基于多注意力特征融合的轻量级低照度图像增强方法

刘 艺[1]，朱佳慧[1]，郑涤尘[2]，张登银[1]

(1. 南京邮电大学物联网学院，南京 210003；2. 南京邮电大学通信与信息工程学院，南京 210003)

摘 要：低照度图像增强是将光照不足条件下获取的图像恢复成正常曝光的图像，现有低照度图像增强算法大多通过设计复杂的网络结构获取良好的增强效果，计算效率较低；增强后的图像仍会存在噪声增多、色彩失真、细节丢失等问题，影响视觉感知和后续高级视觉任务。为此，本文提出一种基于多注意力特征融合的轻量级低照度图像增强方法。采用简单门控注意力模块对低照度图像全局特征进行高效提取，通过简化通道注意力及门控单元减少计算开销并保留图像细节信息；采用多注意力融合模块对全局特征及局部接收场提取的局部特征进行信息整合，通过像素注意力强化通道注意力与空间注意力对于全局及局部特征的表征，更好地恢复图像色彩、抑制噪声。此外，使用联合损失函数对增强任务进行约束，在真实数据集上的广泛实验表明，本文方法的性能超过了当前先进的低照度图像增强方法，并具有良好的计算效率和泛化能力。

关键词：图像增强；低照度图像；注意力机制；轻量级；特征融合

中图分类号：TP391. 41　　**文献标志码：**A

Lightweight Low-light Image Enhancement Method Based on Multi-attention Feature Fusion

LIU Yi[1]，ZHU Jia-hui[1]，ZHENG Di-chen[2]，ZHANG Deng-yin[1]

(1. School of Internet of Things, Nanjing University of Posts and Telecommunications, Nanjing 210003, China；2. School of Communications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China)

Abstract：Low Light Image Enhancement, which is to restore the image acquired under the condition of insufficient light to the normal exposure image. Most of the existing low-light image enhancement algorithms obtain good enhancement effect by designing complex network structure, and the computational efficiency is low. The enhanced image will still have problems such as increased noise, color distortion and detail loss, which will affect visual perception and subsequent advanced visual tasks. Therefore, a lightweight low-light image enhancement method based on multi-attention feature fusion is proposed in this paper. Simple gate attention module is used to extract the global features of low-light images effectively, and the computational overhead is reduced and image details are preserved by simplifying the channel attention and gating unit. The multi-attention fusion module is used to integrate the information of global features and local features extracted from local receiving fields, and enhance the representation of channel attention and spatial attention to global and local features through pixel attention, so as to better restore image color and suppress noise. In addition, the joint loss function is used to constrain the enhancement task, and extensive experiments on real data sets show that the performance of the proposed method exceeds the current advanced low-light image enhancement methods, and has good computational efficiency and generalization ability.

Key words：image enhancement；low-light image；attention；lightweight；feature fusion

引　言

由于环境光的缺乏，在暗光或曝光不足环境下得到的图像通常具有对比度低、亮度低、噪声高、细节模糊等缺点。这些缺陷不仅显著降低了人类的视觉感知质量，还阻碍了高级计算机视觉任务（如目标检测[1]、图像修复[2]等）的应用。因此，有必要对低照度图像进行增强，以提升低光条件下高层视觉任务的精度。

基金项目：国家自然科学基金(62471241)。

收稿日期：20**-**-**；**修订日期**：20**-**-**

传统低照度增强方法多依赖于直方图均衡化和 Retinex 理论。直方图均衡化（Histogram Equalization，HE）[3]通过调节像素值来改变图像的对比度和亮度分布，但全局操作往往会引起图像细节的过度增强或压缩，导致图像失真扭曲，并一定程度上放大噪声。尽管改进方法[4]-[6]尝试通过附加先验约束抑制噪声，仍难以平衡细节保留与噪声控制。Retinex 理论模拟人类对光的视觉感知将图像分解为光照分量（随环境变化）和反射分量（固有特征）。基于 Retinex 理论的方法[7][9]通过增强反射分量提升视觉效果，但依赖于手工设计先验和正则化项，在复杂场景下易产生伪影或颜色偏差，且普遍忽视噪声抑制问题。

