

〈图像处理与仿真〉

基于空洞全局注意力机制的近红外图像彩色化方法

高美玲，段锦，赵伟强，胡奇

(长春理工大学 电子信息工程学院，吉林 长春 130012)

摘要：针对目前卷积神经网络未能充分提取图像的浅层特征信息导致近红外图像彩色化算法存在结果图像局部区域误着色及网络训练不稳定导致结果出现模糊问题，提出了一种新的生成对抗网络方法用于彩色化任务。首先，在生成器残差块中引入自行设计的空洞全局注意力模块，对近红外图像的每个位置理解更加充分，改善局部区域误着色问题；其次，在判别网络中，将批量归一化层替换成梯度归一化层，提升网络判别性能，改善彩色化图像生成过程带来的模糊问题；最后，将本文算法在 RGB_NIR 数据集上进行定性和定量对比。实验表明，本文算法与其他经典算法相比能充分提取近红外图像的浅层信息特征，在指标方面，结构相似性提高了 0.044，峰值信噪比提高了 0.835，感知相似度降低了 0.021。

关键词：彩色化；近红外图像；生成对抗网络；空洞全局注意力；梯度归一化

中图分类号：TP391 文献标识码：A 文章编号：1001-8891(2023)10-1096-10

Near-infrared Image Colorization Method Based on a Dilated Global Attention Mechanism

GAO Meiling, DUAN Jin, ZHAO Weiqiang, HU Qi

(College of Electronic and Information Engineering, Changchun University of Science and Technology, Changchun 130012, China)

Abstract: A new generative adversarial network method is proposed for colorization of near-infrared (NIR) images, because current convolutional neural networks fail to fully extract the shallow feature information of images. This failure leads to miscoloring of the local area of the resultant image and blurring due to unstable network training. First, a self-designed dilated global attention module was introduced into the generator residual block to identify each position of the NIR image accurately and improve the local region miscoloring problem. Second, in the discriminative network, the batch normalization layer was replaced with a gradient normalization layer to enhance the network discriminative performance and improve the blurring problem caused by the colorized image generation process. Finally, the algorithms used in this study are compared qualitatively and quantitatively using the RGB_NIR dataset. Experiments show that the proposed algorithm can fully extract the shallow information features of NIR images and improve the structural similarity by 0.044, PSNR by 0.835, and LPILS by 0.021 compared to other colorization algorithms.

Key words: colorization, near-infrared images, generative adversarial networks, dilated global attention, gradient normalization

0 引言

近红外光是介于可见光和中红外光之间的电磁波，其波长范围为 780~1000 nm。近红外成像技术穿

透性能非常高，即使在恶劣天气下也能获得细节丰富的图像。但是与符合人类视觉系统的彩色图像相比，近红外图像缺少颜色信息，人类对灰度图像的敏感度低于彩色图像敏感度，且近红外图像具有可见光图像

收稿日期：2022-09-07；修订日期：2022-09-29。

作者简介：高美玲（1997-），女，辽宁锦州人，博士研究生，主要研究方向：图像处理等。

通信作者：段锦（1971-），男，吉林长春人，博士，教授，博士生导师，主要研究方向：图像处理与模式识别等。

基金项目：吉林省科技发展计划项目（20210203181SF）。

不具备的纹理细节特征,故需要将其彩色化来增加色彩信息,使近红外图像能更符合人类的视觉感官。

目前针对近红外图像彩色化算法主要分为两大类,一类是基于参考图像的颜色像素匹配方法,另一类是基于深度学习的信息特征映射方法。

基于参考图像的颜色像素匹配方法需要将近红外图像与参考图像转换到特定颜色空间,利用各像素级进行映射实现图像彩色化。主要分为两类,一类是局域颜色拓展着色,另一类是颜色传递着色。针对局域颜色拓展着色,戴康^[1]利用图像的局部马尔可夫性质通过最小化相邻像素间的颜色区别实现着色。Musalski^[2]等人实现个性化用户输入偏好的物体颜色信息以实现图像的彩色化;针对颜色传递着色,Welsh^[3]等人利用查找匹配像素的方式实现近红外图像的颜色迁移。Reinhard^[4]等人提出色彩迁移公式,在CIELAB (commission international eclairage LAB,CIELAB)颜色空间转换进行着色。以上两种方法中均需要一张特定的参考图像作为输入,且颜色传递方法需要人工操作进行上色,生成结果因素复杂且操作量较大。

基于深度学习的信息特征映射方法,通常将灰度图像和可见光图像两个不同域作为网络的输入,两个域的图像不断地通过卷积神经网络进行特征提取,网络不断地进行两个域之间的特征学习,进而实现近红外图像彩色化。主要分为两类,一类是卷积神经网络着色,另一类是生成对抗网络着色。针对卷积神经网络着色,冯佳男^[5]等通过卷积神经网络将低级特征与高级特征进行融合着色实现图像彩色化,但由于近红外图像浅层特征与深层特征进行融合会导致浅层特征信息被覆盖,从而边缘漫色问题依旧存在。Cheng^[6]等利用色度分量和联合双边滤波结合得到彩色化结果,但生成效果不佳;针对生成对抗网络着色,Zhu^[7]等人提出双向循环生成对抗网络模型,引入双向循环一致损失函数更好地实现图像彩色化,但由于生成对抗网络因陡峭的梯度空间导致网络训练不稳定,使得该方法彩色化结果出现模糊现象;Deblina^[8]等人提出DUNIT网络,借用无监督框架完成黑天到白天的转换及目标检测两大视觉任务,但结果图像纹理细节不清晰。万园园^[9]等人提出的UNIT和GAN相结合实现灰度图像彩色化,在一定程度上提升了彩色化效果,但错误着色、模糊问题依旧存在。

