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:TP391.4;TP18　　**文献标识码**:A　　**文章编号**:1006-8961(2026)08-2882-15

论文引用格式:Zhang Q, Miao Z, Wang J B, Ji W Y, Bi X H and Lei X Z. 2026. Dual-branch feature enhancement and calibration network for low-light images. Journal of Image and Graphics, 31(8):2882-2896(张奇, 苗壮, 王家宝, 纪文宇, 毕翔鹤, 雷小舟. 2026. 低光照图像的双分支特征增强与校准网络. 中国图象图形学报, 31(8):2882-2896)[DOI:10.11834/jig.250530]

低光照图像的双分支特征增强与校准网络

张奇,苗壮,王家宝,纪文宇,毕翔鹤,雷小舟*

陆军工程大学指挥控制工程学院,南京　210007

摘　要:**目的**　低光照图像增强(low light image enhancement,LLIE)旨在改善低光照条件下的图像质量,在夜间监控、自动驾驶等领域具有重要应用价值。然而,现有方法在sRGB(standard red green blue)、HSV(hue saturation value)、YCbCr(luminance and chrominance blue and red)等色彩空间中常面临亮度—色彩耦合导致的色彩失真与亮度伪影平衡难题。HSV空间存在极坐标量化引起的色彩不连续,YCbCr空间对色度信息处理不足。**方法**　针对上述问题,提出一种基于HVI(horizontal vertical intensity)色彩空间的门控卷积双分支增强网络(gated convolution dual-branch interaction network,GCDINet)。该网络利用HVI空间的物理解耦特性,将亮度信息与色彩结构信息分离处理。设计了门控卷积特征增强(gated convolutional feature enhancement,GCFE)模块,通过双路门控激活函数与逐元素乘法生成空间注意力掩码,强化细节区域特征响应。同时,提出了双路径自适应特征校准结构,使用其引导亮度与色彩分量的重建,并抑制由亮度调整引入的噪声和色彩的偏移。**结果**　在LOLv2-Synthetic数据集上,本文方法的峰值信噪比(peak signal-to-noise ratio,PSNR)、结构相似度(structural similarity index measure,SSIM)和感知相似性(learned perceptual image patch similarity,LPIPS)分别达到25.698 dB、0.935和0.047。与当前先进LLIE方法Retinexformer相比,3个指标分别提升0.028 dB、5.3%和20.3%。同时,计算量为7.94 GFLOPs,比Retinexformer降低了50%,达到了与最优方法相比速度和性能的有效平衡。**结论**　所提出的低光照增强方法,融合了双分支结构与HVI色彩空间的优势,从而实现低光照图像增强过程中模型色彩保真度与亮度伪影抑制的平衡能力。

关键词:低光照图像增强(LLIE);HVI色彩空间;双分支特征增强;门控卷积;特征校准

Dual-branch feature enhancement and calibration network for low-light images

Zhang Qi, Miao Zhuang, Wang Jiabao, Ji Wenyu, Bi Xianghe, Lei Xiaozhou*

Command and Control Engineering College, Army Engineering University of PLA, Nanjing　210007, China

Abstract: **Objective** Low-light image enhancement (LLIE) is a key challenge in computational photography and computer vision, with critical applications ranging from autonomous navigation and night surveillance to computational imaging. Although existing methods in sRGB, HSV, and YCbCr color spaces have shown effectiveness, they face an irreconcilable trade-off between preserving color fidelity and suppressing luminance artifacts, often introducing chromatic shifts or halo distortions during the enhancement process. Recent studies focus on color space optimization. YCbCr-based methods perform shallow luminance processing (e. g., linear filtering), causing an imbalance between chrominance denoising and

收稿日期:2025-11-05;**修回日期**:2026-01-15;**预印本日期**:2026-01-22

*通信作者:雷小舟　sigmoid@qq.com

基金项目:国家自然科学基金项目(72471240)

Supported by:National Natural Science Foundation of China(72471240)

detail preservation. Accordingly, the performance of these methods is capped. HSV space strategies decouple the hue and saturation components but cause color discontinuities due to polar coordinate quantization. This study proposes a novel lightweight LLIE network to address these limitations. This network fully leverages the channel decoupling properties of the HVI color space, minimizing the mutual interference between luminance information and color structural information during enhancement. **Method** We propose GCDINet, an end-to-end framework leveraging HVI color space's physical decoupling. The architecture comprises: 1) a dual-branch structure that processes value-intensity and hue components independently via the GCFE module, which utilizes dual GELU activations and element-wise multiplication to generate spatial attention masks; 2) a dual-path adaptive feature calibration (DAFC) structure that nonlinearly modulates illumination and chrominance features before reconstruction; 3) a composite loss combining HVI-space and sRGB-space constraints. The novel HVI color space decouples images into luminance and color structure maps at the original resolution, allowing the dual-branch network to process luminance and color structure information independently. This mechanism avoids interference from processing in the same feature stream and provides independent branch information for the GCFE module. A DAFC structure nonlinearly modulates the illumination features from the luminance branch and the chrominance vectors from the color branch before feature reconstruction. This approach guides information reconstruction during decoding, suppressing color fidelity loss and luminance artifact amplification caused by information loss. Although the dual-branch structure and gated convolutions decouple luminance and color structure information for a refined feature enhancement, uneven channel weight distribution during feature fusion can result in the loss of key luminance and chrominance information. During training, we optimize using a composite loss function that combines L1 intensity loss, structural similarity index measure (SSIM) loss, and edge-aware regularization. The model is implemented in PyTorch with the Adam optimizer (β_1 = 0.9, β_2 = 0.999), an initial learning rate of 1×10^{-4}, and a cosine annealing decay strategy that reduces the learning rate to 1×10^{-7}. Training is conducted on an NVIDIA RTX 3090 GPU for 1 000 epochs. The training images are cropped to 256×256 with a batch size of eight. We apply random rotation, horizontal flipping, and vertical flipping for the data augmentation of LOLv2-Real-captured. **Result** Comprehensive evaluations on the LOLv1, LOLv2-Real, and LOLv2-Synthetic datasets demonstrate the state-of-the-art performance of GCDINet. On the LOLv1 dataset, our method achieves the best results in terms of SSIM and learned perceptual image patch similarity (LPIPS) and is only 0.76 dB short of the best method in peak signal-to-noise ratio (PSNR). On the LOLv1-Real dataset, our method attains optimal results in PSNR and SSIM and is only 0.006 short of the best method in LPIPS. Experimental results on the LOLv2-Synthetic dataset demonstrate that the proposed method achieves a PSNR of 25.698 dB, an SSIM of 0.935, and an LPIPS score of 0.047. These metrics show improvements of 0.028 dB, 5.3%, and 20.3% compared with the state-of-the-art LLIE method Retinexformer. Furthermore, the computational complexity is 7.94 GFLOPs, which is 50% lower than that of Retinexformer, achieving an effective balance between speed and performance compared to the existing methods. Ablation studies show that the PSNR, SSIM, and LPIPS can be improved by 1.9%, 0.2%, and 9.4%, respectively, using only the dual-path adaptive feature calibration structure. When integrated with the GCM module, these improvements increase to 2.8%, 0.64%, and 11.3%. The utilization of HVI color space decomposition yields in enhanced PSNR by 18.9% over the HSV baseline, SSIM by 6.4%, and perceptual quality by 64.3%. These three evaluation metrics improve by 4.9%, 2.1%, and 36.4% compared with YCbCr. The use of only the HVI loss lacks pixel-space consistency constraints, resulting in large pixel errors and decreased performance across all three metrics, especially PSNR. By contrast, the use of only sRGB loss focuses on pixel-space enhancement but ignores the low-light probability distribution in the HVI color space, resulting in color imbalance. However, from a visualization perspective, the pixel-level errors and color distribution of using only HVI/sRGB loss functions are not intuitively observable. **Conclusion** The proposed LLIE method in this study effectively combines the advantages of a dual-branch structure and the HVI color space, achieving a balance between color fidelity and luminance artifact suppression in low-light image enhancement. GCDINet resolves the fundamental trade-off between color and artifacts through deterministic component decoupling and adaptive feature calibration, with ablation studies confirming the synergistic benefits of the GCFE and DAFC modules.

