

# 基于结构和纹理感知的井下低光照自监督图像增强方法

潘 杉<sup>1</sup>, 于 挺<sup>1</sup>, 陈 伟<sup>1,2,3</sup>, 田子建<sup>1</sup>, 岳中文<sup>4</sup>

(1. 中国矿业大学(北京) 人工智能学院, 北京 100083; 2. 中国矿业大学计算机科学与技术学院, 江苏 徐州 221000; 3. 煤矿智能化与机器人创新应急管理中心重点实验室, 江苏 徐州 221000; 4. 中国矿业大学(北京) 力学与土木工程学院, 北京 100083)

**摘要:**受井下复杂空间环境和不均匀人造光源影响, 井下视觉设备采集的图像易呈现整体或部分区域光线不足、图像内容可见度差的问题。对于井下低光照图像的增强, 现有图像增强方法的结果易出现对比度差以及部分区域过曝和欠曝的问题。基于此, 提出一种基于结构和纹理感知的井下低光照自监督图像增强方法, 以摆脱训练时对配对的井下低光照/正常光照图像的依赖。首先, 为产生分段平滑的光照图, 设计了一种自监督结构和纹理感知的光照估计网络, 根据自监督训练损失保留场景的边缘结构并平滑纹理细节。为了深入挖掘低光照图像中的局部纹理特征和全局结构特征来提高光照估计网络的性能, 在光照估计网络中引入了局部-全局感知模块。该模块利用卷积操作中较小感受也能够捕获局部特征的能力以及视觉 Transformer 的自注意力机制能够实现全局信息交互的特性来实现低光照图像中局部和全局特征的提取。其次, 为了实现自监督学习的训练方式, 针对光照图分段平滑的特性, 采用了一种结构感知的平滑损失。为了进一步细化光照估计网络产生的光照图, 使其具有合理的亮度和对比度, 引入了伪标签图像生成器来合成具有良好对比度和亮度的伪标签图像。通过重建损失约束, 增强图像与伪标签图像之间亮度和对比度之间的一致性, 从而间接约束光照图。在多个公开的基准数据集和井下真实场景的低光照图像数据集上进行主观和客观评价的试验结果表明: 该方法具有较好的低光照图像增强效果, 在面对井下场景的低光照图像时, 该方法也具有良好的泛化性能。

**关键词:**低光照图像; 自监督; 图像增强; 光照估计网络; 局部-全局感知

中图分类号: TP391 文献标志码: A 文章编号: 0253-9993(2025)04-2310-11

## Underground low-light self-supervised image enhancement method based on structure and texture perception

PAN Shan<sup>1</sup>, YU Ting<sup>1</sup>, CHEN Wei<sup>1,2,3</sup>, TIAN Zijian<sup>1</sup>, YUE Zhongwen<sup>4</sup>

(1. Department of Artificial Intelligence, China University of Mining and Technology(Beijing), Beijing 100083, China; 2. School of Computer Science & Technology, China University of Mining and Technology, Xuzhou 221000, China; 3. Key Laboratory of Coal Mine Intelligent and Robot Innovation Emergency Management Department, Xuzhou 221000, China; 4. Department of Mechanics and Civil Engineering, China University of Mining and Technology(Beijing), Beijing 100083, China)

**Abstract:** Due to the complex spatial environment and uneven artificial lighting underground, images captured by underground visual equipment often suffer from insufficient overall or partial lighting and poor visibility of image content. Existing image enhancement methods for low-light underground images often result in poor contrast and issues with overex-

收稿日期: 2024-03-06 策划编辑: 郭晓炜 责任编辑: 黄小雨 DOI: 10.13225/j.cnki.jecs.2024.0234

基金项目: 国家自然科学基金资助项目(52074305, 51874300); 国家重点研发计划资助项目(2021YFC2902103)

作者简介: 潘 杉(1993—), 女, 安徽淮南人, 博士研究生。E-mail: panshan0416@163.com

通讯作者: 田子建(1964—), 男, 湖南望城人, 教授, 博士生导师。E-mail: tianzj0726@126.com

引用格式: 潘杉, 于挺, 陈伟, 等. 基于结构和纹理感知的井下低光照自监督图像增强方法[J]. 煤炭学报, 2025, 50(4): 2310-2320.

PAN Shan, YU Ting, CHEN Wei, et al. Underground low-light self-supervised image enhancement method based on structure and texture perception[J]. Journal of China Coal Society, 2025, 50(4): 2310-2320.



移动阅读

posure and underexposure in certain areas. Within this article, we propose a self-supervised image enhancement method for low-light underground conditions based on structural and texture perception, aiming to alleviate the dependence on paired low-light/normal-light images during training. Firstly, to generate smoothly segmented illumination maps, we design a self-supervised structural and texture-aware illumination estimation network, which preserves scene edge structures and smooths texture details based on self-supervised training losses. To further exploit local texture features and global structural features in low-light images to improve the performance of the illumination estimation network, we introduce a local-global perception module into the illumination estimation network. This module leverages the ability of convolutional operations with small receptive fields to capture local features and the self-attention mechanism of visual transformers to facilitate global information interaction, thus extracting local and global features from low-light images. Secondly, to facilitate self-supervised learning, we adopt a structure-aware smoothness loss considering the segmented smoothness property of illumination maps. Finally, to refine the illumination maps generated by the illumination estimation network for reasonable brightness and contrast, we introduce a pseudo-label image generator to synthesize pseudo-label images with good contrast and brightness. By constraining the consistency between brightness and contrast of the enhanced images and pseudo-label images through reconstruction loss, we indirectly constrain the illumination maps. Experimental results on multiple public benchmark datasets and a dataset of low-light images in real underground scenes demonstrate the effectiveness of our method in enhancing low-light images, as well as its good generalization performance when faced with low-light images in underground scenarios.

**Key words:** low-light images; self-supervision; image enhancement; illumination estimation network; local-global aware

## 0 引 言

“智慧矿山”的提出,为新一代人工智能技术与能源矿山的深度融合以及构建无人化、智能化绿色矿山提供了契机和有力支撑<sup>[1]</sup>。煤矿井下光照条件差,导致成像设备捕获的图像存在整体或部分区域光线不足的问题,进而导致细节信息损失严重,图像内容可见度差。然而,完成煤矿井下关键任务(如危险行为识别、监控视频分析和炮孔检测识别)都依赖具有良好对比度和亮度的图像作为输入<sup>[2]</sup>。因此,煤矿井下低光照图像增强作为底层计算机视觉任务对提升智能化设备的可靠性具有重要意义。

目前,研究人员提出了大量低光照图像增强的相关工作<sup>[3-4]</sup>,主要分为基于非学习的低光图像增强方法和基于学习的低光图像增强方法。基于非学习的图像增强方法是基于先验知识设计相应的方法,主要包括直方图均衡化、引导滤波的方法和基于Retinex理论的方法。直方图均衡化<sup>[5-7]</sup>通过调整图像中每个灰度级的像素数量来提高图像的对比度和亮度。引导滤波的方法有多种,龚云等<sup>[8]</sup>提出结合HSV空间和加权分布自适应Gamma校正的同态滤波算法,该方法在提高图像亮度的同时,有效保留了煤矿井下环境图像的细节信息。张立亚等<sup>[9]</sup>在HSV空间变换条件下将改进双边滤波算法和多尺度Retinex算法<sup>[10-11]</sup>进行融合,实现了对煤矿井下模糊图像的增强。Retinex理

