E-mail: jig@aircas.ac.cn
Website: www.cjig.cn
Tel: 010-58887035
中国图象图形学报
JOURNAL OF IMAGE AND GRAPHICS
©中国图象图形学报版权所有

中图法分类号:TP391;TP18　　**文献标识码**: A　　**文章编号**: 1006-8961(2026)09-3336-16

论文引用格式:Bao X A, Chen Y J, Zhang N, Hu T B, Xu M Y and Tu X M. 2026. Zero-supervised low-light image enhancement with illumination-guided reflectance estimation. Journal of Image and Graphics, 31(9):3336-3351(包晓安, 陈奕江, 张娜, 胡天缤, 许铭洋, 涂小妹. 2026. 光照引导反射分量估计的零监督弱光图像增强. 中国图象图形学报, 31(9):3336-3351)[DOI:10.11834/jig.250486]

光照引导反射分量估计的零监督弱光图像增强

包晓安[1],陈奕江[1],张娜[1],胡天缤[2],许铭洋[1],涂小妹[3*]

1. 浙江理工大学计算机科学与技术学院,杭州　310018; 2. 河海大学人工智能与自动化学院,常州　213000;
3. 浙江广厦建设职业技术大学建筑工程学院,东阳　322100

摘　要:**目的**　弱光环境下成像质量下降严重制约视觉监控、自动驾驶等高级视觉任务的性能。现有方法虽然可提升亮度,但普遍存在过曝抑制不足且依赖成对正常光照数据的问题,限制了其在真实场景中的应用。**方法**　针对这些局限,提出一种基于光照引导反射分量估计的零监督弱光图像增强方法。首先,通过可逆亮度扰动合成训练样本,设计伪曝光生成策略,在不改变图像固有反射率的情况下,生成多种光照样本以摆脱成对监督数据的依赖;其次,设计双域协同注意力去噪模块以抑制背景噪声;然后,利用多层残差卷积与结构对称的U-Net架构协同生成照明图,并利用其深层光照信息对反射分量进行显式引导;最后,通过引导式反射估计实现结构纹理的高保真重建与过曝抑制。**结果**　在LOL、DICM和WildNight-8K数据集上与9种代表性的最新方法进行对比分析。在LOL数据集上,相比Retinexformer和LightenDiffusion,本文方法的自然图像质量评估指标(natural image quality evaluator,NIQE)分别降低0.62和0.92,感知指数(perceptual index,PI)分别降低1.23和0.25;在DICM数据集上,相比Retinexformer和RUAS(retinex-inspired unrolling with architecture search),本文方法的NIQE分别降低0.11和0.81,PI分别降低0.48和0.72;在WildNight-8K数据集上,相比Retinexformer和EnlightenGAN,本文方法的NIQE分别降低1.32和1.87,PI分别降低0.33和5.98。在YOLOv8目标检测任务中,平均检测精度(mean average precision,mAP@50-95)较原图提升4.60%。**结论**　本文提出的基于光照引导反射分量估计的零监督弱光图像增强方法在多个数据集上取得了领先的客观指标,并能够有效提升下游目标检测任务性能,验证了其在真实弱光应用场景中的价值。

关键词:弱光图像增强;注意力机制;零监督;目标检测;卷积神经网络

Zero-supervised low-light image enhancement with illumination-guided reflectance estimation

Bao Xiaoan[1], Chen Yijiang[1], Zhang Na[1], Hu Tianbin[2], Xu Mingyang[1], Tu Xiaomei[3*]

1. School of Computer Science and Technology, Zhejiang Sci-Tech University, Hangzhou　310018, China; 2. School of Artificial Intelligence and Automation, Hohai University, Changzhou　213000, China; 3. School of Civil Engineering and Architecture, Zhejiang Guangsha Vocational and Technical University of Construction, Dongyang　322100, China

Abstract: **Objective**　Low-light environments in practical computer vision applications, such as security surveillance, autonomous driving, and intelligent urban monitoring, considerably degrade image quality and thus restrict the performance of high-level visual tasks. Images captured under poor illumination often exhibit insufficient brightness, heavy noise

收稿日期:2025-10-15;**修回日期**:2026-01-28;**预印本日期**:2026-02-04

*通信作者:涂小妹　txm_95@163.com

基金项目:浙江省重点研发计划资助(2020C03094);金华市科技局重点课题项目(2024C22592)

Supported by:Zhejiang Provincial Key R&D Program, China(2020C03094);Jinhua Municipal Science and Technology Bureau Key Research Project(2024C22592)

amplification，and severe loss of structural and textural information. These degradations not only harm perceptual quality but also compromise the reliability of downstream applications such as object tracking，object detection，and semantic segmentation. Existing methods have attempted to solve these problems by increasing image brightness，but they frequently suffer from inadequate overexposure suppression and heavy reliance on paired datasets collected under normal lighting conditions. Given that such datasets are expensive and difficult to obtain，the generalization ability of these supervised methods remains limited in real-world complex scenarios. Furthermore，some unsupervised methods attempt to reduce dependency on paired data，but they often fail to achieve noise suppression，exposure correction，and structure preservation simultaneously，leading to unsatisfactory performance in practice. **Method** To address these challenges，this study proposes a novel zero-supervision low-light image enhancement framework that utilizes illumination components to guide reflectance estimation，thereby ensuring high-fidelity reconstruction. The framework is designed to operate without paired training data while simultaneously achieving effective noise suppression，overexposure control，and structure preservation. The process begins with a reversible brightness perturbation strategy combined with pseudoexposure generation，which synthesizes abundant pseudo low-light and pseudo well-lit samples without changing the intrinsic reflectance of the image and thus effectively simulates diverse illumination conditions and eliminates the dependence on paired supervision. A dual-domain collaborative attention denoising module is introduced to improve robustness further. This module operates in the spatial and channel domains：in the spatial domain，contextual dependencies are captured to distinguish true structures from random noise，while in the channel domain，correlations among different feature channels are modeled to suppress redundant information and emphasize critical features. By collaborating across these two domains，the module ensures effective denoising while preserving structural fidelity. For structural representation，multilayer residual convolutions and a structurally symmetric U-Net architecture work in concert to generate the illumination map. The symmetric encoder-decoder structure enables extraction of multiscale features，capturing global illumination and local detail，while the residual connections mitigate vanishing gradients and accelerate convergence. The U-Net is designed not only to generate illumination maps but also to extract illumination-related features that serve as explicit guidance for reflectance recovery. On this basis，a guided reflectance estimation method is introduced. Instead of directly treating reflectance as a residual，illumination priors are fused into the estimation process. This fusion enables the network to separate illumination variations effectively from intrinsic structures，restoring edges and textures with high fidelity while suppressing overexposed regions. The integration of illumination priors makes the restored images appear natural and structurally consistent，thereby addressing the shortcomings of previous enhancement methods. **Result** Extensive comparative experiments were conducted on the LOL，DICM，and WildNight-8K datasets against nine representative state-of-the-art low-light image enhancement methods. On the LOL dataset，the proposed method outperformed Retinexformer and LightenDiffusion，achieving reductions of 0. 62 and 0. 92 in the natural image quality evaluator（NIQE）and decreases of 1. 23 and 0. 25 in the perceptual index（PI），respectively. On the DICM dataset，compared with Retinexformer and retinex-inspired unrolling with architecture search（RUAS），the proposed approach reduced NIQE by 0. 11 and 0. 81，and PI by 0. 48 and 0. 72，respectively. On the WildNight-8K dataset，relative to Retinexformer and EnlightenGAN，NIQE was further reduced by 1. 32 and 1. 87，while PI decreased by 0. 33 and 5. 98，respectively. These results demonstrate that the proposed method consistently delivers enhanced images with improved naturalness and perceptual quality across diverse low-light scenarios. In addition to image quality assessment，the enhanced images were further evaluated in a downstream animal detection task using YOLOv8. Experimental results show that the mean average precision（mAP@50−95）increased by 4. 60% compared with the original low-light images，indicating that the proposed enhancement method not only improves visual quality but also remarkably boosts the reliability of downstream perception tasks in challenging low-light conditions. **Conclusion** This study introduces a zero-supervision low-light image enhancement framework that integrates illumination-guided reflectance estimation for high-fidelity image reconstruction. The key innovations include reversible brightness perturbation with pseudoexposure generation to remove paired dataset dependency，a dual-domain collaborative attention denoising module to reduce noise while preserving structural details，a residual U-Net backbone for effective multiscale representation，and an illumination-fused reflectance estimation mechanism to achieve structure-preserving restoration and overexposure suppression. Experimental results across multiple datasets and real-world applications demonstrate that the proposed method not only achieves state-of-the-art perfor-

mance in quantitative and qualitative metrics but also remarkably improves downstream vision tasks, particularly object detection. These contributions indicate strong practical value and broad application potential. Looking forward, future work could explore extending this framework to dynamic video enhancement with temporal consistency, incorporating adaptive illumination priors derived from scene semantics, and optimizing lightweight deployment for edge devices such as autonomous vehicles and mobile cameras. With these directions, the proposed method could further strengthen its role as a reliable solution for low-light visual perception in safety-critical and real-time applications.