Key words: low-light image enhancement (LLIE); HVI color space; dual-branch feature enhancement; gated convolution; feature calibration

0 引 言

作为计算成像领域的关键技术，低光照图像增强（low light image enhancement，LLIE）能够解决低照度条件下的图像质量退化问题。在低光照环境下，成像设备捕获的图像通常会出现亮度伪影和色彩信息丢失的现象，这会显著降低图像的视觉质量和实用价值。因此，LLIE致力于恢复图像亮度，纠正图像由于低光照环境导致的亮度伪影以及色彩失真。

早期LLIE研究主要基于直方图均衡化（Xiao等，2016；Łoza等，2010；Kim等，2016）、Retinex理论（Land，1977；Fu等，2015；Park等，2017）。直方图均衡化及其改进算法通过扩展灰度分布提升可见性，但常引起过增强和色彩失真。Retinex模型通过分解光照与反射分量实现对比度增强，但易导致光晕伪影。单尺度Retinex（Zhang等，2017）采用高斯滤波估计光照，但产生光晕伪影；多尺度Retinex（Gao和Zhai，2022）融合多尺度滤波结果，计算复杂度较高；带色彩恢复的多尺度视网膜（multi scale Retinex，MSR）理论（王奎和黄福珍，2022）引入色彩校正因子，但参数调节依赖经验，易导致色彩偏差。传统方法的共性问题在于手工设计特征难以建模复杂的光照—噪声耦合退化，且依赖大量人工调参。这些问题推动了基于深度学习方法的兴起。

随着深度学习技术的突破，基于卷积神经网络（convolutional neural network，CNN）的端到端方法（Dudhane等，2025；马龙 等，2022；Liu等，2021；Moran等，2020；Xu等，2022；Yi等，2023）逐渐占据主导地位。RetinexNet（Wu等，2022）将Retinex理论与CNN结合使网络可以分别处理光照调整和反射率修复，但存在光照估计偏差问题。

随着去噪扩散概率模型（diffusion probabilistic model，DDPM）（Ho等，2020）的发展，基于DDPM的生成模型通过隐式表示提升生成质量在LLIE任务中取得了显著的效果。例如，Diff-Retinex（Yi等，2023）提出生成框架，进一步缓解了低光照导致的内容丢失和色彩偏差问题。然而，这些扩散模型往往不能完全解耦亮度和色彩信息且迭代式推理效率较低。

生成对抗网络（generative adversarial network，GAN）（Goodfellow等，2014）的发展为LLIE提供了一个新的思路。通过对抗训练生成低光图像对应的正常光图像，如EnlightenGAN（Jiang等，2021）使用生成器模型将低光图像转换成正常光图像。

现有LLIE方法虽然取得了显著进展，但仍难以平衡色彩保真度与亮度伪影控制之间的矛盾。sRGB色彩空间因其与显示设备的高效兼容性得到广泛采用（Poynton，2003），但该空间中亮度与色彩通道高度耦合，在增强过程中容易导致二者相互干扰。例如，在提升亮度时容易引发色彩偏移（Cai等，2023；Jiang等，2021；Gevers等，2012），而保护色彩信息又可能导致亮度增强不足。基于HSV色彩空间的分解策略（Zhou等，2025）虽然能实现亮度分量独立调节，但其色相—饱和度平面具有极坐标特性，不仅会放大红色不连续噪声，而且在极暗区域易产生黑色平面伪影。这些伪影本质上破坏了亮度的均匀性和自然过渡，成为新的干扰源。YCbCr色彩空间虽然通过亮度—色度分离机制缓解了通道干扰，但现有方法（Brateanu等，2025）忽略了图像的空间结构相关性，导致模型的亮度增强功能与保留细微色彩细节、边缘信息功能两者之间难以取得平衡。

目前，双分支网络已经被证明在图像增强方面更有效，通过子网络将任务分解成子任务，逐步进行图像恢复。因此，许多LLIE方法（Chen等，2023；Wang等，2024；杨微 等，2022）都遵循这种设计理念。然而，这些网络一般在原始分辨率下使用双分支保留空间细节，虽然有利于细节保持，但会引入显著的计算负担和内存开销。同时，在原始分辨率下并行处理两个分支，会导致不同分支处理的信息在关键的空间结构关联性上相对隔离。此外，这些网络通常将图像亮度信息与色彩结构信息耦合在一起进行增强（Guo和Hu，2023；Zhou等，2025；张航和颜佳，2024），这导致在提升亮度分量时易对耦合在同一特征空间内的色彩结构信息产生干扰。还有一些方法通过跳跃连接直接融合多尺度特征（胡海峰和李凤英，2023；Wang等，2024），忽视了不同通道在跨阶段信息传递中的差异性，可能会导致关键色度信息丢失并引入伪影。

针对上述挑战，本研究主要贡献如下：1）针对LLIE过程中亮度和色彩结构信息相互干扰问题设计了一种新的轻量级LLIE网络GCDINet，使用双路门控卷积特征增强（gated convolutional feature enhancement，GCFE）模块降低增强过程中的亮度信息与色彩结构信息互相干扰问题，实现双路信息的

特征融合与关键细节的自适应增强。仅使用较小的参数和计算负载，就可以学习到不同光照条件下的光度映射关系。2）针对双分支网络特征重建中通道权重失衡问题，将通道感知特征校准（channel perceive feature calibration，CPFC）模块嵌入U型网络跨层连接通路，获得双路径自适应特征校准结构。在特征重建前对亮度特征与色彩结构特征进行非线性调制，确保亮度与色彩信息的高保真重建。3）针对LOLv2-Real数据集训练过程中由于场景同构性过高而出现的梯度异常问题，提出基于随机几何变换的自适应数据增强策略。根据一定的概率P对图像进行翻转或旋转的预处理操作，从而降低图像的场景同构性。4）在LOLv1、LOLv2-Real-captured和LOLv2-Synthetic等多个基准数据集上的实验表明，本文方法在性能与效率平衡上具有优势，并且能保持更优的视觉一致性。

图1给出了使用LOLv2-Synthetic数据集评估当前最先进的（state of the art，SOTA）方法在复杂性方面的性能对比分析。可以看到，所提方法在参数量与性能之间取得了较好的平衡。

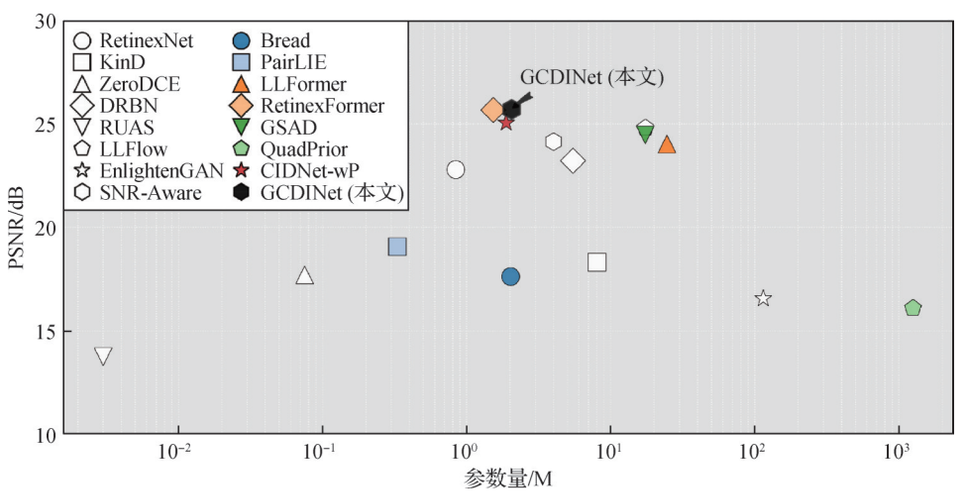


图1　所提模型GCDINet与已有方法在LLIE任务中的性能对比

Fig. 1　Performance comparison between the proposed GCDINet and existing methods on the LLIE task

1　网络结构

1.1　双分支增强网络

图2为所提方法的网络架构。其中，图2（a）为所提方法的整体框架，主要由3部分组成：基于HVI色彩空间的图像转换（HVT）模块、门控卷积特征增强（GCFE）模块和内嵌CPFC模块的双路径自适应特征校准结构。绿色箭头表示亮度分支，粉色箭头表示色彩结构分支。图2（b）为GCFE模块结构图，主要由多头自注意块（multi-head self-attention block，MHSA）和门控卷积块（gated convolutional module，GCM）两个模块组成。该网络建立在U-Net架构上，包括一个编码器和一个解码器，以及一个双路径自适应特征校准结构。编码器和解码器主要由3个门控卷积特征增强（GCFE）模块组成。亮度分支用B_I表示，色彩结构分支用B_{HV}表示。

1.2　门控卷积特征增强模块

为了增强B_I和B_{HV}包含的图像信息，本文提出GCFE模块来学习亮度图$\boldsymbol{I}$和色彩结构图$\boldsymbol{HV}$的信息。如图2（a）所示，GCFE中的B_{HV}和B_I分别处理HV特征和亮度特征。具体而言，GCFE模块由一个多头自注意块（MHSA）和一个门控卷积块（GCM）组成。低光照图像的LLIE任务可以解耦为两部分：低光照区域的亮度增强和色彩结构信息恢复。$\boldsymbol{X}_{\text{Input}} \in \mathbf{R}^{H\times W\times 3}$表示模型的输入，所提方法使用HVI色彩空间（Yan等，2025）将$\boldsymbol{X}_{\text{Input}}$的亮度信息与色彩结构信息解耦为亮度图$\boldsymbol{X}_I \in \mathbf{R}^{H\times W\times 1}$和色彩结构图$\boldsymbol{X}_{HV} \in \mathbf{R}^{H\times W\times 2}$。由于$B_{HV}$和$B_I$在GCFE中呈对称结构，用$B_I$为例对所提方法进行详细说明。如图2（b）所示，$\boldsymbol{X}_I \in \mathbf{R}^{H\times W\times 1}$表示$B_I$的输入，通过1×1卷积和3×3卷积生成查询向量$\boldsymbol{Q}$，进一步使用深度可分离卷积提取空间特征。同时，将键（$\boldsymbol{K}$）和值（$\boldsymbol{V}$）进行拆分。具体计算为

$$\begin{cases} \boldsymbol{Q} = DW(W_q \boldsymbol{X}_I) \\ [\boldsymbol{K},\boldsymbol{V}] = Chunk(DW(W_{kv}\boldsymbol{Y}_I)) \end{cases} \tag{1}$$

