

# 结构细化的神经风格迁移

沈瑜 杨倩\* 陈小朋 苑玉彬 张泓国 王霖

(兰州交通大学电子与信息工程学院 兰州 730070)

**摘要:** 风格迁移过程中风格元素均匀分布在整幅图像中会使风格化图像细节模糊, 现有的迁移方法主要关注迁移风格的多样性, 忽略了风格化图像的内容结构和细节信息。因此, 该文提出结构细化的神经风格迁移方法, 通过增加边缘检测网络对内容图像的轮廓边缘进行提取实现风格化图像内容结构的细化, 凸显内容图像中的主要目标; 通过对转换网络中的常规卷积层的较大卷积核进行替换, 在具有相同的感受野的条件下, 使网络模型参数更少, 提升了迁移速度; 通过对转换网络中的常规卷积层添加自适应归一化层, 利用自适应归一化在特征通道中检测特定样式笔触产生较高的非线性同时保留内容图像的空间结构特性来细化生成图像的结构。该方法能够细化风格化图像的整体结构, 使得风格化图像连贯性更好, 解决了风格纹理均匀分布使得风格化图像细节模糊的问题, 提高了图像风格迁移的质量。

**关键词:** 图像处理; 深度学习; 神经网络; 风格迁移; 边缘检测; 归一化

**中图分类号:** TN911.73; TP391

**文献标识码:** A

**文章编号:** 1009-5896(2021)08-2361-09

**DOI:** 10.11999/JEIT200211

## Structural Refinement of Neural Style Transfer

SHEN Yu YANG Qian CHEN Xiaopeng YUAN Yubin

ZHANG Hongguo WANG Lin

(College of Electronics and Information Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China)

**Abstract:** In the process of style transfer, stylized image details are blurred when style elements are evenly distributed in the whole image. Besides, the existing style transfer methods mainly focus on the diversity of transferred styles, ignoring the content structure and details of the stylized images. To this end, a neural style transfer method of structure refinement is proposed, which refines the content structure of stylized image by adding edge detection network to extract the contour edge of the content image to highlight the main objectives in the content image. By replacing the larger convolution kernel of the conventional convolution layer in the transfer network, the model parameters of the transfer network are reduced, and the transfer speed is improved, while ensuring that the original receptive field is unchanged. Through the adaptive normalization of the conventional convolution layer, the structure of the generated image is refined by using the adaptive normalization to detect certain style of stroke in the feature channel to produce high nonlinearity while preserving the spatial structure of the content image. The method can refine the overall structure of the stylized image, make the stylized image more coherent, that the stylized image details are blurred due to the uniform distribution of style texture, and improve the quality of image style transfer.

**Key words:** Image processing; Deep learning; Neural network; Style transfer; Edge detection; Normalization

### 1 引言

风格迁移是一幅图像的语义内容用力一幅图像的风格纹理进行表示<sup>[1]</sup>。深度神经网络凭借其强大

的图像表示能力<sup>[2]</sup>, 推动了神经风格转移方法的发展。然而, 近几年风格迁移都主要关注风格迁移的速度和多样性, 在保证迁移速度和多样性的前提下, 怎样更好地表现出风格化图像的细节是需要急需解决的问题。开创性的神经风格迁移方法<sup>[3,4]</sup>使用卷积神经网络将图像的内容和风格特征表示进行分离, 并独立处理高层特征来实现图像风格迁移, 获得了非常可观的艺术效果。Luan等人<sup>[5]</sup>将语义分割和风格迁移相结合实现了更逼真地真实图像风格

收稿日期: 2020-03-25; 改回日期: 2021-01-30; 网络出版: 2021-07-21

\*通信作者: 杨倩 13662175532@163.com

基金项目: 国家自然科学基金(61861025)

Foundation Item: The National Natural Science Foundation of China (61861025)

转换。Huang等人<sup>[6]</sup>提出了自适应实例标准化(AdaIN)层与迭代优化相结合的前馈方法快速地实现任意样式的实时转换。Li等人<sup>[7]</sup>提出了能够捕获马尔可夫碎片的特征统计信息的马尔可夫生成对抗网络,该网络直接将内容图像转换成艺术画作。Dumoulin等人<sup>[8]</sup>搭建了可扩展的深层网络实现了任意风格的图像风格转换。Chen等人<sup>[9]</sup>提出的卡通化生成对抗网络生成高质量的卡通图像。Johnson等人<sup>[10]</sup>从预训练网络中提取的高级特征来定义和优化感知损失函数提升风格转换的速度。

风格迁移中,当输入图像具有复杂空间布局时,迁移结果将样式元素均匀地分布在整个图像中,使整体结构不可识别。对于结构变形敏感的输入,纹理均匀分布模糊了细节,破坏了原结构。因此,本文提出了细节细化的风格迁移方法,将检测网络的卷积层进行输出,设置不同步长,获得多尺度多层次的边缘特征图,再选取不同特征图进行加权融合,获得边缘特征图,用边缘特征图对迁移过程进行纹理分布的控制;在转换网络中,在非残差卷积层后面引入AdaIN层,AdaIN将特征图在对应的通道中匹配均值和方差,计算仿射参数,以此保留内容图像的空间结构;用小卷积核替代大卷积核能保证相同的感受野,增加非线性,减少参数和计算量。本文搭建的网络模型能够实现多种风格迁移,风格化图像空间结构能够得到细化。