近年来，深度学习在低照度图像增强中的应用研究取得了长足的发展，相比于传统算法，在精度、鲁棒性和速度上展现出更好的性能。LLNet[10]是最早将深度神经网络应用于低照度图像增强的方法，采用了一种堆叠式稀疏去噪自编码器的变体来提升图像亮度，有效促进了端到端网络在该领域的发展。MBLLEN[11]通过构建多分支增强网络，结合特征提取、增强和融合三个模块，实现了有效的特征表达。MIRNet [错误!未找到引用源.]则利用递归残差来提取深层特征，能够有效恢复图像的细节信息，但其高计算复杂度也带来了较大的计算开销。考虑到 Retinex 理论和深度学习各自的优势，研究者将两者有效结合[错误!未找到引用源。-错误!未找到引用源。]。**RetinexNet 网络[17]通过分别调整图像分解的光照和反射分量实现了图像质量提升，但其浅层的上/下采样设计使得分解结果模糊并引入噪声。**URetinex-Net[18]通过自适应拟合数据驱动的隐式先验，巧妙地将优化问题转化为可学习网络，在最终分解结果中实现了噪声抑制和细节保留，也实现了更准确、更高效的分解。基于 Retinex 的方法虽然具备一定的理论基础，但其多个训练阶段及模型复杂度对计算成本提出了很高的要求。Retinexformer[19]中提出了一种新的单阶段 Retinex 框架，利用 Transformer 模型捕获远程依赖关系，通过向反射和照度分量中引入扰动项以模拟图像损坏情况，进而改进多阶段 Retinex 模型。然而，Transfomer 的计算成本随输入空间大小呈二次方增长，仍然面临计算复杂度过高的问题，难以在低照度图像中高效应用。

基于无监督学习的方法能够有效降低计算成本，并且不依赖于成对数据集训练，对各种环境更具有鲁棒性。SCI[20]提出了一种创新的权重共享照明学习框架，通过引入自校准模块，消除了设计网络结构的繁琐过程，只需要简单的操作即可达到增强的目的，但其结果往往出现亮度不足的情况。Guo 等[21]提出了一种深度曲线估算网络 Zero-DCE，该算法利用高阶曲线来调节图像的动态范围，在像素水平上对图像进行优化，以获得增强后图像。采用轻量级设计将图像映射成曲线，从而减少了网络参数，提高了推理速度。其加速版本 Zero-DCE++[22]进一步提升了运算效率，但在极端低光条件或复杂噪声下效果仍然有限。**此类方法因缺乏噪声与颜色约束，难以支撑后续高级视觉任务。**

生成式人工智能的快速发展也为低照度图像增强提供了新的范式。生成对抗网络（Generative Adversarial Network，GAN）通过对抗训练学习数据分布，能够生成逼真的增强结果。EnlightenGAN[23]从全局和局部两个方面协同增强低光图像，利用低光输入的光照信息生成自正则化注意图嵌入深度特征的每一层中来调节模型学习。然而，GAN 存在模式崩溃和训练不稳定的问题，且低光输入可见性差时，其约束效果有限。扩散模型通过前向扩散添加噪声、反向扩散去噪的马尔可夫链，从随机高斯噪声中重建目标图像，实现高质量图像生成。其生成能力优于 GAN，但需要大量的计算资源和推理成本，对此，研究者们将图像处理迁移至低维潜在空间以减少计算复杂度。DiffLL[24]通过引入小波变换将扩散过程迁移到低频子带，结合高频恢复模块（HFRM）在保持感知保真度的同时加速推理。LightenDiffusion[25]将 Retinex 理论与扩散模型结合，提出潜在空间内容迁移分解网络（CTDN），通过跨模态引导扩散过程实现无监督低光增强，在提升视觉质量的同时增强了模型泛化能力，但对于计算效率的提升仍然有限。