目前已有的配对并标注好的近红外-可见光图像数据集稀少,这为近红外图像彩色化任务增添了难题。但生成对抗网络着色采用的是无监督学习,无需成对图像就能完成基本的着色任务,且独特的网络结

构和训练机制能更好地学习建立近红外图像与可见光图像之间的映射关系,使得生成对抗网络结构在大多数视觉任务中广泛应用,如图像超分辨率^[10-11],图像分割^[12-13]及目标识别^[14-15]等。

受此启发,本文设计了一种新的近红外图像彩色化生成对抗网络结构,将空洞全局注意力机制和梯度归一化相结合完成近红外图像彩色化任务。第一,为解决彩色近红外图像中出现的局部区域误着色问题,本文结合空洞卷积模块和全局注意力模块的优势,自行构建了一个名为空洞全局注意力机制模块,该模块利用空洞卷积模块不同扩张率实现不同感受野下同等特征图的信息,解决近红外图像特征感受野受限制的问题,并结合全局注意力机制模块中通道和空间两路注意力操作来解决上下级近红外图像特征理解不充分、浅层特征提取不足问题。在生成器网络中将空洞全局注意力模块融入到残差块中,用来充分提取近红外图像特征并反卷积进行着色。第二,为解决彩色近红外图像出现的模糊问题,本文在判别网络中将梯度归一化代替批量归一化,克服了批量归一化模块在训练过程中不稳定的问题,提高了彩色近红外图像的生成质量。

1 空洞全局注意力机制彩色化 GAN 模型

本文模型的总体框架如图 1 所示。该模型采用 ResNet 基础模型作为生成网络,采用马尔可夫判别器 (Markovian discriminator)^[16]基础模型作为判别网络。近红外图像域和可见光图像域分别定义为 I 和 V 。主要包括两个生成器结构, $G_{I \rightarrow V} = I \rightarrow V'$ 和 $F_{V \rightarrow I} = V \rightarrow I'$, 两个判别器 $D_V : V \rightarrow V'$ 和 $D_I : I \rightarrow I'$ 。对于原始 GAN, 生成器 $G_{I \rightarrow V}$ 是近红外图像生成彩色图像并且判别器 D_V 是用来鉴别是否是生成的假彩色图像。类似地,对于双向循环 GAN, 生成器 $F_{V \rightarrow I}$ 是将彩色图像生成近红外图像并且判别器 D_I 是用来鉴别是否是生成的假近红外图像。

1.1 生成器

本文模型生成器在 ResNet 网络残差块基础上引入了空洞全局注意力机制模块 (dilated global attention module, DA Block), 该模块充分结合了空洞卷积模块和全局注意力模块的优势,解决了近红外图像浅层特征理解不充分的问题;利用空洞卷积模块的优点解决生成对抗网络浅层特征图感受野受限制,不能获取近红外图像全部细节信息的缺陷;利用全局注意力模块的优点解决近红外图像全局上下文特征信息理解不充分、局部区域误着色问题,通道注意力模块引入全连接进行降维并提取近红外图像中更重要的细节信

息，空间注意力模块引入池化操作提取近红外图像不同特征并增强特征多样性，使得近红外图像更能充分理解上下文信息，着色结果准确率显著提升。生成器结构如图 2 所示，空洞全局注意力机制模块如图 3 所示，网络结构如图 4 所示。

1) 空洞卷积模块

空洞卷积是指在卷积核元素之间填充 0 来扩大卷积核的过程。空洞卷积如图 5 所示。

若 a 为扩张率, k 为原卷积核大小, 则加入空洞卷积后尺寸为 $k=k+(k-1)(a-1)$ 。比如 $a=1, 2, 4$ 时感受野如图 6 所示。

从图 6 中可以看出空洞卷积可以增大输出单元的感受野且不会增大卷积核大小，带有多个不同空洞率的卷积进行叠加时不同的感受野可以带来多尺度信息，获取上下文多尺度信息。本文基于 TridderNet^[17]的思想，通过空洞率分别为 1、2、3 的空洞卷积实现不同感受野下同等特征图的获取。

2) 全局注意力模块

全局注意力模块^[18]用于在减少信息弥散情况下也同时放大全局维度交互特征。给定一个中间特征图作为输入，卷积注意力模块会依次沿通道和空间两个独立维度判断注意力图，然后将其乘以自适应特征进行优化得到结果。

通道注意力模块是将中间特征图作为输入，分别通过全局池化和最大池化两个操作转成 $1 \times 1 \times C$ ，然后经过多层感知机进行全连接逐像素加权操作，再经过 Sigmoid 激活操作生成通道注意力特征图，最后将通道注意力特征图与输入逐像素加乘得到最终的通道注意力特征图。该特征图关注的是该近红外图像上哪些内容有重要作用，通道注意力机制可表达为公式(1)所示：

