

从局部到全局的零参考低照度图像增强方法

杨伟¹, 王帅^{1,2}, 吴佳奇¹, 陈伟^{1,3}, 田子建¹

(1. 中国矿业大学(北京)人工智能学院, 100083, 北京; 2. 国家矿山安全监察局内蒙古局, 010010, 呼和浩特;
3. 中国矿业大学计算机科学与技术学院, 221116, 江苏徐州)

摘要: 为解决现有的低照度图像增强方法存在的色彩失真、细节损失以及暗区增强不足和亮区增强过度导致低照度图像增强效果不理想的问题,提出了一种从局部到全局的零参考低照度图像增强方法。采用局部照度增强对低照度图像进行像素级增强,改进了自适应光照映射估计函数,提升了照度调整能力,避免了生成大量的迭代参数,提高了模型的推理速度;采用基于 Transformer 结构的全局图像调整对局部增强后的图像进行全局调整,解决了亮区照度增强过度的曝光问题和暗区照度增强不足的问题,提升了图像的整体对比度;优化损失函数,对低照度图像特征和增强图像特征进行相似性约束,提升了目标检测精度。实验结果表明,LOL 数据集上的客观指标峰值信噪比和结构相似性达到了 20.18 dB 和 0.80, MIT-Adobe FiveK 数据集上达到了 23.31 dB 和 0.87, ExDark 数据集上增强后图像的目标检测精度提高了 7.6%,有效提升了低照度图像可视化质量和目标检测效果。

关键词: 图像处理;机器视觉;轻量级网络;低照度图像;图像增强;目标检测

中图分类号: TP391 **文献标志码:** A

DOI: 10.7652/xjtuxb202404015 **文章编号:** 0253-987X(2024)04-0158-12

A Zero-Reference Low-Light Image Enhancement Method from Local to Global

YANG Wei¹, WANG Shuai^{1,2}, WU Jiaqi¹, CHEN Wei^{1,3}, TIAN Zijian¹

(1. School of Artificial Intelligence, China University of Mining and Technology (Beijing), Beijing 100083, China;

2. Inner Mongolia Bureau of State Mine Safety Supervision Bureau, Hohhot 010010, China;

3. School of Computer Science and Technology, China University of Mining and Technology, Xuzhou, Jiangsu 221116, China)

Abstract: To address the problem of unsatisfactory low-light image enhancement caused by color distortion, loss of details, insufficient enhancement of dark areas, and excessive enhancement of bright areas in existing low-light image enhancement methods, a zero-reference low-light image enhancement method from local to global is proposed. Initially, the low-illumination image is enhanced at the pixel level through local illumination enhancement, the adaptive light mapping function is improved, and the illumination adjustment ability is enhanced to avoid the generation of massive iterative parameters and improve the inference speed of the model. Subsequently, the global adjustment is performed on the locally enhanced image through Transformer-based global image adjustment to tackle the exposure problem of excessive illumination enhancement in the bright area and insufficient illumination enhancement in the dark area, thereby enhancing the overall contrast of the image. The loss function is optimized to constrain similarities between

features of the low-illumination image and those of the enhanced image, thereby improving the target detection accuracy. The experimental results show that the objective metrics peak signal-to-noise ratio and structural similarity can reach 20.18 dB and 0.80 on the LOL dataset, 23.31 dB and 0.87 on the MIT-Adobe FiveK dataset, and the target detection accuracy of the enhanced image on the ExDark dataset has increased by 7.6%, thus effectively improving the visualization quality of the low-illumination image and the target detection effect.

Keywords: image processing; machine vision; lightweight network; low-light images; image enhancement; object detection

低照度条件下如夜间、阴天和地下空间等,物体表面反射光较弱,拍摄到的图像受到背光、不均匀光照、低对比度和密集噪声的影响,不仅图像质量严重降低,难以获得理想的采集效果和关键信息,而且为后续的高层视觉任务带来巨大挑战,如目标检测、图像分割和物体追踪等^[1-2]。因此,低照度图像增强在计算机视觉领域具有重要的研究意义和应用价值^[3]。

随着计算机视觉的快速发展,低照度图像增强算法在视觉监控、自动驾驶、智能交通和摄影等领域有广泛应用,主要分为基于传统的低照度图像增强方法和基于深度学习的低照度图像增强方法。传统低照度图像增强算法包括直方图均衡化方法^[4]、频域方法^[5]、基于 Retinex 理论的方法^[6-7]和基于去雾模型的方法^[8-9]。此类方法一定程度上提高了主观视觉效果,但优化过程较为复杂,运行时间较长,图像的照度调整不够灵活,增强效果往往伴随着严重的色彩失真和细节损失。

相比于传统低照度图像增强算法,凭借大规模数据中提取先验的强大能力,基于深度学习的低照度图像增强算法具有更强的鲁棒性、适应性和处理速度^[10],根据学习策略可以分为监督学习、强化学习、无监督学习和零次学习。其中,Lü等^[11]基于监督学习设计了端到端的多分枝增强模型。部分基于监督学习的方法采用了 Retinex 理论^[12-14],如 Wei等^[15]通过新的约束条件和先进的模型设计,使得 Retinex-Net 达到了较好的增强性能。Zhang等通过图层分解、反射率恢复和照度调整提出了 KinD^[16]和 KinD++^[17]方法,有效缓解了图像增强后的主观视觉缺陷问题。Yu等^[18]基于强化学习提出了 DeepExposure 方法,在没有配对数据集的情况下,通过融合不同曝光下的多个修饰图像获得增强后的图像。Jiang等^[19]基于无监督学习提出了 EnlightenGAN 方法,通过全局-局部判别器和特征保留损失函数,解决了基于合成数据训练的过拟合和难以泛化的问题。基于监督学习的模型需要基于配对数据进行训练,因此模型泛化性不强,对现实低

照度图像的增强效果不理想;基于无监督学习的模型较大,不利于模型部署;基于强化学习的模型难以实现高效稳定的训练。因此,使用零次学习实现低照度图像增强具有重要意义。零次学习仅从训练数据中学习照度增强过程,应用过程中不需要配对或非配对的数据。Zhang等^[20]通过估算适合输入图像的 S 曲线来调整滤波器分离出来的基础层,融合调整后的基础层和细节层得到增强后的图像。Guo等提出 Zero-DCE^[21],将低光增强作为针对图像光照曲线的估算任务,利用高阶曲线对输入图像光照的动态范围进行调整得到增强图像,并在此基础上提出了轻量化版本 Zero-DCE++^[22],通过共享高阶曲线的权重参数和模型优化达到了增强过程的快速推算,但增强结果存在不同程度的偏色、亮区增强过度 and 暗区增强不足现象。以上低照度增强方法在人眼主观视觉和客观评价指标取得了明显提升,一定程度上提高了目标检测与识别的性能,但低照度增强任务和目标检测任务的目的不同,前者对细节敏感侧重于抽象度低的特征,后者侧重于抽象度高的目标特征,经过低照度增强后的图像可能包含不利于目标检测的噪声,无法保证增强后图像有更好的目标检测效果。

