

doi: 10.12052/gdutxb.220039

基于生成对抗网络的低光照图像增强算法

杨镇雄, 谭台哲

(广东工业大学 计算机学院, 广东 广州 510006)

摘要: 传统的基于深度学习的方法在低照度图像增强中已经有比较好的发挥, 但是这些方法通常需要成对的数据集进行训练, 而相对应的数据集正是目前难以收集的。目前的增强方法在真实的低照度图像增强中也会产生增强效果不完美和出现图像噪声等问题。针对这些问题, 设计了无监督生成对抗网络, 使其可以不用配对训练数据集进行训练, 并且把网络分解为注意力机制网络和增强网络2个子网络。通过注意力机制网络把低照度图像中的低光区域和亮光区域区分开, 并使用残差增强网络结合全局局部判别器, 对图像进行增强。实验结果表明, 本文的方法在低光照图像增强方面优于Enlighten-GAN、Cycle-GAN等方法。

关键词: 生成对抗网络; 低照度图像增强; 注意力机制

中图分类号: TP391

文献标志码: A

文章编号: 1007-7162(2024)01-0055-08

Low Illumination Image Enhancement Algorithm Based on Generative Adversarial Network

Yang Zhen-xiong, Tan Tai-zhe

(School of Computer Science and Technology, Guangdong University of Technology, Guangzhou 510006, China)

Abstract: Traditional deep learning-based methods have achieved promising performance for low-illumination image enhancement. However, these methods usually need to be trained on the pair-wise datasets, which are difficult to collect. Moreover, most existing enhancement methods have the problems of imperfect enhancement effect and image noise in real low illumination image enhancement. To address this, a unsupervised generative adversarial network is designed for low-illumination image enhancement, which has no requirement of training on the pair-wise datasets. The proposed network consists of two subnetworks: attentional mechanism network and enhancement network. The attentional mechanism network is used to distinguish the low-light region from the bright region of the low-illumination image, and the residual enhancement network is used to enhance the image by combining with the global-local discriminator. By doing this, a low-illumination image can be well enhanced. Extensive experimental results show that the proposed method outperforms the baseline Enlighten-GAN and Cycle-GAN for low-light image enhancement.

Key words: generative adversarial network; low-light image enhancement; attentional mechanism

现阶段, 图像处理技术在各个领域都有广泛的应用, 而图像增强是其中的一个重要方向。在不同情况下获取图像时, 环境因素会对图像的质量造成或多或少的影响。而图像增强则是处理这些图像, 使图像中的细节增加, 去除图像中因为环境而出现的噪声, 提高图像对比度, 从而提高图像的整体质量, 使其和现实中的真实场景靠拢。特别是其中的低照度

图像, 在监控、半晚拍照等情况下, 由于光线不足, 形成图像的亮度严重不够, 图像中的许多细节缺失^[1], 所以低照度图像增强方法随之产生。随着时间推移, 越来越多的研究者开始尝试通过深度学习解决图像增强任务^[2-3]。

目前基于深度学习的图像增强任务有LL-Net(Low Light Network)^[4], 通过增强和去噪2个阶段

收稿日期: 2022-03-02

作者简介: 杨镇雄(1997-), 男, 硕士研究生, 主要研究方向为图像处理与计算机视觉, E-mail: yangzx613@foxmail.com

通信作者: 谭台哲(1970-), 男, 副教授, 博士, 主要研究方向为机器学习与大数据处理、图像处理与计算机视觉、区块链技术, E-mail: taizhetan@gdut.edu.cn

对图像进行增强; MSR-Net(Multi-Scale Retinex Network)^[5]引入卷积神经网络(Convolutional Neural Networks, CNN),仿照多尺度的Retinex理论(Multi-Scale Retinex, MSR),分为多尺度对数变换、卷积差分、颜色恢复3个模块; Retinex-Net^[6]参考Retinex理论,其分为2个子网络,解耦网络对输入图像进行解耦得到光照图和反射图,增强网络再对光照图进行增强处理,增强后再与反射图结合得到结果图等。但基于深度学习的图像恢复和增强方法主要依赖于配对的数据集。

但是,配对的数据集在获取方面比较大的困难,在很多场景中,因为现实光照增强不受控制,想要获得相对应的正常光图像和低光图像几乎很难实现^[6-7],所以现有大部分的低照度图像增强方法基本都是使用合成的配对数据集进行训练。例如,LL-Net^[4]使用伽马校正和添加高斯噪声来模拟低光环境,并通过2个阶段把图像进行增强和去噪; LOL数据集中的低光照图像则是用Adobe Lightroom接口调节得到的^[6]。但合成的配对数据集和正常场景相对比就显得不够真实,通过合成数据集训练的模型在现实场景中的表现难以让人满意^[2,8]。

