RAUGAN：基于循环生成对抗网络的红外图像彩色化方法

朴 燕，康继元

（长春理工大学 电子信息工程学院，长春 130022）

摘 要：针对近红外图像彩色化过程中的色彩失真、语义模糊和纹理形状不清晰的问题，提出了一种红外图像彩色化方法（RAUGAN）。该算法首先改进了CycleGAN网络的生成器，设计并融合了一种Res-ASPP-Unet网络，将空洞空间金字塔池化（ASPP）在原始UNet的Skip connection结构处连接，使解码分支中的不同尺度输出特征图都能与编码器中对应的输出特征图相结合；其次，设计了由残差块与通道和空间注意力模块（CBAM）构成的深度瓶颈层块替换UNet网络中的瓶颈层，用于增强局部区域特征，提高其识别能力；最后，在判别网络中引用感知损失函数从而解决色彩恢复失真的问题。实验结果表明：该方法彩色化效果明显优于其他方法。

关键词：计算机应用；红外图像彩色化；循环生成对抗网络；空洞空间金字塔池化；注意力模块

中图分类号：TP391.41 **文献标志码**：A **文章编号**：1671-5497(2025)08-2722-10

DOI：10.13229/j.cnki.jdxbgxb.20240012

RAUGAN：infrared image colorization method based on cycle generative adversarial networks

PIAO Yan，KANG Ji-yuan

（*School of Electronic and Information Engineering*，*Changchun University of Science and Technology*，*Changchun* 130022，*China*）

Abstract：In order to solve the problems of color distortion，semantic ambiguity，unclear texture and shape characteristics in the process of near-infrared image colorization，we propose a method of infrared image colorization（RAUGAN）. Firstly，the CycleGAN network generator is improved，and a Res-ASPP-UNet network is designed and fused，which connects atrous spatial pyramid pooling（ASPP）to Skip connection structure of the original UNet，the output characteristic graphs of different scales in the decoding branch can be combined with the corresponding output characteristic graphs in the encoder. Secondly，a deep bottleneck layer composed of residual block，convolutional block attention module（CBAM）is designed to replace the bottleneck layer in UNET network to enhance the local area feature，improve its recognition ability and prevent gradient explosion. Finally，the perceptual loss function is introduced in the discriminant network to solve the problem of color restoration distortion.Experimental results show that the

收稿日期：2024-01-04.

基金项目：吉林省自然科学基金项目（20210101180JC）.

作者简介：朴燕（1964-），女，教授，博士. 研究方向：三维成像及机器视觉. E-mail：piaoyan@cust.edu.cn

proposed method is superior to the original CycleGAN network.

Key words: computer application; infrared image colorization; cycle-consistent generative adversarial networks; atrous spatial pyramid pooling; attention mechanism

0　引　言

红外图像成像的主要特点是可以在无光照或低光照条件下工作,因为它依赖的是物体的热辐射,而不是光照。因此,它无法提供颜色信息获取清晰图像,尤其在低照度环境下的红外图像亮暗分布不均匀,信息捕捉不够全面,所以红外图像彩色化的结果存在严重的细节模糊、信息丢失和场景过渡不自然问题,这些客观存在的问题增加了实现红外图像彩色化的难度。而近红外与可见光图像间模态差异较大导致其彩色化结果出现颜色失真、细节边缘模糊的问题则是该领域的研究难题。

目前,近红外图像彩色化方法主要分为传统的颜色映射彩色化和基于深度学习的彩色化这两类。传统的颜色映射彩色化主要包括基于颜色传递的彩色化方法和基于用户引导的彩色化方法两种。针对第一种方法,Reinhard 等[1]根据颜色空间特点提出颜色传递算法,在颜色校正方法上借用另一个图像的颜色特征,以不同的颜色应用不同的通道进而实现彩色化;Welsh 等[2]提出用改进的 Retinex 算法增强灰度融合图像的颜色传递;针对第二种方法,Levin 等[3]提出了一种交互式彩色技术,用户通过相邻像素灰度值相似且色彩相近的特点进行上色。上述方法均具有局限性、计算复杂且需要人工标注和对应的参考图像才能对灰度图像上色;实时处理的彩色化方法虽然处理速度很快,但着色效果并不自然。

基于深度学习的彩色化具有极强的泛化性,可以减少人工操作和干预,能自动学习两个域各自的特征信息,将大批红外图像投入神经网络中训练学习,通过提取图像特征实现红外图像的彩色化,改善视觉效果并突出细节内容。目前,大多数基于深度学习的红外图像彩色化算法是通过生成式对抗网络(Generative adversarial networks, GAN)[4]架构模型来实现的,一般的可以将其分为有监督[5-7]和无监督[8-10]两种类型。Suárez 等[11]利用 GAN 架构模型实现近红外图像的彩色化;王晓宇[12]提出了利用 CycleGAN 的对偶思想实现非配对图像间的颜色迁移,并在损失函数中引入边缘损失,抑制了生成图像中的细节扭曲;Li 等[13]在 CycleGAN 的基础上提出实时的近红外视频 I2VGAN 彩色化网络,在视频结果中部分帧图像出现细节丢失问题;高美玲等[14]在 CycleGAN 的基础上设计了一种空洞级联模块,在编解码级联层中引入空洞卷积模块进一步提取不同尺度的近红外图像特征信息。

上述方法在图像去噪[15]、图像生成[16,17]以及图像超分辨率[18,19]等视觉任务中被广泛应用。虽然 GAN 网络能自动学习红外图像和彩色图像之间的复杂映射关系并在一定程度上改善了传统方法对于红外图像彩色化的效果,但在彩色化的过程中,卷积神经网络在逐层卷积时会丢失原有的像素信息,从而导致生成图像特征信息丢失、色彩不均衡的问题。

