

文章编号 1004-924X(2023)14-2111-12

多通道融合注意力网络的低照度图像增强

陈清江, 顾媛*

(西安建筑科技大学理学院, 陕西西安 710055)

摘要:针对低照度图像亮度低、对比度低、颜色失真以及现有增强算法大多没有区别处理不同的通道,不利于提取多层次特征的问题,提出多通道融合注意力网络的低照度图像增强算法。首先,通过将八度卷积(Octave Convolution, OctConv)引入通道拆分后的残差结构中提出多层次特征提取模块;其次,利用注意力机制与交叉残差结构提出跨尺度特征注意力模块;再次,通过大小与通道数不同的模块堆叠的方式获取多层次信息;最后,在通道维度上进行特征融合,并通过重建模块获得输出。实验结果表明,与RISSNet算法相比,在真实图像上的峰值信噪比与结构相似度分别由27.0016 dB和0.8892提升到27.9781 dB和0.9255。所提算法在峰值信噪比、结构相似度、均方误差、视觉信息保真度4种客观评价指标上均获得了最好的结果。该算法能够有效地提高低照度图像的亮度及对比度,且图像纹理细节及色彩保持较好。

关键词:图像增强;低照度;注意力机制;多通道;八度卷积

中图分类号:TP391.4 **文献标识码:**A **doi:**10.37188/OPE.20233114.2111

Low-light image enhancement algorithm based on multi-channel fusion attention network

CHEN Qingjiang, GU Yuan*

(School of Science, Xi'an University of Architecture and Technology, Xi'an 710055, China)

* Corresponding author, E-mail: 2544020739@qq.com

Abstract: Low-light images have low brightness, low contrast, and color distortion, and most existing enhancement algorithms do not deal with different channels differently, which is not conducive to the extraction of multi-level features. Therefore, this study proposes a low-light image enhancement algorithm based on a multi-channel fusion attention network. Firstly, we introduced octave convolution (OctConv) into the residual structure after channel splitting and propose a multi-level feature extraction module. Secondly, we proposed a cross-scale feature attention module using an attention mechanism and cross-residual structure. Thirdly, we obtained multi-level information by stacking modules with different sizes and channels. Finally, we performed feature fusion in the channel dimension and obtained the final output through the reconstruction module. The experimental results showed that compared with the RISSNet algorithm, the peak signal-to-noise ratio and structural similarity of real images were improved from 27.0016 dB and 0.8892 to 27.9781 dB and 0.9255, respectively. The proposed algorithm achieved the best results in four objective evaluation indicators: peak signal-to-noise ratio, structural similarity, mean squared error,

收稿日期:2022-07-05;修订日期:2022-10-13.

基金项目:国家自然科学基金资助项目(No. 61902304);陕西省自然科学基金资助项目(No. 2021JQ-495)

and visual information fidelity. The algorithm can effectively improve the brightness and contrast of low-light images with well-maintained image textures and colors.

Key words: image enhancement; low light; attention mechanism; multi-channel; OctConv

1 引言

近年来,图像增强技术已广泛应用于刑事侦查^[1]、医学成像^[2]等领域。然而,在低照度环境下拍摄的图像会存在亮度低、对比度低、颜色失真等情况,这些现象不仅会降低用户的视觉体验,还会干扰场景分析,对目标跟踪^[3]、图像融合^[4]等后续计算机视觉任务造成困难。因此,对低照度图像增强的研究具有重要的应用价值。

目前,低照度图像增强算法主要分为传统算法与深度学习算法两大类。传统算法可分为:(1)Hummel等提出了基于直方图均衡化(HE)方法^[5],这种方法通过改变直方图的分布来增强图像。针对HE方法只适用于处理动态范围较小的图像的局限性,相继提出了自适应直方图均衡化(AHE)方法^[6]与限制对比度自适应直方图均衡化(CLAHE)方法^[7]。虽在视觉效果上优于HE方法,但这些算法都是直接增强图像的对比度,忽略了图像内容,导致结果不理想。(2)Land等提出了Retinex理论^[8],将图像分解为反射分量与光照分量。基于此理论,Jobson等提出了单尺度Retinex(SSR)算法^[9]与带色彩恢复的多尺度Retinex(MSRCR)算法^[10]。然而,这些算法将反射分量作为最终增强结果,使得图像颜色失真,纹理不清晰。

随着深度学习的发展,出现了许多用于低照度图像增强的网络模型。其中,比较经典的有:(1)Wei等提出了RetinexNet算法^[11],利用Retinex理论与深度学习相结合,通过分解图像,再调整反射分量与光照分量的思想来增强图像;Zhang等在RetinexNet算法的分解和重建结构上增加了调节网络,提出KinD算法^[12]。然而,这些算法不能解决图像对比度弱的问题,还会存在图像锐化过度等问题。(2)随着生成对抗网络(GAN)在图像处理领域的应用,Guo等提出了Zero-DCE算法^[13],利用高阶曲线调整图像的动态范围,以获得增强后图像;Jiang等通过设计一种双鉴别器结构,提出了非配对训练的Enlight-

enGAN算法^[14]。基于GAN增强的图像通常具有较好的对比度,但存在增强结果不均匀的问题。(3)Lv等提出了MBLLEN算法^[15],对不同层次的特征进行提取和融合,但会出现边缘模糊与局部细节不清晰问题;Pan等提出了多尺度融合的残差编解码器的低照度图像增强算法^[16],直接学习明暗图像之间的端到端映射,更好地去除噪声;Zhao等提出了基于图像分解网络和图像增强网络的RISSNet算法^[17],该算法可以应用于任何低照度图像。但是以上算法均忽略了通道的影响,没有针对不同通道进行处理。

针对以上问题,提出了一种多通道融合注意力网络,有区别地处理不同的通道,有利于提取多层次特征。该网络利用不同通道具有不同信息的特性,通过多通道融合方式,聚合更广的空间上下文信息,有效提取多尺度特征。本文的贡献如下:

(1)将OctConv引入通道拆分后构造的分层残差结构中,提出了多层级特征提取模块。实现高低频信息交互,减少空间冗余的同时以更细粒度表示多尺度特征。

(2)利用注意力机制(CBAM)与多个尺寸卷积核提取到的不同感受野的特征构造交叉残差结构,提出了跨尺度特征注意模块。该模块可以关注更重要的特征,提取多样化信息。

(3)通过三条支路上的特征图尺寸的成倍减小,通道数的成倍增加,使得提取到的特征更具有代表性。同时使用VGG16和VGG19混合的感知损失函数,提高模型性能。

2 相关理论

2.1 OctConv 模块

一幅图像可分为保存大体信息的低频部分与包含细节信息的高频部分,对于一个卷积层输出的特征图,也可按通道分为低频通道与高频通道两部分。Chen等提出了一种名为八度卷积(Octave Convolution, OctConv)^[18]的新的卷积方

法,解决了编码过程中空间维度冗余的问题。根据尺度空间理论,将特征图分解成两组不同频率的特征图,并分别对其进行处理,以捕获更多的上下文信息,具体结构如图 1 所示。

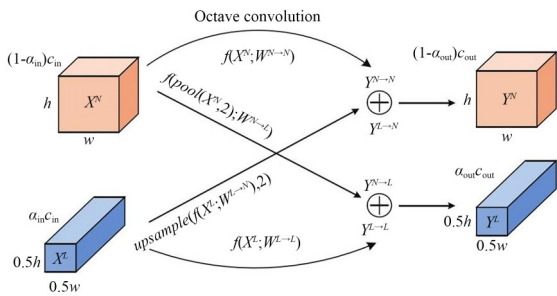


图 1 八度卷积模块
Fig. 1 OctConv module

如图 1 所示,OctConv 将特征图 X 转化成高频分量 X^H 与低频分量 X^L 两部分,卷积核也分为两个分量: $W = (W^H, W^L)$ 。首先对高频分量进行卷积操作,得到特征图 $Y^{H \rightarrow H}$,实现频率内的信息更新,再进行平均池化与卷积操作,得到特征图 $Y^{H \rightarrow L}$,实现高频向低频的信息更新。同样地,对低频分量进行卷积操作,得到特征图 $Y^{L \rightarrow L}$,实现频率内的信息更新,再进行上采样和卷积操

作,得到特征图 $Y^{L \rightarrow H}$,实现频率间的信息更新。最终按元素相加,得到输出特征图 $Y = (Y^H, Y^L)$ 。其具体过程可用式(1)和式(2)表示:

$$Y^H = f(X^H; W^{H \rightarrow H}) + \text{upsample}(f(X^L; W^{L \rightarrow H}), 2), \quad (1)$$

$$Y^L = f(X^L; W^{L \rightarrow L}) + f(\text{pool}(X^H, 2); W^{H \rightarrow L}). \quad (2)$$

2.2 注意力机制

注意力机制的本质是聚焦局部信息,抑制无用信息。CNN 中常见的注意力机制有:SE (Squeeze-and-Excitation)^[19]注意力机制,在通道维度增加注意力机制,关键是 squeeze 和 excitation 操作。通过自动学习的方式,获取特征图的每个通道的重要程度,提升对当前任务有用的特征图的通道,抑制相对不重要的特征通道;ECA (Efficient Channel Attention)^[20]注意力机制,在 SE 模块的基础上,把 SE 中使用的全连接层改为 1×1 卷积,学习通道注意信息,减缩通道维度,降低参数量;CBAM (Convolutional Block Attention Module)^[21]注意力机制,它是解决 SE 只考虑通道而忽略空间信息的弊端提出的结构,如图 2 所示。

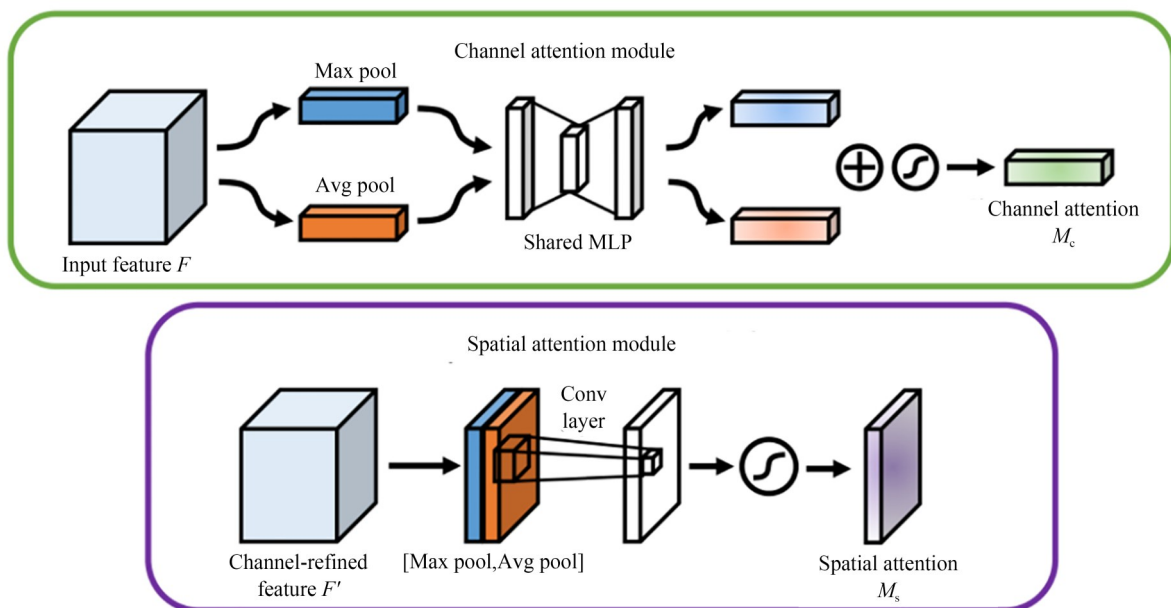


图 2 CBAM 中通道注意力与空间注意力模块

Fig. 2 Channel Attention and Spatial Attention Module in CBAM

CBAM 的计算分为两个阶段:首先,在通道维度上对输入特征图 F 分别进行全局最大池化和全局平均池化,得到两个一维的特征矢量,并通过全连接对特征矢量分配权重,得到权重系数 M_c 。其次,将通道注意力与输入按元素相乘,得到特征图 F' ,实现通道维度上的特征信息加强。计算过程可用式(3)和式(4)表示:

$$M_c = \sigma\left(MLP\left(\text{MaxPool}(F)\right)\right) + \sigma\left(MLP\left(\text{AvgPool}(F)\right)\right), \quad (3)$$

$$F' = M_c(F) \otimes F, \quad (4)$$

式(3)中, $\sigma(\cdot)$ 为 *sigmoid* 激活函数, MLP 表示多层感知机。式(4)中, \otimes 表示对应元素相乘。

再次,在空间维度上对特征图 F' 再次进行全局最大池化与全局平均池化,得到二维特征图,然后对其进行卷积操作,得到权重系数 M_s ,实现空间维度的特征信息加强。最后,将空间注意力与 F' 相乘,得到新特征图 F'' 。计算过程可用下式表示:

$$M_s = \sigma\left(f^{7 \times 7}\left(\left[\text{MaxPool}(F'); \text{AvgPool}(F')\right]\right)\right), \quad (5)$$

$$F'' = M_s(F') \otimes F', \quad (6)$$

式(5)中, $\sigma(\cdot)$ 为 *sigmoid* 激活函数, $f(\cdot)$ 表示卷积运算。式(6)中, \otimes 表示对应元素相乘。

3 本文算法

3.1 网络结构

本文所提模型结构由多层次特征提取、跨尺度特征注意、特征融合与图像重建四部分组成。整个网络分为三条支路, $layer_1$ 支路首先对 256×256 大小的输入特征图进行 3×3 卷积,通道数为 16;其次对其进行 MIFEM 与 CSFAM 操作,获取特征信息;再次进行 1×1 卷积。 $layer_2$ 支路首先对 $layer_1$ 支路经过 MIFEM 操作后的输出进行下采样与 3×3 卷积,通道数为 32;其次对其进行 MIFEM 与 CSFAM 操作;再次通过扩张因子为 2 的反卷积将其恢复至输入大小。同样地, $layer_3$ 支路首先进行下采样与 3×3 卷积,通道数为 64;其次对其进行 MIFEM 与 CSFAM 操作;再次通过扩张因子为 4 的反卷积将其恢复至输入大小。最后对三条支路关注到的重要信息进行融合,并对其使用四个 3×3 卷积获得增强后的图像。具体结构如图 3 所示。

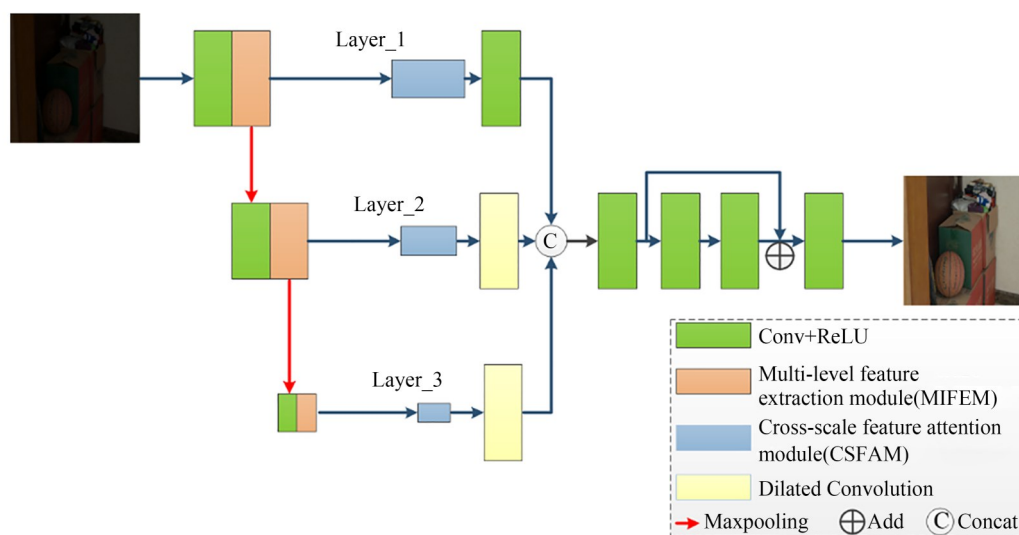


图 3 本文整体的网络结构

Fig. 3 Overall network structure of this paper

3.2 多层次特征提取模块 (MIFEM)

Chen 等^[18]已证明 OctConv 可以降低空间冗余,节省储存和计算,并且捕获更多上下文信息。

Gao 等提出了 Res2Net 结构^[22],利用其在一个残差块内构造多通道残差连接,提高网络多尺度表征能力。本文用 OctConv 替换改进的 Res2Net

结构中的普通卷积,提出多层次特征提取模块(Multi-Level Feature Extraction Module, MIFEM),在更细粒度级别提取多尺度特征,同时实现高低频特征的信息交互,节约计算量,提高网络性能。首先进行 1×1 与 3×3 的卷积,再将两次卷积后的输出按通道平均划分为5组,除 X_1 外,其余子特征均进行OctConv操作。然后对输出子特征进行通道拼接后再进行 3×3 卷积,并与 1×1 卷积后的特征进行残差连接,获得最终输出。该模块第一层卷积的通道数随着最后一层卷积的通道数的改变而改变,分别设置为16,32,64。具体结构如图4所示。

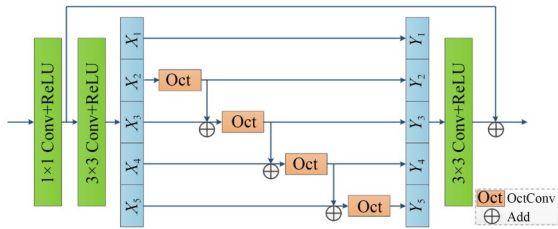


图4 多层次特征提取模块

Fig. 4 Multi-level feature extraction module

通过构造分层残差连接,使得MIFEM中每一个OctConv操作中都含有上一层的特征信息,实现了在更细粒度的层面提取特征,同时每一个输出都能增大感受野。计算过程可用式(7)表示:

$$Y_i = \begin{cases} X_i, & i = 1 \\ Oct(X_i), & i = 2 \\ Oct(X_i + Y_{i-1}), & 2 < i \leq 5 \end{cases}, \quad (7)$$