式中，$\boldsymbol{Y}_I$为正常光照下的图像，W_q、W_{kv}分别为查询向量$\boldsymbol{Q}$和键值向量$[\boldsymbol{K},\boldsymbol{V}]$的权重，$DW(\cdot)$为深度可分离卷积函数，$Chunk(\cdot)$为通道拆分函数。

对$\boldsymbol{Q},\boldsymbol{K}$向量进行L2归一化后，计算通道注意力权重$Attn=\sigma((QK^{\mathrm{T}})\cdot\tau)$。最后，通过$\boldsymbol{V}\cdot Attn$加权融合特征，经过$1\times1$卷积后输出$\tilde{\boldsymbol{X}}_I$。以上过程可描述为

$$\tilde{\boldsymbol{X}}_I = W(\boldsymbol{V} \otimes \sigma((\boldsymbol{Q} \otimes \boldsymbol{K}^{\mathrm{T}})\cdot\tau) + \boldsymbol{X}_I) \tag{2}$$

式中，$\tau\in\mathbf{R}$为可学习参数，σ为softmax函数，$W(\cdot)$表示特征嵌入卷积，$\otimes$表示逐元素相乘。通过参数τ动态调节各头注意力聚焦强度，结合深度卷积增强局部特征提取，实现高效的多源特征交互。

接下来，GCM模块通过全连接层将$\tilde{\boldsymbol{X}}_I$特征通道扩展至原通道的2倍后，将其拆分为两个分支特征$\boldsymbol{x}_1,\boldsymbol{x}_2$，具体计算为

$$[\boldsymbol{x}_1,\boldsymbol{x}_2] = Chunk(FC_1(\tilde{\boldsymbol{X}}_I)) \tag{3}$$

式中，FC_1表示全连接层。

然后，对分支特征$\boldsymbol{x}_1,\boldsymbol{x}_2$分别进行深度可分离卷积，提取空间上下文特征。将从分支特征$\boldsymbol{x}_I$提取到的上下文特征经由GELU(Gaussian error linear unit)激活函数处理，利用GELU函数基于概率的门控机制对特征进行自适应调制，以控制像素值亮度增强和颜色校正程度。具体计算为

$$GELU(\boldsymbol{x}_1) = \boldsymbol{x}_1 \cdot \frac{1}{2}\left[1 + erf\left(\frac{\boldsymbol{x}_1}{\sqrt{2}}\right)\right] \tag{4}$$

式中，$erf(\cdot)$表示误差函数。

接着，将经由GELU激活函数映射后的分支特征$\boldsymbol{x}_1$与分支特征$\boldsymbol{x}_2$逐元素乘积构建空间感知门控机制，具体计算为

$$\hat{\boldsymbol{X}}_I = GELU(DW(\boldsymbol{x}_1)) \otimes DW(\boldsymbol{x}_2) \tag{5}$$

式中，$GELU(\cdot)$为门控卷积激活函数。

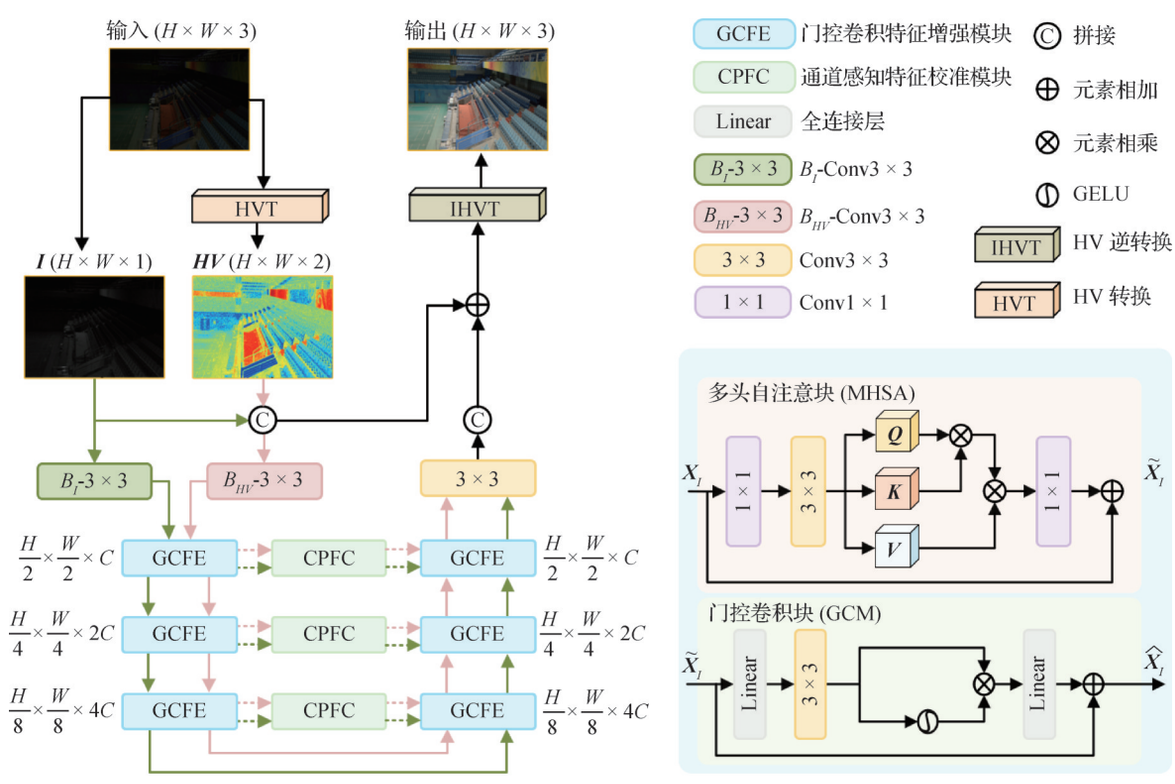
门控机制通过空间位置相关的激活权重动态调节特征响应，可使关键区域激活强度大幅提升。最后，通过全连接层降维并保留关键信息后输出。具体计算为

$$\hat{\boldsymbol{X}}_I = Dropout(FC_2(\mathrm{Dropout}(\hat{\boldsymbol{X}}_I))) \tag{6}$$

式中，$Dropout(\cdot)$是一种正则化操作，用于抑制模型出现过拟合，FC_2表示全连接层。

1.3 双路径自适应特征校准结构

传统双分支网络在跨层特征融合时，往往通过



(a) 本文方法的整体框架　　(b) GCFE 模块结构图

图2　网络结构图

Fig. 2　Network architecture diagram ((a) the overall framework of ours; (b) the structure of the GCFE module)

简单相加或拼接方式传递编码器特征，忽视了B_I与B_{HV}在不同解码阶段对通道信息的差异化需求，如图3所示。

具体而言，在低光照增强场景下，B_I和B_{HV}在不同解码阶段对通道信息的需求存在差异，例如：B_I侧重于光照信息的恢复，而B_{HV}关注色彩细节的保真。传统方法忽略不同通道的侧重需求，通过粗粒度的融合机制直接聚合多通道特征，导致亮度恢复通道与色彩结构通道被赋予同等权重，容易造成亮度和色彩结构信息的无效扩散，特别是在低光照条件下图像质量难以得到可靠提升。

为解决该问题，本文提出一种双路径自适应特征校准结构。其核心机制如图4所示。原始路径保证特征的快速传递，在跳跃连接部分嵌入通道感知特征校准（CPFC）模块，用于精细化的特征过滤和优化。

具体而言，传统跳跃连接的直通路径被替换为

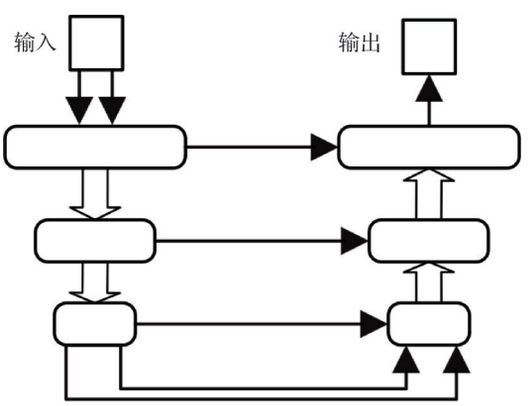


图3　传统双分支网络跨层特征融合结构

Fig. 3　Cross-layer feature fusion structure for traditional two-branch networks

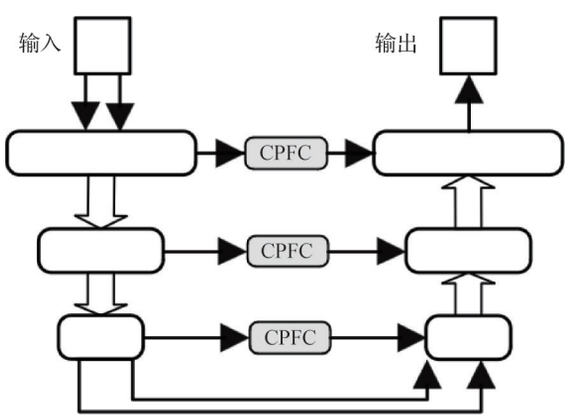


图4　双路径自适应特征校准结构

Fig. 4　Dual-path adaptive feature calibration structure

一种双路径交互结构。该结构一方面保留输入特征的原始传递路径；另一方面新增一个CPFC模块。CPFC模块在压缩激励（squeeze-and-excitation，SE）模块（Hu等，2018）的基础上增加了两个dropout层得到。首先执行通道感知机制，即通过对特征图进行全局平均池化提取通道级光照分布统计量。

