

结合语义分割的灰度图像彩色化

高晓宁 尹丽菊 王峰 邓玉林 秦怡鸣

Colorization of grayscale images combined with semantic segmentation

GAO Xiaoning, YIN Liju, WANG Feng, DENG Yulin, QIN Yiming

引用本文:

高晓宁, 尹丽菊, 王峰, 等. 结合语义分割的灰度图像彩色化[J]. 应用光学, 2025, 46(3): 632–642. DOI: 10.5768/JAO202546.0302003

GAO Xiaoning, YIN Liju, WANG Feng, et al. Colorization of grayscale images combined with semantic segmentation[J]. Journal of Applied Optics, 2025, 46(3): 632–642. DOI: 10.5768/JAO202546.0302003

在线阅读 View online: <https://doi.org/10.5768/JAO202546.0302003>

您可能感兴趣的其他文章

Articles you may be interested in

基于掩码一致性机制的弱监督图像语义分割研究

Weakly supervised image semantic segmentation based on masked consistency mechanism

应用光学. 2024, 45(4): 741–750 <https://doi.org/10.5768/JAO202445.0402003>

基于改进DeepLabv3+的无人车夜间红外图像语义分割

Semantic segmentation of nocturnal infrared images of unmanned vehicles based on improved DeepLabv3+

应用光学. 2020, 41(1): 180–185 <https://doi.org/10.5768/JAO202041.0106002>

基于改进Fast-SCNN的裂缝图像实时分割算法

Real-time segmentation algorithm of crack images based on improved Fast-SCNN

应用光学. 2023, 44(3): 539–547 <https://doi.org/10.5768/JAO202344.0302001>

基于图像分割和局部亮度调整的微光图像颜色传递算法

Low-light image color transfer algorithm based on image segmentation and local brightness adjustment

应用光学. 2020, 41(2): 309–317 <https://doi.org/10.5768/JAO202041.0202004>

基于注意力机制与图卷积神经网络的单目红外图像深度估计

Depth estimation of monocular infrared images based on attention mechanism and graph convolutional neural network

应用光学. 2021, 42(1): 49–56 <https://doi.org/10.5768/JAO202142.0102001>

基于改进BiSeNet的非结构化道路分割算法研究

Unstructured road segmentation algorithm based on improved BiSeNet

应用光学. 2023, 44(3): 556–564 <https://doi.org/10.5768/JAO202344.0302003>



关注微信公众号, 获得更多资讯信息

文章编号: 1002-2082 (2025) 03-0632-11

引用格式: 高晓宁, 尹丽菊, 王峰, 等. 结合语义分割的灰度图像彩色化 [J]. 应用光学, 2025, 46(3): 632-642.

GAO Xiaoning, YIN Liju, WANG Feng, et al. Colorization of grayscale images combined with semantic segmentation[J]. Journal of Applied Optics, 2025, 46(3): 632-642.



在线阅读

结合语义分割的灰度图像彩色化

高晓宁, 尹丽菊, 王 峰, 邓玉林, 秦怡鸣

(山东理工大学 电气与电子工程学院, 山东 淄博 255049)

摘 要: 图像彩色化是计算机视觉领域备受关注的核心问题之一, 近年来吸引了越来越多的研究者。彩色化技术提高了人眼对灰度图像, 特别是微光图像中目标的识别率和场景的理解力, 但目前彩色化方法仍然存在颜色溢出、颜色偏差和边界模糊等问题。为解决上述问题, 提出了一种结合语义分割的灰度图像彩色化方法。该方法设计了一个由分类子网络、语义分割子网络组成的彩色化网络, 将语义分割子网络和分类子网络提取的特征融合到彩色化网络中, 引导彩色化网络为图像中每个对象和背景赋予准确的颜色; 将结合空间和通道的注意力机制引入语义分割子网络, 获得更准确的语义分割结果, 提高彩色化网络对图像边界的感知能力和对细节的处理能力。为验证本文方法的有效性, 通过多像素光子计数器实验平台, 在低照度环境下拍摄了多组不同场景的微光图像, 并对其进行彩色化验证。结果表明, 所提方法在峰值信噪比 (peak signal-to-noise ratio, PSNR)、结构相似性 (structural similarity index measure, SSIM)、感知相似度 (learned perceptual image patch similarity, LPIPS) 上分别提升 6.12%、11.65%、4.99%。

关键词: 图像彩色化; 语义分割; 注意力机制; 微光图像

中图分类号: TN201; TP391.41

文献标志码: A

DOI: 10.5768/JAO202546.0302003

Colorization of grayscale images combined with semantic segmentation

GAO Xiaoning, YIN Liju, WANG Feng, DENG Yulin, QIN Yiming

(School of Electrical and Electronic Engineering, Shandong University of Technology, Zibo 255049, China)

Abstract: Image colorization is one of the core problems of great interest in computer vision and has attracted more and more researchers in recent years. The colorization technique improves the recognition rate and scene comprehension of targets of human eyes in grayscale images, especially in low-light-level (LLL) images. However, the current colorization methods still have problems such as color bleeding, color deviation and boundary blurring. To solve the above problems, a colorization method for grayscale images combined with semantic segmentation was proposed. A colorization network composed of a classification sub-network and a semantic segmentation sub-network was designed to fuse the features extracted from the semantic segmentation sub-network and the classification sub-network into the colorization network, and guide the colorization network to assign an accurate color to each object and background in the image. Then the attention mechanism combining with the space and channel was introduced into the semantic segmentation sub-network to obtain a more accurate semantic segmentation results, thereby improving the perception ability of image

收稿日期: 2024-04-22; 修回日期: 2024-08-03

基金项目: 国家自然科学基金 (62101310); 山东省自然科学基金 (ZR2020MF127)

作者简介: 高晓宁 (1998—), 女, 硕士研究生, 主要从事深度学习、图像处理研究。E-mail: gxn_18520302@163.com

通信作者: 尹丽菊 (1972—), 女, 博士, 教授, 硕士生导师, 主要从事光电检测与成像技术研究。E-mail: Ljyin72@163.com

boundaries and the processing ability of details of the colorization network. To verify the effectiveness of the proposed method, several sets of LLL images of different scenes were captured under LLL environment and colorized by multi-pixel photon counter (MPPC) experimental platform. The experimental results show that the peak signal-to-noise ratio (PSNR), structural similarity index measure (SSIM), and learned perceptual image patch similarity (LPIPS) of the proposed method is improved by 6.12%, 11.65%, and 4.99%, respectively.