为进一步改善红外图像彩色化色彩失真、细节纹理模糊的问题,本文设计了一种新的 RAUGAN 模型,该模型的主体是在 CycleGAN 网络的基础上进行改进,设计并融合了一种 Res-ASPP-Unet 网络在生成器中,网络框架使用 UNet 网络,在其跳跃连接处融合 ASPP[20]模块,这样可以在不增加计算复杂度和参数的情况下通过使用不同大小采样率的空洞卷积来捕获多个尺度的图像语义信息,并可以解决 Unet 编码阶段的最大池化操作导致像素值丢失的问题。为防止网络反向传播时发生梯度爆炸,决定由残差块构成的残差网络 Resnet 替换 UNet 网络中的瓶颈层即 U 型最低端的全连接部分并融合注意力模块 CBAM[21],以加强网络对图像中关键特征信息的学习,减少对不重要特征信息的学习。本文方法使近红外图像彩色化后保留了原始图像更多的细节纹理信息,色彩更加温和,重要物体特征信息更加明显,解决了图像细节信息捕捉不够全面的问题,符合人眼视觉需求。

1　RAUGAN模型

1.1　网络模型结构

RAUGAN 模型的基础网络是循环生成对抗

网络,该网络由两个生成器和两个判别器构成,其中使用Unet网络作为生成器整体框架,判别器使用马尔可夫判别器[22](Markovian discriminator, PatchGAN)。网络的整体框架如图1所示,X和Y分别代表近红外图像域和可见光图像域的图像。生成器$G_{X\to Y}$将X域的图像转换为Y域的图像,然后生成器$F_{Y\to X}$将其恢复为X域的原始输入图像;生成器$F_{Y\to X}$将Y域的图像转换为X域图像,然后生成器$G_{X\to Y}$再将其重构为Y域的原始输入图像。判别器D_X和D_Y用于鉴别生成图像是否为假的彩色或近红外图像,以确保图像的风格转换。

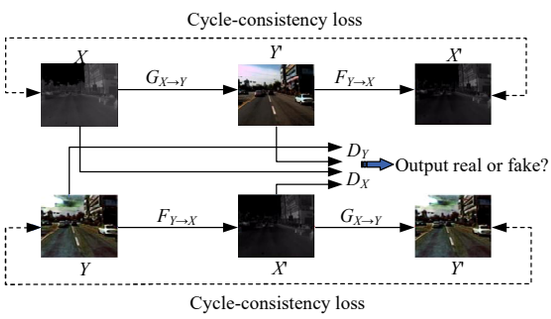


图1 网络整体框架

Fig. 1 Overall framework of the network

1.2 生成器

本文设计了一种Res-ASPP-Unet网络对生成器进行改进,融合Unet网络并在其上采样和下采样的连接处引入ASPP模块使网络获得较大的感受野,不让分辨率损失太多,从而解决红外图像彩色化后出现的边界模糊、原图像细节信息丢失严重的问题。同时设计了由残差块和CBAM模块组成的深度瓶颈层块替换UNet网络中的瓶颈层最低端的位置,不仅可以防止反向传播时发生梯度爆炸,还可以对图像中的重点部位提高训练强度从而加深语义特征理解,改进后的生成器网络框架如图2所示。

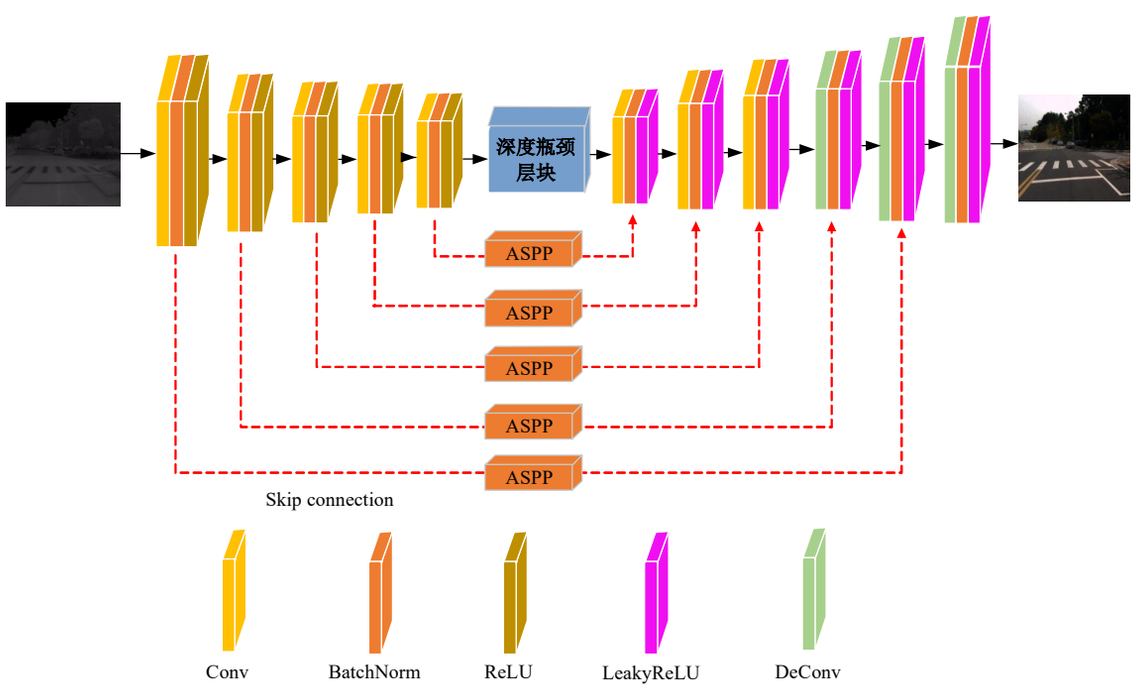


图2 改进后的生成器网络框架

Fig. 2 Our Improved generator network framework

生成器整体结构由卷积块、深度瓶颈层块、残差连接、反卷积块和输出块组成。其中,下采样模块由卷积块一至卷积块五和深度瓶颈层块组成,用于从输入的近红外图像中提取特征,其中,每个卷积块将特征图减小到一半,深度瓶颈层块提取并保留每个卷积块后的信息;深度瓶颈层块由残差块和CBAM模块组成,防止反向传播时梯度消失并加深彩色化后图像的语义特征理解;跳跃连接处融合ASPP模块,用于将下采样模块的各尺度特征细节直接映射到上采样模块;上采样模块由反卷积块一至反卷积块六和输出块组成,用于图像重建,反卷积块对编码器的信息进行解码进而重构输出图像,输出块输出重建的彩色图像。生成器网络配置如表1所示。

