

文章编号 1004-924X(2024)10-1538-14

双路混合注意力的跨层次特征聚合图像增强

袁 姮¹, 王笑雪^{1*}, 颜廷昊¹, 张晟翀²

(1. 辽宁工程技术大学 软件学院, 辽宁 葫芦岛 125100;

2. 光电信息控制和安全技术重点实验室, 天津 300308)

摘要:针对低照度图像存在亮度低、噪声大、颜色偏差和细节纹理丢失等问题,提出一种双路混合注意力的跨层次特征聚合图像增强方法。首先,设计多尺度双路注意力残差模块(Multi-scale Dual-path Attention Residual module, MDAR),MDAR包括并行多尺度特征采样块(Parallel Multi-scale Feature Sampling Block, PMFB)和双路混合注意力块(Dual-path Hybrid Attention Block, DHAB)。其中PMFB用于提取和融合多尺度特征信息,促进局部特征的全局化表示,使图像细节信息得到有效增强,而DHAB能够对图像噪声区域和颜色信息给予更大关注,缓解不同注意力间特征的差异,有效抑制噪声,提高图像质量。此外,设计跨层次特征聚合模块(Cross-level Feature Aggregation Module, CFAM),将不同层次特征进行融合,弥补深层特征与浅层特征之间的差异,强化对浅层特征的感知,实现图像增强。实验结果表明,所提方法在LOL数据集上的PSNR,SSIM,LPIPS和NIQE分别达到了22.347 dB,0.850,0.178和4.153;在MIT-Adobe 5K数据集上的PSNR,SSIM,LPIPS和NIQE分别达到了22.703 dB,0.903,0.137和3.822。与其他算法相比均有较大提升,证明了所提方法的有效性。

关键词:图像增强;多尺度;混合注意力;特征聚合

中图分类号:TP394.1 **文献标识码:**A **doi:**10.37188/OPE.20243210.1538

Cross-level feature aggregation image enhancement with dual-path hybrid attention

YUAN Heng¹, WANG Xiaoxue^{1*}, YAN Tinghao¹, ZHANG Shengchong²

(1. College of Software, Liaoning Technical University, Huludao 125105, China;

2. Key Laboratory of Optoelectronic Information Control and Security Technology, Tianjin 300308, China)

* Corresponding author, E-mail: wxx2585354184@163.com

Abstract: To address the problems of low brightness, high noise, color deviation and loss of detail and texture in low-light images, this study proposed an image enhancement method using dual-channel hybrid attention and cross-level feature aggregation. Firstly, the Multi-scale dual-path attention residual module (MDAR) was designed. MDAR included a Parallel multi-scale feature sampling block (PMFB) and a Dual-path hybrid attention block (DHAB). By extracting and fusing multi-scale feature information, PMFB promoted the global representation of local features, and effectively enhanced image details. DHAB could pay more attention to image noise regions and color information, which not only alleviates

收稿日期:2023-11-21;修订日期:2024-01-05.

基金项目:国防预研基金项目(No. 172068);辽宁省自然科学基金(No. 20170540426);辽宁省教育厅重点基金(No. LJYL049)

the feature differences between different attention spans, but also effectively suppress noise and improve image quality. In addition, this paper designed a Cross-level feature aggregation module (CFAM), which fuses features at different levels to make up for the differences between deep features and shallow features, strengthen the perception of shallow features, and achieve image enhancement. Experimental results indicate that the PSNR, SSIM, LPIPS and NIQE of the proposed method on the LOL dataset reached 22.347 dB, 0.850, 0.178 and 4.153 respectively and the PSNR, SSIM, LPIPS and NIQE of the proposed method on the MIT-Adobe 5K dataset reached 22.703 dB, 0.903, 0.137 and 3.822 respectively. Compared with other algorithms, the algorithm in this paper has been greatly improved, which proves the effectiveness of the proposed method.

Key words: image enhancement; multi-scale; mixed attention; characteristic polymerization

1 引 言

图像增强技术广泛应用在医疗、工业和煤矿勘测等领域。低光照环境拍摄的照片往往存在亮度低、噪声大、颜色偏差和细节纹理丢失等问题,给后续图像分析和识别工作带来了极大的挑战。因此,如何对低照度图像进行处理和增强,得到高质量的清晰图像,是图像处理领域研究的热点问题之一。

低照度图像增强算法主要分为传统方法和基于深度学习的方法。传统方法主要包括基于分布映射的方法和基于模型优化方法。直方图均衡化方法是基于分布映射的代表方法,常规的直方图均衡化是一种通过调整图像的直方图分布,拉伸图像动态范围,从而增强图像对比度的方法。Pisano 等^[1]提出基于直方图均衡化(Histogram Equalization, HE)的方法实现图像增强,但 HE 对处理动态范围较大的图像具有局限性,于是直方图扩展方法 BPDHE^[2]和限制对比度自适应直方图均衡化方法 CLAHE^[3]被提出,虽然视觉效果强于传统 HE 方法,但增强过程中缺乏对图像内容的关注。Retinex 理论假设方法是基于模型优化的代表方法,Land 等^[4]提出 Retinex 理论,利用人眼对颜色感知的不变性原理将图像分解为光照图和反射图,Guo 等^[5]基于 Retinex 理论提出一种简单有效的 LIME 方法,该算法能较好地提升图像视觉质量,但增强图像存在颜色偏差和曝光现象。为解决 LIME 的曝光过度现象,Zhang 等^[6]引入正则项约束对光照进行调整,但算法复杂,速度较慢。Li 等^[7]在 Retinex 基础上考虑在增强图像亮度的同时削弱噪声对图像的影响,然而该模型忽

略对图像色彩的增强,出现颜色失真现象。BI-MEF^[8]在 Retinex 基础上引入相机模型实现图像增强,但图像亮度较低,效果不理想。总体而言,利用传统方法对现实场景中的图像进行增强仍存在很多问题,难以满足实际应用需求。

针对传统方法存在的不足,研究学者们提出利用深度学习方法解决图像增强问题。LLNet^[9]是第一个将深度学习应用到图像增强领域的算法,该算法对图像增强和去噪都有较好效果。Retinex Net^[10]方法将 Retinex 理论和深度学习相结合,将低光图像分解成光照图和反射图,再利用增强网络对图像进行增强和去噪,但该方法缺乏对图像色彩增强的关注,出现颜色失真问题。KinD^[11]在 Retinex Net 算法基础上对网络进行调整,利用分解网络、重建网络和调节网络实现对图像亮度的连续调整。针对图像增强过程中可能出现的曝光过度或曝光不足现象,Wu 等人^[12]提出 URetinex Net 方法将分解公式转化为隐式先验正则化模型,利用三个基于学习的模块实现噪声抑制和细节保留,Wang 等^[13]利用多曝光融合的方法对低照度图像进行增强。但由于 Retinex 理论的局限性,这类方法在图像分解时往往会放大噪声,影响图像增强效果。DSLRL^[14]将拉普拉斯金字塔引入低光照图像增强领域,利用拉普拉斯金字塔的多尺度结构与高阶残差,使图像增强显著。近年来,Transformer 在自然语言处理领域表现优秀,因此学者们也将其引用到图像增强领域并展开研究。Wang 等^[15]提出一种分层编解码网络 UFormer,通过引入局部增强窗口模块和多尺度复原调制模块灵活处理各种图像增强任务。此外,Li 等^[16]观察到图像美学质量评价与人类主观