式中： σ 为 Sigmoid 函数；MLP 是多层感知机；AvgPool 为平均池化；MaxPool 为最大池化。

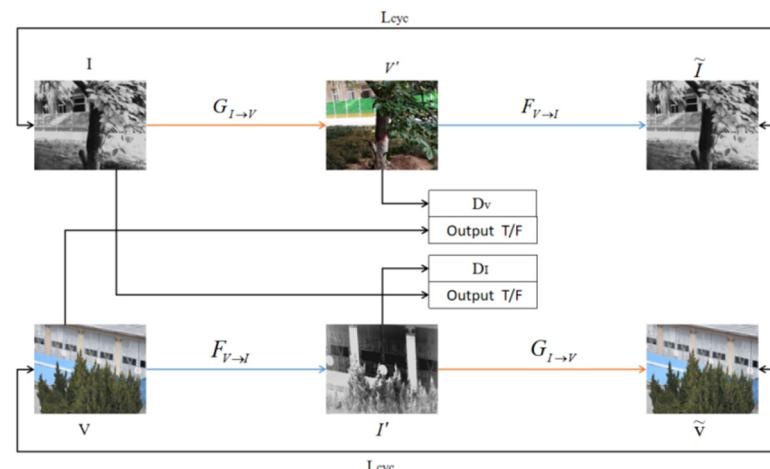


图 1 本文彩色化 GAN 模型总体框架

Fig.1 Overall framework for colorized GAN models in this paper

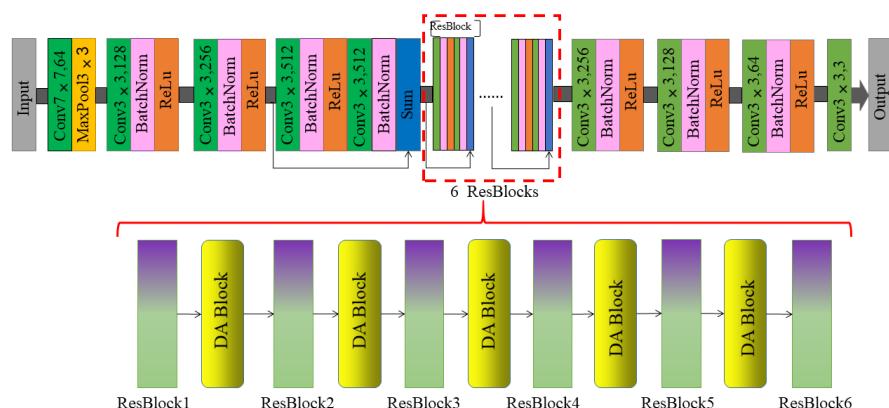


图 2 生成器结构

Fig.2 Structure of generator

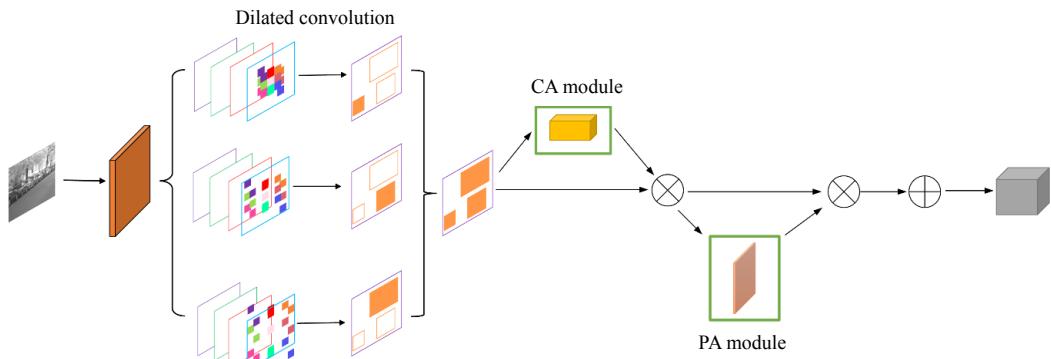


图3 空洞全局注意力机制模块

Fig.3 Dilated global attention module

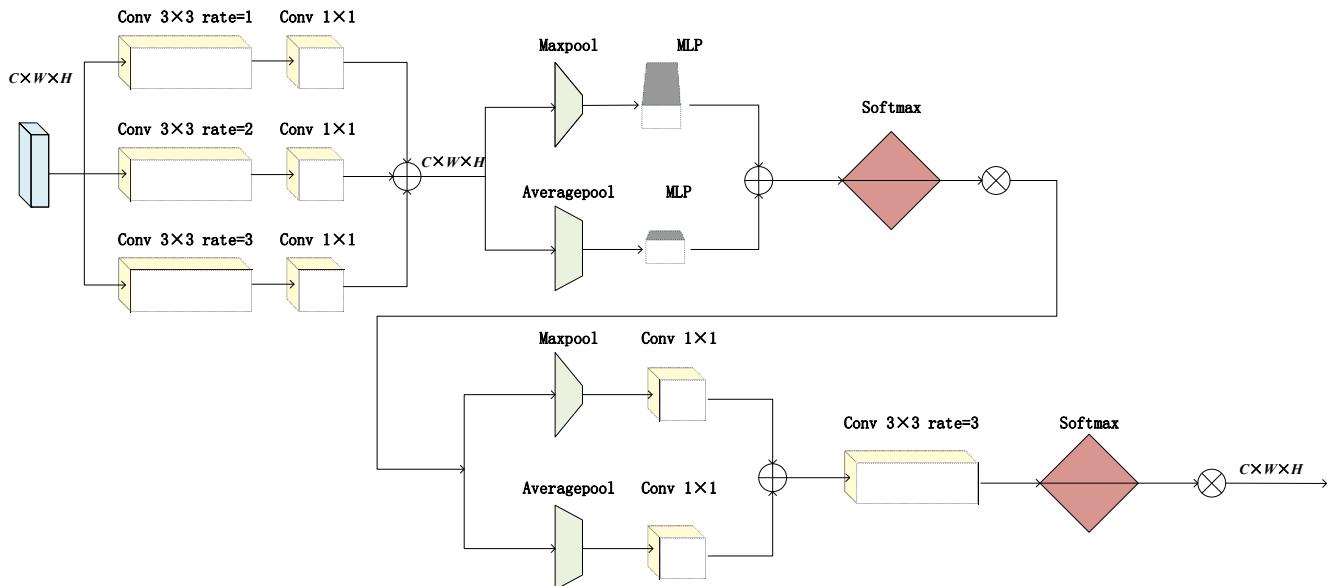
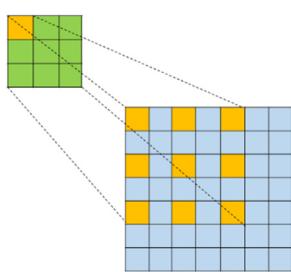


