

文章编号:1007-2780(2024)06-0856-11

跨级特征自适应融合的暗光图像增强算法

梁礼明, 朱晨锷*, 阳 渊, 李仁杰

(江西理工大学 电气工程与自动化学院, 江西 赣州 341000)

摘要:针对低照度环境下采集的图像存在亮度低、对比度低和视觉效果不佳等问题,提出一种跨级特征自适应融合的低照度图像增强算法。该算法首先结合分级采样和广域卷积搭建网络前端,生成大面积感受野的多尺度特征,使浅层信息被充分挖掘。其次引入多头转置注意力模块嵌于网络中部,计算通道间的互协方差以生成注意力映射,隐式地建立全局上下文信息关联。然后构建联合损失函数修正模型收敛方向,辅助模型从对比度和结构等方面进行优化,提高算法的鲁棒性。最后在 LOL 和 LOLv2 数据集上进行实验验证,实验结果表明,所提算法在峰值信噪比 (PSNR) 和结构相似性 (SSIM) 等客观指标上整体优于大部分先进算法,主观上图像亮度自然,噪声和伪影得到有效抑制。

关键词:低照度图像;广域卷积;多尺度;多头转置注意力;联合损失函数

中图分类号:TP391.41 **文献标识码:**A **doi:**10.37188/CJLCD.2023-0210

Cross-level feature adaptive fusion network for low-light image enhancement

LIANG Liming, ZHU Chenkun*, YANG Yuan, LI Renjie

(School of Electrical Engineering and Automation, Jiangxi University of Science and Technology, Ganzhou 341000, China)

Abstract: Aiming at the problems of low brightness, low contrast and poor visual effect in images collected in low-light environment, a low-light image enhancement algorithm based on cross-level adaptive feature fusion is proposed. Firstly, a network frontend is built by combining hierarchical sampling and large receptive field convolution to generate multi-scale features of large-area receptive fields, so that shallow information mining can be fully carried out. Secondly, a multi-head transposed attention module embedded in the middle of the network is introduced, the cross-covariance between channels is calculated to generate attention maps, and global context information associations are implicitly established. Thirdly, a joint loss function is constructed to correct the convergence direction of the model, assist the model optimized from the perspective of contrast and structure, and improve the robustness of the

收稿日期:2023-06-09;修订日期:2023-07-14.

基金项目:国家自然科学基金(No. 51365017, No. 61463018);江西省自然科学基金面上项目(No. 20192BAB205084);江西省教育厅科学技术研究重点项目(No. GJJ170491)

Supported by National Natural Science Foundation of China (No. 51365017, No. 61463018); General Project of Natural Science Foundation of Jiangxi Province (No. 20192BAB205084); Science and Technology Research Key Project of Education Department of Jiangxi Province (No. GJJ170491)

*通信联系人, E-mail: 281392667@qq.com

algorithm. Relevant experiments are carried out on the LOL and LOLv2 datasets. The experimental results show that the proposed algorithm outperforms most advanced algorithms in terms of objective indicators such as peak signal-to-noise ratio (PSNR) and structural similarity (SSIM). Subjectively, the image brightness is natural and the noise is low, and artifacts are effectively suppressed.

Key words: low-light image; large receptive field convolution; multi-scale; transformer; joint loss function

1 引言

图像采集过程中,由于相机的物理限制或者不适当的光照条件,通常会引入不同程度的图像退化,产生低亮度、低对比度、纹理不清晰和带有噪声的图像,这类图像将对后续视觉任务带来阻碍。图像增强作为底层计算机视觉任务的一个分支,发挥着重要作用。因此,研究背光图像的退化原理,提高其视觉效果,具有极大的研究价值和应用潜力。

目前,传统的低照度图像增强方法可分为3类:基于直方图均衡化算法、基于Retinex算法和基于图像去雾算法。基于直方图均衡化^[1]算法通过均匀调整图像动态范围提高图像亮度。Ibrahim等^[2]利用亮度保持的动态直方图均衡化,计算局部最大值划分不同区域并分配新的动态范围,再分别进行直方图均衡化,但该方法丢失了大量的纹理信息。基于Retinex算法通过模拟人的视觉系统,将图像分解为反射分量和照明分量,对照明分量进行调整,实现亮度和对比度同时增强。如Jobson等^[3]构建了单尺度Retinex算法。Wang等^[4]设计了一种利用亮度顺序误差和对数变换的图像增强方法,缓解了Retinex算法对亮度增强不均匀等缺陷。Guo等^[5]提出一种运用光照估计的图像增强方法,先计算RGB 3个通道的最大值,然后附加先验结构细化初始光照图。该方法虽有效提升了亮度,但会造成图像整体对比度不均匀。基于图像去雾的算法^[6]将转置运算后的低照度图像视作雾图,对其运用去雾算法实现图像增强。Li等^[7]基于该思想将图像划分成不同区域并进行降噪处理,有效提升了图像清晰度。传统的图像增强算法对图像亮度的增强成效显著,但需要先验参数对算法进行调整,且图像纹理细节丢失等问题仍比较突出。

