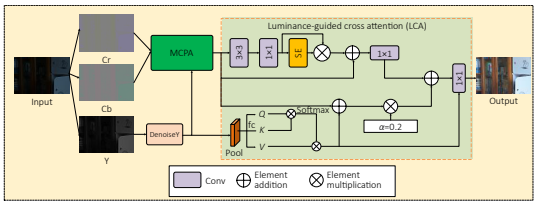
OEE | 光电工程 Opto-Electronic Engineering

Article

2026 年 4 月, 第 53 卷, 第 4 期

结合多通道并行注意力与交叉融合的低光照图像增强



程　浩[1,2], 林坚普[1,2*], 孙　磊[2,3], 张永爱[1,2], 林珊玲[1,2], 郭太良[2], 林志贤[1,2]

[1]福州大学先进制造学院，福建 泉州 362200；
[2]中国福建光电信息科学与技术实验室，福建 福州 350116；
[3]福州大学至诚学院，福建 福州 350002

摘要：针对现有低光照图像增强方法普遍存在模型复杂度高、部署成本大，以及亮度与色彩语义耦合、融合伪影等问题，限制了其在资源受限设备上的应用。为此，本文提出一种结合多通道并行注意力与交叉融合的轻量化低光照图像增强网络。该方法在 YCbCr 颜色空间中分别建模亮度与色彩信息，通过轻量级去噪模块与多头注意力机制提取光照结构特征，同时设计多通道并行注意力模块提升色彩上下文建模能力。此外，引入亮度引导的交叉融合模块，在通道-空间注意力联合优化下实现结构与色彩的协同增强。实验结果表明，该网络在 LOLv1、LOLv2 和 LSRW 数据集上分别实现 PSNR 25.66 dB/SSIM 0.84、PSNR 24.61dB/SSIM 0.85 和 PSNR 20.26 dB/SSIM 0.57，优于 Retinexformer 等方法，且参数量仅 0.059 M，FLOPs 为 10.06 G。

关键词：低照度图像增强；亮度-色彩解耦；注意力机制；多头自注意力；图像感知质量；轻量化网络

The English title and abstract appear at the end of the paper.

DOI: 10.12086/oee.2026.250244 | CSTR: 32245.14.oee.2026.250244 | 中图分类号：TP391.4 | 文献标识码：A

程浩，林坚普，孙磊，等. 结合多通道并行注意力与交叉融合的低光照图像增强 [J]. 光电工程，2026，**53**(4): 250244

Cheng H, Lin J P, Sun L, et al. Low-light image enhancement combined with multi-channel parallel attention and cross-fusion[J]. *Opto-Electron Eng*, 2026, **53**(4): 250244

1　引　言

在复杂光照环境下，图像采集设备获取的视觉图片往往存在严重的亮度衰减、对比度降低、噪声干扰以及细节丢失等问题。这些问题不仅显著降低了图像的视觉感知质量，更可对计算机视觉系统的性能产生严重影响，包括但不限于目标检测准确率下降、图像分割精度降低以及特征提取困难等。因此，开展低光照图像增强研究，探索有效的图像质量提升方法，恢复潜在的细节信息和空间结构特征，对于推动计算机视觉领域的发展具有重要意义[1]。

随着深度学习的迅速发展，基于数据驱动的方法逐渐成为低光照图像增强领域的主流。相较于传统图像处理方法，深度学习模型能够从大量图像中学习复杂的光照映射关系，在保持图像自然感的同时显著提升亮度与细节还原能力。目前，深度学习方法主要可分为两大类：基于 Retinex 理论的方法和非 Retinex 方法。

基于 Retinex 理论的方法源自经典的图像分解思想，其核心在于将图像拆分为照明分量与反射分量，通过调

收稿日期：2025-08-18；修回日期：2025-12-16；录用日期：2026-01-04；网络出版日期：2026-04-25
基金项目：国家重点研发计划 (2023YFB3609400)；国家自然科学基金青年科学基金资助项目 (62101132)
*通信作者：林坚普，ljp@fzu.edu.cn

节照明分量提升整体亮度，同时保持反射细节的稳定性[2]。Wei 等人提出的 RetinexNet 首次实现端到端的图像分解与增强，取得了良好的视觉效果[3]。随后，KinD 系列网络引入噪声建模与颜色修复机制，进一步提升了在复杂低照度场景下的增强鲁棒性[4]。为加强全局建模能力，Retinexformer[5]、Diff-Retinex++[6] 等将 Transformer 与扩散建模方法引入 Retinex 框架，在细节恢复与视觉感知质量方面展现出强大优势。

非 Retinex 方法则摒弃了图像物理分解假设，采用端到端的学习方式直接建立低照度图像到高质量图像的映射。EnlightenGAN 基于生成对抗网络，在无监督设定下展现了出色的性能[7]。SCI 强调网络的轻量化与高效性，在移动设备上具备良好部署优势[8]。HVIMamba 则通过引入 HVI 颜色空间与状态空间建模策略，结合交叉注意力机制，有效缓解了传统 RGB 空间方法在色彩失真与伪影方面的问题[9]。

尽管现有低照度图像增强方法在提升图像亮度和细节恢复方面取得了显著进展，但仍面临三大核心挑战。首先，多数方法依赖复杂深层网络结构，计算开销大、参数冗余，难以在移动设备等资源受限平台上高效部署。其次，亮度与色彩在图像中承载不同的语义信息，传统网络往往采用统一建模方式，忽视二者的语义差异，导致在复杂光照条件下增强图像出现色彩失真或结构模糊等问题。最后，现有方法在特征融合阶段通道间耦合过强，缺乏有效的信息选择机制，容易引入冗余信息和融合伪影，进而影响增强图像的整体质量与视觉一致性。以上问题限制了现有方法在实际应用中的性能与适用性，尤其在追求轻量化、高质量增强效果的场景中亟待解决。

针对上述问题，本文受到文献 [10] 启发，提出了一种融合 Transformer 思想的低照度图像增强网络。该网络在 YCbCr 颜色空间下采用亮度与色彩解耦的双分支结构，通过引入全局建模机制与亮度引导的跨通道交互，实现结构信息与颜色信息的独立感知与协同增强。主要创新包括：

1) 针对 YCbCr 颜色空间中 Y 通道承担亮度结构、Cb/Cr 通道表征色彩特征的属性差异，设计亮度去噪分支与色彩修复分支构成的双分支网络结构，分别对 Y 分量与 Cb/Cr 分量建模，从而实现亮度恢复与色彩恢复的功能协同。

2) 提出多通道并行注意力模块 (multi-channel parallel attention, MCPA)，嵌入色彩路径通道，用于提取不同尺度下的上下文信息并进行特征重加权，提升网络对图像多尺度结构与色彩变化的自适应建模能力。

3) 设计了亮度引导的交叉融合模块 (luminance-guided cross attention，LCA)，在特征重建阶段融合通道注意力与多头自注意力 (multi-head self-attention, MSA)，实现结构与色彩的联合建模。该模块通过 Y 分支提供的结构信息引导色彩特征的动态调整，有效提升增强图像的细节保真度与色彩一致性。

2 本文方法

本节首先介绍整体网络结构，其次详细说明 YCbCr 颜色空间、多通道并行注意力模块 (MCPA) 以及亮度引导的交叉融合模块 (LCA)，最后介绍本文所采用的损失函数设计。

2.1 网络结构

图 1 展示了本文提出的整体架构。该网络基于亮度 – 色彩解耦思想，在 YCbCr 颜色空间下分别对亮度信息 (Y) 与色彩信息 (Cb/Cr) 进行建模，以实现结构与色彩的协同增强。