尽管现有深度学习低照度图像增强算法能够一定程度上改善低光图像的光照问题，但仍存在以下几方面不足：1）由于图像的对比度调整要同时考虑局部和全局信息，现有方法往往通过堆叠卷积层或连续下采样操作获取足够大的感受野来捕捉全局对比，这类方法很容易造成局部图像信息的永久性损失，同时增加了网络复杂度；2）不少算法所使用的损失函数未能对网络训练进行全面约束，造成了处理后图像存在明显的细节损失和颜色偏差；3）很多方法采用合成数据集进行训练及验证，然而实际环境中光照条件和噪声类型多样，这使得这些方法在应对真实低照度场景时表现不佳，泛化能力有限。

为解决上述问题，本文研究了一种新颖的轻量级低照度图像增强方法，能够更高效地利用输入图像的信息，避免增强过程中丢失细节信息，并在此基础上降低了参数量与计算量，加速了模型收敛，获得了高

质量的增强结果。使用 LOL 数据集评估本文方法与最先进方法在复杂度上的性能对比分析结果如图 1 所示，可以看出，本文所提方法综合性能达到最优，仅采用 14.9K 计算量模型达到了增强去噪的目的。

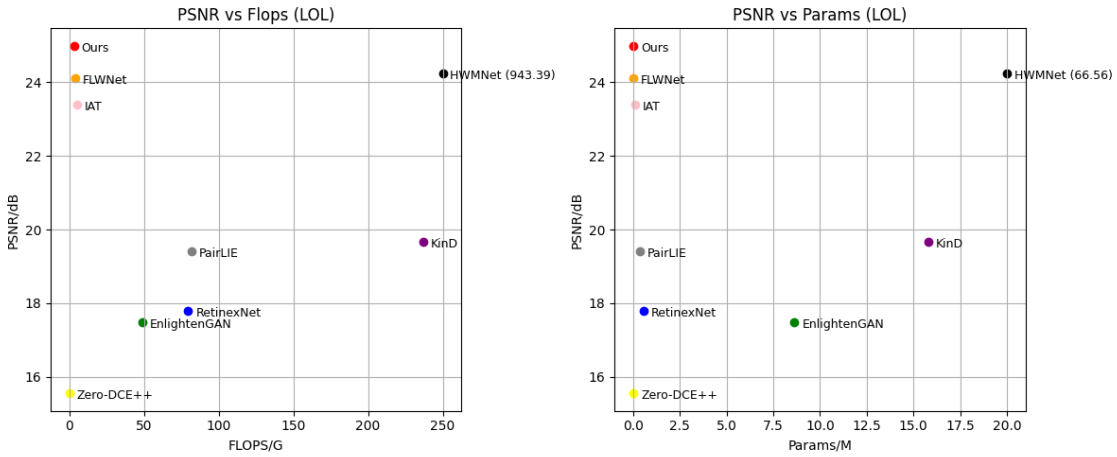


图 1　计算复杂度与模型性能对比

Fig. 1　Comparison of computational complexity and model performance

本文的主要贡献可以概括如下：

1）提出了一个轻量级的模型，设计了一个简单门控注意力模块用于提取全局细节特征，通过简化卷积网络模块，以一个小的接收场高效捕捉图像全局信息；引入 V 通道直方图学习全局亮度特征，共同构成全局特征提取模块，增强图像亮度的同时保留图像纹理细节，有效抑制图像噪声。

2）提出了一个多注意力特征融合模块，结合通道注意力、空间注意力、像素注意力整合图像局部信息与全局信息，避免在融合过程中丢失关键细节，确保特征融合的准确性与完整性。

3）设计了一个联合损失函数，引入 6 项损失函数对模型训练进行全面约束，提高了模型的增强能力；通过在多个真实数据集上进行训练及定量和定性分析，验证了本文方法在现实场景下相较于最先进方法更强的综合性能。