卷积神经网络因其优异的性能以及不依赖

显式先验信息等优势,被广泛应用于计算机底层视觉任务中。Lore^[8]等构建了一种堆叠的稀疏去噪自动编码器,开创了卷积神经网络应用在图像增强领域的先河。Zhang等^[9]提出一种结合Retinex理论的多阶段网络,并设计映射函数替换伽马变换,提高了亮度调整的灵活性,但增强后的图像存在伪影现象。Guo等^[10]设计一种将图像增强任务转换成亮度估计曲线的算法,该算法可以在零参考图像的情况下进行端对端训练,但增强后的图像噪声方差较大。Zamir^[11]等构建了一种多尺度特征融合网络,利用分级采样保留了高分辨率特征的空间细节,并采取稠密的跳跃连接手段复用不同分支的上下文信息,由粗到细地实现图像恢复,但网络结构的复杂度较大。Zhou等^[12]提出一种轻量级网络,通过设计自适应环绕函数结合数据驱动实现图像增强。Cotogni等^[13]创新性地将树搜索理论运用到图像增强算法中,不仅能推理出暗光图和增强图像之间的映射关系,而且有效抑制了伪影的产生。Wu等^[14]引入一种基于Retinex理论深度展开模型,将优化问题展开成具有学习能力的网络。该方法虽有效地保持了图像纹理细节,但对图像颜色的还原仍有提升空间。

基于深度学习的方法较传统方法在照度和细节恢复等方面更具优势,但仍存在图像伪影、色彩还原度不足和噪声等缺陷。针对现有算法的不足,本文提出一种跨级特征自适应融合的低照度图像增强算法。主要贡献有:

(1)构建像素感知模块作为网络的前端,利用大核卷积的优势扩大感受野,提取丰富的特征图信息。

(2)引入多头转置注意力模块以并行连接的形式嵌入网络中部,建立局部与全局特征的联系和抑制冗余信息。

(3)设计联合损失函数约束训练过程,辅助

模型从结构和对比度等方面优化模型误差,提高模型非线性拟合能力。

2 网络结构

本文提出一种跨级特征自适应融合网络(Cross-level Feature Adaptive Fusion Network, CMA-Net)用于低照度图像增强。CMA-Net主要由多级特征生成部分、特征重建部分和特征融

合部分构成,如图 1 所示。其中多级特征生成部分运用像素感知模块生成不同尺度的特征图组成网络前端,加强像素信息的表达和提高网络适应性。特征重建部分将多头转置注意力模块以串联和并联的形式嵌入网络中部,有利于强化图像局部与全局的像素联系和抑制冗余信息。特征融合部分将自适应特征融合模块在网络末端以递归的形式堆叠而成,目的是减少不同层级特征语义差异,动态融合各个分辨率图像的特征信息。

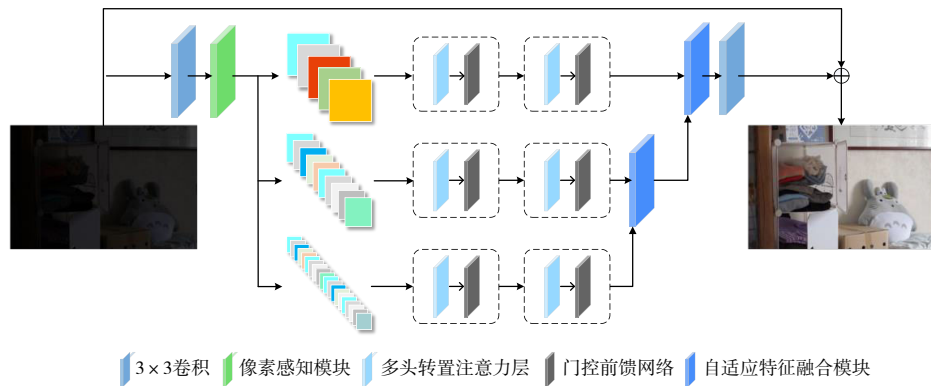


图 1 跨级特征自适应融合网络

Fig. 1 Cross-level feature adaptive fusion network

2.1 像素感知模块

低照度图像整体信息隐晦,常规卷积感受野受限与卷积核的大小,无法捕捉长距离依赖信息,对邻域特征的提取造成了阻碍。文献[15]结合深度可分离卷积和空洞卷积的优点改进了像素注意力模块,有效增大了感受野。受文献[11,15]启发,本文提出一种像素感知模块用于提取低照度图像特征,如图 2 所示。该模块包含采样子模块和像素注意力子模块,二者依次进行多级特征图生成和扩大感受野。首先输入图像流向 $\log_2 N$ (N 分别取 1、2、4) 个采样子模块生成对应下采样倍数的浅层特征 F ,采样方式为抗锯齿