输入图像首先被转换至 YCbCr 空间，分离出 Y、Cb 和 Cr 三个通道。Y 通道经过 DenoiseY 模块去除噪声，并通过多头自注意力模块 (MSA) 提取结构特征。为了增强上下文建模能力，MSA 前添加池化操作以压缩 Token 数量。同时，Cb 和 Cr 通道经 MCPA 模块进行多尺度注意力增强，提取丰富的色彩上下文。随后，在 LCA 模块中，降噪后的 Y 通道作为引导特征与 Cb/Cr 通道融合。该模块包含通道注意力 (SE) 与逐元素交互机制，实现亮度引导下的结构-色彩融合，融合强度由参数 $\alpha = 0.2$ 控制。

最终融合后的特征经卷积层还原至 RGB 空间，得到增强图像。本文通过亮度与色彩的分支建模与交叉引导，有效提升了低照图像在细节、光照与色彩层面的整体质量。

2.2 YCbCr 颜色空间

YCbCr 颜色空间是一种常用于图像和视频处理的颜色代码模型，它通过将图像中的光彩信息分解为亮度分量 Y 和色彩分量 Cb、Cr，实现了光照和色彩特征的解

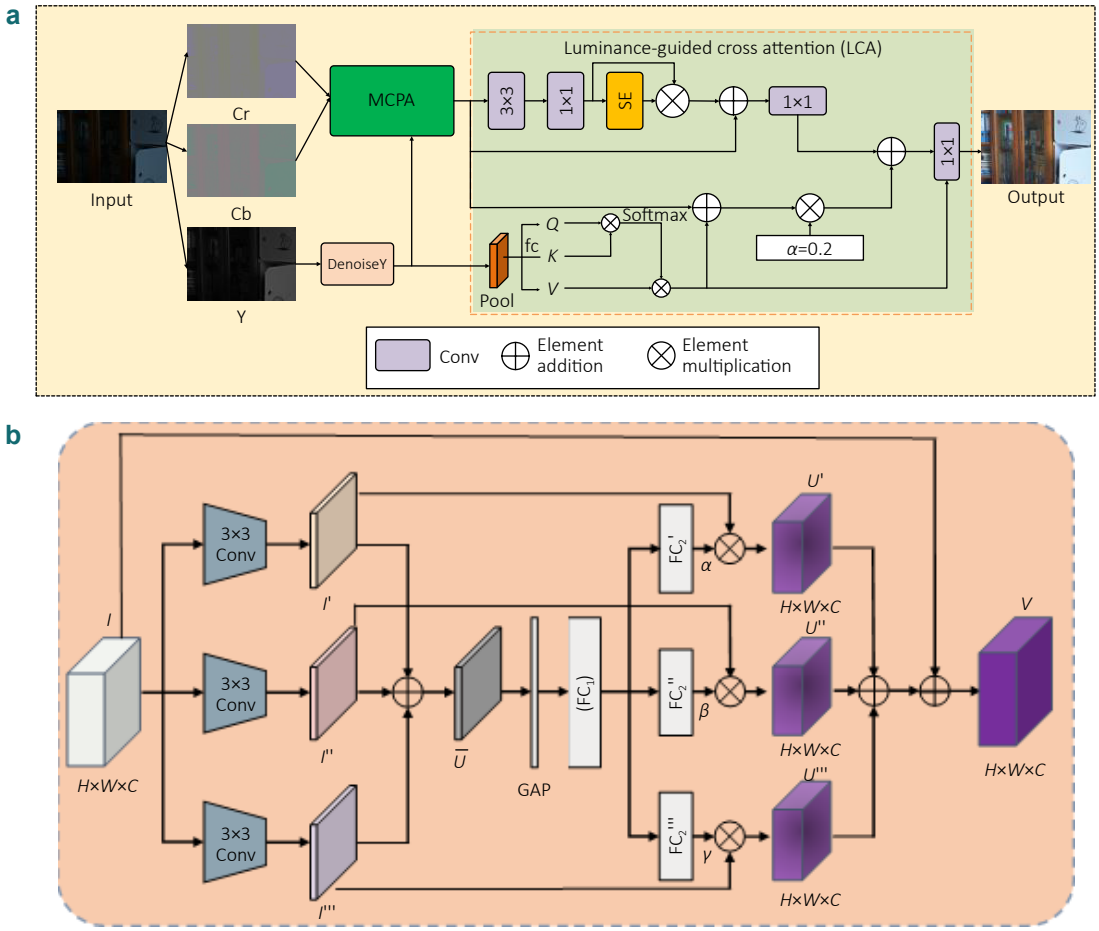


图 1 所提方法的整体网络框架。(a) 整体网络结构图；(b) DenoiseY 网络结构图

Fig. 1 Overall framework of the proposed method. (a) Overall network architecture diagram; (b) DenoiseY network architecture diagram

耦表示[11]，该颜色空间通常通过以下公式将 RGB 颜色值转换为 YCbCr 表示：

$$\begin{cases} Y(i,j) = 0.299R(i,j) + 0.587G(i,j) + 0.114B(i,j) \\ Cb(i,j) = -0.168736R(i,j) - 0.331264G(i,j) \\ \qquad\quad + 0.5B(i,j) + 0.5 \\ Cr(i,j) = 0.5R(i,j) - 0.418688G(i,j) \\ \qquad\quad - 0.081312B(i,j) + 0.5 \end{cases}, \quad (1)$$

式中：(i,j)表示图像中像素的位置；$R(i,j)$、$G(i,j)$、$B(i,j)$分别表示该像素点在 RGB 三个颜色通道上的强度值；$Y(i,j)$为亮度分量，反映了图像的灰度信息和结构细节，是人眼感知中最为敏感的部分；$Cb(i,j)$，$Cr(i,j)$分别代表蓝色色度差和红色色度差，用于描述图像的色彩信息。

Y 通道对应亮度特征，描述图像的基本光照和结构线索；Cb 和 Cr 通道对应色彩偏移维度。由于人眼对图像内容的敏感性主要集中于亮度信息，因此将亮度与色彩解耦可以显著提升图像处理中对不同属性的感知精度与优化空间。YCbCr 分离处理能高效表达图像的光照和色彩信息。在不同光照强度、颜色变化或噪点传播的环境中，YCbCr 保持了良好的抗干扰性和处理稳定性，特别适用于低光照图像增强等应用场景。

在本文中将 Y、Cb、Cr 通道分别形成亮度和色彩分支，进行对应的特征应对和优化。Y 分支根据光照结构进行先驱处理，Cb/Cr 分支则重点对色彩偏移进行选择性维护。这种基于 YCbCr 的解耦网络构筑，有效降低了通道间信息耦混并提升了后续融合效果的维度符合性和对齐性。

2.3 多通道并行注意力模块

为了提升网络对不同通道特征的结构细节与上下文语义的适应建模能力，本文提出了多通道并行注意力模块 (MCPA)，以应对亮度与色彩分支中多尺度信息融合不足的问题。MCPA 模块的设计灵感来自文献 [12]，其主要思想是在保持输入特征结构不变的前提下，从多个