综合以上分析,本文从图像视觉质量和目标检测效果两个角度考虑,提出了一种从局部到全局的零参考低照度图像增强方法,解决了以下几个问题:Zero-DCE 的自适应光照映射估计函数的动态调整范围不足;多次生成像素级的高阶迭代参数需要较多的参数量和运行时间;增强图像的对比度不足、噪点突出和曝光现象;低照度图像的目标检测精度低。

1 从局部到全局的零参考低照度图像增强方法

提出的低照度图像增强方法的总体架构如图 1 所示,主要包含两部分:①局部照度增强,由 7 个卷积组成,每个卷积后采用 ReLU 激活函数,最后 1 个

卷积采用 Tanh 激活函数,并对自适应光照映射估计函数进行改进,采用倒数函数取代原有的二次迭代函数,提高照度的调整范围,降低参数量和运算时间;② 基于 Transformer 的全局图像调整,包含

Transformer 结构、颜色校正矩阵和伽马变换,通过捕获单个像素和周围像素之间的关系,获得更丰富的全局特征信息,解决亮区增强过度 and 暗区增强不足的问题,提高图像的整体对比度。

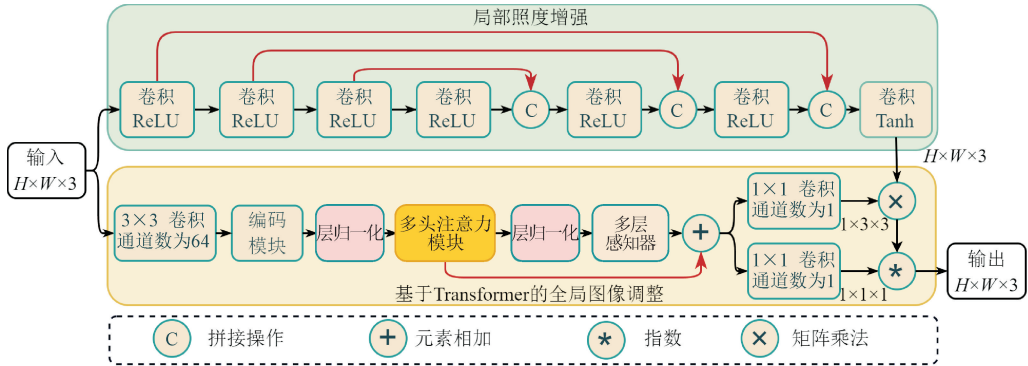


图 1 提出的低照度图像增强方法的总体架构

Fig. 1 Overall architecture diagram of proposed low-light image enhancement method

1.1 局部照度增强

局部照度增强用于估计像素级的照度映射关系以校正光照的影响,结构如图 1 所示,采用了 7 个卷积和 3 次拼接操作,每个卷积后采用 ReLU 激活函数,最后 1 个卷积采用 Tanh 激活函数,未采用下采样和采样操作,尽可能保留图像的细节信息。局部照度增强是在 Zero-DCE 的基础上改进了自适应光照映射估计函数。Zero-DCE 方法提出的基于二次函数迭代的光照映射估计函数为

$$E(X) = X + \alpha X(1 - X) \quad (1)$$

$$E_n(X) = X_{n-1} + \alpha_n X_{n-1}(1 - X_{n-1}) \quad (2)$$

式中: X 表示输入; $E(X)$ 表示为给定输入 X 的增强版; $E_n(X)$ 表示迭代第 n 次产生的增强结果; $\alpha_n \in [-1, 1]$ 是通过模型训练得到的参数,用于第 n 次叠代调整映射曲线,控制曝光水平。

由于调整参数 α 为像素级的参数,即输入图像的每一个像素都有一个最适合的 α 曲线来调整其动态范围,因此基于二次迭代函数的像素级光照映射关系可表示为

$$E_n(X) = E_{n-1}(X) + A_n E_{n-1}(X)(1 - E_{n-1}(X)) \quad (3)$$

式中: A_n 与输入 X 的维度相同,表示第 n 次迭代的权重。

提出的基于倒数函数的光照映射估计函数可表示为

$$E(X) = \frac{X(\alpha + 1)}{X + \alpha} \quad (4)$$

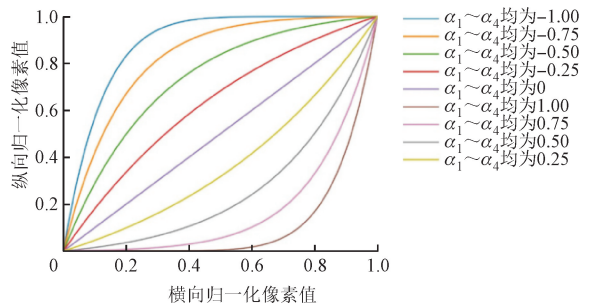
式中: X 表示输入; $E(X)$ 表示为给定输入 X 的增强版; $\alpha \in [-\infty, -1] \cup [0, +\infty]$ 是通过模型训练得到的参数。

由于调整参数 α 为像素级的参数,即输入图像的每一个像素都有一个最适合的 α 曲线来调整其动态范围,因此提出的基于倒数函数的像素级光照映射关系可表示为

$$E(X) = \frac{X(A_\alpha + 1)}{X + A_\alpha} \quad (5)$$

式中: A_α 是与输入 X 维度大小相同的权重。所有像素都被归一化为 $[0, 1]$,并将曲线分别作用于 RGB 的 3 个通道。

图 2 为自适应光照映射估计函数曲线。从图中可以看出,迭代次数越多,光照调整能力越强,但表征范围仍有缺失。同时,迭代次数越多,计算量和参数量越多,模型结构越复杂,模型推理时间越长。提出的基于倒数函数的光照映射估计函数如图 2(c)所示,光照调整范围更大,不需要迭代计算,因此学习参数更少,模型更加轻量化。



(a) 二次迭代函数 ($n=4$)