评价在一定程度上是一致的,因此提出利用美学策略生成模块和美学评估模块提高图像增强任务性能。为满足人们对图像增强效果和处理速度的需求,Qu等^[17]提出一种编解码结构的实时低光图像增强网络DDNet,通过颜色增强模块和梯度增强模块分解和重建图像,具有较好的图像增强效果,但该方法缺乏对图像局部区域的关注,导致增强图像局部亮度不足。针对图像增强过程中特征融合不充分导致的图像亮度不均匀以及细节信息丢失等问题,刘光辉等^[18]提出渐进特征融合的思想,分阶段对特征信息进行补充,缓解了信息冗余问题。陈清江^[19]等引入八度卷积对通道进行拆分,考虑了不同通道对图像增强效果的影响,再通过多通道融合的方式提取多层次特征,有效提高了图像亮度且图像纹理细节保持较好。

随着无监督深度学习方法的兴起,Jiang等^[20]提出Enlighten GAN方法,该方法基于U-Net网络以无监督训练的方式实现低光图像到正常光图像的映射,但GAN在训练过程中容易出现梯度爆炸问题,很难训练。Guo等^[21]提出Zero DCE方法,该方法使用卷积神经网络估计出一组光照曲线的映射参数,并通过迭代光照曲线实现低光照图像增强,但该方法存在细节丢失问题。Li等^[22]在Zero DCE基础上提出Zero DCE++,模型的运算效率进步明显,但性能相差不大。Liu等^[23]提出RUAS对低光图像曝光不足的结构建立模型表示,再对图像整体进行优化,该网络速度较快所需资源较少。Ma等^[24]提出SCI方法实时处理图像增强任务,但该方法存在局部曝光现象,图像整体视觉较暗。这些方法虽能减少网络训练时对数据的依赖,但在图像细节增强等方面具有局限性。

与传统图像增强方法相较,现有深度学习图像增强方法在性能上有了较大提升,但在细节增强、噪声退化和色彩恢复等方面仍有很大进步空间。神经网络中的残差网络和编解码网络对解决网络加深带来的梯度下降和获取多尺度信息非常有效,不仅能解决网络饱和问题,还能以逐步减小空间维度的方式对特征进行提取,在图像对比度增强、噪声退化和细节恢复方面具有很好的效果。因此,本文以残差网络和编解码网络为基础,设计多尺度双路注意力残差模块(Multi-scale Dual-path Attention Residual module,

MDAR)和跨层次特征聚合模块(Cross-level Feature Aggregation Module, CFAM),对编解码网络进行优化,提出一种双路混合注意力的跨层次特征聚合图像增强方法。

首先,构建MDAR,该模块由并行多尺度特征采样块(Parallel Multi-scale Feature sampling Block, PMFB)和双路混合注意力块(Dual-path Hybrid Attention Block, DHAB)构成。为使输入特征信息得到最大化利用,设计PMFB提取多尺度特征信息并融合,使图像全局特征与局部特征建立紧密联系,充分还原图像细节信息和色彩信息。为增强图像亮度的同时去除低光照图像夹杂的噪声,提出DHAB,该模块利用像素位置权重和通道权重对图像噪声区域和颜色信息给予更大关注,通过缓解不同注意力间特征的差异有效抑制噪声,提高图像质量。此外,为学习丰富的特征信息,缓解网络因层数过多丢失部分浅层特征信息问题,除在编解码器之间引入跳跃连接外,本文CFAM在不同层级特征之间建立联系,削弱不同层次、不同尺度特征间的语义鸿沟,多方位聚合不同层次信息,有效避免信息流失。

2 本文方法

本文提出的双路混合注意力的跨层次特征聚合图像增强网络主要包括通道注意力(Channel Attention, CA)、编码器、解码器、MDAR和CFAM。本文网络结构图如图1所示,该网络是一个端到端的图像增强网络,输入为待增强的低光照图像,输出为同尺寸的增强图像。

首先,在编码器进行浅层特征提取前,利用 3×3 卷积和通道注意力对网络提取的原始图像特征进行通道加权,有助于引导网络对包含重要信息的图像区域赋予更大关注度,让网络在图像增强过程中充分学习更多有效特征,提高网络特征提取能力。其次,将赋予权重信息的特征图像作为编码器的输入,编码器由三个 3×3 下采样卷积组成,解码器由三个 3×3 上采样卷积组成,相较于单一上下采样层,本文在编解码器中间设计了MDAR模块。

MDAR模块由PMFB和DHAB构成。PMFB提取多尺度特征信息并融合,充分利用图像的局部特征和全局特征,丰富模型对图像特征

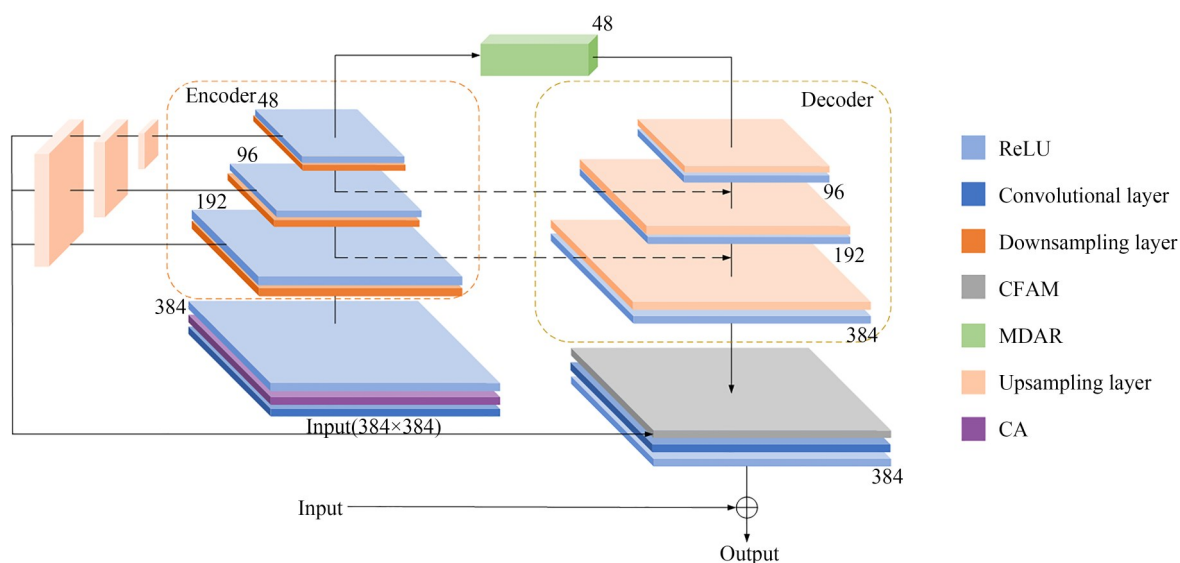


图1 本文网络结构图

Fig. 1 Network structure diagram of this paper

的感知能力,使图像细节信息得到有效增强。DHAB提高网络对多尺度特征有用的颜色信息和光照信息的关注度,强化对高噪声区域的学习,有利于得到高质量低噪声图像,实现性能增益。

同时,本文还设计了CFAM模块,除在相同维度的上下采样卷积间进行跳跃连接外,还加入跨层次连接,将深层特征与浅层特征通过CFAM进行聚合,实现特征再利用,减少冗余特征,达到

更好恢复图像纹理细节和颜色信息的目的。最后,将CFAM输出特征图通过 3×3 卷积与ReLU激活函数进行通道压缩,再与输入图像相加,得到增强图像。

2.1 多尺度双路注意力残差模块(MDAR)

