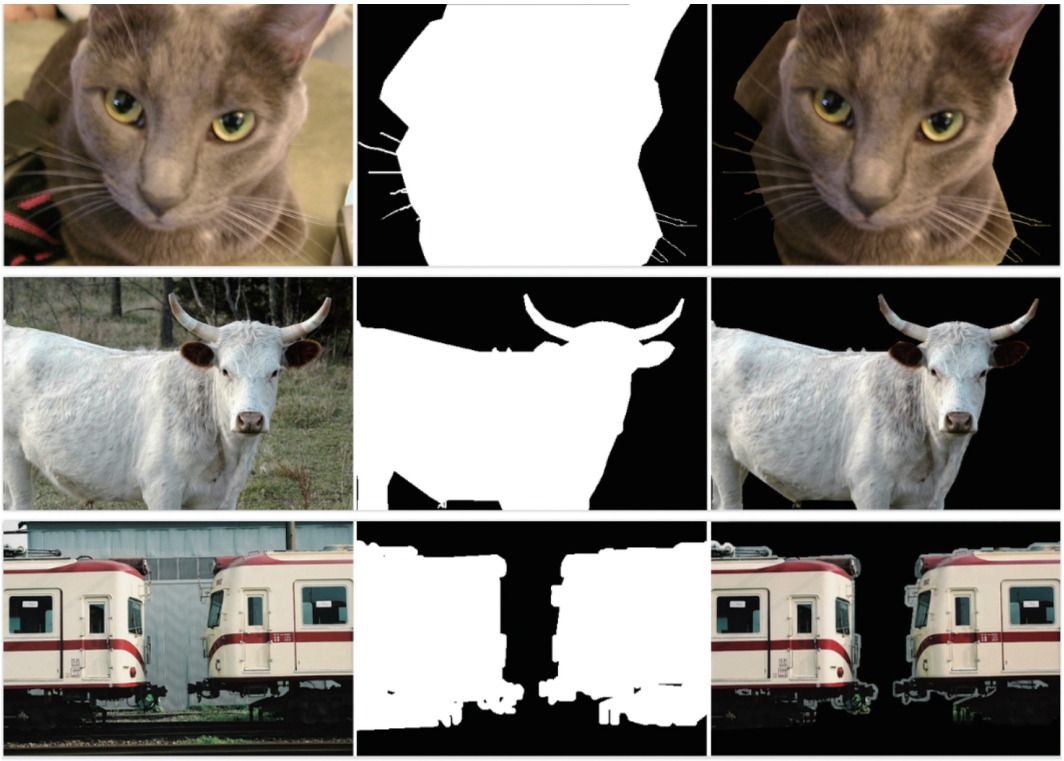

图 2 掩模效果图

Fig. 2 Mask effect maps

2.2 特征提取

卷积神经网络(CNN)是一种前馈神经网络，具有强大的表征学习能力，因此被广泛应用于计算机视觉等领域。CNN 模型较传统的神经网络具有更深的层数，提取的特征更加丰富。本文采用 VGG16[14]卷积神经网络模型提取图片特征，相比于 LeNet、AlexNet 等卷积神经网络模型，其规模是 AlexNet 的 2 倍以上并拥有规律的结构和更少的参数量，简化了神经网络的拓扑结构，使性能更加优越。

VGG16 的隐含层由 13 个卷积层、5 个池化层和 3 个全连接层组成，其网络架构如图 3 所示。卷积层是 VGG16 的核心，仅使用 3×3 的卷积核并保持卷积层中输出特征图尺寸不变，在此层求取输入图片的局部数据与卷积核的内积，以便提取图片的局部特征。池化层的输入来自上一个卷积层，主要作用是对输入图像进行压缩，使输出的特征图尺寸减半。采用最大池化与平均池化两种方法，其中最