因为配对数据集难以获得,为了解决这个问题,在其他图像生成领域里已经有人通过无监督学习来对图像进行转化增强^[9-10],而这些方法大多是基于生成对抗网络(Generative Adversarial Networks, GANs),在对抗中生成图像和目标图像逐渐靠近,从而达到不依赖于相对应配对的图像数据集的目的,也可以在低照度图像和正常亮度图像之间转换。

目前这些无监督生成对抗模型在图像增强任务上可以生成对比度比较好的图像^[11-14],循环生成对抗网络(Cycle-Consistent Generative Adversarial Networks, Cycle-GAN)^[15]在图像转化、超生图像增强等方面也有较好表现,但在低照度图像中就显得有些不足,增强图像的对比度和照度没有那么理想。因为低照度图像中的亮度分布不均匀,在增强过程中可能一些相对高亮度区域被增强得过高,也可能对一些低照度噪声进行加强而产生过多的噪声。

针对低照度图像增强方法的难点,本文进行了以下研究:采用无监督学习网络,基于循环生成对抗网络无监督网络结构,解决配对数据集难以获取的问题;对生成器引入注意力机制,把低光照图像中的低光区域和亮区域区分开来,从而促进网络对图像中的低光区域给予更多的关注;使用残差增强网络,通过加深网络层数,加强低照度图像的增强效果;判

别器采用全局-局部双判别器结构,保证生成的正常亮度图像的所有局部区域看起来更接近于现实环境,避免局部曝光过低。

1 相关工作

1.1 生成对抗网络

生成对抗网络是一个比较特殊的模型^[16],由生成模型和判别模型组成。其中生成器用于处理输入数据,输入图像到生成器后产生数据样本,这些数据不仅适用于图像,在文本甚至语音等方面也可以使用^[17]。判别器则用来判断生成器生成图像相比于真实图像的真实程度。两者互相更新对抗,生成器的目的是处理数据生成判别器都无法判断的图像样本,而判别器的目的则是提升判断能力,以确保判断数据更准确。通过生成器和判别器的互相迭代升级,使输出的数据更好。

其损失函数为

$$\min_D \max_G V(D, G) = -E_{x \sim P_{\text{data}}} \ln D(x) - E_{z \sim P_z} \ln(1 - D(G(z))) \quad (1)$$

式中: x 为真实正常光照图像数据, z 为真实正常光照图像数据, G 为生成器, D 为判别器, z 通过生成器后输出,并且和 x 一起输入到判别器中,判别器判断其真实程度; $E_{x \sim P_{\text{data}}}$ 为真实数据的数学期望,其中服从概率分布 P_{data} ; $E_{z \sim P_z}$ 为真实数据的数学期望,其中服从概率分布 P_z ; $V(G, D)$ 为二元交叉熵函数,通常用于二元分类问题。若把判别器的判定标准转换为数据显示,则把真实的判定真实的正常光照图像定为1,生成器输出的不符合真实数据的伪正常光照图像定为0,而生成器的目的是不断自我升级,使生成的图像能够欺骗判别器,使判别器认为生成器输出的是真实图像,判别器也不断自我升级使其能更好地进行判断。

1.2 注意力机制

视觉注意机制^[18]是类似于人类眼睛的视觉,后者是人类在进化过程中形成的特殊技能,视觉看到的画面从人类的眼睛输入,在脑海故意排除无关的信息,并对关心的对象领域集中更多的注意力。注意力机制可以有效利用人类有限的视觉注意力,排除其他无关因素,专注处理目标数据。因此,把注意力机制引入深度学习中,使深度学习越来越高效化。

注意力机制主要分为柔性注意力和硬性注意力^[19-20],柔性注意力和硬性注意力都可以分为基于输入项和基于位置2个方向。

1.3 残差网络

在网络模型中,因为卷积层的层数加深,会出现各种各样的问题,比如过拟合和梯度消失。由若干的残差块所构成的残差网络(Residual Networks, Res-Net)^[21-22]能够很好地解决以上问题,残差网络已经发展成了一种最为重要的特征提取方法。残差块的数学模型如下。

$$y_l = x_l + F(x_l, W_l) \quad (2)$$

$$x_{l+1} = f(y_l) \quad (3)$$

式中: x_l 为残差块的输入, x_{l+1} 为残差块的输出特征图的映射, $F(x_l, W_l)$ 为运算过程, f 为激活函数。通常残差块是通过卷积操作组成的,每次卷积都有非线性激活函数。

2 方法

2.1 循环生成对抗网络

传统的生成对抗网络都是需要输入图像相对应的真实数据图像,即输入图像通过生成器生成一个假的数据图像,再通过判别器判断这个假数据是否属于真实图像,并把结果发送回给生成器,生成器通过这个结果自我更新,使下次生成的图像能更接近