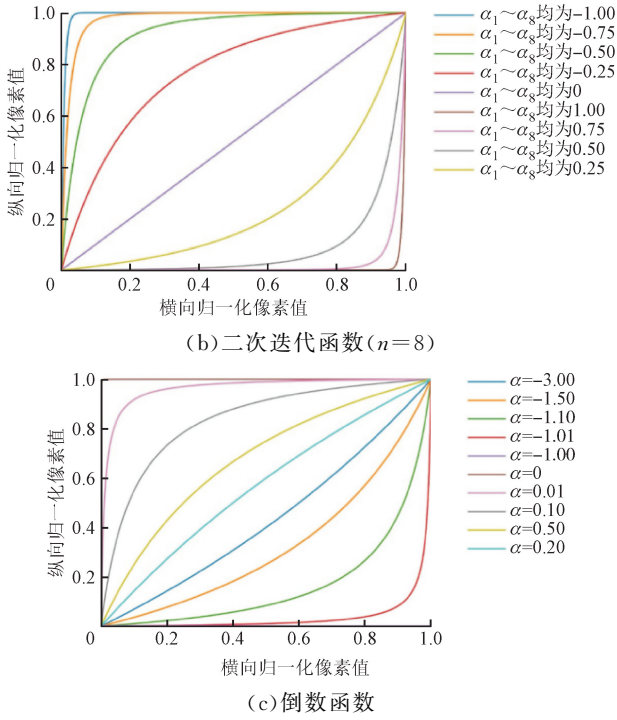
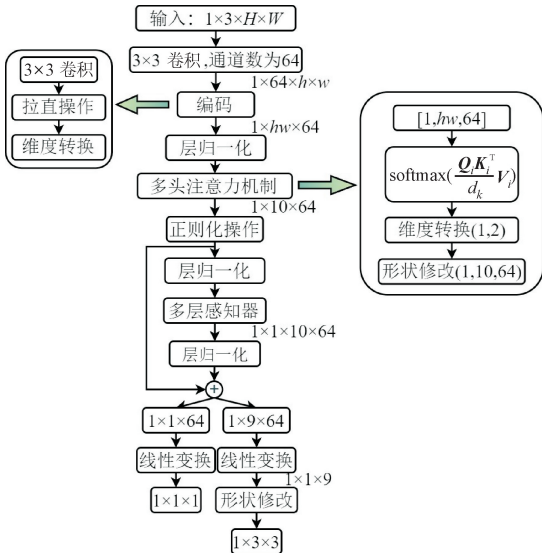


图2 自适应光照映射估计函数曲线

Fig. 2 Adaptive light mapping estimation function curve

1.2 基于 Transformer 的全局图像调整

全局图像调整过程中,首先对输入图像进行卷积,获得高维度和低分辨率的图像特征,低分辨率可以节省计算成本,有利于模型实现轻量化,较高的维度特征有利于提取图像的全局特征。然后,通过特征编码后进入多头注意力模块,捕获周围像素和单个像素之间的全局交互,获得图像的全局特征^[23-25]。全局图像调整的具体计算过程如图3所示。



h, w —第一次卷积后图像的高和宽; Q, K, V —多头注意力机制的3个向量; d_k —向量 k 的维度。

图3 全局图像调整的计算过程

Fig. 3 Calculation process of global image adjustment

进入多头注意力模块之前,首先要进行特征编码,特征序列化过程如图4所示。第一次卷积后的特征图大小为 $64 \times h \times w$, 首先将 $h \times w$ 大小的特征图的每个像素作为一个 token, 并在通道上拉直展开得到 64 维的特征向量, 即 $(h \times w) \times 64$ 的二维矩阵 X_T , 将 X_T 经过线性变换进行编码, 得到满足条件的输入序列 X_T 。

输入序列 X_T 分别乘上模型自行学到的权重 W_q, W_k, W_v 得到 Q, K, V , 并拆分得到 Q_i, K_i, V_i 在多个空间内进行注意力计算, 计算过程可表示为

$$\begin{cases} Q = X_T W_q; & K = X_T W_k; & V = X_T W_v \\ Q_i = Q W_{q_i}; & K_i = K W_{k_i}; & V_i = V W_{v_i} \end{cases} \quad (6)$$

$$Z_i = \text{Attention}(Q_i, K_i, V_i) = \text{softmax}\left(\frac{Q_i K_i^T}{\sqrt{d_k}}\right) V_i \quad (7)$$

式中: Attention 为求取注意力机制操作; softmax 为 softmax 函数; W_q, W_k, W_v 是由模型自行习得的权重; $W_{q_i}, W_{k_i}, W_{v_i}$ 分别为 Q, K, V 的第 i 个子空间的权重; Z_i 为第 i 个子空间的注意力计算结果; d_q, d_k, d_v 分别分 Q, K, V 的维度, 且 $d_q = d_k = d_v$ 。

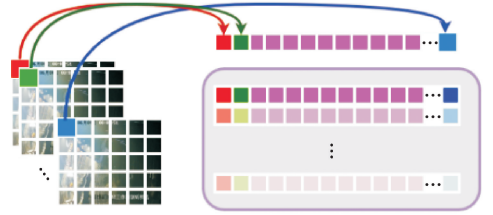


图4 特征图的序列化过程

Fig. 4 Serialization process of feature maps

多头自注意力是在多个子空间进行注意力计算, 该过程可表示为

$$Z = \text{MultiHead}(Q, K, V) = \text{Concat}(Z_1, Z_2, Z_3, Z_4) W_o \quad (8)$$

式中: MultiHead 为求取多头注意力机制操作; Concat 为拼接操作; W_o 为权重矩阵。输入序列 X_T 经过式(6)得到 Q, K, V , 分别进入不同的子空间, 通过式(7)得到所有子空间的注意力权重 $Z_1 \sim Z_4$, 通过式(8)得到多头注意力机制权重 Z 。

多头注意力的输出大小为 $1 \times 10 \times 64$, 经过多层感知机和线性变换生成一个 3×3 的颜色矩阵和一个 1×1 的伽马值。颜色矫正矩阵对 R, G, B 三通道颜色进行色彩调整, 伽马值用于全局图像照度的非线性调节。对局部增强后的图像进行全局调整的过程可表示为

$$G(X) = (\max(\sum W_{col} \times E(X), \epsilon))^{\gamma} \quad (9)$$

式中: $E(X)$ 表示局部增强后的图像; $G(X)$ 表示全局调整后的图像; W_{col} 表示颜色变换矩阵; γ 表示伽马校正的指数值; ϵ 表示一个非负最小值, 实验中设为 $\epsilon = e^{-8}$ 。

1.3 损失函数

低照度图像增强模型的整体损失 L_{total} 由空间一致性损失 L_{sps} 、曝光控制损失 L_{exp} 、色彩不变性损失 L_{col} 和特征相似性损失 L_{fs} 共 4 个损失函数联合计算, 具体可表示为

$$L_{\text{total}} = L_{\text{sps}} + L_{\text{exp}} + L_{\text{col}} + L_{\text{fs}} \quad (10)$$

为引导模型学到适合的低光和正常光的映射关系, 采用了 Zero-DCE 方法提出的 3 个非参考损失函数用于训练低光增强模型。空间一致性损失^[21]通过控制输入图像和增强图像相邻区域的梯度来保持图像的空间一致性, 可表示为

$$L_{\text{sps}} = \frac{1}{m_1} \sum_{i=1}^{m_1} \sum_{j \in \Omega(i)} (|Y_i - Y_j| - |(I_i - I_j)|)^2 \quad (11)$$