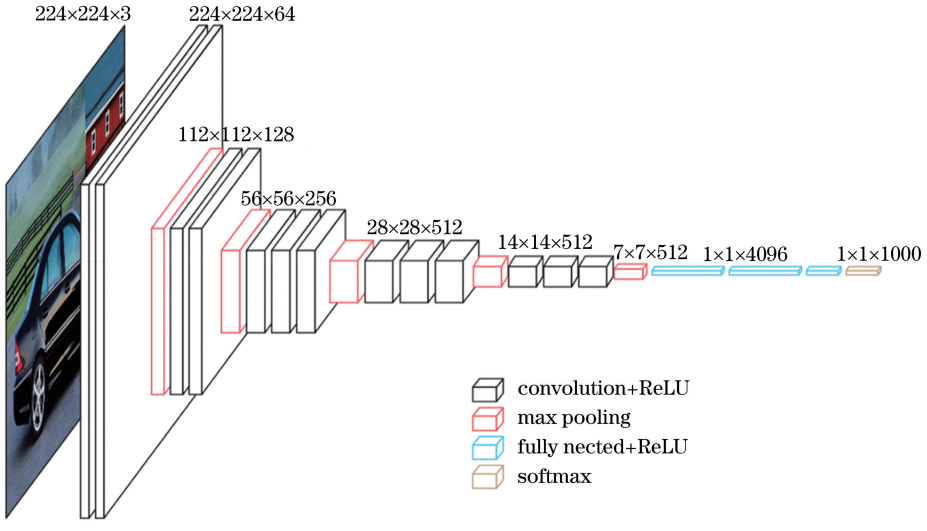

图 3 VGG 网络结构图

Fig. 3 VGG network architecture

大池化主要运用于卷积层，输出每个区域的最大值；平均池化用于全连接层，计算的是区域内的平均值。全连接层用 Softmax 进行激活。本文采用 VGG16 模型提取图像特征，将内容图、风格图及带有目标区域的图片中各点像素值矩阵作为网络输入，通过卷积、池化等操作获得图片特征。

2.3 残差网络

残差网络[8,15-17]用来解决更深层网络难以训练的问题，减轻网络的训练负担。假设把网络结构设计为 $H(x)=F(x)+x$，这样可以将 $H(x)=x$ 转化成学习一个残差函数 $F(x)=H(x)-x$。此时，只需令 $F(x)=0$，则 $H(x)=x$ 即为最优解。残差结构可简单的写成如下形式：

$$x_{l+1}=x_l+F(x_l,W_l)。\tag{1}$$

通过递归可以得到任意深层单元 L 特征的表达式为

$$x_L=x_l+\sum_{i=l}^{L-1}F(x_i,W_i),\tag{2}$$

式中：l 表示任意浅层单元；L 表示任意深层单元；x 表示某一单元的特征；W 表示权重值。浅层特征加上一个 $\sum_{i=l}^{L-1}F$ 残差函数，表明任意单元 L 和 l 之间具有残差特性。残差结构使用了跳跃连接的连接方式，如图 4 所示，计算公式如下：

$$F=W_2\cdot\sigma(W_1x),\tag{3}$$

$$y=F(x,W_i)+x。\tag{4}$$

由图 4 可以看到，x 为输入值，$F(x)$ 是经过第一层线性变化和激活函数 σ 后的输出，在第二层进行线性变化之后，$F(x)$ 加入了这一层输入值 x，然后再进行激活后输出。

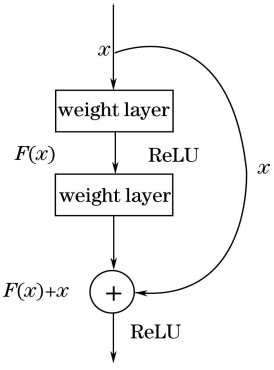

图 4 残差网络示意图

Fig. 4 Residual network schematic

残差网络的使用有助于解决梯度消失和梯度爆炸问题，可以在训练更深网络的同时，保证其良好的性能。

2.4 局部风格迁移损失函数

本文定义了局部风格迁移的损失函数，使用 VGG16 作为损失网络、优化网络权重，使生成图的内容特征更接近于内容图、生成图的风格特征更接近于风格图。

内容损失表示为

$$L_{\text{cont}}=\frac{\boldsymbol{M}_t}{CHW}\left\|\boldsymbol{\varphi}(y)-\boldsymbol{\varphi}(y_c)\right\|_2^2。\tag{5}$$