为提升特征利用率,在编码器、解码器之间构建MDAR,该模块由PMFB与DHAB构成,MDAR结构如图2所示。

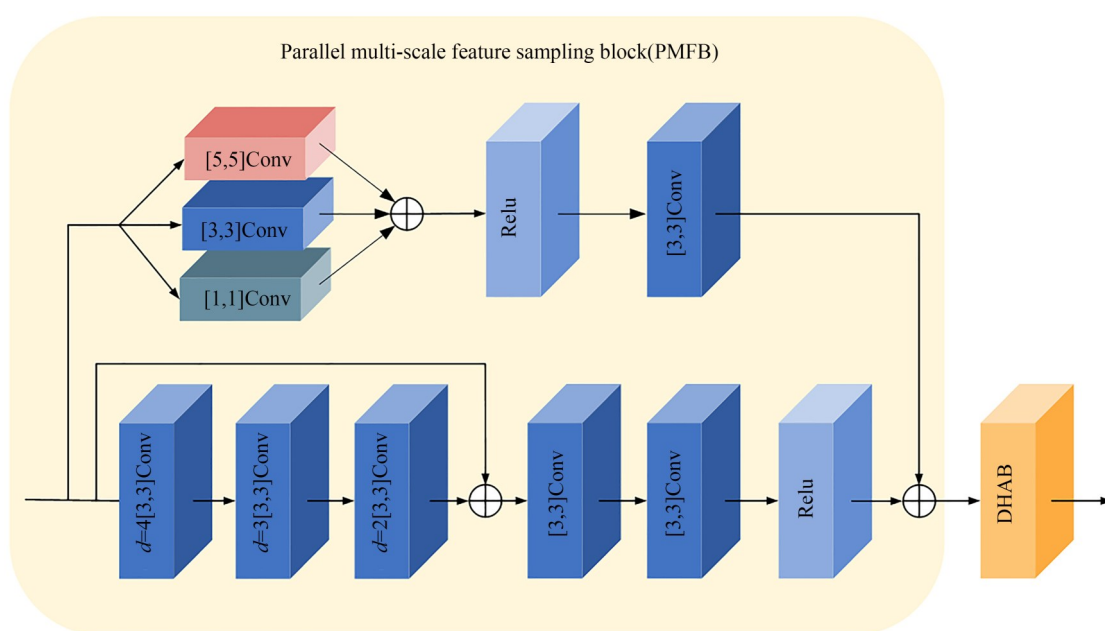


图2 MDAR结构图

Fig. 2 Structure diagram of MDAR

2.1.1 并行多尺度特征采样块(PMFB)

低照度图像不同区域往往存在不同的照度信息,而使用相同感受野的卷积提取图像特征会导致图像增强效果不理想。因此,本文设计PMFB提取多尺度特征。多尺度指经过不同感受野的卷积可以得到不同范围的特征,大感受野可以得到更全面、语义信息更丰富的特征,小感受野可以得到更偏向于局部和细节的特征。因此,在主干支路使用三个不同空洞率(4,3,2)的卷积构建混合空洞卷积,同时采用残差结构将获取特征与原始特征再次融合,增强关键特征信息,使网络更加健壮有效。为避免网格效应带来的信息丢失问题,本文除了选择无大于1公因数关系的空洞率外,还在空洞卷积后放置两个 3×3 卷积对不同感受野获取的特征信息进行补充,保证图像信息的连续性。最后使用ReLU激活函数使特征图像更加稀疏化,从而挖掘更丰富的层次特征。

相较于使用多个小卷积核堆叠获得大感

受野的方式,PMFB在并行支路使用大卷积核能更直接拓展感受野,提高网络多尺度表征能力,同时规避模型深度增加带来的优化难题^[25]。该支路将大小分别为 1×1 , 3×3 和 5×5 的卷积核并行放置,提取特征并进行融合,随后经过ReLU激活函数与 3×3 卷积,再与主干混合空洞卷积残差结构进行融合。PMFB利用不同感受野的卷积核充分挖掘低照度图像的多尺度特征信息,使图像全局特征与局部特征建立紧密联系,充分还原了图像细节信息和色彩信息。

2.1.2 双路混合注意力块(DHAB)

低照度图像往往伴随噪声等问题,单纯对图像进行亮度增强或图像去噪并不能很好地恢复图像质量,且对图像进行亮度增强和噪声去除的顺序不同也会有不同的增强效果,进而影响图像质量。因此,在提升亮度的同时对图像进行去噪是图像增强的关键。本文提出的DHAB结构如图3所示。

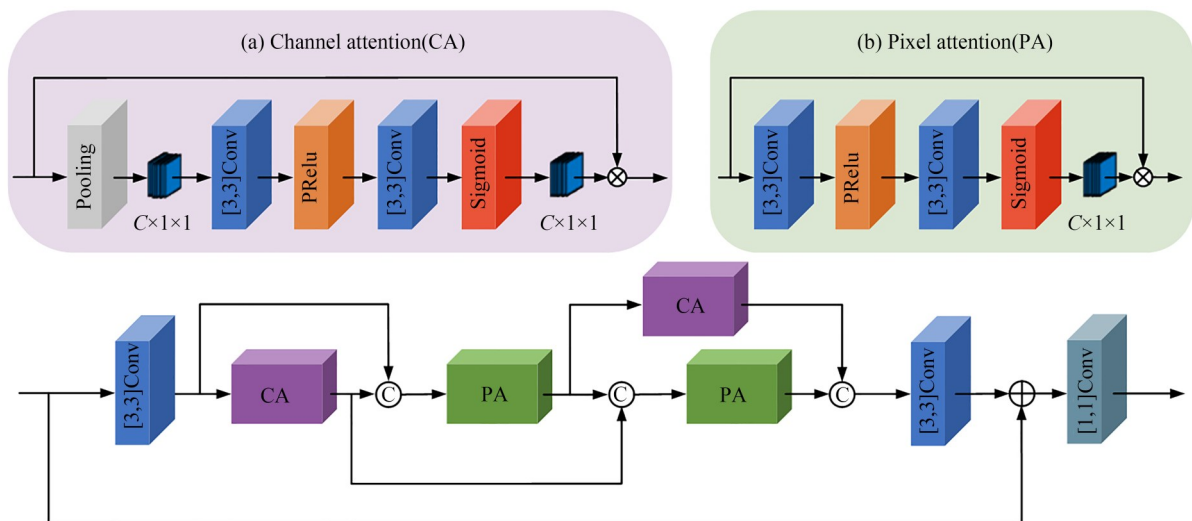


图3 DHAB结构图

Fig. 3 Structure diagram of DHAB

常见注意力方法将不同注意力机制以串联或并联的单一方式排列,这种方法限制了特征间的信息共享。DHAB将通道注意力与像素注意力(Pixel Attention, PA)以串并联混合的方式排列,弥补通道注意力与像素注意力各自的缺陷,补充不同排列方式造成的特征差异,同时每个注意力输出特征都会与相邻注意力特征融合,在得

到像素注意力权重的同时保证通道间的信息交互,有助于网络抑制低光照图像噪声的同时恢复图像颜色与亮度。最后将原始输入特征与经过DHAB输出的特征进行融合,避免图像失真问题。

对绝大部分低光照图像增强网络来说,其输出特征图的每一个通道对增强结果的作用都是

平均的,然而有些通道特征信息不重要,有些通道特征信息更重要,这就要求网络在学习过程中更加关注通道间的相关性并对其赋予更高学习权重。通道注意力首先在通道维度上对大小为 $C \times H \times W$ 的输入图像 X_{input} 进行全局池化操作,得到一个大小为 $C \times 1 \times 1$ 的特征向量 V ,如式(1)所示:

$$V = \frac{1}{H \times W} \sum_{x=1}^H \sum_{y=1}^W X_{\text{input}}(x, y). \quad (1)$$