于答案,通过反复循环提高,生成最贴近于真实图像的输出。这样的设计在有相对应的图像数据集的情况下是没有问题的。

由于现实中成对低照度图像数据集难以获取,并且光照的模式和风格种类繁多,人工合成的光照和噪声通常与自然中的有所不同,导致这些基于合成图像的低照度增强方法对来自真实世界的低照度图像增强能力有限,可以利用无监督学习,从循环生成对抗网络出发,设计一个网络解决当前低照度图像增强方法依赖配对图像数据集的问题。

循环生成对抗网络不仅将低照度图像生成正常亮度图像,而且将生成的正常亮度图像作为输入,让生成器生成低照度图像,并和原输入的低照度图像进行对比,使两者更加接近,以此来摆脱对相应数据集的依赖。

如图1所示,循环生成对抗网络分为2组生成器和判别器,输入分别为低照度图像和正常照度图像,两者数据集不用互相对应。首先,向生成器1输入低照度图像,生成的正常光照度图像通过判别器1判断是否是正常光照度图像,同时输入生成器2,生成伪低照度图像,其要与原输入的低照度图像尽可能相似。另一组原输入为正常照度图像,流程一致。

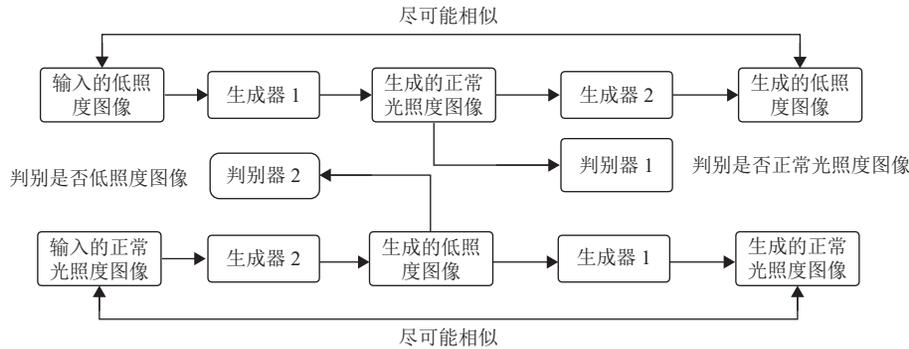


图1 Cycle-Gan网络结构

Fig.1 The network structure of the Cycle-Gan

Cycle-GAN的损失函数总体来说可以分为2部分,一部分是对抗损失,另一部分是循环损失。

$$L(G, F, D_X, D_Y) = L_{GAN}(G, D_Y, X, Y) + L_{GAN}(G, D_Y, X, Y) + \lambda L_{cyc}(G, F) \quad (4)$$

对抗损失函数为

$$L_{GAN}(G, D_Y, X, Y) = E_{y \sim P_{data}(y)} [\ln D_Y(y)] + E_{x \sim P_{data}(x)} [\ln(1 - D_Y(x))] \quad (5)$$

循环损失函数为

$$L_{cyc}(G, F) = E_{x \sim P_{data}(x)} [\|F(G(x)) - x\|_1] + E_{y \sim P_{data}(y)} [\|G(F(y)) - y\|_1] \quad (6)$$

式(4)~(6)中: $E_{x \sim P_{data}(x)}$ 和 $E_{y \sim P_{data}(y)}$ 分别为循环生成对抗网络的X和Y两组输入图像在数据空间中定义的真实数据的数学期望, λ 为用于控制循环一致性损失和对抗损失之间的相对权重。因为需要使循环生成的低照度图像和原输入图像尽可能接近,所以在生成器1生成正常亮度图像后,在生成器2中尽可能恢复成原输入图像,除了对抗损失之外,还需要循环损失来保证循环一致性^[23]。

2.2 改进的网络

为了增强低照度图像的同时解决图像噪声,把网络分解^[24]为注意力机制引导网络和残差增强网络。如图2所示,首先输入图像通过注意力机制网络区分低光区域和亮光区域,生成注意力机制图,以便后续的增强网络能对其进行区分增强。接着把输入图像和注意力机制图通过串接形式一起输入残差网络,通过注意力机制图象的区域划分,对输入图像进行精准亮度增强。因为低照度图像中的光线普遍偏暗,所以图像中的特征信息较少,并且会有部分噪声干扰,在这个基础上,使用残差连接构建基础残差模块,使网络的层数加深。通过增强模型的弱光图像,避免由于网络加深造成特征损失。



图2 生成器网络结构

Fig.2 Generator network diagram

2.2.1 注意力机制网络

传统图像增强学习方法一般针对图像整体发展进行不断增强,而忽略了照片中各个区域亮度不一致的情况,进而容易造成高亮度区域过曝光,低亮度区域相对欠曝光的问题。为解决低照度图像中亮度区域分布不均匀的问题,打算引入注意力机制,通过注意力机制网络区分低光区域和亮光区域,生成注意力机制图,从而促进网络对图像中的弱光区域给予更多的关注和增强。在低照度图像生成正常图像时,尽管图中的不同区域进行了不同增强,各个部分的亮度尽可能靠近,但生成亮度还是有所区别。因此在正常图像中,也有亮度分配不同的情况,通过对亮度情况进行区分,同样也能生成低照度图像。

