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

中图法分类号:TP183;TP391　　**文献标识码**:A　　**文章编号**:1006-8961(2026)02-0448-17

论文引用格式: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):0448-0464(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):0448-0464)[DOI:10.11834/jig.250177]

融合多重空洞卷积与坐标分组的暗光图像增强网络

孙婉倩[1,2],彭春燕[1,2,3,4*],张效娟[1,2]

1. 青海师范大学计算机学院,西宁　810008; 2. 藏语智能全国重点实验室,西宁　810008;
3. 藏文信息处理教育部重点实验室,西宁　810008; 4. 青海省藏文信息处理与机器翻译重点实验室,西宁　810008

摘　要:**目的**　在低光环境下拍摄的图像通常存在亮度不均、细节丢失和色彩失真等问题,导致视觉质量大幅下降。大多数低光图像增强方法更多关注亮度提升,忽略了色彩恢复,容易产生色偏和伪影。此外,现有深度学习方法大多结构复杂、训练成本较高,且难以同时兼顾亮度和色彩恢复。**方法**　提出一种融合多重空洞卷积与坐标分组的暗光图像增强网络(low-light image enhancement network via fused multi-scale dilated convolutions and coordinate grouping,MCCNet)。MCCNet由色彩转换和增强网络两个独立分支组成。其中,基于权重感知的HVI(hue, value, intensity)色彩空间中引入色彩自适应因子,将亮度和颜色信息解耦,从而提高色彩增强的精确性。提出全局分组坐标注意力模块(global grouped coordinate attention module,GGCA),促进颜色空间与色彩恢复分支之间的信息交互,并有效抑制低光图像中的噪声。结合多种空洞率卷积层,提出多重空洞融合注意力模块(multi-scale dilated fusion attention module,MDFA),实现更精细化地亮度恢复。最后通过色彩转换和亮度增强两个分支的共同协作,在提升亮度的同时保留了更多图像细节,解决了复杂场景和极端低光场景中色彩失真的难题。**结果**　将本文方法与RetinexNet等15种方法进行比较,在3个公开数据集、2个极端低光数据集和5个无配对数据集上进行评估,相较于次优值,本文算法的峰值信噪比(peak signal-to-noise ratio,PSNR)提高1.57 dB,在LOLv1数据集上的峰值信噪比(PSNR)和结构相似度(structural similarity index measure,SSIM)均优于其他算法。**结论**　本文提出的MCCNet能够有效恢复亮度和色彩,在解决色偏和伪影问题时表现优异,且能够更好地调整光照、抑制噪声,实现高质量视觉增强效果。此外,本文方法充分考虑了计算复杂度与模型规模在实际应用中的需求,模型参数量仅为1.98 M,FLOPs为8.06 G,在实现领先性能的同时,显著降低了模型规模与计算开销,与经典图像增强方法以及近年来性能突出的先进算法进行对比,模型参数和计算量平均减少约96%和46%,实现了高效轻量的设计。

关键词:低照度图像增强;HVI颜色空间;注意力机制;坐标分组;空洞卷积

Low-light image enhancement via a multiscale dilated convolution with coordinate grouping enhancement network

Sun Wanqian[1,2], Peng Chunyan[1,2,3,4*], Zhang Xiaojuan[1,2]

1. School of Computer, Qinghai Normal University, Xining　810008, China; 2. State Key Laboratory of Tibetan Intelligent, Xining　810008, China; 3. Key Laboratory of Tibetan Information Processing, Ministry of Education, Xining　810008, China; 4. Qinghai Provincial Key Laboratory of Tibetan Information Processing and Machine Translation, Xining　810008, China

收稿日期:2025-05-25;**修回日期**:2025-07-15;**预印本日期**:2025-07-23

*通信作者:彭春燕　pcy@qhnu.edu.cn

基金项目:藏语智能全国重点实验室建设项目(2022-2J-J08);国家自然科学基金项目(62563033,62441609,62262056)

Supported by: The State Key Laboratory of Tibetan Intelligence (2025-2J-J108); National Natural Science Foundation of China (62563033, 62441609, 62262056)

Abstract: **Objective** Images captured under low-light conditions often suffer from various visual quality degradation issues, including uneven illumination, significant loss of structural and textural details, and severe color distortions. These issues not only compromise human visual perception but also pose serious challenges for high-level vision tasks such as object detection and recognition. While numerous low-light image enhancement (LLIE) techniques have been developed to address these problems, most existing methods primarily concentrate on improving brightness or contrast. As a result, they often overlook the critical aspect of accurate color restoration, leading to undesired artifacts such as color casts, over-enhancement, or structural inconsistencies. Conventional approaches, such as Retinex-based models, enhance image brightness by decomposing images into reflectance and illumination components. However, this strategy tends to produce color distortions, particularly in extremely low-light scenarios in which the lack of information causes the reflectance component to be inaccurately estimated. Deep learning-based methods have emerged as a promising alternative given their ability to learn complex mappings between low- and normal-light domains. Nonetheless, many of these models are characterized by large network sizes, high computational costs, and suboptimal generalization to diverse illumination conditions, making them minimally suitable for practical deployment on resource-constrained platforms. **Method** To overcome the aforementioned challenges, we propose a novel LLIE framework named multiscale dilated convolution with coordinate grouping enhancement network (MCCNet). MCCNet is designed with a clear emphasis on brightness restoration and color correction, ensuring visually pleasing and natural-looking outputs. The architecture comprises two independently designed yet interactively trained branches: a color conversion branch and a detail enhancement branch. The color conversion branch operates in a hue-value-intensity (HUI) color space, which is more perceptually aligned with human vision than traditional RGB representations. We introduce a color adaptation factor that is dynamically learned on the basis of pixel-wise weight perception to effectively decouple and modulate the color and brightness components of images. This decoupling enables targeted enhancement of color fidelity without introducing additional artifacts or distortion. In the enhancement branch, we design a global grouped coordinate attention (GGCA) module to improve feature expressiveness and encourage effective cross-branch information exchange. GGCA selectively emphasizes meaningful features while suppressing irrelevant or noisy signals, which is especially beneficial for images with high levels of darkness and low signal-to-noise ratios. Furthermore, we propose a multiscale dilated fusion attention (MDFA) module that aggregates features across multiple dilation rates. This module enables the network to capture global context and fine-grained local structure, leading to precise brightness recovery and edge preservation. By fusing multiscale features, MDFA mitigates the typical blurring and oversmoothing problems that occur in conventional convolutional enhancement networks. The two branches of MCCNet work collaboratively to generate high-quality enhanced images. The color conversion branch ensures accurate chromatic consistency, while the enhancement branch restores luminance and structure details. This dual-branch strategy effectively addresses the common trade-off between brightness enhancement and color fidelity that limits many existing LLIE methods. **Result** To thoroughly evaluate the performance of the proposed MCCNet, we conduct comprehensive experiments on multiple public benchmarks, including LOLv1, LOLv2, and five unpaired real-world low-light datasets. In addition, we test the model on two extremely low-light subsets to assess its robustness under severe lighting degradation. MCCNet is compared against 15 state-of-the-art LLIE methods, including traditional models like RetinexNet and zero-reference deep curve estimation(Zero-DCE), as well as recent deep learning approaches such as enlightening generative adversarial networks for low-light image enhancement(EnlightGAN), unsupervised retinex network(URetinex-Net), generative diffusion prior for unified image restoration and enhancement (GDP), lightening diffusion for low-light image enhancement (LightenDiffusion), and CIDNet. Quantitative results demonstrate that MCCNet achieves superior performance in terms of peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM). Specifically, MCCNet improves PSNR by 1.57 dB compared with the best-performing baseline on the LOLv1 dataset. Our method also achieves the highest SSIM score, indicating better structural preservation and perceptual quality. Qualitative comparisons reveal that MCCNet produces images with more natural lighting, reduced noise, and vivid, realistic colors, even in complex scenes with significant shadows or saturation challenges. In terms of computational complexity, MCCNet contains only 1.98 million parameters, and FLOPs are limited to 8.06 G. This dramatic reduction in model size and computational load makes MCCNet highly suitable for deployment on edge devices, mobile platforms, or real-time applications where efficiency is critical. **Conclusion** The proposed color

space-based MCCNet effectively restores brightness and color, addresses color cast and artifacts, and excels at lighting adjustment and noise suppression, achieving high-quality visual enhancement. Moreover, the method fully considers the practical demands of computational complexity and model size. The MCCNet model contains only 1.98 million parameters and requires 8.06 G FLOPs. While achieving state-of-the-art performance, it significantly reduces model size and computational cost. In particular, compared with classic image enhancement methods and recent advanced algorithms, MCCNet reduces model parameters and computational load by approximately 96% and 46% on average, respectively, demonstrating an efficient and lightweight design.