1.2.1 空间空洞金字塔池化ASPP模块

空间空洞金字塔池化(Atrous spatial pyramid

表 1　生成器网络配置

Table 1　Generator network configuration

network layer	Output	Convolution(Conv)、stride(s)、Padding(p)
Input	256×256	
ConvBlock 1	256×256	{3×3Conv, BatchNorm, ReLU}, $s=1, p=1$
DBlock 1	64×64	{3×3Conv, BatchNorm, LeakyReLU}, $s=2, p=1$
DBlock 2	125×128	{3×3Conv, BatchNorm, LeakyReLU}, $s=2, p=1$
DBlock 3	256×256	{3×3Conv, BatchNorm, LeakyReLU}, $s=2, p=1$
Output	256×256	{3×3Conv, BatchNorm, ReLU}, $s=2, p=2$

pooling, ASPP),其主要操作流程就是使用不同大小采样率的空洞卷积对同一幅顶端特征图并行采样,再将采样后的各结果合并到一起,从而扩大通道数,然后通过一个卷积层将通道数降至所需数值,从而扩大网络的感受野,增强模型对图像中不同尺度目标的感知能力,缓解局部复杂特征类型相互干扰的问题。$H_{1,1}(x)$、$H_{6,3}(x)$、$H_{12,3}(x)$、$H_{18,3}(x)$分别表示x对做采样率为1、6、12、18和卷积核大小为1×1和3×3的空洞卷积,image(x)为对主干网络输出的特征图进行全局平均池化操作后获取的图像级特征。其计算过程表示为:

$$Y=\mathrm{Concat}(\mathrm{Image}(x),H_{1,1}(x),H_{6,3}(x),H_{12,3}(x),H_{18,3}(x)) \tag{1}$$

ASPP模块网络结构如图3所示,在生成器使用的ASPP模块中,包含了aspp 1、aspp 2、aspp 3、aspp 4四个分支和一个平均池化层(avg_pool)。每个分支都包含了一个空洞卷积层(atrous_conv),用于扩大感受野,使用不同的采样率进行空洞卷积,卷积层具有512个输入通道和512个输出通道,使用了一个4×4的卷积核,步长为2,填充为1。

并且,每个分支后面都跟着一个批归一化层(BatchNorm2d)和ReLU激活函数,ReLU激活函数表示为:

$$\mathrm{ReLU}(x)=\max(x,0)=\begin{cases}0, x<0\\ x, x>0\end{cases} \tag{2}$$

1.2.2　深度瓶颈层块

ASPP-Unet架构高度依赖于更多的卷积操作,容易使准确率与增加的卷积操作数量成反比,从而导致原始输入的细节信息会在传递过程中丢失,同时也会造成梯度爆炸。该问题可以通过在堆叠层之间插入残差块的残差方法来解决。因此,本文设计了一个深度瓶颈层块,包括残差块和CBAM模块,网络结构如图4所示,由此替换UNet网络中的瓶颈层最低端的位置。残差块可以帮助模型学习到更深层次的特征,而注意力模块可以在这些特征上进行选择和加权,这样可以进一步增强模型的表达能力,使得模型在学习到的特征基础上,进一步选择和关注重要的特征。残差块网络结构如图5所示。其中,注意力机制采用通道和空间注意力模块(Convolutional block attention module, CBAM)。该模块由通道注意力和