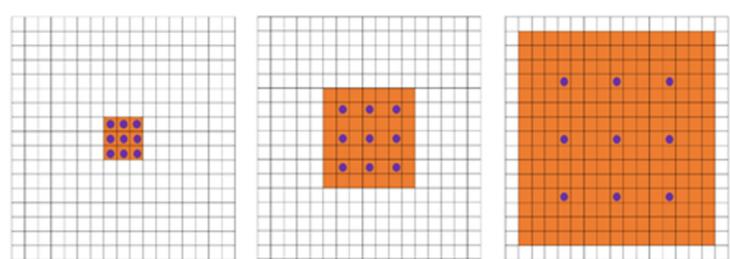
图4 空洞全局注意力机制模块网络结构

Fig.4 The network architecture of dilated global attention mechanism module

图5 空洞卷积
Fig.5 Dilated convolution

空间注意力模块是将最终通道注意力特征图作为输入，经过基于通道的全局最大池化和平均池化操作，接着将这两个结果进行通道融合操作，再降维成1个通道的特征图，再经过Sigmoid激活函数生成空间注意力机制图，最后将该特征图与输入特征图逐像素加乘得到最终的结果。空间注意力机制表达式如公式(2)所示：

$$M_s(F) = \sigma\left(f^{7 \times 7} \left(\text{AvgPool}(F); \text{MaxPool}(F)\right)\right) \quad (2)$$

图6 不同扩张率下感受野
Fig.6 Receptive filed at different expansion rates
式中：7×7代表卷积核。

1.2 判别器

本文模型判别器采用的是原始马尔可夫模型。判别器属于卷积神经网络，为了提取图像特征，最后一层输出一维特征图判断图像的真假。针对训练不稳定带来的彩色化结果模糊问题，本文在原始判别器批量归一化层（batch normalization, BN）^[19]替换成梯度归一化层（gradient norm, GN），为判别函数施加了一个L1-利普希茨约束，解决了判别器因陡峭的空间造成

训练不稳定的问题，提升了判别器性能。本文判别网络模型如图 7 所示。

1) 批量归一化

批量归一化是 2015 年 Google 研究人员提出的一种参数归一化手段。BN 层可使网络超参数的设定更加宽泛，收敛较快。BN 操作在数据输入层，对输入数据进行求均值、方差做归一化，具体操作如下：

输入：输入 $x: \beta = \{x_1, \dots, m\}$

输出：归一化后的网络响应 $\{y_i = BN_{\gamma, \beta}(x_i)\}$

$$\textcircled{1} \text{ 计算批量处理数据均值 } \mu_\beta \leftarrow \frac{1}{m} \sum_{i=1}^m x_i;$$

② 计算批量处理数据方差

$$\sigma_s^2 \leftarrow \frac{1}{m} \sum_{i=1}^m (x_i - \mu_\beta)^2;$$

$$\textcircled{3} \text{ 归一化 } \hat{x}_i \leftarrow \frac{x_i - \mu_\beta}{\sqrt{\sigma_\beta^2 + \epsilon}};$$

$$\textcircled{4} \text{ 尺度变换 } y_i \leftarrow \gamma \hat{x}_i + \beta = BN_{\gamma, \beta}(x_i);$$

⑤ 返回学习参数 γ 和 β 。

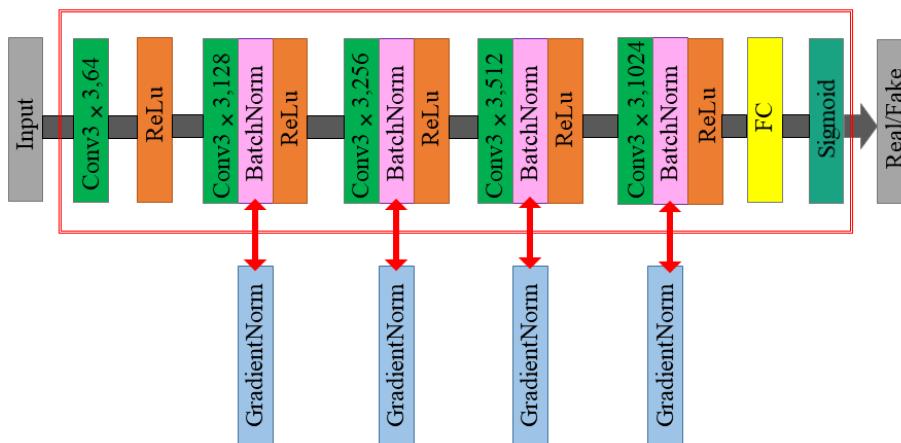


图 7 判别模型

Fig.7 Discriminator model

2) 梯度归一化

梯度归一化 (gradient normalization, GN) 是 2021 年 Wu^[20]提出的带有 L1-利普希茨约束的归一化方式，该文主要解决的问题就是生成对抗网络生成图像质量较模糊问题，并且在 CIFAR10、CelebA-HQ 及 LSUN Church Outdoor 等数据集上进行测试，证明 GN 同时满足模型集、非基于采样及硬约束 3 个特性，且可兼容各种网络架构，带有约束的参数不会损失生成图像的分辨率，不会带有额外的超参数。具体操作如下：

输入：生成器与判别器参数 θ_G, θ_D ，学习率 α_G, α_D ，批处理图片 M ，更新率 N_{dis} ，总迭代次数 N 。

输出：归一化后的网络响应 $\{\theta_G = GN_{\alpha_{\theta_G}}(z_i)\}$

$$\textcircled{1} \hat{D} := D(x) / (\|\nabla_x D(x)\| + |D(x)|);$$

② 对于 $t=1 \sim N$ ；

③ 对于 $t_{dis}=1 \sim N$ ；

④ 对于 $i=1 \sim m$ ；

⑤ $z \sim p_z, x \sim p_r(x)$ ；

$$\textcircled{6} L_D^i \leftarrow \hat{D}[(G(z)) - \hat{D}(x)];$$

$$\textcircled{7} \theta_D \leftarrow Adam(\frac{1}{2M} \sum_{i=1}^M L_D^i, \alpha_D, \theta_D);$$