式中: m_1 表示为局部区域的数量; $\Omega(i)$ 表示以区域 i 为中心的 4 个相邻区域(上、下、左、右); Y 和 I 分别为增强图像和输入图像的局部区域平均强度值。

曝光控制损失通过测量局部区域的平均强度与正常曝光水平 E ^[26] 之间的距离控制曝光水平, 避免曝光过度或曝光不足现象, 可表示为

$$L_{\text{exp}} = \frac{1}{m_2} \sum_{m=1}^{m_2} |Y_m - E| \quad (12)$$

式中: m_2 表示不重叠的局部区域数量; Y_m 表示各个局部区域的平均强度; E 表示正常曝光水平, 根据文献^[21]和文献^[26], E 设定为 0.6。

色彩恒定损失通过建立三通道 R 、 G 、 B 之间的关系矫正潜在的彩色偏差, 使得增强后的颜色尽可能保持一致, 可表示为

$$\begin{cases} L_{\text{col}} = \sum_{\forall (p, q) \in \epsilon'} (J^p - J^q)^2 \\ \epsilon' = \{(R, G), (R, B), (G, B)\} \end{cases} \quad (13)$$

式中: J^p 和 J^q 表示增强图像对应的 p 通道和 q 通道的平均强度值; (p, q) 表示一对通道。

为提升目标检测的精度, 提出了基于 KL 散度的特征相似性损失。KL 散度又叫相对熵或信息熵, 用于衡量两个变量之间的差异^[27]。假设两个特征序列为 F_1 和 F_2 , 对应的概率分布函数为 $f_1(x)$ 和 $f_2(x)$, KL 散度可以表示为

$$\text{KL}(F_1 \| F_2) = \sum f_1(x) \log \frac{f_1(x)}{f_2(x)} \quad (14)$$

模型训练过程中, 通过 YOLOv7 的主干网络分

别抽取原始图像和增强后图像的特征, 对抽取的特征进行全局平均池化, 得到两个 $1 \times 1 \times C$ 大小的向量, 逐点计算两个向量的 KL 散度, 衡量低照度图像和增强图像的特征差异。通过约束增强前后图像的特征相似性, 避免增强过程产生对后续检测不利的影响。基于 KL 散度的特征相似性损失表示为

$$L_{\text{fs}} = \text{GAP}(F_{\text{ori}}) \log \frac{\text{GAP}(F_{\text{en}})}{\text{GAP}(F_{\text{ori}})} \quad (15)$$

式中: $\text{GAP}(F_{\text{ori}})$ 为原始低照度图像特征的全局池化操作; $\text{GAP}(F_{\text{en}})$ 为增强后图像特征的全局池化操作。

2 实验与结果分析

实验操作系统为 64 位 Windows11, CPU 为 AMD Ryzen 7 5800H with Radeon Graphics, 显卡为 NVIDIA RTX 3070 GPU。网络模型基于 Python3.9.0, Pytorch1.12.1 和 Cuda11.6 框架搭建, 开发环境为 PyCharm 2022.3.2。模型训练采用的 SICE 数据集^[28], 包含 589 个室内和室外场景的多重曝光图像, 每个场景包含 3~18 张不同曝光的低对比度图像。模型训练中, 学习率为 0.000 1, batch size 为 8, epoch 为 100, 动量为 0.9, 为避免训练过程中出现过拟合现象, 采用余弦退火降低学习率, 衰减系数为 0.000 05。

2.1 数据集和评价指标

为验证提出的低照度图像增强方法的性能, 从主观视觉和客观指标两方面, 将本文方法与最先进的方法: Retinex-Net^[15] (2018 年)、MIRNetv2^[29] (2023 年)、SCI^[30] (2022 年)、Zero-DCE^[21] (2020 年)、Zero-DCE++^[22] (2022 年) 和 EnlightenGAN^[19] (2021 年) 在 LOL 数据集^[15] 和 MIT-Adobe FiveK 数据集^[31] 上进行对比。其中, Retinex-Net 和 MIRNetv2 方法是监督学习方法, EnlightenGAN 和 SCI 是无监督学习方法, Zero-DCE 和 Zero-DCE++ 是零次学习方法。LOL 数据集是在真实场景下拍摄的配对数据集, 通过改变曝光时间和感光度 ISO 收集了 500 对大小为 400×600 像素的低光/正常光图像。MIT-Adobe FiveK 数据集包含 5 000 张图像, 每张图像由 5 位专业摄影师(A/B/C/D/E)修图并保存为 raw 格式, 本文选取了专家 A 的修图结果作为对照, 并将图像转换为 JPG 格式。

低照度图像增强对比实验采用了峰值信噪比(PSNR, 用符号 P_{SNR} 表示)和结构相似性(SSIM, 用符号 S_{SIM} 表示)作为客观评价指标。其中, 峰值信

噪比对增强前后图像像素点的灰度值进行对比计算,用于评估图像的处理质量,单位为 dB,数值越小表示失真程度越大,计算过程如下

$$P_{\text{SNR}} = 20\lg\left(\frac{I_{\text{max}}}{\sqrt{M_{\text{se}}}}\right)$$

(16)

式中: I_{max} 表示输入数据最大值; M_{se} 表示增强前后图像之间的均方误差。

结构相似性从亮度、对比度和结构信息三方面对增强前后的图像进行相似度对比,取值范围为 $[0, 1]$,数值越大表示相似度越大,具体计算过程如下

表 1 不同方法客观评价指标和模型效率对比结果

Table 1 Comparative results of objective evaluation indicators and model efficiency between different methods

方法	峰值信噪比/dB		结构相似性		计算量/ 10^9	参数量/ 10^3	推理时间/ms
	LOL	MIT-Adobe FiveK	LOL	MIT-Adobe FiveK			
Retinex-Net ^[15]	16.77	12.31	0.46	0.67	587.47	555.21	84.17
MIRNetv2 ^[29]	24.74	23.97	0.85	0.93	140.38	5 858.56	39.27
EnlightenGAN ^[19]	17.48	18.26	0.68	0.75	48.71	8 636.67	78.40
SCI ^[30]	15.00	20.54	0.52	0.76	0.09	0.26	1.75
Zero-DCE++ ^[22]	14.70	13.09	0.57	0.66	<u>2.42</u>	<u>10.56</u>	<u>1.92</u>
Zero-DCE ^[21]	14.86	13.20	0.58	0.71	5.01	79.42	2.44
本文	<u>20.18</u>	<u>23.31</u>	<u>0.80</u>	<u>0.87</u>	3.67	86.51	4.20