其中: X_i 表示第*i*块特征层, $Oct(\cdot)$ 表示OctConv操作, Y_i 表示经OctConv操作后的输出。

3.3 跨尺度特征注意模块(CSFAM)

为了关注更重要的信息,提升网络性能,提出跨尺度特征注意模块(Cross-scale feature attention module, CSFAM)。首先,对输入特征分别进行 3×3 与 5×5 的卷积,获取不同尺度的特征,形成两条分支。其次,在不同分支上分别引入注意力机制(CBAM)与 1×1 的卷积层,并使用交叉残差连接方式捕获跨尺度特征。再次,将CBAM操作后的信息融合后分别进行 5×5 与 7×7 的卷积,继续使用交叉残差连接注意多尺

度特征。最后,将不同分支上注意到的信息按元素相加后进行 1×1 的卷积,实现跨尺度信息交互。该模块所有卷积层的通道数随着最后一层卷积的通道数的改变而改变,分别设置为16,32,64。具体结构如图5所示。

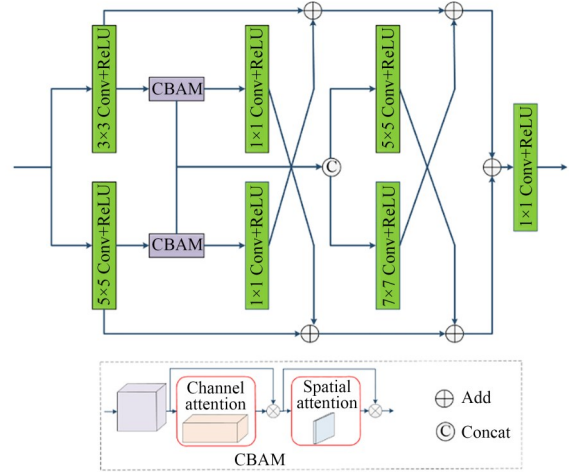


图5 跨尺度特征注意模块

Fig. 5 Cross-scale feature attention module

CBAM结合了特征通道和特征空间两个维度的注意力机制,可以自动获取通道与空间维度的重要程度。CSFAM中包含了四个卷积核大小不同的卷积层,同时引入CBAM,在不同尺度的感受野上关注多尺度特征信息,增加重要信息的多样性。

3.4 特征融合与图像重建模块

由于layer_1, layer_2与layer_3特征图的大小不同,对layer_2与layer_3分别采用扩张因子为2和4的反卷积将其恢复至输入大小并按通道拼接,融合多级特征。图像重建模块由四个 3×3 的卷积层组成,用以获得增强后的图像。其中,前三个卷积层的通道数均为16,最后一个通道数为3,且将第一层与第三层卷积后的输出按元素相加,解决梯度消失问题,避免网络性能的退化。

3.5 损失函数

本文的损失函数由结构相似性损失(Structural Loss)、平均绝对误差损失(MAE Loss)和混合感知损失(Hybrid Perceptual Loss)三部分组成,计算公式如式(8)所示:

$$L = \alpha L_s + \beta L_m + \gamma L_p, \quad (8)$$

其中: L 为总损失, L_s 为结构相似性损失, L_m 为平均绝对误差损失, L_p 为混合感知损失, $\alpha=0.8$, $\beta=0.1$, $\gamma=0.1$ 。

结构相似性损失 (Structural Loss): 衡量真实图像与所提算法增强后的图像之间的差异性,用于改善图像的整体视觉质量。计算公式如式(9)所示:

$$L_s = 1 - \frac{1}{N} \sum_{img} \left(\frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \cdot \frac{2\sigma_{xy} + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \right), \quad (9)$$

其中: μ_x, μ_y 为像素平均值, σ_x^2, σ_y^2 为像素值的方差, σ_{xy} 为像素的协方差, C_1 和 C_2 为常量。

平均绝对误差损失 (MAE Loss): 衡量真实图像与所提算法增强后的图像之差绝对值和的均值,用以表示平均误差幅度。计算公式如式(10)所示:

$$L_m = \frac{1}{N} \sum_{i=1}^N \|y_i - f(x_i)\|_1, \quad (10)$$

其中, y_i 与 $f(x_i)$ 分别表示真实图像与增强后图像。

混合感知损失 (Hybrid Perceptual Loss): Structural Loss 与 MAE Loss 不能有效的关注图像的深层信息,为此提出混合感知损失改善图像质量。VGG 网络^[23]是一种有效的特征提取网络,本文利用训练好的 VGG16 与 VGG19 作为特征提取器,分别测量重建图像的特征和相应真实图像之间的差异,并对其计算 L_2 损失。计算公式如式(11)所示:

$$L_p = 0.5 \left(\left\| \hat{L}_{VGG16} - L_{VGG16} \right\|_1 + \left\| \hat{L}_{VGG19} - L_{VGG19} \right\|_1 \right), \quad (11)$$

其中, \hat{L} 与 L 分别为增强图像和真实图像。

4 实验

4.1 实验条件及参数配置

本文实验基于 Tensorflow2.0 深度学习框架与 python3.7 环境进行训练与测试,在 Windows10, Inter i7-10750H, 2.6 GHz, 16 G RAM, Nvidia GeForce 1660Ti GPU 平台上完成。在训练过程中,对所有输入图像裁剪为 256×256 的

大小。采用 Adam 优化器对模型进行优化,动量参数为 $\beta_1=0.9$, $\beta_2=0.999$,设置固定学习率 $lr=0.0001$, $epoch=3000$ 。