Key words: low-light image enhancement; attention mechanism; zero-supervision; object detection; convolutional neural network

0 引　言

弱光图像增强作为计算机视觉与图像处理中的基础问题，对于提升图像视觉质量并保障下游感知任务的可靠性具有重要意义。低照度成像常伴随亮度不足、补光过曝、噪声放大与细节缺失等问题，严重影响了图像的可用性和可解释性。

针对这一问题，研究者提出了多类方法。早期传统方法主要基于直方图均衡、自适应直方图均衡及对比度拉伸等简单增强技术，通过亮度与局部对比度调整改善成像效果。然而在复杂弱光场景下，这类方法往往难以有效抑制噪声，并可能造成纹理细节的进一步丢失，表现出较强的局限性（Guo等，2017）。

近年来，深度学习驱动的弱光增强方法取得了快速发展。以卷积神经网络（convolutional neural network，CNN）为代表的早期工作，多基于曲线估计或图像分解，学习从低光到正常光的端到端映射。然而，CNN方法在处理光照高度不均匀的自然场景时常面临结构保持不足与泛化性能受限的问题。随后，Retinex理论与深度模型的结合成为重要发展方向：如Retinexformer（Cai等，2023）通过引入变换器结构提升长程依赖建模与非均匀光照下的结构保持能力，RetinexMamba（Bai等，2024）则利用状态空间模型（state space model，SSM）在较低计算成本下获得近乎Transformer的表现。同时，CIDNet（Yan等，2025）提出可训练HVI（horizontal/vertical-intensity）色彩空间，该空间中引入可学习映射来替代固定的RGB表征，从而在弱光图像增强过程中更好地分离亮度与色彩信息，这一设计能够有效缓解常见的色偏与色彩不稳定问题，尤其是在长时间曝光或复杂光照条件下，提升了增强结果的自然感与一致性（孙婉倩 等，2026）。FLOL（fast baselines for real-world low-light enhancement）（Benito等，2025）在频域与空域轻量融合的框架下实现了毫秒级推理，展现了在实际部署中的应用潜力。总体而言，Transformer与SSM的引入显著提升了表征与光照建模能力，而色彩空间建模与轻量化路线改善了效率与稳定性，弥补了传统CNN在真实夜景中的不足。

与此同时，扩散模型的兴起为弱光增强带来了新的思路。去噪扩散概率模型（denoising diffusion probabilistic model，DDPM）（Ho等，2020）最初应用于图像生成，在弱光图像中被引入为从退化的低光图像逐步恢复清晰、高质量图像的途径。Yi等人（2023）提出的DiffusionRetinex方法利用扩散过程建模图像在不同曝光条件下的潜在分布，在去噪的同时增强了亮度与细节信息。Jiang等人（2024）提出的LightenDiffusion方法将Retinex分解迁移至潜空间，并采用扩散模型进行无监督恢复，同时引入自约束一致性损失以抑制内容泄漏。在多个真实场景基准数据集上的结果显示，该方法相较于同类无监督方法表现更优，并接近有监督方法的水平，同时具备更好的场景泛化能力。但是，与Transformer相比，基于扩散的增强方法在推理时延上仍然存在劣势，通常需要额外的加速或蒸馏策略满足实际应用需求（Lin等，2025）。

从近3年（2023—2025年）的研究进展来看，弱光图像增强方法在建模范式与应用目标上呈现出明显分化趋势。以Retinexformer（Cai等，2023）和RetinexMamba（Bai等，2024）为代表的方法，通过引入Transformer或SSM显著增强了全局光照建模能力，在非均匀照明和大尺度亮度变化场景下取得了优于传统CNN的效果。然而，这类方法通常依赖复杂的网络结构与多阶段特征交互，对训练数据分布较为敏感，在零监督或真实复杂环境中仍存在一定的泛

化风险。在此基础上，部分研究开始关注弱光增强中的色彩稳定性与感知一致性问题，CIDNet（Yan等，2025）通过构建可训练的HVI色彩空间，实现了亮度与色彩信息的更有效解耦，在长曝光和复杂光照条件下显著缓解了色偏现象。但该类方法主要从颜色表征层面改进增强效果，对噪声放大与结构退化问题的抑制能力相对有限，往往需要与额外的去噪或结构约束模块协同使用。另一方面，扩散模型为弱光增强提供了更强的分布建模能力。Diffusion-Retinex（Yi 等，2023）与 LightenDiffusion（Jiang 等，2024）通过逐步去噪的生成过程，在细节恢复和噪声抑制方面展现出优势，尤其在无监督条件下具备较好的泛化性能。然而，扩散模型普遍存在推理开销大、亮度调节效率低以及局部过曝等问题，限制了其在实时或轻量化应用场景中的实用性。此外，FLOL（Benito 等，2025）等最新的零监督轻量化方法通过频域与空域的高效融合显著提升了推理速度，但在极端弱光和高噪声条件下仍难以兼顾增强质量与结构保真度。综合来看，近3年的方法在全局建模、色彩一致性或生成质量等单一维度上均取得了进展，但在零监督条件下同时兼顾噪声抑制、结构保持、过曝控制与计算效率方面仍存在不足，进一步凸显了在无需成对数据的前提下，构建具备稳定分解能力与鲁棒增强性能的弱光图像增强框架的研究价值。

在上述对比分析的基础上可以看出，国内外弱光图像增强的研究已从传统直方图与对比度增强，发展到CNN驱动的曲线估计与分解网络，再到融合Retinex先验的Transformer/SSM框架，以及近年的扩散模型（Li等，2025）。它们在亮度提升、噪声抑制与结构保持等方面不断取得突破，但整体仍面临三大挑战：1）非均匀光照与补光噪声导致增强结果易出现局部过曝与失真；2）复杂背景下缺乏有效结构建模，增强过程容易破坏目标边界与纹理细节（Wei等，2018），从而影响下游检测与分割性能；3）成对监督数据采集成本高昂，限制了现有方法在零监督或弱监督条件下的适用性（Lin等，2025）。基于此，亟需一种无需成对数据且能够在真实复杂环境下实现弱光增强，同时兼顾噪声抑制、结构保真与视觉真实性的新型方法。

1　本文方法

本文提出的零监督弱光增强框架如图1所示，首先，低光图像由伪曝光生成策略（pseudo- exposure generation，PEG）处理，通过对输入施加可逆亮

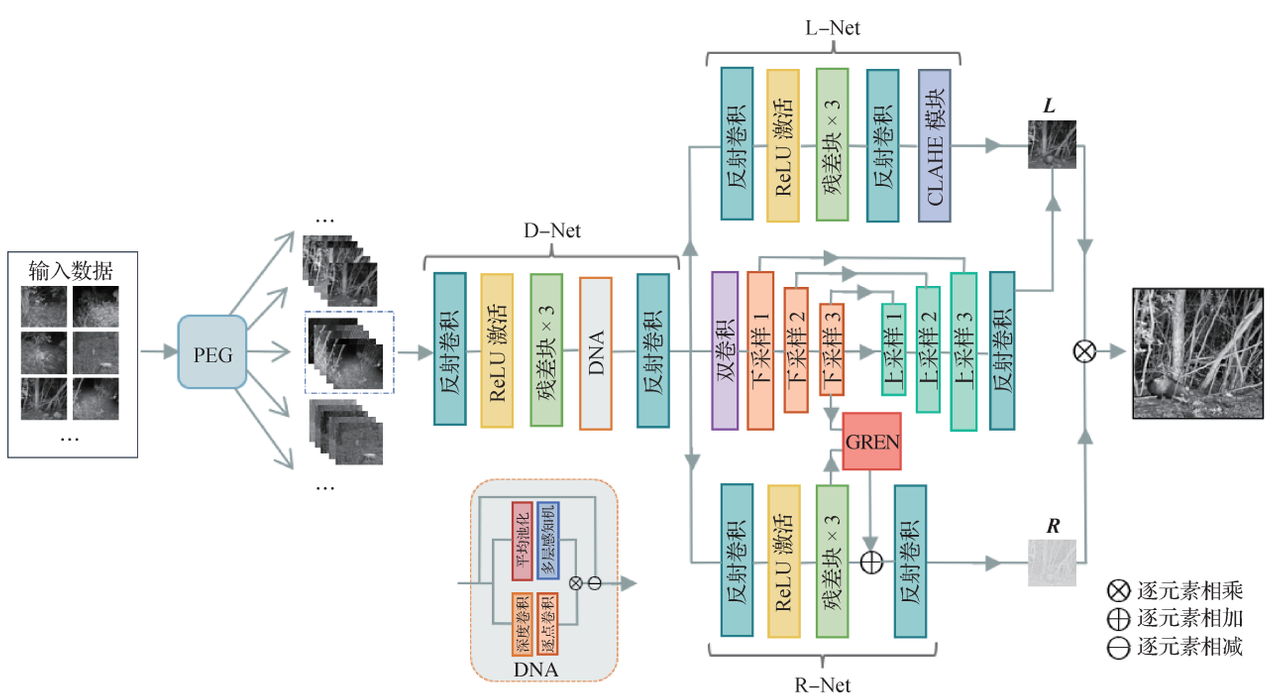


图1　本文网络结构

Fig. 1　The network structure of this article

度变换生成多幅伪曝光图，为后续分解提供无标签训练样本；随后，图像进入带有噪声抑制模块（denoise attention，DNA）的D-Net完成自适应去噪，以获得干净的输入特征；接着，L-Net和U-Net结构估计光照图$\boldsymbol{L}$；R-Net在融合U-Net提取的深层照明特征的基础上，通过引导式反射图增强模块（guided reflectance enhancement network，GREN）重建反射图$\boldsymbol{R}$；最终，将光照与反射逐像素相乘，输出亮度自然且纹理保真的增强结果。

1.1 可逆亮度扰动的伪曝光生成

在Retinex理论框架下，图像$\boldsymbol{I}$可被建模为光照分量$\boldsymbol{L}$与反射分量$\boldsymbol{R}$的乘积，即：$\boldsymbol{I}=\boldsymbol{R}\odot\boldsymbol{L}$（Land和McCann，1971）。其中，$\boldsymbol{R}$表征场景固有的材质与纹理属性，$\boldsymbol{L}$则反映入射光照的分布。传统的监督式Retinex分解方法通常依赖具有严格曝光配对关系的成对图像与正常光图，以此提供对$\boldsymbol{L}$和$\boldsymbol{R}$的监督信息（Guo等，2020）。然而，高质量成对曝光数据在真实拍摄条件下极难获取，严重限制了此类方法在实际场景中的应用范围。为突破成对标签依赖的瓶颈，本文采用了一种零监督的分解策略。具体而言，针对单幅低光输入图像$\boldsymbol{I}$，本文通过伪曝光生成策略（PEG）施加可逆的亮度变换，采用γ校正、线性缩放和对数/指数映射，从而生成多个亮度水平不同的伪曝光图像$\{\boldsymbol{I}_j\}$。在Retinex假设下，这些伪图共享相同的反射信息$\boldsymbol{R}$，但其光照分量$\boldsymbol{L}_j$存在差异。因此，即便缺乏真实曝光标签，也可将它们构造成训练样本，驱动网络学习分离出一致的$\boldsymbol{R}$与各自对应的$\boldsymbol{L}_j$。训练目标通过重建损失、一致性约束和平滑正则等多项损失函数共同优化，逐步引导网络从非成对图像中学习到结构一致、物理意义明确的光照与反射分离（Cai等，2018）。这种策略摆脱了对昂贵成对数据的依赖，在保持训练稳定性的同时显著提升了模型的实际应用能力。