注:加粗表示最优结果;下划线表示次优结果。

从表 1 可以看出,所提算法在 LOL 数据集上的 P_{SNR} 达到了 20.18 dB, S_{SIM} 达到了 0.80,在 MIT-Adobe FiveK 数据集上的 P_{SNR} 达到了 23.31 dB, S_{SIM} 达到了 0.87,仅次于 MIRNetv2 方法的增强效果。同时,相比与 MIRNetv2 方法 140.38×10^9 次的计算量, 5.90×10^6 个参数和 39.27 ms 的测试时间,本文提出的方法仅需要 3.67×10^9 次的计算量, 0.09×10^6 个参数和 4.2 ms 的测试时间,所提方法的综合性能优于其他低照度增强方法。

2.3 主观视觉效果

在 LOL 数据集上将本文方法与其他六种低照度图像增强方法进行对比,选取其中 7 个场景的对比结果进行主观性分析,对比结果如图 5 所示。从图中可以看出,Retinex-Net 方法增强后,图像亮度提升明显,颜色饱和度高,但增强后的图像模糊,出现了颜色和纹理失真,如场景 2、4、6 和 7。MIRNetv2 方法增强后,整体色彩恢复比较真实,亮区的增强效果较好,如场景 1、2、3 和 5,暗区的增强效果较差,如场景 4 和 7。EnlightenGAN 方法增强后的图像颜色偏黄如场景 7,亮度提升有限如场景 4 和 6。SCI 方法增强后的图像未出现颜色和纹理失

$$S_{\text{SIM}}(\mathbf{X}, \mathbf{Y}) = \frac{(2\mu_{\mathbf{X}}\mu_{\mathbf{Y}} + C_1)(2\sigma_{\mathbf{XY}} + C_2)}{(\mu_{\mathbf{X}}^2 + \mu_{\mathbf{Y}}^2 + C_1)(\sigma_{\mathbf{X}}^2 + \sigma_{\mathbf{Y}}^2 + C_2)}$$

(17)

式中: $\mu_{\mathbf{X}}, \mu_{\mathbf{Y}}$ 表示图像均值; $\sigma_{\mathbf{X}}, \sigma_{\mathbf{Y}}$ 表示标准差; $\sigma_{\mathbf{XY}}$ 表示图像 $\mathbf{X}$ 和 $\mathbf{Y}$ 的协方差。

2.2 客观评价指标

选取峰值信噪比和结构相似性作为客观评价指标,将本文方法与其他 6 种低照度增强方法在 LOL 数据集和 MIT-Adobe FiveK 数据集上进行对比。选取参数量、计算量和单张图像测试时间,在 400×600 像素的图像上进行效率对比,结果如表 1 所示。

真现象,但整体亮度和对比度较低,如场景 4 和 7。Zero-DCE++方法的照度调整不足,增强后的图像整体偏暗,出现严重偏色现象,视觉效果较差,如场景 2、3 和 5 的偏蓝现象非常明显。Zero-DCE 方法增强后的图像没有出现严重的偏色现象,但整体亮度提升不足,如场景 2、4 和 7。本文方法增强后的图像未出现颜色偏差和纹理失真现象,颜色饱和度较好,亮度提升明显,整体视觉效果较好。

在 MIT-Adobe FiveK 数据集上将本文方法与其他 6 种低照度图像增强方法进行对比,选取其中 6 个场景的增强结果进行主观性分析,对比结果如图 6 所示。从图中可以看出,Retinex-Net 增强后的图像颜色过度饱和,如场景 1、场景 6 和场景 7,纹理过于突出如场景 2,该方法难以恢复较好的视觉效果。MIRNetv2 方法对亮区的增强效果较好,如场景 1、2 和 5,暗区增强后略有失真,如场景 3 中人脸面部模糊,细节丢失,且不同光源下的颜色失真程度不同,如场景 3 和 4。EnlightenGAN 增强后的图像出现了明显的色彩偏差,整体图像偏黄严重,无法恢复图像原本的色彩,如场景 1 的水面和波光、场景 3 的人脸和衣服条纹、场景 4 的墙面和场景 5 玻璃上的

白色贴纸,并且出现了严重的光晕和细节丢失现象,如场景 6 右上角和场景 3 左上角。SCI 增强后图像的整体亮度提升明显,但暗区域照度提升有限,如场景 3 的人物背景和场景 6 的地面,亮区域容易增强过度产生严重的曝光现象,如场景 1 的喷泉和场景 5 的右侧部分。Zero-DCE++ 增强后图像的偏色现象不明显,但依旧存在,如场景 4 的白色墙壁和场景 5 的整体光线,且整体增强效果较暗。Zero-DCE 的增强效

果略有提升,但整体亮度依旧较暗,如场景 4 的墙壁。暗区增强后噪点明显,如场景 3 左侧阴影区域和场景 7 地面。增强后细节较为模糊,如场景 6 中墙上的字母。本文所提方法增强后的图像未出现明显的光晕、偏色和细节丢失现象,色彩恢复效果较好,照度提升明显。在主观视觉效果方面,本文提出的方法优于其他 6 种低照度图像增强方法,两个数据集上的增强结果表明所提方法的泛化性和鲁棒性较好。