采样^[16],采样子模块内部利用跳跃连接缓解梯度弥散。特征图依次经过点卷积、 5×5 的深度可分离卷积和空洞卷积,利用线性运算低廉的计算成本整合通道之间的信息,同时运用大核卷积的性质稀疏连接广泛邻域的像素信息,从而在不大幅度增加参数的前提下扩大感受野。最后将多形式卷积运算结果与 F 在末端做乘积操作,得到模块输出 Y 。给定输入 X ,其流程如式(1)和式(2)所示:

$$F_{shar} = \text{Conv}_{1 \times 1} \left(\Phi \left(\text{Conv}_{3 \times 3} \left(\text{Conv}_{1 \times 1} (X) \right) \right) \right) + \text{Conv}_{1 \times 1} \left(\Phi (X) \right), \quad (1)$$

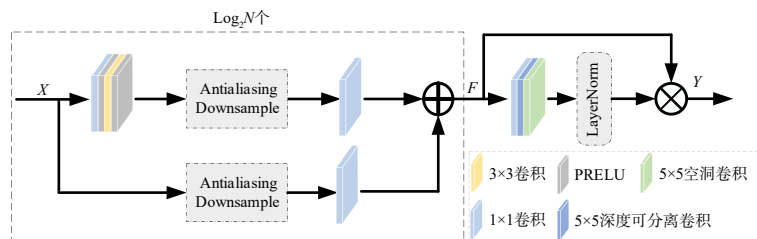


图 2 像素感知模块

Fig. 2 Pixel perception block

$$Y = \left[\text{DCConv}_{5 \times 5} \left(\text{DWConv}_{5 \times 5} \left(\text{Conv}_{1 \times 1} \left(F^{\log_2 N} \right) \right) \right) \right] \cdot F^{\log_2 N}, \quad (2)$$

其中:Conv_{1×1}、Conv_{3×3}、DWConv_{5×5}和DCConv_{5×5}分别为1×1卷积、3×3卷积、5×5深度可分离卷积和5×5空洞卷积,Φ为抗锯齿采样,N为采样倍数。

2.2 多头转置自注意力模块

由于低照度图像存在无用信息冗余和有效信息的表征能力被削弱的现象,导致图像内部特征信息交互受到限制,不利于图像细节的重建。为过滤赘余信息和建立图像局部与全局特征的长期依赖关系,本文引入一种高效的多头转置注意力模块^[17]建模全局信息。该模块由多头转置注意力层(Multi-Dconv Head Transposed Attention,MDTA)和门控前馈网络(Gated-Dconv Feed-Forward Network,GDFN)组成。

2.2.1 MDTA

MDTA结构如图3所示。给定经过层标准

化后的张量输入 $Y \in R^{\hat{H} \times \hat{W} \times \hat{C}}$,首先通过1×1卷积聚合通道之间的像素上下文信息,其次通过3×3深度可分离卷积编码通道之间的空间上下文信息,两者利用卷积运算互补的优势,强化局部空间的表征能力,进而生成查询($Q = W_d^Q \cdot W_p^Q \cdot Y$)、索引($K = W_d^K \cdot W_p^K \cdot Y$)和键值($V = W_d^V \cdot W_p^V \cdot Y$)。其中 W_d 表示3×3深度可分离卷积, W_p 表示1×1卷积。最后运用点积运算重塑Q和K的投影,Q和K相互作用生成大小为 $R^{\hat{C} \times \hat{C}}$ 的转置注意力映射A,隐式地建模像素之间的全局关系。MDTA总流程概括为:

$$\hat{X} = W_p \text{Attention}(\hat{Q}, \hat{K}, \hat{V}) + X, \quad (3)$$

$$\text{Attention}(\hat{Q}, \hat{K}, \hat{V}) = \hat{V} \cdot \text{Softmax}(\hat{K} \cdot \hat{Q} / \alpha), \quad (4)$$

其中: X 和 \hat{X} 分别表示输入和输出特征图, $\hat{Q} \in R^{\hat{H} \times \hat{W} \times \hat{C}}$ 、 $\hat{K} \in R^{\hat{C} \times \hat{H} \times \hat{W}}$ 和 $\hat{V} \in R^{\hat{H} \times \hat{W} \times \hat{C}}$ 分别由原始输入 $R^{\hat{H} \times \hat{W} \times \hat{C}}$ 分解而成, α 为控制 \hat{Q} 、 \hat{K} 点积大小的参数。