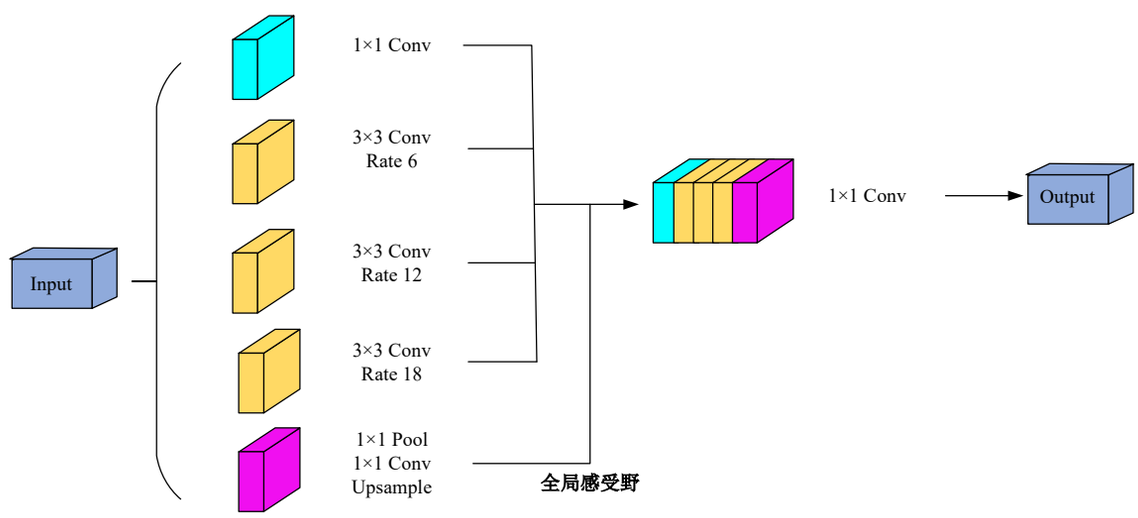


图3　ASPP模块网络结构

Fig. 3　Network architecture of ASPP model

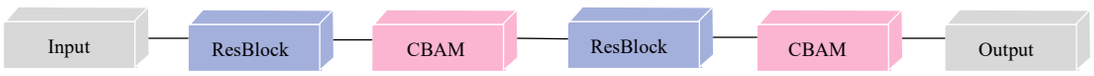


图4 深度瓶颈层块网络结构

Fig. 4 Network architecture of deep bottleneck layer block

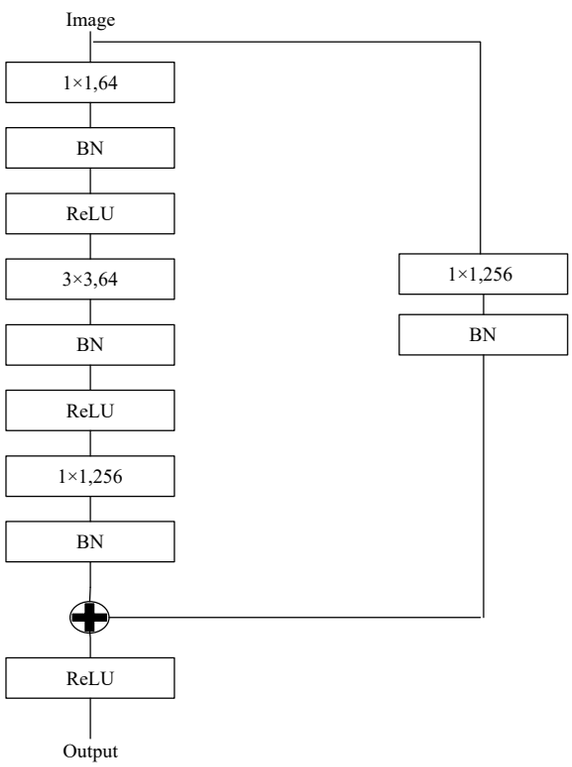


图5 残差块网络结构

Fig. 5 Residual block network structure

空间注意力两个子模块构成,网络结构图如图6所示。通道注意力可以让网络更多地关注有意义的特征,空间注意力可以让网络找到需要关注的重要区域。因此,研究在深度瓶颈层块中引入CBAM模块以提高网络的表达能力,从而对图像中的前景物体进行高度关注,同时降低对天空等背景的注意程度,解决红外图像彩色化过程中图像前景与背景模糊一片,重点物体没有突出的问题。

CBAM模块操作过程分为两步,首先是计算

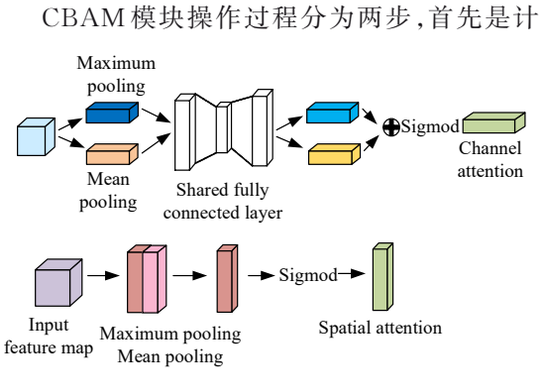


图6 CBAM模块网络结构

Fig. 6 Structure of the CBAM module

通道注意力Mc,将输入特征图F分别进行一个空间的全局平均池化和最大池化,然后通过对特征图的通道维度进行加权,以提取重要的特征信息。第二步是计算空间注意力Ms,它与通道注意力Mc相似。空间注意力对特征图的空间维度进行加权,以提高特征图中不同位置的特征关注度。最终,将两个模块的加权结果相乘,得到最终的特征表示F''。整个过程可表述为:

$$\begin{aligned}\mathrm{Mc}(F)&=\sigma(\mathrm{MLP}(\mathrm{AvgPool}(F))+\\&\mathrm{MLP}(\mathrm{MaxPool}(F)))=\\&\sigma(W_1(W_0(F^{\mathrm{c}}_{\mathrm{avg}}))+(W_1(W_0(F^{\mathrm{c}}_{\mathrm{max}})))\end{aligned}\tag{3}$$

$$\begin{aligned}Ms(F')&=\sigma(f^{7\times7}([\mathrm{AvgPool}(F'),\\&\mathrm{MaxPool}(F')])=\sigma(f^{7\times7}([F^{\mathrm{s}}_{\mathrm{avg}};F^{\mathrm{s}}_{\mathrm{max}}]))\end{aligned}\tag{4}$$

$$F''=\mathrm{Ms}(F')\otimes F'\tag{5}$$

1.3 判别器

判别器采用马尔可夫鉴别器,通过将输入图像进行逐层卷积后映射为$N\times N$的矩阵块形式来判别彩色图像的真假。在判别器中融合注意力机制,用于增强局部区域特征,提高其识别能力。该网络由5个卷积层构成,前4个卷积层用来提取生成图像的特征,最后一个卷积层产生输出图像用于判别图像的真假,并在最后一个卷积层之前添加了CBAM模块,这样可以使提取出的图像特征首先经过通道注意力进行加权突出重要特征,再经过空间注意力对突出的重要特征再进行筛选,从而使得判别器专注于重要区域的特征而忽略不必要的特征。判别器模型如图7所示。第一个和第五个卷积层去掉了批归一化层,卷积核(Conv)大小为3×3,步长(Stride)为1,激活函数采用

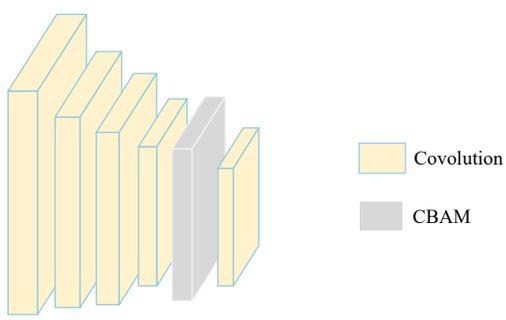


图7 判别器模型

Fig. 7 Discriminator model

Leaky ReLU 函数,表示为:

$$f(x)=\max(0.01x,x) \tag{6}$$

1.4　损失函数

损失函数采用对抗损失、循环一致性损失并引入感知损失函数。使得在白天情况下彩色化后的图像色度和亮度方面都接近真实的可见光图像,并进行了有效的内容转换和风格保存。

次级生成器需要在一个周期内尽可能地恢复所需图像,以保持输入图像和重建输出图像之间的一致性,为此,使用循环一致性损失,其计算公式定义为:

$$L_{\text{cyc}}=\|x-G_2(G_1(x))\|_1 \tag{7}$$

式中:x 为输入的红外图像;$G_2(G_1(x))$ 为先经过生成器一再经过生成器二生成的假红外图像。

为了鼓励网络输出具有更真实细节的彩色化结果,采用了对抗损失。对抗损失用于使彩色化结果与真实可见光图像无法区分,定义对抗损失的计算公式为:

$$L_{\text{adv}}=E_x[-\log(D(G(x),x))] \tag{8}$$

式中:x 不仅是生成器的输入,也是鉴别器的输入,作为条件项。

使用对抗损失容易产生失真纹理,为了缓解这个问题,使用基于 VGG-19 网络 ReLU 函数的感知损失,感知损失有利于在高层空间中保持纹理,定义感知损失的计算公式为:

$$L_{\text{perceptual}}=\sum_k\frac{1}{C_kH_kW_k}\sum_{i=1}^{H_k}\sum_{j=1}^{W_k}\|\Phi_k(y)_{ij}-\Phi_k(G(x))_{ij}\| \tag{9}$$

式中:Φ_k 为 VGG-19 网络中第 k 个最大池化层的特征;$C_kH_kW_k$ 为这些特征的大小。

因此,模型总损失为:

$$L_{\text{total}}=\lambda_1L_{\text{cyc}}+\lambda_2L_{\text{adv}}+\lambda_3L_{\text{perceptual}} \tag{10}$$

式中:λ_1、λ_2 和 λ_3 分别为在整个总损失函数中控制不同损失份额的权重,通过实验调试将 λ_1 设置为 1,λ_2 设置为 10,λ_3 设置为 5,实验效果最佳。

2　实　验

2.1　实验条件

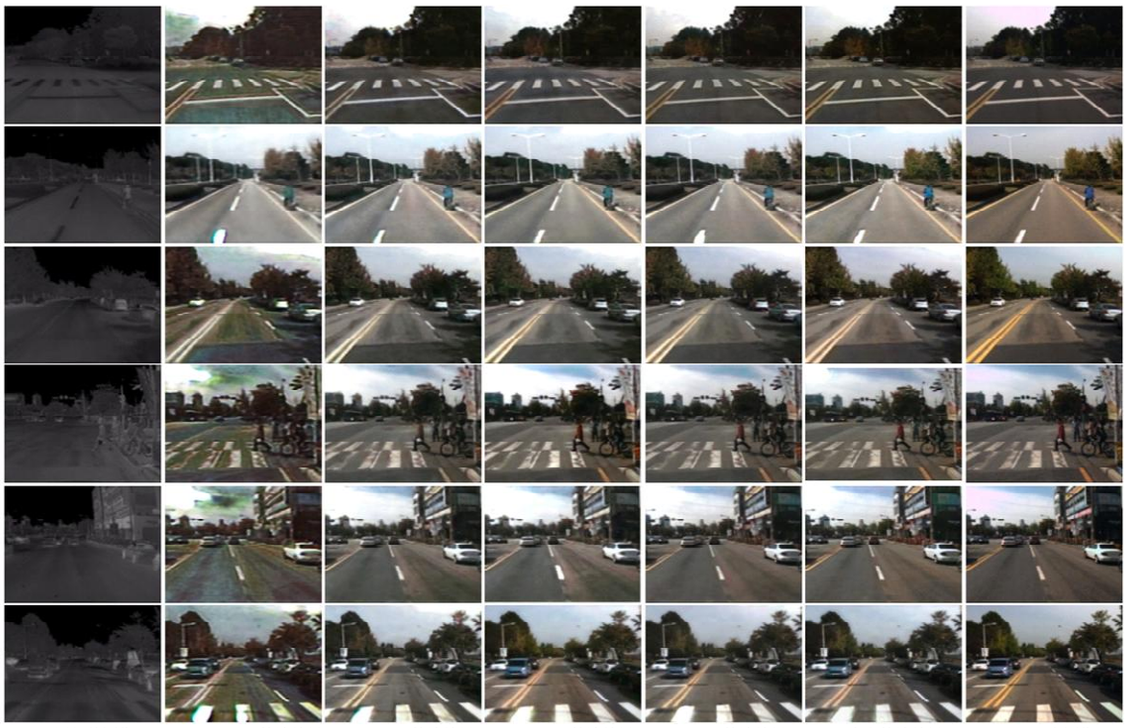
本文在 Windows 10 系统上进行实验,编程语言采用 Python,深度学习框架采用 Pytorch1.10.1。实验数据集使用公开的 KAIST[23] 数据集和 IRVI[24] 数据集,KAIST 数据集包含校园、道路和市中心 3 类场景,其中包含 33 399 对训练图像和 29 178 对验证图像,每幅图像的分辨率为 640×512。IRVI 数据集包含交通和监控两个数据集子集,交通场景包含 17 000 对训练图像和 1 000 对验证图像,监控场景包含 1 731 对训练图像和 150 对验证图像,图像的分辨率为 256×256。

在训练前,将数据集中的每一张图像大小裁剪为 256×256。网络训练过程中的初始学习率设置为 0.000 2,权重衰减设置为 0.000 1,训练的周期设定为 200,批处理大小设定为 64。