⑧ 潜在空间的随机样本 $\{z_i\}_{i=1}^{2M} \sim p_z(z)$ ；

$$\textcircled{9} \theta_G \leftarrow Adam[\frac{1}{2M} \sum_{i=1}^M (-\hat{D}(G(z_i))), \alpha_G, \theta_G].$$

批量归一化虽有优点，但在实际网络训练中，训练批尺寸参数通常设置为 1，无法有效地处理所有批次图像，进而导致网络在训练过程中出现不稳定，出现生成图像模糊问题，而原梯度归一化论文的思想就是为了解决生成对抗网络结果图像生成模糊问题，故本文模型将批量归一化层替换成梯度归一化层进行验证。

1.3 损失函数设计

本文算法的损失函数基于生成对抗网络进行设计，为了使生成图像的细节和边缘纹理更接近原始近红外图像，在对抗网络损失函数基础上引入 SSIM (structure similarity index measure, SSIM) 损失函数，SSIM 损失函数在亮度、对比度和图像结构三方面考虑了两幅图像之间的细节。为了实现上述目标，本文损失函数如下所示。

1) 对抗损失。该损失目的是使生成的图像更加真实, 生成更加生动的结果, 是生成对抗网络中基础的损失函数。

$$L_{\text{GAN}}(G, D_V, I, V) = E_{v \sim P_{\text{data}(v)}} [\lg D_{v(v)}] + E_{i \sim P_{\text{data}(i)}} [\lg (1 - D_v(G(i)))] \quad (3)$$

式中: G 是生成函数; D 是判别函数; I 为近红外图像域; V 为可见光图像域。

2) 循环一致损失。该损失用来增强整体结构信息, 降低近红外图像域和可见光图像域间错误映射的概率。

$$L_{\text{cycle}}(G, F) = E_{i \sim P_{\text{data}(i)}} [\|F(G(i) - i)\|_1] + E_{v \sim P_{\text{data}(v)}} [\|G(F(v) - v)\|_1] \quad (4)$$

$F(G(i))$ 及 $G(F(v))$ 为前向预测, 差异越大, 预测与原始之间差别越大。

3) 识别损失。该损失旨在控制生成器随意修改生成的色调及样式, 增强生成图像的色调准确性等。

$$L_{\text{identity}}(G, F) = E_{v \sim P_{\text{data}(v)}} [\|G(v) - v\|_1] + E_{i \sim P_{\text{data}(i)}} [\|F(i) - i\|_1] \quad (5)$$

4) 结构相似性损失。用于衡量近红外图像与彩色近红外图像之间的差异, 提升生成器对输入图像结构的学习能力。

$$L_{\text{SSIM}(i, v)} = \frac{(2\mu_i\mu_v + c_1)(2\sigma_i\sigma_v + c_2)}{(\mu_i^2 + \mu_v^2 + c_1)(\sigma_i^2 + \sigma_v^2 + c_2)} \quad (6)$$

式中: μ 和 μ_v 分别是近红外图像和生成的彩色图像的像素平均值; σ 和 σ_v 分别是近红外图像和生成的彩色图像的方差。

5) 总损失函数。联合上述所有损失函数, 本文目标是优化总损失函数。

$$L_{\text{total}} = L_{\text{GAN}}(G, D_V, I, V) + L_{\text{GAN}}(F, D_I, V, I) + \lambda_1 L_{\text{cycle}}(G, F) + \lambda_2 L_{\text{identity}}(G, F) + \lambda_3 L_{\text{SSIM}}(i, v) \quad (7)$$

式中: λ_1 、 λ_2 、 λ_3 为超参数, 调整 λ_1 可以调节循环一致损失函数的权重, $\lambda_1=10$; 调整 λ_2 可以调节识别损失函数的权重, $\lambda_2=10$; 调节 λ_3 可以调节结构相似性损失函数的权重, $\lambda_3=0.6$ 。

2 实验结果分析

2.1 实验条件

本实验在 Windows 10 操作系统下进行, 深度学习框架选取 PyTorch, CUDA 版本为 11.1, 开源视觉库为 Python-Open CV4.5.1, 图像工作站配置为: Intel Core i7-6700 CPU, 16 GB 内存, NVIDIA GeForce

RTX2060S (8G) 显卡, 选用 RGB_NIR^[21]数据集进行训练, 统一尺寸为 500×500。在构建网络时, 通过随机梯度下降更新生成器的和判别器的权重, 优化器采用 Adam, 学习率初始设置为 2×10^{-4} , 权重衰减设置为 1×10^{-4} , 批处理大小设置为 32, 训练周期数量为 200, 训练前 100 周期保持学习率不变, 训练时间为 9 h 左右。

2.2 评价指标

为了验证本文算法的彩色化效果, 本文从定性和定量两个方面进行评估。定性分析主要是从主观角度进行评价, 定量分析主要是从客观的评价指标作为参考进行评价。本文选用峰值信噪比 (peak signal to noise ratio, PSNR)、结构相似性 (structural similarity image similarity, SSIM)、感知相似度 (learned perceptual image patch similarity, LPIPS) 作为评价指标。结构相似性是用来衡量两幅图像在结构上的相似程度, 其值越大说明两者之间的结构未发生质变。峰值信噪比是为了衡量图像的失真程度, 数值越大说明生成图像的失真越小。感知相似度是用来衡量两幅图像之间的纹理结构相似程度, 数值越小说明近红外图像与彩色近红外图像纹理结构越相似, 画质越好。