感受野尺度提取不同粒度的上下文信息，并通过注意力机制进行动态融合。

设输入特征图为 X，将其分别送入多个不同尺度的下采样分支中，每个分支使用不同的池化尺度$s\in\{1,2,3,6\}$，构造金字塔式的多尺度特征表示：

$$F_i = Upsample(\sigma(Conv_{1\times1}(Norm(Pool_s(X))))) , \quad (2)$$

式中：$Pool_s(\cdot)$表示自适应平均池化至$s\times s$尺寸；$Norm(\cdot)$为实例归一化层；$Conv_{1\times1}(\cdot)$为通道变换操作；σ为非线性激活函数；$Upsample$ 表示上采样至原始空间大小。最后所有尺度分支输出与原始输入共同进行加权融合：

$$X' = \sum_{i=1}^{S} \alpha_i F_i + \alpha_0 X , \quad (3)$$

式中：$\alpha_i \in \mathbb{R}^{1\times1\times H\times W}$为注意力分支学习的动态权重，表示第 i 尺度对融合结果的重要性，满足$\sum_{i=0}^{S}\alpha_i = 1$。为保证训练初期融合稳定性，所有动态权重$\alpha_i$通过对输入特征进行$1\times1$卷积生成，并经 Softmax 函数归一化处理，从而实现权重在空间维度上的自适应分配。具体地，将原始输入 X 经一组并行的1×1卷积核分别生成$S+1$个注意力图，表示不同尺度及残差路径的融合权重：

$$[\alpha_0,\alpha_1,\cdots,\alpha_s] = Softmax(\phi_a([F_0,F_1,F_2])) , \quad (4)$$

式中：$Conv_{1x1}^{(att)}$表示用于权重学习的注意力生成模块；Softmax 操作沿通道方向进行，确保各权重之和为 1。该设计既避免了显式初始化的敏感性问题，又使得网络在训练过程中能够自动调节多尺度融合的权重分布。

最终，将所有加权特征进行通道拼接并使用1×1卷积进行压缩，获得融合后的表达：

$$\hat{X} = Conv_{1\times1}([\alpha_0\odot F_0,\alpha_1\odot F_1,\alpha_2\odot F_2]) , \quad (5)$$

其中，$\odot$表示逐元素乘法。MCPA 模块分别嵌入 Cb、Cr 通道的特征路径中，使每一路径在保持原始结构信息的同时引入丰富的上下文辅助，如图 2 所示，MCPA 模块通过多尺度池化分支与通道注意力机制联合构建，实现对不同空间尺度下结构与色彩信息的自适应建模。实现对不同空间区域间依赖关系的自适应捕获。具体而言，输入特征经线性投影后，被划分为多个注意力头并行计算，从而在不同的表示子空间中协同聚焦于关键的结构与色彩信息。特别是在低照度场景下，该模块能够凭借其强大的长程上下文建模能力，有效缓解由于区域暗弱带来的结构模糊和色彩漂移现象，为后续亮度−色彩交叉融合提供更稳定、更具判别性的特征支撑。

2.4 亮度引导交叉融合模块

为了实现亮度与色彩特征之间的协同优化，本文提出了一种亮度引导的交叉融合模块 (LCA)。该模块整合了通道注意力机制 (SE) 与多头自注意力机制 (MSA)，以亮度特征为引导信息，有效调节色彩通道的响应模式，从而提升增强图像的结构保真度与色彩一致性。

设亮度分支输出特征为$F_Y \in \mathbb{R}^{C\times H\times W}$，色彩分支融合特征为$F_C \in \mathbb{R}^{C\times H\times W}$，我们首先将色彩特征$F_{Cb}$和$F_{Cr}$通过$1\times1$卷积进行通道压缩并融合：

$$F_{ref} = Conv_{1\times1}(Concat(F_{Cb},F_{Cr})) . \quad (6)$$

接着，将亮度特征F_Y引导映射为辅助调控信号：

$$G_Y = Conv_{1\times1}(F_Y) . \quad (7)$$

再通过通道注意力机制对融合特征进行调节，定义如下：

$$F_{mod} = SE(F_{ref} + \beta G_Y) , \quad (8)$$

式中：β是可学习的引导系数；$SE(\cdot)$表示通道注意力模块，用于加强关键通道的响应能力。

为进一步建模全局上下文依赖性，LCA 模块引入多头自注意力机制对调控特征F_{mod}进行空间维度建模。设输入特征$F \in \mathbb{R}^{H\times W\times C}$，通过三个无偏置线性变换获得查

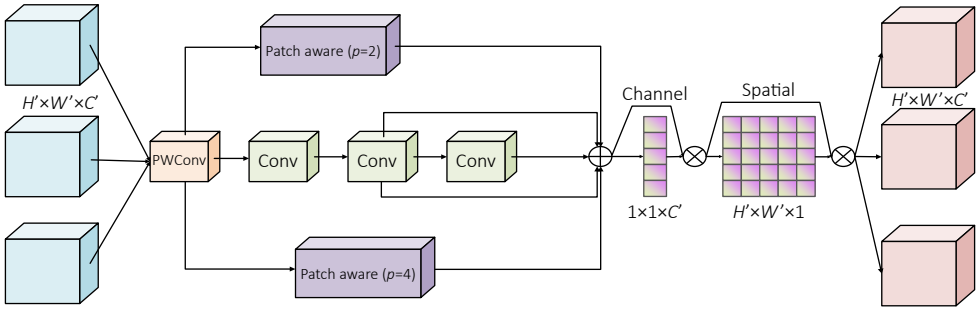


图 2 MCPA 模块

Fig. 2 Architecture of the MCPA module

询 ($\boldsymbol{Q}$)、键 ($\boldsymbol{K}$) 和值 ($\boldsymbol{V}$) 向量：

$$\boldsymbol{Q}=FW_Q,\ \boldsymbol{K}=FW_K,\ \boldsymbol{V}=FW_V,\quad \boldsymbol{Q},\boldsymbol{K},\boldsymbol{V}\in\mathbb{R}^{H\times W\times C}. \tag{9}$$

将$\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V}$拆分为 k 个注意力头，每个头的维度为 $d_k=\dfrac{C}{k}$，则第 i 个注意力头的计算为

$$Attention(\boldsymbol{Q}_i,\boldsymbol{K}_i,\boldsymbol{V}_i)=Softmax\left(\frac{\boldsymbol{Q}_i\boldsymbol{K}_i}{\sqrt{d_k}}\right)\boldsymbol{V}_i. \tag{10}$$

多个头的输出在通道维度拼接后，再通过一个线性层恢复至原始维度，得到输出特征$F_{\text{out}}\in\mathbb{R}^{H\times W\times C}$。

最终，将该全局融合特征与亮度特征再次融合用于图像重建，过程如下：

$$F_{\text{fuse}}=Conv_{3\times3}(Concat(F_{\text{out}},F_{\text{Y}})). \tag{11}$$

该融合机制不仅利用亮度信息引导色彩调整，还通过空间上下文建模弥合色彩与结构的差异，有效提升了增强图像在边缘、纹理与色调上的整体感知质量。

2.5 损失函数

为了提升低照度图像提高结果的综合质量，本文引入了多项复合损失函数，从结构、感知、色彩与分布等多个维度对重建图像进行联合约束。整体损失函数定义如下：

$$\begin{aligned}L_{\text{total}}=&\alpha_1L_{\text{SmoothL1}}+\alpha_2L_{\text{Perceptual}}+\alpha_3L_{\text{Histogram}}\\&+\alpha_4L_{\text{MS-SSIM}}+\alpha_5L_{\text{PSNR}}+\alpha_6L_{\text{Color}},\end{aligned} \tag{12}$$