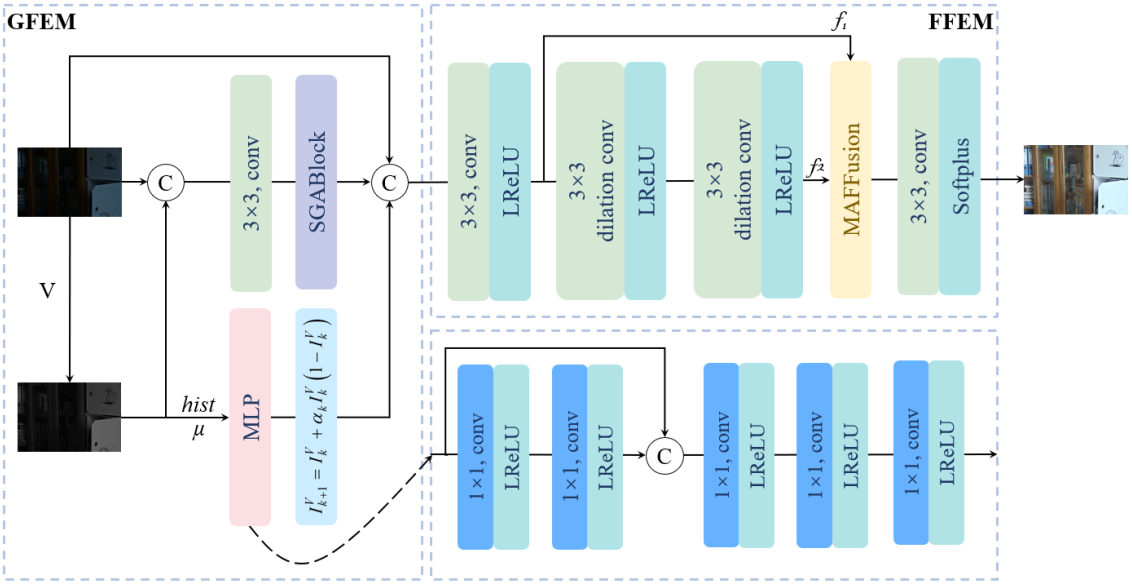


图 2　整体网络结构图

Fig. 2　Overall framework of the proposed method

1　网络结构

在图 2 中，展示了本文所提方法的整体架构，该模型包括两个主要模块：全局特征提取模块（Global Feature Extraction Module，GFEM）和特征融合增强模块（Feature Fusion Enhancement Module，FFEM）。GFEM 以低照度图像为输入，分别提取全局细节特征和全局亮度特征；将低照度图像及其 V 通道连接以提高细节表征，而后输入简单门控注意力模块（Simple Gate Attention Block，SGABlock）生成全局细节特征图；从 V 通道图像直方图中提取亮度信息，通过高阶曲线调整方法生成全局亮度特征图。FFEM 将低光输入、全局细节特征图以及全局亮度特征图作为输入，通过一个局部接收场提取局部特征信息，而后将全局和局部特征输入多注意力特征融合模块（Multi-attention Feature Fusion，MAFFusion）以融合图像信息并生成高频细节。

1.1　全局特征提取模块

在低照度图像增强中，全局信息提取至关重要，然而现有方法大多数使用大接收场或 Transformer 结构提取全局信息，这不可避免地增加了网络的复杂性。因此，本文通过简化门控单元和通道注意力机制提取全局细节特征，利用 V 通道图像直方图中亮度信息表征全局亮度特征，构建了一个全局特征提取的简易网络。如图 2 所示，该模块首先将低照度输入图像与其 V 分量连接，通过一个 3×3 的卷积将 4 通道调整为 16 通道，而后通过简单门控注意力模块对图像全局细节特征进行提取。同时，从低照度图像 V 通道的直方图中通过多层感知提取全局亮度信息，并通过高阶曲线调整方法生成全局亮度特征图。最后，将低光输入及其全局细节特征以及全局亮度特征图进行连接，得到全局特征，输入特征融合增强模块进一步处理。

1.1.1　全局细节特征提取

在计算机视觉领域，注意力机制和归一化策略是两种常用的技术。注意力机制通过提高对有用信息的注意水平，生成更具区别性的特征表示；归一化方法能够有效缓解内部协变量偏移，提升学习率并加速模型收敛。其中，通道注意力（Channel Attention，CA）可以提取图像特征中各通道间的相互关系，从而帮助网络更好地聚焦于重要的特征信息。CA 通过挤压（Squeeze）和激励（Excitation）操作[25]学习每个通道的权重，并利用卷积特征映射通道间的关系。具体如图 3（a）所示，挤压操作通过全局平均池化对全局上下文进行编码，激励操作通过 2 个卷积层实现，这 2 个卷积层分别采用 ReLU 与 Sigmoid 进行激活，从而生成通道注意力权重。随后，将生成的权重与原始特征图相乘，得到通道注意力图。为了降低计算量，Chen 等[27]将通道注意力进行简化，提出了一种简单通道注意力（Simple Channel Attention，SCA），具体如图 3（b）所示。SCA 仅使用 1 个卷积层进行激励操作，相比于 CA 更加简单，同时没有性能损失。