U-Net^[25]在图像生成领域特别是图像恢复和增强方面有着不错的表现,可以在输入图像中获取不同深度的特征信息,保留比较完整的特征信息,并且利用多层次信息多尺度地生成效果较好的图像,因此使用类U-Net网络作为主要网络。

注意力机制网络作为辅助网络,其训练得到的注意力图与原输入图像相结合,以待后续为加强网络对低光区域进行特别加强,确保低照度图像中低光区域的增强效果。

图3为注意力机制网络,采取类U型网络,首先输入低照度图像通过下采样层的池化和卷积,对输入图像进行分解,加深维度。卷积操作增加图像的

维度,池化操作减小图像每一维度的尺寸,以达到提取特征有效地收集光照信息的目的。再结合上采样层得到的光照信息,对得到的信息进行建模,得到注意力图,如图4所示。图中的白色区域代表输入图像的低照度区域,黑色区域代表输入图像的较亮区域。

$$A = \frac{|\max(I_x) - \text{pixel}(I_x)|}{\max(I_x)} \quad (7)$$

式中: I_x 为低照度图像, $\text{pixel}(I_x)$ 为输入图像本身的像素值, $\max(I_x)$ 为输入图像中的最大像素值。

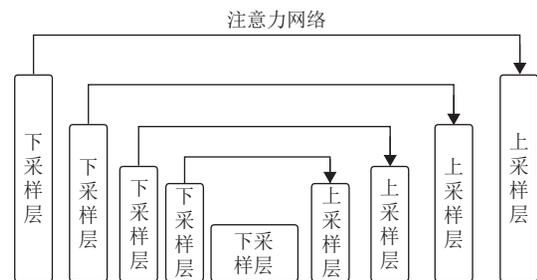


图3 注意力网络结构

Fig.3 Attention network structure



(a) 低照度图像 (b) 注意力图像

图4 低照度图像和注意力图像

Fig.4 Low illumination images and attention-based images

2.2.2 增强网络模块

如图5所示,残差增强网络分为3个部分,其中第1部分为卷积层,首先将图像合并在一起,然后进行卷积运算,步长为2,即将每一维度的尺寸大小减半。中间的残差单元部分为第2部分,通过建立连接多个残差单元提高低照度图像的增强能力,通过一系列的残差块,利用上采样恢复图像的尺寸以及维度。第3部分为普通卷积层,把前面处理的多特征转换成RGB图像。



图5 残差增强网络结构

Fig.5 Residuals enhance network architecture

2.3 判别器

在传统的生成对抗网络方法中^[26],判别器不能很好地对光线变化大的低照度图像发挥作用,和上述注意力机制的作用相比,不能很好地处理低照度图像中特别亮或者特别暗的区域,原因在于传统方法中使用的是全局判别器,对于局部信息变化大的部分不能拥有比较好的自适应能力。为了提高判别器对于低照度图像任务的作用,使用了全局加局部的双判别器结构^[27],其中全局判别器在工作的期间,局部判别器也同时工作。从生成器输出的图像和真实的光照图像当作输入,输入全局判别器中,同时把输入分解成多个局部区域输入局部判别器中,局部判别器使用PatchGAN对真实图像和输出图像进行判别,通过对抗,使输出图像包括其局部区域看起来更加接近于真实图像,避免局部过度曝光。

如图6所示,把生成器生成的图像和正常亮度的图像作为输入在全局判别器中进行判定,从全局的角度来保证完整图像的真实性;同时把两者裁剪数量相等的若干局部区域输入局部判别器中,并输出每份局部区域的真实度,保证图像中每个区域都更靠近真实图像。

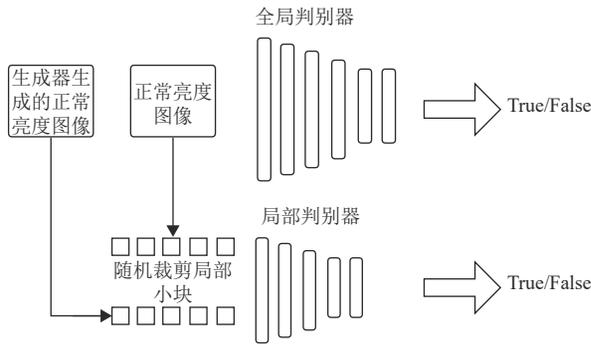


图6 全局-局部判别器

Fig.6 Global-local discriminator

在全局判别器中,使用LSGAN版本的相对论平均GAN损失^[28],损失函数如下。

$$L_D^{\text{global}} = E_{x_r \sim P_r} \left[(D_g(x_r) - E_{x_r \sim P_r} D_g(x_r) - 1)^2 \right] + E_{x_f \sim P_f} \left[(D_g(x_f) - E_{x_r \sim P_r} D_g(x_r))^2 \right] \quad (8)$$

$$L_G^{\text{global}} = E_{x_r \sim P_r} \left[(D_g(x_r) - E_{x_r \sim P_r} D_g(x_r) - 1)^2 \right] + E_{x_f \sim P_f} \left[(D_g(x_f) - E_{x_r \sim P_r} D_g(x_r))^2 \right] \quad (9)$$