Key words: image colorization; semantic segmentation; attention mechanism; low-light-level images

引言

图像彩色化是指为灰度图像的每一个像素预测合理颜色,它能够延伸人眼视觉感知能力,在当前图像处理中备受关注^[1-3]。灰度图像,尤其是微光图像^[4],在实际应用中受到了可视化效果和人眼获取图像信息的限制。这是因为在低光照环境下,传统的可见光摄像头在这种情况下无法正常工作,而多像素光子计数器 (multi-pixel photon counter, MPPC) 则能够准确地捕捉到低光照环境中的细节,不会造成欠曝的情况, MPPC 依赖于雪崩效应成像,能够在极微弱的环境中检测图像,在捕获微光图像方面具有很大优势。然而,由于 MPPC 捕获的微光图像通常是灰度图像,缺乏颜色信息,造成很多重要细节的缺失,这限制了其在实际应用中的可视化效果和人对图像信息的获取程度。因此,对微光图像进行彩色化具有重要的现实意义,尤其是在军事、生物医学、夜间自动驾驶和海洋探索等领域,彩色化技术能够将原来低对比度的微光图像转换为符合或近似人类视觉感知的彩色图像,显著改善图像质量,并提升人眼对图像中目标的识别能力,从而更好地理解场景的细节^[5-6]。这一技术对于修复老照片、老电影、多光谱图像、遥感图像,以及各种计算机视觉任务(如图像压缩、辅助图像分类和分割任务等)方面都具有重要的应用价值^[7-10],彩色化后的微光图像在航空航天、军事作战、军用以及民用等领域也都得到了广泛的应用。

图像彩色化技术分为传统彩色化方法和基于深度学习的彩色化方法。传统的彩色化方法依靠用户的干预,如局部颜色扩展^[11-12]或颜色传递^[13-15]的方法。这些方法需要大量人力参与,并且对参考图像要求较高。在处理复杂的图像结构和颜色时,传统彩色化方法的效果并不理想。目前,基于深度学习的彩色化方法利用大规模数据集对神经网络进行训练,使网络可以自主地学习图像的颜色分布和特征,逐渐取代了传统人工彩色化方法,

在彩色化技术中得到了广泛的应用。这种端到端的网络不受人为或其他因素的影响,实现了高效且易于操作的彩色化过程。现有的彩色化方法大多选择借助从灰度图像中再次提取特征作为辅助信息,改善彩色化网络的着色效果。CHENG Z 等人^[16]提出了第 1 个使用深度神经网络对灰度图像进行彩色化的方法,从不同图像中提取低级、中级、高级 3 种不同层次的特征作为网络预测色度通道的依据,再利用联合双边滤波进一步细化输出的色度值,从而减少神经网络引入的伪影;LIZUKA S 等人^[17]提出了双分支网络架构,在彩色化网络中融合局部特征和全局特征先验信息,提升了网络对图像整体语义信息的理解,提高了彩色化的效果;LARSSON G 等人^[18]利用分类任务预训练网络,用 CNN(convolutional neural networks)学习逐像素颜色直方图,通过 VGG(visual geometry group)网络来解释场景的语义组成和物体的定位,然后预测每个像素的颜色直方图;ZHANG R 等人^[19]将图像彩色化问题转换为了对像素点进行颜色分类的问题,提取局部和全局特征来预测颜色信息,该网络在标记数据集上进行分类和彩色化的联合训练,学习每个像素的颜色分布,这缓解了使用回归损失训练彩色化模型导致的图像颜色暗淡的问题。

另外,生成对抗网络 (generative adversarial network, GAN) 具有学习高位空间数据概率分布的能力,在图像生成、图像分割和风格转换等领域被广泛应用^[20]。受此启发, GAN 网络也被应用于图像彩色化。TRENESKA S 等人^[21]首次利用生成对抗网络,通过图像彩色化实现自监督视觉特征学习; ISOLA P 等人^[22]提出使用基于 U-Net 生成器的条件生成对抗网络 (conditional generative adversarial network, CGAN) 将输入图像映射到输出图像,并结合 $L1$ 损失和对抗损失对网络进行训练; CAO Y 等人^[23]也使用了生成对抗网络,通过对输入噪声进行多次采样,实现了彩色化结果的多样性; Deoldify^[24]使用 NoGAN 训练方法使网络能够生成

逼真的彩色图像; NAZERRI K 等人^[25]的彩色化网络通过在回归损失的基础上引入对抗损失,有效缓解了回归损失引起的颜色暗淡问题; VITORIA P 等人^[26]提出了一种将 GAN 与语义类分布学习相结合的端到端自动对抗方法。但这些基于 GAN 的方法都没有使用语义分割等附加信息,使得彩色化效果不理想。

理论上,灰度图像中不同物体的颜色在灰度级上有所差异,因此基于图像的灰度值,神经网络可以大致推测出物体的颜色。然而,在处理相似灰度的像素时,例如绿色的草地可能被彩色化成蓝色,蓝色的大海可能被彩色化为绿色,这会导致结果不准确。此外,在图像彩色化过程中,当物体多、颜色丰富时,容易出现颜色溢出和边界模糊等问题。因此,我们需要让神经网络具备一定的常识,使其能够辨别物体的边界位置,并在不同场景下准确判断物体应该呈现的真实颜色。针对上述问题,本文提出一种基于语义分割的灰度图像彩色化网络。设计了一种由分类子网络、语义分割子网络组成的彩色化网络,通过将语义分割子网络和分类子网络提取到的特征融合到彩色化网络中,更准确地获取图像中物体的语义和类别信息,使图像中物体边界更清晰,语义信息更准确;引入注意力机制,可以获得更准确的图像语义分割结果,提高对图像边界和细节的处理能力,同时增强图像的颜色特征,从而指导彩色化网络生成更加