对于生成图，只计算目标区域的局部损失，将生成图和内容图通过 VGG16 网络得到特征图。$\boldsymbol{\varphi}(y)$、$\boldsymbol{\varphi}(y_c)$ 分别表示生成图和内容图在 VGG16 的 conv3_3 层的特征，C、H、W 分别表示对应层特征的通道数、高、宽，$\boldsymbol{M}_t$ 表示掩模矩阵。计算两特征图之间的欧氏距离，即 2 范数，目的是防止模型过

拟合、提高模型的泛化能力。然后，最小化该距离，使生成图和内容图在内容上保持一致。

在损失网络中低层的特征损失较小，但不能与生成图产生较大区分，所以在高层提取内容图特征。内容图中的内容和结构会被保留，但颜色、纹理等细节将不再存在，这一问题将在风格损失中得到解决。

风格损失：为了解决内容损失中在损失网络高层提取特征会使颜色、纹理等细节消失的问题，本文通过优化风格损失函数，使目标区域的风格与风格图保持一致，非目标区域的风格与内容图保持一致，公式为

$$L_{\text{style}} = \| \boldsymbol{G}(y) - \boldsymbol{G}(y_{\text{s}}) \|_{\text{F}}^{2}, \tag{6}$$

式中：$\boldsymbol{G}$ 表示 Gram 矩阵，计算生成图 y 和风格图 y_{s} 两特征之间的偏心协方差，即没有减去均值的协方差矩阵。既能体现出特征，又能体现出不同特征间的紧密程度。Gram 矩阵定义为 n 维欧氏距离中任意 k 个向量 $(\boldsymbol{a}_1 \quad \boldsymbol{a}_2 \quad \cdots \quad \boldsymbol{a}_k)$ 的内积所组成的矩阵，公式为

$$G = \begin{pmatrix} \boldsymbol{a}_1^{\text{T}} \\ \boldsymbol{a}_2^{\text{T}} \\ \vdots \\ \boldsymbol{a}_k^{\text{T}} \end{pmatrix} \cdot (\boldsymbol{a}_1 \quad \boldsymbol{a}_2 \quad \cdots \quad \boldsymbol{a}_k) = \begin{pmatrix} \boldsymbol{a}_1^{\text{T}}\boldsymbol{a}_1 & \boldsymbol{a}_1^{\text{T}}\boldsymbol{a}_2 & \cdots & \boldsymbol{a}_1^{\text{T}}\boldsymbol{a}_k \\ \boldsymbol{a}_2^{\text{T}}\boldsymbol{a}_1 & \boldsymbol{a}_2^{\text{T}}\boldsymbol{a}_2 & \cdots & \boldsymbol{a}_2^{\text{T}}\boldsymbol{a}_k \\ \vdots & \vdots & & \vdots \\ \boldsymbol{a}_k^{\text{T}}\boldsymbol{a}_1 & \boldsymbol{a}_k^{\text{T}}\boldsymbol{a}_2 & \cdots & \boldsymbol{a}_k^{\text{T}}\boldsymbol{a}_k \end{pmatrix}。 \tag{7}$$

总损失为

$$L_{\text{total}} = \lambda_{\text{c}} L_{\text{cont}} + \lambda_{\text{s}} L_{\text{style}}, \tag{8}$$

式中：λ_{c} 和 λ_{s} 分别表示内容图和风格图的权重因子。

通过本文定义的局部风格迁移损失函数，可以训练局部风格迁移模型，捕捉到目标区域的空间位置信息，进而只对目标区域进行风格迁移，保持非目标区域的风格不变。

3 实　　验

3.1 实验设置

本文的实验环境是 Python3.6.4，PC 处理器为 i5 8600K，内存 32 GB，Win10 操作系统，同时配备 GTX1060 显卡，主要基于开源的机器学习框架 TensorFlow-gpu、CUDA 和 CuDNN[18] 来实现 GPU 加速。本文使用 Microsoft Coco2014 数据集[19]进行训练，将 80000 张训练图片的大小均调整为 256×256。实验中设 batch-size=4，初始学习速率 0.001，迭代 40000 次，并用 Adam[20] 优化，在此训练数据上运行大约两个周期，耗时约为 4 h。