式中: G 为生成器, D 为判别器, P_r 为真实图像的正态光分布, x_r 为其相对应的样本; P_f 为生成图像的正态光分布, x_f 为其相对应的样本, $E_{x_r \sim P_r}$ 和 $E_{x_f \sim P_f}$ 分别为真实数据和生成数据的数学期望。

对于局部判别器,也采用原始LSGAN损失来进行训练,生成器和判别器的损失函数如下。

$$L_D^{\text{local}} = E_{x_r \sim P_r \text{-patches}} \left[(D_l(x_r) - 1)^2 \right] + E_{x_f \sim P_f \text{-patches}} \left[(D_l(x_f))^2 \right] \quad (10)$$

$$L_G^{\text{local}} = E_{x_r \sim P_r \text{-patches}} \left[(D_l(x_r) - 1)^2 \right] \quad (11)$$

3 实验

在本节,首先比较每个模型在Enlighten-GAN的未配对增强数据集上的对比度和照明增强性能。然后,报告合成的低光(Low-Light, LOL)数据集的对比度增强后的结果,这些数据集都包含具有明显噪声的低光图像。

3.1 数据集

未配对的增强(Enlighten)数据集:Jiang等^[26]收集了一个未配对的数据集,用于训练对比度增强模型。训练集由914张较暗但没有显著噪声的低光图像和1016张来自公共数据集的正光图像组成。本文使用这个数据集来比较每个模型的对比度增强性能。该评价集由公共数据集的148对低光/正光图像对组成。所有来自训练和评估集的图像的分辨率都被重新调整到400×600。

LOL数据集^[6]:LOL数据集由500对弱光和正光图像对组成,分为485对训练对和15对测试对。低光图像包含了在照片捕获过程中产生的噪声。大部分的图像都是室内的场景。为了使数据集适应本文的无监督设置,这里采用485张训练图像作为本文的弱光训练集,并采用未配对增强Enlighten数据集的正光图像作为正光训练集。测试映像与LOL数据集保持相同。所有图像的分辨率都为400×600。

3.2 实施细节及评价指标

使用Adam优化器,学习率设置为 10^{-4} ,批量大小设置为32。为了对增强后的图像进行质量评价,本文采用峰值信噪比(Peak Signal to Noise Ratio, PSNR)和结构相似度(Structural Similarity, SSIM)^[29]2个指标进行定量分析。SSIM表示了一种全参考的图像评价质量指标,代表结构相似性指标,包括亮度、对比度以及图像具体的结构3个部分,PSNR是峰值信噪比指标。

3.3 Enlighten数据集上的实验

这里比较了本文模型以及包括UNIT^[25]、CycleGAN^[15]和Enlighten-GAN^[26]在内的低光照图像增强模型在Enlighten数据集上的表现。所有的模型都在训

练集上进行训练,并在测试集上进行评估。

在图7中可以看出UNIT模型和Cycle-GAN模型提升了一定的亮度,但对比其他模型,其色彩有一定的缺失,产生沉重的伪影和遭受轻微的颜色扭曲;而Enlighten-GAN模型的增强效果很明显,并且色彩鲜艳,在天空和地上行人部分都能得到很好的增强,和

本文的模型一样可以在全局和局部区域生成具有合理对比度和清晰度的图像。

除了主观视觉感官外,本文还报告了生成图像的PSNR和SSIM,作为对未配对增强数据集的数据评价。结果如表1所示,本文的模型比现有的模型表现得更好,这与视觉结果是一致的。