接近真实图像的颜色分布。在低照度环境下,通过多像素光子计数器(MPPC)实验平台拍摄了多组不同场景下的微光图像,将本文方法应用于微光图像,改善图像质量。

1 语义分割的灰度图像彩色化方法

图1展示了本文提出的结合语义分割的灰度图像彩色化网络框架,该网络由生成器和鉴别器组成。生成器的网络设计参考了文献[17],主要包括3个部分:主干网络、语义分割子网络和分类子网络。其中,主干网络采用编码器-解码器结构,为了简化训练,将U-Net左侧的结构作为编码器,编码器在进行卷积的同时,额外拼接对应尺度的语义特征和分类特征,同时使用VGG-16提取全局特征,与编码器共享底层特征,这为彩色化网络提供了丰富的空间信息和细节信息,解码器根据语义分割信息和分类信息对灰度图像的每个像素点进行准确的颜色预测,有效解决由于对图像中物体的背景轮廓和类别无法正确判断所引起的颜色偏差和边缘模糊问题;分类子网络通过额外的卷积层和全连接层获取全局特征,在网络训练时,目标图像需归一化到 224×224 像素大小,以便网络能够固定全连接层大小进行特征融合和分类;语义分割子网络则专注于分割任务,使网络能够更好地处理图像细节,从而提高边缘清晰度和彩色化效果。鉴别器采用 70×70 的PatchGAN,用于评估生

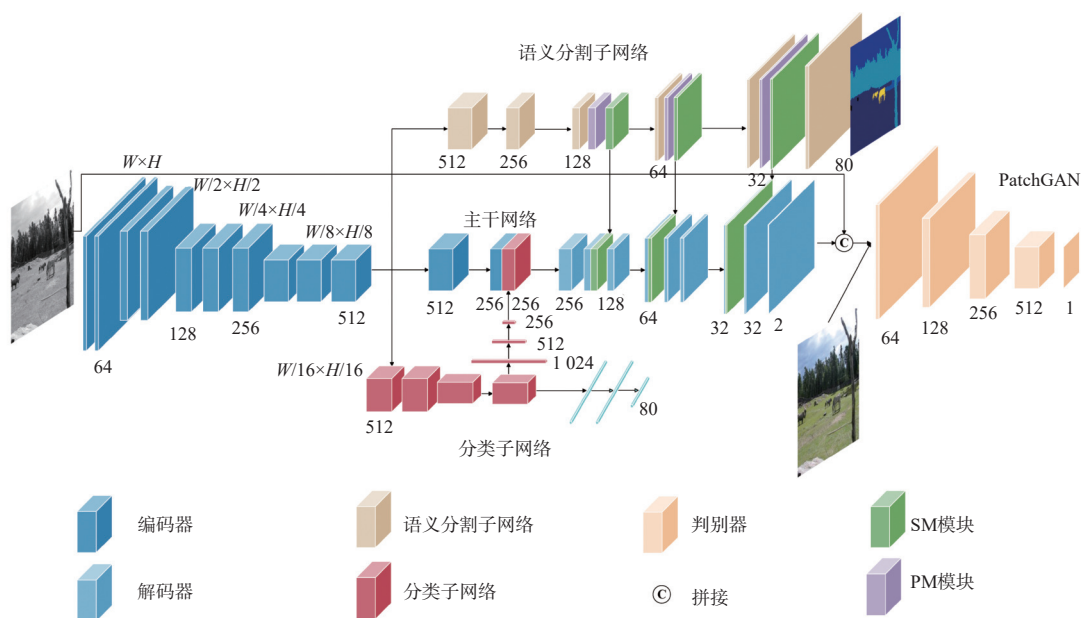


图1 本文彩色化方法的整体框架

Fig. 1 Overall framework of proposed colorization method

成彩色图像的真实性。生成器与鉴别器之间的对抗训练有助于网络学习从灰度图像到彩色图像的最佳映射。特别地,在图像彩色化任务中,使用亮度-色度颜色空间有助于分离图像的亮度信息和颜色信息。本文使用 CIELab 颜色空间,该空间是一种基于人类色彩感知的颜色空间。其中, L 分量表示感知亮度范围为 $[0,100]$, a 分量表示红色和绿色之间的颜色范围, b 分量表示蓝色和黄色之间的颜色范围,两个分量的范围均为 $[-128,127]$ 。在 Lab 颜色空间下,彩色化网络以 L 通道作为输入,仅预测 a 、 b 通道,最后将预测的 a 、 b 通道与 L 通道拼

接得到彩色图像。

1.1 语义分割子网络

为了进一步提升网络的彩色化效果,本文引入语义分割子网络来提取图像的语义信息。将分割网络提取到的特征与彩色化网络融合,以提高网络对图像中物体的边缘划分能力,帮助网络学习物体的颜色和轮廓,确保彩色化边界清晰准确。同时在语义分割网络中引入注意力机制,获取更准确的语义分割结果,从而辅助彩色化网络产生边界更加清晰和准确的彩色化效果。注意力机制的结构如图 2 所示。

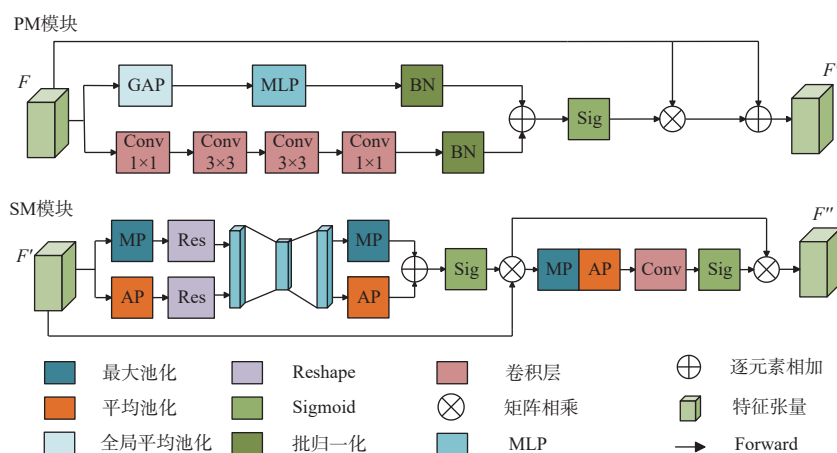


图 2 注意力机制结构

Fig. 2 Structure diagram of attention mechanism

注意力机制包括并行模块（parallel module, PM）和串行模块（serial module, SM）。在 PM 模块中,同时对通道和空间注意力进行计算。通道注意力分支利用通道间的关系,强调关注图像的边缘特征。空间注意力分支通过连续卷积获取一个大的感受野范围来有效利用图像的上下文信息,了解图像的空间位置。从两个注意分支中获取通道注意 $M_c(F)$ 和空间注意 $M_s(F)$ 后,将它们组合生成注意映射 F' 。这种组合使得图像颜色和位置能够精确对齐,进而减少了彩色化中的颜色偏差。PM 模块的映射函数如式(1)~式(4)所示:

$$M_c(F) = BN(MLP(AvgPool(F))) \quad (1)$$

$$M_s(F) = BN(f_3^{1 \times 1}(f_2^{3 \times 3}(f_1^{3 \times 3}(f_0^{1 \times 1} AvgPool(F)))) \quad (2)$$

$$M(F) = \sigma(M_c(F) + M_s(F)) \quad (3)$$

$$F' = F + F \otimes M(F) \quad (4)$$

式中: F 表示输入特征图; f 表示卷积操作,其上标表示卷积核的大小; BN 表示批量归一化层; $\otimes$ 表示

逐元素相乘; σ 表示 Sigmoid 激活函数。

在 SM 模块中,分别对通道维和空间维进行特征提取,通道注意力分支重点关注什么样的特征是有意义的,空间注意力分支则关注哪里特征是有意义的。最后将两个分支以串联的方式结合,精准地捕捉物体的边界和细节。SM 模块的映射函数如式(5)~式(7)所示:

$$M_{cc}(F') = \sigma(MLP(AvgPool(F')) \oplus MLP(MaxPool(F'))) \quad (5)$$

$$M_{ss}(F') = \sigma(f^{7 \times 7}([AvgPool(M_{cc}(F')); MaxPool(M_{ss}(F'))])) \quad (6)$$

$$F'' = M_{ss}(F') \otimes (M_{cc}(F') \otimes F') \quad (7)$$

式中: F' 表示 PM 模块的输出特征图,将其作为 SM 模块的输入特征图; F'' 表示 SM 模块的输出; $\oplus$ 表示逐元素相加; AvgPool、MaxPool 分别表示平均池化和最大池化。

1.2 分类子网络

本文构造了分类子网络,同时引入类别损失提

高着色准确性。分类子网络旨在获取图像的全局特征和类别标签信息。具体地,将图1中分类子网络的4个卷积层的输出分别处理。一方面,分类子网络将卷积结果解析为图片类别概率分布 $\hat{Y} \in R^{n \times 80 \times 1 \times 1}$,通过全连接层获得一个80维向量,表示数据集中的80个分类标签,如斑马、橘子、羊、卧室等不同物体类别,这些详细的分类标签提高了图像彩色化的准确性;另一方面,将卷积结果通过另外3个全连接层后得到一个256维的全局特征,然后将256维向量复制 $W/8 \times H/8$ 次后与主干网络卷积层的输出特征图结合,用于引导颜色正确分类。

1.3 损失函数

在生成对抗网络中,生成器的目标是生成与真实图像数据分布尽可能接近的图像,判别器评估生成图像与真实图像之间的差距。根据网络结构的描述,本文定义了一个新的损失函数,该损失函数包括像素损失、分类损失、分割损失和鉴别器损失,总的损失函数可以表示为

$$L(G_\theta, D) = L_r(G_{\theta_1}^1) + \lambda_{cls} L_{cls}(G_{\theta_2}^2) + \lambda_{seg} L_{seg} + \lambda_g L(G_{\theta_1}^1, D) \quad (8)$$

式中:前3项为生成器损失; $G_\theta = (G_{\theta_1}^1, G_{\theta_2}^2)$, θ_1 、 θ_2 为生成器的参数。

$L_r(G_{\theta_1}^1)$ 像素损失的计算式为

$$L_r(G_{\theta_1}^1) = E_{(L, a_r, b_r) \sim p_r} [\|G_{\theta_1}^1(L) - a_r, b_r\|_2^2] \quad (9)$$

式中: (L, a_r, b_r) 为彩色图像在Lab空间中的表示; $\|\cdot\|_2$ 为欧几里得范数; p_r 为真实彩色图像的分布。Lab色彩空间中的欧几里得距离更适合于感知色彩。

通过分类损失辅助网络更好地进行彩色化任务。分类网络的损失定义如式(10):

$$L_{cls}(G_{\theta_2}^2) = E_{L \sim p_{rg}} [\text{KL}(y_v \| G_{\theta_2}^2(L))] \quad (10)$$

式中: p_{rg} 为输入灰度图像的分布; $y_v \in R^m$ 是对数据集中图像进行分类得到的分类向量; m 为图像分类的次数。KL散度计算公式如式(11):

$$D_{KL}(p \| q) = \sum_{i=1}^n p(x_i) \log \left(\frac{p(x_i)}{q(x_i)} \right) = \sum_{i=1}^n p(x_i) \log(p(x_i)) - \sum_{i=1}^n p(x_i) \log(q(x_i)) \quad (11)$$

式中前一部分是 p 的熵,后一部分是 p 、 q 的交叉熵。同样地,在模型训练过程中,最小化真实数据与生成数据分布之间的KL散度,从而使得生成数据尽可能接近真实数据分布。

在网络中引入语义分割子网络可以通过每层的语义特征引导解码器着色,同时引入分割损失辅助编码器解析更加全面的全局特征和局部特征。语义分割的损失函数定义如式(12):

$$L_{seg} = - \sum_c p_i \log(q_i) \quad (12)$$

式中: C 代表语义分割的类别数; p 代表真实图像的语义分布; q 代表语义分割子网络的输出结果。

判别器损失使用ARJOVSKY等人提出的WGAN方法对真实图像与生成图像之间的差距进行度量,避免了传统GAN中常见的梯度消失与梯度爆炸问题。WGAN方法应用线性插值方法在生成图像和真实图像分布之间构建新的数据分布,对每个输入数据进行梯度惩罚(gradient penalty, GP),通过引入梯度惩罚项,WGAN能有效地约束判别器,提升网络模型的收敛速度和稳定性,最终实现更稳定的训练。鉴别器损失定义如式(13):

$$L(G_{\theta_1}^1, D) = E_{\tilde{x} \sim p_g} [D(\tilde{x})] - E_{x \sim p_r} [D(x)] + \lambda_{GP} E_{\tilde{x} \sim p_{\tilde{x}}} [(\|\nabla_{\tilde{x}} D(\tilde{x})\|_2 - 1)^2] \quad (13)$$

式中: λ_{GP} 是梯度惩罚项的权重,取值为10; $\tilde{x}$ 为使用真实分布和生成分布构造的新图像。在训练本文的彩色化模型时,使用Adam优化器来最小化损失函数。在训练过程中,本文将式(8)中的参数 λ_{cls} 、 λ_{seg} 、 λ_g 分别设置为0.003、0.003、0.1。