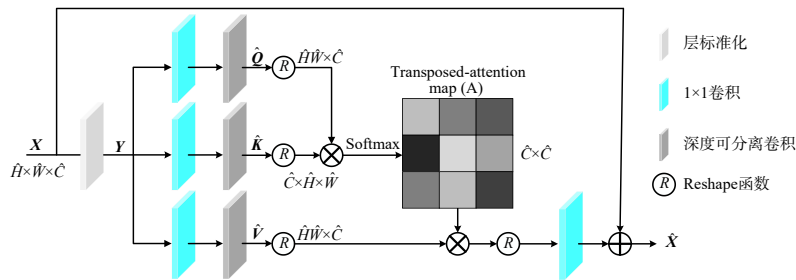


图3 多头转置注意力层

Fig. 3 Multi-Dconv head transposed attention

2.2.2 GDFN

GDFN结构如图4所示,由常规的前馈网络^[18]改进而来。一是附加门控机制,二是引入深度可分离卷积替换普通卷积。其中门控机制表现为两条并行路径线性变换层的乘积,其中一路

调用了GELU激活函数^[19]实现控制信息流动。与MDTA相似之处在于GDFN同样使用了深度可分离卷积编码邻域像素位置信息,优化空间上下文信息的联系,有助于图像结构的恢复。给定张量输入 $X \in R^{\hat{H} \times \hat{W} \times \hat{C}}$,GDFN定义为:

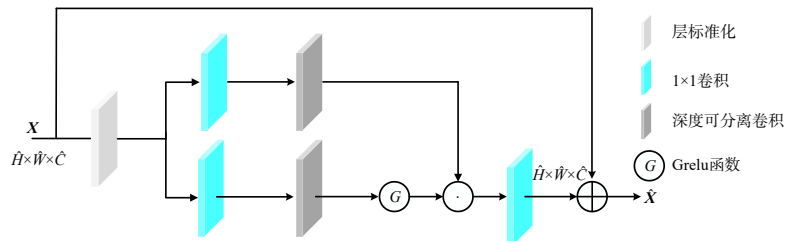


图4 门控前馈网络

Fig. 4 Gated-Dconv feed-forward network

$$\hat{X} = W_p^0 \text{Gating}(X) + X, \quad (4)$$

$$\text{Gating}(X) = \phi(W_d^1 W_p^1 (\text{LN}(X))) \odot W_d^2 W_p^2 (\text{LN}(X)), \quad (5)$$

其中: \odot 表示逐元素累乘, ϕ 表示 GELU 函数, LN 表示层标准化。

与 MDTA 不同之处在于 GDFN 通过控制分支中各个层次的信息流达到信息交互和互补的目的, 而 MDTA 运用并行结构和多形式卷积丰富上下文信息。

2.3 自适应特征融合模块

多尺度网络常用的特征融合手段有拼接与求和, 该类方法具备运算简易和高效的特点, 但文献[20]指出该类运算限制了网络的特征表达能力, 不利于模型性能的提升。为充分利用网络提取的优质特征, 在本文网络框架中,

引入一种自适应特征融合模块(Selective Kernel Feature Fusion, SKFF)^[11], 运用自注意力机制非线性融合不同层级特征, 模块结构如图 5 所示。SKFF 可视作两个阶段。第一阶段为特征聚合(Fuse): 首先依次对张量采用拼接、重塑和累加操作进行合并, 过程定义为: $L = L_1 + L_2$ 。其次运用全局平均池化从空间维度上聚合特征 $L \in R^{H \times W \times C}$ 得到通道信息 $s \in R^{1 \times 1 \times C}$ 。最后采用逐点卷积强化特征相关性, 生成张量 $z \in R^{1 \times 1 \times r}$, 其中 r 为 C 的 $1/8$ 。第二阶段为加权融合(Select), 张量 z 分两路前向传播, 经过 1×1 卷积分别生成特征描述符 v_1 和 v_2 ($v_1, v_2 \in R^{1 \times 1 \times C}$)。其次利用 softmax 函数计算得到注意力权值 s_1 和 s_2 。最后 L_1 和 L_2 分别和注意力权值相乘, 在末端求和得到模块输出。总流程可概括为: $U = L_1 \cdot s_1 + L_2 \cdot s_2$ 。

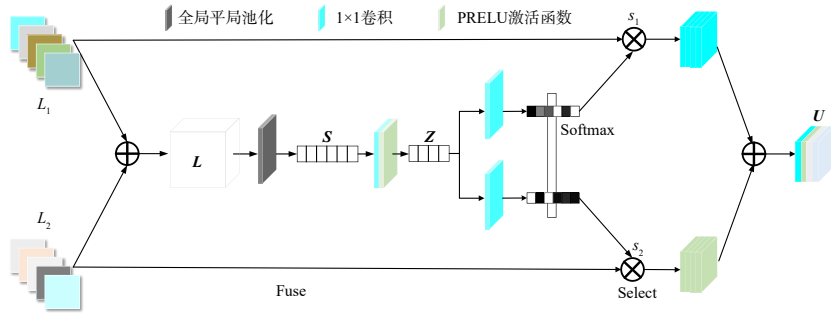


图 5 自适应特征融合模块

Fig. 5 Selective kernel feature fusion module

2.4 损失函数

Charbonnier 损失函数因具备简易运算和收敛性强等优势, 被广泛运用在图像恢复领域。其表达式为:

$$L_{\text{con}} = \sqrt{(I_{\text{label}} - I_{\text{predict}})^2 + \epsilon^2}, \quad (6)$$