随后，由两个全连接层构建的瓶颈结构建模通道间非线性关系，生成通道级的注意力权重，从而动态评估不同通道的重要性。强化与亮度和色彩相关的关键特征，同时抑制噪声相关通道。最后，CPFC模块通过自适应校准操作，应用学习到的权重对输入特征进行校准后输出至解码器。该结构能够感知解码阶段的需求变化，在浅层解码阶段侧重全局亮度增强，深层阶段侧重局部色彩恢复，实现通道级别的差异化加权。

1.4　HVI色彩空间

HVI色彩空间是在HSV色彩空间的基础上建立的，旨在解决HSV空间因解耦图像时产生的色彩空间噪声问题。具体而言，HSV色彩空间对图像的亮度与色彩结构信息进行解耦时，会产生红色不连续和黑色平面噪声，而HVI色彩空间利用极化操作和可学习的坍缩函数，有效解决了这些问题。

HVI色彩空间由亮度图$\boldsymbol{I}$和色彩结构图$\boldsymbol{HV}$两部分组成。亮度图$\boldsymbol{I}$由输入图像$\boldsymbol{R}$，$\boldsymbol{G}$，$\boldsymbol{B}$ 3个通道中的最大像素值组成，具体计算为

$$\boldsymbol{I} = \max_{c \in \{R,G,B\}} \boldsymbol{I}_c(x) \tag{7}$$

式中，x表示单个像素值，c为图像通道。

色彩结构图$\boldsymbol{HV}$由传统HSV色彩空间中的色相（hue，H）和饱和度（saturation，S）分量组成，其中$\boldsymbol{H}$分量是经过极化处理的。为了解决传统HSV色彩空间解耦图像时带来的红色不连续噪声，Yan等人（2025）对传统HSV色彩空间中的$\boldsymbol{H}$分量进行极化操作，具体计算为

$$\begin{cases} \tilde{\boldsymbol{H}} = \cos(\pi h/3) \\ \tilde{\boldsymbol{V}} = \sin(\pi h/3) \end{cases} \tag{8}$$

式中，h为$\boldsymbol{H}$中的任意像素值。

对于传统HSV色彩空间解耦时带来的黑色平面噪声问题，HVI色彩空间中构建一个自适应的坍缩函数C_k，将黑色平面噪声坍缩为一个噪声点。具体计算为

$$C_k(x)=\sqrt[k]{\sin\left(\frac{\pi \boldsymbol{I}_{\max}(x)}{2}\right)+\varepsilon} \tag{9}$$

式中，$k\in \boldsymbol{Q}^{+}$是一个可训练参数，用于控制黑色点密度；ε是一个极小值用于防止梯度爆炸。

然后，利用式(8)得到极化后的$\tilde{\boldsymbol{H}}$和$\tilde{\boldsymbol{V}}$分量、坍缩函数C_k以及传统HSV色彩空间中的$\boldsymbol{S}$分量进行拼接，得到最终的$\boldsymbol{HV}$色彩结构图。具体计算为

$$\begin{cases}\hat{\boldsymbol{H}}=C_k\otimes \boldsymbol{S}\otimes \tilde{\boldsymbol{H}}\\ \hat{\boldsymbol{V}}=C_k\otimes \boldsymbol{S}\otimes \tilde{\boldsymbol{V}}\end{cases} \tag{10}$$

式中，⊗表示逐元素相乘。最后，将亮度图$\boldsymbol{I}$与$\boldsymbol{HV}$色彩结构图拼接起来构成HVI色彩空间。

1.5 损失函数

为了适应双分支增强网络以及充分发挥HVI色彩空间和sRGB色彩空间各自的优点，构建了由HVI损失$l(\hat{\boldsymbol{X}}_{HVI},\boldsymbol{Y}_{HVI})$和RGB损失$l(\hat{\boldsymbol{X}}_I,\boldsymbol{Y}_I)$组成的联合损失，具体计算为

$$L=\lambda_c\cdot l(\hat{\boldsymbol{X}}_{HVI},\boldsymbol{Y}_{HVI})+l(\hat{\boldsymbol{X}}_I,\boldsymbol{Y}_I) \tag{11}$$

式中，λ_c是平衡HVI色彩空间与RGB色彩空间的权重，$\hat{\boldsymbol{X}}_{HVI}$是HVI色彩空间下最终增强后的图像，$\boldsymbol{Y}_{HVI}$是HVI色彩空间下的正常光照图像，$\hat{\boldsymbol{X}}_I$是RGB色彩空间下最终增强后的图像，$\boldsymbol{Y}_I$是RGB色彩空间下的正常光照图像，损失$l(\hat{\boldsymbol{X}}_{HVI},\boldsymbol{Y}_{HVI})$和$l(\hat{\boldsymbol{X}}_I,\boldsymbol{Y}_I)$计算方式相同，具体为

$$\begin{aligned}l(\hat{\boldsymbol{X}}_{HVI},\boldsymbol{Y}_{HVI})=\lambda_1\cdot L_1(\hat{\boldsymbol{X}}_{HVI},\boldsymbol{Y}_{HVI})+\\ \lambda_d\cdot L_d(\hat{\boldsymbol{X}}_{HVI},\boldsymbol{Y}_{HVI})+\lambda_e\cdot L_e(\hat{\boldsymbol{X}}_{HVI},\boldsymbol{Y}_{HVI})+\\ \lambda_p\cdot L_p(\hat{\boldsymbol{X}}_{HVI},\boldsymbol{Y}_{HVI})\end{aligned} \tag{12}$$

式中，L_1表示L1损失、L_d表示结构相似性损失(Wang等，2004)、L_e表示边缘损失(Seif和Androutsos，2018)、L_p表示感知损失(Johnson等，2016)，λ_1，λ_d，λ_e，λ_p是各损失的权重系数。

L_1损失计算预测像素值与真实像素值之间绝对差的平均值，引导网络在增强过程中减少像素级误差，从而使输出图像在亮度、对比度等方面更接近真实值。具体计算为

$$L_1=\frac{1}{N}\sum_{n=1}^{N}\left\|\boldsymbol{I}_{\text{enhanced}}-\boldsymbol{I}_{\text{gt}}\right\| \tag{13}$$

式中，N为像素数，$\boldsymbol{I}_{\text{enhanced}}$和$\boldsymbol{I}_{\text{gt}}$分别表示增强后图像与真实光照图像。

L_d损失为结构相似性损失，通过结构相似度(structural similarity index measure，SSIM)指标度量两幅图像之间的结构相似程度，包括亮度、对比度和结构3个维度的相似度。SSIM值越接近1，表示两幅图像的结构相似性越高。具体计算为

$$\begin{cases}L_d=1-SSIM\left(\boldsymbol{I}_{\text{enhanced}},\boldsymbol{I}_{\text{gt}}\right)\\ SSIM(x,y)=\left(\dfrac{2\mu_x\mu_y+c_1}{\mu_x^2+\mu_y^2+c_1}\right)\times\left(\dfrac{2\sigma_{xy}+c_2}{\sigma_x^2+\sigma_y^2+c_2}\right)\end{cases} \tag{14}$$

式中，μ_x和μ_y为两幅图像的均值，σ_x和σ_y为两幅图像的方差。c_1和c_2是两个极小的常数，防止分母为零。

L_e为边缘损失，其目的在于引导网络在增强图像亮度和对比度的同时，保留图像的边缘结构信息。该损失计算预测图像与真实图像在边缘空间上的绝对差，防止增强图像出现边缘模糊。具体计算为

$$L_e=\frac{1}{N}\sum_{n=1}^{N}\left\|G\left(\boldsymbol{I}_{\text{enhanced}}\right)-G\left(\boldsymbol{I}_{\text{gt}}\right)\right\|_1 \tag{15}$$

式中，$G(\cdot)$表示边缘检测算子。

L_p为感知损失，旨在提升高级特征空间中度量图像的相似性。该损失利用预训练的深度卷积网络作为特征提取器，比较两幅图像在这些特征空间中的表达差异。引导增强图像在语义内容、纹理和全局结构上与真实图像保持一致。具体计算为

$$L_p=\frac{1}{C_jH_jW_j}\left\|\phi_j\left(\boldsymbol{I}_{\text{enhanced}}\right)-\phi_j\left(\boldsymbol{I}_{\text{gt}}\right)\right\|_2^2 \tag{16}$$

式中，ϕ_j表示预训练网络的第j个卷积层所输出的特征图，C_j，H_j，W_j分别为该层特征图的通道数、高度和宽度。

使用以上4种损失函数作为最终损失函数，旨在让模型同时实现像素级保真、结构保持、细节增强和视觉自然的功能。

GCDINet的伪代码如下：

输入：低光照图像$\boldsymbol{I}_{\text{low}}$、正常光照图像$\boldsymbol{I}_{\text{normal}}$。

输出：Model_GCDINet。

1)初始化阶段：批大小为8、迭代总次数1 000，采用余弦模拟退火学习率调整策略，初始学习率lr为1E-4。

2)循环训练阶段：

3)for i = 1 to 1 000 do;

4) #前向传播计算预测输出；

5) $\boldsymbol{I}_{\text{enhanced}}=\text{GCDINet}(\boldsymbol{I}_{\text{low}})$；

6) #按式(11)计算损失；

7) $\boldsymbol{I}_{\text{enhanced\_hvi}}=\text{HVT}(\boldsymbol{I}_{\text{enhanced}})$；