2 实验结果与分析

2.1 数据集

COCO数据集^[27]是一个大型、丰富的物体检测、分割数据集,包含80类语义信息(人类生活的常见物体与场景),训练集有118287张彩色图像,测试集采用COCO验证集的5000张彩色图像测试彩色化效果,同时测试集还包括在低照度环境下,通过多像素光子计数器(MPPC)实验平台拍摄的多组不同场景下的微光图像,并将微光图像调整为 256×256 像素大小。

2.2 实验参数

在实验预处理阶段,本文将训练数据集中的图像尺寸均调整为 224×224 像素大小,并将图片从RGB颜色空间转化为Lab颜色空间,提取亮度通道 L 作为网络的输入通道,同时将亮度通道 L 复制3次,以生成三通道灰度图像。在测试阶段,交替训练生成器 G 和判别器 D ,共训练25个epoch,

批次大小设置为 16, 使用 Adam 优化器^[28] 最小化目标损失, 学习率设为 2×10^{-5} , 动量参数 $\beta_1=0.5$, $\beta_2=0.999$ 。

2.3 评价指标

现有的图像彩色化评价指标主要分为主观评价和客观评价。主观评价即为人类根据个人的主观感受和审美标准对生成的彩色图像进行评价。客观评价指标包括峰值信噪比 (peak signal-to-noise ratio, PSNR)、结构相似性 (structural similarity index measure, SSIM)、感知相似度 (learned perceptual image patch similarity, LPIPS) 和信息熵 (information entropy, IE)^[29-30]。PSNR 和 SSIM 是图像质量评价中常用的两个指标。PSNR 是峰值信噪比, 通常用于评估图像压缩前后的质量, PSNR 值越高, 生成的图像质量越好; SSIM 是结构相似度指数, 用于衡量两张图像之间相似度; LPIPS 是一种感知质量量化评价指标, 可在视觉预测等任务中提供良好的感知评估, LPIPS 值越低, 生成图像与原始图像之间的感知差异越小; IE 是信息熵, 可客观评价彩色化中信息的丰富程度, IE 值越高, 图像包含的信息越丰富。

2.4 不同算法实验结果对比

为了验证本文方法的有效性, 将本文方法与 5 种先进的彩色化方法 (LIZUKA S 等人^[17]、LARSSON G 等人^[18]、ANTIC J 等人^[24]、KANG X 等人^[31]、VITORIA P 等人^[26]) 进行比较, 并从主观和客观两个角度评估本文方法的彩色化效果。

本文方法与其他 5 种方法的对比实验结果如图 3 所示。通过对比这些方法的图像彩色化效果可以看出, 第 1 种方法在训练神经网络时采用了回归损失函数, 导致生成的彩色化结果偏棕色且不饱和; 第 2 种方法处理虽然较第 1 种有所改善, 但仍存在物体颜色偏浅和整体偏棕色的问题; 第 3 种和第 4 种方法的彩色化结果存在明显的颜色溢出; 第 5 种方法处理的图像也存在轻微的颜色溢出; 相比之下, 本文方法使用分割损失和对抗损失训练神经网络, 能够生成更加鲜艳的彩色效果, 同时通过利用图像本身的高级语义分割信息引导网络对图像进行准确着色, 使彩色化图像中的物体边界清晰, 减少颜色溢出。本文方法在所有对比方法中表现最好, 适用于不同场景下的自然图像; 另一方面, 无论是低复杂度还是高复杂度的图像, 本文



图 3 不同彩色化方法的效果对比

Fig. 3 Comparison of effects of different colorization methods

方法均能更好地逼近真实图像,表明网络具有较好的鲁棒性。

为了客观地对比每种方法的彩色化效果,本文使用 PSNR、SSIM、LPIPS 和 IE 4 个指标进行定量分析。表 1 和表 2 给出了本文与其他方法在 4 个评价指标上的结果。最优结果用加粗与箭头来表示,表格中加粗数据后的箭头向上表示数值最大,箭头向下表示数值最小。从表中可以看出,本文方法在这些指标上有更好的表现。其中,由于存在少部分图像的 PSNR 值没有达到最优,PSNR 提升不大,因此引入 SSIM 评价图像的结构相似性和

生成图像的失真程度,弥补 PSNR 无法完全反映图像质量与人眼视觉效果一致性的缺陷。本文方法的 SSIM 指标明显优于其他方法,生成的图像更接近真实图像;同时信息熵值也较其他方法更高,这表示生成图像的颜色分布更加多样,具有更丰富的颜色表现能力,相反,较低的信息熵值意味着生成图像的颜色分布较为集中,缺乏多样性。因此,无论是从主观视觉还是客观指标来看,本文方法都表现出了明显的优势,能在保持图像细节的同时生成丰富多样的颜色,且生成的彩色图像更加令人信服。

表 1 不同彩色化方法的图像质量指标(PSNR 和 SSIM)对比

Table 1 Comparison of image quality metrics (PSNR and SSIM) for different colorization methods

图像	LIZUKA S等人		LARSSON G等人		ANTIC J等人		KANG X等人		VITORIA P等人		本文方法	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
orange	17.868	0.778	20.027	0.751	20.803	0.767	20.335	0.608	22.482	0.781	23.231↑	0.829↑
grassland	21.166	0.899	21.789	0.947	27.990↑	0.959	26.95	0.960	26.822	0.895	27.145	0.963↑
desktop	16.555	0.808	17.589	0.811	17.772	0.767	17.634	0.787	19.734	0.816	20.818↑	0.891↑
sheep	27.118	0.918	25.506	0.932	27.001	0.948	26.044	0.901	29.985	0.952	30.343↑	0.960↑
pavement	26.429	0.908	23.613	0.940	27.681	0.953	27.504	0.895	30.135	0.965	32.386↑	0.973↑
zebra	19.586	0.894	21.227	0.951	23.729	0.895	27.595	0.967↑	28.583↑	0.959	27.372	0.964

表 2 不同彩色化方法的图像质量指标(LPIPS 和 IE)对比

Table 2 Comparison of image quality metrics (LPIPS and IE) for different colorization methods