取 VGG16 网络的 conv3_3 层计算内容损失，conv2_2，conv3_2，conv4_2，conv5_2 层计算风格损失，$\lambda_{\text{c}}/\lambda_{\text{s}}=10^{-2}$，与文献[2]的实验效果进行了对比，训练数据均从 Microsoft Coco2014 数据集里抽取，实验结果如图 5 所示。

通过这些实验结果可以看出，局部风格迁移网络模型可以很好地完成图片目标区域的局部风格迁移工作。例如，图 5 中以汽车为目标区域的图像，文献[2]的方法只能将整张图片的风格进行转换，而不能识别汽车区域与非汽车区域，更不能完成只对汽车区域进行风格转换的工作。本文方法则可以很好地识别二者，并将汽车区域的风格转换海浪等风格，非汽车区域保持原有风格。

3.2 内容与风格的权重关系

在生成图片时，最小化的局部风格迁移损失函数是内容损失和风格损失的线性组合，可以表示内容或风格的约束。如果过多强调内容，虽然可以清晰地表现照片内容、结构等细节信息，但是图片的风格却无法很好地绘制，例如，图 6 中 $\lambda_{\text{c}}/\lambda_{\text{s}}=10^{-1}$。但强调风格过多就会导致图片只包含有效的纹理特征，而几乎没有图片的内容信息，例如，图 6 中 $\lambda_{\text{c}}/\lambda_{\text{s}}=10^{-4}$。

调整内容和风格的权重关系，寻找最优权重关系定量值。以图 6 中猫咪图像为例，当 $\lambda_{\text{c}}/\lambda_{\text{s}}=10^{-2}$ 时，图片既可以清晰表现出猫咪的脸型、眼睛、胡须等细节信息，又包含有效的海浪特征，视觉效果令人满意。所以，本文将 $\lambda_{\text{c}}/\lambda_{\text{s}}=10^{-2}$ 作为最优权重关系定量值进行接下来的实验。

3.3 不同层的生成效果

图片生成过程中风格的表现需要网络层多尺度的表示，这些网络层的位置和数量决定了风格的视觉体验。在实验中可以发现，网络的高层可以在最大尺度上保留图片的纹理特征，生成更平滑、更生动的视觉效果，因此，本文在 conv2_2，conv3_2，conv4_2，conv5_2 上提取风格特征。

为了分析不同层生成图的效果，设置相同图片和参数配置，$\lambda_{\text{c}}/\lambda_{\text{s}}=10^{-2}$ 情况下对照片目标区域风格化，如图 7 所示。当在网络的低层上进行风格化时，目标区域保留了内容图更多的结构等细节信息，纹理几乎不融合进内容图中。相反，在网络的高层上进行风格化时，内容图的细节信息依旧被保留，