其次利用两层 3×3 卷积与 ReLU 激活函数对特征向量进行学习,使用 Sigmoid 函数校准每个通道的权重并得到权重向量 V_{wca} ,如式(2)所示:

$$V_{\text{wca}} = S\left(\text{Conv}\left(R\left(\text{Conv}(V)\right)\right)\right), \quad (2)$$

其中: S 表示 Sigmoid 激活函数, R 表示 ReLU 激活函数, Conv 表示卷积核大小为 3×3 的卷积操作。最后将 V_{wca} 与输入特征按元素相乘,得到具有通道关注的特征图 X_{ca} 。如式(3)所示:

$$X_{\text{ca}} = V_{\text{wca}} \otimes X_{\text{input}}, \quad (3)$$

其中: $\otimes$ 表示元素级乘法。 CA 通过重新定义每个通道的权重信息来区分每个特征通道的重要性,使模型能够灵活对特征进行取舍,筛选出包含图像亮度信息和颜色信息的重要特征,从而使网络对任务相关信息的获取更加敏感。

低照度图像像素间存在很强的依赖关系,可以通过像素注意力合理分配各像素点权重进而提高网络对图像像素信息的关注度,使网络借助生成的像素注意力图弱化噪声在图像增强过程中产生的影响。像素注意力模块依次通过 3×3 卷积、ReLU 激活函数、 3×3 卷积和 Sigmoid 激活函数对输入图像 X_{input} 进行学习,通过压缩通道数增强单通道内图像的信息交流,进而得到权重向量 V_{wpa} 。之后,使用生成的权重向量 V_{wpa} 与输入特征图以逐元素相乘的方式进行融合,得到与输入特征图相同大小的像素注意力映射 X_{pa} , PA 增强了图像像素信息的有效性,提高网络对图像中光线分布不均匀区域的关注度,弱化噪声的影响。如式(4)和式(5)所示:

$$V_{\text{wpa}} = S\left(\text{Conv}\left(R\left(\text{Conv}\left(X_{\text{input}}\right)\right)\right)\right), \quad (4)$$

$$X_{\text{pa}} = V_{\text{wpa}} \otimes X_{\text{input}}, \quad (5)$$

其中: S 表示 Sigmoid 激活函数, R 表示 ReLU 激

活函数, Conv 表示卷积核大小为 3×3 的卷积操作, $\otimes$ 表示元素级乘法。

2.2 跨层次特征聚合模块(CFAM)

在网络卷积层中,浅层卷积得到的特征往往具有更丰富的细节信息,但由于卷积层数少,其包含的语义信息相对匮乏。深层卷积得到的特征具有更丰富的语义信息和更少的噪声,但包含的细节信息也相对较少。因此,将不同层次特征进行融合能够弥补因经过的卷积层数不同所造成的特征差异,改善图像增强效果。然而传统图像增强网络只在编解码器中间引入跳跃连接对相同大小的特征进行聚合,不仅会导致局部特征融合不充分,而且由于不同特征的重要性不同,简单地以相同权重对特征进行聚合会造成特征冗余,使有用特征被无用特征削弱,丢失部分特征。因此,本文设计 CFAM,利用上采样卷积将编码器层输出特征恢复至原始图像大小,并将编解码器上不同尺度特征以不同权重灵活的聚合在一起,更好地恢复图像纹理信息和颜色信息。

本文设计的 CFAM 结构如图 4 所示。层归一化操作将同一样本内所有特征规范化,一方面保持特征间依赖关系,另一方面使进入 Sigmoid 激活函数的输入特征值落在函数较为敏感的区域,从而避免梯度消失现象,减少训练时间。CFAM 首先将网络提取的不同层次特征连接在一起,经过卷积、BN 层和 Sigmoid 激活函数对不同层次特征赋予不同的权重。其次,不同层次特征与各自对应权重相乘并将权重引导的特征进行聚合,充分弥补深浅层特征信息间的差异,获得更具鲁棒性的特征图像。

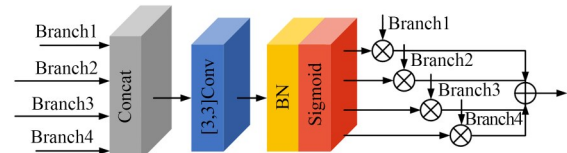


图4 CFAM结构图

Fig. 4 Structure diagram of CFAM

2.3 损失函数

为降低训练复杂度,本文采用均方误差损失(Mean-Square Error loss, MSE loss)监督网络训练过程,该损失对于异常值更加敏感,收敛速度

更快。均方误差损失如式(6)所示:

$$MSE = \frac{1}{N} \sum_{i=1}^N (Y_i - f(x_i))^2, \quad (6)$$

其中: N 为训练集的样本总数, Y_i 表示获得的增强图像, $f(x_i)$ 表示参考图像,通过计算 Y_i 与 $f(x_i)$ 差值的平方优化模型的训练过程。

3 实验及结果分析

3.1 实验细节

本文采用端到端的方式在Tensorflow 2.12.0框架上进行训练,实验在Intel(R) Core (TM) i5-7300HQ CPU 和 NVIDIA Tesla P100 的设备进行加速,批处理大小为4。本文使用初始学习率为0.000 1的Nadam优化器动态调整学习率,进而对网络进行优化。为验证本文算法有效性,选取两种不同的低光照图像数据集分别对网络进行训练与测试。训练数据集和测试数据集的具体细节如下所示。

LOL^[26]是从真实场景中获取的成对低光照图像数据集,其通过改变曝光时间和ISO获取弱光图像。LOL共500对低照度与正常照度图像,其中485对训练图像,15对测试图像,所有图像尺寸均为600×400。

MIT-Adobe FiveK^[27]是使用单反相机拍摄的真实场景中的有参考低光照图像,数据集共5 000幅不同亮度条件下的室内室外图像和分别由五个(A,B,C,D,E)专家手工修饰后的图像,本文将专家C处理后的图像作为参考图像。本文选取前4 500对图像作为训练集,其余500对图像作为测试集。

3.2 对比实验结果

本文选择低光照图像增强领域内16种具有代表性的算法分别在LOL数据集和MIT-Adobe 5K数据集上进行评估与分析,其中选取经典传统算法BPDHE^[2],CLAHE^[3],LIME^[5]和BIMEF^[8],选取深度学习领域代表性算法Retinex Net^[10],KinD^[11],URetinex Net^[12],DSLRL^[14],UFormer^[15],ALL-E^[16],DDNet^[17],Enlighten GAN^[20],Zero DCE^[21],Zero DCE++^[22],RUAS^[23]和SCI^[24]算法与本文算法进行对比分析。

3.2.1 实验结果主观评估与分析

首先在LOL数据集和MIT-Adobe 5K数据集上,将本文算法与对比算法增强后的视觉效果进行比较,不同算法在LOL数据集上的视觉效果对比如图5所示,不同算法在MIT-Adobe 5K数据集上的视觉效果对比如图6所示。