图像	LIZUKA S等人		LARSSON G等人		ANTIC J等人		KANG X等人		VITORIA P等人		本文方法	
	LPIPS	IE	LPIPS	IE	LPIPS	IE	LPIPS	IE	LPIPS	IE	LPIPS	IE
orange	0.152	7.134	0.188	6.979	0.169	7.248	0.190	6.873	0.126	7.286↑	0.106↓	6.922
grassland	0.074	7.218	0.057	7.269	0.065	7.285	0.069	7.144	0.052	7.237	0.051↓	7.305↑
desktop	0.139	7.428	0.115	7.372	0.101	7.191	0.106	7.459	0.129	7.502	0.079↓	7.515↑
sheep	0.063	7.635	0.096	7.625	0.059↓	7.667	0.095	7.619	0.062	7.671↑	0.060	7.641
pavement	0.156	7.682	0.162	7.676	0.073	7.722	0.123	7.671	0.075	7.639	0.051↓	7.694↑
zebra	0.093	7.215	0.067	7.189	0.066	7.173	0.079	7.222	0.061	7.261	0.059↓	7.287↑

2.5 消融实验

为了进一步验证所提算法的性能,本文设置了 3 组消融实验来测试每个模块对网络性能的影响。本文设置了 4 种不同的网络结构来分析语义分割子网络、分类子网络对网络性能的影响: 1) Baseline; 2) +分割子网络; 3) +分类子网络; 4) 本文方法。

图 4 展示了不同网络结构的彩色化效果图和

对应的 LPIPS 映射图。映射图中颜色的变化表示 LPIPS 值的变化,同时也反映了生成图像与真实图像之间的感知相似度(颜色越接近深紫色,对应的 LPIPS 值越低,生成图像与真实图像之间的差异越小;颜色越接近黄色,对应的 LPIPS 值越高,生成图像与真实图像之间的差异越大)。通过可视化 LPIPS,可以直观地了解生成图像与真实图像之间在哪些特征或结构上存在显著差异。

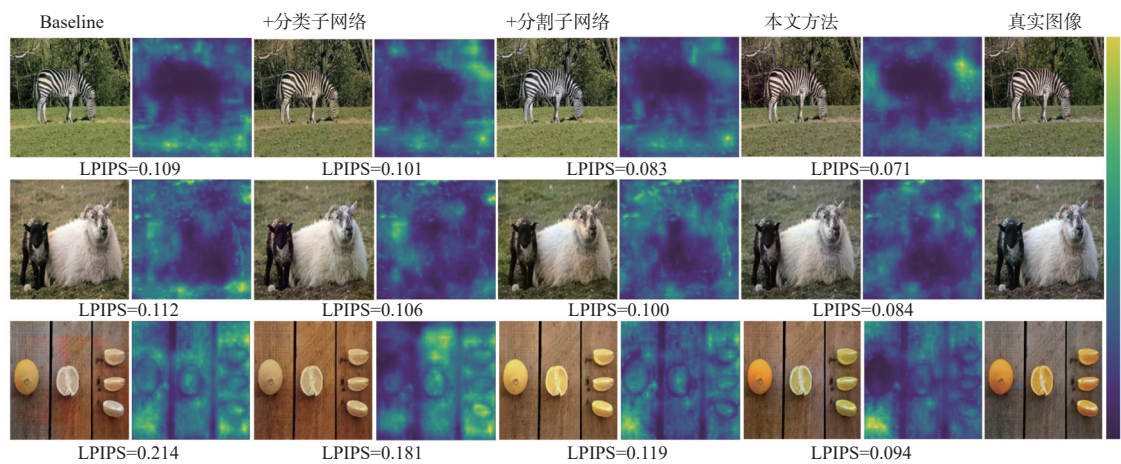


图4 不同网络结构的彩色化效果对比

Fig. 4 Comparison of colorization effects of different network structures

引入语义分割子网络增强了网络对图像中物体边界的识别能力,使得彩色化后的图像边界更加清晰,着色误差显著降低;引入分类子网络会使网络产生更精准、更自然的彩色化效果。表3中展示了不同网络结构的客观结果。可以看出,与Baseline相比,引入语义分割子网络和分类子网络后 LPIPS 值减小, SSIM、PSNR 值均有所增加,说明本文方法中添加的两个子网络对彩色化效果有明显提升。

表3 不同网络结构的图像质量指标

Table 3 Image quality metrics for experiments with different network structures

方法	PSNR	SSIM	LPIPS	IE
Baseline	22.731	0.894	0.138	7.422
+分类子网络	25.042	0.922	0.133	7.415
+分割子网络	25.949	0.931	0.128	7.494
本文方法	26.793↑	0.938↑	0.073↓	7.579↑

其次,本文在分割子网络中引入了注意力机制,以增强分割网络对语义信息的提取能力,并辅助提升网络的彩色化效果。为了进一步验证所提注意力机制对彩色化网络的有效性,本节设置了相关的对比实验,在分割网络中将所提注意力机制层替换为正常的卷积层与本文方法进行对比,以此说明加入注意力机制的彩色化效果更好,实验结果如图5所示。

从图5可以看出,加入注意力机制后的彩色化效果在细节方面有显著提升。本文的注意力机制将通道和空间位置注意力进行并行和串行的组合。串行的连接方式先提取颜色特征,然后学习

图像位置信息;而并行的连接方式同时学习图像的颜色和位置信息的非线性关系。这样的设计可使网络学习如何将颜色信息与位置信息对齐,同时并行和串行的组合提高了网络的彩色化准确率,使生成的彩色图像更加接近真实图像。此外,如表4所示,引入注意力机制使网络在大多数指标上都表现良好。

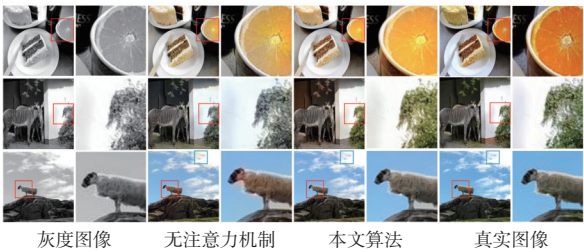


图5 分割网络有无注意力机制的彩色化效果对比

Fig. 5 Comparison of colorization effects of segmented networks with and without attention mechanisms

表4 有无注意力机制实验的图像质量指标

Table 4 Image quality metrics for experiments with and without attention mechanisms

方法	PSNR	SSIM	LPIPS	IE
无注意力机制	25.652	0.915	0.132	7.458
本文方法	26.793↑	0.938↑	0.125↓	7.579↑