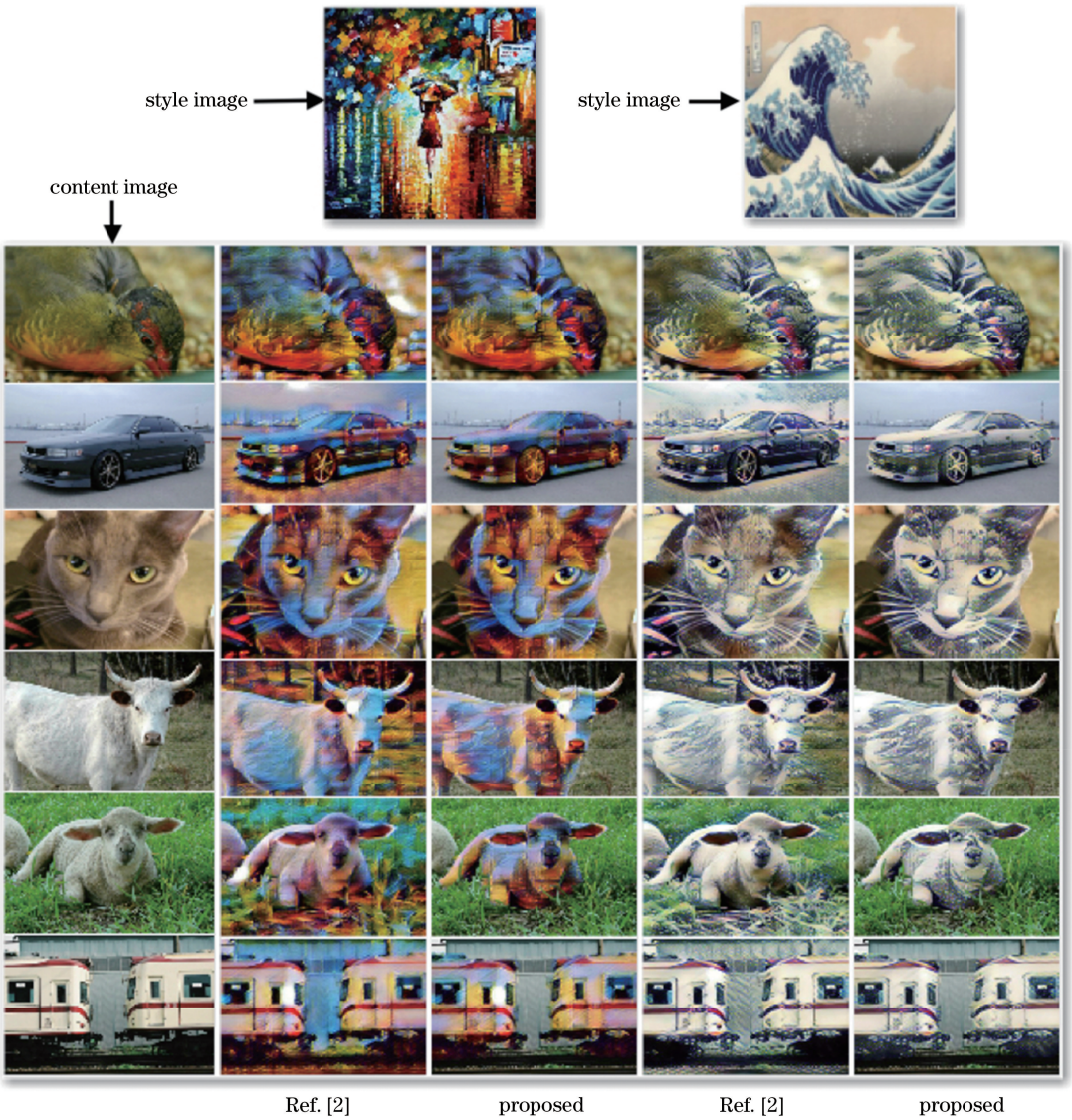

图 5 实验结果

Fig. 5 Experimental results

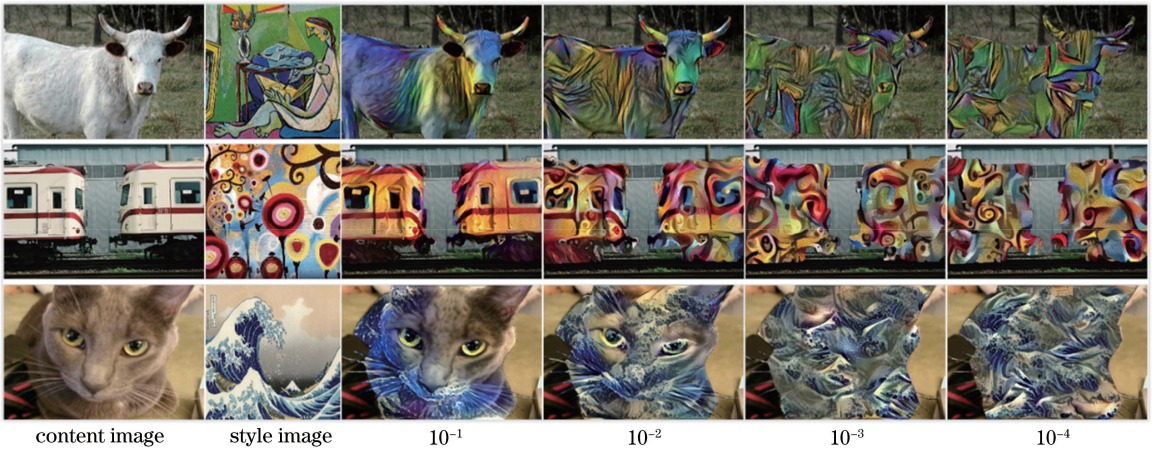

图 6 内容与风格的权重关系图

Fig. 6 Weight diagrams of content and style

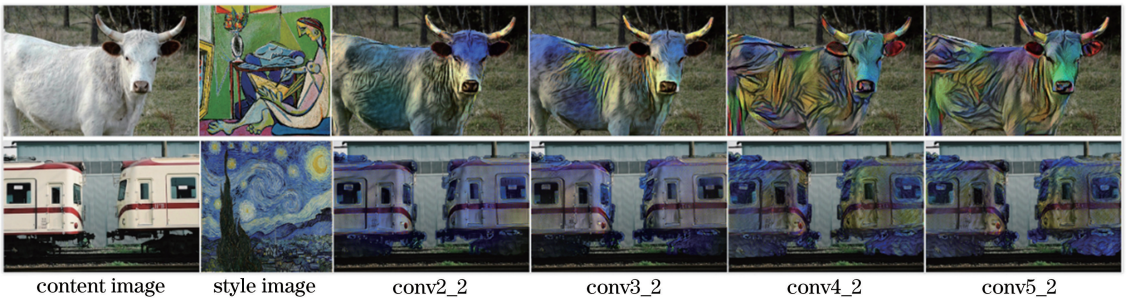

图 7 不同层的生成效果图

Fig. 7 Generation effect of different layers

风格图的纹理和内容图目标区域的内容恰当地融合在一起。

3.4 运行速度的比较

尽管本文的网络是用 256×256 尺寸训练的，但在 512×512 和 1024×1024 的情况下均能成功完成图片的局部风格迁移工作。本文方法不需要对网络模型进行重新训练，仅需进行一次前馈计算即可，减少了时间成本。表 1 给出了本文方法和其他方法在 256×256、512×512 和 1024×1024 尺寸下的运行速度的比较数据。因为文献[1]采用迭代优化的方法，所以本文以文献[1]方法迭代 500 次的时间与其他方法进行比较。

表 1 时间对比表

Table 1 Time comparison s

Method	256×256	512×512	1024×1024
Ref. [1]	14.327	51.192	200.354
Ref. [2]	0.014	0.045	0.166
Ref. [21]	0.017	0.064	0.254
Ref. [4]	0.620	1.139	2.947
Proposed	0.023	0.061	0.215

实验结果表明，本文方法在时间上优于大部分风格迁移方法，唯独略慢于文献[2]方法。这是因为文献[2]中使用的风格图是预先训练好的，并且只能使用预先训练好的风格图，不能处理实时风格图像。本文方法可以使用任意一张图片作为目标区域风格迁移的风格图，不需要预训练，节省时间成本。如此看来，本文在综合时间效率上优于文献[2]方法。

4 结 论