需要指出的是，真实相机曝光变化本质上由入射光通量、曝光时间及传感器响应函数等因素共同决定，其对图像亮度分布的影响通常表现为近似的非线性映射关系。在未知具体相机参数的情况下，γ校正、线性强度缩放以及对数/指数映射等操作能够在统计意义上有效近似不同曝光条件下的亮度变化过程（Wang等，2023）。在此基础上，本文所提出的伪曝光生成策略（PEG）并非试图精确复现某一具体相机的物理曝光模型，而是从数据驱动与Retinex分解一致性的角度出发，通过多种可逆亮度变换构造一组覆盖不同亮度分布的伪曝光样本，以最大化光照分量的多样性。由于这些变换对反射分量的影响可视为一致，从而满足Retinex假设中“反射不随光照变化”的基本前提。

为进一步验证PEG在模拟曝光变化方面的有效性，本文将其与多种常见曝光变换策略进行了对比分析，包括单一γ校正、线性增益调整和基于直方图重映射的亮度增强方法。不同策略生成的伪曝光图像分别用于Retinex分解网络训练，结果如表1所示，相较于单一形式的曝光变换，PEG所生成的多级伪曝光在自然图像质量评估指标（natural image quality evaluator，NIQE）、基于深度残差网络（deep residual network，DRN）指标和感知指数（perceptual index，PI）等指标上取得了更优的表现，说明其在提升光照建模多样性与训练稳定性方面具有明显优势。

表1 曝光变化方法对比分析表

Table 1 Comparative analysis table of exposure variation

曝光变化方法	NIQE↓	DRN↓	PI↓
单一γ校正	23.502 7	0.227 8	17.598 4
线性增益调整	24.150 2	0.240 1	17.450 2
基于直方图重映射	24.857 8	0.254 1	19.248 1
PEG	**22.261 0**	**0.181 4**	**15.566 5**

注：加粗字体表示各列最优结果，↓表示负向指标，数值越低表示越优。

同时，PEG数量的选择直接影响生成伪曝光图像的覆盖范围及模型的泛化能力。当PEG数量过少时，生成的伪曝光样本亮度分布有限，光照分量多样性不足，网络在分离$\boldsymbol{R}$与$\boldsymbol{L}$时容易陷入欠拟合，导致增强后的低光图像仍保留暗区噪声或亮度不均匀；相反，当PEG数量过多时，虽然理论上可以覆盖更广亮度范围，但过多的伪曝光图像会引入大量冗余样本，增加网络训练负担，降低收敛速度，且在极端低光、结构复杂且噪声繁多的图像中可能出现伪影，从而降低图像自然度和NIQE指标表现。为验证PEG数量对不同光照条件下增强效果的影响，本文进行了多级别PEG数量的实验，并通过量化指标对各级别的增强效果进行评估，结果如图2所示。对

于极低光照、噪声多以及结构复杂的场景（如 WildNight-8K 数据集），PEG 数量设置为 5 时，网络能够在保证光照分量多样性的同时避免冗余样本带来的训练负担，因此在 NIQE 指标上表现最佳，这表明适度的伪曝光级别可在极端低光条件下最大化模型的泛化能力与图像自然度。相比之下，对于亮度稍高、噪声较少且结构相对简单的场景，增加 PEG 数量至 7 可以进一步扩展训练样本覆盖范围，使网络在更宽亮度分布下学习到更稳健的光照表示，从而获得略优的增强效果。

这一趋势表明，PEG 数量的最优设置具有一定的条件依赖性：极暗且噪声多的图像倾向于较少的 PEG 级别；而亮度较好、场景简单的图像则可适当增加 PEG 数量以获取更丰富的光照多样性。在本研究中，PEG = 5 可作为低光增强任务的默认优选值，兼顾训练效率与泛化性能；同时，对于特定光照条件或不同场景复杂度的图像，可根据亮度分布与噪声水平微调 PEG 数量，以进一步提升增强效果与模型稳健性。

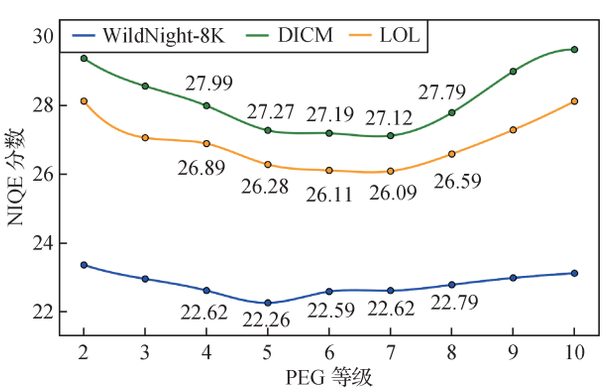


图 2　PEG 分级下的 NIQE 表现曲线

Fig. 2　NIQE performance across PEG levels

1.2　双域协同注意力去噪

为缓解低照图像中普遍存在的噪声干扰问题（Zhang 等，2017），本文在分解前引入噪声抑制模块（DNA），以提升后续照明图与反射图建模的准确性与稳定性。DNA 模块所关注的噪声主要表现为以信号无关的加性高斯噪声为主，并混合少量由光子统计波动引起的泊松噪声，这符合多数 CMOS/CCD（complementary metal-oxide-semiconductor/charge-coupled device）传感器在不同 ISO（international drganization for standardization）设置下的典型低光噪声特性。需要指出的是，本文并未对噪声分布形式进行显式建模，而是通过卷积特征提取与注意力加权机制，在特征域中自适应学习噪声与结构信号之间的统计差异（Foi 等，2008）。该设计使 DNA 模块无需依赖具体传感器型号或 ISO 参数，即可在端到端训练过程中隐式适配不同光照条件下的混合噪声分布，从而实现稳健的低光去噪效果。该模块以浅层卷积和多层残差单元为基础，并结合去噪注意力机制对图像进行自适应去噪。输入灰度图像 $\boldsymbol{X} \in \mathbf{R}^{1\times H\times W}$ 首先通过卷积层提取初始特征。具体为

$$\boldsymbol{F}_0 = \delta(\boldsymbol{C}(\boldsymbol{X})) \tag{1}$$

式中，$\boldsymbol{F}_0$ 表示由卷积层提取的初始特征图，$\delta(\cdot)$ 表示带泄露的修正线性单元激活函数（leaky rectified linear unit，Leaky ReLU），$\boldsymbol{C}$ 表示卷积运算，$\boldsymbol{X}$ 表示输入的灰度低照度图像。随后，特征图 $\boldsymbol{F}_0$ 依次通过 3 个带残差连接的卷积模块，捕捉局部结构信息，形成中间特征 $\boldsymbol{F}_{\mathrm{r}}$，具体为

$$\boldsymbol{F}_{\mathrm{r}} = \boldsymbol{F}_0 + \sum_{i=1}^{3} \boldsymbol{R}_i(\boldsymbol{F}_i) \tag{2}$$

式中，$\boldsymbol{F}_{\mathrm{r}}$ 表示经过多层残差卷积模块后的输出特征图，$\boldsymbol{R}_i$ 表示第 i 个残差卷积模块，$\boldsymbol{F}_i$ 表示第 i 个残差模块的输入特征。为了进一步增强网络对重要区域的关注能力，引入了去噪注意力机制，融合通道注意力（channel attention）与空间注意力（spatial attention）的联合注意力机制（Woo 等，2018）。通道注意力部分首先对输入特征进行全局平均池化，再通过两层逐点卷积建模通道间依赖关系，得到通道注意力权重，具体为

$$\boldsymbol{M}_{\mathrm{c}}(\boldsymbol{F}) = \sigma\left(\boldsymbol{W}_2 \cdot \delta\left(\boldsymbol{W}_1 \cdot P(\boldsymbol{F})\right)\right) \tag{3}$$

式中，$\boldsymbol{M}_{\mathrm{c}}(\boldsymbol{F})$ 表示由通道注意力机制生成的特征图，$\boldsymbol{F}$ 表示输入特征图，σ 表示 sigmoid 函数，$\boldsymbol{W}_1$、$\boldsymbol{W}_2$ 为 1×1 卷积核，P 表示全局平均池化。空间注意力部分通过深度可分离卷积挖掘局部显著性区域，其表达式为

$$\boldsymbol{M}_{\mathrm{s}}(\boldsymbol{F}) = \sigma(\boldsymbol{C}_{1\times 1}(\boldsymbol{C}^{\mathrm{dw}}_{3\times 3}(\boldsymbol{F}))) \tag{4}$$

式中，$\boldsymbol{M}_{\mathrm{s}}(\boldsymbol{F})$ 表示由空间注意力机制生成的特征图，$\boldsymbol{C}^{\mathrm{dw}}$ 表示深度可分离卷积。两种注意力权重经逐点相乘后作用于原始特征图，实现通道与空间维度的联合选择性增强。具体为

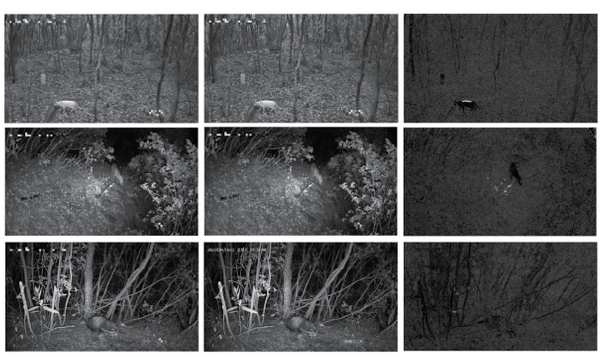
$$\boldsymbol{F}_{\mathrm{a}} = \boldsymbol{F} + \boldsymbol{F} \cdot (\boldsymbol{M}_{\mathrm{c}}(\boldsymbol{F}) \cdot \boldsymbol{M}_{\mathrm{s}}(\boldsymbol{F})) \tag{5}$$

式中，$\boldsymbol{F}_{\mathrm{a}}$ 表示通道与空间维度融合后的特征图。最终，输出特征经尾部卷积层和 sigmoid 归一化后得到

去噪后的图像。具体为

$$\tilde{\boldsymbol{X}} = \sigma(C(\boldsymbol{F}_{\mathrm{a}})) \tag{6}$$

式中，输出将作为后续照明图和反射图估计的输入。图3中的噪声主要表现为低照条件下由传感器读出过程引入的高频随机扰动，并在物体轮廓等高梯度区域更加显著。该噪声在真实成像中通常与泊松噪声混合，但在本实验所示场景下，其统计特性更接近加性高斯噪声，因此DNA模块对该类噪声的有效抑制为后续Retinex分解提供了更加稳定和可靠的输入基础。



(a) 原图　(b) 去噪后的图　(c) 噪声

图3　去噪结果可视化

Fig. 3　Visualization of denoising results

((a)original images; (b)denoised images; (c)noise)