4.2 评价指标

本文使用峰值信噪比 (Peak signal-to-noise ratio, PSNR)^[24]、结构相似度 (Structural similarity, SSIM)^[25]、均方误差 (Mean squared error, MSE)、视觉信息保真度 (Visual information fidelity, VIF)^[26] 与信息熵 (Information entropy, IE) 五项评价指标,对合成的与真实的低照度图像进行质量评估,说明本文算法的有效性。其中,PSNR 的值越大表明图像质量越好;SSIM 是一种衡量两幅图像相似度的指标,其值越大表明增强图像越接近参考图像;MSE 是一种反映两幅图像之间差异性的指标,其值越小表明图像失真越小;VIF 是一种衡量两幅图像之间互信息的指标,其值越大表明图像细节信息保留越好;IE 用来反应图像信息量的大小,其值越大图像信息越丰富。

4.3 对合成的低照度图像实验结果分析

对人工合成的低照度图像进行实验,从公开数据集 Berkeley Segmentation Dataset^[27] 中选取 330 幅正常光照图像,利用 Retinex 理论^[8] 对光照分量随机取值,合成 330 幅低照度图像。为验证本文模型的有效性,将本文算法与两种传统的 HE^[4] 与 SSR^[9] 算法及五种基于深度学习的 RetinexNet^[11], KinD^[12], Zero-DCE^[13], MBLLEN^[15] 与 RISSNet^[17] 算法进行对比,本文随机选取了 15 幅图像作为测试集,主观视觉对比结果如图 6 所示。

图 6 中,HE 算法出现了严重的颜色失真问题;SSR 算法存在过曝光现象;RetinexNet 算法增强后的图像整体亮度偏暗;KinD 算法在图像 3 与图像 4 中的部分区域出现了噪声与颜色失真现象;Zero-DCE 算法虽提升了图像的亮度,但与参考图像相比,整体亮度依然偏暗;MBLLEN 算法虽亮度方面较接近参考图像,但图像 3 中的花朵与图像 4 的房顶出现了明显的噪声;RISSNet 算法增强后的图像虽然亮度恢复较好,但纹理信息增强不佳,图像 2 中出现伪影,图像 4 中存在噪声;本文算法主观视觉效果最好,在图像亮度提升与色彩保持方面最接近参考图像。



图 6 不同算法在合成低照度图像上的主观视觉对比

Fig. 6 Subjective visual contrast of different algorithms on synthesizing low-light images

进一步对本文算法进行客观评价,对每种方法进行 PSNR^[24], SSIM^[25], MSE、VIF^[26]与 IE 值计算,测试集上的平均值如表 1 所示。

表 1 中,本文算法除了在 IE 指标上低于 HE

算法,在其他的 PSNR, SSIM, MSE 与 VIF 四项客观评价指标上均优于其他对比算法。说明本文算法在低照度图像亮度提升、色彩及纹理细节保持方面较好,图像失真程度较小。

表 1 不同算法在合成低照度图像上的评价指标平均值

Tab. 1 Average value of evaluation indicators for different algorithms on synthesizing low-light images

方法	HE	SSR	RetinexNet	KinD	Zero-DCE	MBLLEN	RISSNet	本文算法
PSNR ↑	13.661 1	9.137 5	18.280 4	21.779 5	22.978 8	25.986 4	27.341 6	28.346 5
SSIM ↑	0.660 8	0.564 2	0.816 4	0.825 9	0.849 1	0.874 0	0.884 2	0.906 5
MSE ↓	0.037 6	0.084 3	0.023 3	0.006 2	0.004 8	0.002 9	0.001 2	0.000 7
VIF ↑	0.237 5	0.410 1	0.519 2	0.536 7	0.571 2	0.578 5	0.631 0	0.682 0
IE ↑	7.771 5	6.158 1	6.444 0	6.642 1	6.658 5	6.767 3	0.671 6	6.962 3

4.4 对真实的低照度图像实验结果分析

使用 LOL dataset^[28] 公开数据集对模型进行测试, LOL dataset 数据集中包含 500 对低照度图像和正常照度图像。为验证本文模型的有效性,将本文算法与两种传统的 HE^[4]与 SSR^[9]算法及四种基于深度学习的 RetinexNet^[11], KinD^[12], Zero-DCE^[13], MBLLEN^[15]与 RISSNet^[17]算法进行对比。本文随机选取了 15 幅图像作为测试集,主观视觉对比结果如图 7 所示。

图 7 中, HE 算法出现了颜色失真问题; SSR 算法存在颜色失真与过曝光的现象; RetinexNet 与 KinD 算法不仅整体亮度提升不足,且图像模糊,存在黑色伪影; Zero-DCE 算法基本可以提升图像的亮度,但图像 5 与图像 6 出现了偏色情

况; MBLLEN 算法在图像亮度提升方面较接近参考图像,但在细节恢复方面存在不足。例如,图像 5 中墙面上的花纹比较模糊; RISSNet 算法与本文算法在主观视觉方面较为接近,但本文算法在图像亮度提升与色彩保持方面更接近参考图像。

进一步对本文算法进行评价,对每种方法进行客观评价指标值计算,测试集上的平均值如表 2 所示。

表 2 中,由于 HE 算法的原理是将原始图像的灰度直方图从比较集中的某个灰度区间变成全部灰度内的均匀分布。当灰度分布更均匀时,图像所含信息量越大,因此本文算法在 IE 指标上略低于 HE 算法,但在其它四项客观评价指标上