为了验证所提方法中损失函数的有效性,本文对分割损失(L_{seg})和分类损失(L_{cls})进行了实验对比,基线网络是没有分割损失和分类损失约束的网络,对比实验在基线网络上分别加入分类损失和分割损失,并与本文方法进行对比,实验结果如图6和表5所示。

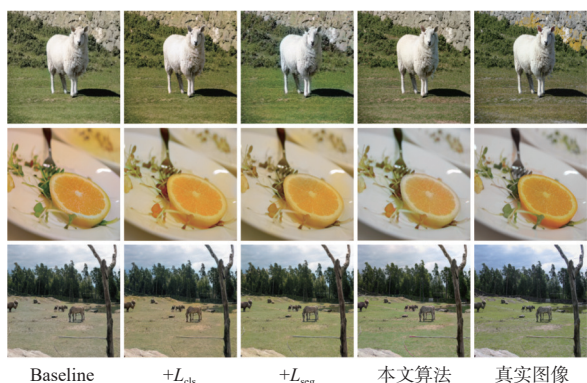


图 6 不同损失函数的彩色化效果对比

Fig. 6 Comparison of colorization effects with different loss functions

表 5 不同损失函数实验的图像质量指标

Table 5 Image quality metrics for experiments with different loss functions

方法	PSNR	SSIM	LPIPS	IE
Baseline	25.853	0.898	0.088	7.136
+ L_{cls}	26.343	0.905	0.084	7.242
+ L_{seg}	26.291	0.913	0.079	7.307
本文方法	26.793↑	0.938↑	0.073↓	7.579↑

从图 6 可以看出, 加入分割损失可帮助网络更好地理解图像的结构和边界信息, 减少了彩色化结果中的颜色溢出, 从而提高生成图像与真实图像的结构相似性, 使得彩色化结果更加自然, 其对应的定量指标 PSNR、SSIM 和 IE 相较于 Baseline 网络均有提升, LPIPS 指标也有明显下降。通过分类损失, 网络可以学习识别图像中的不同类别, 有助于指导网络正确地为图像中的不同物体赋予合适的颜色, 进而提升彩色化的质量和准确度。添加损失函数后的彩色化结果在客观评价指标和主



图 8 不同彩色化方法的微光图像效果对比

Fig. 8 Comparison of LLL image effects of different colorization methods

观感受上均表现优异。

2.6 微光图像

为了验证所提方法在实际应用中的先进性, 用本文所提方法对微光图像进行彩色化处理。在低照度环境下, 通过多像素光子计数器 (MPPC) 实验平台拍摄了多组不同场景下的微光图像, 同时调整图像大小为 256×256 像素。该微光实验平台主要由 MPPC 探测器、步进电机控制器、二维电控导轨、宽量程微光照度计、计算机、光纤以及连接各器件的电缆线等组成, 实验平台实物图如图 7 所示。

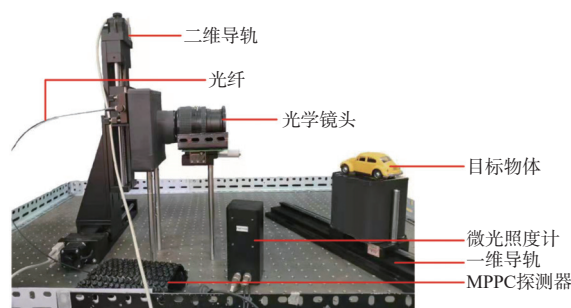


图 7 多像素光子计数器实验平台实物图

Fig. 7 Physical drawing of multi-pixel photon counter experimental platform

本文在微光图像上进行验证, 并且选择几种表现优异的算法进行对比实验。通常, 为了取得较好的微光图像彩色化效果, 网络模型训练时应该采用微光图像和对应彩色图像, 但由于条件限制, 本文采用公开数据集作为训练集训练网络模型。虽然微光图像也属于灰度图像, 但和数据集中的灰度图像有一定区别, 如果用数据集中的灰度图像训练得到的神经网络模型能够实现微光图像的彩色化, 这更能证明本文方法的先进性。微光图像彩色化效果如图 8 所示。

从图 8 中可以看出, 经过彩色化后的微光图像不仅在视觉感知上更加舒适, 同时图像所含信息也更易于辨认, 显著提高了人眼对目标的分辨能力。彩色化结果表明本文提出的彩色化网络具有良好的适应性和较高的实用价值。此外, 表 6 展示了微光图像彩色化在 4 个评价指标上的数值结果, 可以看出, 本文方法在所有评价指标上均取得了最优结果。这些指标的提升与分割子网络和分类子网络有关, 它们能够引导彩色化网络更好地识别物体的类别和边界信息, 有效提高目标探测和识别的效率。

表 6 不同彩色化方法对微光图像的质量指标

Table 6 Quality metrics of LLL images by different colorization methods

方法	PSNR	SSIM	LPIPS	IE
LARSSON G等人	10.847	0.207	0.747	7.332
ANTIC J等人	10.550	0.206	0.739	7.332
VITORIA P等人	10.762	0.212	0.737	7.337
本文方法	11.194↑	0.236↑	0.709↓	7.408↑

3 结论

本文提出一种结合语义分割的灰度图像彩色化网络, 引入语义分割子网络引导网络学习如何正确地对图像进行分割, 引入分类子网络引导网络学习区分不同类别的颜色分布和特征, 从而实现准确、饱和、自然、连贯的彩色化效果, 分割损失函数和分类损失使生成数据尽可能接近真实数据分布, 提升了图像整体的彩色化效果。实验表明, 本文提出的灰度图像彩色化方法在提升图像质量和保留图像结构信息方面表现更优异。同时, 本文方法也可以使生成的彩色微光图像更符合人眼的视觉感知。相比于以往的方法, 本文方法在避免颜色溢出、颜色偏差和边界模糊问题方面具有较高的优势。

参考文献:

- [1] 谯帅, 孙韶媛, 谷小婧, 等. 基于轮廓跟踪的车载红外视频彩色化方法[J]. 应用光学, 2012, 33(4): 727-732.
QIAO Shuai, SUN Shaoyuan, GU Xiaojing, et al. Colorization method of in-vehicle infrared video based on contour tracking[J]. Applied Optics, 2012, 33(4): 727-732.
- [2] 杨晨, 张智勇, 王继红, 等. 红外伪彩色图像的去彩化研

究[J]. 应用光学, 2015, 36(3): 403-407.