Key words: low-light image enhancement; HVI color space; attention mechanism; coordinate grouping; dilated convolution

0 引 言

图像增强是利用算法改善在暗光条件下拍摄的图像或视频质量的技术，通过提升图像的亮度、对比度、清晰度和色彩还原度，使其更符合人眼视觉或机器分析的需求。该技术广泛应用于摄影、安防监控、自动驾驶和医学成像等领域。近年来，深度学习在低光增强任务中取得显著进展，但仍面临亮度和对比度恢复、细节保留及计算效率等多重挑战(赵明华等，2024)。

具体而言，现有方案在亮度与对比度的自适应精准调节、微弱细节的精准还原等方面仍存在较大优化空间，且计算复杂度偏高的核心问题尚未得到有效解决(程龙昊 等，2025)。Yao等人(2024)提出的SDFNet(spatial-frequency dual-domain feature fusion network)采用空间域和频率域特征融合进行低光遥感图像增强，这种方法在处理高分辨率遥感图像时能够捕捉长距离空间相关性，然而在高空复杂环境下进行特征融合时，会引入频率域的噪声和伪影，且计算复杂度高，资源需求量大。Jiang等人(2021)提出的EnlightenGAN在无配对数据集的情况下利用对抗学习生成更自然的增强结果，但对抗生成网络(generative adversarial network，GAN)在训练时无法很好地抑制图像中的噪声，缺乏对图像细节的恢复。Zamir等人(2022)提出Restormer模型，需要依赖大量的训练数据学习图像的全局特征，当训练数据集不够多样化或包含的场景过于单一时，则无法在未知场景或特定环境下进行有效增强。极端低光条件下有效恢复图像质量较为困难，恢复后的图像色彩偏差较大。Hsu等人(2022)提出一种结合极端低光图像增强与场景文本恢复的框架，该方法在训练过程中依赖于精细标注的文本区域和复杂损失设计，对数据集质量和模型参数调优要求高。

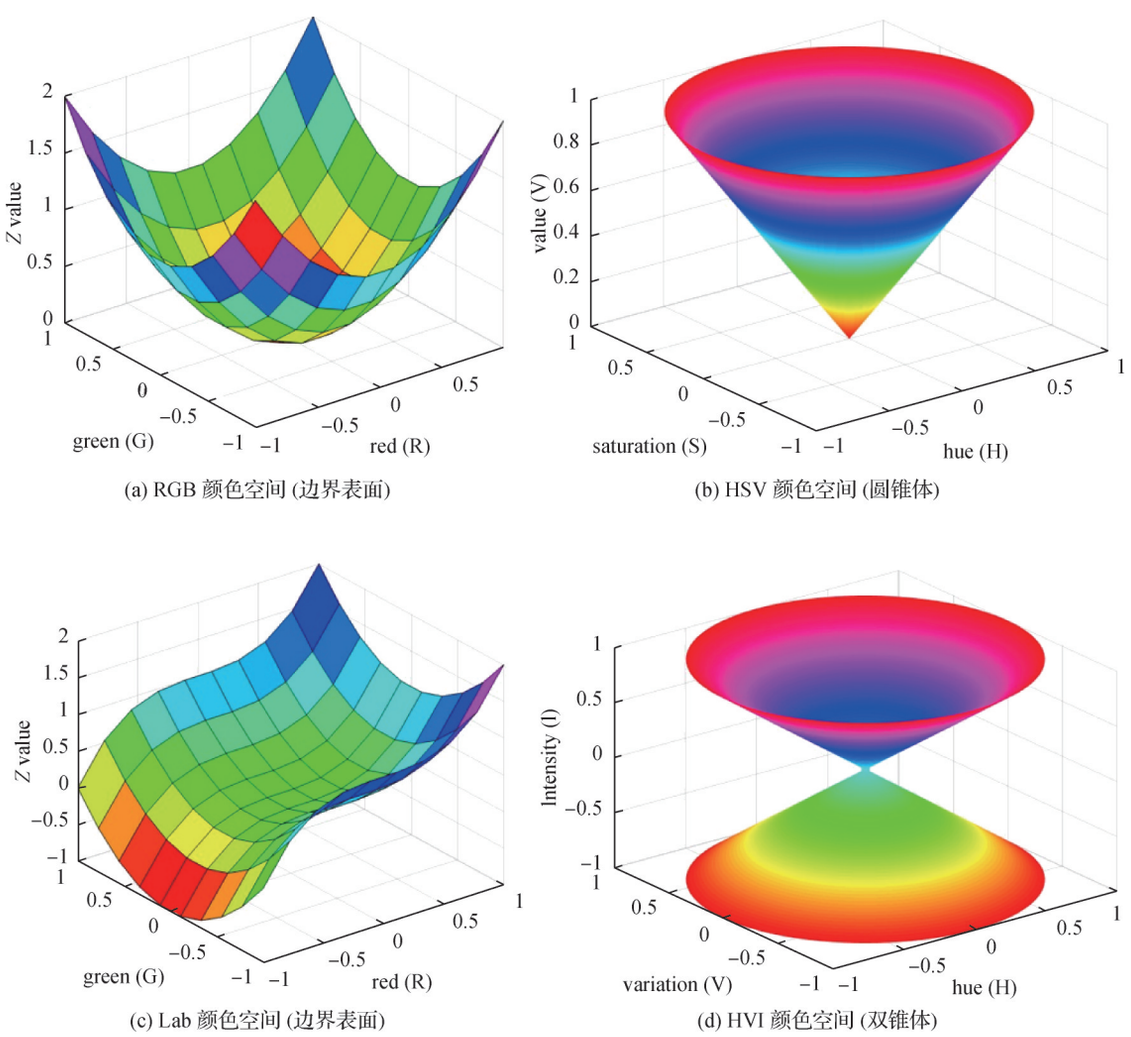
Cai等人(2023)提出的Retinexformer模型通过Transformer架构增强光照估计能力，然而该方法对计算资源要求较高，在处理高分辨率图像时存在推理速度慢的问题。Jiang等人(2024)提出LightenDiffusion，利用扩散模型与Retinex理论相结合，提出无监督的图像增强框架，该方法训练成本较高，对数据分布的依赖性较强，在处理非均匀光照图像时存在噪声过度增强的问题。

在无监督学习方面，马龙等人(2022)在低光照图像增强综述中已指出，无监督方法因无需成对训练数据的优势成为研究热点，但仍面临稳定性、泛化性等核心挑战。Yu等人(2023)开发了一种基于对抗学习的低光增强网络，该网络通过生成对抗网络(GAN)增强图像亮度，同时抑制噪声与色偏问题，由于GAN的训练不稳定性，该方法会引入过度增强或色彩失真的情况，导致某些图像区域出现异常伪影。一种无监督的自蒸馏低光增强模型(陈祖国 等，2024)通过跨尺度特征融合学习亮度调整策略，有效减少细节损失和增强伪影，但该方法在不同数据集上的泛化能力仍需提升，且在某些低光极端情况下无法恢复足够的细节信息，边缘处理效果不佳。

许多研究者对现有的低光增强模型进行了优化。Zhang等人(2021)对KinD(kinematics-aware dehazing)模型进行优化，提出KinD++，引入多尺度照明注意力模块减少视觉伪影，该模型的复杂性较高，推理时间相较于原始KinD有所增加，在移动端或资源受限的设备上运行存在一定困难。此外，结合频域信息，Ren和Miao(2024)提出基于频域特征分解的低光增强方法，该方法在极端暗光场景下会导致过度平滑，使图像细节丢失，影响最终视觉效果。

颜色空间作为用于定量描述和表征颜色属性的数学框架，在低光照成像系统中对图像增强算法的性能具有基础性制约作用（张亚邦 等，2024）。不同颜色空间通过定义独特的色彩分量组合与感知映射关系，直接影响低光图像中色彩传递、噪声分布以及动态范围调整等关键环节的优化效果，在低光环境下，选择合适的颜色空间对于图像增强效果至关重要。图 1 列举了常见的 4 种颜色空间，分别展示了 RGB、HSV、HVI 和 Lab 空间对图像颜色信息的不同表达方式。

RGB 颜色空间对光照变化极为敏感，尤其在低光环境下，容易导致图像亮度的过度增强或色彩失真，He 等人（2024）提出一种基于 RGB 空间的深度学习低光增强方法，在处理极端低光图像时仍存在色彩失真问题，在不同光照条件下，模型的泛化能力明显不同。HSV 颜色空间通过分离色调（hue）、饱和度（saturation）和亮度（value），能够单独调整图像的亮度信息，避免对色彩的直接影响，于成康和韩广良（2025）提出一种改进的 HSV 颜色空间转换方法，在极端低光场景下仍导致局部色彩漂移，影响图像质量（张航和颜佳，2024）。Lab 颜色空间基于人眼视觉的感知特性，具有较好的感知均匀性，然而，Lab 颜色空间的计算较为复杂，并且在极端低光或非均匀光照条件下可能会导致色彩漂移，使得图像的色彩恢复较为困难（Dang 等，2017）。HVI（hue，value，intensity）颜色空间通过将 HSV 空间中的色调和饱和度转化为直角坐标系中的 X、Y 分量，并结合亮度信息计算颜色敏感度，Wang 等人（2025）研究了一种结合 HVI 颜色空间与注意力机制的低光增强算法，但该方法在极端低光条件下，仍存在细节过度平滑的问题，影响图像清晰度。



(a) RGB 颜色空间 (边界表面)　(b) HSV 颜色空间 (圆锥体)

(c) Lab 颜色空间 (边界表面)　(d) HVI 颜色空间 (双锥体)

图 1　颜色空间对比图

Fig. 1　Color space comparison diagram ((a) RGB color space(surface); (b) HSV color space(cylinder); (c) Lab color space(surface); (d) HVI color space(double cone))

1 网络模型

为突破现有复杂场景及极端低光场景图像增强方法在亮度与色彩恢复方面存在的局限，解决特征表达能力弱、细节建模不足以及亮度与色彩耦合干扰等问题，本文提出一种基于颜色空间的多重空洞卷积与坐标分组的暗光图像增强网络（low-light image enhancement network via fused multi-scale dilated convolutions and coordinate grouping，MCCNet），旨在有效提升低光图像中亮度与色彩的建模与恢复能力。整体网络由颜色空间转换分支与增强网络分支两部分构成，分别对亮度与色彩信息进行分离建模与协同增强。

在颜色空间转换分支中，本文提出颜色空间解耦机制，将输入图像同时映射至RGB与HVI两种具有互补特性的颜色空间，有效分离亮度与色彩信息，以避免建模过程中的相互干扰。同时引入了色彩自适应调整因子（color adaptation factor），根据图像局部亮度信息动态调节颜色通道的权重，实现更加自然且均衡的色彩增强。在低光区域中，色彩自适应调整因子自动提升通道权重，增强颜色饱和度并改善色彩保真度；在高光区域，适当降低通道权重，避免过度增强引发颜色失真，从而保持自然的色彩分布。相较于传统采用静态映射策略的做法，本文提出的动态调节机制具备更强的适应性与调控能力，显著提升增强图像在不同亮度区域中的色彩表现与视觉一致性。