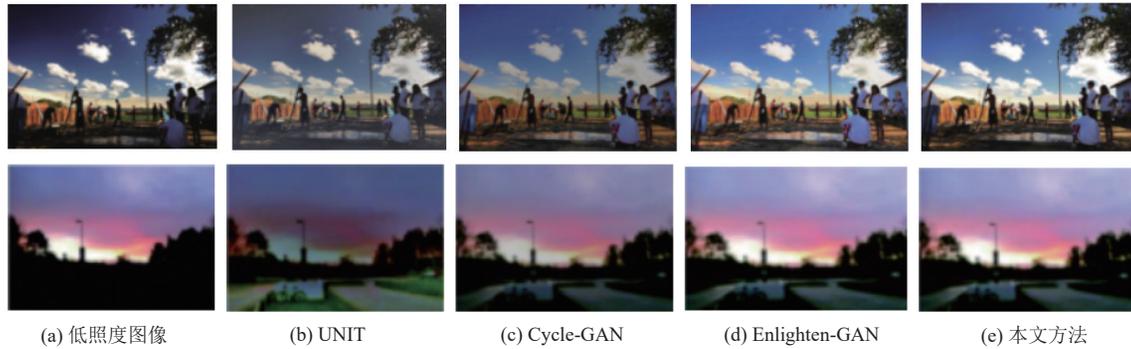


图7 Enlighten数据集上的结果
Fig.7 Results on the Enlighten dataset

表1 在Enlighten数据集和LOL数据集上的结果
Table 1 Results on the Enlighten and LOL datasets

Model	Enlighten数据集		LOL数据集	
	PSNR	SSIM	PSNR	SSIM
UNIT	13.74	0.462 7	15.72	0.739 1
Cycle-GAN	15.59	0.683 6	16.38	0.758 3
Enlighten-GAN	17.53	0.804 1	18.48	0.796 3
本文方法	18.86	0.818 5	19.26	0.820 7

3.4 LOL低光图像增强实验

在合成数据集的LOL弱光图像增强数据集上评

估了本文的方法,并将其与先进的无监督图像转换进行比较,包括Cycle-GAN和Enlighten-GAN。图8显示了LOL数据集的表现结果。Cycle-GAN的数据存在严重的颜色失真,无法保留LOL数据集的细节。Enlighten-GAN能够提高图像的对比度,但图像上仍存在噪声和伪影。

放大图8中保龄球道增强的细节,如图9所示,Enlighten-GAN中的球瓶显示情况并没有本文方法的清晰,从此细节看出本文方法在单一模型的真实低光图像增强中同样具有优势。

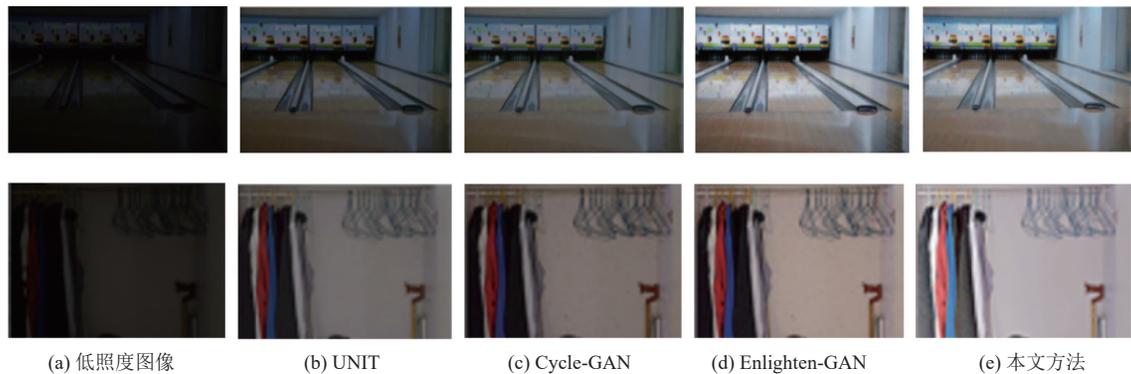


图8 LOL数据集上的结果
Fig.8 Results on the LOL dataset

同样,对于LOL数据集的实验,本文也报告了生成图像的PSNR和SSIM,作为LOL数据集的数据评价。结果如表1所示,本文的模型同样领先,进一步证明了本文模型的优越性。

3.5 消融研究

为研究本文模型中的每个组件对最终结果的贡献设计了2个实验,分析了使用残差网络而不用注意力机制网络的情况,和使用了注意力机制网络并且



图 9 放大的细节图像

Fig.9 Magnified details of the image

调整残差块数量的情况,对比整个模型,明确看出其发挥的关键作用。

具体来说,第1个实验不使用注意力机制网络来进行训练并和完整的模型进行对比,以此看出注意力机制网络在本文模型中的作用。第2个实验则是因

为本文使用了残差块进行增强,并且在残差增强网络中残差块的数量也对实验结果有着一定的差异影响,所以进行额外对比,分别选取数量为4,6,8,10的残差块数量进行对比实验。

实验结果如图10所示,在没有注意力机制网络时,结果与完整模型相比,在两侧的玻璃和玻璃柜子上的亮度偏低,并且出现了一些亮度噪声。而残差块部分,几种残差块数量的结果在感官上大致相差不大。随着残差块数量的增加,增强网络的层数也增加时,在一些局部复杂部分,部分线条会出现重叠,出现过拟合的现象。残差块数量为6块时,得到的图像最为平滑。