1.3　混合残差与U-Net架构

本文方法用于表征图像全局与局部光照分布的照明图$\boldsymbol{L}$应当具有平滑连续的空间特性，同时能够覆盖强光、阴影等高动态区域。为此，本文设计了多层残差卷积网络作为照明图估计子网络L-Net，并加入结构对称、感受野充足的U-Net架构辅佐多尺度光照信息的有效建模与融合，同时U-Net也对后续R-Net网络建模反射图进行引导。

轻量级的网络L-Net由3部分组成：头部、主体和尾部。头部包含一个反射填充层、一个卷积层和ReLU激活函数；主体由3个残差块组成，用于深层特征提取（潘晓英 等，2022）；尾部通过一个反射填充层和一个卷积层将通道数降为1，最后通过sigmoid函数得到归一化的光照图估计。

U-Net的输入为去噪后的图像$\boldsymbol{X}$，网络首先通过一个双卷积模块提取初始特征，随后依次经过3层下采样编码器逐步获取高层次语义信息，最终再通过对称结构的3层上采样解码器逐级还原分辨率并融合低层特征，实现光照图的重建。具体为

$$\begin{aligned}&\boldsymbol{F}_0 = C(\boldsymbol{X}), \boldsymbol{F}_1 = D_1(\boldsymbol{F}_0), \cdots, \boldsymbol{F}_3 = D_3(\boldsymbol{F}_2)\\&\boldsymbol{F}_4 = Up_1(\boldsymbol{F}_4, \boldsymbol{F}_3), \cdots, \boldsymbol{F}_6 = Up_3(\boldsymbol{F}_5, \boldsymbol{F}_0)\\&\boldsymbol{L} = \sigma(C(\boldsymbol{F}_6))\end{aligned} \tag{7}$$

式中，$\boldsymbol{X}$表示输入的特征图，D表示下采样模块（最大池化+双卷积），Up表示上采样模块（反卷积+跳跃连接+双卷积），$\boldsymbol{F}_i$表示采样后的特征图。每层跳跃连接允许低层的细节信息参与最终输出的恢复，确保在估计光照图时保留边缘平滑性与区域一致性。为了进一步提升感受野并抑制边界伪影，网络中所有卷积操作均配合反射填充（Reflection Padding）实现。此外，L-Net的最终输出经过sigmoid函数归一化至[0, 1]区间，满足Retinex理论对光照分量的定义。在L-Net的训练过程中，为了确保分解的准确性和稳定性，网络采用了多种损失函数约束照明图$\boldsymbol{L}$的学习过程，如图4所示，主要有Retinex Loss和Guidance Loss。其中，与照明图生成有关的损失函数共同构成了Retinex Loss，照明图本身的平滑性通过Total Variation Loss约束。具体为

$$L_{\mathrm{TV}}(\boldsymbol{L}) = \left\| \nabla_x \boldsymbol{L} \right\|_1 + \left\| \nabla_y \boldsymbol{L} \right\|_1 \tag{8}$$

式中，L_{TV}表示损失函数，$\boldsymbol{L}$表示光照分量，包括使用最大通道拟合来近似图像的明度，具体为

$$L_{\mathrm{l}} = MSE(\boldsymbol{L}, \max(\boldsymbol{X})) \tag{9}$$

式中，L_{l}表示损失函数，MSE表示均方误差损失函数，$\boldsymbol{L}$表示预测光照图，$\boldsymbol{X}$表示去噪后的输入图像，max(·)表示沿通道维度取最大值以近似图像亮度。此外，为保证Retinex分解满足理论上的乘积一致性，网络在训练过程中引入了反射与照明的重构一致性约束。根据Retinex公式，得

$$\boldsymbol{X} = \boldsymbol{L} \cdot \boldsymbol{R} \Rightarrow \boldsymbol{R} = \frac{\boldsymbol{X}}{\boldsymbol{L}} \tag{10}$$

式中，$\boldsymbol{X}$是原始特征图，$\boldsymbol{L}$是光照信息分量，$\boldsymbol{R}$是反射图分量。因此，网络在分解结果满足物理模型的同时也对反射图与反演值之间引入均方误差约束项，具体为

$$L_{\mathrm{r}} = \left\| \boldsymbol{L} \cdot \boldsymbol{R} - \boldsymbol{I} \right\|_2^2 \tag{11}$$

$$L_{\mathrm{d}} = \left\| \boldsymbol{R} - \frac{\boldsymbol{I}}{\boldsymbol{L} + \epsilon} \right\|_2^2 \tag{12}$$

式中，L_{d}表示损失函数，ϵ为一个很小的常数（如1×10^{-6}），用于防止除零异常。

鉴于U-Net的下采样部分凭借多尺度特征提取机制，能够高效捕捉到丰富的语义信息与精细的空间结构，本文决定利用这一优势为R-Net提供引导。

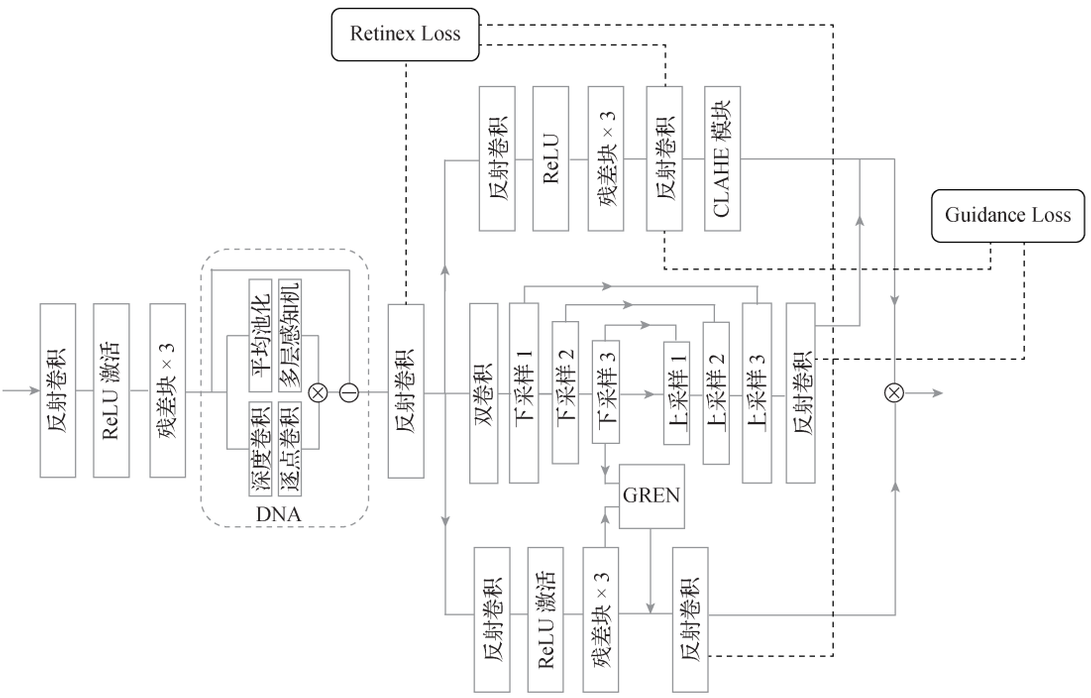


图 4　损失函数设计示意图

Fig. 4　Overview of the loss function design

为了确保 U-Net 的下采样部分能够充分吸收 L-Net 所蕴含的关键信息，进而为 R-Net 提供有效的指导，引入了一个额外的损失函数 Guidance Loss。具体为

$$L_g = \frac{1}{N}\sum_{i=1}^{N}\left| \boldsymbol{L}_{u_i} - \boldsymbol{L}_i \right| \tag{13}$$

式中，L_g 表示损失函数，N 表示参与损失计算的样本总数，$\boldsymbol{L}_{u_i}$ 为 U-Net 网络输出特征，$\boldsymbol{L}_i$ 为 L-Net 网络输出的特征。多种损失函数设计能够促使 U-Net 的输出尽可能地贴近 L-Net 的输出，在模型训练过程中，实现二者输出的高度一致性，确保信息在二者之间的顺畅传递，为 R-Net 的优化训练奠定了坚实的基础。

1.4　过曝感知的引导式反射估计

在本文分解框架中，反射图含有图像中与光照无关的固有属性，包括结构纹理和材质特征，但分解时其中携带的不必要光照信息会引起局部区域过曝。为了抑制过曝现象并生成高质量的反射图估计，本文设计了一个带有照明定向信息的引导式反射图增强模块(GREN)。如图 5 所示，R-Net 接收由去噪子网络(D-Net)输出的图像 $\boldsymbol{X}'$，并通过多层残差卷积提取了 GREN 模块的初始输入特征之一 $\boldsymbol{F}_R \in \mathbf{R}^{64\times H\times W}$。

同时，从与照明估计网络对齐的 U-Net 中提取深层语义特征 $\boldsymbol{F}_L \in \mathbf{R}^{256\times H/8\times W/8}$，通过两层 1 × 1 卷积将

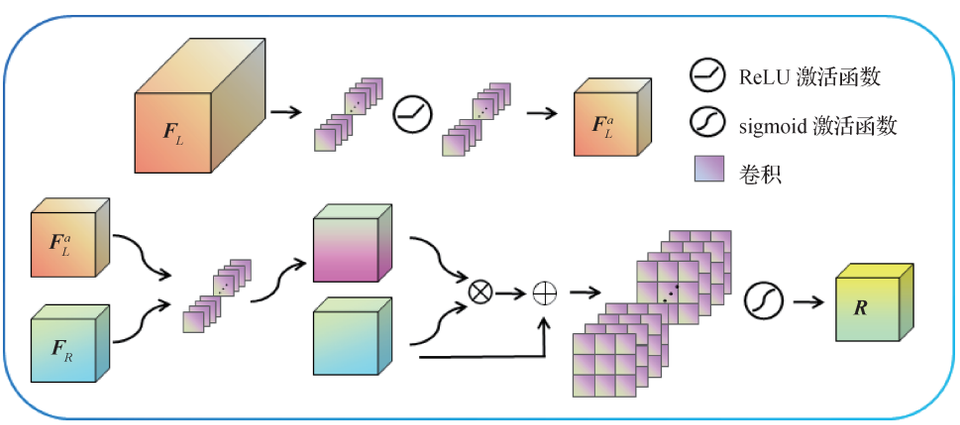


图 5　GREN 模块细节

Fig. 5　Detailed illustration of GREN

其通道数从256降至64，实现通道压缩，并通过双线性插值对降维后的特征进行空间对齐，确保与$\boldsymbol{F}_R$在空间维度的一致性。具体为

$$\boldsymbol{F}_L^d = \sigma\left(\delta\left(C_{1\times1}\left(\boldsymbol{F}_L\right)\right)\right) \tag{14}$$

$$\boldsymbol{F}_L^a = I\left(\boldsymbol{F}_L^d, \boldsymbol{F}_R\right) \tag{15}$$