式中：L_{SmoothL1}用于保持结构一致性，具有良好的鲁棒性；$L_{\text{Perceptual}}$基于 VGG 网络特征表示，衡量图像的语义感知误差；$L_{\text{Histogram}}$从像素分布角度约束整体亮度形态，提升图像的全局一致性；$L_{\text{MS-SSIM}}$评估图像在多尺度下的结构保留能力；L_{PSNR}反映数值重建精度；L_{Color}则保证图像的色彩均衡与自然性。

在实验中，本文采用的损失权重设置为$\alpha_1=1.00$, $\alpha_2=0.06$, $\alpha_3=0.05$, $\alpha_4=0.5$, $\alpha_5=0.0083$, $\alpha_6=0.25$。该加权策略在多个真实场景下表现出良好的增强效果，兼顾结构保真、感知质量与色彩一致性。综上所述，如算法 1 所示。

算法 1 多损失联合优化算法

输入：低光照图像 I，真实标注 G

1. 初始化：

- 网络参数$\theta\sim N(0,0.01)$

- 损失权重$\alpha=[1.0,0.06,0.05,0.5,0.0083,0.25]$

2. 迭代优化 (共 N 轮)：

步骤 1 前向传播：

$O=f_\theta(I)$

步骤 2 损失计算：

$$L_1=\frac{1}{n}\sum_{k=1}^{n}L_{\text{Smooth1}}(O_k,G_k),$$

$$L_2=\|\phi_{\text{VGG19}}(O)-\phi_{\text{VGG19}}(G)\|_2^2,$$

$$L_3=1-\frac{\sum\min(H_O,H_G)}{\sum H_G},$$

$$L_4=1-\prod_{s=1}^{5}[SSIM(O,G)_s]^{\beta},$$

$$L_5=10^{-PSNR(O,G)/10},$$

$$L_6=\Delta E_{00}(O,G).$$

步骤 3 加权融合：

$$L_{\text{total}}=\sum_{i=1}^{6}\alpha_iL_i$$

步骤 4 反向传播：

$\theta\leftarrow\theta-\eta\cdot Adam(\nabla_\theta L_{\text{total}})$

3. 输出最优模型参数θ^*及增强结果O^*

3 实验及结果

为系统地评估本文在低照度图像增强任务中的综合性能，实验部分选取了 LOLv1[3]、LOLv2[13]、LSRW[14] 三个成对监督数据集，以及 LIME[15]、NPE[16]、MEF[17] 三个非成对真实低照度数据集，全面覆盖不同类型与分布特征的图像场景。对比方法涵盖当前具有代表性的主流增强算法，包括基于 Retinex 理论的 KinD、Retinexformer 和 RetinexMamba[18]，轻量级自校准增强方法 SCI，多尺度特征学习模型 MIRNet[19]，生成对抗机制驱动的 EnlightenGAN，以及基于概率建模的扩散式增强方法 LightenDiffusion[20]。在评价标准方面，本文采用峰值信噪比[21](peak signal-to-noise ratio, PSNR)、结构相似性指数[22](structural similarity index, SSIM)、自然图像质量评估指标[23](natural image quality evaluator, NIQE) 以及盲参考图像空间质量评估指标[24](blind/referenceless image spatial quality evaluator, BRISQUE) 四项指标，从图像的亮度还原、结构保持、色彩自然性与无参考质量等角度进行多维度对比。同时，为综合考虑模型在实际部署中的计算开销，本文还统计了各方法的参数量 (params) 与浮点运算量 (floating point operations, FLOPs)。通过全面的定量与定性实验，验证本文在增强质量、模型效率与泛化能力等方面的优越性。为了全面评估图像增强算法的性能，本实验采用了 PSNR、SSIM、NIQE、

BRISQUE 四项评价指标。PSNR 用于衡量增强图像与原始图像之间的信噪比，PSNR 值越高，表示图像的质量越好，噪声越小。

3.1 实验环境和数据集

本文实验的硬件与软件环境配置如下：采用 12 核 vCPU 的 Intel (R) Xeon (R) Platinum 8352VCPU@2.10 GHz 作为处理器，配备一张 vGPU-32 GB (32 GB 显存) 的显卡用于模型训练。操作系统为 Ubuntu 20.04，深度学习框架使用 PyTorch 2.0.0，Python 版本为 3.8，CUDA 版本为 11.8。

在训练阶段，本文使用 Adam 优化器对网络进行端到端优化，其中超参数设置为 $\beta_1 = 0.9$，$\beta_2 = 0.999$，训练总轮数为 500 轮。初始学习率设为 2×10^{-4}，并采用余弦退火策略逐步衰减至 1×10^{-6}，以提升训练稳定性并避免陷入局部最优。损失函数为由结构、感知、分布与色彩组成的混合损失，其各项权重分别设为 $\alpha_1 = 1.00$，$\alpha_2 = 0.06$，$\alpha_3 = 0.05$，$\alpha_4 = 0.5$，$\alpha_5 = 0.0083$。在数据配置方面，使用 LOLv1 数据集中 485 对图像作为训练集，标准测试集中的 15 对图像用于测试，训练批量大小 (batch size) 设为 16。所有实验均在统一的随机初始化条件下进行，不使用任何预训练模型，亦未采用额外的数据增强策略，以确保实验结果的公平性、可重复性与模型泛化性能的真实反映。

3.2 配对数据集对比实验

为验证本文在结构建模与色彩一致性方面的效果，本文在有参考数据集 LOLv1 的 15 对测试图像、LOLv2-real 的 344 对测试图像、LSRW 数据集中的 200 对图像进行了定量分析和定性分析，全面评估其在亮度恢复、细节重建与色彩校正方面的性能。

1) 定量分析

为了全面评估低照度图像增强方法在实际应用中的表现，本文从色彩饱和度、细节纹理恢复、去噪能力等多个方面进行综合考量，并选取近年来具有代表性的先进算法作为对比对象；如表 1 所示，表中加粗的数值表示该指标下的最优结果。具体包括：基于 Retinex 理论的 KinD，代表传统分解增强方法；基于生成对抗网络的 EnlightenGAN，体现无监督增强策略；轻量化设计的 SCI，适用于资源受限场景；高性能深度网络 MIRNet，代表多阶段卷积增强模型；融合 Retinex 理论与高效注意力机制的 RetinexMamba；基于扩散生成模型的 LightenDiffusion；以及结合 Retinex 分解与 Transformer 架构的 Retinexformer。我们在公开的 LOLv1、LOLv2 和 LSRW 数据集上对各方法进行定量对比，涵盖 PSNR、SSIM、参数量 (params) 和计算量 (FLOPs) 等评价指标，以系统验证本文所提方法在不同光照条件和场景下的增强效果与模型效率。

从表 1 中可以看出，本文在三个数据集上的 PSNR 与 SSIM 指标均取得领先性能，分别在 LOLv1、LOLv2 与 LSRW 上达到了 25.66 dB/0.84、24.61 dB/0.85 与 20.26 dB/0.57，均优于当前最佳方法 Retinexformer 与 RetinexMamba。与此同时，本文在模型复杂度上也展现出显著优势，其参数量仅为 0.059 M，FLOPs 为 10.06 G，远低于如 MIRNet (785 G)、EnlightenGAN (61.01 G) 等方法，体现出极高的计算效率与轻量化部署潜力。该结果充分验证了本文在亮度 – 色彩解耦建模机制与多分支注意力设计下，不仅能有效提升增强质量，同时具备良好的实用性与适应性。

2) 定性分析

为进一步从主观视觉层面评估各方法在低照度图像增强任务中的实际表现，本文在 LOLv1 数据集中选取了多个具有代表性的测试样本进行可视化对比，图 3 展示了各方法的增强结果及其对应的亮度表面图，从结构