为了进一步提升跨通道协同建模能力与空间一致性，在颜色空间转换分支中，本文设计了全局分组坐标注意力模块（global grouped coordinate attention module，GGCA）。其主要创新为：引入坐标嵌入机制，显式编码图像的空间位置信息为可学习特征，增强模型对结构连续性和几何稳定性的建模能力；设计跨通道分组策略，将颜色通道进行语义分组处理，提升模型对通道间色彩语义关联及局部区域特征的感知能力；融合全局上下文建模机制，实现对图像整体结构、光照趋势与色彩一致性的多层次建模。与现有注意力机制相比，GGCA是首次面向低光图像增强中亮度与色彩分支协同建模需求所设计，显著提升了多分支间信息交互效率与建模表现，增强了复杂场景下的结构对齐能力与语义一致性。

在增强网络分支中，针对传统卷积及传统空洞卷积模块在多尺度感受野整合与细节保持方面的不足，本文提出多重空洞融合注意力模块（multi-scale dilated fusion attention module，MDFA）。传统空洞卷积虽然扩展了感受野，但通常缺乏对多尺度特征的充分融合与细节关注，且注意力机制应用有限，导致对极低照区域的细节恢复效果欠佳。为此，MDFA通过构建多分支结构，分别采用不同空洞率（小、中、大）实现多尺度上下文信息的捕获，突破了单一尺度感受野的限制，增强了模型对不同尺度特征的适应能力；同时，在每个分支中并行引入通道注意力机制与空间注意力机制，协同提升特征表达的选择性与细节保持能力；该设计不仅有效整合了全局语义信息，还特别强化了极暗区域的细节恢复，显著改善了传统空洞卷积在极低光条件下图像增强效果的不足，体现了本文方法在多尺度建模与细节增强方面的创新优势。

最终，颜色空间转换分支输出的色彩增强特征与增强网络分支的亮度增强特征在网络尾部进行融合，生成高质量、自然感强的增强图像。整体流程如图2所示。

1.1 基于权重感知的HVI颜色空间

HVI色彩空间的核心目标是对输入图像进行色彩空间转换，并在转换后的空间中进行亮度增强和色彩恢复。转换过程首先将图像的颜色信息拆分为色调H、亮度V和照明分量I这3个独立分量，使亮度增强和色彩调整能够分别进行。这种分离方式有助于在增强亮度的同时保持色彩的相对稳定性，避免整体调整时可能引入的色偏问题，并且能够更有效地模拟人眼对亮度和色彩的感知特性，使增强后的图像视觉效果更加自然、真实，能够适用于多种复杂光照环境下的图像处理任务。

然而，传统HVI变换存在以下问题，固定的色彩敏感度无法适应不同光照条件，在低光环境下增强色彩的需求较大，而在高光环境下，如果增强幅度过大，容易导致色彩失真，无法针对不同区域的光照水平自适应调整，导致色彩增强不均衡。尽管HVI变换在理论上能够减少色偏，但由于色调H与亮度V之间的关系没有得到有效控制，在某些极端低光下，仍然可能出现颜色不均或不自然的问题。

为了解决上述问题，本文设计了一种自适应的HVI颜色空间变换方法，使色彩增强能够针对不同

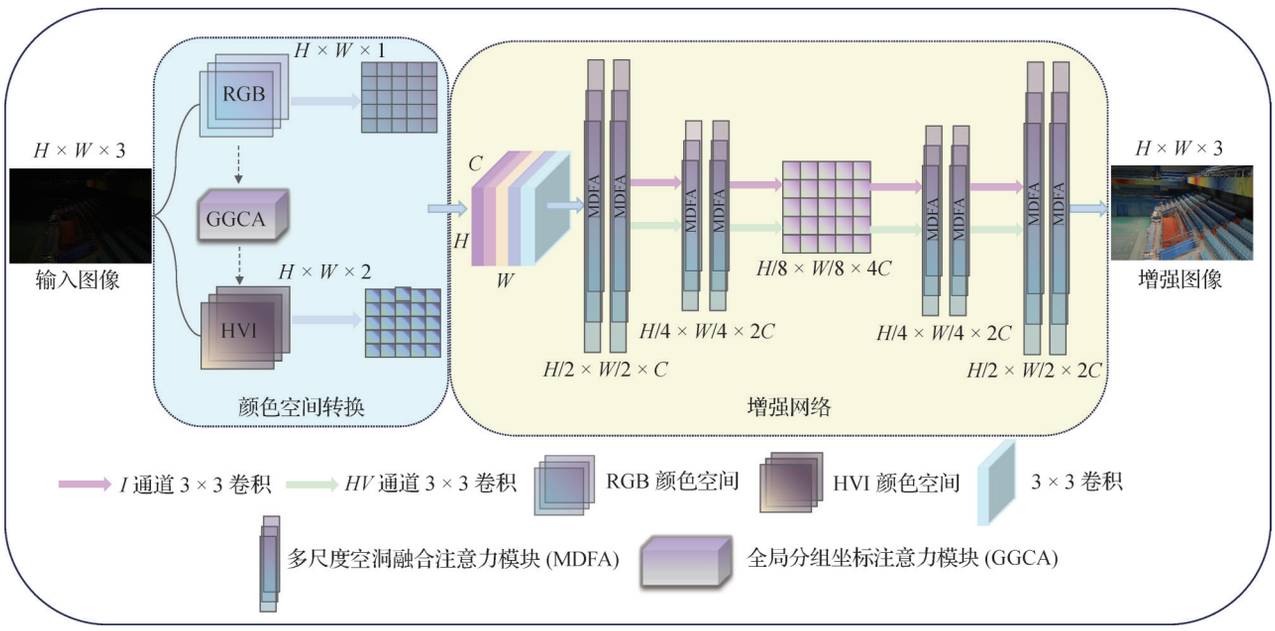


图 2　MCCNet 网络整体架构图

Fig. 2　MCCNet network overall architecture diagram

光照条件进行智能调整，从而提高低光区域的色彩保真度，避免高光区域的过度增强。基于视觉感知的动态权重调整策略优化 HVI 颜色空间变换，提出色彩自适应因子（color adaptation factor）来进行自适应调节，从而改善色彩增强的效果。在低光区域，提高色彩自适应因子值，使色彩增强更强，从而提升色彩保真度，减少低光环境下的色彩损失。另外，在高光区域，可以降低色彩自适应因子值，减少色彩增强幅度，防止过度增强导致的颜色失真，保持自然的色彩分布。该策略能够根据图像的局部亮度动态调整，使得整个图像的色彩增强更加平衡，避免过度增强或不足，提升图像整体视觉效果。

首先，计算每个像素点的最大值（$\max(R, G, B)$）和最小值（$\min(R, G, B)$）；然后，利用这些值计算该像素的色调。通过色调，可以调整图像的颜色分布，使图像在视觉上更具层次感和对比度。色调 H 计算为

$$H = \frac{1}{6} \times (\max(R, G, B) - \min(R, G, B)) \tag{1}$$

式中，R、G、B 分别表示红色、绿色和蓝色通道的值。$\max(R, G, B)$ 是像素的最大值，表示该像素的最亮颜色。$\min(R, G, B)$ 是像素的最小值，表示该像素的最暗颜色。

饱和度 S 衡量颜色的鲜艳程度，计算为

$$S = \frac{\max(R, G, B) - \min(R, G, B)}{\max(R, G, B)} \tag{2}$$

式中，S 计算了色彩的纯度，值越高代表颜色越鲜艳，值越低则代表颜色越接近灰色。通过调整饱和度，可以进一步增强或柔化图像的色彩表现。

在 HVI 变换中，利用色彩自适应因子 f_{cadp} 调节色调和亮度的映射，色调 H 会根据图像的亮度动态调整，使低光区域的色彩增强更强，增强色彩的真实感，而高光区域的色彩增强幅度适当减小，避免色彩失真。特别是在不同光照条件下，使图像的色彩表现更加自然。HVI 色彩变换计算为

$$H' = H + k \times f_{\text{cadp}} \tag{3}$$

式中，H 是计算得到的原始色调。H' 是经过自适应调整后的色调，k 为可训练参数，用于调节色彩自适应因子，根据不同光照条件灵活调整色彩增强的强度，以便让图像在不同环境下均能保持自然色彩。

HVI 空间中的最终颜色由色调 H、饱和度 S 和亮度 V 的组合决定。进一步优化 H 和 V 的关系，HVI 变换能够更好地控制图像的亮度和色彩平衡。空间转换式为

$$\begin{bmatrix} \boldsymbol{X} \\ \boldsymbol{Y} \\ \boldsymbol{Z} \end{bmatrix} = f(H, S, V) \tag{4}$$

式中，$\boldsymbol{X}$、$\boldsymbol{Y}$ 和 $\boldsymbol{Z}$ 是 HVI 空间中的分量，代表了图像的颜色信息。f 是一个转换函数，考虑了 H 和 V 的优化，并结合色彩自适应因子调整亮度和色彩之间的平衡。

在图像处理完成后，需要将HVI空间中的数据反向转换为RGB格式。反向转换计算色调H和饱和度S后，恢复到RGB图像格式，确保增强后的图像能够以正确的色彩和亮度呈现。HVI反向转换为RGB格式计算为

$$(R, G, B) = inverse\_HVI(H', S, V) \tag{5}$$

式中，$inverse_HVI$是将调整后的色调H'和饱和度S反向转换为RGB。

最后，在HVI变换的输出层，使用1×1卷积层进行降维，减少通道数，提高计算效率，同时调整图像的最终色彩输出，计算为

$$\boldsymbol{X} = Conv_{1\times 1}(HVI) \tag{6}$$

式中，$\boldsymbol{X}$为最终输出。

1.2 全局分组坐标注意力

全局分组坐标注意力（GGCA）是一种轻量级注意力机制，通过通道分组和全局坐标注意力增强特征表示能力。结合全局池化和共享特征变换，用于提取空间信息，并在高度（H）方向和宽度（W）方向计算独立的注意力权重。图3是GGCA模块的结构图。