论认为,拍摄的图像由照度分量和反射分量共同决定,反射分量由物体本身的反射性质决定,照度分量则决定了图像的亮度和对比度。Retinex分解后调整照度分量的动态范围,可以实现低光图像增强。最近,GUO等<sup>[12]</sup>基于Retinex理论提出了LIME方法。通过选取输入图像中各像素通道的最大值构建初始光照图,再利用结构化先验知识对初始光照图进行细化,最后基于Retinex理论变换将反射图作为增强结果。虽然LIME方法采用了Gamma校正的方式对细化后的光照图进行了亮度提升,但面对不同的场景需要调整参数设置,参数设置不当会产生过曝光现象。

基于学习的低光照图像增强方法通过学习图像的潜在特征,并利用这些特征来增强图像,避免了依据经验调整参数所带来的问题。WEI等<sup>[13]</sup>基于Retinex理论提出了用于低光照图像增强的深度学习模型RetinexNet。该模型设计了2个网络结构,分别实现了照度分量和反射分量的分解阶段,照度分量和反射分量的调整阶段以及照度分量和反射分量的重建阶段。作为将Retinex理论嵌入深度学习网络的首次尝试,该方法有效提高了图像的对比度和亮度,恢复了低光照图像的细节内容,但存在对训练数据依赖性大的问题,且容易产生过冲伪影效果。为解决这一问题,YANG等<sup>[14]</sup>为RetinexNet添加了新的约束,进一步优化了RetinexNet的结构,获得了比RetinexNet更好的结果。然而优化的RetinexNet结构中的可训练

参数少,模型性能受限。ZHANG 等<sup>[15]</sup>提出了由 3 个子网络组成的 KinD 算法,分别用于图像分解、反射分量纠正和照度分量调整,通过对分量进行纠正来得到更清晰的增强图像。但该方法在网络分解时容易造成反射分量和照度分量之间的不一致性,从而产生伪影。为解决 KinD 增强结果中的视觉伪影,ZHANG 等<sup>[16]</sup>又提出了 KinD++ 网络。该网络将原空间解耦为 2 个更小的空间,使用多尺度光照注意模块减少了视觉伪影,同时可以根据不同的需求调整亮度。为了学习输入图像到光照的映射,WANG 等<sup>[17]</sup>提出了一个欠曝光图像增强网络 DeepUPE,通过约束增强图像与标签图像的一致性,来间接约束光照估计网络。但该方法受限于标签图像的亮度和对比度,泛化能力较差。受深度图像先验知识的启发<sup>[18]</sup>,ZHAO 等<sup>[19]</sup>提出的 RetinexDIP 将高斯噪声图像作为输入,分别生成反射分量和照度分量,将 Retinex 的解耦问题转化为了图像生成问题。但该方法进行低光照图像增强时需对每个输入图像进行一定次数的迭代,消耗大量的计算时间,网络的实时性较差。LI 等<sup>[20]</sup>提出了一种不依赖于配对或非配对训练数据集的无参考图像增强算法,称为 Zero-DEC,该方法通过为给定的低光图像学习特定的高阶曲线,来实现低光图像增强。尽管通过卷积网络学习高阶曲线的参数是一种巧妙的自监督学习方法,但设计的高阶曲线还需要进一步优化以解决面对较暗的井下低光照图像时亮度提升不足的问题。为提高高阶曲线参数估计的准确性,ZHANG 等<sup>[21]</sup>提出了一种轻量级变压器来提高特征提取能力,并预测 Zero-DEC 的高阶曲线。MA 等<sup>[22]</sup>提出了一种自校准学习框架 SCI,该方法是通过构建自校准模块实现级联照明的学习过程,并通过权重共享来处理这项任务,确保了增强图像保持良好的平滑特性,同时抑制了过度曝光。

尽管很多先进的研究方法不断涌现,但井下缺乏自然光且人造光源不均匀,使得基于非学习的方法需要频繁调整参数以适应不同场景需求。此外,尽管有监督学习可以根据正常光照数据反馈模型性能并进行优化,但井下正常光照图像的获取以及与低光照图像的对齐处理难度大,难以产生大规模高质量的配对数据以满足有监督学习的训练,导致基于监督学习的方法存在泛化性差的问题。现有的基于自监督学习的低光照图像增强方法产生的增强效果仍需进一步提升,以满足实际井下应用的需求。由于井下复杂的空间环境以及人造光源无法产生均匀的光照,低光照图像中纹理细节和边缘结构特征不够明显,此处的边缘结构主要指受井下人造光源的影响,在图像中反映

为场景中明暗交界的边缘;而细节纹理主要指在场景中,受光源影响时物体表面存在的粗糙纹理特性。因此,进行 Retinex 分解时产生理想的具有分段平滑的光照图是一个挑战。

针对以上存在的问题,提出一种基于结构和纹理感知的井下低光照自监督图像增强方法。该方法首先设计了一种光照估计网络,利用自监督训练损失保留场景的边缘结构并平滑纹理细节,以产生分段平滑的光照图。受井下复杂光照环境的影响,统一的特征提取方法对结构和纹理感知不够显著。为了深入挖掘低光照图像中的局部纹理特征和全局结构特征以提高光照估计网络的性能,在光照估计网络中引入了局部-全局感知模块,简称为“LA-GA”模块,其中 LA 分支表示局部特征提取分支,GA 分支表示全局特征提取分支。为了实现自监督学习的训练方式,摆脱模型训练对标签图像的依赖,针对光照分段光滑的特性,采用了一种结构感知的平滑损失对模型进行训练。为了进一步细化光照估计网络产生的光照图,使其具有合理的亮度和对比度,引入了伪标签图像生成器来合成具有良好对比度和亮度的伪标签图像。通过重建损失约束增强图像与伪标签图像之间亮度和对比度间的一致性,从而间接约束光照图。

## 1 井下低光照自监督图像增强网络

为摆脱模型训练对配对的低光照/正常光照图像的依赖,提出一种井下低光照自监督图像增强网络,从而实现复杂矿井场景下的低光照图像增强。图 1 为网络整体结构图,其中具有可训练参数的为结构和纹理感知的光照估计网络,黑色箭头和红色箭头表示训练过程中数据的流向,虚线箭头表示损失函数约束下的梯度传播方向。结构和纹理感知的光照估计网络模型训练好以后,仅保留红色箭头部分,即可获得井下低光照图像增强结果。

在模型训练过程中,对于给定的井下低光照图像  $S$ ,一方面在结构感知的平滑损失约束下,结构和纹理感知的光照估计网络生成具有分段平滑的光照图  $L_n$ ,并基于 Retinex 模型计算得到增强图像  $Y_n$ 。另一方面,通过随机生成一组 Gamma 值,对给定的井下低光照图像  $S$  进行 Gamma 变换,以生成具有不同对比度和亮度的参考图像,其中包括  $N$  张比输入图像对比度和亮度高的参考图像和  $N$  张比输入图像对比度和亮度低的参考图像。此外,为了更好地计算参考序列中不同区域的对比度和亮度,将给定的井下低光照图像  $S$  和上一次迭代取得的增强图像  $Y_{n-1}$  加入参考图像序列。将包含  $(2N+2)$  张图像的参考图像序列作为伪标