其中: I_{label} 和 I_{predict} 分别为标签和样本, ϵ 设定为常数, 其值为 0.001。

但 Charbonnier 损失函数仅对像素之间进行误差计算, 并未考虑到全局信息, 因此通常会出现过度平滑问题。为使图像在恢复的同时进一步提升图像高频细节的真实性, 文献[21]提出在 Charbonnier 损失函数的基础上附加边缘损失函数, 用于约束标签与样本之间的高频分量, 其定义为:

$$L_{\text{edge}} = \sqrt{(\text{Lap}(I_{\text{label}}) - \text{Lap}(I_{\text{predict}}))^2 + \epsilon^2}, \quad (7)$$

其中: $\text{Lap}(I_{\text{label}})$ 和 $\text{Lap}(I_{\text{low}})$ 分别表示拉普拉斯算子对 I_{label} 和 I_{low} 提取的边缘特征; ϵ 是一个常数, 通常为 0.001。

上述损失函数的组合在图像恢复任务中有良好的功效, 但对恢复图像结构等隐含信息的能力有限。为进一步提升图像的视觉效果, 本文引入结构相似性损失函数^[22]辅助模型从对比度和结构等方向进行优化。综上, 将三者组合构造成联合损失函数:

$$\text{Loss} = L_{\text{con}}(X, Y) + a \cdot L_{\text{edge}}(X, Y) + L_{\text{SSIM}}(X, Y), \quad (8)$$

其中: X 和 Y 分别表示样本和标签, a 为常数, 其在参考文献[21]设定为 0.05。

3 实验结果分析

3.1 实验设计

为验证模型的可行性,本文基于Pytorch和Ubuntu16.04系统搭建实验环境,CPU型号为Intel Core TM i7-6700H,16G内存,GPU型号为Nvidia GeForce GTX2070,CUDA版本为10.1。选用Wei等^[23]提出的LOL和LOLv2数据集验证算法,其中LOL数据集包含低照度图像和与之对应的正常照度图像共500对,其中485对用作训练,15对用作测试。LOLv2数据集包含1589对真实场景和合成图像训练集,100对真实图像测试集。图像分辨率均为 400×600 。对训练集采用裁剪、随机旋转和镜像等数据增强方法。选取Adam优化器,批次数量为2,批次大小为 256×256 ,训练次数为100轮,初始学习率为 $1e-4$,终

止学习率为 $1e-6$ 。

3.2 评价指标

本文选用峰值信噪比(Peak Signal to Noise Ratio,PSNR)、结构相似性(Structural Similarity,SSIM)、图像感知相似度(Learned Perceptual Image Patch Similarity,LPIPS)和均方误差(Mean Square Error,MSE)指标度量模型性能。其中PSNR值越高则图像失真程度越低;SSIM结合图像亮度、对比度和结构信息评估图像质量,其值的大小反映两张图像相似程度的高低;LPIPS指标更加符合人眼视觉习惯,其值与图像之间的差异成正相关。

3.3 实验结果定性分析

为验证本文算法的先进性,选取直方图均衡化算法BPDHE、Retinex算法LIME、去雾算法Dong和基于深度学习的算法kinD、ZERO-DCE、



图6 LOL和LOLv2数据集上的对比结果

Fig. 6 Comparative results on LOL and LOLv2 datasets

DCCNet^[24]、SurroundNet、LLformer^[25]与本文算法进行对比。将上述算法置于同一实验环境,在LOL和LOLv2数据集上的实验结果如图6所示,其中第1行和最后1行分别为输入样本和标签,2~10行为不同算法结果,(i)和(ii)列图像为LOL数据集增强结果,(iii)和(iv)列图像为LOLv2数据集增强结果。

BPDHE和Dong算法增强图像特征信息隐晦,整体可视度较差,如图6中第2行(i)、(ii)和(iii)图左侧内容难以辨认。本文算法通过设计像素感知模块提取丰富的特征,增强图像蕴含的信息较对比算法更为全面。LIME算法照度明显提升,但亮度过于饱和且引入了大量的噪声,增强图像不符合人眼视觉习惯,如图6中第4行(i)图左侧亮度不自然并且存在大量噪声。本文算法引入结构相似性损失函数辅助模型从视觉感官等方向进行优化,增强图像亮度更接近于参考图像。ZERO-DCE算法和KinD算法抑制了色彩过于饱和的现象,但同时引入了噪声等冗余信息并造成图像局部失真,如图6中第5行(i)、(ii)和(iv)图红框细节处存在部分噪声,第6行(iii)图存在明显失真和伪影现象。本文算法引入多头转置注意力模块,利用其门控机制筛选出优质特征,提高了图像可读性。DCCNet、SurroundNet、LLformer在增强亮度和还原色彩方面表现出色,整体视觉效果良好。但DCCNet算法增强图像丢失了部分信息,如图6第7行(iii)图右上角出现了不属于原图的光圈,SurroundNet和LLformer