GGCA主要由以下几个部分组成，分组机制将输入特征图按通道数分为多个组，使不同组的特征可以独立提取空间注意力信息；全局池化模块在高度H方向和宽度W方向分别进行全局平均池化和全局最大池化，获取全局空间信息；共享特征变换模块采用两个1×1卷积层进行特征降维和恢复，增强特征表示能力；最后通过sigmoid激活函数计算H方向和W方向的注意力权重，计算得到的H方向和W方向的注意力权重作用于输入特征图，实现自适应增强。

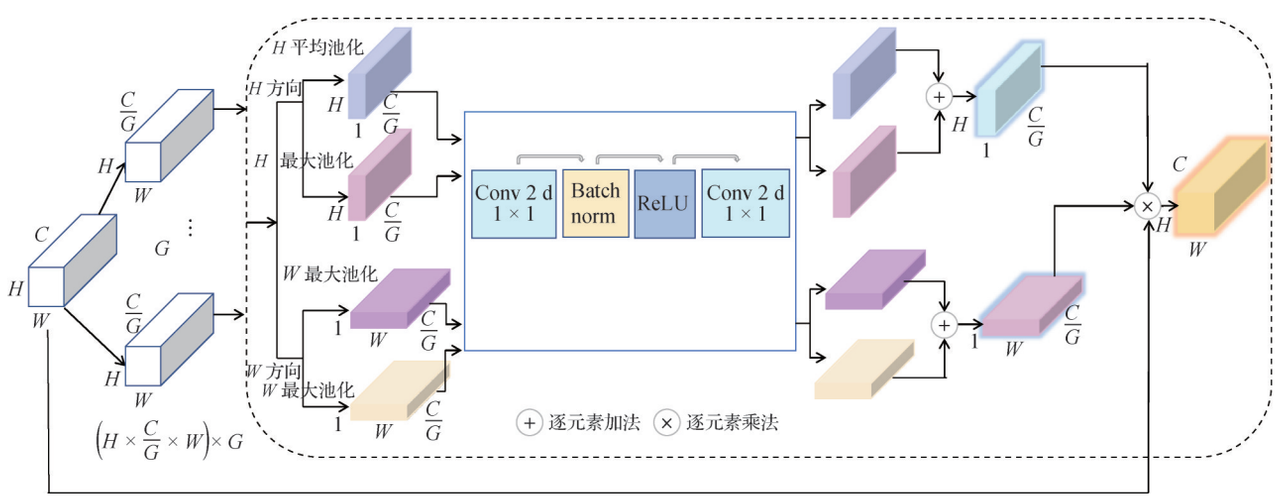


图3 全局分组坐标注意力模块

Fig. 3 Global grouped coordinate attention module (GGCA)

GGCA作用于经过HVI变换的输入特征图$\boldsymbol{X}$，其形状为

$$\boldsymbol{X} \in \mathbf{R}^{B\times C\times H\times W} \tag{7}$$

式中，B是批次大小，C是通道数，H，W分别是高度和宽度。

GGCA采用通道分组机制，将输入特征图$\boldsymbol{X}$按通道数C进行分组，每组包含C/G个通道，计算为

$$C_g = \frac{C}{G} \tag{8}$$

式中，G是通道分组数，C_g是每组的通道数，可以减少计算量，让不同组的特征独立地提取空间注意力信息，同时降低计算量，避免直接在整个通道维度上计算注意力带来的高计算成本。

将输入的特征图$\boldsymbol{X}$按通道维度进行分组，表示为

$$\boldsymbol{X}_g \in \mathbf{R}^{B\times G\times C_g\times H\times W} \tag{9}$$

式中，$\boldsymbol{X}_g$是分组后的特征图，表示第g组的特征。

对第g组的特征$\boldsymbol{X}$在高度H方向和W方向进行全局平均池化和最大池化，计算为

$$\begin{cases} \boldsymbol{X}_d^{\text{avg}} = \dfrac{1}{D}\displaystyle\sum_{i=1}^{D}\boldsymbol{X}_g(:,:,k,:) \\ \boldsymbol{X}_d^{\max} = \displaystyle\max_{k=1}^{D}\boldsymbol{X}_g(:,:,k,:) \\ \boldsymbol{X}_d^{\text{avg}}, \boldsymbol{X}_d^{\max} \in \mathbf{R}^{B\times G\times C_g\times H\times 1} \end{cases} \tag{10}$$

式中，当$D = W$时，对应高度H方向的池化，当$D = H$时，对应宽度W方向的池化，提取全局信息。$\boldsymbol{X}_d^{\max}$和

$\boldsymbol{X}_d^{\text{avg}}$表示全局最大池化和全局平均池化。使模型能够感知全局亮度趋势并且防止增强后高亮区域过曝。将全局池化后的输出通过两个 1 × 1 卷积层进行特征变换，计算为

$$\boldsymbol{Y} = \boldsymbol{W}_2\sigma(\boldsymbol{W}_1(\boldsymbol{X}_D^{\text{avg}} + \boldsymbol{X}_D^{\max})) \tag{11}$$

式中，$\boldsymbol{W}_1$和$\boldsymbol{W}_2$分别是第 1 个卷积层和第 2 个卷积层的权重矩阵，分别用来进行通道降维和恢复通道维度。$\sigma(\cdot)$代表 ReLU（rectified linear unit）激活函数。$\boldsymbol{Y}$是最终的注意力特征结果。

对$\boldsymbol{Y}$特征进行加法融合，并通 sigmoid 获取归一化权重，计算为

$$\begin{cases} A_h = \sigma(\boldsymbol{Y}_h^{\text{avg}} + \boldsymbol{Y}_h^{\max}) \\ A_w = \sigma(\boldsymbol{Y}_w^{\text{avg}} + \boldsymbol{Y}_w^{\max}) \end{cases} \tag{12}$$

式中，A_h是H方向的注意力权重，A_w是W方向的注意力权重，$\sigma(\cdot)$是 sigmoid 激活函数，使权重归一化到[0,1]，避免过度增强或失真，使低光区域的特征得到加强的同时，而不过度改变亮度均匀的区域。

最后，将H方向和W方向的注意力权重分别作用于输入特征图$\boldsymbol{X}$，并且恢复原始输入的形状$\boldsymbol{X}'$，计算为

$$\boldsymbol{X}' = \boldsymbol{X}_g \times A_h \times A_w,\ \boldsymbol{X}' \in \mathbf{R}^{B\times C\times H\times W} \tag{13}$$

式中，$\boldsymbol{X}'$是应用注意力后的特征图，$\boldsymbol{X}_g$是分组后的输入特征图，A_h是H方向的注意力权重，A_w是W方向的注意力权重。

1.3　多重空洞卷积

本文方法中的多重空洞融合注意力模块（MDFA）的主要目标是捕捉不同尺度的图像特征，并利用注意力机制增强重要区域的表达能力，不仅需要恢复亮度细节，还需要去除噪声，MDFA 模块通过多个分支和不同尺度的空洞卷积来达到这一目标。

核心组成部分是多个卷积分支，第 1 分支采用 1 × 1 卷积操作，未设置空洞率，保持输入图像的通道维度不变，主要用于提取局部细节特征，不改变感受野大小，也能有效处理细节信息；第 2 分支采用 3 × 3 卷积，空洞率为 9，增加感受野的大小，以便捕捉到中等尺度的上下文信息，获取更多上下文信息，同时保留局部细节；第 3 分支采用 3 × 3 卷积，空洞率为 18，进一步扩大感受野，提取较大尺度的图像特征，捕捉到更广泛的信息；第 4 分支采用 3 × 3 卷积，空洞率为 27，最大化感受野，用来提取图像的全局信息，帮助恢复全局亮度和色彩。此外，还包含一个全局特征提取分支，通过全局平均池化和 1 × 1 卷积提取图像的全局信息，有助于模型理解整个图像的亮度和色彩分布，并与其他分支的特征图进行融合。图 4 是 MDFA 模块的结构图。

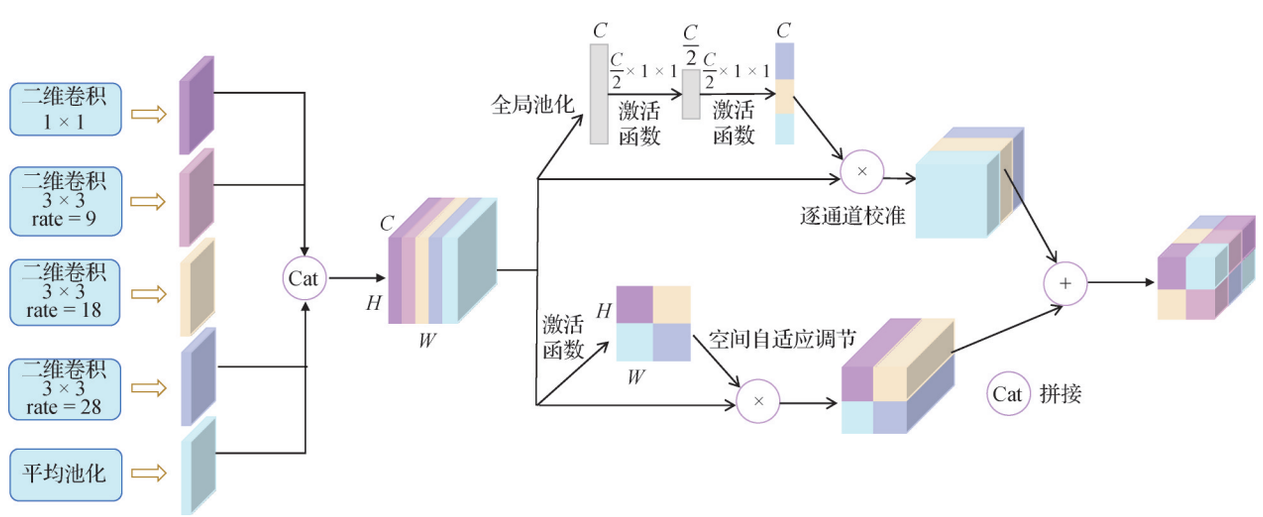


图 4　多重空洞融合注意力模块

Fig. 4　Multi-scale dilated fusion attention module（MDFA）

通过 GGCA 坐标分组模块后的特征图$\boldsymbol{X}'$作为 MDFA 模块的输入特征$\boldsymbol{X}' \in \mathbf{R}^{B\times C\times H\times W}$。MDFA 采用的 5 个分支中，4 个分支是空洞卷积，1 个是全局特征提取，计算为

$$\boldsymbol{X}_i = \sigma(W_i * \boldsymbol{X}' + b_i),\ i \in \{1,2,3,4\} \tag{14}$$