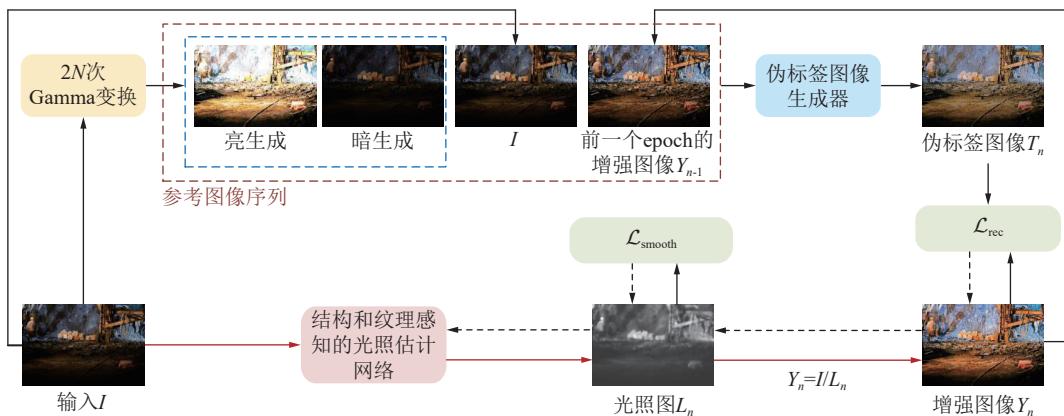


图1 井下低光照自监督图像增强网络结构

Fig.1 Underground low-light self-supervised image enhancement network structure

签生成器的输入, 经过计算获得当前迭代训练的伪标签图像  $T_n$ 。利用重建损失约束伪标签图像  $T_n$  与增强图像  $Y_n$  在亮度和对比度上的一致性, 从而间接约束结构和纹理感知的光照估计网络, 产生具有良好亮度和对比度的光照图。在模型测试和推理阶段, 仅保留红色箭头部分, 则结构和纹理感知的光照估计网络可以根据给定的输入生成具有分段平滑以及良好亮度和对比度的光照图。

### 1.1 伪标签图像

受图像融合算法<sup>[23]</sup>启发, 不同曝光程度的图像可以按照每像素亮度、局部对比度和每像素的色彩饱和度计算图像融合权重, 根据融合权重选择参考图像序列中的像素, 从而得到具有良好亮度、对比度和饱和度的图像。为此, 采用非线性的 Gamma 变换来改变低光照图像的亮度和对比度。

$$G = I^\gamma \quad (1)$$

式中:  $G$  为 Gamma 变换后的图像;  $I$  为输入的低光照图像;  $\gamma$  为调整参数。当  $\gamma < 1$  时, 拉伸低光照图像中灰度级较低的像素, 使图像整体亮度提升, 对比度降低。当  $\gamma > 1$  时, 拉伸低光照图像中灰度级较高的像素, 使图像整体亮度降低, 对比度提升。

Gamma 变换是对每个像素应用非线性变换, 提升全局亮度, 在一定程度上扩大低光照图像的亮度范围, 但忽略了像素之间的关系, 且参数设置不当, 容易造成局部区域的过曝光, 需要根据具体情况调整 Gamma 值。当 Gamma 取值在  $(0, 1)$  时, 增强输入图像中曝光不足区域亮度, 提升暗部细节; 当 Gamma 取值在  $[1, 25]$  时, 降低输入图像中过曝光区域亮度, 提过高曝光区域的可视性。此外, 在选择区间时参考了文献<sup>[24]</sup>。因此, 分别在  $(0, 1)$  和区间  $[1, 25]$  内随机生成  $N$  个 Gamma 值, 对输入的低光照图像进行  $N$  次 Gamma 变换, 再将生成的  $2N$  张 Gamma 变换的图像

进行图像融合, 以生成伪标签图像。此外, 将未经变换的低光照图像和训练过程中产生的增强图像添加至用于进行图像融合的参考图像序列, 将参考图像序列表示为  $[G_0, \dots, G_{N-1}, S, Y_{n-1}]$ , 其中,  $(n-1)$  表示模型训练中前一次产生的增强图像。由于  $n=0$  次训练过程中未产生增强图像, 因此  $n=0$  次训练中用于计算伪标签的参考图像序列长度为  $2N+1$ , 其他训练过程中的参考图像序列长度均为  $2N+2$ 。

采用图像融合策略生成伪标签时, 用于计算融合权重的指标包含每像素亮度度量  $E$ 、局部对比度度量  $C$  和每像素色彩饱和度度量  $S$ 。每像素的亮度度量定义为

$$E(i) = |\mu_{h \times h} - 0.5| \quad (2)$$

式中:  $i$  为图像中的独立像素,  $px$ ;  $\mu_{h \times h}$  为图像中  $h \times h$  区域内的所有像素的平均强度; 参数 0.5 表示较好的亮度级别。局部对比度用图像中  $h \times h$  区域内所有像素的局部方差表示, 计算公式如下:

$$C(i) = \frac{1}{h \times h} \sum_p [I(p) - \mu_{h \times h}]^2 \quad (3)$$

式中:  $I(p)$  为图像中  $h \times h$  区域内的所有像素。每像素色彩饱和度用 HSV 颜色空间中的饱和通道来度量图像的色彩饱和度, 使用 RGB 颜色空间定义饱和度, 表示为

$$S(i) = \frac{\max(R, G, B) - \min(R, G, B)}{\max(R, G, B)}, \quad (4)$$

式中:  $S(i)$  为 HSV 颜色空间中的色彩饱和度;  $R$ 、 $G$  和  $B$  为像素  $i$  在 RGB 颜色空间上 3 个颜色通道的像素值。

通过计算参考图像序列  $[G_0, \dots, G_{N-1}, S, Y_{n-1}]$  中每张图像  $k$  的像素亮度、局部对比度和像素饱和度, 按照式(5)生成伪标签图像:

$$T_n(i) = \operatorname{argmax}_{k=1,2,\dots,2N+2} \frac{C_k(i) - S_k(i)}{E_k(i)} \quad (5)$$

## 1.2 自监督训练损失

采用结构感知的平滑损失约束光照估计网络产生分段平滑的光照图。此外,采用重建损失约束增强图像与伪标签图像的一致性,从而间接约束光照图估计网络以产生具有合理对比度和亮度的光照图。结构感知的平滑损失表示为

$$\mathcal{L}_{\text{smooth}} = \|L_n^{\text{init}} - L_n\|_2 + \sum_{i=1}^N \sum_{j \in Q_{h(i)}} w_{i,j} \|L_n^i - L_n^j\|_2 \quad (6)$$

式中:  $L_n^{\text{init}}$  为构建的初始光照,即选择低光照图像  $S$  的 YCbCr 颜色空间中  $Y$  通道作为初始光照;  $L_n$  为估计的光照图;  $Q_{h(i)}$  为以像素  $i$  为中心,  $h \times h$  区域内不包含像素  $i$  的所有像素;  $w_{i,j}$  为结构引导权重。 $w_{i,j}$  的计算公式如下:

$$w_{i,j} = \exp \left( -\frac{\|L_n^i - L_n^j\|_2}{2\sigma^2} \right) \quad (7)$$