2.3 不同算法之间的对比

为了验证本文算法的有效性, 将本文算法与 4 种典型的基于生成对抗网络的彩色化算法进行对比, 算法分别是算法 1: Deoldify^[22]、算法 2: Wei^[23]、算法 3: In2i^[24]、算法 4: CycleGAN 算法^[7]。算法结果对比如图 8 所示。指标比对如表 1 所示。

从图 8 可以看出, Deoldify 算法结果只针对天空上色非常好, 但是对于其他场景效果并不符合人类视觉感官。Wei 及 In2i 结果针对草坪和树木上色效果几乎呈现荧绿色, 存在很多不合理的细节颜色信息, 饱和度较差。CycleGAN 算法出现了严重的误着色现象, 草坪图像中几乎所有物体均呈现出红色, 无法区分物体颜色。本文算法相较于其他算法在牧场、大树、雕像图中没有明显的局部区域误着色现象, 且结果图像颜色整体分布合理。从表 1 可以看出, 本文算法除了在山峰图中 PSNR 较低于 CycleGAN 算法, 其他情况均取得了最佳的结果, 在牧场图中相较于其他算法 SSIM 提高了 0.09, PSNR 提高了 1.744, LPIPS 降低了 0.017, 在山峰图中相较于其他算法 SSIM 提高了 0.037, PSNR 减少了 0.467, LPIPS 降低了 0.016, 在雕像图中相较于其他算法 SSIM 提高了 0.004, PSNR 提高了 1.228, LPIPS 降低了 0.03。

2.4 消融实验

本实验采取不同的网络结构进行训练, 来验证

GAN 结构和各模块对算法性能的影响。共采取 4 组实验分别进行验证，分别是实验一：原始 GAN；实验二：GAN+空洞卷积注意力模块；实验三：GAN+梯度归一化；实验四：GAN+空洞卷积注意力模块+梯度归一化，即本文算法。从测试集中随机挑选 3 组实验结果如图 9 所示，指标对比如表 2 所示。

评价指标选取 IS 和 FID 作为衡量生成图像清晰度和质量的考核标准。Inception 分数 (Inception score, IS) 是对生成图片清晰度和多样性的衡量，IS 值越大越好。特征度量距离 (Fréchet inception distance, FID) 是用来衡量两组图像的相似度，是计算真实图像和生成图像的特征向量之间距离的一种度量。FID 分数常被用于评估由生成对抗网络生成的图像的质量，分数越低说明图像的质量越高。

综合图 9 和表 2 指标的结果可以看出，实验一结果会出现误着色、上色模糊问题，实验二在实验一基础上引入了空洞全局注意力机制模块，效果上可以清晰看出错误着色现象明显改善，实验三是在实验一基础上引入了梯度归一化模块，可以明显看出生成图像较实验一更为清晰，但误着色现象依旧存在，实验四也就是本文算法，在实验一基础上引入空洞全局注意力机制模块和梯度归一化模块，从结果图可以明显看出图像细节处颜色界限更加分明，模糊现象已明显改善，并且从指标中也可以看出，IS 平均提高了 0.491，FID 平均降低了 1.055。为此说明本文引入的空洞全局注意力机制模块解决了近红外图像局部区域误着色问题，引入的梯度归一化较好地改善了上色模糊现象。



图 8 各个算法对比结果：(a) 近红外图像；(b) Deoldify^[22]结果；(c) Wei^[23]结果；(d) In2i^[24]结果；
(e) CycleGAN 算法^[7]结果；(f) 本文算法结果；(g) 可见光图像

Fig.8 Comparison results of each algorithm: (a) NIR images; (b) Deoldify^[22] results; (c) Wei^[23] results;
(d) In2i^[24] results; (e) CycleGAN^[7] Results; (f) Our method results; and (g) Visible images

2.5 验证 SSIM 损失函数

为验证本文引入的 SSIM 损失函数着色效果是否有提升，采用均方误差 (mean-square error, MSE)、峰值信噪比作为评价指标。均方误差指的是真实图像与彩色近红外图像之间的误差，数值越大说明两者

相差越大。峰值信噪比是为了衡量图像的失真程度，数值越大说明生成图像的失真越小。

共采取两组实验进行验证。实验一：本文模型+原始总损失函数。实验二：本文模型+改进后总损失函数。将两组实验分别进行 200 周期训练，且每 25 周

期随机抽取 15 组图像进行 MSE 及 PSNR 测试，并取平均值绘制出最终曲线。其均方误差趋势曲线图和峰值信噪比趋势曲线图分别如图 10、图 11 所示。

图 10、11 中黑色线代表实验一结果，红色线代表实验二结果。从两个曲线趋势图中可以明显看出，引

入 SSIM 损失函数后，本文算法在 200 周期中，实验二中 MSE 较实验一提高了 3.0001，实验二 PSNR 较实验一提高了 2.71 dB，并且无论迭代次数处于什么阶段，MSE 和 PSNR 指标均优于实验一，进一步证明了引入 SSIM 损失后确实提升了近红外图像彩色化效果。

表 1 指标对比
Table 1 Comparison of indicators

Algorithm	Ranch image			Mountain image			Statue image		
	SSIM	PSNR/dB	LPIPS	SSIM	PSNR/dB	LPIPS	SSIM	PSNR/dB	LPIPS
Deoldify ^[22]	0.718	17.088	0.447	0.781	16.682	0.451	0.844	17.397	0.450
Wei ^[23]	0.743	17.533	0.419	0.699	16.054	0.354	0.705	17.316	0.377
In2i ^[24]	0.823	16.987	0.371	0.703	18.271	0.435	0.759	17.746	0.362
CycleGAN ^[7]	0.715	19.787	0.341	0.867	20.738	0.326	0.831	17.681	0.222
Our method	0.832	21.531	0.324	0.904	20.271	0.310	0.835	18.974	0.219