2.2　不同算法结果比较

为验证本文算法在不同场景下的有效性,将本文算法与 4 种基于深度学习的代表性彩色化算法分别在 KAIST 与 IRVI 两个数据集上进行实验对比,分别是算法 1:CycleGAN[25]、算法 2:CDGAN[26]、算法 3:I2VGAN[13]、算法 4:MCL[27]。算法结果对比如图 8 和图 9 所示。图 8 展示了 KAIST 数据集中 6 组市中心场景的马路和行人的图像比较,从图 8 可以看出 CycleGan 算法着色效果并不好,图像中建筑物和车辆融合在一起区分不清,天空等背景呈现灰绿色,图像失真严重。其他 3 种算法着色效果明显比 CycleGan 算法更好,车辆、人物、树木以及建筑物都能很好地与天空背景区分,但是 CDGAN 算法和 I2VGAN 算法对于饱和度高的颜色会有失真的情况,如身穿蓝色衣服的人物,并且图片中天空以及斑马线处都有曝光现象,斑马线的边缘细节是存在缺失的。同样,虽然 MCL 算法相较于前 3 种算法着色分布更均匀,图像更为清晰,但仍存在局部区域颜色失真、建筑物与斑马线等细节之处不清晰的问题。与其他算法相比,本文算法在局部区域避免了误着色、纹理模糊和细节扭曲等问题,同时使内部物体的边界更加清晰。

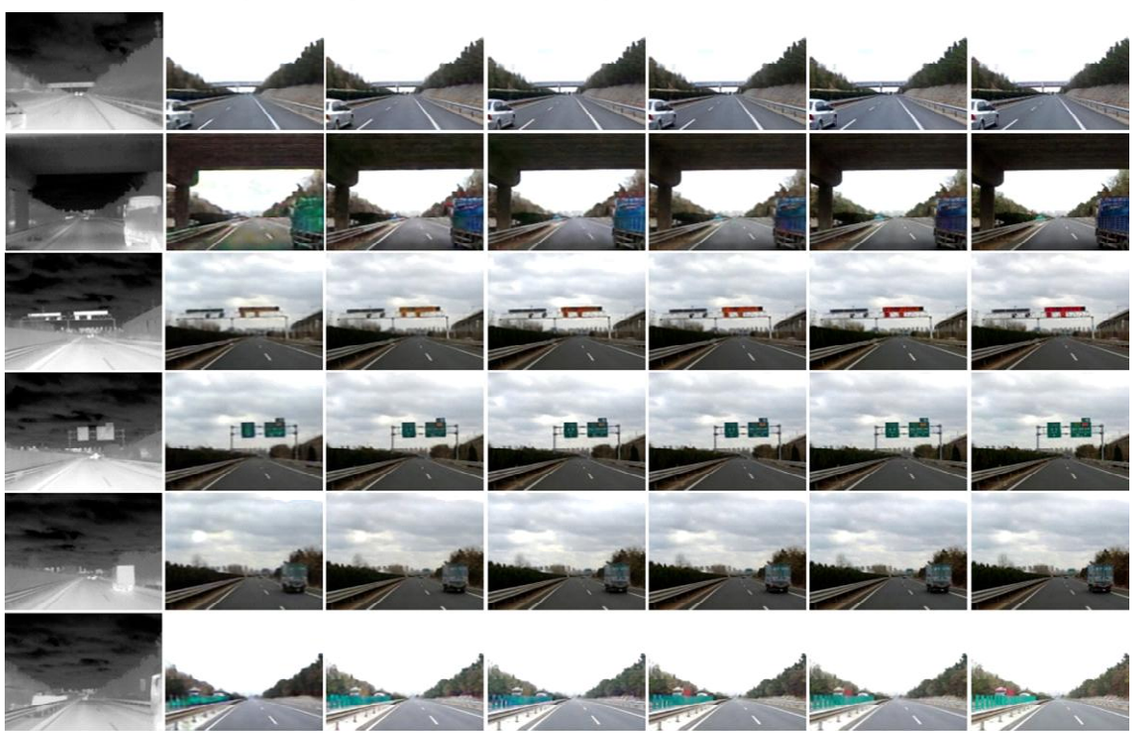
图 9 展示了 IRVI 数据集中 6 组交通场景的图像比较,该数据集包含的图像较少,且验证图像与部分训练图像相似。从图 9 可以看出,除 CycleGAN 算法的结果存在严重色差以及物体之间边界线融合外,大部分方法生成的图像质量较高。但是,与其他方法相比,本文算法在细节恢复方面取得了最好的性能。从第一和二组图像可以看到,广告牌上的字体和卡车车身上的纹理都得到了很好的恢复,并且在第六组图像中本文算法恢复的栅栏根根分明,细节明显。第五组图像中天空上的光晕并无伪影,去噪效果良好。



(a) 近红外图像 (b) CycleGAN算法结果 (c) CDGAN算法结果 (d) I2VGAN算法结果 (e) MCL算法结果 (f) 本文算法结果 (g) 可见光图像

图8 KAIST数据集上各算法对比结果

Fig. 8 Comparison results of each algorithm for the KAIST dataset



(a) 近红外图像 (b) CycleGAN算法结果 (c) CDGAN算法结果 (d) I2VGAN算法结果 (e) MCL算法结果 (f) 本文算法结果 (g) 可见光图像

图9 IRVI数据集上各算法对比结果

Fig. 9 Comparison results of each algorithm for the IRVI dataset

综上,本文方法在不同数据集上的彩色化效果明显优于其他算法,更贴近真实场景图像颜色。

2.3 评价指标比较

选用峰值信噪比PSNR(Peak signal to noise ratio)和结构相似性SSIM(Structure similarity in-

dex measure）作为各图像的客观评价指标进行对比，来验证本文算法的有效性。峰值信噪比PSNR是一种衡量图像质量的指标，单位为dB，其公式为：

$$\mathrm{PSNR}=20\log_{10}\left(\frac{\mathrm{MAX}_I}{\sqrt{\mathrm{MSE}}}\right) \tag{11}$$

式中：n为每个像素；MAX_I为图像颜色的最大值，MSE表示为：

$$\mathrm{MSE}=\frac{\sum_{i=1}^{n}(f(x)-y)^2}{n} \tag{12}$$

结构相似性SSIM是衡量两张图像相似度的指标，表示为：

$$\mathrm{SSIM}(X,Y)=\frac{(2\mu_x\mu_y+C_1)(2\delta_{xy}+C_2)}{(\mu_x^2+\mu_y^2+C_1)(\delta_x^2+\delta_y^2+C_2)} \tag{13}$$