式中:  $\sigma$  为标准差,试验中设置为 1。重建损失使用  $L_1$  范数计算增强图像与伪标签图像之间的距离,使增强图像与伪标签图像的像素值之间的绝对差值最小化,可以有效保留图像细节。此外,使用 SSIM 损失函数衡量亮度和对比度方面的一致性,使增强图像与伪标签图像的结构相似性最大化,可以有效保留图像的亮度和对比度。重建损失表示为

$$\mathcal{L}_{\text{rec}} = (1 - \alpha) \|Y_n - T_n\| + \alpha \mathcal{L}_{\text{SSIM}}(Y_n, T_n) \quad (8)$$

式中:  $\alpha$  为平衡参数,用于控制平滑损失和原始损失之间的权重,通常为一个  $[0, 1]$  的正数。依据文献[25]提供的较为详细的试验数据且参考了大部分关于图像增强的文献中关于平衡参数的设置,将平衡参数  $\alpha$  设置为 0.84。 $\mathcal{L}_{\text{rec}}$  为重建损失, $\mathcal{L}_{\text{SSIM}}$  为 SSIM 损失,表示为

$$\mathcal{L}_{\text{SSIM}} = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (9)$$

式中: $x$  和  $y$  分别为增强图像  $Y_n$  与伪标签图像  $T_n$  的窗口数据,  $\mu_x$  和  $\mu_y$  分别为  $x$  和  $y$  的均值;  $\sigma_x^2$  和  $\sigma_y^2$  分别为  $x$  和  $y$  的方差;  $\sigma_{xy}$  为  $x$  和  $y$  的协方差;  $c_1$  和  $c_2$  为常数。

最终,用于模型训练的总损失可以表示为

$$\mathcal{L} = \mathcal{L}_{\text{smooth}} + \mathcal{L}_{\text{rec}} \quad (10)$$

## 1.3 结构和纹理感知的光照估计网络

Retinex 理论指出,给定一张输入图像  $S$ ,可以分解为反射率  $R$  和照度  $L$ 。反射率  $R$  描述了场景物体的内在颜色属性;照度  $L$  描述了场景照明的光强属性,

决定了图像的对比度和亮度。该分解过程表示为

$$S = R \odot L \quad (11)$$

式中:  $\odot$  表示逐元素相乘。Retinex 的解耦具有非适定性,为了缩小解空间,可以将式(1)变换为  $R=S/L$ 。此时仅需要构建光照估计网络以产生合理的光照图  $L_n$ ,将  $R$  作为最终的增强图像。

在理想情况下,照度分量是连续且光滑的,即在图像的每个像素处,照度值都应与相邻像素的照度值相近。然而,在实际情况中,受人造光源以及物体表面的边缘和纹理影响,低光照图像会形成明显的阴影和高光区域以及出现细小的亮度变化。这些变化使得照度分量在局部范围出现不连续现象,因此照度分量在实际中表现为分段光滑的特性。

为了从大规模低光照图像中学习鲁棒的内在特征规律,并为给定的低光照图像匹配合理的光照图,构建了一个结构和纹理感知的光照估计网络,其结构如图 2 所示。该网络基于编码器-解码器结构进行设计。在编码器部分采用了 4 层卷积结构,其中 2 层为步长为 2 的  $3 \times 3$  卷积,主要对特征图进行采样,缩小特征图的空间分辨率。在解码器部分采用了 2 个反卷积层来提高特征图的空间分辨率,以恢复至与输入图像相同的分辨率。为了深入挖掘低光照图像中的局部纹理特征和全局结构特征,在编码器和解码器中间,引入了局部-全局感知模块。

## 1.4 局部-全局感知模块

边缘结构主要指受井下矿灯等不均匀的人造光源影响在图像中反映为场景中明暗交界的边缘。纹理细节主要指场景中物体表面存在的粗糙特性,受光源影响时产生的粗糙纹理。受井下复杂光照环境影响,统一的特征提取方法对结构和纹理感知不明显,为了深入挖掘低光照图像中的局部纹理特征和全局结构特征以提高光照估计网络的性能,在光照估计网络中引入了局部-全局感知模块。

局部-全局感知模块结构如图 2 所示。该模块首先通过 Split 操作将输入特征按通道维度平均划分为 2 部分:一部分利用局部特征提取分支感知场景中的纹理细节特征,另一部分利用全局特征提取分支感知边缘结构特征。将 2 部分特征通过 Concat 操作进行特征拼接,并采用 Fusion 模块中的通道注意力机制<sup>[26]</sup>控制特征图之间的通道注意力权重,以融合局部特征和全局特征。

### 1.4.1 局部特征提取分支

传统卷积中卷积核的权重是固定的,通常需要多层卷积的堆叠才能很好地适应图像内容。为了充分

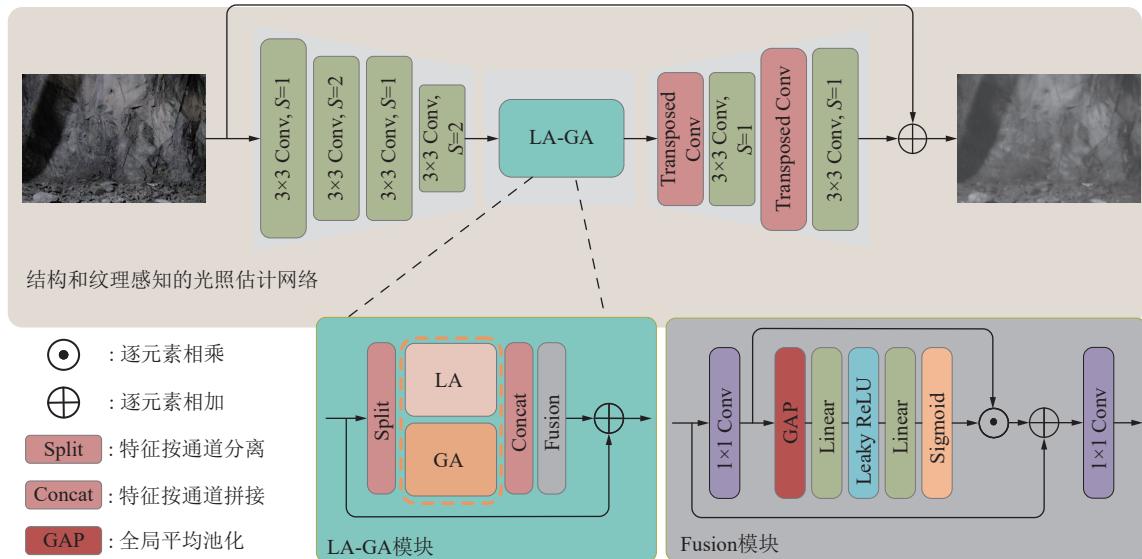


图2 结构和纹理感知的光照估计网络结构

Fig.2 Structure and texture-aware illumination estimation network architecture

挖掘低光照图像中的局部纹理特征,在局部-全局感知模块中引入了局部特征提取分支。如图3a所示为局部特征提取分支的结构示意图,在设计局部特征提取分支时,参考自注意力机制范式,通过线性变换和深度可分离卷积得到 $Q$ 、 $K$ 和 $V$ ,其中 $Q$ 和 $K$ 用于产生上下文感知权重。具体表示为:

$$W = \text{Tanh}\left(\frac{FC(\text{Swish}(FC(Q \odot K)))}{\sqrt{d}}\right) \quad (12)$$