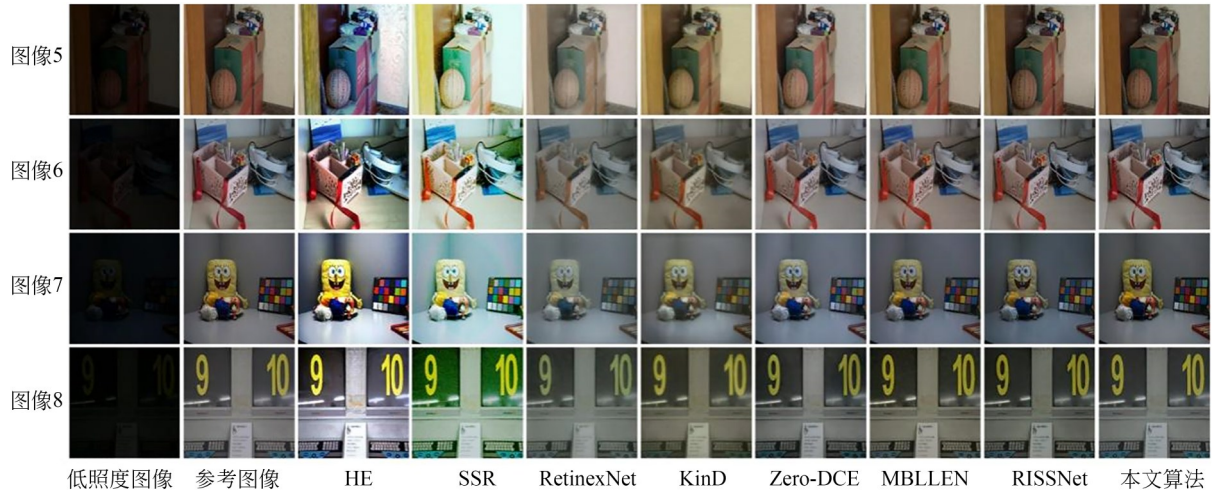


图 7 不同算法在真实低照度图像上的主观视觉对比

Fig. 7 Subjective visual comparison of different algorithms on real low-light images

表 2 不同算法在真实低照度图像上的评价指标平均值

Tab. 2 The average value of evaluation indicators for different algorithms on real low-light images

方法	HE	SSR	RetinexNet	KinD	Zero-DCE	MBLLEN	RISSNet	本文算法
PSNR ↑	14.056 9	10.359 3	19.853 3	21.468 5	22.250 1	26.046 8	27.001 6	27.978 1
SSIM ↑	0.712 8	0.593 2	0.812 8	0.840 9	0.853 7	0.888 5	0.889 2	0.925 5
MSE ↓	0.041 2	0.103 3	0.028 9	0.018 3	0.011 3	0.003 5	0.001 3	0.000 7
VIF ↑	0.301 3	0.308 2	0.440 2	0.480 6	0.483 3	0.576 0	0.634 6	0.669 9
IE ↑	7.944 7	6.826 6	6.607 4	6.799 3	6.710 8	6.856 7	6.876 7	6.984 2

均优于其他对比算法。说明本文算法增强后的图像色彩及纹理细节保持方面较好,最接近参考图像。

4.5 消融实验

为验证所提算法的有效性,本文分别对不同的模型结构及损失函数进行消融实验。实验选用 PSNR^[24], SSIM^[25] 与 VIF^[26] 三项客观评价指标对 LOL dataset 测试结果进行评价。

4.5.1 网络结构评估

为评估 OctConv、CBAM 模块的作用及 layer_1、layer_2 与 layer_3 支路使用不同通道数的增强效果,分别对不使用 OctConv 的 MIFEM、不使用 CBAM 的 CSFAM、将 CSFAM 结构中的 CBAM 替换为 SE、ECA 模块以及 layer_1、layer_2 与 layer_3 支路的通道数均为 16、32 或 64 的网络结构进行消融实验。以图像 7 和图像 8 为例,主观视觉效果如图 8 所示。

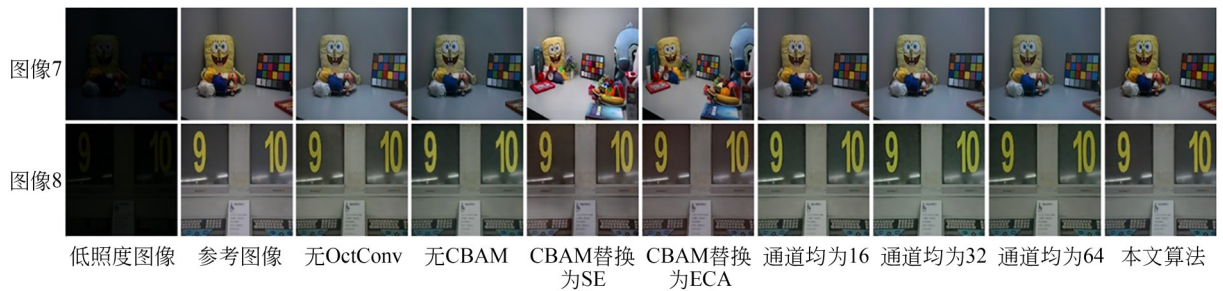


图 8 不同网络结构增强效果的主观视觉对比

Fig. 8 Subjective visual contrast of enhancement effects of different network structures

图 8 中,无 OctConv 增强后的图像局部区域出现过曝光现象;无 CBAM 增强后的图像整体亮度提升不足;将 CBAM 替换为 SE 模块后的图像出现过曝光现象,且有噪声;将 CBAM 替换为 ECA 模块后的图像局部区域出现过曝光;layer_1、layer_2 与 layer_3 支路通道数均为 16 时,增强后的图像出现偏色情况;通道数均为 32 或 64 时,增强后的图像色彩恢复不足。本文算法色彩及细节保持方面较好,视觉效果最接近参考图像。为进一步衡量本文模型的有效性,进行客观评价,测试集上的平均值如表 3 所示。

表 3 不同网络结构增强效果的评价指标平均值

Tab. 3 The average value of the evaluation index for the enhancement effect of different network structures

模型	PSNR ↑	SSIM ↑	VIF ↑
无 OctConv	25.091 5	0.890 1	0.499 7
无 CBAM	24.633 8	0.892 5	0.505 1
CBAM 替换为 SE	24.064 1	0.879 1	0.489 2
CBAM 替换为 ECA	26.318 2	0.893 7	0.536 6
通道数均为 16	26.553 4	0.899 8	0.515 4
通道数均为 32	25.911 7	0.899 9	0.497 7
通道数均为 64	24.918 5	0.891 0	0.509 9
本文算法	27.978 1	0.925 5	0.669 9