式中，$\boldsymbol{F}_L$表示通过多层残差卷积提取的特征图，I表示对特征图在空间维度进行对齐。随后将两个特征图在通道维度上拼接，并通过一层1×1卷积进行压缩与融合，得到引导增强后的特征表示为

$$F_t = C_{1\times1}\left(\left[\boldsymbol{F}_R; \boldsymbol{F}_L^a\right]\right) \tag{16}$$

式中，$\boldsymbol{F}_L^a$表示$\boldsymbol{F}_L^d$与$\boldsymbol{F}_R$空间维度对齐的特征图。该引导融合机制能够显式地向R-Net提供照明上下文信息，从而增强模型对边缘、纹理与阴影区域的判别能力，降低照明变化对反射估计的影响。融合特征$\boldsymbol{F}_f$随后被送入尾部输出卷积进行编码，进一步提取高阶结构信息后输出单通道反射图，并使用sigmoid激活函数归一化至[0,1]区间，具体为

$$\boldsymbol{R} = \sigma(C(\boldsymbol{F}_R + \boldsymbol{F}_R * \boldsymbol{F}_f)) \tag{17}$$

式中，$\boldsymbol{F}_f$表示压缩融合后的特征图，$\boldsymbol{R}$表示GREN模块处理后的输出特征图。该网络设计通过引入照明引导特征(Guo等，2025)，显著增强了反射图对结构与纹理的表达能力，有效提升了光照与反射分量的分离质量，为后续图像增强提供了稳定的反射信息支撑。

2 实 验

2.1 数据集和实施细节

2.1.1 数据集介绍

本研究使用了3个数据集进行实验评估：1个私有弱光图像数据集WildNight-8K(Wildlife Night-time 8000-image dataset)和2个公开数据集LOL和DICM(Lee等，2013)。

其中，WildNight-8K，包含8 000幅图像，来源于自然动物保护区的10个监测摄像头，分辨率为3 200×1 800像素。由于该数据集涵盖真实的野生动物夜间监控场景，具有较强的实际代表性和复杂光照条件，因此对本文研究具有重要意义。数据质量受限于监控设备性能和环境光照变化，存在噪声与纹理复杂性，且不包含对应的正常光照质量参考图，因此本文在该数据集上采用无监督指标进行评估。表2给出了详细统计信息，图6展示了部分典型样例。

表2 WildNight-8K数据集统计信息

Table 2 WildNight-8K dataset statistics

属性	统计信息
图像来源	浙江省东阳市东白山自然保护区
图像数量	8 000幅
图像类别	夜间短波红外成像图
分辨率	3 200×1 800像素
拍摄时段	20:00—次日5:00
拍摄场景	密林、灌木带、草地、藤蔓覆盖区等
摄像头总数	108
动物类别	果子狸、黄鹿、猪獾、白鹇、野猪等
亮度平均值	74.75
亮度最小值	64.82
亮度最大值	93.02



图6 WildNight-8K数据集典型样例图

Fig. 6 Typical examples from the WildNight-8K dataset

此外，为验证本文方法在典型无参考弱光增强场景中的适用性，实验在DICM和LOL数据集上进行。DICM包含64幅来自多种数码相机的真实低光图像，涵盖室内昏暗环境、室外夜景以及复杂光照条件，如背光和阴影区域；LOL数据集则提供500幅配对的低光/正常光图像。两者的结合能够全面评估算法在不同实际场景下的鲁棒性与适应性。

2.1.2 实施细节

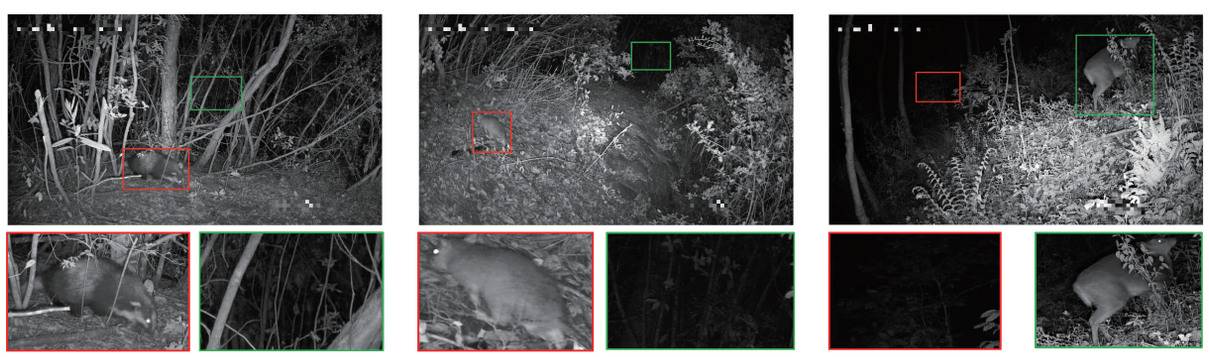
本文使用自适应矩估计优化器(adaptive moment estimation，Adam)，设置$\beta_1=0.9$，$\beta_2=0.999$，初始学习率设置为1×10^{-4}。训练过程中采用多步学习率衰减策略，每训练100个epoch将学习率衰减为原来的1/2。训练批次大小设为4，总共训练300个epoch。训练使用单张NVIDIA GPU 4060Ti(启用了CUDA加速)。模型的训练过程包括4种损失函数的联合优化，为防止训练过程中的不稳定情况，加入了NaN(not-a-number)检测逻辑，并在每次反向传播后应用了最大范数为1.0的梯度裁剪，以稳定模型更新过程。

2.2 评价指标

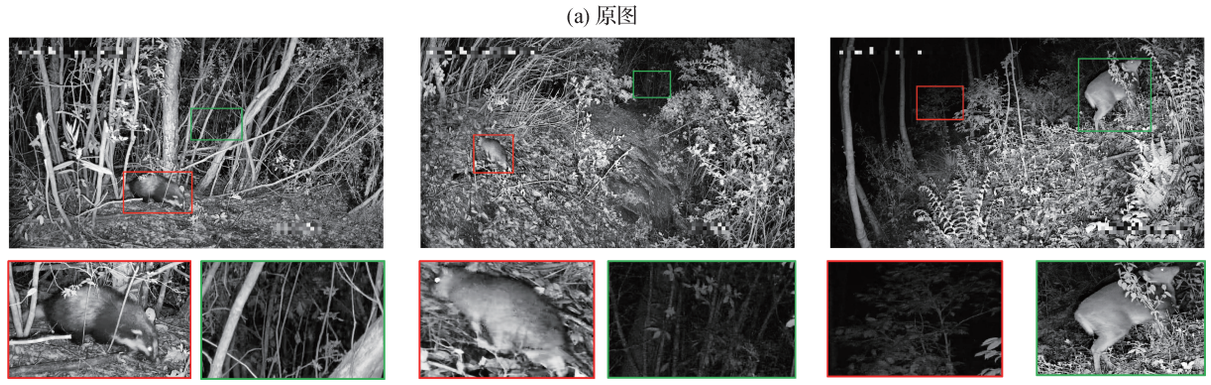
为了客观且全面地评估所提出图像增强算法在弱光条件下的性能表现，本文选用了5种广泛应用且具有代表性的图像质量评价指标，包括3种传统无参考指标：自然图像质量评估指标（NIQE）（Mittal等，2013）、无参考图像空间质量评估器（blind/referenceless image spatial quality evaluator，BRISQUE）（Mittal等，2012）和基于深度残差网络DRN指标（Zhang等，2017），以及两种感知导向指标：感知指数（PI）（Blau等，2018）和学习感知图像块相似度（learned perceptual image patch similarity，LPIPS）（Zhang等，2017）。其中，NIQE通过衡量图像统计特征与自然图像模型的差异评估自然度，BRISQUE侧重提取局部空间特征以评估失真程度，DRN利用深度学习模型模拟人类主观感知以更精准地评价视觉质量，而PI通过无参考指标提供更符合人类视觉感知的整体评分，LPIPS基于深度特征空间的感知相似性度量，可以有效反映增强后图像在结构与纹理上的变化。通过这5种指标的综合分析，能够更加全面地反映算法在亮度改善、细节保留以及感知自然度等方面的性能表现。

2.3 对比实验

本研究算法在真实弱光场景下的增强效果如图7所示，在有效提升图像亮度的同时，避免了过曝现象的出现，并显著增强了图像的可见性。增强后的图像在保持自然亮度分布的基础上，展现出更丰富的纹理和细节信息，且未引入明显的噪声或伪影，整体视觉质量得到显著改善。



(a) 原图



(b) 弱光增强图像

图7 本文方法增强效果细节图

Fig. 7 Detailed images of the enhanced effect of this method ((a) original images; (b) low-light enhanced images)

为了验证本文方法的有效性和鲁棒性，分别在WildNight-8K、DICM和LOL数据集上与9种代表性弱光图像增强方法进行了定量与定性对比分析。所选方法涵盖了经典的传统算法与近年来提出的主流深度学习模型，包括ZERO-DCE（zero-reference deep curve estimation）（Guo等，2020）、RetinexNet（Wei等，2018）、RUAS（retinex-inspired unrolling with architecture search）（Liu等，2021）、EnlightenGAN（Jiang等，2021）、SCI（self-calibrated illumination）（Ma等，2022）、RetinexFormer（Cai等，2023）、LightenDiffusion（Jiang等，2024）、FLOL（fast baselines for real-world low-light enhancement）（Benito等，2025）以及CIDNet（color and intensity decoupling network）（Yan等，2025）。本文采用NIQE、BRISQUE和DRN 3种无参考图像质量评价指标，对各方法增强结果的客观性能进行评估，结果见表3和表4。