表 1 各方法在成对数据集上的性能与复杂度对比

Table 1 Comparison of performance and complexity for different methods on the paired dataset

Method	complexity		LOLv1		LOLv2		LSRW	
	Params/M	FLOPs/G	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
KinD	8.02	34.99	20.38	0.82	18.94	0.78	16.25	0.47
SCI	0.0013	0.03	14.78	0.52	15.56	0.58	14.95	0.41
MIRNet	31.76	785	24.14	0.84	21.93	0.80	15.53	0.43
EnlightenGAN	114.35	61.01	17.48	0.65	17.61	0.72	16.89	0.45
RetinexMamba	4.59	42.82	24.02	0.82	24.17	**0.88**	17.05	0.47
LightenDiffusion	27.83	211.5	20.29	0.80	22.17	0.87	17.01	0.37
Retinexformer	1.61	15.57	24.65	**0.84**	24.23	**0.88**	17.45	0.49
Ours	0.059	10.06	**25.66**	**0.84**	**24.61**	0.85	**20.26**	**0.57**

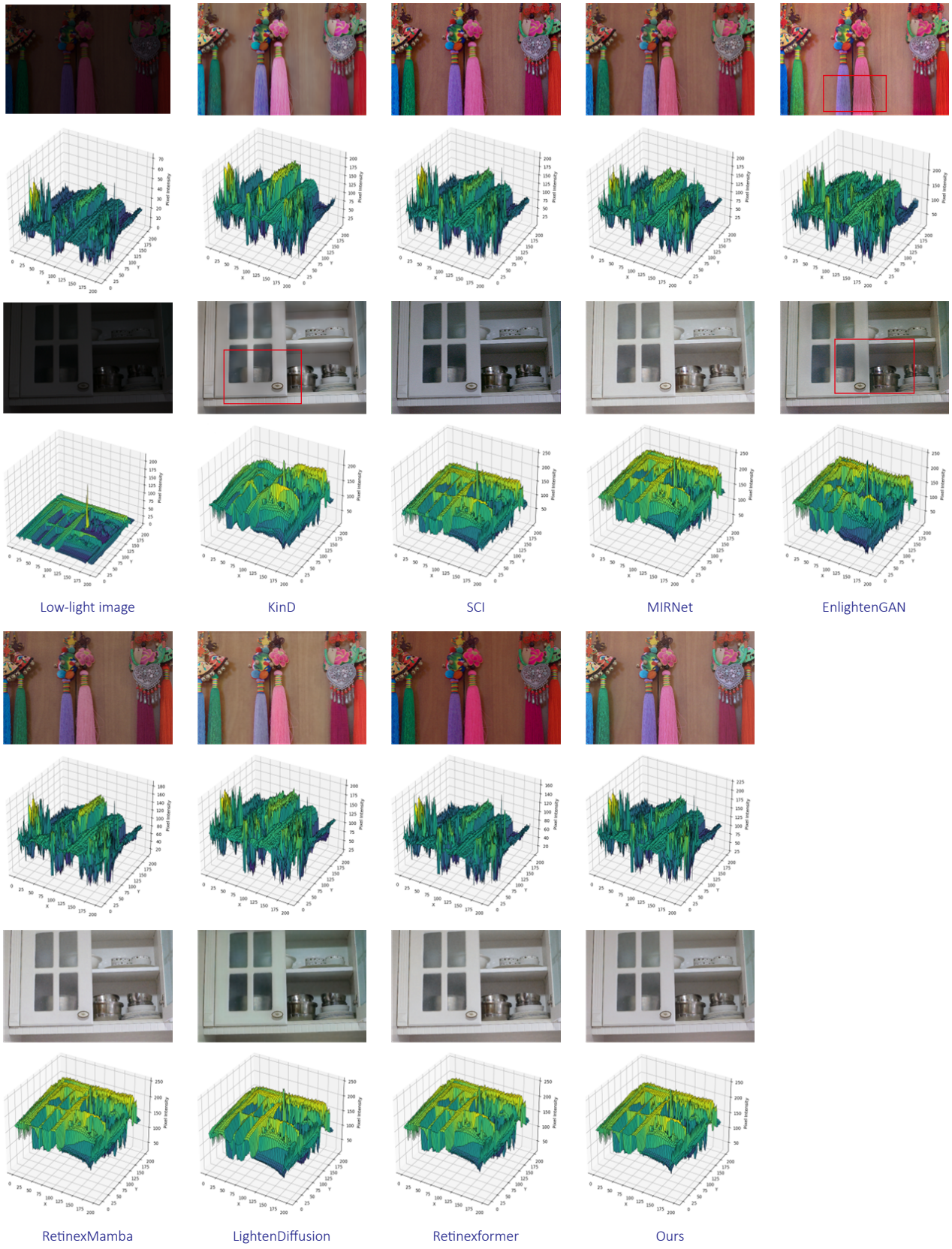


图 3 各方法在 LOLv1 测试图像上的增强效果及亮度表面图对比

Fig. 3 Comparison of enhancement results and brightness surface maps for different methods on the LOLv1 test images

恢复与亮度重建两个维度直观分析模型性能差异。

从图 3 中结果可以观察到，SCI 方法虽具备轻量化特性，但由于缺乏对亮度与结构信息的有效建模，在亮度提升过程中未能充分保留细节特征，导致图像出现色偏现象，边缘结构模糊不清。MIRNet 和 EnlightenGAN 在增强过程中借助多尺度感知与对抗学习分别实现亮度提升，但前者因其深层特征重建中缺乏显式结构约束，在暗部区域容易出现细节缺失，而后者则受限于 GAN 训练不稳定性，在不同区域表现出亮度不均匀，影响整体层次感。RetinexMamba 与 Retinexformer 两种方法依赖于 Retinex 理论进行亮度与反射分离，因而在色彩还原方面表现较为稳定，但在高光或阴影区域，由于对局部极端亮度变化的适应能力不足，仍可能出现伪影或纹理残缺问题。LightenDiffusion 借助扩散模型生成机制提升整体亮度分布，但缺乏细粒度调控能力，增强图像往往显得饱和度偏高，细节层次感减弱。

相较之下，本文在设计上通过亮度 – 色彩解耦与亮度引导的跨分支融合机制，有效实现结构信息与色彩特征的协同增强，既避免了过度增强引起的伪影，也提升了细节还原的完整性。其增强图像在亮部与暗部区域均表现出良好的层次过渡，边缘结构清晰，色彩自然协调；亮度表面图也呈现出更平滑且连续的变化趋势，进一步验证了所提方法在结构感知与光照建模方面的优势。

3.3 非配对数据集对比实验

为进一步验证本文在无监督场景下的泛化能力，本文在四个公开的非成对低照度图像数据集 LIME、NPE、MEF 和 DICM 上进行了定量分析与定性分析。这些数据集包含多样化的真实拍摄低照图像，覆盖不同光照条件与场景分布，具有较强的代表性。

1) 定量分析

针对非成对低照度图像增强场景，本文进一步在 LIME、NPE、MEF 与 DICM 四个公开数据集上开展了定量实验，评估本文在缺乏配对监督条件下的性能表现。我们采用无参考图像质量评价指标 NIQE 与 BRISQUE 对增强图像进行客观评估，数值越低表示图像质量越好。如表 2 所示，其中加粗表示各项指标中的最优结果。