8) $\boldsymbol{I}_{\text{normal\_hvi}}=\text{HVT}(\boldsymbol{I}_{\text{normal}})$；

9) $L=L(\boldsymbol{I}_{\text{enhanced}},\boldsymbol{I}_{\text{normal}})+L(\boldsymbol{I}_{\text{enhanced\_hvi}},\boldsymbol{I}_{\text{normal\_hvi}})$；

10） #反向传播计算网络参数梯度；
11） L = backward(L)；
12） #采用Adam算法更新模型参数；
13） Adam(model. paramsters())；
14） #采用余弦退火策略调整学习率lr；
15） If $i \leqslant 250$ lr余弦增加至0. 000 2；
16） Else if $i = 251$ lr重置为 0. 000 1；
17） Else lr 余弦减少至0. 000 000 1；
18）End for。

2 实验结果和讨论

2.1 数据集和设置

采用LOLv1(Chen等,2018)、LOLv2-Real-captured和LOLv2-Synthetic(Yang等,2021)这3个公开且常用的LLIE基准数据集进行评估。LOL数据集有LOLv1和LOLv2两个版本。与既包含真实数据又包含合成数据的LOLv1相比,LOL-v2分为真实子集和合成子集。对与LOLv1和LOLv2-Synthetic,将训练图像裁剪成256×256像素大小,批处理大小设置为8,而后对GCDINet训练1 000轮次。对于LOLv2-Real-captured,将训练图像进行随机旋转、水平翻转以及垂直翻转用于数据增强。批大小设置为8,训练GCDINet 1 000次。

实验设置使用PyTorch实现GCDINet。使用单个NVIDIA RTX 3090 GPU,采用Adam(Kingma和Ba,2015)优化器($\beta_1 = 0.9$和$\beta_2 = 0.999$)对模型进行训练。将初始学习率设置为1×10^{-4},然后在训练过程中使用余弦退火策略(Loshchilov和Hutter,2017)对学习率进行调整,使其稳步降低到1×10^{-7}。

评价指标采用图像评估中常用的峰值信噪比(peak signal-to-noise ratio, PSNR),结构相似度(SSIM)(Wang等,2004)作为失真指标。为了全面评估恢复图像的感知质量,本文以AlexNet(Krizhevsky等,2017)为参考,使用学习感知图像块相似度(learned perceptual image patch similarity, LPIPS)(Zhang等,2018),用于评估图像感知质量。

2.2 实验

为验证GCDINet方法的有效性与优越性,本节通过在多个基准数据集上与主流方法进行定量与定性比较,以及数据增强策略分析两个方面开展实验。首先,通过在LOLv1、LOLv2等多个基准数据集上与主流方法进行客观指标与主观视觉的全面对比,旨在综合评价模型在图像增强质量、细节恢复及色彩保真方面的综合性能,并深入分析其在高效率约束下所实现的卓越性能平衡;此外,针对LOLv2-Real数据集训练中出现的梯度异常问题,专项设计了数据增强策略分析实验,通过系统调整预处理参数以优化训练稳定性,并最终确定最佳配置以提升模型在该场景下的泛化能力与鲁棒性。

2.2.1 主要实验结果

从表1可以看出,在LOLv1数据集上,所提方法在SSIM和LPIPS指标上取得最优效果,在PSNR指标上仅比最优方法和次优方法相差0.766 dB和0.223 dB。在LOLv2-Real数据集上,所提方法在PSNR和SSIM两个指标上均取得最优效果,在LPIPS指标上比最优方法相差0.006。在LOLv2-Synthetic数据集上,所提方法的所有指标都是最优的。另外,所提方法在PSNR、SSIM、LPIPS指标方面优于基于RGB的最佳方法GSAD,而且仅利用GSAD参数量的10%。与基于Retinex的SOTA(state of the art)方法Retinexformer相比,GCDINet在表1中的PSNR指标中没有优于Retinexformer。这主要是因为两种模型设计初衷具有差异,Retinexformer的参数量大、计算复杂度高。这种大的模型容量使其具有较强的拟合能力,因此,能够更精细地重建图像像素,从而在PSNR这种像素级保真度指标上的效果更好。而GCDINet的设计目标之一是性能与效率的平衡,GCDINet的计算复杂度仅为Retinexformer的50%。这种对效率的追求,会限制模型的拟合能力,导致模型在PSNR指标上有一定的牺牲。

可视化的结果如图5所示,结果图下方彩色方框图表示局部细节放大图。现有方法存在的亮度伪影、局部色偏和结构信息丢失等问题,在GCDINet的处理结果中得到了有效的抑制。通过所提的GCFE模块分别单独处理B_I与B_{HV},从而让各分支专注于各自的任务,避免在LLIE过程中因亮度信息与色彩结构信息耦合带来的亮度伪影和色偏问题。由于解码阶段双路自适应特征校准结构的引导,有效避免了因色彩结构信息缺失导致的色偏和伪影。另外,在消融实验部分也验证了GCFE模块和双路径自适应特征校准结构的关键作用。

2.2.2 数据增强策略分析

在训练过程中观察到使用LOLv2-Real数据集

表1　在LOLv1和LOLv2数据集的定量结果比较

Table 1　Quantitative comparison of results on LOLv1 and LOLv2 datasets

方法	颜色模型	参数量/M	FLOPs/G	LOLv1			LOLv2-Real			LOLv2-Synthetic		
				PSNR/dB ↑	SSIM ↑	LPIPS ↓	PSNR/dB ↑	SSIM ↑	LPIPS ↓	PSNR/dB ↑	SSIM ↑	LPIPS ↓
RetinexNet	Retinex	0.840	584.47	16.774	0.419	0.742	16.097	0.410	0.543	22.794	0.840	0.171
KinD	Retinex	8.020	34.99	20.860	0.802	0.160	17.540	0.669	0.375	18.320	0.796	0.252
ZeroDCE	RGB	0.075	4.83	14.861	0.559	0.222	16.059	0.580	0.313	17.712	0.815	0.169
DRBN	Retinex	5.470	48.61	16.290	0.617	0.160	20.290	0.831	/	23.220	0.927	/
RUAS	Retinex	0.003	0.83	16.405	0.500	/	15.326	0.488	0.310	13.765	0.638	0.305
LLFlow	RGB	17.420	358.40	21.149	0.854	0.119	17.433	0.831	0.176	24.807	0.919	0.067
EnlightenGAN	RGB	114.350	61.01	17.480	0.651	0.153	18.230	0.617	0.309	16.570	0.734	0.220
SNR-Aware	SNR+RGB	4.010	26.35	24.610	0.842	/	21.480	0.849	0.163	24.140	0.928	0.056
Bread	YCbCr	2.020	19.85	22.960	0.838	/	20.830	0.847	0.174	17.630	0.919	0.091
PairLIE	Retinex	0.330	20.81	19.510	0.736	0.248	19.885	0.778	0.317	19.074	0.794	0.230
LLFormer	RGB	24.550	22.52	23.649	0.816	0.169	20.056	0.792	0.211	24.038	0.909	0.066
Retinexformer	Retinex	1.530	15.85	**25.153**	0.846	0.131	22.794	0.840	0.171	25.670	0.930	0.059
GSAD	RGB	17.360	442.02	22.770	0.852	0.102	20.153	0.846	**0.113**	24.472	0.929	0.051
QuadPrior	Kubelka-Munk	1 252.750	1 103.20	20.310	0.808	0.202	20.592	0.811	0.202	16.108	0.758	0.114
CIDNet-wP	HVI	1.880	7.57	23.809	0.857	0.101	23.690	0.861	0.135	25.049	**0.935**	0.048
GCDI (本文)	HVI	2.050	7.94	24.387	**0.858**	**0.096**	**23.891**	**0.866**	0.119	**25.698**	**0.935**	**0.047**

注：加粗字体表示各列最优结果。“/”表示未在该数据上报告结果。浮点运算次数FLOPs（floating point operations）在单幅256×256像素图像上测试。“↑”与“↓”分别表示值越高越好、越低越好。

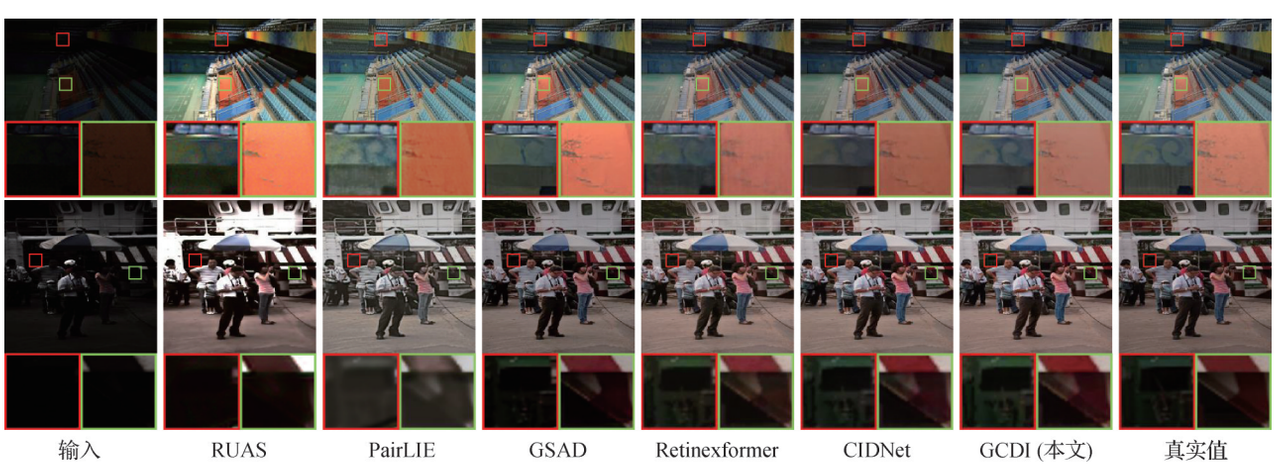


图5　可视化结果图

Fig. 5　Visualization comparison of enhancement results

会出现梯度异常问题，通过数据分布分析发现：该数据集虽然包含不同光照条件的配对样本，但场景同构性过高，导致输入数据的色彩空间分布呈现高度一致性。这种数据冗余现象容易引发参数更新过程中的梯度方向趋同，进而导致训练过程中梯度爆炸问题。为改善该问题，本文提出基于随机几何变换的自适应数据增强策略。具体而言，在模型训练之前，先根据一定的概率P对数据进行水平/垂直翻转或者随机角度旋转的预处理操作。通过自适应数据增强策略，有效缓解了LOLv2-Real数据集训练过程