式中，W_1是 1 × 1 卷积核（无空洞），W_2是 3 × 3 卷积核，空洞率为 9，W_3是 3 × 3 卷积核，空洞率为 18，W_4是 3 × 3 卷积核，空洞率为 27，b_i是对应的偏置，* 表

示空洞卷积操作(当$i=1$时为普通卷积),$\sigma(\cdot)$是ReLU激活函数,计算式为

$$\boldsymbol{X}_5=\sigma(W_5*AvgPool(\boldsymbol{X}')+b_5) \tag{15}$$

式中,$AvgPool(\boldsymbol{X}')$计算全局均值,W_5是全局平均池化,通过1×1卷积提取图像的全局信息,$\boldsymbol{X}_5$经过上采样,恢复到原始大小。再合并所有通道,然后通过通道注意力计算,计算为

$$\begin{aligned}&\boldsymbol{X}_{\text{cat}}=[\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{X}_3,\boldsymbol{X}_4,\boldsymbol{X}_5]\in\mathbf{R}^{B\times 5C\times H\times W}\\&\boldsymbol{X}_{\text{channel}}=ReLU(W_c*AvgPool(\boldsymbol{X}_{\text{cat}}))\\&\boldsymbol{X}_{\text{spatial}}=\sigma(W_s*\boldsymbol{X}_{\text{cat}})\end{aligned} \tag{16}$$

式中,$\boldsymbol{X}_{\text{cat}}$,$\boldsymbol{X}_{\text{channel}}$,$\boldsymbol{X}_{\text{spatial}}$分别是合并后的通道进行通道拼接、通道注意力机制和空间注意力机制的输出,最终将注意力融合,计算为

$$\begin{aligned}&\boldsymbol{A}=\max(\boldsymbol{X}_{\text{channel}},\boldsymbol{X}_{\text{spatial}})\\&\boldsymbol{X}_{\text{weighted}}=\boldsymbol{A}\odot\boldsymbol{X}_{\text{cat}}\end{aligned} \tag{17}$$

式中,max(·)取两个注意力的最大值,⊙表示逐元素乘法。

最终,再使用1×1卷积进行降维,计算为

$$\boldsymbol{Y}=\sigma(W_{\text{out}}*\boldsymbol{X}_{\text{weighted}}+b_{\text{out}}) \tag{18}$$

式中,W_{out}是1×1卷积核,用于降维,b_{out}是可学习的偏置。

2 实 验

2.1 数据集设置

1)LOL(low-light dataset)数据集。包括两个版本:LOL-v1和LOL-v2。其中,LOL-v2进一步分为真实子集(real)和合成子集(syn)。训练集和测试集的划分比例分别为485∶15(LOL-v1)、689∶100(LOL-v2-real)和900∶100(LOL-v2-syn)。在训练过程中,LOL-v1和LOL-v2-real数据集将训练图像裁剪为400×400像素的小块,并采用batchsize=8,训练1 500轮。LOL-v2-syn数据集不进行图像裁剪,batchsize=1,训练500轮。2)SICE数据集。原始SICE(Singapore image contrast enhancement)数据集总共包含589组低光和过曝图像,训练集、验证集和测试集按照7∶1∶2的比例划分。本文在SICE训练集上进行训练,并在SICE-Mix和SICE-Grad数据集上进行测试。将原始SICE图像裁剪为160×160像素大小,并在此基础上进行了1 000轮训练。

2.2 不同增强方法评估

为了定量地评估15种方法和本文方法的增强效果,本文采用峰值信噪比(peak signal-to-noise ratio,PSNR)、结构相似性指数(structural similarity index measure,SSIM)、感知图像质量评估(learned perceptual image patch similarity,LPIPS)、无参考图像质量评估(natural image quality evaluator,NIQE)4项指标进行测试。

表1展示了LOLv1数据集上,不同方法在PSNR、SSIM、LPIPS和NIQE指标下的定量比较结果。参与对比的模型包括RetinexNet、DRBN、RUAS、RetinexDIP、URetinex-Net、SCRnet、Kind++等15种算法。由表1可以看出,本文提出的MCCNet在PSNR、SSIM、LPIPS评价指标上取得了最优性能。MCCNet的PSNR达到了24.57 dB,相比经典方法RetinexNet提升了46.5%,相较于近期的强基线模型SCNet也提升了6.1%。在SSIM方面,MCCNet的结果达到0.86,较RetinexNet提升了53.6%。在LPIPS指标上,MCCNet取得0.09的最低值,较LightenDiffusion降低了30.8%,说明该方法生成的增强图像在感知上更贴近真实图像。在NIQE评价指标上,MCCNet结果比NNPriors方法略差,但NNPriors在PSNR和SSIM等关键指标上表现差异较大,这表明其更低的NIQE值源于图像过度平滑和细节丢失,未能有效提升图像整体质量。相比之下,MCCNet在实现高PSNR和SSIM的同时,仍保持了较低的NIQE值,在增强质量与图像自然度之间实现了良好平衡,更加契合于实际应用中的主观视觉期望。

为验证本文MCCNet在低光照图像增强任务中的有效性,在LOLv2-real与LOLv2-syn两个子数据集上进行了定量性能评估,并与近年来多种代表性方法进行了对比。实验结果如表2所示。从表2中可以看出,本文方法在两个子数据集上均取得了最佳性能。在真实拍摄的LOLv2-real数据集上,MCCNet显著优于Restormer与LLFormer等方法。与Restormer的PSNR相比,提升了3.98 dB,增幅达21.3%;同时SSIM值较LLFormer提升了0.07,增幅为8.9%。这说明MCCNet在亮度恢复和结构信息保持方面表现更加稳定,尤其适用于真实场景下的图像增强任务。

在合成低光场景LOLv2-syn数据集上,MCCNet同样展现出显著优势。与QuadPrior方法相比,PSNR提升了9.35 dB,增幅高达58.0%;SSIM提升了0.18,增幅为23.7%。这一结果表明,MCCNet能

表 1　LOLv1 数据集定量比较结果

Table 1　Quantitative comparisons on the LOLv1 dataset

方法	LOLv1				复杂度	
	PSNR/dB ↑	SSIM ↑	LPIPS ↓	NIQE ↓	Params/M	Flops/G
RetinexNet(Wei 等,2018)	16.77	0.56	0.47	8.87	0.84	148.54
DRBN(Yang 等,2020)	15.32	0.71	0.39	4.96	0.58	42.41
ZeroDCE(Guo 等,2020)	16.79	0.64	0.36	7.77	**0.08**	5.21
Kind++(Zhang 等,2021)	21.31	0.82	0.16	3.88	8.54	36.57
RUAS(Liu 等,2021)	18.23	0.72	0.35	6.34	0.36	233.09
EnlightenGAN(Jiang 等,2021)	17.48	0.65	0.32	4.86	8.37	72.61
RetinexDIP(Zhao 等,2022)	9.65	0.38	0.44	7.07	18.22	**0.79**
URetinex-Net(Wu 等,2022)	21.33	0.83	1.22	5.45	0.36	233.09
NNPriors(Liang 等,2022a)	15.49	0.65	0.38	**3.73**	2.19	64.37
Bread-ME(Pan 等,2023)	22.96	0.83	0.18	3.94	0.84	6.17
GDP(Fei 等,2023)	15.90	0.54	0.43	8.49	0.37	536.40
CLIP-LIT(Liang 等,2023)	12.39	0.49	0.39	6.73	48.72	7.36
LightenDiffusion(Jiang 等,2024)	20.45	0.81	0.13	4.13	64.96	49.58
SCNet(Chen 等,2025)	23.16	0.84	0.21	7.32	0.37	24.22
CIDNet(Yan 等,2025)	23.50	0.85	0.12	5.32	1.88	7.57
MCCNet(本文)	**24.57**	**0.86**	**0.09**	4.83	1.98	8.06

注:加粗字体表示各列最优结果。“↑”和“↓”分别表示值越高越好,值越低越好。

表 2　LOLv2-real 和 LOLv2-syn 数据集上的定量比较结果

Table 2　Quantitative comparison results on LOLv2-real and LOLv2-syn datasets

方法	LOLv2-real		LOLv2-syn		复杂度	
	PSNR/dB ↑	SSIM ↑	PSNR/dB ↑	SSIM ↑	Params/M	Flops/G
RetinexNet(Wei 等,2018)	16.10	0.40	17.14	0.76	0.84	584.47
LEDNet(Zhou 等,2022)	19.94	0.83	23.71	0.91	6.34	35.90
DRBN(Yang 等,2020)	20.29	0.83	23.22	0.93	5.47	48.61
ZeroDCE(Guo 等,2020)	16.06	0.58	17.71	0.82	0.08	4.83
Kind++(Zhang 等,2021)	14.74	0.64	13.29	0.58	8.02	34.99
RUAS(Liu 等,2021)	15.33	0.49	13.77	0.64	**0.03**	**0.83**
EnlightenGAN(Jiang 等,2021)	18.23	0.62	16.57	0.73	114.35	61.01
Restormer(Zamir 等,2022)	18.69	0.84	21.41	0.83	26.13	144.25
LLFormer(Wang 等,2023)	20.06	0.79	24.04	0.91	24.55	22.52
QuadPrior(Wang 等,2024b)	20.48	0.81	16.11	0.76	6.82	6.80
HAIR(Cao 等,2024)	21.44	0.84	24.71	0.91	29.45	158.62
NalAper(Tang 等,2025)	21.12	0.84	24.48	0.93	26.78	59.43
MCCNet(本文)	**22.67**	**0.86**	**25.46**	**0.94**	1.98	8.06