算法增强图像细节处稍显模糊且存在少量噪声,如图6中第8、9行(i)和(ii)图的红框细节处。综上,本文算法相较于对比算法在提升亮度和对比度之间取得了更好的平衡,增强图像的纹理细节得到了有效保持。如图6中第10行(i)图的玩具娃娃面部纹理、(ii)图各色衣物的对比度、(iii)图的彩色环靶和白色墙壁以及(iv)图的窗户和全局亮度。这些图像更接近于参考图像,相较于其他算法有更佳的视觉感受。

3.4 实验结果定量分析

本文选用PSNR、SSIM、LPIPS和MSE评价指标与其他算法进行定量对比,在LOL和LOLv2数据集上的客观评价结果分别如表1和表2所示,其中最优值被加粗。综合分析表1和表2结果,本文算法在LOL数据集上的PSNR、SSIM和LPIPS指标均最优,分别为23.65、0.857和0.113;在LOLv2数据集上SSIM和LPIPS均最优,分别为0.853和0.161,整体指标显著优于传统算法BPDHE、LIME和Dong。LLformer虽在LOL数据集上的MSE指标取得最优值0.00771,但SSIM和LPIPS指标均劣于本文算法。SurroundNet在LOLv2数据集上的PSNR和MSE分别为21.30和0.01011,为两项指标的最优值,但仅少量高于本文算法,且SSIM和LPIPS指标大幅度劣于本文算法,同时在LOL数据集上的整体指标均不如本文算法。

LOL和LOLv2测试集上单幅图像的PSNR和SSIM指标分别如图7和图8所示。综合分析

表1 LOL数据集实验结果

Tab. 1 Experimental results on the LOL dataset

Method	PSNR ↑	SSIM ↑	LPIPS ↓	MSE ↓
BPDHE ^[2]	12.22	0.344	0.547	0.0848
LIME ^[5]	14.15	0.366	0.500	0.0429
Dong ^[6]	15.96	0.541	0.330	0.0448
ZERO-DCE ^[10]	15.58	0.511	0.431	0.03637
KinD ^[9]	19.29	0.805	0.207	0.0164
URetinexNet ^[14]	21.33	0.833	0.121	0.01050
TreEnhance ^[13]	21.96	0.81	—	—
DCCNet ^[24]	22.72	0.81	0.143	0.01008
SurroundNet ^[12]	22.81	0.853	0.190	—
LLformer ^[25]	23.65	0.816	0.169	0.00771
Ours	23.65	0.857	0.113	0.00831

表 2 LOLv2数据集实验结果

Tab. 2 Experimental results on the LOLv2 dataset

Method	PSNR \uparrow	SSIM \uparrow	LPIPS \downarrow	MSE \downarrow
BPDHE	13.77	0.389	0.458	0.05035
LIME	16.95	0.535	0.348	0.03033
Dong	17.90	0.546	0.335	0.02661
ZERO-DCE	12.00	0.438	0.510	0.0768
KinD	17.58	0.735	0.448	0.02114
LLformer	20.23	0.782	0.261	0.01325
DCCNet	20.49	0.814	0.220	0.01288
SurroundNet	21.30	0.817	0.200	0.01011
Ours	21.17	0.853	0.161	0.01261

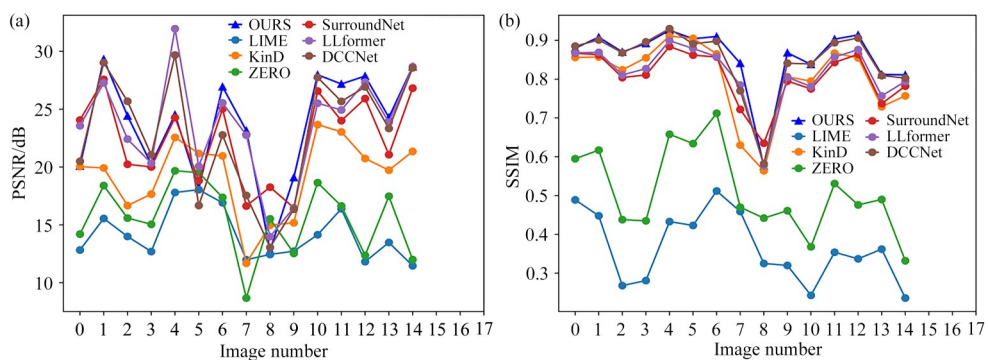


图 7 LOL数据集单幅图像对比结果。(a)PSNR;(b)SSIM。

Fig. 7 Comparison results of single image on the LOL dataset. (a)PSNR;(b)SSIM.