表 3 中,本文算法增强效果最优,说明在 MI-FEM 中使用 OctConv 代替普通卷积,以及在 CS-FAM 中加入 CBAM 的有效性。此外,说明 lay-

er_1、layer_2 与 layer_3 支路的通道数分别为 16、32、64 的多通道融合网络的效果优于单一通道数的网络。

4.5.2 损失函数评估

本文所用总损失函数为结构相似性损失 (Structural Loss)、平均绝对误差损失 (MAE Loss) 和混合感知损失 (Hybrid Perceptual Loss) 三部分的加权和,其中不同的权重系数对实验结果的影响不同,故本文分别验证 α, β, γ 取不同值时对实验结果的影响。表 4 为 α, β, γ 取不同值时网络模型在 LOL dataset 测试集上的指标平均值对比结果。表 4 结果表明,当 $\alpha = 0.8, \beta = 0.1, \gamma = 0.1$ 时,网络性能最好。

表 4 不同权重系数的损失函数的评价指标平均值

Tab. 4 Average of the evaluation indicators for loss functions with different weight coefficients

损失函数	PSNR ↑	SSIM ↑	VIF ↑
$\alpha = 0.1, \beta = 0.1, \gamma = 0.8$	23.613 4	0.831 3	0.486 7
$\alpha = 0.3, \beta = 0.3, \gamma = 0.4$	24.921 3	0.856 2	0.509 3
$\alpha = 0.5, \beta = 0.3, \gamma = 0.2$	26.613 8	0.873 7	0.550 9
$\alpha = 0.7, \beta = 0.1, \gamma = 0.2$	27.122 6	0.899 5	0.621 0
$\alpha = 0.8, \beta = 0.1, \gamma = 0.1$	27.978 1	0.925 5	0.669 9

本文使用了 VGG16 与 VGG19 的混合感知损失函数,为验证所提损失函数的有效性,分别对仅含 VGG16、仅含 VGG19 及无 VGG 损失函数的网络模型进行消融实验。以图像 7 和图像 8 为例,主观视觉效果如图 9 所示。

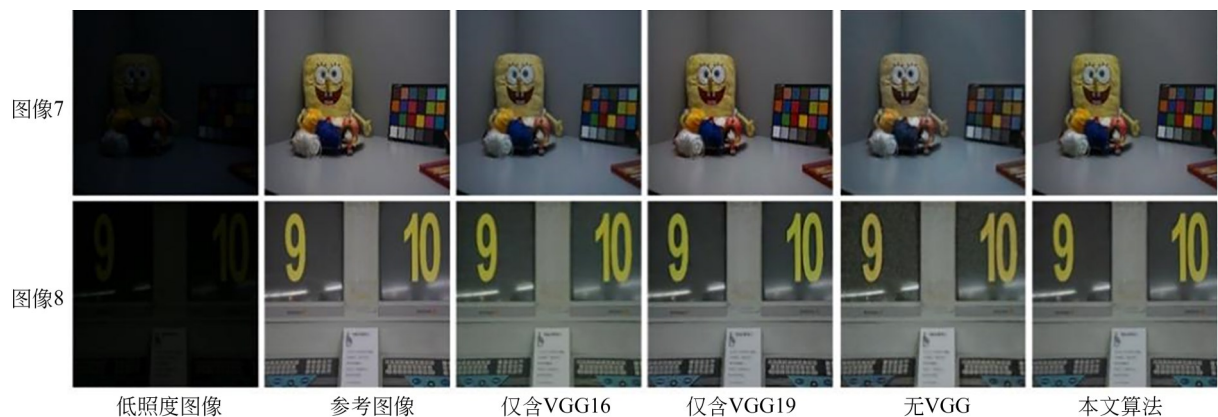


图 9 不同损失函数增强效果的主观视觉对比

Fig. 9 Subjective visual contrast of the enhancement effects of different loss functions

图 9 中, 损失函数中仅含 VGG16 与仅含 VGG19 时, 增强后的图像出现轻微的偏色情况; 无 VGG 时, 增强后的图像存在噪声。本文算法的视觉效果最接近参考图像。为进一步衡量所提损失函数的有效性, 进行客观评价, 测试集上的平均值如表 5 所示。

表 5 中, 本文算法增强效果最优, 说明所提混合感知损失能够较好地保留图像的色彩及纹理细节信息。

表 5 不同损失函数增强效果的评价指标平均值

Tab. 5 Average of the evaluation indicators for the enhancement effect of different loss functions

损失函数	PSNR ↑	SSIM ↑	VIF ↑
仅含 VGG16	27.106 2	0.902 5	0.558 1
仅含 VGG19	27.003 8	0.901 6	0.531 6
无 VGG	26.181 0	0.886 9	0.463 9
本文算法	27.978 1	0.925 5	0.669 9

参考文献:

- [1] ELIZA JACOB J, SARITHA S. *Video Enhancement and Low-Resolution Facial Image Reconstruction for Crime Investigation* [M]. Intelligent Data Communication Technologies and Internet of Things. Singapore: Springer Singapore, 2021: 773-788.
- [2] MA Y H, LIU J, LIU Y H, *et al.* Structure and illumination constrained GAN for medical image enhancement[J]. *IEEE Transactions on Medical Imaging*, 2021, 40(12): 3955-3967.
- [3] 朱均安, 陈涛, 曹景太. 基于显著性区域加权的相关滤波目标跟踪[J]. *光学精密工程*, 2021, 29(2):363-373.
ZHU J A, CHEN T, CAO J T., CHEN T, CAO J T. Salient region weighted correlation filter for object tracking [J]. *Opt. Precision Eng.*, 2021, 29(2):363-373. (in Chinese)
- [4] 杨艳春, 裴佩佩, 党建武, 等. 基于交替梯度滤波器和改进 PCNN 的红外与可见光图像融合[J]. *光学精密工程*, 2022, 30(9):1123-1138.
YANG Y C, PEI P P, DANG J W, *et al.* Infrared

5 结 论

针对低照度图像现存的问题, 本文提出了一种多通道融合注意力网络的低照度图像增强算法。首先, 将八度卷积引入通道拆分后的残差结构中, 提出 MIFEM, 以更细粒度提取多层次特征; 其次, 利用注意力机制与交叉残差结构, 提出 CSFAM, 使网络关注更重要的特征; 再次, 对不同通道数的三条支路上提取到的信息进行融合; 最后, 在通道维度融合, 获得输出。实验结果表明, 所提算法在真实图像上的峰值信噪比与结构相似度分别达到 27.978 1 dB 和 0.925 5, 均优于其他对比算法。本文所提算法针对不同的通道进行区别处理, 有效提高网络的特征提取能力, 从而增强后的图像色彩及纹理细节保持较好。现实生活中许多低照度场景是以视频的形式存在的, 所提算法未来将应用于低照度视频增强。