注:加粗字体表示各列最优结果。“↑”表示值越高越好。

够在极端低光条件下有效恢复图像细节并保持结构一致性，提升图像整体感知质量。

除了性能上的提升，MCCNet在模型参数量和计算复杂度方面也实现了良好的平衡。在近3年方法中，MCCNet的参数量仅为1.98 M，显示出其在保证高性能的同时具备良好的轻量化特性。尽管没有达到最优值，但相比性能提升幅度，展现了其优良的效率与性能比。

为进一步直观展示不同方法在低照图像增强任务中的性能差异，图5展示了在LOLv1数据集上多组典型场景的视觉对比结果。图5(a)为低照度图像，原始输入图像存在严重的曝光不足。图5(b)是CLIP-LIT的结果，增强结果出现明显的色彩偏差，导致视觉一致性差。图5(c)是LightenDiffusion的结果，该方法虽然增强了整体亮度，但伪影问题较为明显，尤其在边缘区域，出现了模糊或虚假的纹理。图5(d)是GDP的结果，在阴影区域处理不平滑，细节恢复能力较弱，图像灰暗、缺乏鲜明的层次感。图5(e)是RetinexDIP的结果，整体图像较为昏暗，色彩偏暗且缺乏鲜艳感，未能有效还原真实场景的色彩和明亮度。图5(f)是本文方法(MCCNet)的结果，在亮度提升、细节还原、色彩恢复和视觉自然性方面均表现出色，图像的色调真实自然，边缘清晰且无明显伪影，增强后的视觉效果与高质量图像接近，体现了优越的增强能力。图5(g)是真值图像，为高曝光参考图像。通过上述对比，可以看出，本文方法在图像增强效果、色彩一致性和视觉自然性方面具有明显优势，能够更好地恢复低照图像的细节并减少伪影。

(a) 低照度图像　(b) CLIP-LIT　(c) LightenDiffusion　(d) GDP　(e) RetinexDIP　(f) MCCNet (本文)　(g) 真值

图5　5种方法在LOLv1数据集上的视觉对比

Fig. 5　Visual comparison of five methods on the LOLv1 clataset ((a) low-light images; (b) CLIP-LIT; (c) lightenDiffusion; (d) GDP; (e) RetinexDIP; (f) MCCNet(ours); (g) ground truth)

为验证本文方法在极端低光照图像增强任务中的有效性，本文在SICE_Mix与SICE_Grad两个挑战性极高的数据集上进行了定量性能评估，并与近年来多种具有代表性的主流方法进行了对比。实验结果如表3所示。在SICE_Mix数据集上，本文方法较表现较好的LLFlow，PSNR增幅为1.8%，LPIPS降幅为12.5%；SSIM较LEDNet增幅为8.6%；为所有方法中最低值，表明在图像结构还原与感知质量方面具备更优表现。

在SICE_Grad数据集上，本文方法的PSNR为12.96 dB，与SG-LLIE方法相比，增幅为4.0%；SSIM较SCNet增幅为20.8%；LPIPS为0.37，虽然略高于ZeroDCE的0.33，但SSIM和RSNR表现良好，进一步验证了其在亮度恢复同时有效抑制感知失真的能力。本文方法在SICE_Mix与SICE_Grad两个极端低光数据集上均表现出卓越的图像增强能力，不仅在亮度恢复、结构还原方面表现突出，同时兼顾良好的感知一致性，验证了其在极端低照度环境下的稳定性，

表 3　在两个极端低光数据集(SICE-Mix 和 SICE-Grad)上的定量比较

Table 3　Quantitative comparison on two extremely low-light datasets(SICE-Mix and SICE-Grad)

方法	SICE_Mix			SICE_Grad		
	PSNR/dB ↑	SSIM ↑	LPIPS ↓	PSNR/dB ↑	SSIM ↑	LPIPS ↓
RetinexNet(Wei 等,2018)	12.39	0.60	0.40	12.45	0.62	0.36
ZeroDCE(Guo 等,2020)	12.42	0.63	0.38	12.47	0.64	**0.33**
RUAS(Liu 等,2021)	8.68	0.49	0.52	8.62	0.49	0.50
URetinexNet(Wu 等,2022)	10.90	0.60	0.40	10.89	0.61	0.36
SGZ(Zheng 和 Gupta,2022)	10.87	0.61	0.42	10.99	0.62	0.36
LLFlow (Wang 等,2022)	12.74	0.61	0.39	12.74	0.62	0.39
SCI(Ma 等,2022)	8.64	0.53	0.51	8.56	0.53	0.48
LEDNet(Zhou 等,2022)	12.67	0.58	0.41	12.55	0.58	0.38
CLIP-LIT(Liang 等,2023)	12.47	0.62	0.39	12.47	0.62	0.39
SCNet(Chen 等,2025)	11.15	0.51	0.40	11.24	0.53	0.36
HAIR(Cao 等,2024)	10.94	0.40	0.48	11.03	0.40	0.42
SG-LLIE(Dong 等,2025)	11.87	0.62	0.41	12.46	0.64	0.37
本文	**12.97**	**0.63**	**0.35**	**12.96**	**0.64**	0.37

注:加粗字体表示各列最优结果。“↑”和“↓”分别表示值越高越好,值越低越好。

在当前低光图像增强研究中具有明显的性能优势。

2.3　无配对数据集实验

为进一步验证所提出方法在非配对条件下的泛化能力,在 5 个广泛使用的无配对低光图像增强数据集 DICM (digital imaging and communications in medicine)、LIME(low-light image enhancement)、MEF (multi-exposure fusion)、NPE (nighttime photograph enhancement) 和 VV (VV dataset low-light image enhancement)上进行评估,使用 BRI(blind/reference-less image spatial guality evaluator)和 NIQE 两种无参考图像质量评估指标,结果如表 4 所示。所有实验均基于在 LOLv1 上训练的模型。在 LIME 数据集上,相较于 PSENet (progressive self-enhancement network),BRI 降低了 47.7%,与 GSAD(global structure-aware diffusion process)相比,BRI 和 NIQE 分别降低了 42.6% 与 40.5%,展现了优异的结构保留与自然感恢复能力。在 MEF 数据集上,相较于 ZeroDCE (zero-reference deep curve estimation),本文方法 BRI 和 NIQE 分别提升 19.1% 与 30.2%;相较于 SCI(self-correction image enhancement network),在 BRI 指标降幅高达 50.3%,充分验证了本文在极端光照条件下的稳定增强能力。在 DICM 数据集上,相较当前性能最接近的 GSAD 降低了 16.1%,显示出更佳的纹理结构还原和去伪影能力。在 NPE 数据集上,较主流方法如 RUAS 与 SNR-Aware 的 BRI 值,分别降低了 64.4% 和 36.1%;在 VV 数据集上,相较于 QuadPrior 的 NIQE 值降低了 27.4%。本文方法在无配对增强任务中表现出一致性的跨数据集高性能,整体在 BRI 和 NIQE 指标上的平均提升幅度达 20%～60%,验证了本文方法在不同光照条件下的广泛适应性与卓越性能。

图 6—图 8 展示了本文方法在 5 个无配对低光数据集上的典型增强结果。可以看出,本文方法能够显著提升图像亮度,同时保持色彩自然、细节清晰,阴影区域得到有效还原,且无明显过曝或偏色。面对极暗或复杂纹理场景,依然表现出良好的稳定性和结构一致性,增强图像整体更加自然、真实,视觉质量优于常见方法。在多个非配对真实场景中均具备优越的性能表现与高度的实用价值,有效解决了低光图像增强中长期存在的亮度不足、细节缺失和色彩失真等问题。

2.4　消融实验

为全面验证所提出各模块在低光图像增强任务中的有效性,设计了两组消融实验:1)针对色彩自适

表4　在5个非配对数据集上的定量比较

Table 4　Quantitative comparison on five unpaired datasets

方法	DICM		LIME		MEF		NPE		VV	
	BRI↓	NIQE↓	BRI↓	NIQE↓	BRI↓	NIQE↓	BRI↓	NIQE↓	BRI↓	NIQE↓
KinD(Zhang等,2019)	48.72	5.15	39.91	5.03	49.94	5.47	36.85	4.98	50.56	4.30
RUAS(Liu等,2021)	38.75	5.21	27.59	4.26	23.68	3.83	47.85	5.53	38.37	4.29
ZeroDCE(Guo等,2020)	27.56	4.58	20.44	5.82	17.32	4.93	20.72	4.53	34.66	4.81
SNR-Aware(Xu等,2022)	37.35	4.71	39.22	5.74	31.28	4.18	26.65	4.32	78.72	9.87
SCI(Ma等,2022)	36.51	4.25	24.19	4.09	28.18	3.81	32.12	4.28	34.53	4.13
pairLIE(Fu等,2023)	33.31	4.03	25.23	4.58	27.53	4.06	28.27	4.18	39.13	3.57
PSENet(Nguyen等,2023)	29.87	3.90	21.75	2.58	32.47	3.66	30.69	3.75	30.46	3.63
GSAD(Hou等,2023)	24.31	3.88	19.83	4.32	30.46	3.90	23.79	3.85	31.54	3.60
QuadPrior(Wang等,2024b)	26.32	4.50	25.35	3.55	30.74	3.64	31.60	**3.19**	31.73	4.12
MSDFAR(Ji等,2025)	35.20	4.20	27.72	4.39	32.36	4.08	25.33	4.20	39.54	3.76
本文	**20.40**	**3.62**	**11.38**	**2.57**	**14.01**	**3.44**	**17.03**	3.70	**30.33**	**2.99**