表3　不同模型在LOL数据集上的对比实验

Table 3　Comparative experiments of different models on LOL dataset

模型	类型	LOL						
		参数量/M	FLOPs/G	NIQE ↓	BRISQUE ↓	DRN ↓	LPIPS ↓	PI ↓
RetinexNet(Wei等,2018)	有监督	0.84	587.47	32.966 2	25.669 4	0.412 8	0.397 9	19.194 2
RUAS(Liu等,2021)	有监督	0.003	0.83	27.200 1	20.361 8	0.780 8	0.270 7	14.857 7
SCI(Ma等,2022)	有监督	0.003 5	0.69	30.188 6	22.902 2	0.601 4	0.349 7	18.417 4
Retinexformer(Cai等,2023)	有监督	1.61	15.57	26.903 3	14.729 5	0.358 3	0.282 1	15.021 1
CIDNet(Yan等,2025)	有监督	1.88	7.57	28.541 1	19.967 7	0.400 5	0.297 9	13.837 6
ZERO-DCE(Guo等,2020)	无监督	0.075	4.83	27.974 5	20.250 4	0.598 2	0.330 5	15.502 1
EnlightenGAN(Jiang等,2021)	无监督	114.35	61.01	28.102 4	16.226 8	0.620 1	0.309 7	14.998 1
LightenDiffusion(Jiang等,2024)	无监督	27.81	59.32	27.205 8	15.085 9	0.525 6	0.338 3	14.044 5
FLOL(Benito等,2025)	无监督	0.81	8.36	28.337 1	21.120 3	0.342 2	0.296 3	15.240 2
本文	无监督	2.64	17.32	**26.281 2**	**14.551 1**	**0.337 6**	**0.261 1**	**13.791 7**

注:加粗字体表示各列最优结果。↓表示负向指标,数值越低表示越优。

表4　不同模型在WildNight-8K和DICM数据集的对比实验

Table 4　Comparative experiments of different models on WildNight-8K and DICM datasets

模型	类型	WildNight-8K				DICM			
		NIQE ↓	BRISQUE ↓	DRN ↓	PI ↓	NIQE ↓	BRISQUE ↓	DRN ↓	PI ↓
RetinexNet(Wei等,2018)	有监督	47.772 3	33.018 1	0.198 6	27.580 2	32.805 6	30.395 6	0.665 2	21.157 9
RUAS(Liu等,2021)	有监督	25.330 8	37.946 3	1.196 7	21.409 2	28.080 8	24.085 3	1.228 8	18.910 2
SCI(Ma等,2022)	有监督	39.476 6	31.340 6	0.998 4	24.25 7	35.797 7	29.384 3	0.665 3	22.467 3
Retinexformer(Cai等,2023)	有监督	23.585 8	28.491 4	0.204 0	15.901 2	27.381 7	16.167 6	0.574 9	18.665 2
CIDNet(Yan等,2025)	有监督	26.087 5	33.457 9	0.246 3	16.102 4	28.456 9	24.882 2	0.596 7	18.735 5
ZERO-DCE(Guo等,2020)	无监督	25.658 7	35.287 5	0.552 4	21.980 2	30.221 7	24.258 8	0.699 5	20.145 2
EnlightenGAN(Jiang等,2021)	无监督	24.128 7	44.548 7	0.356 5	21.543 3	28.717 7	18.954 8	0.780 3	19.354 2
LightenDiffusion(Jiang等,2024)	无监督	24.035 3	31.574 3	0.189 6	16.580 1	28.183 5	26.512 1	0.686 7	18.882 3
FLOL(Benito等,2025)	无监督	26.456 5	31.106 5	0.651 3	17.258 1	29.470 9	23.218 6	0.578 8	19.421 1
本文	无监督	**22.261 0**	**27.191 1**	**0.181 4**	**15.566 5**	**27.272 8**	**15.104 9**	**0.564 2**	**18.187 5**

注:加粗字体表示各列最优结果。↓表示负向指标,数值越低表示越优。

同时,图8—图10在WildNight-8K、DICM和LOL数据集给出了各方法的视觉对比。在WildNight-8K数据集上,原始图像整体黯淡,仅零星区域出现亮斑,RetinexNet虽然抬升局部暗部亮度,但整体偏暗且伴随明显糊化与噪声放大;RUAS、ZERO-DCE、Retinexformer与FLOL在增亮方面表现积极,但易引入大面积过曝。EnlightenGAN、LightenDiffusion及FLOL虽然抑制了高光溢出,却整体色调偏黄偏绿;CIDNet在降噪与亮度提升上较为均衡,然而纹理细节仍显粗糙。在DICM数据集上的实验结果表明,RetinexNet在提升图像亮度的同时引入了较多噪声;RUAS、ZERO-DCE和Retinexformer均存在明显的过曝光现象;EnlightenGAN和FLOL虽能有效增强亮度,但整体色调偏黄,存在色差问题;LightenDiffusion和CIDNet虽然未出现上述问题,但与本文方法相比,对图像中后段背景的亮度提升效果有限,且在

树木等纹理细节的还原上不够自然。在LOL数据集上，各方法在亮度恢复与色彩一致性方面的差异明显，LOL场景为室内弱光日常环境，RetinexNet在该数据集上依旧存在分解不稳定的问题，增强后图像出现明显的颜色偏移与伪彩现象，且暗部噪声被进一步放大；RUAS与ZERO-DCE虽然能显著提升整体亮度，但伴随局部区域过曝，尤其在前景高反射物体处细节丢失严重；Retinexformer在亮度提升上较为充分，但仍存在亮区泛白和柜子颜色失真的问题。FLOL在一定程度上缓解了过曝现象，但增强结果整体偏黄偏暖，色彩还原准确性不足；EnlightenGAN、LightenDiffusion和CIDNet在色彩稳定性方面表现较好，噪声控制相对理想，但对深暗区域（书本部分）的提亮幅度有限，导致整体观感仍偏暗，部分低对比纹理难以辨识。综上所述，本文方法在显著提亮全局的同时，精细还原了远景植被、近景纹理，有效抑制过曝与噪声，亮度自然、细节丰富，视觉效果最接近真实光照场景（Dong等，2025）。

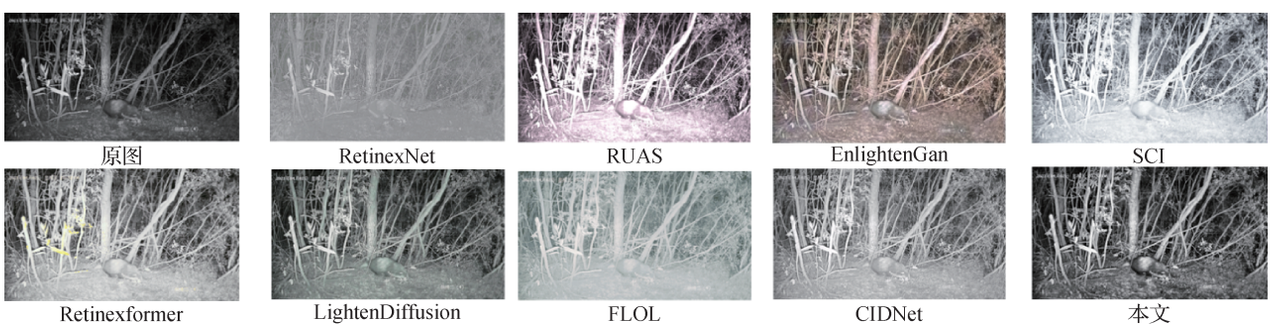


图8　不同模型在WildNight-8K数据集上图像增强任务的可视化结果对比

Fig. 8　Visual comparison of enhancement results across different models on WildNight-8K dataset

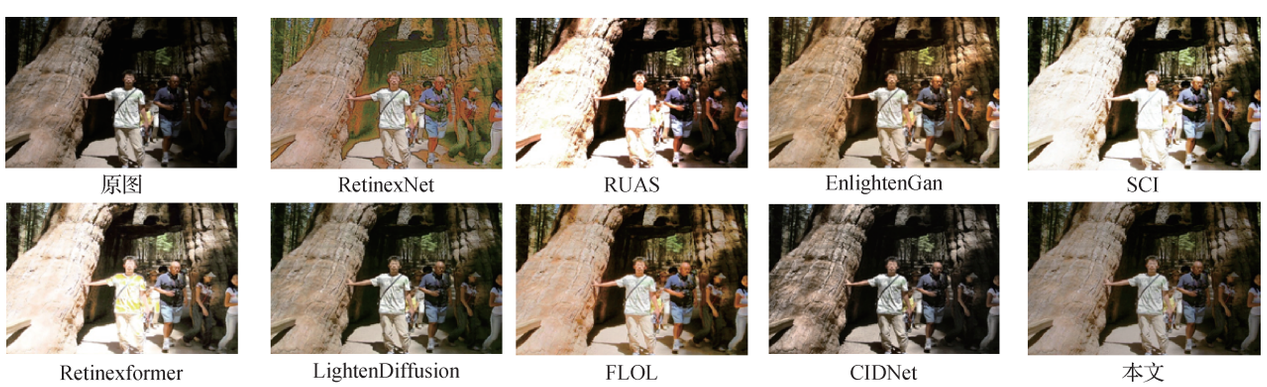


图9　不同模型在DICM数据集上图像增强任务的可视化结果对比

Fig. 9　Visual comparison of enhancement results across different models on DICM dataset

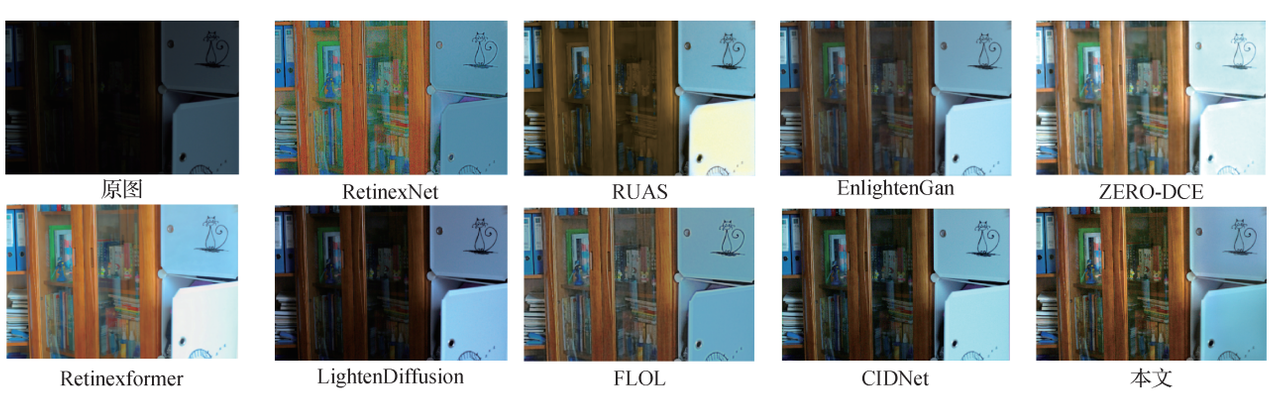


图10　不同模型在LOL数据集上图像增强任务的可视化结果对比

Fig. 10　Visual comparison of enhancement results across different models on LOL dataset