中的梯度爆炸问题。

为了确定自适应增强策略的最佳参数，对LOLv2-Real数据集的预处理参数进行分析。如表2所示，当旋转/翻转概率P = 0.30时，模型在测试集取得最优指标。PSNR达到23.891 dB，比次优值提升0.434 dB，SSIM提升至0.866，同时LPIPS指标下降至0.119，表明增强后的图像在结构相似性和感知质量方面均有显著改善。值得注意的是，当P > 0.30时各指标出现明显波动（SSIM下降0.46%，LPIPS上升0.84%），这可能源于过度的空间变换破坏了低光照区域的噪声分布特性，导致模型学习到伪影模式。

2.3 消融研究

为深入验证本文所提GCDINet中各个核心组件设计的必要性与有效性，设计了3个层次的消融实验。首先，通过模块消融实验系统评估了MHSA模块、GCM模块以及双路径自适应特征校准结构对模型性能的独立与联合贡献，旨在定量与定性地揭示各模块的功能；其次，通过色彩空间对比实验，旨在分析HVI色彩空间相较于传统HSV、YCbCr空间在解耦亮度与色彩信息、抑制伪影方面的固有优势；最后，通过损失函数分析实验，旨在验证联合使用HVI空间与sRGB空间损失函数的必要性，并辅以训练损失曲线证明其优化过程的稳定性。为了快速收敛和稳定的性能，实验都是在LOLv2-Synthetic数据集上进行的。

表2 在LOLv2-Real数据集上预处理参数分析

Table 2 Analysis of preprocessing parameters on the LOLv2-Real dataset

旋转/翻转概率P	PSNR/dB ↑	SSIM ↑	LPIPS ↓
0.15	23.205	0.850	0.138
0.20	23.311	0.861	**0.119**
0.25	23.329	0.862	0.121
0.30	**23.891**	**0.866**	**0.119**
0.35	23.457	0.862	0.120

注：加粗字体表示各列最优结果，“↑”表示值越高越好，“↓”表示值越低越好。

2.3.1 模块消融实验分析

使用定量结果（表3）和可视化结果（图6）验证了GCDNet中的关键模块。图6中的“△”表示双路径自适应特征校准结构，结果图下方彩色方框图表示局部细节放大图。

表3 GCDINet的主要模块消融实验结果

Table 3 Ablation study results of main modules in GCDINet

U-Net Baseline	MHSA	GCM	△	PSNR/dB ↑	SSIM ↑	LPIPS ↓
√	–	–	–	15.399	0.730	0.207
√	√	–	–	24.979	0.929	0.053
√	√	√	–	25.456	0.931	0.048
√	√	–	√	25.176	0.928	0.053
√	√	√	√	**25.698**	**0.935**	**0.047**

注：加粗字体表示各列最优结果。“△”表示双路径自适应特征校结构。“√”和“–”分别表示采用和未采用，“↑”表示值越高越好，“↓”表示值越低越好。

表3展示了所提方法主要模块的消融实验结果，其中，第1行表示未使用注意力机制和特征增强模块的U-Net。将MHSA模块加入基础U-Net网络中，PSNR、SSIM、LPIPS这3个指标分别提升62.2%、27.2%和74.4%，这说明注意力机制在双分支增强网络中具有很大的应用潜力。通过图6中MHSA模块的可视化结果可以看到，MHSA模块通过建立长程依赖关系，有效恢复图像的细节纹理。

对比表3中第1、3行的实验结果显示，在MHSA模块的基础上使用GCM模块组成GCFE模块，3个指标分别提升了65.3%、27.5%和76.8%。从图6中的GCFE模块的可视化结果可以看到，MHSA与GCM模块的结合重点优化亮度分布的均衡性，有效缓解了亮度调整带来的亮度伪影和色彩保真度不足问题，验证了所设计GCFE模块的有效性。

对比表3第2、4行的实验结果显示，在MHSA模块的基础上使用双路径自适应特征校准结构，3个指标变化很小。但从图6中使用双路径自适应特征校准结构的可视化结果来看，图像的色彩结构信息更加丰富，亮度重建更加接近Ground Truth。3个指标没有明显提升说明双路径自适应特征校准结构与MHSA模块并没有很好的适配效果，可视化结果验证了所提双路径自适应特征校准结构的有效性。

对比表3中第1、5行的实验结果显示，同时使用双路径自适应特征校准结构和GCFE模块，融合各模块优势得到最佳的模型性能，3个指标分别提升

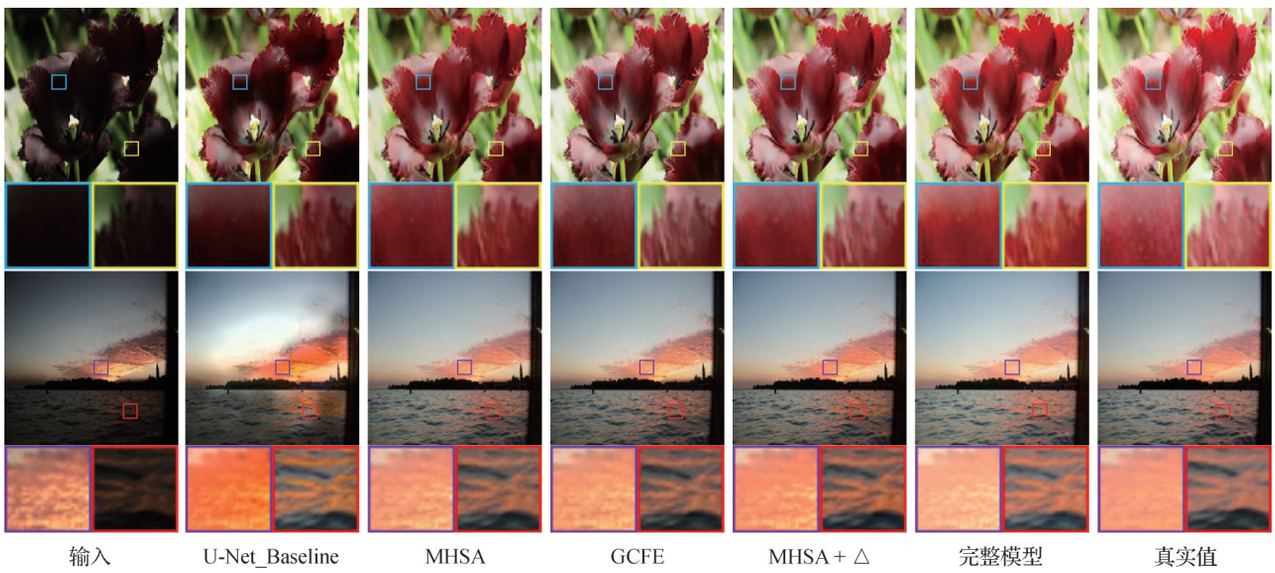


图6　模块消融可视化结果图

Fig. 6　Visual ablation study of core modules

66.8%、28.1%、77.3%。图6显示，全模型不仅在亮度分布（红色方框）和色彩恢复（蓝色方框）方面效果最佳，同时完整地保留了边缘纹理（黄色方框）和背景光照的渐进式过渡（紫色方框）。通过量化和可视化的结果验证了所提方法的有效性。

2.3.2　MHSA模块注意力头数分析

为优化模型效率并验证MHSA模块中注意力头数设置对模型性能的影响，进行了注意力头数的对比实验，结果如表4所示。“2-2-2”，“4-4-4”，“8-8-8”和“2-4-8”分别表示3次使用MHSA模块的注意力计算头数数量。

表4　MHSA模块多头数量对比实验结果

Table 4　Comparative experimental results of multi-head numbers for MHSA module

多头数量	PSNR/dB ↑	SSIM ↑	LPIPS ↓
2-2-2	25.137	0.921	0.069
4-4-4	25.161	0.928	0.052
8-8-8	25.274	0.928	0.051
2-4-8	**25.698**	**0.935**	**0.047**

注：加粗字体为各列最优值，“↑”表示值越高越好，“↓”表示值越低越好。

对比表4中第1、2行的实验结果显示，将头数配置从2-2-2提升至4-4-4，模型的3个评价指标分别提升了0.1%、0.8%和24.6%。这说明适度增加注意力头数可以有效增强模型的特征捕捉能力，从而显著提升性能，尤其是在感知质量（LPIPS）上改善明显。