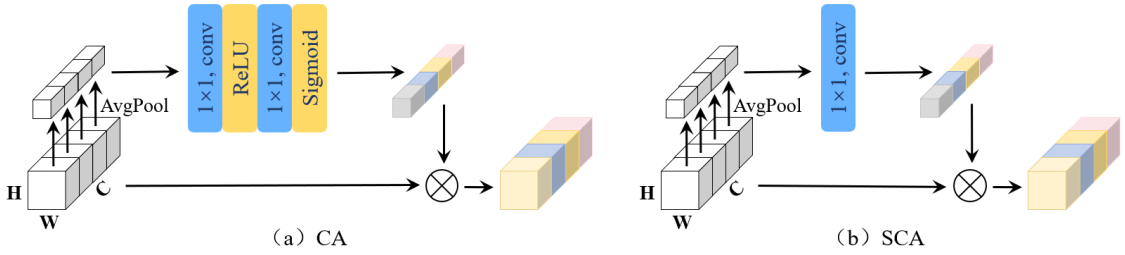


图 3　通道注意力简化前后对比

Fig. 3 Comparison before and after simplification of channel attention

鉴于门控单元在图像恢复任务中的性能增益作用[27]，本文引入了一种简单门控以降低网络复杂性，直接将特征映射在通道维度上分为两部分而后相乘，具体表示为：

$$SimpleGate(X,Y)=X\cdot Y \tag{1}$$

式中，X 和 Y 分别表示在通道维度上划分的两个相同大小的特征图。该门控单元通过逐元素相乘使得模型能够选择性地增强或抑制输入特征，在处理复杂数据时能够动态调整特征的重要性，提高特征表征。同时，通过将输入特征分为两部分而后相乘，实现了特征的非线性组合，能够捕获输入特征之间的交互关系，并

且代替非线性激活函数，进一步简化网络。基础的张量计算也使得该门控单元能够以较低的计算成本实现网络性能的提升。

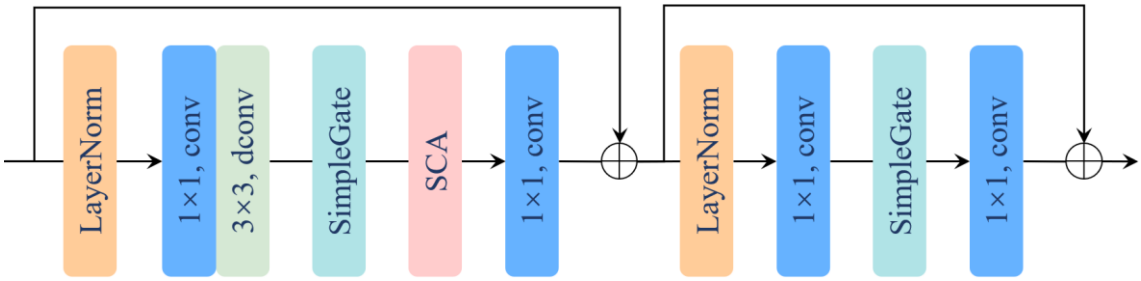


图 4　SGABlock

Fig. 4　Framework of the SGABlock

通过将归一化、简单门控单元（SimpleGate）、简单通道注意力（SCA）以及残差连接结合，构建了简单门控注意力模块（SGABlock），具体如图 4 所示。该模块采用层归一化方法，使得特征间的相关性得以保留；通过简单门控引入非线性机制，提升特征表达；SCA 的引入使得网络能够更好地考虑到图像中的黑暗区域，有效降低了噪声与伪影的影响；残差连接确保了深层特征提取的同时保持浅层特征的完整性，提高了训练过程的稳定性。