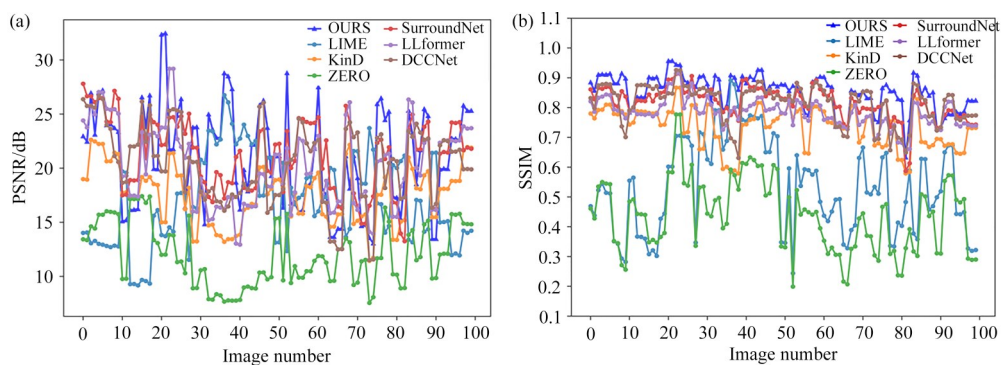


图 8 LOLv2数据集单幅图像对比结果。(a)PSNR;(b)SSIM。

Fig. 8 Comparison results of single image on the LOLv2 dataset. (a)PSNR;(b)SSIM.

图 7 和图 8 结果,大部分单幅图像取得了最佳 PSNR 和 SSIM 值,如图 7 中横坐标为 6 和 11 的图像所示。本文算法折线整体位于对比算法折线上沿,仅有少数坐标点位于对比算法折线下沿,表明 CMA-Net 在增强全局像素信息和保留图像结构方面功效明显,相较于对比算法更具优势。综上分析,本文算法的 4 项指标整体优于对比算

法,说明 CMA-Net 性能优越,可以良好地推理出低照度图像与正常照度图像之间的映射关系。

4 消融实验

为验证所提网络框架各组成部分的有效性,本文基于 LOL 数据集设计两组消融实验,分别

对模型和损失函数进行评估。一是不改变联合损失函数,对模块进行消融实验;二是保持网络结构不变,对联合损失函数进行消融实验。

4.1 模块消融实验

为验证各个模块对图像增强性能的影响,对 3 种模块依次进行消融实验,共有以下 3 种组合:①M1:仅消除像素感知模块;②M2:仅消除多头转置注意力模块;③M3:仅消除选择核特征融合模

块,用常规的求和运算代替。实验定量结果如表 3 所示,最优结果用粗体表示。结果表明,消除像素感知模块后,均方误差明显增加,说明该模块能有效捕获像素信息;消除多头转置注意力模块后,整体性能大幅减弱,说明该模块建模全局信息和过滤冗余特征效果显著;用求和运算替换选择核特征融合模块后,各项评价指标稍劣于原算法,说明该模块在一定程度上提高了模型表征能力。

表 3 模块消融实验结果

Tab. 3 Block ablation results

	PPM	MTAM	SKFF	PSNR ↑	SSIM ↑	LPIPS ↓	MSE ↓
M1		✓	✓	22.36	0.841	0.125	0.011 43
M2	✓		✓	22.14	0.833	0.143	0.009 79
M3	✓	✓		22.92	0.850	0.122	0.009 07
M1+M2+M3	✓	✓	✓	23.65	0.857	0.113	0.008 31

4.2 联合损失函数消融实验

针对联合损失函数的消融实验共有 3 种组合,①M1:仅消除 CharLoss 损失函数;②M2:仅消除边缘损失函数;③M3:仅消除结构相似性损

失函数。实验结果如表 4 所示,其中最优值被加粗。结果表明,无论舍弃三者其中之一,模型性能均有不同程度的降低。三者组合指标最优,说明本文损失函数能良好地估量模型偏差。

表 4 损失函数消融实验结果

Tab. 4 Loss function ablation results

	CharLoss	EdgeLoss	SSIMLoss	PSNR ↑	SSIM ↑	LPIPS ↓	MSE ↓
M1		✓	✓	21.88	0.852	0.121	0.011 44
M2	✓		✓	23.16	0.855	0.117	0.009 08
M3	✓	✓		23.45	0.844	0.119	0.008 44
M1+M2+M3	✓	✓	✓	23.65	0.857	0.113	0.008 31

5 结 论

针对低照度图像亮度低和对比度不均匀等问题,本文提出一种跨级特征自适应融合网络用于低照度图像增强。通过构造像素感知模块提取更大感受野的图像特征,并引入多头转置注意力模块隐式建立局部和全局的长期依赖关系,特化网络学习

的语义信息。在网络末端设计自适应特征融合模块,跨级融合多级分辨率特征。最后构造联合损失函数,综合考虑图像亮度和结构等方面辅助模型优化。实验结果表明,本文算法增强图像清晰度效果显著,有效抑制了纹理细节丢失和噪声问题,对工业、生活等领域有一定的应用价值,未来将研究模型进一步轻量化,并提高网络的泛化性能。