## 2 风格迁移模型

图像风格迁移主要通过以下两个方面实现:(1)对不同的色彩通道进行不同的处理实现对颜色的控制。(2)为了对内容图像先进行语义分割再对其进行风格转换或者对不同的区域先进行标记再进行不同纹理迁移,使迁移后的图像符合自然图像的语义内容。风格迁移产生不合理的空间布局会使生

成图像的语义内容扭曲,从而无法识别出目标。因此,本文对风格迁移的纹理分布进行了细化,网络模型如图1所示。

### 2.1 边缘检测网络

多尺度是对信号不同程度采样<sup>[11]</sup>,具有不同参数和接受域大小,将数据输入到多个流中,然后将各个流产生串联的特征响应输入到全局输出层中,获得不同尺度下的不同特征。常见的多尺度特征融合网络有并行多分支网络<sup>[12]</sup>和串行跳跃连接结构<sup>[13]</sup>,两者都是在不同的感受野下进行特征提取。本文采用并行多分支网络结构<sup>[14]</sup>,将边缘检测中将边缘映射组合在一起,结构如图2所示。本文的边缘检测网络基于VGG-19网络,并做了如下改进:(1)将Conv1和Conv2的第2个卷积层的边缘特征图与Conv3,Conv4和Conv5的后两个卷积层的边缘特征图进行输出。(2)本文不使用第5个池化层和3个全连接层。因为随着段卷积步长增大,产生的边缘检测图太模糊,不利于生成更精确的边缘特征图。(3)添加加权融合层,对多尺度多层次的特征图进行融合,获得整体边缘特征图,并运用多路径反向传播优化获得最终误差最小的边缘特征图。

本文的边缘检测网络分为5个阶段,每个阶段设置不同的卷积步长(表1所示),能够获得多层次多尺度的边缘特征图,选取不同阶段的8层边缘特征输出进行加权融合,不同融合程度的边缘检测效果如图3所示。边缘检测精度使用固定轮廓阈值(ODS)、图像最佳阈值(OIS)和平均精度(AP)进行评估,如表2所示。本文的边缘检测图融合了更多的细节信息,全局结构边缘更加清晰,客观指标均取得较好的结果。

### 2.2 迁移网络

本文搭建的风格迁移网络分为编码器、转换网络和解码器3个部分。编解码器是用VGG-19网络

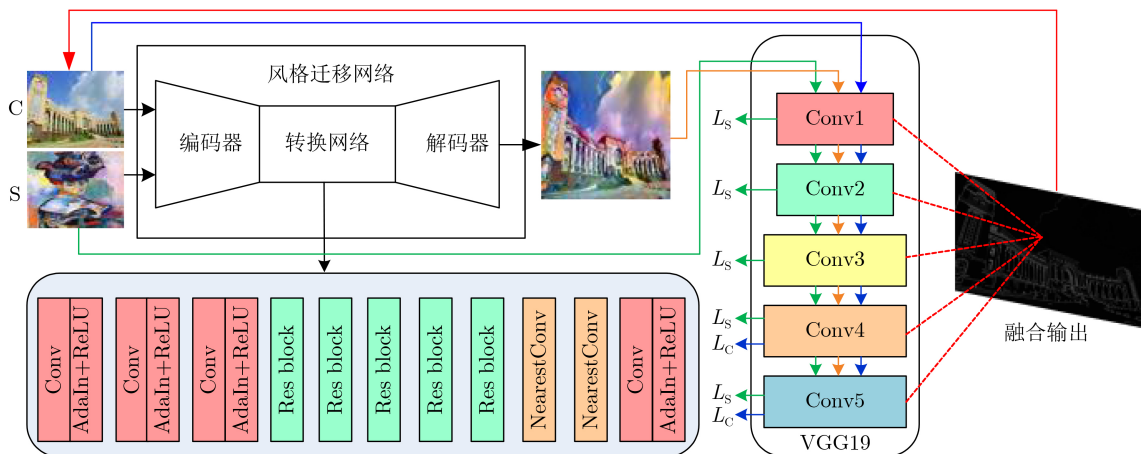


图1 风格迁移模型

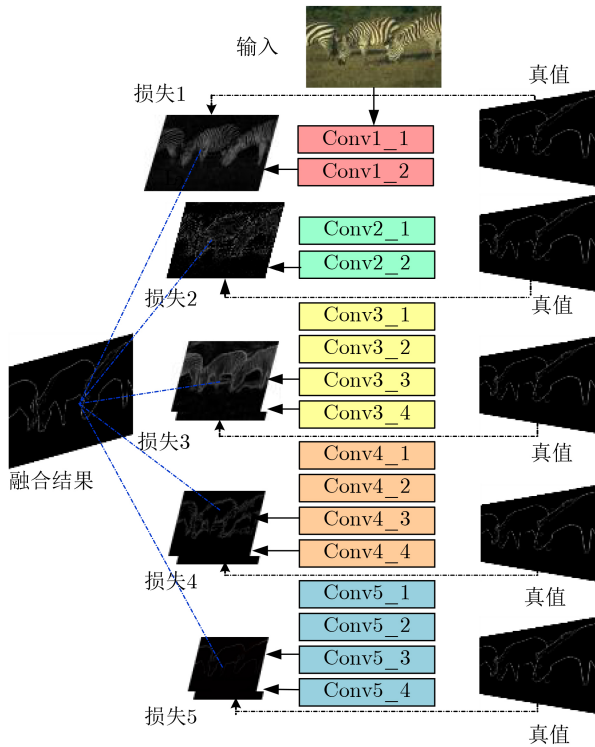


图 2 边缘提取过程示意图

表 2 在BSDS500数据集上的客观评价指标

指标	ODS	OIS	AP
5层融合边缘检测图	0.760	0.784	0.800
6层融合边缘检测图	0.774	0.797	0.798
7层融合边缘检测图	0.777	0.788	0.814
8层融合边缘检测图	0.786	0.802	0.822