- and visible image fusion based on alternating gradient filter and improved PCNN [J]. *Opt. Precision Eng.*, 2022, 30(9):1123-1138. (in Chinese)
- [5] HUMMEL R. Image enhancement by histogram transformation [J]. *Computer Graphics and Image Processing*, 1977, 6(2): 184-195.
- [6] PIZER SM, AMBURN EP, AUSTIN JD, *et al.* Adaptive histogram equalization and its variations [J]. *Computer Vision, Graphics, and Image Processing*, 1987, 39(3): 355-368.
- [7] REZA A M. Realization of the contrast limited adaptive histogram equalization (CLAHE) for real-time image enhancement [J]. *Journal of VLSI Signal Processing Systems for Signal, Image and Video Technology*, 2004, 38(1): 35-44.
- [8] LAND E H. The retinex theory of color vision [J]. *Scientific American*, 1977, 237(6): 108-128.
- [9] JOBSON D J, RAHMAN Z, WOODDELL G A. Properties and performance of a center/surround retinex [J]. *IEEE Transactions on Image Processing: a Publication of the IEEE Signal Processing Society*, 1997, 6(3): 451-462.

- [10] JOBSON D J, RAHMAN Z, WOODSELL G A. A multiscale retinex for bridging the gap between color images and the human observation of scenes [J]. *IEEE Transactions on Image Processing: a Publication of the IEEE Signal Processing Society*, 1997, 6(7): 965-976.
- [11] WEI C, WANG W, YANG W, *et al.* Deep Retinex Decomposition for Low-Light Enhancement [EB/OL]. 2018: *arXiv*: 1808.04560. <https://arxiv.org/abs/1808.04560>
- [12] 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*. October 21-25, 2019, Nice, France. New York: ACM, 2019: 1632-1640.
- [13] 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)*. 13-19, 2020, Seattle, WA, USA. IEEE, 2020: 1777-1786.
- [14] JIANG Y F, GONG X Y, LIU D, *et al.* EnlightenGAN: deep light enhancement without paired supervision [J]. *IEEE Transactions on Image Processing*, 2021, 30: 2340-2349.
- [15] LV F F, LU F, WU J H, *et al.* MBLLEN: Low-Light Image/Video Enhancement Using CNNs [C]. *BMVC*. 2018, 220(1): 4.
- [16] 潘晓英, 魏苗, 王昊, 等. 多尺度融合残差编解码器的低照度图像增强方法[J]. *计算机辅助设计与图形学学报*, 2022, 34(1): 104-112.
PAN X Y, WEI M, WANG H, *et al.* A multi-scale fusion residual encoder-decoder approach for low illumination image enhancement [J]. *Journal of Computer-Aided Design & Computer Graphics*, 2022, 34(1): 104-112. (in Chinese)
- [17] ZHAO J B, CHEN H Y, ZENG S Y, *et al.* RISSNet: Retain low-light image details and improve the structural similarity net [J]. *IET Image Processing*, 2022, 16(7): 1793-1806.
- [18] CHEN Y P, FAN H Q, XU B, *et al.* Drop an octave: reducing spatial redundancy in convolutional neural networks with octave convolution [C]. 2019 *IEEE/CVF International Conference on Computer Vision (ICCV)*. October 27-November 2, 2019, Seoul, Korea (South). IEEE, 2020: 3434-3443.
- [19] HU J, SHEN L, SUN G. Squeeze-and-Excitation Networks [C]. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 18-23, 2018, Salt Lake City, UT, USA. IEEE, 2018: 7132-7141.
- [20] WANG Q L, WU B G, ZHU P F, *et al.* ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks [C]. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 13-19, 2020, Seattle, WA, USA. IEEE, 2020: 11531-11539.
- [22] WOO S, PARK J, LEE J Y, *et al.* CBAM: Convolutional Block Attention Module [M]. *Computer Vision - ECCV 2018*. Cham: Springer International Publishing, 2018: 3-19.
- GAO S H, CHENG M M, ZHAO K, *et al.* Res2Net: a new multi-scale backbone architecture [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021, 43(2): 652-662.
- [23] SIMONYAN K, ZISSERMAN A. Very Deep Convolutional Networks for Largescale Image Recognition [C]. *Proc of the 3rd International Conference on Learning Representations*. La Jolla, CA: *ICLR Press*, 2015: 1-14.
- [24] MANNOS J, SAKRISON D. The effects of a visual fidelity criterion of the encoding of images [J]. *IEEE Transactions on Information Theory*, 1974, 20(4): 525-536.
- [25] WANG Z, BOVIK A C, SHEIKH H R, *et al.* Image quality assessment: from error visibility to structural similarity [J]. *IEEE Transactions on Image Processing: a Publication of the IEEE Signal Processing Society*, 2004, 13(4): 600-612.
- [26] SHEIKH H R, BOVIK A C. Image information and visual quality [J]. *IEEE Transactions on Image Processing*, 2006, 15(2): 430-444.
- [27] CHEN X Y, WANG S A. Superpixel segmentation based on delaunay triangulation [C]. 2016

23rd International Conference on Mechatronics and Machine Vision in Practice (M2VIP). 28-30, 2016, Nanjing, China. IEEE, 2017: 1-6.

[28] CHEN C, CHEN Q F, XU J, *et al.* Learning to

see in the dark[C]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. 18-23, 2018, Salt Lake City, UT, USA. IEEE, 2018: 3291-3300.

作者简介:



陈清江(1966—),男,河南信阳人,博士,教授,硕士生导师,2006年于西安交通大学获博士学位,主要从事小波分析,图像处理与信号处理方面的研究。E-mail:qjchen66xytu@126.com

通讯作者:



顾媛(1998—),女,陕西汉中,人,硕士研究生,主要从事图像处理方面的研究。E-mail:2544020739@qq.com