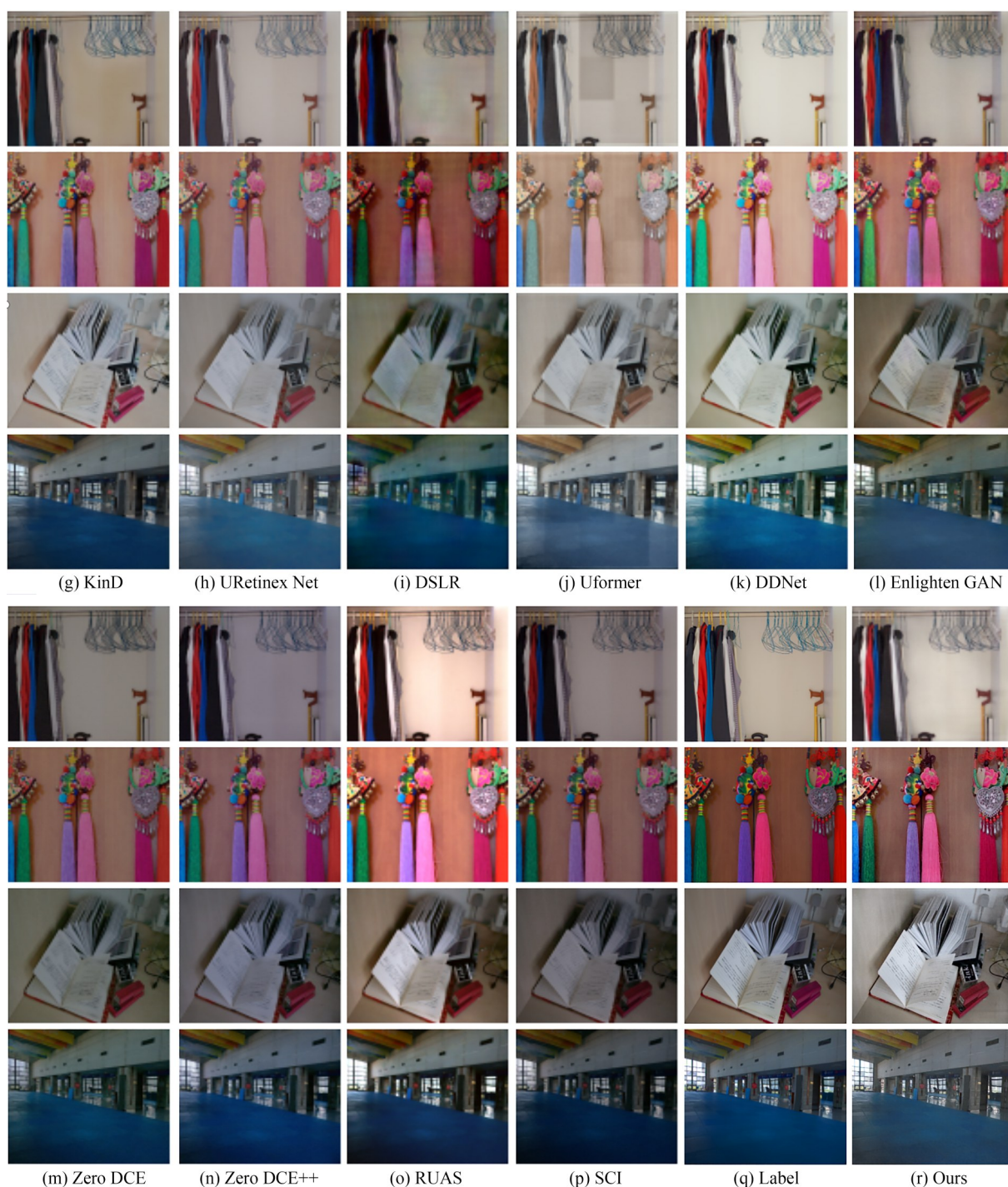


图 5 不同算法在 LOL 数据集上的视觉效果对比图

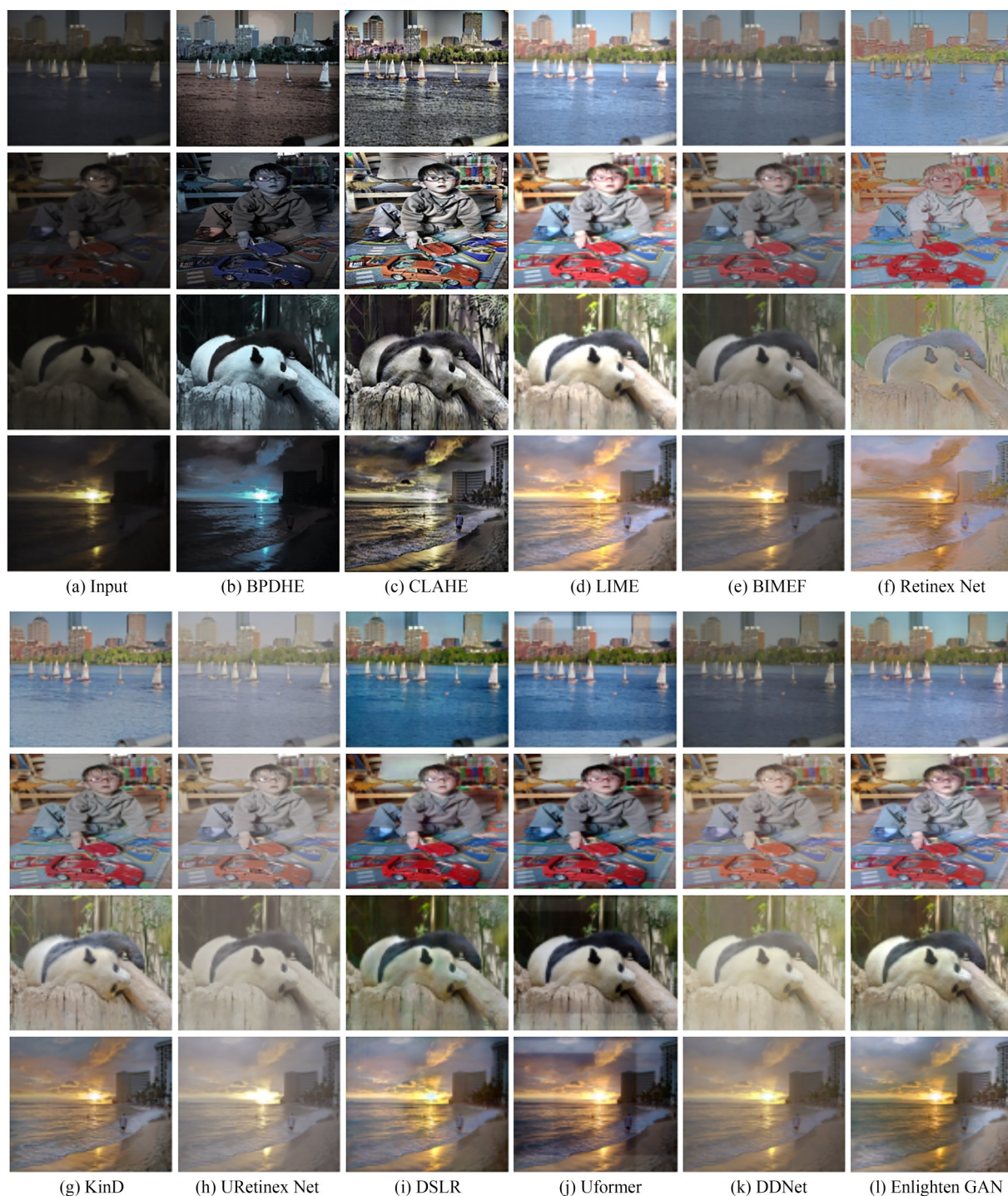
Fig. 5 Visual comparison of different algorithms on the LOL dataset

从图 5 和图 6 可以看出,传统方法 LIME, BPDHE, CLAHE 和 BIMEF 的图像增强效果无法在图像色彩、噪声和亮度方面做到很好的均衡, LIME 虽然对图像色彩还原较为明显但图像含有大量噪声,且 BPDHE 和 CLAHE 方法出现局部曝光现象。Retinex Net 和 Enlighten GAN