本文结合了卷积神经网络、掩模分割技术和残差网络的方法，提出基于残差式神经网络的局部风格迁移方法，使用局部风格迁移损失函数来训练网络。对图片目标区域的局部风格转换应用该方法，达到了很好的生成效果，在时间效率上有明显的提高。

在未来的工作中，将加强对目标区域边缘的优化，加快生成速度，并期望把局部风格迁移网络模型应用在更多其他的局部转换工作中，如视频等。

参 考 文 献

[1] Gatys L A, Ecker A S, Bethge M. Image style transfer using convolutional neural networks[C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE, 2016: 2414-2423.

[2] Johnson J, Alahi A, Li F F. Perceptual losses for real-time style transfer and super-resolution[C]// Leibe B, Matas J, Sebe N, et al. Computer vision-ECCV 2016. Lecture notes in computer science. Cham: Springer, 2016: 694-711.

[3] Yanai K, Tanno R. Conditional fast style transfer network[C]// Proceedings of the 2017 ACM on International Conference on Multimedia Retrieval, June 6-9, 2017, Bucharest, Romania. New York: ACM, 2017: 434-437.

[4] Li Y, Fang C, Yang J, et al. Universal style transfer via feature transforms[C]// Advances in Neural Information Processing Systems, December 4-9, 2017, Long Beach, CA, USA. San Diego: NIPS, 2017: 386-396.