1.1.2　全局亮度特征提取

全局亮度特征对于低照度图像增强至关重要，通过分析全局亮度特征，可以有效区分图像中的噪声和实际信息，从而降低噪声对增强过程的干扰，提升最终图像整体质量。Zhang 等[28]采用仅 1.4K 参数的五层多层感知从图像的直方图中提取全局亮度信息，实验证明了直方图中亮度信息对于图像增强的有效性，并且图像尺寸的增大对计算成本的影响保持在较低水平，不会显著提高。同时，Di 等[29]证明了 HSV 色彩空间中的 V 通道足以表示输入图像的亮度，因此，本文从低照度图像 V 通道的直方图中提取全局信息，并通过高阶曲线调整方法生成全局亮度特征图[21]。具体可以表示为：

$$\{\alpha_{0,1,\ldots,t}\} = MLP(hist(I^V), \mu) \tag{2}$$

式中，I^V 表示低光图像 I 在 HSV 色彩空间中的 V 通道，$hist(I^V)$ 表示获取图像 I^V 直方图的操作，μ 表示期望的平均亮度值，$MLP(\cdot)$ 表示采用五层多层感知获取直方图中全局信息的操作，将全局信息表示为高阶曲线系数，记为 $\alpha_{0,1,\ldots,t}$。其中，μ 包含在直方图中，用于引导模型调整图像亮度，在训练过程中，将 μ 设置为参考图像 V 通道的平均值；在测试过程中，将 μ 固定为恒定值 0.55。

然后，利用高阶曲线系数迭代调整低光图像的 V 通道，从而得到全局亮度特征图。具体表示为：

$$I^V_{k+1} = I^V_k + \alpha_k I^V_k \left(1 - I^V_k\right) \tag{3}$$

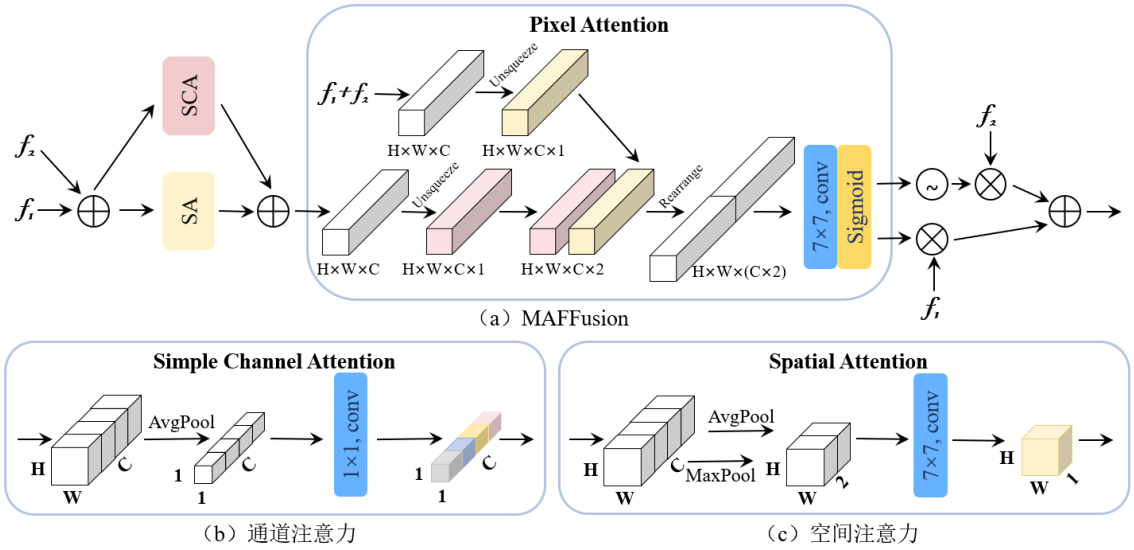
式中，k 表示迭代次数，I^V_0 表示低光输入的 V 通道。全局亮度特征图能够使网络模型在图像增强过程中更好地掌握和调整图像的亮度分布，从而提升增强图像的视觉效果和质量。