式中：X为可见光图像；Y为生成的彩色化图像。

从表2和表3的数据可以看出，本文算法相较于其他算法PSNR和SSIM均有所提升，由此说明本文提出的彩色化算法具有较低的失真度、较高的着色准确性，相较于其他算法表现更优。

2.4　消融实验

为验证本文算法对CycleGAN网络改进后添加的空洞空间金字塔池化ASPP、注意力模块CBAM和感知损失函数对红外图像彩色化的有效性，使用前面提到的峰值信噪比和结构相似性作为评价指标，从表4中可以看出，添加ASPP模块和CBAM模块后PSNR和SSIM值均高于原始CyclGAN网络，并且添加ASPP模块比添加CBAM模块对算法进行了明显的优化。对改进算法添加感知损失函数后，上述两个指标也有些许提升，其中同时加入CBAM和ASPP融合并引入感知损失函数的网络效果最好，两种对比指标都有较高的值。

其中，针对感知损失函数对本文算法着色效果有效性的验证，分别在引入感知损失函数和未引入感知损失函数两种情况下计算200个迭代周期的评价指标并进行对比，变化折线图如图10所示，其中灰线代表添加感知损失函数，黑线代表未添加感知损失函数。从图10中可以看出，引入感知损失函数的LPIPS值在大部分迭代周期要小于未引入感知损失函数，二者的整体趋势呈下降趋势并且在140周期以后LPIPS相差值变大，由此证明迭代周期越长感知相似度值越优越，图像质量越接近人类的视觉感知，并且引入感知损失

表2　图8中各图像客观评价指标对比

Table 2　Comparison of objective evaluation indexes of each image in Fig. 8

Image	Evaluation indicator	CycleGAN	CDGAN	I2VGAN	MCL	Ours
Group 1	PSNR/dB	19.80	26.18	29.45	28.84	**33.80**
	SSIM	0.72	0.85	0.90	0.93	**0.96**
Group 2	PSNR/dB	17.38	20.25	22.82	24.03	27.52
	SSIM	0.84	0.86	0.89	0.90	0.92
Group 3	PSNR/dB	22.50	25.87	25.67	28.05	31.29
	SSIM	0.83	0.90	0.94	0.94	0.95
Group 4	PSNR/dB	21.21	24.21	25.70	27.38	28.03
	SSIM	0.78	0.90	0.91	0.92	0.94
Group 5	PSNR/dB	19.42	23.93	24.02	25.76	27.86
	SSIM	0.77	0.89	0.90	0.93	0.93
Group 6	PSNR/dB	22.31	23.11	25.24	27.34	29.00
	SSIM	0.80	0.87	0.92	0.91	0.93

表3　图9中各图像客观评价指标对比

Table 3　Comparison of objective evaluation indexes of each image in Fig. 9

Image	Evaluation indicator	CycleGAN	CDGAN	I2VGAN	MCL	Ours
Group 1	PSNR/dB	22.32	27.55	29.27	30.27	32.97
	SSIM	0.88	0.92	0.93	0.94	0.96
Group 2	PSNR/dB	22.22	25.33	24.89	26.68	29.99
	SSIM	0.67	0.87	0.88	0.90	0.93
Group 3	PSNR/dB	23.22	29.26	29.46	31.35	31.73
	SSIM	0.88	0.90	0.93	0.95	0.96
Group 4	PSNR/dB	21.49	29.61	29.48	31.65	32.09
	SSIM	0.88	0.87	0.91	0.93	0.97
Group 5	PSNR/dB	19.55	30.50	29.10	30.46	**33.19**
	SSIM	0.83	0.91	0.89	0.90	**0.97**
Group 6	PSNR/dB	22.17	27.76	28.10	29.03	31.68
	SSIM	0.82	0.87	0.90	0.93	0.95

表4　近红外图像彩色化的客观评价指标对比

Table 4　Comparison of evaluation indexes for colorization of NIR images

算法	PSNR/dB	SSIM
CycleGAN	27.64	0.82
CycleGAN＋ASPP	31.34	0.91
CycleGAN＋CBAM	29.65	0.87
CycleGAN＋ASPP＋CBAM	33.27	0.93
CycleGAN＋ASPP＋CBAM＋感知损失函数	**33.12**	**0.96**

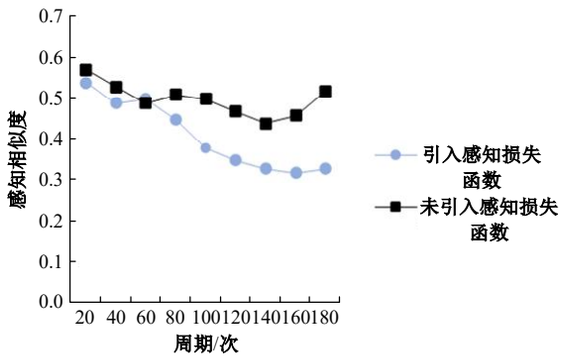


图10 改进前后感知相似度变化趋势

Fig. 10 Learned perceptual image patch similarity trend before and after improvement

函数以后图像着色效果更好。

实验表明:本文对CycleGAN网络进行改进后的算法显著改善了红外图像彩色化后图像色彩失真的情况,彩色化图像质量有较大的提升,效果更好。

3 结束语

本文提出了一种适用于近红外图像彩色化的RAUGAN模型,利用空洞空间金字塔池化和注意力模块的优势解决了红外图像彩色化过程中图像色彩失真、细节信息丢失和重点物体不突出的问题,在提高判别器性能的同时引入了感知损失函数,提高了色彩的饱和度。实验结果证明:本文提出的RAUGAN模型方法无论在主观人类感知还是客观评价指标方面相比于其他基于深度学习的彩色化算法,都显著改善了红外图像彩色化后色彩失真、语义模糊和纹理形状不清晰的问题,峰值信噪比(PSNR)和结构相似性指数(SSIM)分别提高了6%和15%。生成的图像能很好地体现细节信息并且色彩均衡,边界区分明显。

参考文献:

[1] Reinhard E, Adhikhmin M, Gooch B, et al. Color transfer between images[J]. IEEE Computer Graphics and Applications, 2001, 21(5): 34-41.