式中: $FC(\cdot)$ 为线性变换层; $d$ 为特征的通道维度; $\odot$ 为逐元素相乘的 Hadamard 积;Swish 和 Tanh 为激活函数,Tanh 主要将生成的上下文感知权重的值域映射至 $[0, 1]$ 。

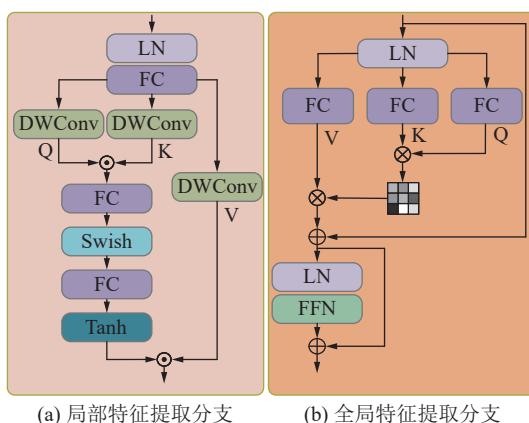


图3 局部-全局感知模块结构

Fig.3 Local-Global aware block structure

不同于自注意力机制中的矩阵乘法,局部特征提取分支通过具有非线性能力的门控机制生成上下文感知权重,以空间变化的方式关注并提取 $V$ 中重要的局部特征信息。该过程可以表示为

#### 1.4.2 全局特征提取分支

视觉 Transformer 的自注意力机制具有较强的全局信息交互能力,使得其全局特征提取能力远高于其他现有方法。然而,光照估计任务需要处理高分辨率图像,视觉 Transformer 的自注意力机制的计算复杂度与输入图像的分辨率呈二次方增长,易导致显存资源的严重消耗。为此,采用一种线性注意力机制来提取全局特征,以降低计算复杂度,其结构如图3b所示。

相比于视觉 Transformer 的自注意力机制,该模块改变了 $Q$ 、 $K$ 和 $V$ 进行自注意力计算时矩阵乘法的顺序,如下表示:

$$F_o = F + V \cdot \text{Softmax}\left(\frac{K^T \cdot Q}{\sqrt{d}}\right) \quad (14)$$

$$X_{\text{gloable}} = F_o + \text{FFN}(F_o) \quad (15)$$

式中: $F \in \mathbb{R}^{[HW, C]}$ 为给定的输入特征; $\text{FFN}(\cdot)$ 为全连接层; $d$ 为特征的通道维度; $Q$ 、 $K$ 和 $V$ 的形状为 $[HW, C]$ , $\text{Softmax}$ 得到注意力权重的形状为 $[C, C]$ ,该线性注意力的计算复杂度与通道维度相关。特征图的通道维度远小于特征图分辨率的二次方,因此该模块的计算复杂度远低于原始视觉 Transformer 的计算复杂度。

## 2 试验结果与分析

### 2.1 试验配置

#### 2.1.1 试验参数设置

在 NVIDIA 3090 GPU 和 Intel Core i9-10900K

CPU 上使用 PyTorch 1.8.2 深度学习框架实现结构和纹理感知的光照估计网络搭建。在训练过程中, 使用 Adam 优化器对模型进行优化, 参数  $\beta_1 = 0.9$ ,  $\beta_2 = 0.999$ ,  $\varepsilon = 10^{-8}$ , 学习率初始化设置为  $5e^{-4}$ , 批量大小设置为 8, 训练次数为 80, 将输入图像大小调整为  $512 \times 512$ 。

### 2.1.2 数据集

为探索该方法对井下低光照图像的适用性和优势, 视频数据取自安徽省淮南市谢桥矿煤矿数据中心, 主要是使用本安型摄像机进行录制。按照间隔 23 帧

的方式对视频进行等间隔采样, 得到了 997 张大小为  $1920 \times 1280$  的低光照图像, 制作了井下真实场景的低光照图像数据集 CUMTB-IE, 部分样本图像如图 4 所示。此外, 为验证方法的普适性, 在公开数据集 LOL<sup>[13]</sup> 和 4 个非参考数据集 LIME、NPE<sup>[27]</sup>、MEF<sup>[28]</sup>、和 DICM<sup>[29]</sup> 上分别进行了性能评估。LOL 数据集是一个常用的低光照图像数据集, 包含 500 对多种场景中不同低光条件下的真实图像, 分辨率为  $400 \times 600$ , 由于该方法为自监督学习, 不受配对和非配对数据集约束, 因此, 随机选取 485 张低光照图像作为训练集。

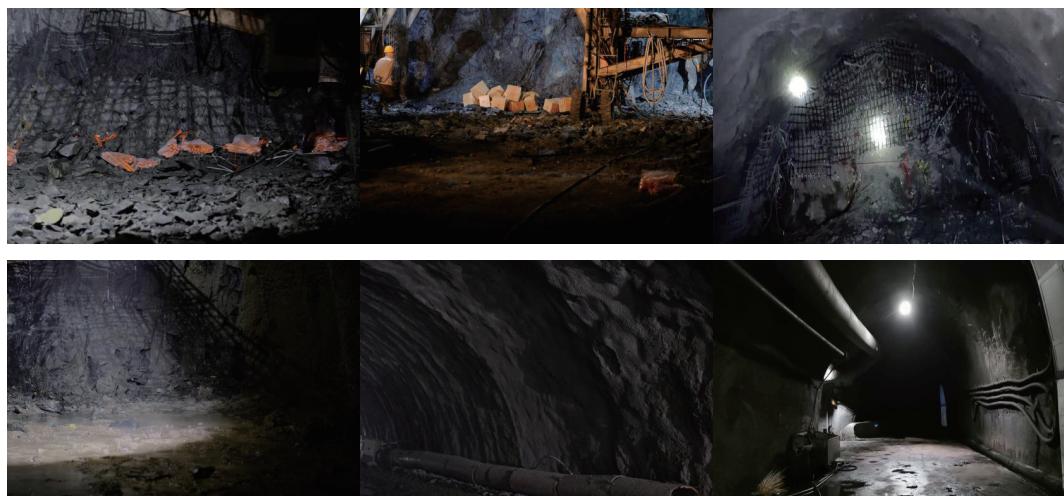


图 4 CUMTB-IE 数据集中部分样本

Fig.4 Partial sample plot in CUMTB-IE dataset

### 2.1.3 评估指标

对带有标签的 LOL 基准数据集, 采用峰值信噪比 PSNR 和结构相似性 SSIM 这 2 个指标进行客观评价。其中, PSNR 和 SSIM 值越高, 表示图像的质量越好。对其他 4 个不具有标签的非参考基准数据集, 采用自然度评价指标 NIQE 和无参考图像空间质量评价指标 BRISQUE 来进行客观评价。其中, NIQE 和 BRISQUE 的值越低, 表示图像的质量越好。