来实现对输入图像的特征提取。转换网络主体由5个残差块组成，为了在特征通道中检测样式图像的笔触并产生较高的平均激活度，在非残差卷积层后增加AdaIN和ReLU层，从而实现纹理合成和色彩迁移，网络主体如图4所示。在转换网络中，本文在Conv1和Conv4中以两个5×5和一个1×1的卷积核代替9×9的卷积核，其他卷积层都使用3×3的卷积核。1×1的卷积核能够在不影响输入输出维数的情况下实现跨通道的信息交互整合，还可以进行通道数的升维和降维。两个5×5卷积核堆叠，则能够增加多层非线性组合，提高网络学习复杂内容的能力，并且使判决函数更具判决性，起到隐式正则化的作用。

表 1 步长和感受野参数设置

Layer	Conv1_2	Conv2_2	Conv3_3	Conv3_4	Conv4_3	Conv4_4	Conv5_3	Conv5_4
步长	1	2	4	4	8	8	16	16
接受域	5	14	40	44	92	100	196	212

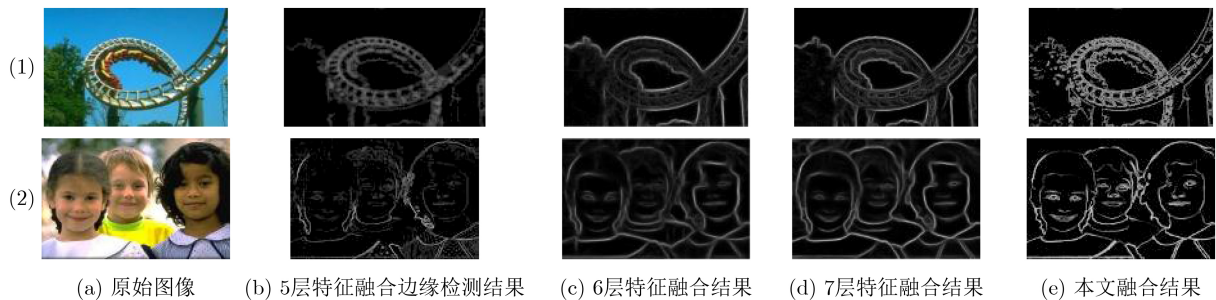


图 3 不同深度的边缘检测图

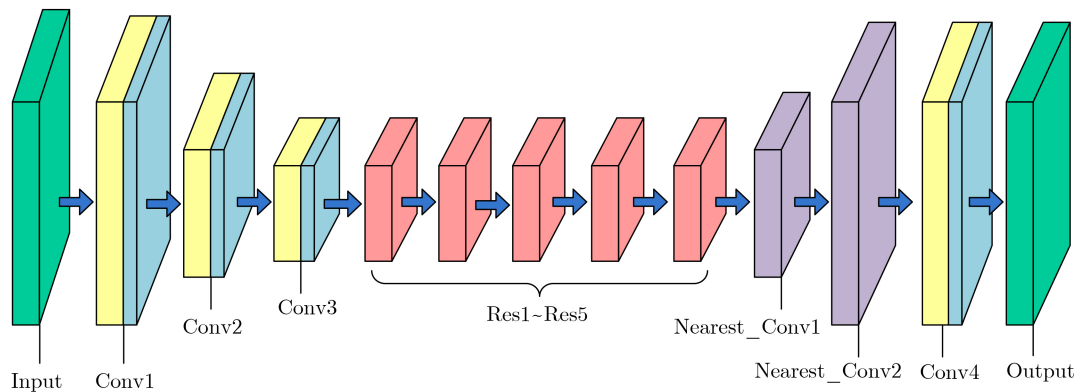


图 4 转换网络结构

网络训练的实质是学习各种参量的过程,其参数量由网络深度、卷积核尺寸以及通道数决定。在单个卷积层中,使用两个 $5 \times 5$ 卷积核的参数量为 $2 \times 5 \times 5 \times C \times C = 50 \times C^2$ ,使用 $9 \times 9$ 卷积核的参数量为 $9 \times 9 \times C \times C = 81 \times C^2$ ( $C$ 为通道数)。由此可见,随着卷积层增多与特征图通道数增加,使用2个 $5 \times 5$ 卷积核的参数量只有 $9 \times 9$ 的一半左右,前者有2个非线性操作,后者有1个非线性操作,其网络具有更强的特征学习能力。转换网络的整体参数量如表3所示,卷积核替换前后,转换网络整体的参数量减少了31.1%,当使用计算能力为7.5的GeForce RTX 2080 Ti进行训练时,改进后的网络比改进前节省了1.61%的计算资源。

大卷积核具有更大的感受野,生成较大的特征图,经过池化去除冗余信息的同时也会损失很多细节信息,使得生成图像损失了很多细节,如图5(c)所示。本文用小卷积核代替大卷积核,能够保证相同的感受野,产生更加准确的特征响应,并且能够增加网络深度,使得网络中的函数能够更好地逼近输入图像的特征。捕捉自然图像的统计属性,让目标轮廓更加清晰,细节纹理更精细,提升视觉效果(图5(d)所示)。