对比表4中第2、3行的实验结果显示，继续将头数配置从4-4-4提升至8-8-8，模型的PSNR指标提升了0.4%，但SSIM指标没有变化，LPIPS指标仅提升了1.9%。这表明在网络各阶段均使用大量注意力头会面临性能提升有限，仅依靠增加计算资源并不能让模型性能对应提高。

对比表4中第3、4行的实验结果显示，使用2-4-8注意力头数配置，模型的PSNR、SSIM和LPIPS指标相较于效果最好的8-8-8头数配置分别提升了1.7%、0.8%和7.8%。这是由于2-4-8注意力头数配置使模型能够在浅层、中层和深层分别适配最合适的感受野，因此该头数配置有效提升了模型性能。

2.3.3　色彩空间分析

为了比较模型使用不同色彩空间的性能差异，将所提方法使用的色彩空间与流行的两种色彩空间在3个评价指标上进行定量比较，并使用不同色彩空间进行可视化结果对比。实验结果如表5和图7所示，其中，图7结果图下方彩色方框图表示局部细节放大图。

从表5中第1行的实验结果可以看到，模型使用HSV色彩空间时，3个评价指标数值分别为21.601 dB、0.871和0.132；从图7中使用HSV色彩空间的可视化结果也可以看到，增强后的图像在色差和亮度方面存在严重的偏差。这也对应上文所述，虽然HSV

表 5　模型使用不同色彩空间的定量比较

Table 5　Quantitative comparison of the model using different color spaces

色彩空间	PSNR/dB ↑	SSIM ↑	LPIPS ↓
HSV	21.601	0.871	0.132
YCbCr	24.488	0.915	0.074
HVI	**25.698**	**0.935**	**0.047**

注：加粗字体表示各列最优结果，“↑”表示值越高越好，“↓”表示值越低越好。

色彩空间可以将亮度与色彩解耦，但在解耦过程中由于 HSV 色彩空间带有的红色不连续和黑色平面噪声，导致了引入更多伪影（黄色方框）和黑色噪声（红色和绿色方框）的缺陷。

对比表 5 第 1、2 行的实验结果显示，仅将 HSV 色彩空间换为 YCbCr 色彩空间，模型的 3 个评价指标分别提升 13.4%、5% 和 43.9%。从图 7 中使用了 YCbCr 色彩空间的可视化结果可以看到，相较于 HSV 色彩空间增强后的图像在色彩保真度（红色方框和绿色方框）方面有很大提升，表明 YCbCr 色彩空间缓解了 HSV 色相维度不连续的问题。

对比表 5 第 1、3 行的实验结果显示，仅将 HSV 色彩空间替换为 HVI 色彩空间，所得模型的 3 个评价指标分别提升 18.9%、7.3% 和 64.4%。同时，从图 7 使用 HVI 色彩空间（full model）的可视化实验结果可以看到，使用 HVI 色彩空间增强后的图像更接近 Ground Truth。可以验证，使用 HVI 色彩空间可以有效解耦 LLIE 过程中的亮度与色彩信息。

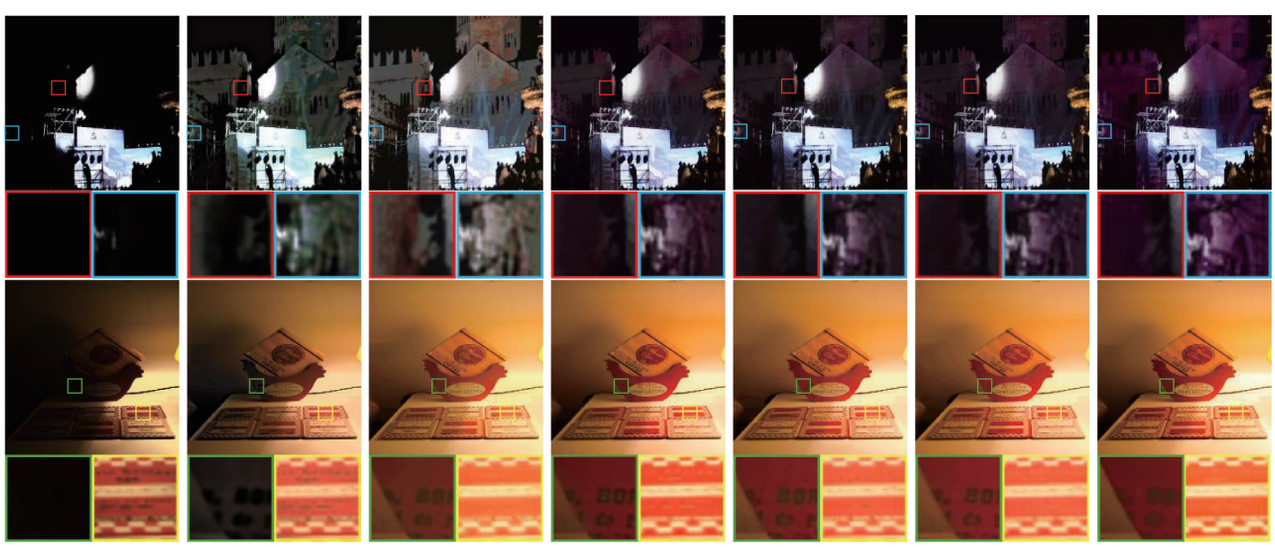


图 7　色彩模型和损失函数消融可视化结果图

Fig. 7　Ablation visualization of color models and loss functions

2.3.4　损失函数分析

表 6 展示了 GCDINet 所使用损失函数的消融实验结果。

对比表 6 中第 1、3 行的实验结果显示，与同时使用 HVI 和 sRGB 损失相比，仅依靠 HVI 损失所得模型的 PSNR 和 SSIM 指标分别下降 1.6% 和 0.2%，但是 LPIPS 指标并没有变化。从图 6 仅使用 HVI 损失的可视化实验结果也可看到，从人眼感知的角度很难观察到与使用两个损失函数模型增强结果的差别，这也对应了表 5 中 LPIPS 指标没有变化的结果。

对比表 6 中第 2、3 行的实验结果显示，仅使用 sRGB 损失所得模型的 3 个指标分别下降 4.8%，1.5% 和 17.5%。从图 7 中仅使用 sRGB 损失的可视

表 6　GCDINet 的损失函数消融实验结果

Table 6　Ablation study results of loss functions in GCDINet

损失	PSNR/dB ↑	SSIM ↑	LPIPS ↓
HVI Only	25.279	0.933	**0.047**
sRGB Only	24.512	0.921	0.057
同时使用	**25.698**	**0.935**	**0.047**

注：加粗字体表示各列最优结果，“↑”表示值越高越好，“↓”表示值越低越好。

化实验结果来看，单独使用sRGB损失函数的像素级误差和色彩分布无法直接观察到。

为了进一步验证使用HVI损失和sRGB损失作为联合损失函数在训练过程中的有效性和稳定性，本文绘制了联合损失函数随训练迭代次数的变化曲线，如图8所示。联合损失函数的值随着迭代次数的增加呈现出显著且平滑的下降趋势。损失值从初始的1.0迅速下降，并在约400个迭代周期后进入相对平稳期，最终损失值收敛至0.2附近。这表明所提GCDINet网络能够有效地从训练数据中学习，并且所设计的联合损失函数为模型优化提供了稳定、有效的指导。

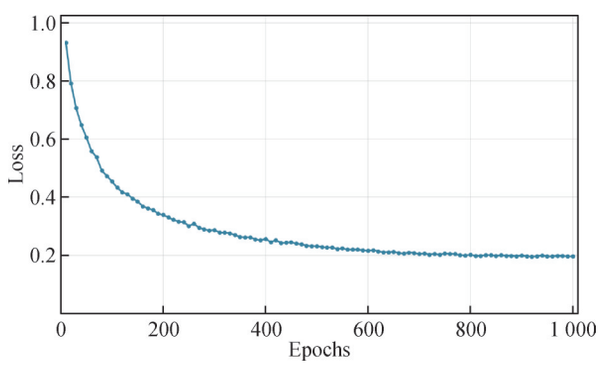


图8　损失函数随训练迭代次数的变化曲线

Fig. 8　The variation curve of the joint loss function with training epochs

另外，整个损失曲线平滑，未出现剧烈的振荡或回升现象。这证明了所使用的训练策略、优化器选择和学习率设置是有效的，模型训练过程稳定，没有出现梯度爆炸或难以收敛等问题。

3　结　论

本文基于HVI色彩空间构建的GCDINet方法解决了LLIE过程中存在的亮度伪影扩散与色彩失真相互耦合的问题。首先通过引入HVI色彩空间解耦图像亮度与色彩信息的方式，突破了传统色彩空间存在的色相不连续与通道耦合限制；然后使用双分支门控卷积机制，实现了亮度增强路径与色彩修复路径的独立建模；进而将通道感知特征校准模块嵌入U型网络跨层连接通路，提出了双路径自适应特征校准结构，解决了双分支网络特征重建中通道权重失衡的问题；最后利用双路径自适应特征校准结构的通道感知机校准功能，成功指导了解码阶段的跨阶段特征融合。实验表明，在计算资源较少的情况下，本文方法与15种先进的主流方法相比，在LOLv1、LOLv2等标准数据集上的多个指标表现出了优越的性能，均取得了最优或次优结果。同时，本文方法具备参数量少的优势，本文方法的计算量相比先进的Retinexformer方法减少了1/2，是一种轻量化的LLIE解决方案。