## 2.2 试验结果与分析

与其他 6 种低光照图像增强方法进行了比较, 分别包括基于监督学习的低光照图像增强方法 RetinexNet、KinD 和 DeepUPE 以及基于自监督学习的低光照图像增强方法 Zero-DCE、RetinexDIP 和 SCI。不同方法的定量评估结果见表 1。由于一些方法提供的官方代码使用的 CUDA 版本太旧, 无法在 NVIDIA 3090 上运行, 因此, 将在 NVIDIA V100 上运行这些

表 1 不同数据集上的低光照图像增强方法的定量比较

Table 1 Quantitative comparison of low-light image enhancement methods on different datasets

方法	LOL		LIME		NPE		MEF		DICM	
	PSNR↑	SSIM↑	NIQE↓	BRI↓	NIQE↓	BRI↓	NIQE↓	BRI↓	NIQE↓	BRI↓
RetinexNet <sup>[13]</sup>	16.770	0.461	4.597	25.57	4.567	24.81	4.389	23.74	4.500	30.82
KinD <sup>[15]</sup>	17.640	<b>0.778</b>	4.763	26.77	3.529	24.25	3.846	30.43	3.565	30.49
DeepUPE <sup>[17]</sup>	11.680	0.502	3.959	23.43	3.994	28.25	3.527	22.55	3.884	26.31
Zero-DCE <sup>[20]</sup>	14.830	0.531	<b>3.789</b>	23.33	3.590	29.59	<b>3.308</b>	25.48	3.624	28.11
RetinexDIP <sup>[19]</sup>	9.442	0.322	3.815	<b>21.28</b>	3.604	25.06	3.657	22.35	3.424	25.32
SCI <sup>[22]</sup>	14.780	0.525	4.138	22.83	4.162	34.89	3.433	26.48	3.609	31.87
提出方法	<b>18.310</b>	0.598	4.055	22.44	<b>3.468</b>	<b>23.37</b>	3.398	<b>21.56</b>	<b>3.401</b>	<b>25.19</b>

注: ↑表示数值越大越好; ↓表示数值越小越好; 最优的结果进行了加粗表示。

代码并产生结果。

分析表1中的数据可知,基于监督学习的低光照图像增强方法受数据集差距的影响,在不同数据集上显示出较大的性能差异,表明基于监督学习的低光照图像增强方法严重依赖配对的低光照/正常光照图像数据集,在真实场景中进行低光照图像增强的泛化能力较差。例如,KinD方法在LOL数据集上进行训练,因此在LOL数据集上取得了较好的PSNR和SSIM,但在面对其他4个真实场景的图像时,则表现出一般

的低光照图像增强性能。提出的方法在5个基准数据集上表现出较为均衡的低光照图像增强性能,表明该方法在不依赖标签图像监督的情况下,产生了较强的泛化性能。这对于解决井下难以制作配对的低光照/正常光照数据集来说是非常重要的。

为验证所提方法面对井下低光照图像的有效性和适用性,在井下真实场景的低光照图像数据集CUMTB-IE上进行了泛化性试验,低光照图像增强效果对比如图5所示。

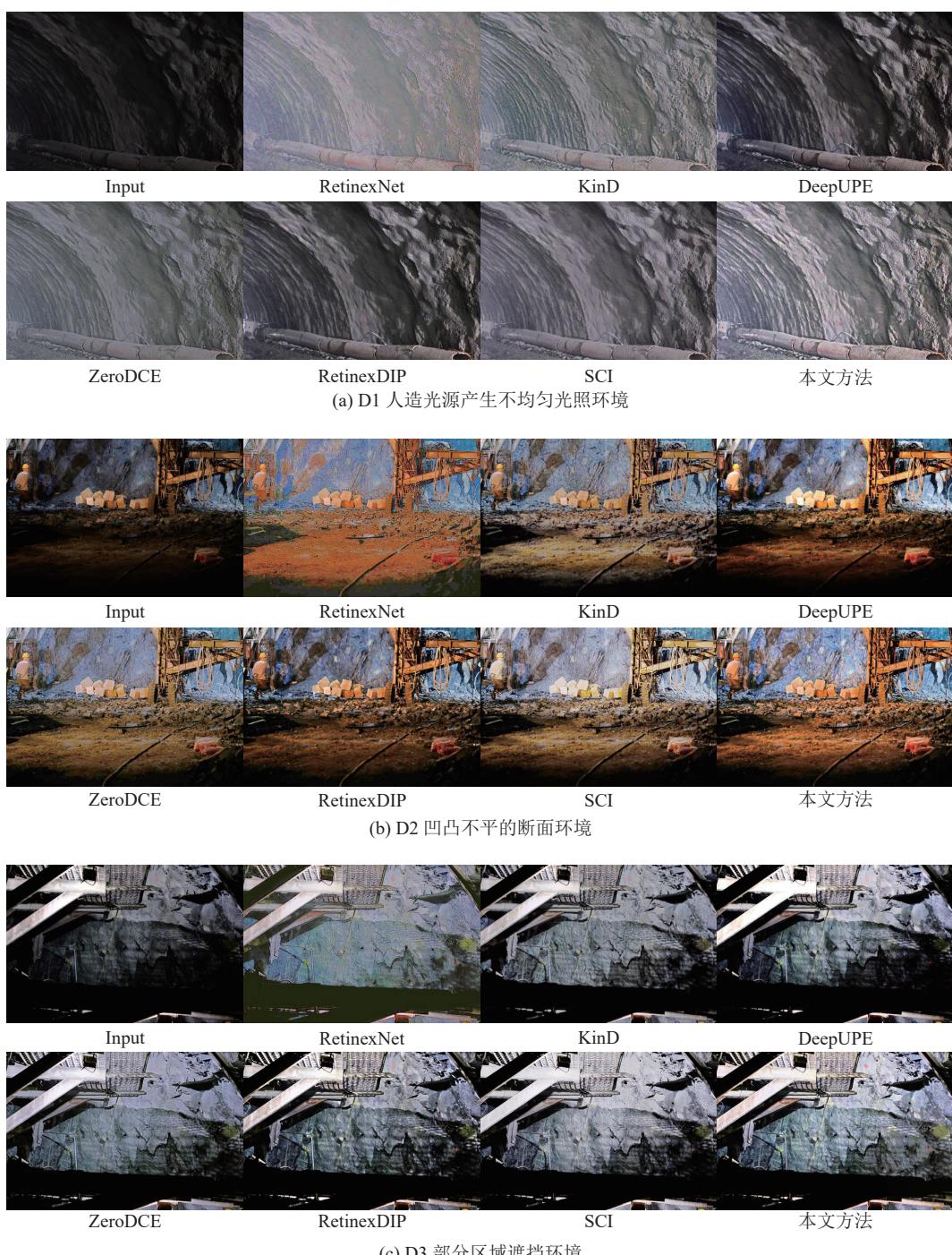


图5 数据集CUMTB-IE上测试图像增强结果对比

Fig.5 Comparison of test image enhancement results on the CUMTB-IE dataset

通过视觉分析图 5 内 3 个不同矿井场景下的低光图像增强效果可以看出,由于光照平滑度约束的限制,RetinexNet 方法在高对比度和局部边缘处产生了过冲伪影,增强结果与原图风格不符,光照不够真实。KinD、ZeroDCE 和 SCI 方法的增强结果改变了原有图像的色彩饱和度。虽然图像的整体亮度和对比度得到提升,但面对局部暗区时显示出亮度和对比度提升不足的现象。DeepUPE 方法的增强结果对比度较高,色彩饱和度较高,但局部区域的亮度提升不足。RetinexDIP 方法面对不同矿井场景的图像时产生了较好的增强效果,但在试验过程中,需要对给定的输入图像进行一定次数的迭代,无法应用于实时性的低光照图像增强,限制了其应用范围。得益于伪标签的