### 3 风格迁移误差

#### 3.1 损失函数

假设网络的第 $l$ 层有 $N^l$ 不同的滤波器,每个特征响应的大小为 $H^l \times W^l$ ,该层的响应为 $M^l \in \mathbb{R}^{(H^l \times W^l) \times N^l}$ , $H^l$ 和 $W^l$ 分别为特征映射的高度和宽度, $M_{(i,j),k}^l$ 为 $l$ 层 $(i,j)$ 出第 $k$ 个滤波器的激活函数值。

内容图像 $x$ 与风格图像 $y$ 和风格化图像 $\hat{y}$ 在第 $l$ 层的损失函数为

$$L_c^l(x, \hat{y}) = \frac{1}{H^l W^l N^l} \|M^l(x) - M^l(\hat{y})\|_2^2 \quad (1)$$

$$L_s^l(y, \hat{y}) = \|G^l(y) - G^l(\hat{y})\|_F^2 \quad (2)$$

其中 $G^l$ 是格莱姆矩阵,一个 $N^l \times N^l$ 的对称矩阵, $G_{i,j}^l$ 是第 $l$ 层的第 $i$ 个和第 $j$ 个矢量化特征图的归一化内积 $G_{i,j}^l = (1/H^l W^l N^l) F_{(h,w),i}^l(x) F_{(h,w),j}^l(x)$ , $F_{i,j}^l$ 表示第 $l$ 层第 $j$ 个位置的第 $i$ 个滤波器的激活值。

神经风格化产生的整体损失函数为

$$L_t = \alpha \sum_{l=1}^L L_c^l(x, \hat{y}) + \beta \sum_{l=1}^L L_s^l(y, \hat{y}) \quad (3)$$

其中, $\alpha, \beta$ 分别为内容损失和风格损失占的权重。



图5 不同卷积核风格迁移纹理对比

表3 迁移网络改进前后参数量对比

对应卷积层参数量	特征图通道数	步长	卷积核尺寸, 参数量	卷积核尺寸, 参数量
Conv1	32	1	<b>9×9, 159418368</b>	<b>2×5×5, 98406400</b>
Conv2	64	2	3×3, 8856576	3×3, 8856576
Conv3	128	2	3×3, 4428288	3×3, 4428288
Resblock1-Resblock5	128		23454552	23454552
Nearest_Conv1	64	1/2	3×3, 57600	3×3, 57600
Nearest_Conv2	32	1/2	3×3, 73728	3×3, 73728
Conv4	3	1	<b>9×9, 15552</b>	<b>2×5×5, 9600</b>
总参数量			<b>196.30×10<sup>6</sup></b>	<b>135.29×10<sup>6</sup></b>

### 3.2 改进的损失函数

本文在转换网络的常规卷积层后增加了AdaIN<sup>[5]</sup>, 修改特征匹配参数, AdaIN通过学习仿射参数, 自动地对内容图像和风格图像的特征统计信息进行匹配。图像通过编码器 $d$ 映射在特征空间中, 将同一层的内容特征映射的均值和方差与风格特征映射的均值和方差对齐, 生成目标特征映射 $h$

$$h = \text{AdaIN}(d(x), d(y)) = \sigma(d(y)) / \sigma(d(x)) (d(x) - \mu(d(x))) + |\mu(d(x)) - \mu(d(y))| \quad (4)$$

其中,  $\mu(d(x)), \sigma(d(x)) \in \mathbb{R}^C$  分别为各特征通道的批处理大小和空间维数的平均值和标准差,

$$\mu(x) = \frac{1}{HW} \sum_{h=1}^H \sum_{w=1}^W d(x),$$

$$\sigma(x) = \sqrt{\frac{1}{HW} \sum_{h=1}^H \sum_{w=1}^W (d(x) - \mu(d(x)))^2 + \varepsilon}.$$

解码器 $g$ 将生成的特征映射转换到图像空间, 经归一化操作后产生的内容损失和风格损失分别为

$$L_{C, \text{AdaIN}} = \|d(g(h)) - h\|_2 \quad (5)$$

$$L_{S, \text{AdaIN}} = \|\mu(\mathbf{G}(\hat{y})) - \mu(\mathbf{G}(y))\|_2 + \|\sigma(\mathbf{G}(\hat{y})) - \sigma(\mathbf{G}(y))\|_2 \quad (6)$$

则图像风格化产生的总体的内容感知损失函数和风格感知损失分别为

$$L_C(x, \hat{y}) = \sum_{l=1}^L \frac{1}{H^l W^l N^l} \|\mathbf{M}^l(x) - \mathbf{M}^l(\hat{y})\|_2^2 + \sum_{l=1}^l \|d(\hat{y}^l) - h^l\|_2 + L_{\text{edge}}(Q, q, p) \quad (7)$$

$$L_S(y, \hat{y}) = \sum_{l=1}^L \|\mathbf{G}^l(y) - \mathbf{G}^l(\hat{y})\|_F^2 + \sum_{l=1}^L (\|\mu(\mathbf{G}^l(\hat{y})) - \mu(\mathbf{G}^l(y))\|_2 + \|\sigma(\mathbf{G}^l(\hat{y})) - \sigma(\mathbf{G}^l(y))\|_2) \quad (8)$$