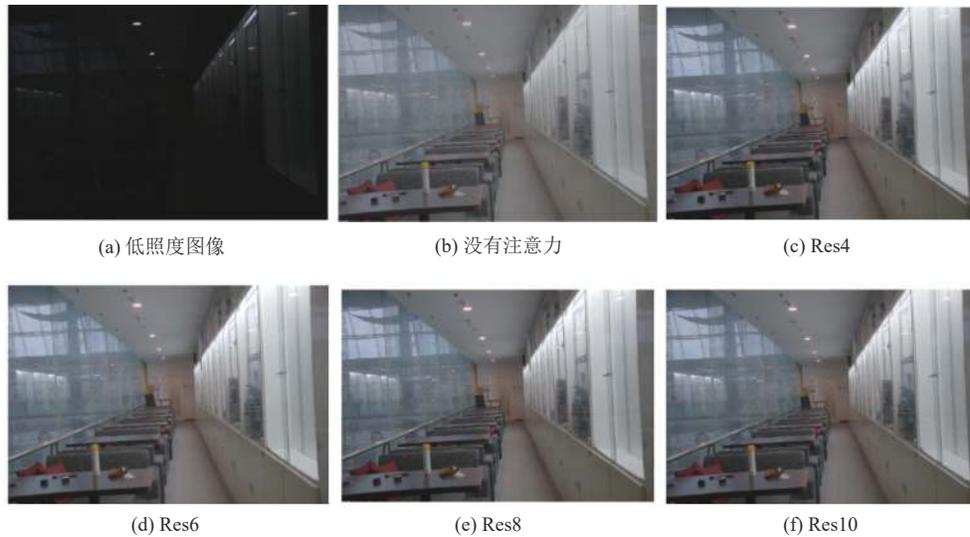


图 10 在Enlighten数据集上的消融研究

Fig.10 Ablation studies on the Enlighten dataset

同时,对于消融实验,本文报告了生成图像的PSNR和SSIM,作为消融实验的数据评价。结果如表2所示,在残差块数量为6块时取得的效果最好,虽然随着数量增加,PSNR和SSIM数值同时也在增加,但是在超过6块时,残差块数量增加会导致模型深度增加,从而使网络的复杂程度和时间消耗增加。所以选取残差块数量为6块的残差网络作为增强网络是最好的选择。

表 2 在Enlighten数据集上消融实验的结果

Table 2 Results of ablation studies on the Enlighten dataset

Model	PSNR	SSIM
Without-Attention- Mechanism	16.87	0.769 2
Res4	18.64	0.801 5
Res6	19.34	0.831 2
Res8	18.89	0.809 5
Res10	19.03	0.812 4

4 结论

本文所介绍的方法采用无监督生成对抗网络作为主要网络,并把网络分解为注意力网络和残差增强网络,对低照度图像进行增强,达到比较理想的效果。相比较传统的方法,不仅能适应多种低照度环境下的图像增强,同时也提高了增强后图像的视觉效果。通过实验表明,本文所提方法,得到的结果细节描述更准确,色彩深度更丰富,局部对比度更强,在主观和客观验证下都优于Enlighten-GAN、CycleGAN等方法。

参考文献:

- [1] ZHOU Z, FENG Z, LIU J, *et al.* Single-image low-light enhancement via generating and fusing multiple sources [J]. *Neural Computing and Applications*, 2020, 32: 6455-6465.
- [2] GHARBI M, CHEN J, BARRON J T, *et al.* Deep bilateral