生成方式以及光照估计网络生成分段光滑的光照图的能力,所提方法面对图 5 内等 3 个不同矿井场景的低光照图像,在整体亮度和局部亮度以及对比度和色彩饱和度方面均产生了相对较好的增强效果,图像内容也得到了有效恢复,没有产生较为明显的伪影以及色彩失真的问题。

表 2 给出不同方法在数据集 CUMTB-IE 上的定量评估结果。由于该数据集没有配对的标签图像,无法实现 PSNR 和 SSIM 指标的计算,因此只进行了无参考评价指标 NIQE 和 BRISQUE 的计算。在 CUMTB-IE 矿井场景低光图像数据集上,相比于其他几种方法,所提方法取得了相对较好的 NIQE 和 BRISQUE。

表 2 CUMTB-IE 数据集上增强方法的定量比较

Table 2 Quantitative comparison of augmentation methods on CUMTB-IE dataset

方法	RetinexNet	KinD	DeepUPE	ZeroDCE	RetinexDIP	SCI	所提方法
NIQE↓	4.742	3.811	4.546	3.928	4.155	4.338	<b>3.752</b>
BRISQUE↓	24.854	24.232	26.479	27.583	23.548	25.378	<b>22.886</b>

由表 2 数据可知,所提方法在 NIQE 和 BRISQUE 指标上均达到了最好。该方法的 NIQE 指标比 RetinexNet、KinD、DeepUPE、ZeroDCE、RetinexDIP 和 SCI 分别降低了 20.88%、1.55%、17.47%、4.48%、9.70% 和 13.51%。BRISQUE 指标比 RetinexNet、KinD、DeepUPE、ZeroDCE、RetinexDIP 和 SCI 分别降低了 7.91%、5.55%、13.57%、17.03%、2.81% 和 9.82%。综合分析,所提方法在视觉效果和评价指标分析上均有显著的优势,因此,更适用于井下低光照场景的图像增强。

### 2.3 消融试验

为验证所提模块对最终生成的增强图像的有效性,通过控制模块变量的方式进行了相关的消融试验。所选测试图像取自数据集 CUMTB-IE,比较了局部特征提取分支、全局特征提取分支以及局部-全局感知模块 3 种情况下的参数量、NIQE 和 BRISQUE 指标。基线模型表示未引入局部-全局感知模块的光照估计网络,试验中仅改变了深层特征提取结构,其他结构保持不变,消融试验结果见表 3。

从表 3 的消融试验结果可以看出,在基线模型上,通过使用局部特征提取分支,NIQE 指标降低了 38.12%,BRISQUE 指标降低了 6.93%,表明局部特征提取分支能够感知场景中的纹理细节特征。通过使用全局特征提取分支,NIQE 指标降低了 41.79%,

BRISQUE 指标降低了 13.23%,表明全局特征提取分支能够感知边缘结构特征。使用局部-全局感知模块,通过具有非线性能力的门控机制生成上下文感知权重,以空间变化的方式关注并提取重要的局部特征信息,同时采用一种线性注意力机制来提取全局特征,NIQE 指标降低了 45.22%,BRISQUE 指标降低了 17.04%,局部-全局感知模块在数据集 CUMTB-IE 上表现出最好的性能。

表 3 消融试验结果

Table 3 Results of ablation test

方法	参数量/M	NIQE↓	BRISQUE↓
基线	—	6.850	27.586
局部特征提取分支	0.087	4.239	25.675
全局特征提取分支	0.060	3.987	23.937
局部-全局模块	0.076	3.752	22.886

为了验证所使用的重建损失  $\mathcal{L}_{\text{rec}}$  和结构感知的平滑损失  $\mathcal{L}_{\text{smooth}}$  在方法上的有效性,通过控制损失函数变量的方式进行相关的消融试验,所选测试图像取自数据集 CUMTB-IE。其中仅使用结构感知的平滑损失  $\mathcal{L}_{\text{smooth}}$  对光照估计网络进行训练,不包含伪标签生成过程;使用重建损失  $\mathcal{L}_{\text{rec}}$  和结构感知的平滑损失  $\mathcal{L}_{\text{smooth}}$  的组合对光照估计网络进行训练,包含了伪标签的生成过程。消融试验增强效果如图 6 所示。

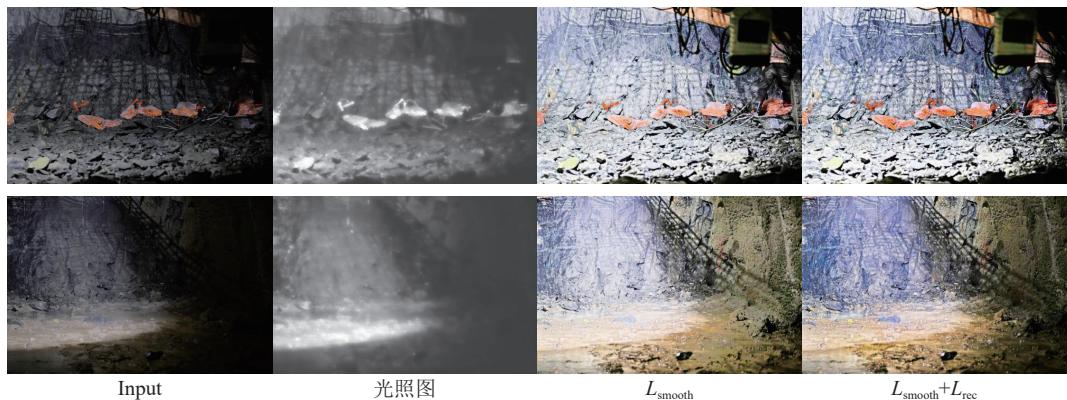


图 6 消融试验结果  
Fig.6 Results of ablation test

由图 6 的消融试验结果可以看出,只使用结构感知的平滑损失进行训练产生的增强图像,虽然能够约束光照估计网络产生分段光滑的光照图,使得暗光区域可以获得较好的对比度,但亮光区域会产生过曝现象。重建损失可以约束增强图像与伪标签图像之间亮度和对比度之间的一致性。重建损失提升了暗光区域的对比度和亮度,且避免了亮光区域的过曝现象产生。

#### 2.4 方法耗时性分析

表 4 为提出的方法与其他几种方法在数据集 CUMTB-IE 下增强单幅图像所需要的时间比较。

表 4 方法计算时间比较  
Table 4 Method calculation time comparison

模型	计算时间/ms
RetinexNet	183.71
KinD	168.32
DeepUPE	113.36
ZeroDCE	88.31
RetinexDIP	—
SCI	76.90
提出方法	94.17