其中,  $L_{\text{edge}}(Q, q, p)$  为边缘检测<sup>[15]</sup>优化后的损失函数。网络学习通过随机梯度下降来优化损失函数

$$L = \alpha L_C + \beta L_S + \gamma L_R \quad (9)$$

其中,  $L_R$ 是正则化<sup>[4]</sup>,  $\gamma$ 表示正则化项的权重, 以增加生成图像的平滑性。

在风格迁移模型中, 每一层的参数更新导致上层的输入数据分布发生变化, 使用AdaIN, 把数据分布映射到一个确定的区间, 并在整体损失函数中增加了边缘检测损失和归一化损失, 与经典算法Gatys<sup>[4]</sup>进行对比, 网络性能更好, 准确率更高, 收敛速度更快(图6)。特征图各通道的均值和方差会影响最终生成图像的风格, 通过式(4)实现风格图像和内容图像均值和方差的匹配, 并在编解码时分别进行归一化与去归一化获得目标图像的风格, 最终实现风格迁移。通过式(9)中内容损失和风格损失的权重比来控制样式转移的程度。图7中 $\eta$ 表示内容损失和风格损失的权重比, 可以看出,  $\eta=0.5$ 时, 两组图像风格化不完全, 语义扭曲。 $\eta=1$ 时, 风格化程度较大, 语义内容和风格纹理很好地结合。归一化前纹理笔触较大, 细节太平滑, 前后景边缘轮廓模糊; 归一化后生成图像的笔触较小, 纹理精细, 边缘轮廓得以增强, 使得风格化后的图像整体视觉效果更有层次感。

## 4 实验与结果分析

### 4.1 实验设置

本文使用MS-COCO<sup>[16]</sup>作为内容图像, Wiki-Art<sup>[17]</sup>的绘画数据集作为风格图像来训练网络。在训练过程中, 利用VGG-19网络中Relu4\_2和Relu5\_2层的激活计算 $L_{\text{content}}$ , Relu1\_2, Relu2\_2, Relu3\_4, Relu4\_4和Relu5\_4层计算 $L_{\text{style}}$ , 在生成网络的常规卷积后计算 $L_{C, \text{AdaIN}}$ 和 $L_{S, \text{AdaIN}}$ 。式(9)中默认参数 $\alpha, \beta, \gamma$ 分别为 $1 \times 10^{-2}, 1 \times 10^{-4}, 1 \times 10^{-3}$ 。本文使用Adam优化器进行优化, 内容样式图像对批处理大小为16, 学习率为 $1 \times 10^{-4}$ , 迭代次数为40000。实验设备为搭载Win10系统的台式机, CPU为Intel I9 9900K 5.0GHz, RAM为64G, 实验环境为PyCharm+TensorFlow, 使用NVIDIA RTX 2080Ti进行GPU加速。

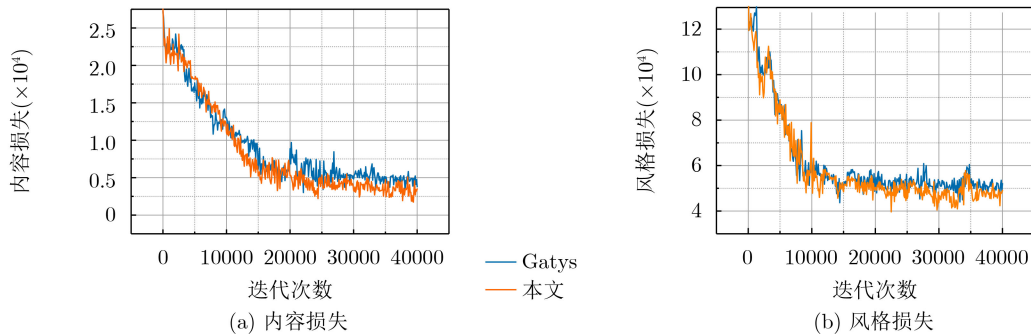


图6 损失函数对比图

## 4.2 结果分析

### 4.2.1 主观评价分析

本文改进的算法在检测的边缘图的引导与约束下以及经AdaIN处理, 迁移效果如图8所示, 图8(a)为风格图像, 图8(b)为内容图像, 图8(c)为边缘检测图, 图8(d)为迁移效果图。将本文算法与文献[4], 文献[6], 文献[10], 文献[18]的算法对比, 风格迁移效果如图9, 图10所示。在图9(c)中部分迁移结果风格均匀分布(图9(c)(1)), 覆盖了语义内容, 结构复杂的输入迁移结果出现原结构轻微破坏, 细节信息模糊(图9(c)(2))。图9(d)普遍存在细小的颗粒覆盖在迁移结果上, 弱化了迁移图像的某些细节, 也使得主要目标变得模糊(图9(d)(2))。图10(c)语义信息覆盖较为严重, 主要目标无法识别, 前后景对

比度低。图10(d)存在少量的多余纹理分布(图10(d)(4)和(6)), 部分迁移结果出现白晕(图10(d)(5)), 弱化了内容的细节。图9(e)和图10(e)为本文算法的迁移结果, 可以看出, 不论是风格化图像的结构还是语义信息保留程度都较出色, 迁移结果无纹理分布覆盖语义信息的现象, 并且前背景的边界也很清晰, 主要目标清晰可辨, 风格化图像结构保持较好, 结构细节损失较小, 细节信息得以表达, 没有光斑。因此, 从视觉效果上看, 本文的迁移方法产生的结果更加出色。