从表中可以观察到，本文在大多数指标下均实现了最优或次优性能，尤其在 LIME 与 NPE 数据集中，NIQE 与 BRISQUE 分别为 3.356/19.38 与 3.268/20.93，显著优于现有方法。在 MEF 与 DICM 数据集上，尽管个别方法在部分指标上略有优势，本文方法依然保持在最低值附近，展示出极强的鲁棒性和稳定性。该结果进一步表明，本文不仅具备出色的低照图像增强能力，同时在无监督条件下依然能够生成高质量、自然真实的图像，验证了其良好的泛化性能。

2) 定性分析

为进一步评估各方法在非配对真实场景下的图像增强性能，本文在 LIME、NPE、MEF 与 DICM 四个典型数据集中选取代表性样本进行视觉效果对比，结果如图 4 所示。图中红色框选区域标出了增强结果中常见的伪影、颜色偏差或过曝光区域，便于直观比较不同方法的局部表现。

从图 4 可以看出，各方法在不同场景下的增强效果存在明显差异。第一行图像为复杂背景下的低照度场景，KinD 和 SCI 方法由于未对亮度与色彩进行有效解耦，

表 2 在非成对低照度图像增强数据集上的无参考指标对比

Table 2 Comparison of no-reference metrics for different methods on the unpaired low-light image enhancement dataset

Method	LIME		NPE		MEF		DICM	
	NIQE↓	BRISQUE↓	NIQE↓	BRISQUE↓	NIQE↓	BRISQUE↓	NIQE↓	BRISQUE↓
KinD	4.77	26.653	3.536	27.212	5.397	35.003	3.559	28.006
SCI	4.206	22.834	3.999	31.603	4.533	30.267	2.917	22.382
MIRNet	3.975	26.166	3.604	26.771	4.064	32.554	3.126	25.213
EnlightenGAN	3.657	21.910	4.018	33.599	3.783	27.598	2.917	22.382
RetinexMamba	3.982	26.553	3.394	23.180	3.764	27.557	**2.890**	**20.825**
LightenDiffusion	3.865	23.076	3.371	24.385	3.926	24.130	3.364	21.020
Retinexformer	4.072	26.223	3.477	24.647	3.789	**23.899**	2.975	22.465
Ours	**3.356**	**19.38**	**3.268**	**20.93**	**3.74**	25.78	2.936	22.90

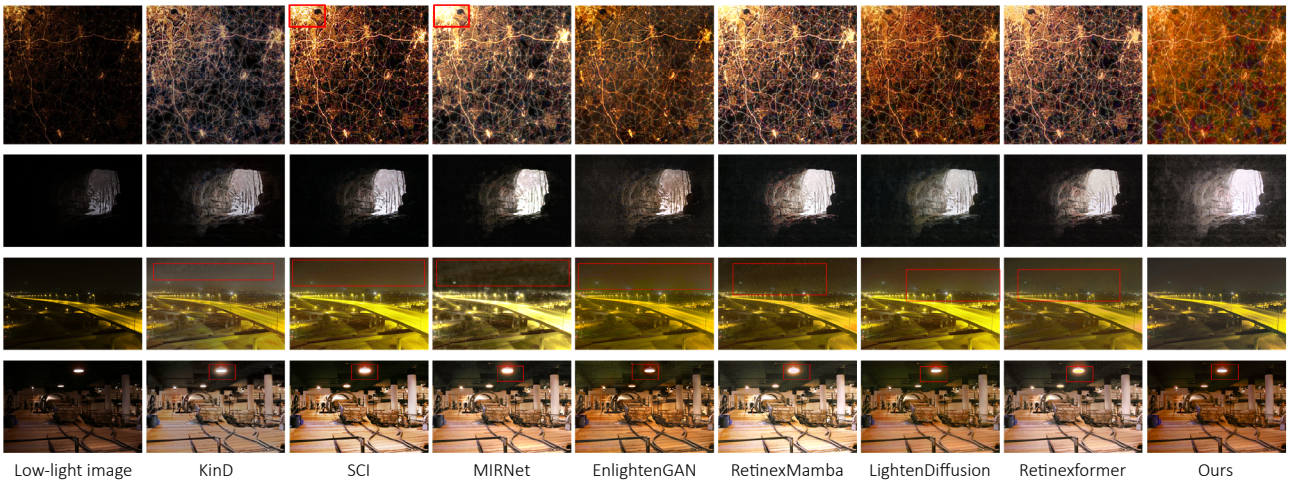


图 4 不同方法在非成对数据集上的视觉增强对比

Fig. 4 Visual comparison of enhancement results for different methods on the unpaired dataset

易造成结构特征与色彩信息相互干扰，导致增强后图像出现颜色漂移和局部区域过曝，整体观感不自然；Retinexformer 虽引入 Transformer 结构强化全局建模能力，但在高频细节保留方面表现不足，造成纹理模糊、颜色增强不协调的问题。EnlightenGAN 与 LightenDiffusion 方法由于对全局亮度分布建模能力较弱，增强后的图像整体偏暗，缺乏足够的对比度与结构感知，细节表现力有限。而本文得益于亮度 – 色彩解耦设计与结构 – 色彩协同增强机制，在提升亮度的同时，有效保留了背景纹理和边缘轮廓，使增强结果更为自然真实。第二行隧道场景光照条件复杂，RetinexMamba 与 SCI 增强后图像亮度分布不均，主要由于其通道注意力建模不足，难以捕捉局部照明变化，导致图像局部区域过暗或过亮；MIRNet 虽在亮度提升方面表现较好，但其深层特征重建结构难以对复杂边缘细节进行充分恢复，出现边界模糊现象。相比之下，本文在交叉融合模块的引导下，有效提升了通道间的光照响应能力，使得增强结果亮度过渡更平滑，隧道边界清晰，整体更贴近真实感知。第三行与第四行图像进一步验证了各方法在极端光照与色温场景中的适应能力。部分方法 (如 KinD、EnlightenGAN) 在高光区域处理时缺乏对结构约束的控制，易出现光晕扩散现象；而在偏色场景中，Retinexformer 等方法由于未充分建模色彩上下文关系，导致图像色温偏黄、色彩失真明显。而本文通过多通道并行注意力增强色彩建模能力，并联合亮度信息引导结构增强，从而在色彩饱和度与结构清晰度之间取得良好平衡，有效抑制了色偏、过曝等问题，增强效果更为稳定可靠。

综上所述，本文在不同复杂低照环境中均展现出良好的视觉一致性、细节保真与伪影抑制能力，进一步验证了其在无监督增强任务中的实用价值与鲁棒性。

3.4 消融实验

为全面评估本文中各核心模块的有效性，本文在 LOLv1 数据集上进行了逐步组件添加的消融实验，指标包括 PSNR、SSIM。消融结果如表 3 所示。

表 3 本文各模块设计的消融实验结果对比

Table 3 Ablation study results of different module designs in the proposed method

模型	模型组合	PSNR/dB	SSIM
模型A	轻量主干+MSA	23.89	0.783
模型B	A+ YCbCr	24.61	0.812
模型C	B + MCPA	25.04	0.831
模型D	C + LCA	25.48	0.837
本章算法	完整结构	**25.66**	**0.844**

可以看出，随着网络模块的逐层引入，增强质量持续提升，验证了本文亮度-色彩解耦建模、MCPA 以及 LCA 的有效性。模型 A 采用轻量化卷积主干，并保留单层多头自注意力（MSA）以维持基本的全局建模能力，该设置作为后续所有实验的基础参照，其 PSNR 和 SSIM 分别为 23.89 dB 与 0.783。

在此基础上，模型 B 引入 YCbCr 颜色空间解耦策略，分别建立亮度 – 色彩子网络进行协同建模，从而缓解了统一建模造成的特征干扰，PSNR 提升至 24.61 dB，