通过表 4 的试验结果可知,在相同的试验数据集下,该方法的计算时间是 94.17 ms,与较先进的算法 ZeroDCE 和 SCI 基本持平。

### 3 结 论

1) 设计了一种光照估计网络,通过结构感知的平滑损失约束光照估计网络,保留了场景的边缘结构并平滑纹理细节,以产生分段平滑的光照图。引入了伪标签图像生成器来合成具有良好对比度和亮度的伪标签图像,并利用重建损失约束增强图像与伪标签图

像之间亮度和对比度之间的一致性,从而间接约束光照估计网络以产生具有合理对比度和亮度的光照图,有效改善了井下低光照图像的视觉效果。在井下真实场景的低光照图像数据集 CUMTB-IE 上,NIQE 和 BRISQUE 指标分别达到了 3.752 和 22.886。

2) 为深入挖掘低光照图像中的局部纹理特征和全局结构特征以提高光照估计网络性能,在光照估计网络中引入局部-全局感知模块。试验表明 NIQE 指标降低了 45.22%, BRISQUE 指标降低了 17.04%,有效保留了场景的边缘结构并平滑纹理细节。

#### 参考文献(References):

- [1] 王国法, 张良, 李首滨, 等. 煤矿无人化智能开采系统理论与技术研发进展[J]. 煤炭学报, 2023, 48(1): 34–53.  
WANG Guofa, ZHANG Liang, LI Shoubin, et al. Progresses in theory and technological development of unmanned smart mining system[J]. Journal of China Coal Society, 2023, 48(1): 34–53.
- [2] 孙林, 陈圣, 姚旭龙, 等. 矿井智能监控目标识别的图像增强方法与应用[J]. 煤炭学报, 2024, 49: 495–504.  
SUN Lin, CHEN Sheng, YAO Xulong, et al. Image enhancement methods and applications for target recognition in intelligent mine monitoring[J]. Journal of China Coal Society, 2024, 49: 495–504.
- [3] 智宁, 毛善君, 李梅, 等. 基于深度融合网络的煤矿图像尘雾清晰化算法[J]. 煤炭学报, 2019, 44(2): 655–666.  
ZHI Ning, MAO Shanjun, LI Mei, et al. Coal mine image dust and fog clearing algorithm based on deep fusion network[J]. Journal of China Coal Society, 2019, 44(2): 655–666.
- [4] 范伟强, 刘毅. 基于自适应小波变换的煤矿降质图像模糊增强算法[J]. 煤炭学报, 2020, 45(12): 4248–4260.  
FAN Weiqiang, LIU Yi. Fuzzy enhancement algorithm of coal mine degradation image based on adaptive wavelet transform[J]. Journal of China Coal Society, 2020, 45(12): 4248–4260.
- [5] PILZER S M, AMBURN E P, AUSTIN J D, et al. Adaptive histogram equalization and its variations[J]. Computer Vision, Graphics, and Image Processing, 1987, 39(3): 355–368.
- [6] KIM J Y, KIM L S, HWANG S H. An advanced contrast enhance-

- ment using partially overlapped sub-block histogram equalization[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2001, 11(4): 475–484.
- [7] PARK G H, CHO H H, CHOI M R. A contrast enhancement method using dynamic range separate histogram equalization[J]. *IEEE Transactions on Consumer Electronics*, 2008, 54(4): 1981–1987.
- [8] 龚云, 颜昕宇. 基于同态滤波方法的煤矿井下图像增强技术研究[J]. 煤炭科学技术, 2023, 51(3): 241–250.  
GONG Yun, XIE Xinyu. Research on coal mine underground image recognition technology based on homomorphic filtering method[J]. Coal Science and Technology, 2023, 51(3): 241–250.
- [9] 张立亚, 郝博南, 孟庆勇, 等. 基于 HSV 空间改进融合 Retinex 算法的井下图像增强方法[J]. 煤炭学报, 2020, 45(S1): 532–540.  
ZHANG Liya, HAO Bonan, MENG Qingyong, et al. Underground image enhancement method based on HSV space improved fusion Retinex algorithm[J]. Journal of China Coal Society, 2020, 45(S1): 532–540.
- [10] JOBSON D J, RAHMAN Z, WOODELL G A. Properties and performance of a center/surround retinex[J]. *IEEE Transactions on Image Processing*, 1997, 6(3): 451–462.
- [11] JOBSON D J, RAHMAN Z, WOODELL G A. A multiscale retinex for bridging the gap between color images and the human observation of scenes[J]. *IEEE Transactions on Image Processing*, 1997, 6(7): 965–976.
- [12] 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.
- [13] WEI C, WANG W J, YANG W H, et al. Deep retinex decomposition for low-light enhancement[EB/OL]. 2018: 1808.04560. <https://arxiv.org/abs/1808.04560v1>.
- [14] YANG W H, WANG W J, HUANG H F, et al. Sparse gradient regularized deep retinex network for robust low-light image enhancement[J]. *IEEE Transactions on Image Processing*, 2021, 30: 2072–2086.
- [15] 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. New York: ACM, 2019: 1632–1640.
- [16] ZHANG Y H, GUO X J, MA J Y, et al. Beyond brightening low-light images[J]. *International Journal of Computer Vision*, 2021, 129(4): 1013–1037.
- [17] WANG R X, ZHANG Q, FU C W, et al. Underexposed photo enhancement using deep illumination estimation[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2019: 6849–6857.
- [18] PAN X G, ZHAN X H, DAI B, et al. Exploiting deep generative prior or for versatile image restoration and manipulation[C]// Computer Vision – ECCV 2020. Cham: Springer International Publishing, 2020: 262–277.
- [19] ZHAO Z J, XIONG B S, WANG L, et al. RetinexDIP: A unified deep framework for low-light image enhancement[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2021, 32(3): 1076–1088.
- [20] 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.
- [21] ZHANG Z Y, JIANG Y T, JIANG J, et al. STAR: A structure-aware lightweight transformer for real-time image enhancement[C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway, NJ: IEEE, 2021: 4086–4095.
- [22] 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). Piscataway, NJ: IEEE, 2022: 5627–5636.
- [23] MERTENS T, KAUTZ J, VAN REETH F. Exposure fusion[C]//15th Pacific Conference on Computer Graphics and Applications (PG'07). Piscataway, NJ: IEEE, 2007: 382–390.
- [24] 刘志成, 王殿伟, 刘颖, 等. 基于二维伽马函数的光照不均匀图像自适应校正算法[J]. 北京理工大学学报, 2016, 36(2): 191–196, 214.  
LIU Zhicheng, WANG Dianwei, LIU Ying, et al. Adaptive adjustment algorithm for non-uniform illumination images based on 2D gamma function[J]. Transactions of Beijing Institute of Technology, 2016, 36(2): 191–196, 214.
- [25] 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.
- [26] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2018: 7132–7141.
- [27] 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.
- [28] MA K D, ZENG K, WANG Z. Perceptual quality assessment for multi-exposure image fusion[J]. *IEEE Transactions on Image Processing*, 2015, 24(11): 3345–3356.
- [29] LEE C, LEE C, KIM C S. Contrast enhancement based on layered difference representation of 2D histograms[J]. *IEEE Transactions on Image Processing*, 2013, 22(12): 5372–5384.