[5] Iizuka S, Simo-Serra E, Ishikawa H. Let there be color!: joint end-to-end learning of global and local image priors for automatic image colorization with simultaneous classification[J]. ACM Transactions on Graphics (TOG), 2016, 35(4): 110.

[6] Zhu D R, Xu L, Wang F B, et al. Multi-focus image fusion algorithm based on fast finite shearlet transform and guided filter [J]. Laser & Optoelectronics Progress, 2018, 55(1): 011001.
朱达荣, 许露, 汪方斌, 等. 基于快速有限剪切波变换与引导滤波的多聚焦图像融合算法[J]. 激光与光

电子学进展, 2018, 55(1): 011001.

[7] Zhang D X, Tan Y Q. Natural texture synthesis algorithm based on convolutional neural network and edge detection [J]. Laser & Optoelectronics Progress, 2019, 56(13): 131001.
张定祥, 谭永前. 基于卷积神经网络和边缘检测的自然纹理合成算法[J]. 激光与光电子学进展, 2019, 56(13): 131001.

[8] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition [C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE, 2016: 770-778.

[9] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[C] // 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 7-12, 2015, Boston, MA, USA. New York: IEEE, 2016: 3431-3440.

[10] Zhou L L, Jiang F. Survey on image segmentation methods [J]. Application Research of Computers, 2017, 34(7): 1921-1928.
周莉莉, 姜枫. 图像分割方法综述研究[J]. 计算机应用研究, 2017, 34(7): 1921-1928.

[11] Guo C C, Yu F Q, Chen Y. Image semantic segmentation based on convolutional neural network feature and improved superpixel matching[J]. Laser & Optoelectronics Progress, 2018, 55(8): 081005.
郭呈呈, 于凤芹, 陈莹. 基于卷积神经网络特征和改进超像素匹配的图像语义分割[J]. 激光与光电子学进展, 2018, 55(8): 081005.

[12] Feng C X, Wang X L. Convolution-deconvolution image segmentation model for fusion features and decision [J]. Laser & Optoelectronics Progress, 2019, 56(1): 011008.
冯晨霄, 汪西莉. 融合特征和决策的卷积-反卷积图像分割模型[J]. 激光与光电子学进展, 2019, 56(1): 011008.

[13] Li M J. Logic operation in arithmetic[J]. Journal of Electrical & Electronic Education, 2017, 39(6): 115-117.
黎明杰. 逻辑运算的算术实现[J]. 电气电子教学学报, 2017, 39(6): 115-117.

[14] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[J/OL]. (2015-04-10)[2019-07-29]. https://arxiv.xilesou.top/abs/1409.1556.

[15] Zhang K, Sun M, Han T X, et al. Residual networks of residual networks: multilevel residual networks[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2018, 28(6): 1303-1314.

[16] Han D, Kim J, Kim J. Deep pyramidal residual networks[C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2016: 5927-5935.

[17] Shen F L, Gan R, Zeng G. Weighted residuals for very deep networks[C] // The 2016 3rd International Conference on Systems and Informatics (ICSAI), November 19-21, 2016, Shanghai, China. New York: IEEE, 2017: 936-941.

[18] Chetlur S, Woolley C, Vandermersch P, et al. cuDNN: efficient primitives for deep learning [J/OL]. (2014-12-18)[2019-07-29]. https://arxiv.xilesou.top/abs/1410.0759.

[19] Lin T Y, Maire M, Belongie S, et al. Microsoft COCO: common objects in context[C] // Fleet D, Pajdla T, Schiele B, et al. Computer vision-ECCV 2014. Lecture notes in computer science. Cham: Springer, 2014: 740-755.

[20] Kingma D P, Ba J. Adam: a method for stochastic optimization[J/OL]. (2017-01-30)[2019-07-29]. https://arxiv.xilesou.top/abs/1412.6980.

[21] Zhang H, Dana K. Multi-style generative network for real-time transfer [C] // Proceedings of the European Conference on Computer Vision (ECCV), September 8-14, 2018, Munich, Germany. New York: IEEE, 2018.