### 4.2.2 主观评价分析

在客观评价中, 本文从峰值信噪比、内容和风格的平均结构相似(MSSIM)进行比较。PSNR由图像信号峰值与均方误差决定, 表示图像风格迁移质

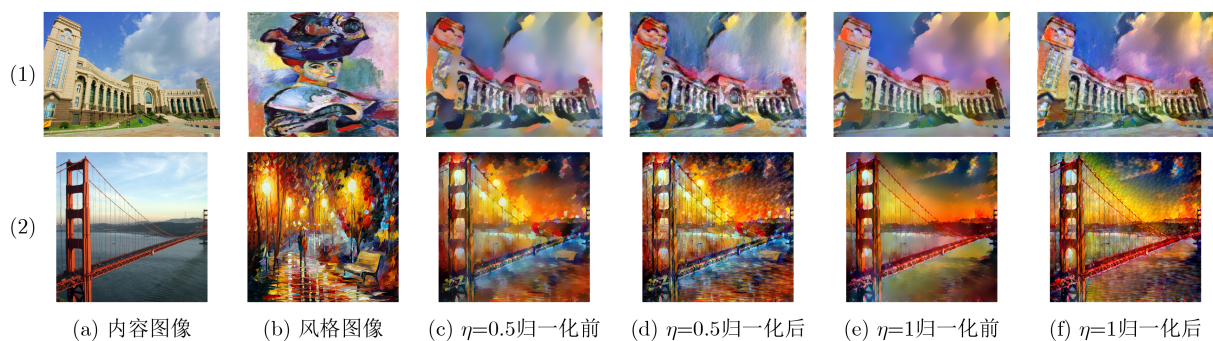


图7 纹理比较

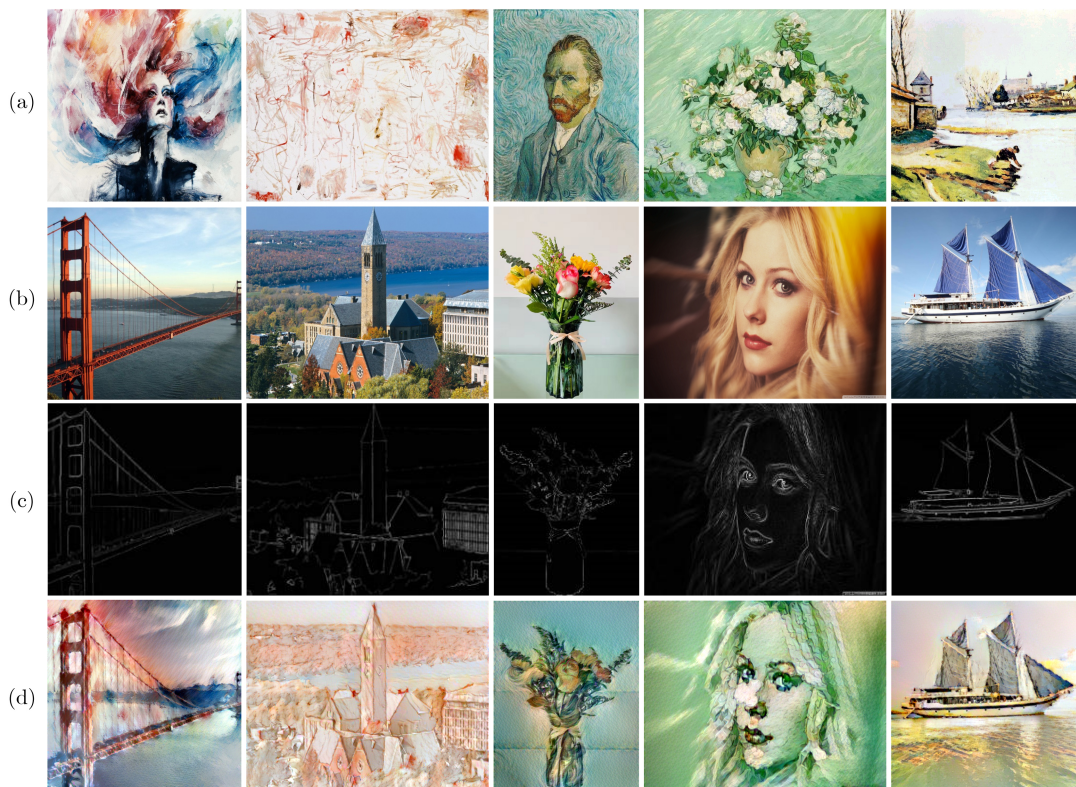


图8 本文算法迁移效果展示

量的好坏；SSIM将图像的亮度、对比度和结构3个因素进行组合。以均值估计亮度，标准差估计对比度，协方差估计结构相似程度。客观数据对比如图11所示。

由图11可知，本文方法在峰值信噪比与平均结构相似度都有较好的表现，说明本文算法风格化图像质量较好，内容结构保留程度高，风格化图像迁移了更多的纹理信息。

#### 4.2.3 运行时间比较

本文将本文算法的运行时间与表4中的文献[4]，文献[6]，文献[10]，文献[18]的不同尺寸的图像进行了对比。总体上看，本文通过增加自适应实例归一化层在一定程度上提升了风格迁移的迁移效率。