YANG Chen, ZHANG Zhiyong, WANG Jihong, et al. Research on decolorization of infrared pseudo-color images[J]. Applied Optics, 2015, 36(3): 403-407.

- [3] 吴强, 王新赛, 王卫平, 等. 基于 LLE 的红外成像彩色化处理算法研究[J]. 应用光学, 2012, 33(4): 733-737.

WU Qiang, WANG Xinsai, WANG Weiping, et al. Research on colorization processing algorithm of infrared imaging based on LLE[J]. Applied Optics, 2012, 33(4): 733-737.

- [4] 陈一超, 胡文刚, 武东生, 等. 三波段微光彩色夜视方法研究[J]. 应用光学, 2015, 36(3): 430-434.

CHEN Yichao, HU Wengang, WU Dongsheng, et al. Research on three-waveband microlight color night vision method[J]. Applied Optics, 2015, 36(3): 430-434.

- [5] FLOREA L, FLOREA C, IONASCU C. Avoiding the deconvolution: framework oriented color transfer for enhancing low-light images[J]. IEEE Computer Society, 2016: 27-35.

- [6] HOGERVORST M A, TOET A. Fast natural color mapping for night-time imagery[J]. Information Fusion, 2010, 11(2): 69-77.

- [7] KHAN M U G, GOTOH Y, NIDA N. Medical image colorization for better visualization and segmentation[C]//Annual Conference on Medical Image Understanding and Analysis, MIUA 2017. Cham: Springer, 2017: 571-580.

- [8] LIU L, JIANG Q, JIN X, et al. CASR-Net: a color-aware super-resolution network for panchromatic image[J]. Engineering Applications of Artificial Intelligence, 2022, 114: 105084.

- [9] 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 Jianan, JIANG Qian, JIN Xin, et al. Remote sensing image colorization method based on deep neural network[J]. Journal of Computer-Aided Design & Computer Graphics, 2021, 33(11): 1658-1667.

- [10] DESHPANDE A, LU J, YEH M C, et al. Learning diverse Image Colorization[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). New York: IEEE. 2017: 2877-2885.

- [11] HORIUCHI T, HIRANO S. Colorization algorithm for grayscale image by propagating seed pixels[C]//2003 IEEE International Conference on Image Processing (ICIP). New York: IEEE, 2003: 03CH37429.

- [12] LEVIN A, LISCHINSKI D, WEISS Y. Colorization using optimization[J]. *ACM Transactions on Graphics*, 2004, 23(3): 689-694.
- [13] WU Z, XUE R. Color transfer with salient features mapping via attention maps between images[J]. *IEEE Access*, 2020, 8: 104884-104892.
- [14] 辛浪, 刘钧, 袁渊. 基于图像分割和局部亮度调整的微光图像颜色传递算法[J]. *应用光学*, 2020, 41(2): 309-317.
XIN Lang, LIU Jun, YUAN Yuan. A color transfer algorithm for ember images based on image segmentation and local brightness adjustment[J]. *Applied Optics*, 2020, 41(2): 309-317.
- [15] 仲红玉, 尹丽菊, 高明亮, 等. 多参数亮度值重映射的颜色传递方法[J]. *红外与激光工程*, 2018, 47(5): 159-166.
ZHONG Hongyu, YIN Liju, GAO Mingliang, et al. A colour transfer method for remapping multiparameter luminance values[J]. *Infrared and Laser Engineering*, 2018, 47(5): 159-166.
- [16] CHENG Z, YANG Q, SHENG B. Deep colorization[C]// 2015 IEEE International Conference on Computer Vision (ICCV). New York: IEEE, 2015: 415-423.
- [17] LIZUKA 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): 1-11.
- [18] LARSSON G, MAIRE M, SHAKHNAROVICH G. Learning representations for automatic colorization[C]// European Conference on Computer Vision. Cham: Springer International Publishing, 2016: 577-593.
- [19] ZHANG R, ISOLA P, EFROS A A. Colorful image colorization[J]. Cham: Springer, 2016: 649-666.
- [20] 余思泉, 韩志, 唐延东, 等. 基于对抗生成网络的纹理合成方法[J]. *红外与激光工程*, 2018, 47(2): 43-48.
YU Siqun, HAN Zhi, TANG Yandong, et al. Texture synthesis method based on adversarial generative network[J]. *Infrared and Laser Engineering*, 2018, 47(2): 43-48.
- [21] TRENESKA S, ZDRAVEVSKI E, PIRES P G S. GAN-based image colorization for self-supervised visual feature learning[J]. *Sensors*, 2022, 22(4): 1599-1615.
- [22] 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 (CVPR). New York: IEEE, 2017: 5967-5976.
- [23] CAO Y, ZHOU Z, ZHANG W, et al. Unsupervised diverse colorization via generative adversarial networks[C]// ECML/PKDD 2017. Cham: Springer, 2017: 151-166.
- [24] ANTIC J. DeOldify [EB/OL]. (2018-10-31) [2021-12-10]. <https://github.com/jantic/DeOldify>.
- [25] NAZERRI K, NG E, EBRAHIMI M. Image colorization using generative adversarial networks[C]// 2020 International Conference on Machine Vision and Image Processing (MVIP). Cham: Springer, 2018: 85-94.
- [26] VITORIA P, RAAD L, BALLESTER C. ChromaGAN: adversarial picture colorization with semantic class distribution[C]// Workshop on Applications of Computer Vision. New York: IEEE, 2020: 2445-2454.
- [27] LIN T Y, MAIRE M, BELONGIE S, et al. Microsoft COCO: common objects in context[C]// 2014 IEEE International Conference on Computer Vision (ICCV). New York: IEEE, 2014: 740-755.
- [28] KINGMA D, BA J. Adam: a method for stochastic optimization[C]// International Conference on Learning Representations. New York: IEEE, 2015: 6980-6994.
- [29] HORE A, ZIOU D. Image quality metrics: PSNR vs. SSIM[C]// 2010 20th International Conference on Pattern Recognition. Istanbul: IEEE, 2010: 2366-2369.
- [30] ZHANG R, ISOLA P, EFROS A A, et al. The unreasonable effectiveness of deep features as a perceptual metric [C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington DC: IEEE Computer Society, 2018: 586-595.
- [31] KANG X, YANG T, OUYANG W, et al. Ddcolor: towards photo-realistic image colorization via dual decoders[C]// Proceedings of IEEE/CVF International Conference on Computer Vision. New York: IEEE, 2023: 328-338.