2.4 消融实验

为了验证本文提出的3项关键模块——伪曝光图像生成(PEG)、去噪注意力机制(DNA)以及光照图引导的反射图增强(GREN)在模型中的有效性，如表5所示，以去除全部模块的模型为基准，在WildNight-8K上依次加入单一模块重新训练：加入PEG后反射细节明显，NIQE与DRN指标均改善，说明该模块可以引导网络学习更有效的结构细节，有助于重建低光条件下模糊或缺失的反射信息；加入DNA后噪声与过曝明显减少，NIQE和DRN指标分别下降11.6%和15.4%，验证了该注意力机制对噪声抑制的有效性；加入GREN则促进了光照与反射的平衡，图像结构和细节还原方面变得自然，过曝区域显著减少。相比之下，加入全部模块的完整模型在所有无参考指标上均保持最优，增强图像在亮度、清晰度和自然感方面表现最优，充分证明各模块协同增益的重要性。

表5 增量消融实验表

Table 5 Incremental ablation study table

基准	PEG	DNA	GREN	NIQE↓	DRN↓
√	–	–	–	27.409 3	0.310 9
√	√	–	–	25.355 7	0.282 2
√	–	√	–	24.237 0	0.263 1
√	–	–	√	24.290 9	0.262 0
√	√	√	√	**22.261 0**	**0.181 4**

注：加粗字体表示各列最优结果；√表示采用相应部分，–表示未采用；↓表示负向指标，数值越低表示越优。

为了进一步系统性验证本文提出的伪曝光图像生成(PEG)、去噪注意力机制(DNA)以及光照图引导的反射图增强(GREN)在整体框架中的独立贡献与协同作用，本文开展了更加完善的消融实验，如表6所示，以包含全部模块的完整模型为参照，分别构建w/o PEG、w/o DNA和w/o GREN的对比模型(w/o)表示移除。结果表明，去除PEG后，模型在暗区结构恢复能力明显下降，反射细节模糊，NIQE与DRN分别上升至24.210 7和0.240 2，说明伪曝光样本在引导网络学习稳定反射结构方面起到了关键作用；去除DNA后，噪声抑制能力显著减弱，高亮区域出现明显过曝，NIQE与DRN均明显劣化，验证了去噪注意力机制在抑制噪声放大和高光失真中的核心作用；去除GREN则导致光照与反射分离不充分，图像整体结构虽仍保持，但局部亮区过曝、纹理不自然，NIQE和DRN分别为23.856 6和0.225 8。相比之下，完整模型在两项无参考指标上均取得最优结果，充分说明三者在亮度提升、噪声抑制与结构恢复之间形成了有效的协同增益。

表6 消融实验表

Table 6 Ablation study table

模型	PEG	DNA	GREN	NIQE↓	DRN↓
w/o PEG	–	√	√	24.210 7	0.240 2
w/o DNA	√	–	√	23.962 7	0.235 5
w/o GREN	√	√	–	23.856 6	0.225 8
完整模型	√	√	√	**22.261 0**	**0.181 4**

注：加粗字体表示各列最优结果；√表示采用相应部分，–表示未采用；↓表示负向指标，数值越低表示越优。

其次，为排除性能提升仅来源于模型规模扩大的可能性，本文进一步对DNA与GREN模块引入前后的参数量与计算复杂度进行了对比分析，PEG属于数据层面的伪曝光生成策略，而非可学习的网络模块，在推理阶段不会引入额外的可训练参数或计算开销。结果如表7显示，相比w/o DNA与w/o GREN的模型，完整模型参数量(Params)仅增加至2.64 M，FLOPs(floating point operations per second)为17.32 G，增幅有限，却带来了显著的NIQE与DRN改善，说明本文方法在计算开销可控的前提下实现了更高的增强质量，具有良好的效率—性能平衡。

表7 复杂度消融分析表

Table 7 Ablation study on model efficiency (Params/FLOPs)

模型	参数量/M	FLOPs/G
w/o DNA	2.42	13.67
w/o GREN	2.59	13.54
完整模型	2.64	17.32

最后，本文进一步对backbone中U-Net的编码器—解码器结构配置进行了系统对比，不仅考察了不同网络深度的对称结构，还分析了非对称结构对低光增强性能的影响。实验结果如表8所示，当编码器/解码器均为2层(E2-D2)时，网络表达能力受限，难以充分建模复杂低光场景中光照分布与噪声

耦合关系，整体增强效果较弱；随着网络加深至对称的3层结构（E3–D3），模型在NIQE和DRN指标上取得显著提升（22.261 0 / 0.181 4），在亮度恢复、噪声抑制与细节保持之间达到最佳平衡；进一步加深至对称的4层结构（E4–D4）时，尽管参数量和计算量明显增加，但性能提升趋于饱和甚至出现回退，说明过深网络容易引入冗余特征建模，并可能干扰光照和反射分解的稳定性。

与此同时，对非对称U-Net结构编码器4层—解码器3层（（E4–D3））和编码器3层—解码器4层（E3–D4）进行对比实验。结果表明，编码器与解码器层级匹配对于多尺度光照信息的稳定传递至关重要，对称的U-Net结构能够在下采样与上采样过程中保持光照语义与空间尺度的一致性，使不同尺度的光照特征得以逐级、有效地恢复；而非对称结构或过深网络易破坏这种尺度对齐，导致光照建模冗余或局部失真。

综上所述，多角度消融实验从模块有效性、复杂度控制以及网络结构设计等方面全面验证了本文方法的合理性与鲁棒性，进一步说明提出的PEG、DNA与GREN模块在低光增强任务中各司其职、相互协同，共同促成了整体性能的显著提升。

2.5　基于下游任务的评估

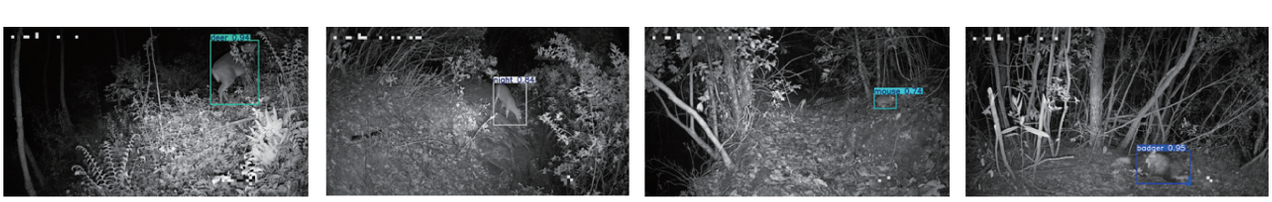
为了进一步验证所提出弱光图像增强方法在真实应用场景中的实用价值，开展了基于任务的评估实验（task-driven evaluation），将增强后的图像应用于下游的动物目标检测任务，考察其对目标识别性能的影响。具体而言，选用当前性能领先的检测模型YOLOv8（you only look once version 8），并以Wild-Night-8K数据集中经不同增强方法处理后的图像作为训练样本，构建动物检测任务数据集。所有检测模型均在统一配置下训练，唯一变量为图像增强策略。本文采用主流的目标检测评价指标平均精度均值（mean average precision，mAP）在测试集上对检测性能进行评估。实验结果如表9所示，使用本文方法增强的图像作为训练数据，YOLOv8获得了显著更高的检测精度，在各项性能指标上均优于原始弱光图像及其他增强方法。本文增强图像在检测中的可视化示例如图11所示，进一步从任务导向的角度验证了本文方法在提升图像可见性、结构可读性与真实感方面的有效性，展现出其在实际应用中的强大潜力和泛化能力。

表8　U-Net网络深度与编码/解码对称性的消融实验

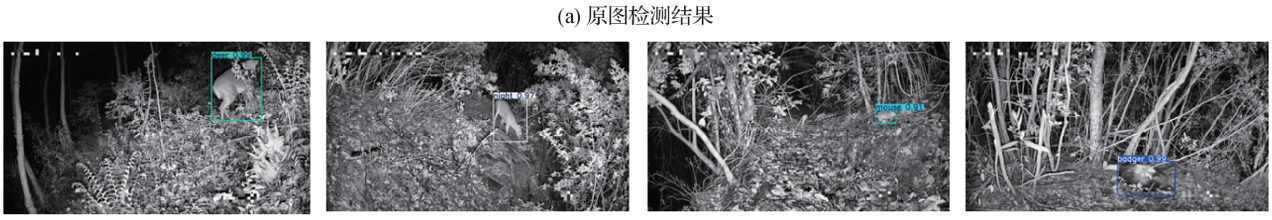
Table 8　Ablation study on U-Net depth and encoder/decoder symmetry

U-Net结构	NIQE ↓	DRN ↓	参数量/M	FLOPs/G
E2–D2	24.742 1	0.236 6	**1.46**	**16.82**
E3–D3	**22.261 0**	**0.181 4**	2.64	17.32
E4–D4	23.021 1	0.189 9	8.51	18.13
E3–D4	23.180 1	0.197 7	2.65	17.56
E4–D3	23.578 9	0.214 3	8.47	17.54

注：加粗字体表示各列最优结果；↓表示负向指标，数值越低表示越优。



(a) 原图检测结果



(b) 增强图像检测结果

图11　YOLOv8检测结果的可视化对比

Fig. 11　Visual comparison of YOLOv8 detection results

((a) detection results on original images; (b) detection results on enhanced images)

表9　不同增强模型在YOLOv8检测中的性能对比

Table 9　Comparison of YOLOv8 detection performance on images enhanced by different models

模型	mAP@50-95(YOLOv8)
原图	0.827
RetinexNet(Dong等,2022)	0.838
RUAS(Liu等,2021)	0.863
EnlightenGAN(Jiang等,2021)	0.847
SCI(Ma等,2022)	0.849
Retinexformer(Cai等,2023)	0.841
Retinexmamba(Bai等,2024)	0.859
LightenDiffusion(Jiang等,2024)	0.842
FLOL(Benito等,2025)	0.837
Cidnet(Yan等,2025)	0.843
本文	**0.873**