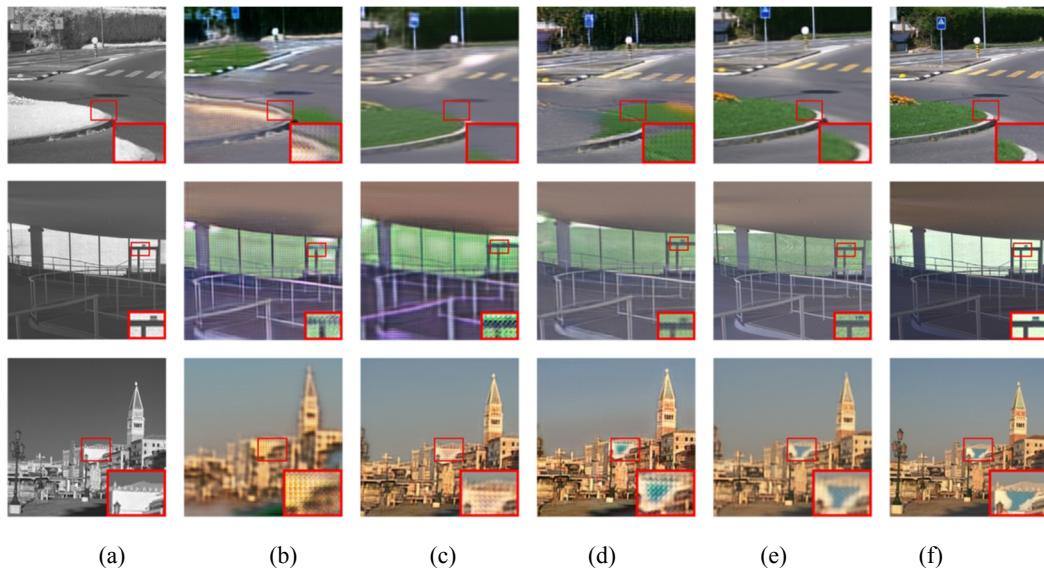


图 9 消融实验一对比算法结果: (a) 近红外图像; (b) 实验一结果; (c) 实验二结果;
(d) 实验三结果; (e) 实验四结果; (f) 可见光图像

Fig.9 Results of ablation exp.1 comparison algorithm: (a) NIR images; (b) Exp.1 results;
(c) Exp.2 results; (d) Exp.3 results; (e) Exp.4 results and (f) Visible images

表 2 消融实验一指标比对
Table 2 Comparison of ablation experiment 1 metrics

	Street		Football		Building	
	IS	FID	IS	FID	IS	FID
Exp.1	8.263	14.824	7.263	10.942	9.823	14.234
Exp.2	9.072	13.495	8.752	10.234	9.102	11.293
Exp.3	8.273	11.293	9.528	11.293	10.583	14.245
Exp.4	9.210	10.056	9.473	9.351	11.973	10.248

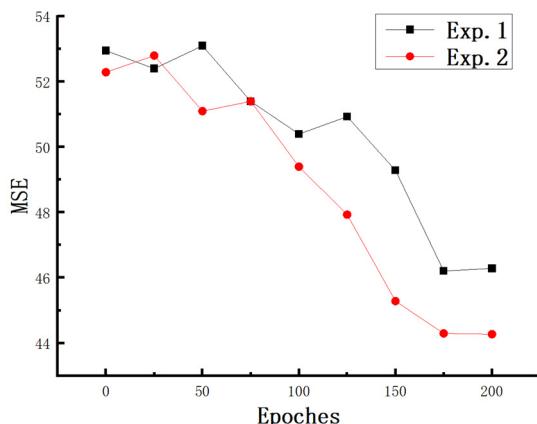


图 10 均方误差趋势

Fig.10 The trend of mean square error

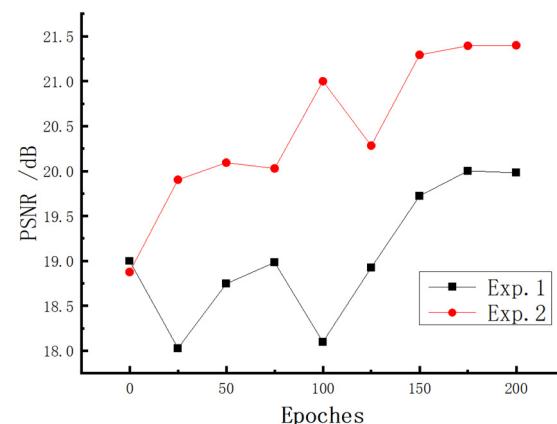


图 11 峰值信噪比趋势

Fig.11 The trend of peak signal-to-noise ratio

3 结论

本文提出一种适用于近红外图像的彩色化生成对抗网络模型，利用空洞全局注意力模块的优势解决了生成的结果图像出现的局部区域误着色问题，在判别网络中将梯度归一化代替原批量归一化，提高了判别器性能的同时，也解决了生成对抗网络因陡峭的梯度空间导致训练不稳定问题。在生成对抗网络损失的基础上引入了SSIM损失函数，便于网络的训练。实验结果表明，本文算法保留了近红外图像更多的细节信息，生成了色彩饱和度更高的彩色结果图像。通过消融实验证明了空洞全局注意力模块和梯度归一化解决了近红外彩色化任务中出现的结果图像局部区域误着色及模糊问题。相比于其他彩色化算法，本文算法结果图在图像细节、边缘纹理及清晰度方面有较好的优势。