## 5 结论

本文将基于神经网络的边缘检测网络添加到神经风格迁移模型上，通过对边缘检测网络的多层输出进行融合获得内容图像的边缘轮廓图，并将其输入风格转换网络引导约束风格迁移；在迁移网络中，本文将常规卷积层的大卷积核进行了替换，减少了网络模型的参数，提升了运行的速度；在常规卷层后添加了AdaIN层，修改了特征匹配参数，增强了风格化结果的轮廓，一定程度上保留了图像的结构布局。本文算法解决了纹理均匀分布破坏结果图像的结构布局，造成细节信息丢失和前后景边界模糊的问题。通过客观数据表明，本文算法在能更好地保留内容结构，细化语义信息，结合风格纹理。



图 9 实验结果对比

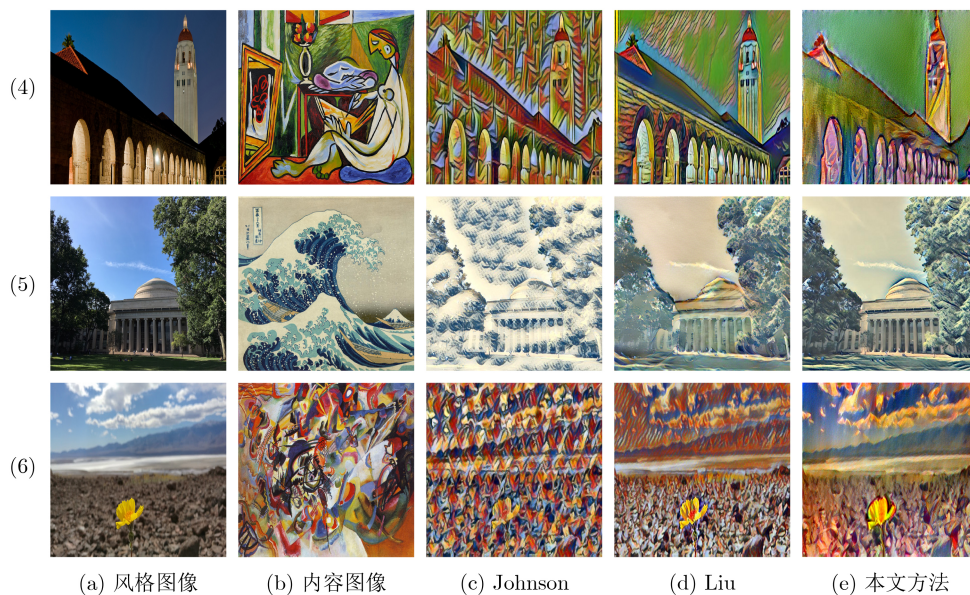


图 10 实验结果对比

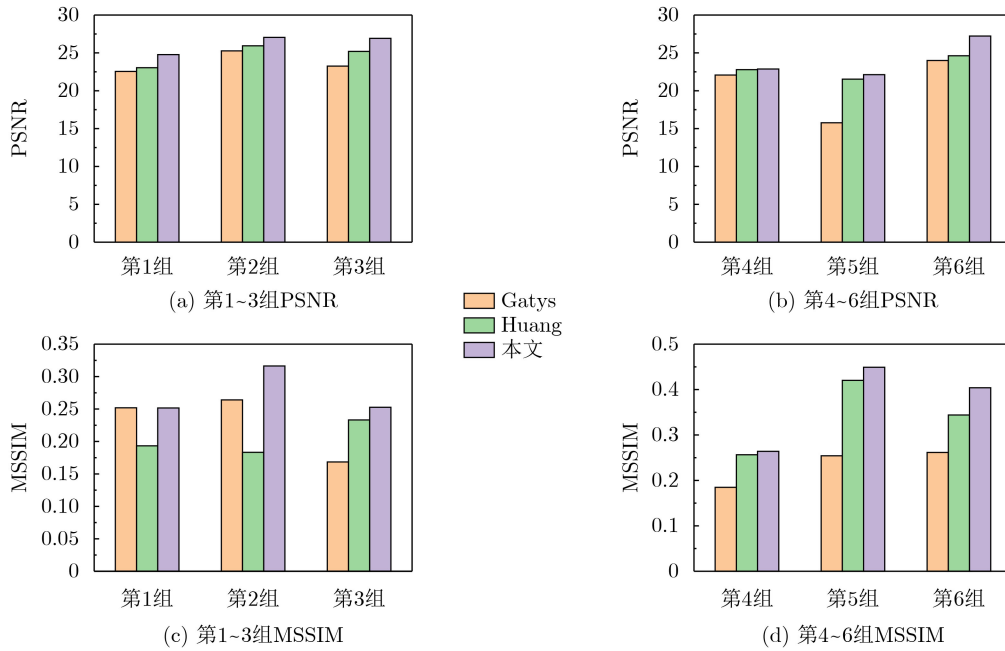


图 11 客观评价指标

表 4 风格迁移算法运行时间比较(s)

方法	Gatys <sup>[4]</sup>	Huang <sup>[6]</sup>	Johnson <sup>[10]</sup>	Liu <sup>[18]</sup>	本文	
图像尺寸	256×256	15.86	0.018	0.015	0.083	0.013
	512×512	54.85	0.065	0.05	0.141	0.038
	1024×1024	214.44	0.275	0.21	0.37	0.255