SSIM 提升至 0.812。随后，模型 C 添加了 MCPA，以提取多尺度上下文结构和颜色信息，有效增强模型对复杂纹理与颜色变化的鲁棒性，性能进一步提升。最后，模型 D 集成了 LCA，通过亮度特征引导色彩修复路径，实现了结构感知下的动态融合，有效抑制伪影与颜色偏差。完整模型区别在于添加了提出的损失函数，在结构效率与增强性能上取得双重优势，其 PSNR 与 SSIM 分别达到 25.66 dB 与 0.84，验证了本文所提解耦建模策略与模块设计的有效性。

4 低光环境下人脸检测性能

低光照增强的主要目标之一是为诸如面部识别、路径识别、智能驾驶等高级视觉任务奠定基础。为验证本文提出的方法在高级视觉任务中的有效性，我们采用了 DSFD[25] 方法与 DARK FACE[26] 进行实验。DARK FACE 数据集的特点在于其涵盖了多种不同场景与条件下的人脸图像，内容包括但不限于光照变化、姿态变化、表情变化、遮挡变化等。该数据集旨在提供具有挑战性的样本，以推动人脸检测算法在复杂场景中的研究与进展。本文的研究重点是针对光照变化的人脸图像。

尽管所有方法在一定程度上都能提高人脸识别的数量，但其识别准确率存在差异。如图 5 所示，原始图像中共有 13 张人脸，但在使用 EnlightenGAN 方法进行增强后，仅能识别出 6 张人脸，MIRNet 和 RetinexMamba 方法的识别结果为 8 张人脸。而本文算法与 SCI 的表现最佳，均能够识别出 12 张人脸，这表明了我们方法在夜间人像识别中的显著优势。接着，本文从 DARK FACE 数据集中随机选取了 50 张图像进行人脸识别，并采用人脸识别中常用的准确率 (ACC)、TAR 和 AP 等指标进行客观评估，实验结果如表 4 所示。

从表 4 可以更直观地看出，增强后的图像在性能上

图 5 对比增强方法在黑暗环境下人脸检测性能的表现

Fig. 5 Performance of different enhancement methods for face detection in dark environments

表 4 不同方法下 ACC、TAR 和 AP 指标的比较

Table 4 Comparison of ACC, TAR, and AP metrics across different methods

Method	ACC↑	TAR↑	AP↑
Input	0.1712	0.1721	0.1699
KinD	0.2468	0.2603	0.2473
SCI	0.3699	0.3767	0.34
MIRNet	0.3082	0.3082	0.3083
EnlightenGAN	0.4315	0.4452	0.4319
Retinexmamba	0.3151	0.3219	0.3152
LightenDiffusion	0.1844	0.1843	0.1843
Retinexformer	0.4109	0.4041	0.4041
Ours	**0.5068**	**0.5137**	**0.5069**

优于原始图像的直接人脸检测。其中，本文方法在所有方法中表现最佳，ACC、TAR 和 AP 均超过了 0.5，这表明其在夜间人脸识别方面具有显著优势。

5 结　论

低照度图像增强任务在保持结构细节与色彩一致性方面仍具有较大挑战。本文围绕轻量化与亮度 – 色彩解耦两个核心问题，提出了一种结合多通道并行注意力与交叉融合的轻量化低照度图像增强网络，实现了在资源受限条件下图像的高质量增强。实验结果验证了其在多个公开数据集上的优越性能，特别是在结构保真、色彩自然性和计算效率之间取得了良好平衡。本研究表明，亮度与色彩的分离建模及亮度引导的交叉融合策略能够有效缓解传统方法中亮度与色彩耦合带来的伪影问题，为轻量化图像增强网络设计提供了新的思路。

未来工作将聚焦于两个方向：一是进一步提升模型的泛化性与实时性能，探索其在视频增强及多模态夜视任务中的应用潜力；二是结合生成式学习和视觉先验信息，构建更高效、更具适应性的低照度视觉增强框架。

作者贡献

程浩：改进方法的提出，论文构思和撰写，实验数据整理与分析；林坚普：改进方法的研讨，论文审核与编辑写作；孙磊：论文审核与编辑写作；张永爱：调查研究；林珊玲：论文思路方案和实验执行进度总体指导，论文审核；郭太良：提供课题实验条件与资源。林志贤：论文审核与编辑写作

利益冲突

所有作者声明无利益冲突

参考文献

1. Sun F Y, Lyu Z, Lyu Z W. Review of low light image enhancement based on deep learning[J]. *Appl Res Comput*, 2025, **42**(1): 19−27.
孙福艳, 吕准, 吕宗旺. 基于深度学习的低光照图像增强研究综述[J]. 计算机应用研究, 2025, **42**(1): 19−27.
2. Land E H, McCann J J. Lightness and retinex theory[J]. *J Opt Soc Am*, 1971, **61**(1): 1−11.
3. Wei C, Wang W J, Yang W H, et al. Deep retinex decomposition for low-light enhancement[Z]. arXiv: 1808.04560, 2018. https://doi.org/10.48550/arXiv.1808.04560.
4. Zhang Y H, Zhang J W, Guo X J. Kindling the darkness: a practical low-light image enhancer[C]//*Proceedings of the 27th ACM International Conference on Multimedia*, 2019: 1632–1640. https://doi.org/10.1145/3343031.3350926.
5. Cai Y H, Bian H, Lin J, et al. Retinexformer: one-stage retinex-based transformer for low-light image enhancement[C]//*Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2023: 12470–12479. https://doi.org/10.1109/ICCV51070.2023.01149.
6. Yi X P, Xu H, Zhang H, et al. Diff-Retinex++: retinex-driven reinforced diffusion model for low-light image enhancement[J]. *IEEE Trans Pattern Anal Mach Intell*, 2025, **47**(8): 6823−6841.
7. Jiang Y F, Gong X Y, Liu D, et al. EnlightenGAN: deep light enhancement without paired supervision[J]. *IEEE Trans Image Process*, 2021, **30**: 2340−2349.
8. Ma L, Ma T Y, Liu R S, et al. Toward fast, flexible, and robust low-light image enhancement[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022: 5627–5636. https://doi.org/10.1109/CVPR52688.2022.00555.
9. Chen J X, Li Y Z, Li W. Low-light image enhancement via fusion of HVI color space and mamba[J]. *Comput Eng Appl*, 2026, **62**(04): 250−260.
陈景霞, 李燚卓, 李婉. 融合 HVI 颜色空间和 Mamba 的低照度图像增强[J]. 计算机工程与应用, 2026, **62**(04): 250−260.
10. Brateanu A, Balmez R, Avram A, et al. LYT-Net: Lightweight yuv transformer-based network for low-light image enhancement[EB/OL]. (2025-09-10)[2026-04-21]. https://doi.org/10.48550/arXiv.2401.15204.
11. Yang Y, Peng Y H, Liu Z G. A fast algorithm for YCbCr to RGB conversion[J]. *IEEE Trans Consum Electron*, 2007, **53**(4): 1490−1493.
12. Xu S, Zheng S, Xu W, et al. Hcf-net: Hierarchical context fusion network for infrared small object detection[C]//2024 IEEE International Conference on Multimedia and Expo (ICME). IEEE, 2024: 1-6.
13. Yang W H, Wang W J, Huang H F, et al. Sparse gradient regularized deep retinex network for robust low-light image enhancement[J]. *IEEE Trans Image Process*, 2021, **30**: 2072−2086.
14. Hai J, Xuan Z, Yang R, et al. R2RNet: low-light image enhancement via real-low to real-normal network[J]. *J Visual Commun Image Represent*, 2023, **90**: 103712.
15. Guo X J, Li Y, Ling H B. LIME: low-light image enhancement via illumination map estimation[J]. *IEEE Trans Image Process*, 2017, **26**(2): 982−993.
16. Wang S H, Zheng J, Hu H M, et al. Naturalness preserved enhancement algorithm for non-uniform illumination images[J]. *IEEE Trans Image Process*, 2013, **22**(9): 3538−3548.
17. Ma K D, Zeng K, Wang Z. Perceptual quality assessment for multi-exposure image fusion[J]. *IEEE Trans Image Process*, 2015, **24**(11): 3345−3356.
18. Bai J S, Yin Y H, He Q Y, et al. Retinexmamba: retinex-based mamba for low-light image enhancement[C]//*Proceedings of the 31st International Conference on Neural Information Processing*, 2024: 427–442. https://doi.org/10.1007/978-981-96-6596-9_30.
19. Zamir S W, Arora A, Khan S, et al. Learning enriched features for real image restoration and enhancement[C]//*Proceedings of the 16th European Conference on Computer Vision*, 2020: 492–511. https://doi.org/10.1007/978-3-030-58595-2_30.
20. Jiang H, Luo A, Liu X H, et al. LightenDiffusion: unsupervised low-light image enhancement with latent-retinex diffusion models[C]//*Proceedings of the 18th European Conference on Computer Vision*, 2024: 161–179. https://doi.org/10.1007/978-3-031-73195-2_10.
21. Huynh-Thu Q, Ghanbari M. Scope of validity of PSNR in image/video quality assessment[J]. *Electron Lett*, 2008, **44**(13): 800−801