[2] Welsh T, Ashikhmin M, Mueller K. Transferring color to greyscale images[J]. ACM Transactions on Graphics, 2002, 21(3): 277-280.

[3] Levin A, Lischinski D, Weiss Y. Colorization using optimization[J]. ACM Transactions on Graphics, 2004, 23(3): 689-694.

[4] Goodfellow I J, Pouget A J, Mirza M, et al. Generative adversarial nets[C]//Advances in Neural Information Processing Systems, Montréal, Canada, 2014: 2672-2680.

[5] Isola P, Zhu J, Zhou T, et al. Image-to-image translation with conditional adversarial networks[C]//Proceedings of the IEEE Conference on ComputerVision and Pattern Recognition, Las Vegas, USA, 2016: 1125-1134.

[6] Park T, Liu M Y, Wang T C, et al. Semantic imagesynthesis with spatially-adaptive normalization[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Long Beach, USA, 2019: 2337-2346.

[7] Wang T C, Liu M Y, Zhu J Y, et al. High-resolutionimage synthesis and semantic manipulation withconditional gans[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, USA, 2018: 8798-8807.

[8] Huang X, Liu M Y, Belongie S, et al. Multimodal unsupervised image-to-image translation[C]//Proceedings of the European Conference on Computer Vision, Munichi, Germany, 2018:172-189.

[9] Lee H Y, Tseng H Y, Huang J B, et al. Diverse image-to-image translation via disentangled representations[C]//Proceedings of the European Conference on Computer Vision, Munichi, Germany, 2019: 35-51.

[10] Yi Z, Zhang H, Tan P, et al. Dualgan: unsupervised dual learning for image-to-image translation[C]//Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 2017: 2849-2857.

[11] Suárez P L, Sappa A D, Vintimilla B X. Infrared image colorization based on a triplet dcgan architecture [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Honolulu, USA, 2017: 212-217.

[12] 王晓宇. 基于颜色迁移的图像彩色化算法研究[D]. 长春: 长春理工大学电子信息工程学院, 2020.
Wang Xiao-yu. Research on image colorization algorithm based on color migration[D]. Changchun: School of Electronic and Information Engineering, Changchun University of Science and Technology, 2020.

[13] Li S, Han B F, Yu Z J, et al. I2VGAN:unpaired Infrared to-visible video translation[C]//Proceedings of

the 29th ACM International Conference on Multimedia, New York,USA, 2021:3061-3069.

[14] 高美玲, 段锦, 莫苏新, 等. 基于空洞循环卷积的近红外图像彩色化方法[J]. 光学技术, 2022, 48(6): 742-748.
Gao Mei-ling, Duan Jin, Mo Su-xin, et al. Near infrared image colorization method based on dilated-cycle convolution[J]. Optical Technique, 2022, 48(6): 742-748.

[15] Chen J, Chen J, Chao H, et al. Image blind denoising with generative adversarial network based noise modeling[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, USA, 2018:3155-3164.

[16] 陈雪云, 许韬, 黄小巧. 基于条件生成对抗网络的医学细胞图像生成检测方法[J]. 吉林大学学报: 工学版, 2021, 51(4): 1414-1419.
Chen Xue-yun, Xu Tao, Huang Xiao-qiao. Detection method of medical cell image generation basedon conditional generative adversarial network[J]. Journal of Jilin University (Engineering and Technology Edition), 2021, 51(4): 1414-1419.

[17] 王小玉, 胡鑫豪, 韩昌林. 基于生成对抗网络的人脸铅笔画算法[J]. 吉林大学学报: 工学版, 2021, 51(1): 285-292.
Wang Xiao-yu, Hu Xin-hao, Han Chang-lin. Face pencil drawing algorithms based on generative adversarial network[J]. Journal of Jilin University (Engineering nd Technology Edition), 2021, 51(1): 285-292.

[18] Monday H N, Li J, Nneji G U, et al. A wavelet convolutional capsule network with modified super resolution generative adversarial network for fault diagnosis and classification[J]. Complex & Intelligent Systems, 2022, 8: 4831-4847.

[19] 彭晏飞, 张平甲, 高艺, 等. 融合注意力的生成式对抗网络单图像超分辨率重建[J]. 激光与光电子学进展, 2021, 58(20): 182-191.
Peng Yan-fei, Zhang Ping-jia, Gao Yi, et al. Attention fusion generative adversarial network for single-image super-resolution reconstruction[J]. Laser & Optoelectronics Progress, 2021, 58(20): 182-191.

[20] Liang C C, George P, Iasonas K, et al. DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40(4): 834-848.

[21] Woo S, Park J, Lee J Y, et al. CBAM: convolutional block attention module[C]//Proceedings of the 15thEurope an Conference on Computer Visdom, Munich, Germany, 2018: 3-19.

[22] Li C, Wan D M. Precomputed real-time texture synthesis with markovian generative adversarial networks [C]//European Conference on Computer Vision of IEEE, Amsterdam, The Netherlands, 2016: 702-716.

[23] Wang S, Park J, Kim N, et al. Multispectral pedestrian detection: benchmark dataset and baseline[C]// IEEE Conference on Computer Vision and Pattern Recognition, Boston, USA, 2015: 1037-1045.

[24] Park T, Liu M Y, Wang T C, et al. Semantic image synthesis with spatially-adaptive normalization[J]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Long Beach, USA, 2019:2337-2346.

[25] Zhu J Y, Park T, Isola P, et al. Unpaired image-to-image translation using cycle-consistent adversarial network[C]//IEEE International Conference on Computer Vision, Venice, Italy, 2017: 2223-2232.

[26] Kancharagunta B K, Ram S D. CDGAN: cyclic discriminative generative adversarial networks forimage-to-image transformation[J]. Journal of Visual Communication and Image Representation, 2022, 82: 103382.

[27] Gou Y. Multi-feature contrastive learning for unpaired image-to-image translation[J]. Intelligent Systems, 2023, 9: 4111-4122.