注:加粗字体表示各列最优结果。“↓”表示值越低越好。

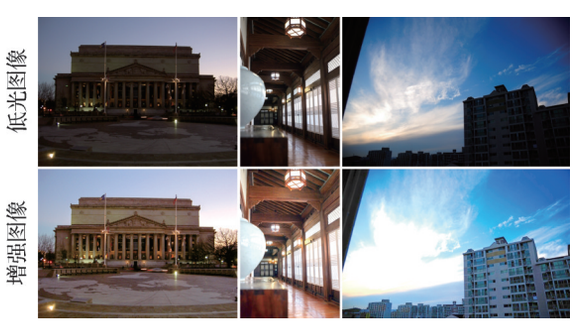


图6　DICM数据集评估效果

Fig. 6　Evaluation results on the DICM dataset

图7　LIME和NPE数据集评估效果

Fig. 7　Evaluation results on the LIME and NPE datasets

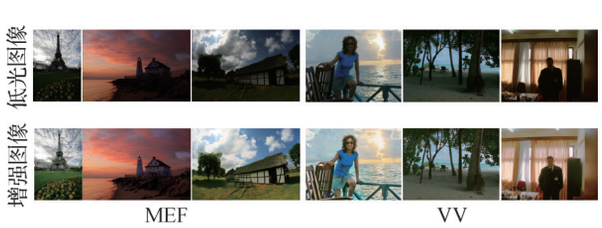


图8　MEF和VV数据集评估效果

Fig. 8　Evaluation results on the MEF and VV datasets

应因子(color factor)、GGCA与MDFA 3个子模块的组合效果进行系统评估;2)深入分析MDFA模块内部的空洞率参数对最终增强性能的影响。

从表5中可以观察到,逐步引入各模块均对性能有正向提升。相比基础模型,模型A和模型B的PSNR分别提升了0.95 dB和1.03 dB,表明颜色适应机制与通道注意力机制对增强质量的正面影响。同时,引入MDFA模块带来了更显著的感知质量提升,NIQE降至4.835,展示了多重空洞卷积结构在局部纹理恢复中具有独特优势。当多个模块联合使用时,增强效果进一步提高,模型F的PSNR达到23.435 dB,表现出较强的互补性。最终,本文将3种模块集成为完整模型G,取得了最优性能,验证了所提出模块设计在提升增强效果方面的有效性与协同优势。

为了进一步挖掘MDFA模块内部结构参数的影响机制,在模型G的基础上固定其他模块结构,仅调整MDFA中的空洞率设定,得到如表6所示的对比结果。本文设计了4组空洞率组合:(3, 6, 9)、(6, 12, 18)、(9, 18, 27)、(12, 24, 36)。实验表明,空洞率配置对模型感受野与多尺度信息聚合能力具有关键作用。当使用(9, 18, 27)这一空洞率组合在PSNR和LPIPS指标上表现最佳,表明该配置能够在细节恢复和感知质量方面达到最优。若采用较小的

表 5　LOL 数据集上的各模块消融实验

Table 5　Ablation study of each module on the LOL dataset

模型	色彩自适应因子	GGCA	MDFA	PSNR/dB ↑	SSIM ↑	LPIPS ↓	NIQE ↓
Base	–	–	–	20.423	0.825	0.137	5.327
A	√	–	–	21.372	0.834	0.132	5.143
B	–	√	–	21.453	0.827	0.143	5.343
C	–	–	√	21.234	0.839	0.133	4.835
D	√	√	–	22.834	0.842	0.128	4.973
E	√	–	√	22.972	0.857	0.109	4.952
F	–	√	√	23.435	0.849	0.104	4.876
G	√	√	√	**24.573**	**0.862**	**0.094**	**4.833**

注：加粗字体表示各列最优结果。“√”表示使用对应模块，“–”表示未使用。“↑”和“↓”分别表示值越高越好，值越低越好。

空洞率虽然组合能够保持较多的图像细节，但由于感受野受限，难以有效捕捉图像中低光区域的整体结构。采用过大的空洞率组合容易引入空洞效应，导致特征提取过程中图像细节的丢失。综合考虑空洞感受野的覆盖能力与图像细节的保持能力，本文选用的(9, 18, 27)空洞率组合在局部精细纹理建模和全局结构表达之间取得了良好平衡，使得模型在多个质量指标上表现优越。

表 6　空洞率消融实验

Table 6　The ablation experiment on dilation rate

方法	LOLv1			
	PSNR /dB ↑	SSIM ↑	LPIPS ↓	NIQE ↓
空洞率 = (3, 6, 9)	21.23	0.84	0.13	4.83
空洞率 = (6, 12, 18)	22.72	0.86	0.14	4.37
空洞率 = (9, 18, 27)(本文)	**24.57**	**0.86**	**0.09**	**4.25**
空洞率 = (12, 24, 36)	23.59	0.86	0.13	4.37

注：加粗字体表示各列最优结果。“↑”和“↓”分别表示值越高越好，值越低越好。

空洞率的优化在一定程度上提升了 MDFA 的空间建模能力，但其性能的充分发挥依赖于 color factor 和 GGCA 模块提供的颜色与语义引导。以模型 C 为例，即使采用了最优空洞率，其性能依然明显低于包含全部模块的模型 G，表明单靠空洞率难以弥补缺失的语义和颜色调节机制，无法达到本文所提出的最优效果。三大核心模块在增强框架中紧密协同，构建了结构严谨且层次分明的增强体系，融合了语义信息与纹理特征的双重驱动机制，从而显著提升了模型在复杂低光环境下的泛化能力与感知性能，确保了增强效果的稳定性与可靠性。

3　结　论

基于颜色空间的多重空洞卷积与坐标分组的暗光图像增强网络(MCCNet)有效改善了低光图像的亮度不均、色彩失真及伪影问题。通过引入 HVI 颜色空间模型并提出色彩自适应因子实现亮度增强与色彩恢复的解耦，提高色彩增强效果。与此同时，结合 MDFA 空洞卷积模块提取全局和局部特征，融合通道注意力和空间注意力机制，以增强特征表达能力。此外，提出的 GGCA 模块在亮度增强与色彩恢复分支之间建立有效信息流通，从而提升整体图像增强质量，并有效抑制低光环境下的噪声干扰。实验结果表明，本文方法在多个公开数据集上的定量与定性评估效果表现优异。相比 SCNet 方法，在 LOLv1 数据集上，PSNR 提升 1. 57 dB，LPIPS 值降至 0. 09，在 SLCE-Mix 和 SICE-Grad 等极端低光数据集上的 PSNR 与 SSIM 均优于现有主流方法，证明了所提方法在亮度增强与细节恢复方面的卓越性能。此外，在 DICM、LIME、MEF、NPE、VV 等无配对数据集上的 BRISQUE 和 NIQE 指标均显著优于已有方法，整体提升幅度最高达 50%～60%，进一步验证了本文方法在不同场景下的泛化能力和稳健性。因此，本文方法不仅在亮度增强方面取得了显著进展，同时在色彩恢复、伪影抑制及视觉感知质量方面均表

现优越。

参考文献(References)

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. Paris, France: IEEE: 12470-12479 [DOI: 10.1109/ICCV51070.2023.01149]

Cao J, Cao Y, Pang L, Meng D Y and Cao X Y. 2024. HAIR: hypernetworks-based all-in-one image restoration [EB/OL]. [2025-05-25]. https://arxiv.org/pdf/2408.08091.pdf

Chen Z G, Li J J, Lu M, Chen C Y, Zou Y and Chen J. 2024. Underwater image enhancement based on adaptive color balancing and improved IBLA. Journal of Optoelectronics·Laser, 35(2): 226-232 (陈祖国,李俊杰,卢明,陈超祥,邹莹,陈娟.2024.基于自适应色彩均衡及改进 IBLA 的水下图像增强. 光电子·激光, 35(2): 226-232) [DOI:10.37188/J.issn.1001-0730.2024.02.005]

Chen B, Zhang X Y, Liu S, Zhang Y B and Zhang J. 2025. Self-supervised scalable deep compressed sensing. International Journal of Computer Vision, 133(2): 688-723 [DOI: 10.1007/s11263-024-02209-1]

Cheng L H, Li C H, Hu R Z and Liu L G. 2025. 3D Gaussian splatting model for low-light enhancement. Journal of Image and Graphics, 30(10):3289-3301 (程龙昊, 李常颢, 胡瑞珍, 刘利刚. 2025. 面向低光增强的三维高斯泼溅模型. 中国图象图形学报, 30(10): 3289-3301) [DOI:10.11834/jig.240598]

Dang R, Yuan Y, Luo C and Liu J. 2017. Chromaticity shifts due to light exposure of inorganic pigments used in traditional Chinese painting. Lighting Research and Technology, 49(7): 818-828 [DOI: 10.1177/1477153516644866]

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: 1460-1470

Fei B, Lyu Z Y, Pan L, Zhang J Z, Yang W D, Luo T Y, et al. 2023. Generative diffusion prior for unified image restoration and enhancement//Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Vancouver, Canada: IEEE: 9935-9946 [DOI: 10.1109/CVPR52729.2023.00958]

Fu Z Q, Yang Y, Tu X T, Huang Y, Ding X H and Ma K K. 2023. Learning a simple low-light image enhancer from paired low-light instances//Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Vancouver, Canada: IEEE: 22252-22261 [DOI: 10.1109/CVPR52729.2023.02131]

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. Seattle, USA: IEEE: 1777-1786 [DOI: 10.1109/CVPR42600.2020.00185]