方法增强的图像噪声严重,且从图 6(f)小男孩面部和熊猫毛色可以明显看出 Retinex Net 方法出现了严重的色偏现象, KinD 方法增强的图像部分区域有伪影,且从图 5(g)衣柜背景板可以看出图像出现了色彩不均匀现象,视觉效果不自然。两幅图中 SCI, Zero DCE 和 Zero DCE++ 方法

增强的图像亮度均相对较暗,URetinex Net、Zero DCE 和 SCI 方法对图像色彩恢复较差,图 6(h)、图 6(m)和图 6(p)整体泛白,且 Zero DCE++方法在图像纹理细节方面有待加强。图 5(i)和图 6(o)方法增强的图像表面模糊类似有雾图,可以看出 DSLR 和 RUAS 方法对图像平滑过渡,造成图像纹理丢失,且图 6(o)中天空曝

光过度。URetinex Net,UFormer 和 ALL-E 方法对图像亮度的增强效果较好,但从图 5(j)中衣服、饰品和地面都失去了原有的颜色可知 UFormer 方法增强的图像整体效果混乱,且图中出现较多明显色块。URetinex Net 和 ALL-E 方法与本文所提方法都使图像达到了很好的增强效果,但本文方法在纹理细节增强和色彩还原方



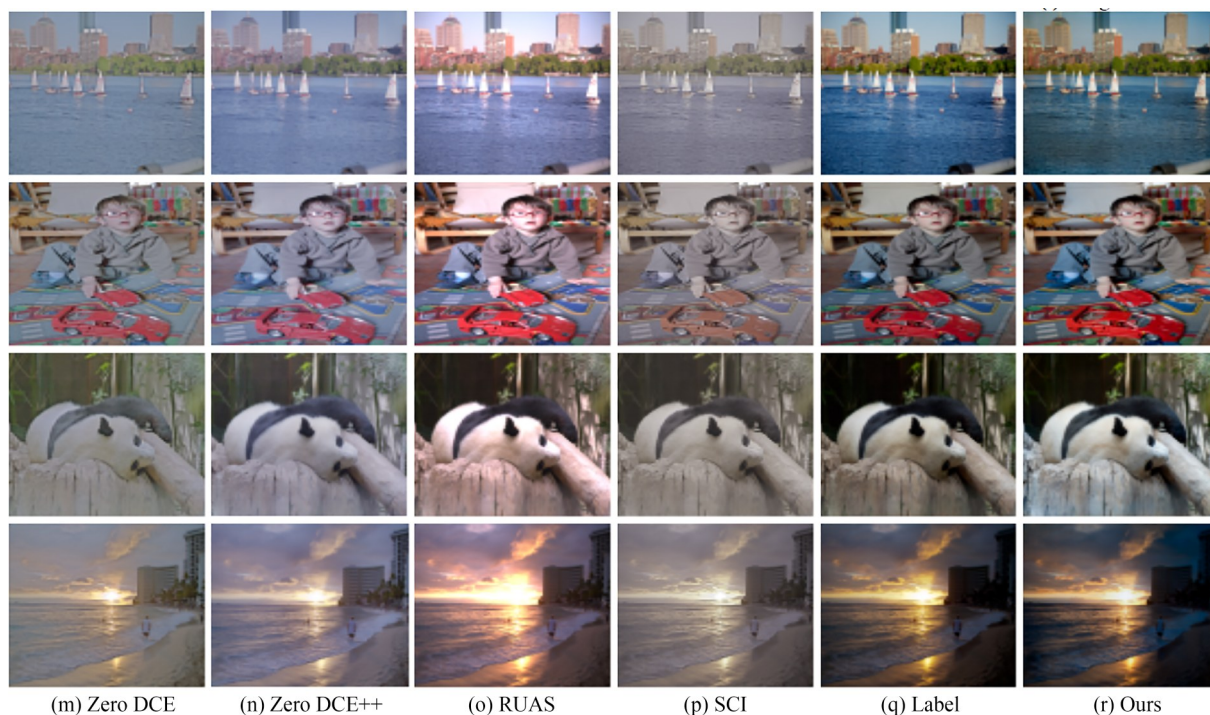


图 6 不同算法在 MIT-Adobe 5K 数据集上的实验结果对比图

Fig. 6 Visual comparison of different algorithms on the MIT-Adobe 5K dataset

面更胜一筹,且本文方法在噪声和伪影方面都有很好的抑制效果。整体来看,本文方法在主观感受方面效果更优。

3.2.2 实验结果客观评估与分析

为进一步测试本文所提方法,除主观视觉效果外,还采用峰值信噪比(Peak Signal-to-Noise Ratio, PSNR)、结构相似性(Structure Similarity Index Measure, SSIM)和图像感知相似度^[28](Learned Perceptual Image Patch Similarity, LPIPS)三种有参考图像质量评价指标和自然图像无参考评估器^[29](Natural Image Quality Evaluator, NIQE)在 LOL 数据集和 MIT-Adobe 5K 数据集上对模型进行定量评估。

PSNR 表示增强后图像的峰值功率与噪声功率的比值,PSNR 值越高,图像质量越好,表示网络能够较好抑制低光照图像的噪声并调整光照。SSIM 是衡量两张图像相似度的指标,SSIM 数值越大,增强图像质量越好,表明两张图像的纹理细节越相似。LPIPS 通过计算特征之间的距离估计纹理相似度,它可以更好地模拟人类视觉系统的感知能力,LPIPS 的值越小代表两张图像越相似。NIQE 是一种无需参考图像的评价指

标,NIQE 值越小,图像越自然,图像质量越高。

不同算法在 LOL 数据集和 MIT-Adobe 5K 数据集上的客观评价指标对比结果如表 1 和表 2 所示。表中各项指标值均为在测试集上求得平均值,其中加粗表示最优值,下划线表示次优值。

从表 1 和表 2 可以看出,与 BPDHE, CLAHE, LIME 和 BIMEF 这四种传统算法相比,本文算法优势明显。LIME 方法相比 BPDHE 方法在 PSNR 和 SSIM 指标对比中提升明显,说明 LIME 对图像纹理细节的增强效果较好,但该方法对图像色彩增强较为生硬,色彩不自然导致其 NIQE 指标评价较差。与其他基于深度学习的对比算法相比,本文算法提升显著,有较好的图像增强效果。Retinex Net 方法和 KinD 方法都是基于 Retinex 思想利用深度卷积神经网络实现图像增强,但 KinD 方法相比 Retinex Net 方法在性能上有显著增强,各项指标对比结果也明显提升,说明 KinD 方法对图像噪声抑制明显,图像色彩和亮度较为自然。URetinex Net 方法、ALL-E 方法和 DDNet 方法的评价指标值与本文所提方法的评价指标值相对较为接近,也位居

表 1 不同算法在 LOL 数据集上的客观评价结果。

Tab. 1 Objective evaluation results of different algorithms on the LOL dataset

方法	PSNR/ dB ↑	SSIM ↑	LPIPS ↓	NIQE ↓
BPDHE	10.936	0.415	0.439	7.985
CLAHE	13.414	0.551	0.361	6.636
LIME	16.758	0.564	0.395	8.058
BIMEF	13.875	0.575	0.326	7.699
RetinexNet	16.774	0.425	0.474	8.879
KinD	17.647	0.771	0.514	5.189
URetinex Net	19.841	0.826	0.235	4.722
DSLR	14.982	0.596	0.376	4.416
Uformer	18.547	0.721	0.321	4.443
ALL-E	18.216	0.763	0.219	4.200
EnlightenGAN	17.482	0.651	0.322	5.807
ZeroDCE	14.860	0.562	0.335	7.767
ZeroDCE++	16.570	0.591	0.328	5.311
RUAS	18.226	0.717	0.270	6.340
SCI	14.780	0.522	0.339	6.617
Ours	22.347	0.850	0.178	4.153