## 参考文献

- [1] KYPRIANIDIS J E, COLLOMOSSE J, WANG Tinghui, et al. State of the “Art”: a taxonomy of artistic stylization techniques for images and video[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2013, 19(5): 866–885. doi: [10.1109/TVCG.2012.160](https://doi.org/10.1109/TVCG.2012.160).
- [2] 袁野, 贾克斌, 刘鹏宇. 基于深度卷积神经网络的多元医学信号多级上下文自编码器[J]. *电子与信息学报*, 2020, 42(2): 371–378. doi: [10.11999/JEIT190135](https://doi.org/10.11999/JEIT190135).  
YUAN Ye, JIA Kebin, and LIU Pengyu. Multi-context autoencoders for multivariate medical signals based on deep convolutional neural networks[J]. *Journal of Electronics & Information Technology*, 2020, 42(2): 371–378. doi: [10.11999/JEIT190135](https://doi.org/10.11999/JEIT190135).
- [3] GATYS L A, ECKER A S, and BETHGE M. A neural algorithm of artistic style[J]. arXiv preprint arXiv: 1508.06576, 2015.
- [4] GATYS L A, ECKER A S, and BETHGE M. Image style transfer using convolutional neural networks[C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, USA, 2016: 2414–2423. doi: [10.1109/cvpr.2016.265](https://doi.org/10.1109/cvpr.2016.265).
- [5] LUAN Fujun, PARIS S, SHECHTMAN E, et al. Deep photo style transfer[C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, USA, 2017: 6997–7005. doi: [10.1109/CVPR.2017.740](https://doi.org/10.1109/CVPR.2017.740).
- [6] HUANG Xun and BELONGIE S. Arbitrary style transfer in real-time with adaptive instance normalization[C]. 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 2017: 1510–1519. doi: [10.1109/iccv.2017.167](https://doi.org/10.1109/iccv.2017.167).
- [7] LI Chuan and WAND M. Combining Markov random fields and convolutional neural networks for image synthesis[C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, USA, 2016: 2479–2486. doi: [10.1109/CVPR.2016.272](https://doi.org/10.1109/CVPR.2016.272).
- [8] DUMOULIN V, SHLENS J, and KUDLUR M. A learned representation for artistic style[J]. arXiv preprint arXiv: 1610.07629, 2016.
- [9] CHEN Yang, LAI Yukun, and LIU Yongjing. CartoonGAN: generative adversarial networks for photo cartoonization[C]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, USA, 2018: 9465–9474. doi: [10.1109/CVPR.2018.00986](https://doi.org/10.1109/CVPR.2018.00986).
- [10] JOHNSON J, ALEXANDRE A, and LI Feifei. Perceptual losses for real-time style transfer and super-resolution[C]. The 14th European Conference on Computer Vision, Amsterdam, Netherlands, 2016: 694–711. doi: [10.1007/978-](https://doi.org/10.1007/978-)

- 3-319-46475-6\_43.
- [11] 王鑫, 李可, 宁晨, 等. 基于深度卷积神经网络和多核学习的遥感图像分类方法[J]. 电子与信息学报, 2019, 41(5): 1098–1105. doi: [10.11999/JEIT180628](https://doi.org/10.11999/JEIT180628).  
WANG Xin, LI Ke, NING Chen, *et al.* Remote sensing image classification method based on deep convolution neural network and multi-kernel learning[J]. *Journal of Electronics & Information Technology*, 2019, 41(5): 1098–1105. doi: [10.11999/JEIT180628](https://doi.org/10.11999/JEIT180628).
- [12] CHEN Chunfu, FAN Quanfu, MALLINAR N, *et al.* Big-little net: an efficient multi-scale feature representation for visual and speech recognition[J]. arXiv preprint arXiv: 1807.03848, 2018.
- [13] WANG Xin, YU F, DOU Ziyi, *et al.* SkipNet: Learning dynamic routing in convolutional networks[C]. The 15th European Conference on Computer Vision, Munich, Germany, 2018: 420–436. doi: [10.1007/978-3-030-01261-8\\_25](https://doi.org/10.1007/978-3-030-01261-8_25).
- [14] LIU Yun, CHENG Mingming, HU Xiaowei, *et al.* Richer convolutional features for edge detection[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019, 41(8): 1939–1946. doi: [10.1109/TPAMI.2018.2878849](https://doi.org/10.1109/TPAMI.2018.2878849).
- [15] XIE Saining and TU Zhuowen. Holistically-nested edge detection[J]. *International Journal of Computer Vision*, 2017, 125(1): 3–18. doi: [10.1007/s11263-017-1004-z](https://doi.org/10.1007/s11263-017-1004-z).
- [16] LIN T Y, MAIRE M, BELONGIE S, *et al.* Microsoft COCO: common objects in context[C]. The 13th European Conference on Computer Vision, Zurich, Switzerland, 2014: 740–755. doi: [10.1007/978-3-319-10602-1\\_48](https://doi.org/10.1007/978-3-319-10602-1_48).
- [17] SAIF M M and SVETLANA K. WikiArt emotions: an annotated dataset of emotions evoked by art[C]. The 11th International Conference on Language Resources and Evaluation, Miyazaki, Japan, 2018: 1225–1238.
- [18] LIU Xiaochang, CHENG Mingming, LAI Yukun, *et al.* Depth-aware neural style transfer[C]. The Symposium on Non-Photorealistic Animation and Rendering, California, USA, 2017: 4. doi: [10.1145/3092919.3092924](https://doi.org/10.1145/3092919.3092924).
- 沈 瑜: 女, 1982年生, 教授, 硕士生导师, 研究方向为深度学习、神经网络、图像处理.
- 杨 倩: 女, 1995年生, 硕士, 研究方向为神经网络、风格迁移.

责任编辑: 陈 倩