22. Wang Z, Bovik A C, Sheikh H R, et al. Image quality assessment: from error visibility to structural similarity[J]. *IEEE Trans Image Process*, 2004, **13**(4): 600−612.
23. Mittal A, Soundararajan R, Bovik A C. Making a "completely blind" image quality analyzer[J]. *IEEE Signal Process Lett*, 2013, **20**(3): 209−212.
24. Venkatanath N, Praneeth D, Bh M C, et al. Blind image quality evaluation using perception based features[C]//*Proceedings of 2015 Twenty First National Conference on Communications (NCC)*, 2015: 1–6. https://doi.org/10.1109/NCC.2015.7084843.
25. Li J, Wang Y B, Wang C G, et al. DSFD: dual shot face detector[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019: 5055–5064. https://doi.org/10.1109/CVPR.2019.00520.
26. Yang W, Yuan Y, Ren W, et al. Advancing image understanding in poor visibility environments: A collective benchmark study[J]. *IEEE Transactions on Image Processing*, 2020, **29**: 5737−5752

作者简介

程浩 (1998-)，男，江西南昌人，硕士研究生，2023年于江西农业大学获得学士学位，主要从事图像处理方面的研究。

E- mail：1035757231@qq.com

【通信作者】林坚普 (1989-)，男，福建泉州人，博士，讲师，硕士研究生导师，福州大学先进制造学院电子信息系教师，主要从事新型显示技术、图像处理技术、电子纸驱动与集成等方面的研究。

E-mail：ljp@fzu.edu.cn

Low-light image enhancement combined with multi-channel parallel attention and cross-fusion

Cheng Hao[1,2], Lin Jianpu[1,2*], Sun Lei[2,3], Zhang Yongai[1,2], Lin Shanling[1,2], Guo Tailiang[2], Lin Zhixian[1,2]

[1] School of Advanced Manufacturing, Fuzhou University, Quanzhou, Fujian 362200, China;
[2] Fujian Science & Technology Innovation Laboratory for Optoelectronic Information of China, Fuzhou, Fujian 350116, China;
[3] Fuzhou University Zhicheng College, Fuzhou, Fujian 350002, China

Abstract

Objective Existing low-light image enhancement methods generally suffer from high model complexity, considerable deployment cost, and semantic coupling between luminance and chrominance information, which often leads to fusion artifacts and restricts their application on resource-constrained devices. To address these issues, this paper proposes a lightweight low-light image enhancement network that combines multi-channel parallel attention and cross-fusion mechanisms. In the YCbCr color space, luminance and chrominance information are modeled separately. A lightweight denoising module and a multi-head self-attention mechanism are introduced to extract illumination structure features, while a multi-channel parallel attention module is introduced to enhance color-context modeling. In addition, a luminance-guided cross-fusion module is introduced to achieve collaborative enhancement of structure and color under joint channel-spatial attention optimization.

Methods The input image is first converted into the YCbCr color space and decomposed into the Y, Cb, and Cr channels. The Y channel is processed by the DenoiseY module to suppress noise, and then passed through a multi-head self-attention module (MSA) to extract structural features. To improve contextual modeling efficiency, a pooling operation is inserted before MSA to reduce the number of tokens. Meanwhile, the Cb and Cr channels are enhanced by the MCPA module, which performs multi-scale attention modeling to extract richer color-context information.

Subsequently, in the LCA module, the denoised Y channel is used as a guidance feature to fuse with the Cb and Cr channels. This module incorporates channel attention based on the squeeze-and-excitation (SE) mechanism together with element-wise interaction,

Foundation item: Project supported by the National Key R&D Program of China Fund (2023YFB3609400), and the National Natural Science Foundation of China (62101132)

*Correspondence: Lin J P, E-mail: ljp@fzu.edu.cn

enabling luminance-guided structure-color fusion.

Finally, the fused features are restored to the RGB space through convolution layers to generate the enhanced image. By separately modeling luminance and chrominance information and introducing cross-guided fusion between the two branches, the proposed method effectively improves the overall quality of low-light images in terms of detail preservation, illumination enhancement, and color restoration.

Results and Discussions The proposed network achieves 25.66 dB PSNR / 0.84 SSIM on LOLv1, 24.61 dB PSNR / 0.85 SSIM on LOLv2, and 20.26 dB PSNR / 0.57 SSIM on LSRW, outperforming recent methods such as Retinexformer. Meanwhile, the model contains only 0.059 million parameters and requires 10.06 GFLOPs, demonstrating its lightweight nature and computational efficiency. In addition, low-light face detection experiments conducted on the DARK FACE dataset show that all three detection metrics exceed 50%, indicating good generalization capability under low-light conditions.

Experimental results further show that MCPA is embedded into both the luminance and chrominance branches to extract contextual information at different scales and perform feature reweighting, thereby improving the network's adaptive modeling ability for multi-scale structural patterns and color variations. The LCA module enables joint modeling of structure and color, where structural information provided by the Y branch dynamically guides the adjustment of chrominance features, effectively improving detail fidelity and color consistency in the enhanced images.

Conclusions The proposed method achieves high-quality low-light image enhancement under resource-constrained conditions. Experimental results on multiple public datasets verify its superior performance, especially in achieving a favorable balance among structural fidelity, color naturalness, and computational efficiency. This study demonstrates that separate modeling of luminance and chrominance information, together with a luminance-guided cross-fusion strategy, can effectively alleviate the fusion artifacts caused by luminance-color coupling in conventional methods.

Keywords: low-light image enhancement; luminance-color decoupling; attention mechanism; multi-head self-attention; image perceptual quality; lightweight network

Cheng H, Lin J P, Sun L, et al. Low-light image enhancement combined with multi-channel parallel attention and cross-fusion[J]. *Opto-Electron Eng*, 2026, **53**(4): 250244; DOI: 10.12086/oee.2026.250244