对比算法前列,说明这些方法增强的图像质量更好,图像更加清晰自然。综合对比表 1-2 中本文所提方法在 LOL 和 MIT-Adobe 5K 两种数据集上的各项指标可知,本文算法在有参考图像质量评价指标和无参考图像质量评价指标中均取得最优值,说明本文算法增强的图像与参考图像最为相似,能够较好恢复图像色彩与纹理细节,对图像噪声有很好的抑制作用,且增强的图像更接近自然,充分证明了本文所提算法的

表 2 不同算法在 MIT-Adobe 5K 数据集上的客观评价结果

Tab. 2 Objective evaluation results of different algorithms on the MIT-Adobe 5K dataset

方法	PSNR/ dB ↑	SSIM ↑	LPIPS ↓	NIQE ↓
BPDHE	12.855	0.572	0.441	5.280
CLAHE	15.955	0.608	0.375	5.353
LIME	13.303	0.749	0.319	4.172
BIMEF	17.968	0.797	0.298	4.598
RetinexNet	12.514	0.670	0.365	4.841
KinD	16.203	0.784	0.254	4.242
URetinex Net	14.184	0.754	0.242	3.867
DSLR	20.243	0.828	0.153	4.352
UFormer	21.917	0.870	0.185	3.961
DDNet	18.308	0.776	0.220	3.830
EnlightenGAN	17.905	0.836	0.238	3.865
ZeroDCE	15.931	0.766	0.219	3.882
ZeroDCE++	14.611	0.405	0.231	3.850
RUAS	15.995	0.786	0.141	4.048
SCI	17.477	0.840	0.216	4.122
Ours	22.703	0.903	0.137	3.822

有效性。

3.3 消融实验结果与模型时间对比

为优化增强图像质量,本文设计了 MDAR, PMFB, DHAB 和 CFAM 模块,并在 PMFB 中采用混合空洞卷积提升网络的图像增强性能。因此,为验证本文算法中各个模块的有效性,本小节在 MIT-Adobe 5K 数据集上对上述所提模块进行消融实验,分别去除各个模块的消融实验可视化效果如图 7 所示。

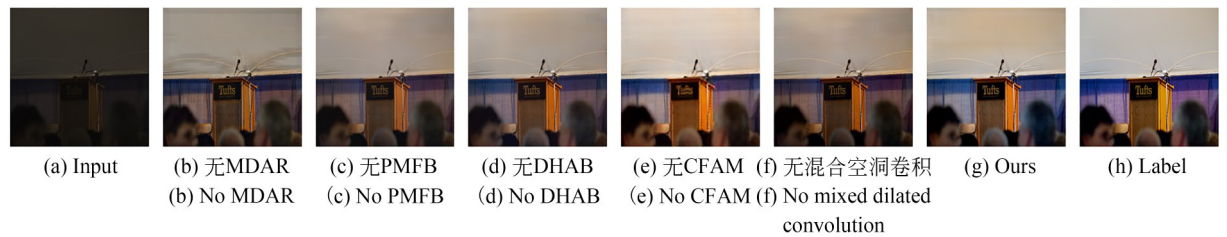


图 7 不同网络模块的消融实验可视化效果

Fig. 7 Ablation experiment visualization of different network modules

图 7(b)为去除 MDAR 模块的增强图像,从图中墙面纹理可以看出增强后图像纹理细节混乱,伪影严重,且图中窗帘有颜色失真现象,噪声明显。去除 PMFB 时,从图 7(c)可以看出墙面纹

理和演讲台文字牌模糊不清,说明增强后图像丢失细节现象严重,也充分证明了 PMFB 模块对图像细节增强的有效性。去除 DHAB 时,从图 7(d)可以看出图像在噪声抑制和色彩增强方面有

所欠缺,说明 DHAB 能有效提高网络对图像色彩的关注度,削弱噪声对图像的影响。去除 CFAM 时,从图 7(e)可以看出图像表面有轻微细纹,演讲台文字牌上的文字出现扭曲模糊现象,整幅图像出现不同程度的细节缺失,说明 CFAM 模块能多方位聚合不同层次信息,有效避免信息流失。图 7(g)为本文所提方法增强的图像效果图,图像整体效果与图 7(h)参考图像更加接近,图像细节纹理和色彩对比度较为自然,对噪声有一定程度的抑制作用。

为进一步验证本文模型的有效性,采用 PSNR,SSIM 指标评估去除单一模块后的网络性能。不同网络模块的评价指标对比结果如表 3 所示。

从表 3 的评价指标对比结果可以看出,去除任一模块的网络模型都会对图像增强效果产生一定程度的影响,充分证明了网络中每一个模块的有效性。

此外,本文在模型运行时间方面对所提方法进行了测试,从时间指标也可以看出,任一模块

表 3 不同网络模块的评价指标对比结果
Tab. 3 Comparison results of evaluation indicators of different network modules

模块	PSNR/dB ↑	SSIM ↑	Time/s ↓
无 MDAR	21.505	0.861	0.100
无 PMFB	21.862	0.878	0.108
无 DHAB	22.512	0.896	0.109
无 CFAM	21.173	0.732	0.106
无混合空洞卷积	21.901	0.890	0.110
Ours	22.703	0.903	0.112

的增删对模型整体运行时间并没有很大影响,说明模型在进行多尺度采样以及特征融合时是相对高效的,进而说明每个模块的高效性。同时,本文与不同主流方法也进行了对比,对比结果如表 4 所示。从表 4 可以看出,本文方法在运行时间方面并不占优势,其原因在于本文模型在深浅层特征间进行多分支跨层次特征融合时,有效利用了提取到的丰富特征,尽管在图像质量增强方面占有更大优势,但一定程度上增加了模型工作量,加大了模型运行时间。

表 4 不同方法的平均增强时间
Tab. 4 Average enhancement time of different method

	BPDHE	CLAHE	LIME	BIMEF	Retinex Net	KinD	URetinex Net	DSLRL
	0.1041	0.1156	0.4914	0.1280	0.1200	0.0500	0.400	0.125
Running time ↓	UFormer	ALL-E	Enlighten GAN	Zero DCE	Zero DCE++	RUAS	SCI	Ours
	0.121	1.055	0.057	0.003	0.0019	0.006	0.001	0.112

4 结 论

本文针对低照度图像亮度低、噪声大、颜色偏差和细节纹理丢失等问题,提出一种双路混合注意力的跨层次特征聚合图像增强方法。首先,设计 MDAR 模块,MDAR 模块包括 PMFB 和 DHAB。将混合空洞卷积引入到 PMFB,利用 PMFB 挖掘低照度图像的多尺度特征信息,充分还原图像细节信息和亮度信息,再利用 DHAB 缓解不同注意力间的特征差异,通过特征加权的方式筛选出包含图像亮度信息和颜色信息的重要特征,增强网络对任务相关信息的敏感度。最后,设计 CFAM,将多尺度特征以不同权重灵活

地聚合在一起,弥补深层特征与浅层特征之间的差异,更好地恢复图像纹理信息和颜色信息,实现图像增强。实验结果表明,所提方法在 LOL 数据集上的 PSNR,SSIM,LPIPS 和 NIQE 分别达到了 22.347 dB,0.850,0.178 和 4.153;在 MIT-Adobe 5K 数据集上的 PSNR,SSIM,LPIPS 和 NIQE 分别达到了 22.703 dB,0.903,0.137 和 3.822。相较于其他先进算法,本文算法不仅有效提高了图像质量,而且在图像细节增强、噪声抑制和色彩重建等方面具有一定优势。实时进行低光图像增强的需求日益增大,如何提高模型运行速度将会是后续研究的重点方向。