通过理论分析发现，当输入图像整体亮度极低时，亮度通道B_I的数值普遍接近0，这将导致亮度图$\boldsymbol{I}$无法准确计算，进而使得算法性能在极端低光场景下受到限制。因此，未来将着重研究极端低光环境下图像亮度和色彩恢复的有效方法和措施，解决亮度通道数值极低甚至为零情况下的局部亮度和色彩有效引导及重建问题。

参考文献（References）

Brateanu A, Balmez R, Avram A, Orhei C and Ancuti C. 2025. LYT-NET: lightweight YUV transformer-based network for low-light image enhancement. IEEE Signal Processing Letters, 32: 2065-2069 [DOI: 10.1109/LSP.2025.3563125]

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]

Chen W, 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. Newcastle, UK: BMVA Press: 155-166

Chen Y Q, Wen C L, Liu W F and He W. 2023. DBENet: dual-branch brightness enhancement fusion network for low-light image enhancement. Electronics, 12(18): #3907 [DOI: 10.3390/electronics12183907]

Dudhane A, Zamir S W, Khan S, Khan F S and Yang M H. 2025. Burst image restoration and enhancement. IEEE Transactions on Pattern Analysis and Machine Intelligence, 47(11): 9454-9467 [DOI: 10.1109/TPAMI.2024.3356188]

Fu X Y, Liao Y H, Zeng D L, Huang Y, Zhang X P and Ding X H. 2015. A probabilistic method for image enhancement with simultaneous illumination and reflectance estimation. IEEE Transactions on Image Processing, 24(12): 4965-4977 [DOI: 10.1109/TIP.2015.2474701]

Gao Z H and Zhai Y S. 2022. Image dehazing based on multi-scale retinex and guided filtering//Proceedings of 2022 International Confer-

ence on Image Processing, Computer Vision and Machine Learning. Xi'an, China: IEEE: 123-126 [DOI: 10.1109/ICICML57342.2022.10009889]

Goodfellow I J, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. 2014. Generative adversarial nets//Proceedings of the 28th International Conference on Neural Information Processing Systems. Montreal, Canada: MIT Press: 2672-2680

Gevers T, Gijsenij A, Van de Weijer J and Geusebroek J M. 2012. Color in computer vision: fundamentals and applications. Hoboken: John Wiley & Sons, 1-300 [DOI: 10.1002/9781118350089]

Guo X J and Hu Q M. 2023. Low-light image enhancement via breaking down the darkness. International Journal of Computer Vision, 131(1): 48-66 [DOI: 10.1007/s11263-022-01667-9]

Ho J, Jain A and Abbeel P. 2020. Denoising diffusion probabilistic models//Proceedings of the 34th International Conference on Neural Information Processing Systems. Vancouver, Canada: Curran Associates Inc.: #574

Hu H F and Li F Y. 2023. Dual attention fusion generative adversarial network for underwater image enhancement. Journal of Guilin University of Electronic Technology, 43(5): 371-380 (胡海峰,李凤英. 2023. 一种双注意力融合生成对抗网络的水下图像增强模型. 桂林电子科技大学学报, 43(5): 371-380) [DOI: 10.3969/1673-808X.2022339]

Hu J, Shen L and Sun G. 2018. Squeeze-and-excitation networks//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE: 7132-7141 [DOI: 10.1109/CVPR.2018.00745]

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]

Johnson J, Alahi A, Li F F. 2016. Perceptual losses for real-time style transfer and super-resolution. Lecture Notes in Computer Science (ECCV 2016), 9906: 694-711 [DOI: 10.1007/978-3-319-46475-6_43]

Kim S E, Jeon J J and Eom I K. 2016. Image contrast enhancement using entropy scaling in wavelet domain. Signal Processing, 127: 1-11 [DOI: 10.1016/j.sigpro.2016.02.016]

Kingma D P and Ba J. 2015. Adam: a method for stochastic optimization//Proceedings of the 3rd International Conference on Learning Representations. San Diego, USA: ICLR: 1-15

Krizhevsky A, Sutskever I and Hinton G E. 2017. ImageNet classification with deep convolutional neural networks. Communications of the ACM, 60(6): 84-90 [DOI: 10.1145/3065386]

Land E H. 1977. The retinex theory of color vision. Scientific American, 237(6): 108-129

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. Nashville, USA: IEEE: 10556-10565 [DOI: 10.1109/CVPR46437.2021.01042]

Loshchilov I and Hutter F. 2017. SGDR: stochastic gradient descent with warm restarts//Proceedings of the 5th International Conference on Learning Representations. Toulon, France: OpenReview.net: 1-16

Loza A, Bull D and Achim A. 2010. Automatic contrast enhancement of low-light images based on local statistics of wavelet coefficients//Proceedings of 2010 IEEE International Conference on Image Processing. Hong Kong, China: IEEE: 3553-3556 [DOI: 10.1109/ICIP.2010.5651173]

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]

Moran S, Marza P, McDonagh S, Parisot S and Slabaugh G. 2020. DeepLPF: deep local parametric filters for image enhancement//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE: 12823-12832 [DOI: 10.1109/CVPR42600.2020.01284]

Park S, Yu S, Moon B, Ko S and Paik J. 2017. Low-light image enhancement using variational optimization-based retinex model. IEEE Transactions on Consumer Electronics, 63(2): 178-184 [DOI: 10.1109/TCE.2017.014847]

Poynton C. 2003. Image digitization and reconstruction. In: Digital Video and HDTV. San Francisco: Morgan Kaufmann, 187-194 [DOI: 10.1016/B978-155860792-7/50020-8]

Seif G and Androutsos D. 2018. Edge-based loss function for single image super-resolution//Proceedings of 2018 IEEE International Conference on Acoustics, Speech and Signal Processing. Calgary, Canada: IEEE: 1468-1472 [DOI: 10.1109/ICASSP.2018.8461664]

Wang K and Huang F Z. 2021. Multiscale retinex image enhancement in HSV space based on illumination compensation. Laser and Optoelectronics Progress, 59(10): #1010004 (王奎, 黄福珍. 2022. 基于光照补偿的HSV空间多尺度Retinex图像增强. 激光与光电子学进展, 59(10): #1010004) [DOI: 10.3788/LOP202259.1010004]

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: New Orleans, USA: IEEE: 5901-5910 [DOI: 10.1109/CVPR52688.2022.00581]

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

Wang Z, Bovik A C, Sheikh H R and Simoncelli E P. 2004. Image quality assessment: from error visibility to structural similarity IEEE

Transactions on Image Processing, 13(4): 600-612 [DOI: 10.1109/TIP.2003.819861]

Xiao L M, Li C, Wu Z Z and Wang T. 2016. An enhancement method for X-ray image via fuzzy noise removal and homomorphic filtering. Neurocomputing, 195: 56-64 [DOI: 10.1016/j. neucom. 2015.08.113]

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. 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, Zhang Z W and Cheng H X. 2022. PNet: multi-level low-illumination image enhancement network based on attention mechanism. Application Research of Computers, 39(5): 1579-1585 (杨微, 张志威, 成海秀. 2022. PNet: 融合注意力机制的多级低照度图像增强网络. 计算机应用研究, 39(5): 1579-1585) [DOI: 10.19734/j.issn.1001-3695.2021.09.0384]

Yang W H, Wang W J, Huang H F, Wang S Q and Liu J Y. 2021. Sparse gradient regularized deep retinex network for robust low-light image enhancement. IEEE Transactions on Image Processing, 30: 2072-2086 [DOI: 10.1109/TIP.2021.3050850]

Yi X P, Xu H, Zhang H, Tang L F and Ma J Y. 2023. Diff-Retinex: rethinking low-light image enhancement with a generative diffusion model//Proceedings of 2023 IEEE/CVF International Conference on Computer Vision. Paris, France: IEEE: 12268-12277 [DOI: 10.1109/ICCV51070.2023.01130]

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]

Zhang S, Wang T, Dong J Y and Yu H. 2017. Underwater image enhancement via extended multi-scale Retinex. Neurocomputing_, 245: 1-9 [DOI: 10.1016/j.neucom.2017.03.029]

Zhang R, Isola P, Efros A A, Shechtman E and Wang O. 2018. The unreasonable effectiveness of deep features as a perceptual metric//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE: 586-595 [DOI: 10.1109/CVPR.2018.00068]

Zhou L B, Chen X J, Ye B S, Jiang X L, Zou S, Ji L, et al. 2025. A low-light image enhancement method based on HSV space. The Imaging Science Journal, 71(3): 16-29 [DOI: 10.1080/13682199.2023.2266308]

作者简介

张奇,男,硕士研究生,主要研究方向为深度学习与低光照图像增强。E-mail: 19901411896@163.com

雷小舟,通信作者,男,讲师,主要研究方向为计算机视觉与机器学习。E-mail: sigmoid@qq.com

苗壮,男,教授,主要研究方向为图像视频处理。E-mail: emiao_beyond@163.com

王家宝,男,副教授,主要研究方向为计算机视觉与机器学习。E-mail: jiabao_1108@163.com

纪文宇,男,硕士研究生,主要研究方向为深度学习与计算机视觉。E-mail: wenyuji_edu@foxmail.com

毕翔鹤,男,硕士研究生,主要研究方向为深度学习与图像融合。E-mail: 614307276@qq.com