He Z Q, Ran W, Liu S L, Li K H, Lu J W, Xie C Y, et al. 2024. Low-light image enhancement with multi-scale attention and frequency-domain optimization. IEEE Transactions on Circuits and Systems for Video Technology, 34(4): 2861-2875 [DOI: 10.1109/TCSVT.2023.3313348]

Hou J H, Zhu Z Y, Hou J H, Liu H, Zeng H Q and Yuan H. 2023. Global structure-aware diffusion process for low-light image enhancement//Proceedings of the 37th International Conference on Neural Information Processing Systems. New Orleans, USA: Curran Associates Inc.: #3490

Hsu P H, Lin C T, Ng C C, Kew J L, Tan M L, Lai S H, et al. 2022. Extremely low-light image enhancement with scene text restoration//Proceedings of the 26th International Conference on Pattern Recognition. Montreal, Canada: IEEE: 317-323 [DOI: 10.1109/ICPR56361.2022.9956716]

Ji P, Lyu Z Y and Zhang Z. 2025. A Retinex-based network for low-light image enhancement with multi-scale denoising and focal-aware reflections. IET Image Processing, 19(1): #e70059 [DOI: 10.1049/ipr2.70059]

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 (ECCV). 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]

Liang J X, Xu Y, Quan Y H, Shi B X and Ji H. 2022a. Self-supervised low-light image enhancement using discrepant untrained network priors. IEEE Transactions on Circuits and Systems for Video Technology, 32(10): 7332-7345 [DOI: 10.1109/TCSVT.2022.3181781]

Liang Y, Li X, Li Y, Li Y and Xu Z. 2022b. low-light image enhancement via neural network priors. IEEE Transactions on Image Processing, 31: 1234-1245 [DOI: 10.1109/TIP.2022.3145678]

Liang Z X, Li C Y, Zhou S C, Feng R C and Loy C C. 2023. Iterative prompt learning for unsupervised backlit image enhancement//Proceedings of 2023 IEEE/CVF International Conference on Computer Vision (ICCV). Paris, France: IEEE: 8060-8069 [DOI: 10.1109/ICCV51070.2023.00743]

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: 10556-10565 [DOI: 10.1109/CVPR46437.2021.

01042]

Ma L, Ma T Y and Liu R S. 2022. The review of low-light image enhancement. Journal of Image and Graphics, 27(5): 1392-1409 (马龙, 马腾宇, 刘日升. 2022. 低光照图像增强算法综述. 中国图象图形学报, 27(5): 1392-1409) [DOI: 10.11834/jig.210852]

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]

Nguyen H, Tran D, Nguyen K and Nguyen R. 2023. PSENet: progressive self-enhancement network for unsupervised extreme-light image enhancement//Proceedings of 2023 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV). Waikoloa, USA: IEEE: 1756-1765 [DOI: 10.1109/WACV56688.2023.00180]

Pan N, Zhang M H, Yang X H, Chen S Y and Guo X J. 2023. Track planning for damage detection on the building enclosure by UAV considering image detection. Journal of Imaging Science and Technology, 67(4): #040401 [DOI: 10.2352/J. ImagingSci. Technol. 2023.67.4.040401]

Ren K and Miao X. Low-light image enhancement via adaptive frequency decomposition network. Journal of Optoelectronics · Laser, 2024, 35(6): 612-620

Tang J H, Zhou K H, Luo Z J and Hou Y E. 2025. Natural language supervision for low-light image enhancement [EB/OL]. [2025-05-25]. https://arxiv.org/pdf/2501.06546.pdf

Wang T, Zhang K H, Shen T R, Luo W H, Stenger B and Lu T. 2023. Ultra-high-definition low-light image enhancement: a benchmark and transformer-based method//Proceedings of the 37th AAAI Conference on Artificial Intelligence. Washington DC, USA: AAAI: 2654-2662 [DOI: 10.1609/aaai.v37i3.25364]

Wang W J, Yang H, Fu J L and Liu J Y. 2024b. Zero-reference low-light enhancement via physical quadruple priors//Proceedings of 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, USA: IEEE: 26057-26066 [DOI: 10.1109/CVPR52733.2024.02462]

Wang Y, Chen J, Han Y J and Miao D Q. 2025. Combining attention mechanism and Retinex model to enhance low-light images. Computers and Graphics, 104: 95-105 [DOI: 10.1016/j. cag. 2022.04.002]

Wang Y F, Wan R J, Yang W H, Li H L, Chau L P and Kot A. 2022. Low-light image enhancement with normalizing flow//Proceedings of the 36th AAAI Conference on Artificial Intelligence. [s.l.]: AAAI: 2604-2612 [DOI: 10.1609/aaai.v36i3.20162]

Wei C, Wang W J, Yang W H and Liu J Y. 2018. Deep Retinex decomposition for low-light enhancement//Proceedings of 2018 British Machine Vision Conference (BMVC). Newcastle, UK: BMVA: #155 [DOI:10.48550/arXiv.1808.04560]

Wu W H, Weng J, Zhang P P, Wang X, Yang W H and Jiang J M. 2022. URetinex-Net: Retinex-based deep unfolding network for low-light image enhancement//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans, USA: IEEE: 5891-5900 [DOI: 10.1109/CVPR52688.2022.00581]

Xu X G, Wang R X, Fu C W and Jia J Y. 2022. SNR-aware low-light image enhancement//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans, USA: IEEE: 17693-17703 [DOI: 10.1109/CVPR52688.2022.01719]

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. Nashville, USA: IEEE: 5678-5687 [DOI: 10.1109/CVPR52734.2025.00533]

Yang W, Wang S, Fang Y, Wang Y and Liu J. 2020. From fidelity to perceptual quality: a semi-supervised approach for low-light image enhancement//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition [s.l.]: [s.n.]: 3060-3069 [DOI: 10.1109/CVPR42600.2020.00313]

Yao Z S, Fan G D, Fan J F, Gan M and Chen C L P. 2024. Spatial-frequency dual-domain feature fusion network for low-light remote sensing image enhancement. IEEE Transactions on Geoscience and Remote Sensing, 62: #4706516 [DOI: 10.1109/TGRS. 2024.3434416]

Yu C K and Han G L. 2025. HSV space-based nonlinear adaptive low-light image enhancement algorithm. Laser and Optoelectronics Progress, 62(8): #0837001 (于成康, 韩广良. 2025. HSV空间的非线性自适应低光照图像增强算法. 激光与光电子学进展, 62(8): #0837001) [DOI: 10.3788/LOP241817]

Yu W S, Zhao L Q and Zhong T. 2023. Unsupervised low-light image enhancement based on generative adversarial network. Entropy, 25(6): #932 [DOI: 10.3390/e25060932]

Zamir S W, Arora A, Khan S, Hayat M, Khan F S and Yang M H. 2022. Restormer: efficient transformer for high-resolution image restoration//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans, USA: IEEE: 5718-5729 [DOI: 10.1109/CVPR52688.2022.00564]

Zhang Y, Zhang J and Guo X. 2019. Kindling the darkness: a practical low-light image enhancer//Proceedings of the 27th ACM International Conference on Multimedia. [s.l.]: ACM: 1632-1640 [DOI: 10.1145/3343031.3350926]

Zhang Y H, Guo X J, Ma J Y, Liu W and Zhang L W. 2021. Beyond brightening low-light images. International Journal of Computer Vision, 129(4): 1013-1037 [DOI: 10.1007/s11263-020-01407-x]

Zhang Y B, Li J Y and Wang M L. 2024. An algorithm for low-light image enhancement in coal mines based on HSV space. Infrared Technology, 46(1): 74-83 (张亚邦,李佳悦,王满利. 2024. 基于

HSV 空间的煤矿井下低光照图像增强方法. 红外技术, 46(1): 74-83) [DOI: 10.11835/j.issn.1001-8891.20240109]

Zhang H and Yan J. 2024. Low-light image enhancement guided by semantic segmentation and HSV color space. Journal of Image and Graphics, 29(4): 966-977 (张航, 颜佳. 2024. 语义分割和HSV色彩空间引导的低光照图像增强. 中国图象图形学报, 29(4): 966-977) [DOI: 10.11834/jig.230182]

Zhao M H, Wen Y C, Du S L, Hu J, Shi C and Li P. 2024. Low-light image enhancement algorithm based on illumination and scene texture attention map. Journal of Image and Graphics, 29(4): 862-874 (赵明华, 汶怡春, 都双丽, 胡静, 石程, 李鹏. 2024. 基于照度与场景纹理注意力图的低光图像增强. 中国图象图形学报, 29(4): 862-874) [DOI: 10.11834/jig.230271]

Zhao Z J, Xiong B S, Wang L, Ou Q F, Yu L and Kuang F. 2022. RetinexDIP: a unified deep framework for low-light image enhancement. IEEE Transactions on Circuits and Systems for Video Technology, 32(3): 1076-1088 [DOI: 10.1109/TCSVT.2021.3073371]

Zheng S and Gupta G. 2022. Semantic-guided zero-shot learning for low-light image/video enhancement//Proceedings of 2022 IEEE/CVF Winter Conference on Applications of Computer Vision Workshops (WACVW). Waikoloa, USA: IEEE: 581-590 [DOI: 10.1109/WACVW54805.2022.00064]

Zhou S C, Li C Y and Loy C C. 2022. LEDNet: joint low-light enhancement and deblurring in the dark//Proceedings of the 17th European Conference on Computer Vision (ECCV). Tel Aviv, Israel: Springer: 573-589 [DOI: 10.1007/978-3-031-20068-7_33]

作者简介

孙婉倩, 女, 硕士研究生, 主要研究方向为非遗数字化与智能处理、图像处理。E-mail: 1029166216@qq.com

彭春燕, 通信作者, 女, 教授, 主要研究方向为文化计算和计算机视觉。E-mail: pcy@qhnu.edu.cn

张效娟, 女, 教授, 主要研究方向为非遗数字化保护及智能处理。E-mail: zhangxj@qhnu.edu.cn