参 考 文 献:

- [1] DONG X B, XU W Y, MIAO Z H, *et al.* Abandoning the Bayer-filter to see in the dark [C]//*Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. New Orleans: IEEE, 2022: 17431-17440.
- [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] LIANG D, LI L, WEI M Q, *et al.* Semantically contrastive learning for low-light image enhancement [C]//*Proceedings of the 36th AAAI Conference on Artificial Intelligence*. Online: AAAI Press, 2022: 1555-1563.
- [4] WANG S H, ZHENG J, HU H M, *et al.* Naturalness preserved enhancement algorithm for non-uniform illumination images [J]. *IEEE Transactions on Image Processing*, 2013, 22(9): 3538-3548.
- [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] DONG X, PANG Y, WEN J T. Fast efficient algorithm for enhancement of low lighting video [C]//*Proceedings of the ACM SIGGRAPH 2010 Posters*. Los Angeles: ACM, 2010: 69.
- [7] LI L, WANG R G, WANG W M, *et al.* A low-light image enhancement method for both denoising and contrast enlarging [C]//*Proceedings of 2015 IEEE International Conference on Image Processing*. Quebec City: IEEE, 2015: 3730-3734.
- [8] 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.
- [9] 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: ACM, 2019: 1632-1640.
- [10] GUO C L, LI C Y, GUO J C, *et al.* Zero-reference deep curve estimation for low-light image enhancement [C]//*Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle: IEEE, 2020: 1780-1789.
- [11] ZAMIR S W, ARORA A, KHAN S, *et al.* Learning enriched features for fast image restoration and enhancement [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023, 45(2): 1934-1948.
- [12] ZHOU F, SUN X, DONG J Y, *et al.* SurroundNet: towards effective low-light image enhancement [J]. *Pattern Recognition*, 2023, 141: 109602.
- [13] COTOGNI M, CUSANO C. TreEnhance: a tree search method for low-light image enhancement [J]. *Pattern Recognition*, 2023, 136: 109249.
- [14] WU W H, WENG J, ZHANG P P, *et al.* URetinex-Net: retinex-based deep unfolding network for low-light image enhancement [C]//*Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. New Orleans: IEEE, 2022: 5901-5910.
- [15] ZHOU L, CAI H M, GU J J, *et al.* Efficient image super-resolution using vast-receptive-field attention [C]//*Proceedings of the Computer Vision-ECCV 2022 Workshops*. Tel Aviv: Springer, 2023: 256-272.
- [16] ZHANG R. Making convolutional networks shift-invariant again [C]//*Proceedings of the 36th International Conference on Machine Learning*. Long Beach: PMLR, 2019: 7324-7334.
- [17] ZAMIR S W, ARORA A, KHAN S, *et al.* Restormer: efficient transformer for high-resolution image restoration [C]//*Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. New Orleans: IEEE, 2022: 5728-5739.
- [18] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, *et al.* An image is worth 16×16 words: transformers for image recognition at scale [J/OL]. *arXiv*, 2020: 2010.11929.
- [19] HENDRYCKS D, GIMPEL K. Gaussian error linear units (GELUs) [J/OL]. *arXiv*, 2016: 1606.08415.
- [20] LI X, WANG W H, HU X L, *et al.* Selective kernel networks [C]//*Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Long Beach: IEEE, 2019: 510-519.
- [21] JIANG K, WANG Z Y, YI P, *et al.* Multi-scale progressive fusion network for single image deraining [C]//*Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle: IEEE, 2020: 8346-8355.
- [22] ZHAO H, GALLO O, FROSIO I, *et al.* Loss functions for image restoration with neural networks [J]. *IEEE Transactions on Computational Imaging*, 2017, 3(1): 47-57.
- [23] WEI C, WANG W J, YANG W H, *et al.* Deep retinex decomposition for low-light enhancement [J/OL]. *arXiv*, 2018: 1808.04560.

- [24] ZHANG Z, ZHENG H, HONG R C, *et al.* Deep color consistent network for low-light image enhancement [C]// *Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. New Orleans: IEEE, 2022: 1899-1908.
- [25] WANG T, ZHANG K H, SHEN T R, *et al.* Ultra-high-definition low-light image enhancement: a benchmark and transformer-based method [C]// *Proceedings of the 37th AAAI Conference on Artificial Intelligence*. Washington: AAAI Press, 2023: 2654-2662.

作者简介:



梁礼明,男,硕士,教授,2005年于江西理工大学获得硕士学位,主要从事机器视觉、医学影像和系统建模等方面的研究。E-mail: lianglm67@163.com



朱晨锟,男,硕士研究生,2016年于湖南师范大学树达学院获得学士学位,主要从事机器学习、人工智能等方面的研究。E-mail:281392667@qq.com