参考文献:

- [1] PISANO E D, ZONG S Q, HEMMINGER B M, *et al.* Contrast Limited Adaptive Histogram Equalization image processing to improve the detection of simulated spiculations in dense mammograms [J]. *Journal of Digital Imaging*, 1998, 11(4): 193.
- [2] IBRAHIM H, KONG N S P. Brightness preserving dynamic histogram equalization for image contrast enhancement[J]. *IEEE Transactions on Consumer Electronics*, 2007, 53(4): 1752-1758.
- [3] ZUIDERVELD K. *Contrast Limited Adaptive Histogram Equalization*[M]. Graphics Gems. Amsterdam: Elsevier, 1994: 474-485.
- [4] LAND E H, MCCANN J J. Lightness and retinex theory[J]. *Journal of the Optical Society of America*, 1971, 61(1): 1-11.
- [5] GUO X J, LI Y, LING H B. LIME: low-light image enhancement via illumination map estimation [J]. *IEEE Transactions on Image Processing*, 2017, 26(2): 982-993.
- [6] ZHANG Q, NIE Y W, ZHU L, *et al.* Enhancing underexposed photos using perceptually bidirectional similarity [J]. *IEEE Transactions on Multimedia*, 2045, 23: 189-202.
- [7] LI M D, LIU J Y, YANG W H, *et al.* Structure-revealing low-light image enhancement via robust retinex model[J]. *IEEE Transactions on Image Processing*, 2018, 27(6): 2828-2841.
- [8] YING Z Q, LI G, GAO W. A bio-inspired multi-exposure fusion framework for low-light image enhancement [J]. *ArXiv e-Prints*, 2017: arXiv: 1711.00591.
- [9] LORE K G, AKINTAYO A, SARKAR S. LL-Net: a deep autoencoder approach to natural low-light image enhancement[J]. *Pattern Recognition*, 2017, 61: 650-662.
- [10] WEI C, WANG W J, YANG W H, *et al.* Deep Retinex Decomposition for Low-Light Enhancement [EB/OL]. 2018: arXiv: 1808.04560. <http://arxiv.org/abs/1808.04560>
- [11] ZHANG Y H, ZHANG J W, GUO X J. Kindling the darkness: a practical low-light image enhancer [C]. *Proceedings of the 27th ACM International Conference on Multimedia. Nice France*. ACM, 2019: 1632-1640.
- [12] WU W H, WENG J, ZHANG P P, *et al.* URet-
inex-Net: retinex-based deep unfolding network for low-light image enhancement [C]. 2022 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. New Orleans, LA, USA. IEEE, 2022: 5891-5900.
- [13] 王殿伟, 邢质斌, 韩鹏飞, 等. 基于模拟多曝光融合的低照度全景图像增强[J]. *光学精密工程*, 2021, 29(2): 349-362.
WANG D W, XING Z B, HAN P F, *et al.* Low illumination panoramic image enhancement algorithm based on simulated multi-exposure fusion [J]. *Opt. Precision Eng.*, 2021, 29(2): 349-362. (in Chinese)
- [14] LIM S, KIM W. DSLR: deep stacked Laplacian restorer for low-light image enhancement [J]. *IEEE Transactions on Multimedia*, 2021, 23: 4272-4284.
- [15] WANG Z D, CUN X D, BAO J M, *et al.* Uformer: a general u-shaped transformer for image restoration[C]. 2022 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. New Orleans, LA, USA. IEEE, 2022: 17662-17672.
- [16] LI L, LIANG D, GAO Y, *et al.* ALL-E: Aesthetics-Guided Low-Light Image Enhancement [J]. *ArXiv preprint*, 2023, arXiv:2304.14610.
- [17] QU J X, LIU R W, GAO Y, *et al.* Double domain guided real-time low-light image enhancement for ultra-high-definition transportation surveillance [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2024, PP(99): 1-13.
- [18] 刘光辉, 杨琦, 孟月波, 等. 一种并行混合注意力的渐进融合图像增强方法[J]. *光电工程*, 2023, 50(4): 220231.
LIU G H, YANG Q, MENG Y B, *et al.* A progressive fusion image enhancement method with parallel hybrid attention[J]. *Opto-Electronic Engineering*, 2023, 50(4): 220231. (in Chinese)
- [19] 陈清江, 顾媛. 多通道融合注意力网络的低照度图像增强[J]. *光学精密工程*, 2023, 31(14): 2111-2122.
CHEN Q J, GU Y. Low-light image enhancement algorithm based on multi-channel fusion attention network[J]. *Opt. Precision Eng.*, 2023, 31(14): 2111-2122. (in Chinese)
- [20] JIANG Y F, GONG X Y, LIU D, *et al.* EnlightenGAN: deep light enhancement without paired su-

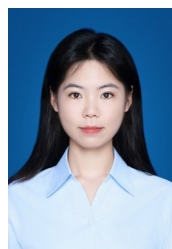
- pervision [J]. *IEEE Transactions on Image Processing*, 2021, 30: 2340-2349.
- [21] GUO C L, LI C Y, GUO J C, *et al.* Zero-reference deep curve estimation for low-light image enhancement [C]. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Seattle, WA, USA. IEEE, 2020: 1777-1786.
- [22] LI C Y, GUO C L, LOY C C. Learning to enhance low-light image via zero-reference deep curve estimation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, 44(8): 4225-4238.
- [23] LIU R S, MA L, ZHANG J A, *et al.* Retinex-inspired unrolling with cooperative prior architecture search for low-light image enhancement [C]. *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Nashville, TN, USA. IEEE, 2021: 10556-10565.
- [24] MA L, MA T Y, LIU R S, *et al.* Toward fast, flexible, and robust low-light image enhancement [C]. *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. New Orleans, LA, USA. IEEE, 2022: 5627-5636.
- [25] DING X H, ZHANG X Y, HAN J G, *et al.* Scaling up Your Kernels to 31×31 : Revisiting large kernel design in CNNs [C]. *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. New Orleans, LA, USA. IEEE, 2022: 11953-11965.
- [26] WEI C, WANG W J, YANG W H, *et al.* *Deep Retinex Decomposition for Low-Light Enhancement* [EB/OL]. 2018: *arXiv*: 1808.04560. <http://arxiv.org/abs/1808.04560>
- [27] BYCHKOVSKY V, PARIS S, CHAN E, *et al.* Learning photographic global tonal adjustment with a database of input/output image pairs [C]. *CVPR*. Colorado Springs, CO, USA. IEEE, 2011: 97-104.
- [28] ZHANG R, ISOLA P, EFROS A A, *et al.* The unreasonable effectiveness of deep features as a perceptual metric [C]. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City, UT, USA. IEEE, 2018: 586-595.
- [29] MITTAL A, SOUNDARARAJAN R, BOVIK A C. Making a “completely blind” image quality analyzer [J]. *IEEE Signal Processing Letters*, 2013, 20(3): 209-212.

作者简介:



袁 姮(1988—),女,湖北黄冈人,博士,副教授,硕士生导师,2010年、2013年、2017年于辽宁工程技术大学分别获得学士学位、硕士学位和博士学位,主要从事人工智能、模式识别、计算机视觉等方面的研究工作。E-mail: yuanheng@lntu.edu.cn

通讯作者:



王笑雪(1999—),女,辽宁阜新人,硕士研究生,2021年于辽宁工程技术大学获得学士学位,主要从事图像与视觉信息计算、模式识别与人工智能方面的研究。E-mail: wx2585354184@163.com