- learning for real-time image enhancement [J]. *ACM Transactions on Graphics (TOG)*, 2017, 36(4): 118.
- [3] GUO X, LI Y, LING H. LIME: Low-light image enhancement via illumination map estimation [J]. *IEEE Transactions on Image Processing*, 2016, 26(2): 982-993.
- [4] LORE K G, AKINTAYO A, SARKAR S. LLNet: a deep autoencoder approach to natural low-light image enhancement [J]. *Pattern Recognition*, 2017, 61: 650-662.
- [5] SHEN L, YUE Z, FENG F, *et al.* MSR-net: low-light image enhancement using deep convolutional network [EB/OL]. arXiv: 1711.02488 (2017-11-07) [2022-03-25]. <https://doi.org/10.48550/arXiv.1711.02488>.
- [6] WEI C, WANG W, YANG W, *et al.* Deep retinex decomposition for low-light enhancement[EB/OL]. arXiv: 1808.04560 (2018-08-14) [2022-03-25]. <https://doi.org/10.48550/arXiv.1808.04560>.
- [7] CHEN C, CHEN Q, XU J, *et al.* Learning to see in the dark[EB/OL]. arXiv: 1805.01934 (2018-05-04) [2022-03-25]. <https://doi.org/10.48550/arXiv.1805.01934>.
- [8] KALANTARI N K, RAMAMOORTHI R. Deep high dynamic range imaging of dynamic scenes[J]. *ACM Transactions on Graphics*, 2017, 36(4): 1-12.
- [9] WU X, SHAO J, GAO L, *et al.* Unpaired image-to-image translation from shared deep space[C]//2018 25th IEEE International Conference on Image Processing (ICIP). Athens: IEEE, 2018: 2127-2131.
- [10] LIU M Y, BREUEL T, KAUTZ J. Unsupervised image-to-image translation networks[EB/OL]. arXiv: 1703.00848 (2018-07-23) [2022-03-25]. <https://doi.org/10.48550/arXiv.1703.00848>.
- [11] MADAM N T, KUMAR S, RAJAGOPALAN A N. Unsupervised class-specific deblurring[C]//Computer Vision—ECCV 2018: 15th European Conference. Munich: Springer International Publishing, 2018: 358-374.
- [12] HUANG X, LIU M Y, BELONGIE S, *et al.* Multimodal unsupervised image-to-image translation[EB/OL]. arXiv: 1804.04732 (2018-08-14) [2022-03-25]. <https://doi.org/10.48550/arXiv.1804.04732>.
- [13] CHOI Y, CHOI M, KIM M, *et al.* Stargan: unified generative adversarial networks for multi-domain image-to-image translation[EB/OL]. arXiv: 1711.09020 (2018-09-21) [2022-03-25]: <https://doi.org/10.48550/arXiv.1711.09020>.
- [14] HOFFMAN J, TZENG E, PARK T, *et al.* CyCADA: cycle-consistent adversarial domain adaptation[EB/OL].arXiv: 1711.03213 (2017-12-29) [2022-03-25]. <https://doi.org/10.48550/arXiv.1711.03213>.
- [15] ZHU J Y, PARK T, ISOLA P, *et al.* Unpaired image-to-image translation using cycle-consistent adversarial networks [EB/OL].arXiv: 1703.10593 (2020-08-24) [2022-03-25]. <https://doi.org/10.48550/arXiv.1703.10593>.
- [16] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, *et al.* Generative adversarial networks [J]. *Communications of the ACM*, 2020, 63(11): 139-144.
- [17] GAWADE A, PANDHARKAR R, DEOLEKAR S. Gantoon: creative cartoons using generative adversarial network[C]//Information, Communication and Computing Technology: 5th International Conference, ICICCT 2020. Singapore: Springer, 2020: 222-230.
- [18] HU D. An introductory survey on attention mechanisms in NLP problems[EB/OL]. arXiv: 1811.05544 (2018-11-12) [2022-03-25]. <https://doi.org/10.48550/arXiv.1811.05544>.
- [19] WANG F, TAX D M J. Survey on the attention based RNN model and its applications in computer vision[EB/OL]. arXiv: 1601.06823 (2016-01-25) [2022-03-25]. <https://doi.org/10.48550/arXiv.1601.06823>.
- [20] JIANG X, ZHANG L, XU M, *et al.* Attention scaling for crowd counting[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, WA: IEEE, 2020: 4706-4715.
- [21] HAN Y S, YOO J, YE J C. Deep residual learning for compressed sensing CT reconstruction via persistent homology analysis[EB/OL]. arXiv: 1611.06391 (2016-11-25) [2022-03-25]. <https://doi.org/10.48550/arXiv.1611.06391>.
- [22] SHIBATA N, TANITO M, MITSUHASHI K, *et al.* Development of a deep residual learning algorithm to screen for glaucoma from fundus photography [J]. *Scientific Reports*, 2018, 8(1): 14665.
- [23] KUANG D Y. On reducing negative jacobian determinant of the deformation predicted by deep registration networks [EB/OL]. arXiv: 1907.00068 (2019-06-28) [2022-03-25]. <https://doi.org/10.48550/arXiv.1907.00068>.
- [24] XIONG W, LIU D, SHEN X, *et al.* Unsupervised real-world low-light image enhancement with decoupled networks. [EB/OL]. arXiv: 2005.02818 (2020-05-06) [2022-03-25]. <https://doi.org/10.48550/arXiv.2005.02818>.
- [25] RONNEBERGER O, FISCHER P, BROX T. U-net: convolutional networks for biomedical image segmentation [EB/OL]. arXiv: 1505.04597 (2015-05-18) [2022-03-25]. <https://doi.org/10.48550/arXiv.1505.04597>
- [26] JIANG Y, GONG X, LIU D, *et al.* Enlightengan: deep light enhancement without paired supervision [J]. *IEEE Transactions on Image Processing*, 2021, 30: 2340-2349.
- [27] IIZUKA S, SIMO-SERRA E, ISHIKAWA H. Globally and locally consistent image completion [J]. *ACM Transactions on Graphics (TOG)*, 2017, 36(4): 107.
- [28] MAO X, LI Q, XIE H, *et al.* Least squares generative adversarial networks[EB/OL]. arXiv: 1611.04076 (2017-04-05) [2022-03-25]. <https://doi.org/10.48550/arXiv.1611.04076>.
- [29] HUYNH-THU Q, GHANBARI M. Scope of validity of PSNR in image/video quality assessment [J]. *Electronics letters*, 2008, 44(13): 800-801.

(责任编辑: 杨耀辉 英文审核: 费伦科)