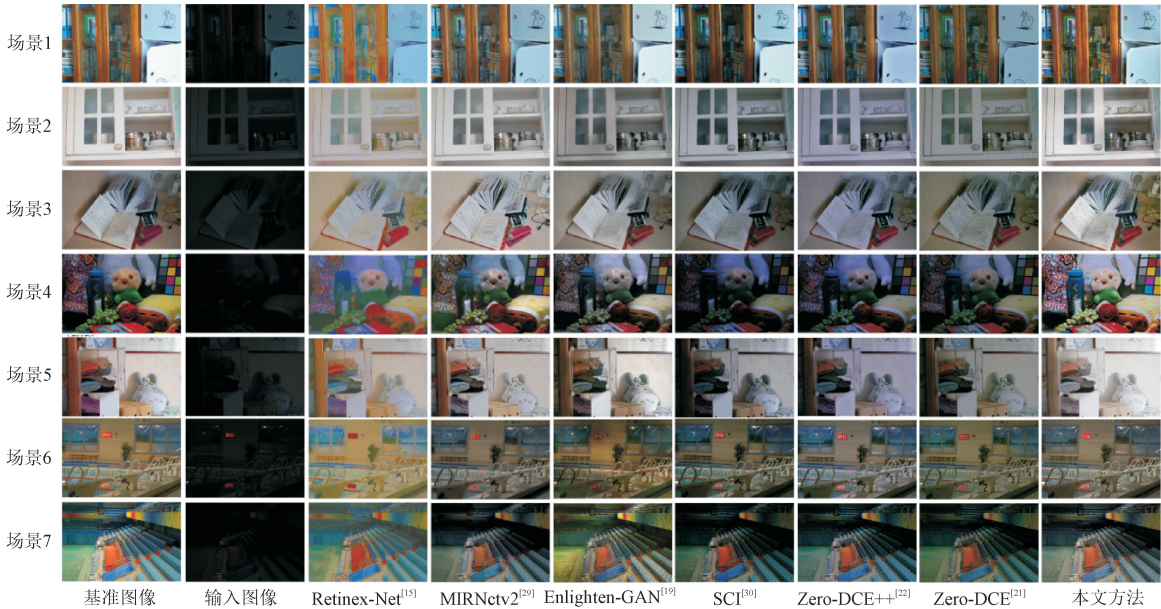


图 5 LOL 数据集上的低照度图像增强结果

Fig. 5 Low-light image enhancement results on LOL dataset



图 6 MIT-Adobe FiveK 数据集上的低照度图像增强结果

Fig. 6 Low-light image enhancement results on MIT-Adobe FiveK dataset

为验证本文提出的方法在真实低照度条件下室内和室外实拍图像上的增强效果,将本文方法与其他 6 种低照度图像增强方法进行对比。在夜间,于

中国矿业大学(北京)校园内 8 处景点拍摄真实低照度图像,拍摄设备为 iPhone13,图像最大分辨率为 $4\,032\times2\,024$ 像素,这些图像没有参与模型的训练,

仅用以验证提出方法的适应性,增强对比结果如图 7 所示。从图中可以看出,Retinex-Net 增强后图像的亮度和对比度明显提升,但颜色饱和度过高如场景 3 和场景 5,纹理过于突出导致图像失真严重如场景 1 和场景 4,场景 3 增强后黑暗处出现明显光晕。MIRNetv2 方法增强后的图像整体泛白如场景 4 和场景 6,部分细节纹理丢失如场景 2 中的桌面和场景 5 中的台阶,场景 3 增强后的天空产生了明显光晕。EnlightenGAN 增强后图像的整体亮度和对比度提升明显,但仍有不足如场景 3 和场景 5,颜色偏黄如场景 2 和场景 7,场景 3 和场景 6 增强后的天空以及场景 8 中的窗户均产生了明显的光晕现象。SCI 方法增强后的图像颜色保真度

较好,但明亮区域过度增强产生了严重的曝光现象,导致部分特征丢失,如场景 3 中的窗户和场景 4 中水塔,图像暗处增强不足如场景 7 和场景 8 的地面。Zero-DCE++ 增强后图像的整体亮度不足,偏色(蓝)现象明显,如场景 4 中的白色石柱和场景 7 中的白色房子等。Zero-DCE 增强后图像的细节纹理保留较好,未出现纹理失真现象,偏色(蓝)现象有所缓解,但依然存在,如场景 2 和场景 7,图像的整体亮度提升不足如场景 1、场景 5 和场景 7。本文所提方法增强后图像的亮度和对比度提升明显,颜色饱和度较好,未出现颜色失真和纹理失真现象,在室内和室外的真实低照度条件下有较好的增强效果。



图 7 低照度条件下实拍图像的增强结果

Fig. 7 Enhancement results of real images under low illumination conditions

2.4 目标检测及特征可视化

为探究本文提出的增强方法对目标检测性能的影响,基于 YOLOv7 在 Exdark 数据集上与其他 6 种方法增强后的检测结果进行对比。低照度图像增强模型联合 YOLOv7 检测模型的应用框架如图 8 所示,表 2 中的“原始图像+YOLOv7”表示使用图 8 中的方式 a,“增强方法+YOLOv7”表示使用图 8 中的方式 b,即原始图像首先经过低照度图像增强

模型进行增强,增强后的结果作为 YOLOv7 检测网的输入。对比结果如表 2 所示,可以看出,Retinex-Net 和 EnlightenGAN 增强后图像的检测精度分别降低了 1.6%和 0.7%,在人眼视觉上取得了一定的提升,但对于目标检测损失了重要的特征信息。MIRNetv2、Zero-DCE++ 和 Zero-DCE 增强后的检测精度分别提升了 5.6%、4.2%和 4.4%,本文方法增强后的目标检测精度提升了 7.6%,检测时间

仅需 42.23 ms,因此本文所提方法更有助于低照度条件下的目标检测任务。

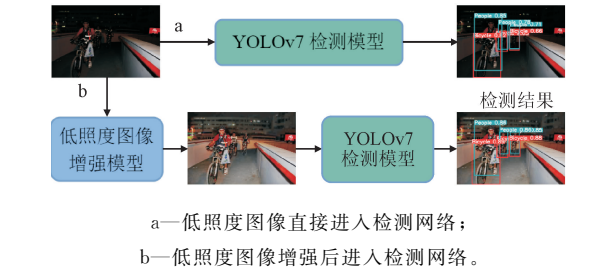


图 8 低照度图像增强模型联合检测网络的应用架构
Fig. 8 Application architecture diagram of joint detection network for low-light image enhancement networks

表 2 ExDark 数据集上的目标检测对比结果
Table 2 Comparison results of target detection on ExDark dataset

方法	平均检测精度/%	推理时间/ms
原始图像+YOLOv7	71.2	38.01
Retinex-Net+YOLOv7	69.6	123.34
MIRNetv2+YOLOv7	<u>76.8</u>	77.37
EnlightenGAN+YOLOv7	70.5	116.21
SCI+YOLOv7	76.3	<u>39.36</u>
Zero-DCE++YOLOv7	75.4	39.94
Zero-DCE+YOLOv7	75.6	40.56
本文+YOLOv7	78.8	42.23

注:加粗表示最优的结果;下划线表示次优的结果。

除此之外,为进一步探究提出的低照度增强方法对目标检测的影响,在 ExDark 数据集上,基于 YOLOv7 对原始图像和增强后图像进行目标检测,并对检测目标进行特征可视化,直观显示出低照度图像和增强图像对特征提取和目标识别的影响,结果如图 9 所示。可以看出,正常亮度区域内的目标,图像增强前后的检测结果差距不大,如场景 1 中的第一个目标(增强前检测精度为 85%,增强后为 86%)和场景 2 中的前 3 个目标(增强前检测精度为 89%、89%和 83%,增强后为 89%、86%和 88%)。低照度区域内,增强前后的目标检测结果差异很大,主要表现在 3 个方面:黑暗中的目标无法识别、相邻目标无法区分以及目标特征边界不明显。例如,场景 2 中低照度图像的目标 7 和 8,场景 3 中低照度图像的目标 2 和 3 无法识别,对应的增强后的图像可以准确检测出目标。场景 4 中,低照度图像中的目标 1 与目标 2 的特征边界无法区分、检测框定位错误,对应的增强后的图像中的目标特征定位准确。场景 1 和 2 中原始图像的对比度低,目标特征无法区分,对应可以准确检测出目标。场景 4 中,低照度图像中的目标 1 与目标 2 的特征边界无法区分、检测框定位错误,对应的增强后的图像中的目标特征定位准确。场景 1 和 2 中原始图像的对比度低,目标特征无法区分,对应的增强后的图像的特征位置集中且边界明显。