注:加粗字体表示最优结果。@50-95表示IoU阈值为0.50到0.95。

3　结　论

针对弱光条件下图像易出现亮度不足、噪声放大、局部过曝及结构细节丢失等问题,本文提出了一种基于光照分量引导反射分量估计的零监督弱光图像增强方法。在无需成对正常光照数据的前提下,引入伪曝光生成策略构建多亮度训练样本,结合去噪注意力机制与光照引导的反射增强模块,构建了一个弱光增强框架,实现了光照与反射分量的稳定分解与高质量重建。

在WildNight-8K、DICM和LOL 3个弱光图像数据集上的实验结果表明,本文方法在NIQE、PI和DRN等多项评价指标上整体优于RetinexNet、RUAS和ZERO-DCE等代表性方法,能够在提升图像亮度的同时有效抑制过曝与噪声放大,并较好地保持结构与纹理细节。进一步的YOLOv8目标检测实验显示,基于本文方法增强的图像在mAP@50-95指标上较原始弱光图像取得明显提升,验证了该方法对下游感知任务的有效促进作用。

尽管本文方法在增强质量和鲁棒性方面取得了较好效果,但仍存在一定的局限性,如模型结构相对复杂、计算开销较高,以及在极端噪声或强局部光源条件下反射与光照分离仍不完全。未来将重点开展模型轻量化设计,进一步提升复杂弱光场景下的稳定性,并拓展其在更多下游视觉任务和实际应用中的适用性。

参考文献(References)

Bai J S, Yin Y H, He Q Y, Li Y X and Zhang X F. 2024. Retinexmamba: retinex-based Mamba for low-light image enhancement//Proceedings of the 31st International Conference on Neural Information Processing. Auckland, New Zealand: Springer: 427-442 [DOI: 10.1007/978-981-96-6596-9_30]

Benito J C, Feijoo D, Garcia A and Conde M V. 2025. FLOL: fast baselines for real-world low-light enhancement [EB/OL]. [2025-10-01]. https://arxiv.org/pdf/2501.09718.pdf

Blau Y, Mechrez R, Timofte R, Michaeli T and Zelnik-Manor L. 2018. The 2018 PIRM challenge on perceptual image super-resolution//Leal-Taixé L, Roth S, eds. Computer Vision —— ECCV 2018 Workshops. Cham, Germany: Springer: 1-12 [DOI: 10.1007/978-3-030-11021-5_21]

Cai J R, Gu S H and Zhang L. 2018. Learning a deep single image contrast enhancer from multi-exposure images. IEEE Transactions on Image Processing, 27(4): 2049-2062 [DOI: 10.1109/TIP.2018.2794218]

Cai Y H, Bian H, Lin J, Wang H Q, Timofte R and Zhang Y L. 2023. Retinexformer: one-stage Retinex-based transformer for low-light image enhancement//Proceedings of 2023 IEEE/CVF International Conference on Computer Vision (ICCV). Paris, France: IEEE: 12470-12479 [DOI: 10.1109/ICCV51070.2023.01149]

Dong W, Min Y, Zhou H and Chen J. 2025. Towards scale-aware low-light enhancement via structure-guided transformer design//Proceedings of 2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Nashville, USA: IEEE: 1454-1461 [DOI: 10.1109/CVPRW67362.2025.00135]

Foi A, Trimeche M, Katkovnik V and Egiazarian K. 2008. Practical Poissonian-Gaussian noise modeling and fitting for single-image raw-data. IEEE Transactions on Image Processing, 17(10): 1737-1754 [DOI: 10.1109/TIP.2008.2001399]

Guo C L, Li C Y, Guo J C, Loy C C, Hou J H, Kwong S, et al . 2020. Zero-reference deep curve estimation for low-light image enhancement//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, USA: IEEE: 1777-1786 [DOI: 10.1109/CVPR42600.2020.00185]

Guo X J, Li Y and Ling H B. 2017. LIME: low-light image enhancement via illumination map estimation. IEEE Transactions on Image Processing, 26(2): 982-993 [DOI: 10.1109/TIP.2016.2639450]

Guo X J, Tian Z Q, Wang Y H, Li S Q, Jiang Y, Du S Y, et al. 2025. ERetinex: event camera meets retinex theory for low-light image enhancement//Proceedings of 2025 IEEE International Conference on Robotics and Automation (ICRA). Atlanta, USA: IEEE: 3759-3765 [DOI: 10.1109/ICRA55743.2025.11127340]

Ho J, Jain A and Abbeel P. 2020. Denoising diffusion probabilistic mod-

els//Proceedings of the 34th International Conference on Neural Information Processing Systems. Vancouver, Canada: Curran Associates, Inc.: #574

Jiang H, Luo A, Liu X H, Han S C and Liu S C. 2024. Lightendiffusion: unsupervised low-light image enhancement with latent-Retinex diffusion models//Proceedings of the 18th European Conference on Computer Vision. Milan, Italy: Springer: 161-179 [DOI: 10.1007/978-3-031-73195-2_10]

Jiang Y F, Gong X Y, Liu D, Cheng Y, Fang C, Shen X H, et al. 2021. EnlightenGAN: deep light enhancement without paired supervision. IEEE Transactions on Image Processing, 30: 2340-2349 [DOI: 10.1109/TIP.2021.3051462]

Land E H and McCann J J. 1971. Lightness and Retinex theory. Journal of the Optical Society of America, 61(1): 1-11 [DOI: 10.1364/JOSA.61.000001]

Lee C, Lee C and Kim C S. 2013. Contrast enhancement based on layered difference representation of 2D histograms. IEEE Transactions on Image Processing, 22(12): 5372-5384 [DOI: 10.1109/TIP.2013.2284059]

Li Z Y, Yang S, Yang H X, Tang X X, Wu F G and Xu F J. 2025. BIAWDiff: enhancing low-light images with bio-inspired attention and wavelet diffusion//Proceedings of 2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Hyderabad, India: IEEE: 1-5 [DOI: 10.1109/icassp49660.2025.10888793]

Lin Y L, Fu Z Q, Wen K R, Ye T, Chen S X, Meng G, et al. 2025. DPLUT: unsupervised low-light image enhancement with lookup tables and diffusion priors//Proceedings of the 39th AAAI Conference on Artificial Intelligence. Philadelphia, USA: AAAI Press: 5316-5324 [DOI: 10.1609/aaai.v39i5.32565]

Liu R S, Ma L, Zhang J A, Fan X and Luo Z X. 2021. Retinex-inspired unrolling with cooperative prior architecture search for low-light image enhancement//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville, USA: IEEE: 10561-10570

Ma L, Ma T Y, Liu R S, Fan X and Luo Z X. 2022. Toward fast, flexible, and robust low-light image enhancement//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans, USA: IEEE: 5627-5636 [DOI: 10.1109/CVPR52688.2022.00555]

Mittal A, Moorthy A K and Bovik A C. 2012. No-reference image quality assessment in the spatial domain. IEEE Transactions on Image Processing, 21(12): 4695-4708 [DOI: 10.1109/TIP.2012.2214050]

Mittal A, Soundararajan R and Bovik A C. 2013. Making a "completely blind" image quality analyzer. IEEE Signal Processing Letters, 20(3): 209-212 [DOI: 10.1109/LSP.2012.2227726]

Pan X Y, Wei M, Wang H and Jia F Z. 2022. A multi-scale fusion residual encoder-decoder approach for low illumination image enhancement. Journal of Computer-Aided Design and Computer Graphics, 34(1): 104-112 (潘晓英, 魏苗, 王昊, 贾丰竹. 2022. 多尺度融合残差编解码器的低照度图像增强方法. 计算机辅助设计与图形学学报, 34(1): 104-112) [DOI: 10.3724/SP.J.1089.2022.18833]

Sun W Q, Peng C Y and Zhang X J. 2026. Low-light image enhancement via a multiscale dilated convolution with coordinate grouping enhancement network. Journal of Image and Graphics, 31(2): 448-464 (孙婉倩, 彭春燕, 张效娟. 2026. 融合多重空洞卷积与坐标分组的暗光图像增强网络. 中国图象图形学报, 31(2): 448-464) [DOI: 10.11834/jig.250177]

Wang Y L, Liu Z, Liu J Z, Xu S C and Liu S C. 2023. Low-light image enhancement with illumination-aware gamma correction and complete image modelling network//Proceedings of 2023 IEEE/CVF International Conference on Computer Vision (ICCV). Paris, France: IEEE: 13082-13091 [DOI: 10.1109/ICCV51070.2023.01207]

Wei C, Wang W, Yang W, Liu J and Guo Y. 2018. Deep Retinex decomposition for low-light enhancement//Proceedings of 2018 British Machine Vision Conference. Newcastle, UK: BMVA Press: #122

Woo S, Park J, Lee J Y and Kweon I S. 2018. CBAM: convolutional block attention module//Proceedings of the 15th European Conference on Computer Vision. Munich, Germany: Springer: 3-19 [DOI: 10.1007/978-3-030-01234-2_1]

Yan Q S, Feng Y X, Zhang C, Pang G S, Shi K B, Wu P, et al. 2025. HVI: a new color space for low-light image enhancement//Proceedings of 2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville, USA: IEEE: 5678-5687 [DOI: 10.1109/CVPR52734.2025.00533]

Yi X, Xu H, Zhang H, Tang L and Ma J. 2023. Diff-Retinex: rethinking low-light image enhancement with a generative diffusion model [EB/OL]. [2025-10-01]. https://arxiv.org/pdf/2308.13164.pdf

Zhang K, Zuo W M, Chen Y J, Meng D Y and Zhang L. 2017. Beyond a Gaussian denoiser: residual learning of deep CNN for image denoising. IEEE Transactions on Image Processing, 26(7): 3142-3155 [DOI: 10.1109/TIP.2017.2662206]

作者简介

包晓安，男，教授，主要研究方向为计算机视觉。
E-mail: baoxiaoan@zstu.edu.cn
涂小妹，通信作者，女，讲师，主要研究方向为图像处理和模式识别。E-mail: txm_95@163.com
陈奕江，男，硕士研究生，主要研究方向为图像处理与机器视觉。E-mail: 1281520471@qq.com
张娜，女，教授，主要研究方向为图像识别。
E-mail: zhangna@zstu.edu.cn
胡天缤，女，博士研究生，主要研究方向为人工智能。
E-mail: 230222010006@hhu.edu.cn
许铭洋，男，硕士研究生，主要研究方向为图像处理与机器视觉。E-mail: 1791595405@qq.com