参考文献：

- [1] 戴康. 基于超像素提取和级联匹配的灰度图像自动彩色化[J]. 计算机与数字工程, 2019, 47(12): 3169-3172.
DAI K. Automatic colorization of grayscale images based on superpixel extraction and cascade matching[J]. *Computer and Digital Engineering*, 2019, 47(12): 3169-3172.
- [2] 曹丽琴, 商永星, 刘婷婷, 等. 局部自适应的灰度图像彩色化[J]. 中国图象图形学报, 2019, 24(8): 1249-1257.
CAO L Q, SHANG Y X, LIU T T, et al. Locally adaptive grayscale image colorization[J]. *Journal of Image and Graphics*, 2019, 24(8): 1249-1257.
- [3] Tomohisa Welsh, Michael Ashikhmin, Klaus Mueller. Transferring color to greyscale images[C]//ACM TOG, 2002, 21(3): 277-280.
- [4] Reinhard E, Adhikmin M, Gooch B, et al. Color transfer between images[J]. *IEEE Computer Graphics and Applications*, 2001, 21(5): 34-41.
- [5] 冯佳男, 江倩, 金鑫, 等. 基于深度神经网络的遥感图像彩色化方法[J]. 计算机辅助设计与图形学学报, 2021, 33(11): 1658-1667.
FENG J N, JIANG Q, JINX, et al. Colorization method of remote sensing image based on deep neural network[J]. *Journal of Computer-Aided Design and Computer Graphics*, 2021, 33(11): 1658-1667.
- [6] CHENG Z, YANG Q, SHENG B. Deep colorization[C]//Proceedings of the IEEE International Conference on Computer Vision, 2015: 415-423.
- [7] Isola P, ZHU J Y, ZHOU T, et al. Image-to-image translation with conditional adversarial networks[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition, 2017: 1125-1134.
- [8] Deblina Bhattacharjee, Seungryong Kim, Guillaume Vizier, et al. DUNIT: detection based unsupervised image-to-image translation[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020: 4787-4796.
- [9] 万园园, 王雨青, 张晓宁, 等. 结合全局语义优化的对抗性灰度图像彩色化[J]. 液晶与显示, 2021, 36(9): 1305-1313.
WAN Y Y, WANG Y Q, ZHANG X N, et al. Adversarial grayscale image colorization combined with global semantic optimization[J]. *Chinese Journal of Liquid Crystals and Displays*, 2021, 36(9): 1305-1313.
- [10] 左岑, 杨秀杰, 张捷, 等. 基于轻量级金字塔密集残差网络的红外图像超分辨率增强[J]. 红外技术, 2021, 43(3): 251-257.
ZUO Q, YANG X J, ZHANG J, et al. Super-resolution enhancement of infrared images based on lightweight pyramidal dense residual networks[J]. *Infrared Technology*, 2021, 43(3): 251-257.
- [11] 姜玉宁, 李劲华, 赵俊莉. 基于生成式对抗网络的图像超分辨率重建算法[J]. 计算机工程, 2021, 47(3): 249-255.
JIANG Y N, LI J H, ZHAO J L. Image super resolution reconstruction algorithm based on generative adversarial networks[J]. *Computer Engineering*, 2021, 47(3): 249-255.
- [12] 张振江, 张宝金, 刘伟新, 等. 基于深度卷积网络的矿岩图像分割算法研究[J]. 采矿技术, 2021, 21(5): 149-152, 171.

- ZHANG Z J, ZHANG B J, LIU W X, et al. Research on mining rock image segmentation algorithm based on deep convolutional network[J]. *Mining Technology*, 2021, **21**(5): 149-152, 171.
- [13] 姚永康. 基于对抗式迁移学习的皮肤病变图像分割方法研究[D]. 西安: 西京学院, 2021.
- YAO Y K. Research on Skin Lesion Image Segmentation Method Based on Adversarial Transfer Learning[D]. Xi'an: Xijing University, 2021.
- [14] 吴杰, 段锦, 董锁芹, 等. DFM-GAN 网络在跨年龄模拟的人脸识别技术研究[J]. *计算机工程与应用*, 2021, **57**(10): 117-124.
- WU J, DUAN J, TONG S Q, et al. DFM-GAN networks in cross-age simulation for face recognition[J]. *Computer Engineering and Applications*, 2021, **57**(10): 117-124.
- [15] 刘高天, 段锦, 范祺, 等. 基于改进 RFBNet 算法的遥感图像目标检测[J]. *吉林大学学报: 理学版*, 2021, **59**(5): 1188-1198.
- LIU G T, DUAN J, FAN Q, et al. Remote sensing image target detection based on improved RFBNet algorithm[J]. *Journal of Jilin University(Science Edition)*, 2021, **59**(5): 1188-1198.
- [16] LI C, WAN D M. Precomputed real-time texture synthesis with markovian generative adversarial networks[C]//2016 European Conference on Computer Vision of IEEE, 2016: 702-716.
- [17] LI Y, CHEN Y, WANG N, et al. Scale aware trident networks for object detection[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV), 2019: 6053-6062.
- [18] Woo S, Park J, Lee J Y, et al. CBAM: convolutional block attention module[C]//2018 European Conference on Computer Vision, 2018: 3-19.
- [19] Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift[C]//2015 Proceedings of the 32nd International Conference on International Conference on Machine Learning, 2015: 448-456.
- [20] Bhaskara V S, Aumentado-Armstrong T, Jepson A D, et al. GraN-GAN: piecewise gradient normalization for generative adversarial networks[C]//Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2022: 3821-3830.
- [21] Brown M, Süsstrunk S. Multi-spectral sift for scene category recognition[C]//2011 IEEE Conference on Computer Vision and Pattern Recognition, 2011: 177-184.
- [22] Jason Antic. jantic/deoldify: a deep learning based project for colorizing and restoring old images (and video!)[J/OL] [2019-10-16] <https://github.com/jantic/DeOldify>.
- [23] LIANG W, DING D, WEI G. An improved dual GAN for near infrared image colorization[J]. *Infrared Physics & Technology*, 2021, **116**(4): 103764-103777.
- [24] Perera P, Abavisani M, Patel V M. In2i: Unsupervised multi-image-to-image translation using generative adversarial networks[C]//2018 International Conference on Pattern Recognition, 2018: 140-146.