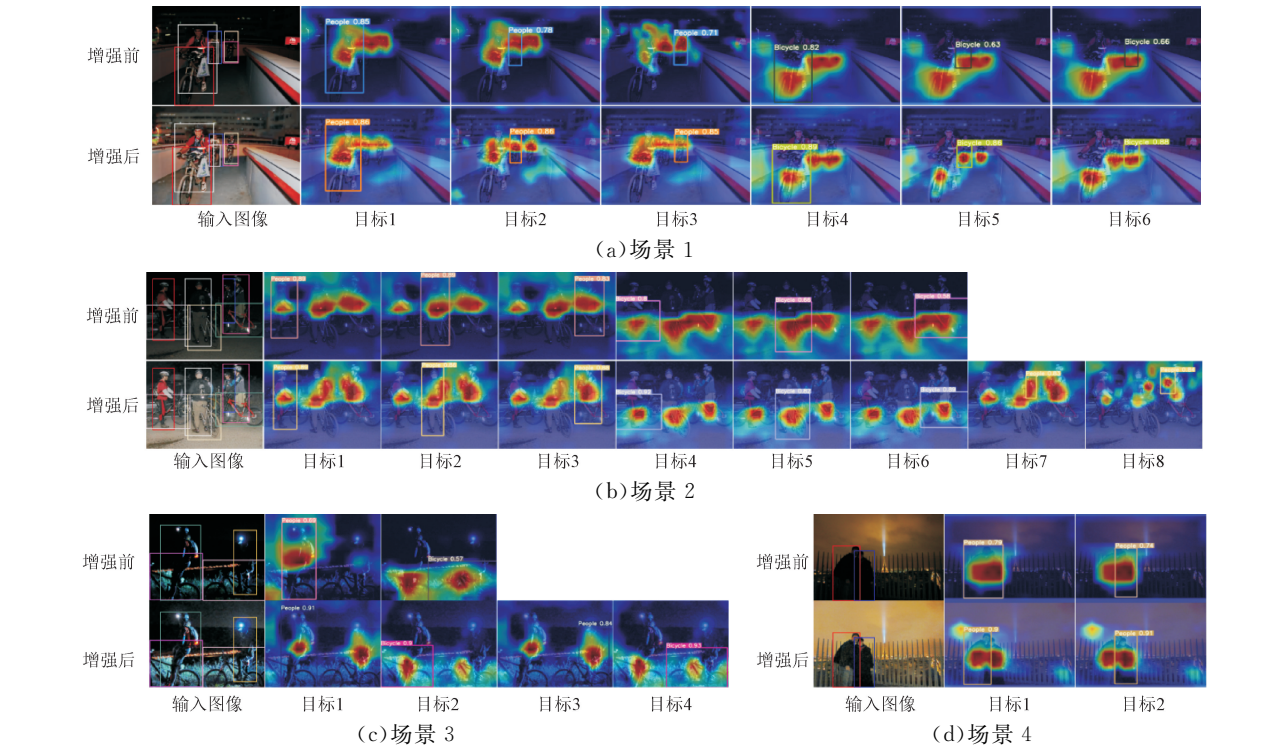


图 9 不同场景下图像增强前后的目标检测和特征可视化结果
Fig. 9 Target detection and feature visualization results before and after image enhancement in different scenes

2.5 消融实验分析

为探究光照映射估计曲线和全局调整的必要性 和有效性,进行了相关的消融实验,测试图像是在真 实场景下拍摄的低照度图像,拍摄设备为 iPhone13, 图像最大分辨率为 $4\,032\times2\,024$ 像素。主观视觉对 比结果如图 10 所示。其中:Global-false 表示去掉 全局图像调整,仅保留局部增强;Curve-quadratic iterative 表示局部增强采用二次迭代函数作为自适

应光照映射估计函数。从图中可以看出,去除全局 图像调整后,整体亮度提升依旧明显,但部分区域过 度增强,如场景 2 和 3 中的天空增强过度,噪点明 显,细节失真,场景 4 中白色框内的路灯被过度增 强导致严重曝光。光照映射估计曲线采用二次迭 代函数增强后,亮度增强依旧明显,但边缘纹理突 出,细节信息丢失严重,亮度对比度不足,如场景 2 和 3。



图 10 消融实验对比结果

Fig. 10 Comparative results of ablation experiments

为验证特征相似性损失函数对目标检测性能的 影响,在 Exdark 数据集上进行了消融实验,结果如 表 3 所示。 L_{dec} 为 YOLOv7 的检测损失函数, L_{spa} 、 L_{exp} 、 L_{col} 分别为空间一致性损失、曝光控制损失和色 彩不变性损失函数, L_{fs} 为特征相似性损失函数。从 结果可以看出,照度增强后目标检测的性能提升明 显(5.3%),特征相似性损失通过约束目标特征,使 得增强后图像的目标特征更有利于目标检测任务, 目标检测的精度提升了 2.3%。

表 3 损失函数对目标检测性能的影响

Table 3 Impact of loss function on object detection performance

L_{dec}	$L_{spa}+L_{exp}+L_{col}$	L_{fs}	平均检测精度/%
✓			71.2
✓	✓		<u>76.5</u>
✓	✓	✓	78.8

注:加粗表示最优的结果;下划线表示次优的结果。

3 结 论

针对低照度图像存在的多种退化现象,现有的 图像增强方法存在颜色失真、亮度提升有限、细节损 失严重和噪声放大的问题,严重影响了图像的可视 化效果和目标检测精度。为解决上述问题,提出了一 种从局部到全局的零参考低照度图像增强方法。

(1)提出的低照度图像增强方法仅需要 3.67×10^9 次的计算量和 0.09×10^6 个参数,单张图像测试 时间仅 4.2 ms,具有轻量化优势,便于模型部署。

(2) LOL 数据集、MIT-Adobe FiveK 和真实低 照度图像上的实验结果表明,所提方法有效提高了 低照度图像的客观评价指标和主观视觉效果,未出 现伪影和噪声放大问题,增强了图像的整体亮度和 对比度,避免了颜色失真和细节损失现象。

(3) ExDark 数据集上的目标检测和特征可视 化结果表明,所提方法增强后图像的目标轮廓更加

清晰,定位更加准确,提高了低照度图像的目标检测的精度,证明了所提方法的有效性。

参考文献:

- [1] LIU Jiaying, XU Dejie, YANG Wenhan, et al. Benchmarking low-light image enhancement and beyond [J]. *International Journal of Computer Vision*, 2021, 129(4): 1153-1184.
- [2] LOH Y P, CHAN C S. Getting to know low-light images with the exclusively dark dataset [J]. *Computer Vision and Image Understanding*, 2019, 178: 30-42.
- [3] CHEN Chen, CHEN Qifeng, XU Jia, et al. Learning to see in the dark [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2018: 3291-3300.
- [4] IBRAHIM H, PIK KONG N S. Brightness preserving dynamic histogram equalization for image contrast enhancement [J]. *IEEE Transactions on Consumer Electronics*, 2007, 53(4): 1752-1758.
- [5] 李天钢, 王素品, 秦辰. 基于信息熵窗的小波低频子带弱目标图像的增强 [J]. *西安交通大学学报*, 2006, 40(2): 187-190.
LI Tiangang, WANG Supin, QIN Chen. Enhancement of weak objective image with wavelet low frequency sub-band based on information entropy window [J]. *Journal of Xi'an Jiaotong University*, 2006, 40(2): 187-190.
- [6] LI Mading, LIU Jiaying, YANG Wenhan, et al. Structure-revealing low-light image enhancement via robust Retinex model [J]. *IEEE Transactions on Image Processing*, 2018, 27(6): 2828-2841.
- [7] GU Zhihao, LI Fang, FANG Faming, et al. A novel Retinex-based fractional-order variational model for images with severely low light [J]. *IEEE Transactions on Image Processing*, 2020, 29: 3239-3253.
- [8] LI Lin, WANG Ronggang, WANG Wenmin, et al. A low-light image enhancement method for both denoising and contrast enlarging [C]//2015 IEEE International Conference on Image Processing (ICIP). Piscataway, NJ, USA: IEEE, 2015: 3730-3734.
- [9] DONG Xuan, WANG Guan, PANG Yi, et al. Fast efficient algorithm for enhancement of low lighting video [C]//2011 IEEE International Conference on Multimedia and Expo. Piscataway, NJ, USA: IEEE, 2011: 1-6.
- [10] WANG Wencheng, WU Xiaojin, YUAN Xiaohui, et al. An experiment-based review of low-light image enhancement methods [J]. *IEEE Access*, 2020, 8: 87884-87917.
- [11] LÜ Feifan, LU Feng, WU Jianhua, et al. MBLLEN: low-light image/video enhancement using CNNs [C]//British Machine Vision Conference 2018. Guildford, UK: BMVA Press, 2018: 20193807460225.
- [12] LI Chongyi, GUO Jichang, PORIKLI F, et al. LightenNet: a convolutional neural network for weakly illuminated image enhancement [J]. *Pattern Recognition Letters*, 2018, 104: 15-22.
- [13] YANG Wenhan, WANG Wenjing, HUANG Haofeng, et al. Sparse gradient regularized deep Retinex network for robust low-light image enhancement [J]. *IEEE Transactions on Image Processing*, 2021, 30: 2072-2086.
- [14] WANG Ruixing, ZHANG Qing, FU C W, et al. Underexposed photo enhancement using deep illumination estimation [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2019: 6842-6850.
- [15] WEI Chen, WANG Wenjing, YANG Wenhan, et al. Deep Retinex decomposition for low-light enhancement [C]//British Machine Vision Conference 2018. Guildford, UK: BMVA Press, 2018: 20193807460275.
- [16] ZHANG Yonghua, ZHANG Jiawan, GUO Xiaojie. Kindling the darkness: a practical low-light image enhancer [C]//Proceedings of the 27th ACM International Conference on Multimedia. New York, NY, USA: Association for Computing Machinery, 2019: 1632-1640.
- [17] ZHANG Yonghua, GUO Xiaojie, MA Jiayi, et al. Beyond brightening low-light images [J]. *International Journal of Computer Vision*, 2021, 129(4): 1013-1037.
- [18] YU Runsheng, LIU Wenyu, ZHANG Yasen, et al. DeepExposure: learning to expose photos with asynchronously reinforced adversarial learning [C]//Proceedings of the 32nd International Conference on Neural Information Processing Systems. Red Hook, NY, USA: Curran Associates Inc., 2018: 2153-2163.
- [19] JIANG Yifan, GONG Xinyu, LIU Ding, et al. EnlightenGAN: deep light enhancement without paired supervision [J]. *IEEE Transactions on Image Processing*, 2021, 30: 2340-2349.
- [20] ZHANG Lin, ZHANG Lijun, LIU Xiao, et al. Zero-shot restoration of back-lit images using deep internal learning [C]//Proceedings of the 27th ACM International Conference on Multimedia. New York, NY, USA: Association for Computing Machinery, 2019: 1623-1631.

- [21] GUO Chunle, LI Chongyi, GUO Jichang, et al. Zero-reference deep curve estimation for low-light image enhancement [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2020: 1777-1786.
- [22] LI Chongyi, GUO Chunle, 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] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16x16 words; transformers for image recognition at scale [C]//ICLR 2021. New York, USA: ICLR, 2021: 3458.
- [24] 高峰, 汲胜昌, 郭洁, 等. 采用对比学习的多阶段 Transformer 图像去雾方法 [J]. 西安交通大学学报, 2023, 57(1): 195-210.
- GAO Feng, JI Shengchang, GUO Jie, et al. A multi-stage transformer network for image dehazing based on contrastive learning [J]. Journal of Xi'an Jiaotong University, 2023, 57(1): 195-210.
- [25] 杨媛玮, 王华珂, 侯兴松. 从全局到局部: 双注意力融合去雾网络 [J]. 西安交通大学学报, 2023, 57(7): 191-200.
- YANG Aiwei, WANG Huake, HOU Xingsong. From global to local: a dual-attention fusion dehazing network [J]. Journal of Xi'an Jiaotong University, 2023, 57(7): 191-200.
- [26] MERTENS T, KAUTZ J, VAN REETH F. Exposure fusion: a simple and practical alternative to high dynamic range photography [J]. Computer Graphics Forum, 2009, 28(1): 161-171.
- [27] 甄志龙, 张居晓. 卡方统计中基于 KL 散度的高维文本数据特征筛选 [J]. 统计与决策, 2022, 38(17): 43-46.
- ZHEN Zhilong, ZHANG Juxiao. Feature screening for high dimensional text data based on KL divergence in chi-squared statistics [J]. Statistics & Decision, 2022, 38(17): 43-46.
- [28] CAI Jianrui, GU Shuhang, ZHANG Lei, et al. Learning a deep single image contrast enhancer from multi-exposure images [J]. IEEE Transactions on Image Processing, 2018, 27(4): 2049-2062.
- [29] 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.
- [30] MA Long, MA Tengyu, LIU Risheng, et al. Toward fast, flexible, and robust low-light image enhancement [C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2022: 5627-5636.
- [31] BYCHKOVSKY V, PARIS S, CHAN E, et al. Learning photographic global tonal adjustment with a database of input/output image pairs [C]//CVPR 2011. Piscataway, NJ, USA: IEEE, 2011: 97-104.

(编辑 刘杨 陶晴)