1.2　特征融合增强模块

由于图像中通常包含大量冗余信息，注意力机制能够帮助模型聚焦于关键内容并忽略无用信息。因此，本文在特征融合增强模块中整合了多种注意力机制，从而生成更具区分度的特征表示。如图 2 所示，该模块首先将连接后的低照度输入图像及其全局特征通过一个 3×3 的卷积将 20 通道调整为 16 通道，记为 f_1。然后通过一个局部接收场提取局部特征，记为 f_2，该局部接收场由两个膨胀因子为 2 的 3×3 膨胀卷积组成，在不增加计算开销的基础上扩展感受野，逐步整合精炼局部特征。最后将全局特征 f_1 和局部特征 f_2 输入多注意力特征融合模块，以实现全面的信息整合和特征融合，而后利用一个 3×3 的卷积将通道数恢复为 3，从而生成增强后图像。

为了使全局和局部特征得到最大融合，本文提出了一种融合简单通道注意力（Simple Channel Attention，SCA）、空间注意力（Spatial Attention，SA）和像素注意力（Pixel Attention，PA）的多注意力特征融合模块，该模块能够将全局特征 f_1 和局部特征 f_2 有效融合，从而提升图像亮度、保留图像细节。

SCA 采用通道权重分配机制，通过为各通道赋权，增强了全局特征的表达能力，使模型能够自动选择对任务最为重要的特征通道，具体如图 5（b）所示。SA 通过为各空间位置分配权重，聚焦图像关键区域，增强模型对局部空间特征的关注，使得模型在处理包含复杂场景或多物体的图像时能够更有效地定位关键区域，具体如图 5（c）所示。SA 首先在通道维度上对输入特征分别进行全局平均池化和全局最大池化，将输出进行拼接后通过一个 7×7 卷积运算生成空间注意力权重。



（a）MAFFusion

（b）通道注意力

（c）空间注意力

图 5　MAFFusion

Fig. 5　Framework of the MAFFusion

然而，通道注意力和空间注意力在某种程度上可能导致信息的“平均化”，会忽略一些局部细节的重要特征，因此本文引入了像素注意力以弥补全局和局部信息处理过程中可能遗漏的细节信息，加强细节特征提取。具体如图 5（a）所示，将特征图 f_1 和 f_2 相加，通过 SCA 块和 SA 块生成通道注意力权重和空间注意力权重，将两个权重相加输入像素注意力块。PA 首先分别为空间注意力和通道注意力结合后的权重以及 f_1、f_2 相加后的特征图通过 Unsqueeze 操作增加一个新维度，然后将得到的结果沿着新增加的维度进行拼接，通过 Rearrange 操作进行维度转换，方便卷积操作，最后通过 7×7 卷积及 Sigmoid 激活生成像素注意力权重。像素注意力块有助于融合和强化通道注意力与空间注意力的输出，使得模型能够更好地结合全局与局部特征，实现对输入数据更加全面深入的理解。将生成的像素注意力权重与特征图 f_1、f_2 进行互补加权，从而获得特征融合注意力图。

1.3　损失函数

本文设计了一个联合损失函数，由平滑损失（Smooth Loss）、感知损失（Perceptual Loss）、直方图损失（Histogram Loss）、色彩损失（Color Loss）、PSNR 损失和 SSIM 损失 6 个部分组成，计算公式如式（4）所示，其中，$\alpha_1 \sim \alpha_6$ 为各个损失对应的权重贡献系数。

$$L = \alpha_1 L_S + \alpha_2 L_{prec} + \alpha_3 L_{hist} + \alpha_4 L_{PSNR} + \alpha_5 L_{color} + \alpha_6 L_{SSIM} \tag{4}$$

平滑损失（Smooth Loss）：为了保证预测图像的平滑度，本文引入 L_1 损失的鲁棒变体平滑 L_1 损失。该损失结合 L_1 和 L_2 损失的优势，能够处理小误差，并对异常值保持鲁棒性，确保模型训练的稳定性。具体表示为：

$$L_S = \begin{cases} 0.5 \cdot \left(y_{true} - y_{pred} \right)^2 & , \left| y_{true} - y_{pred} \right| < 1 \\ \left| y_{true} - y_{pred} \right| - 0.5 & , otherwise \end{cases} \tag{5}$$

感知损失（Perceptual Loss）：为了利用图像的深层语义信息提升视觉质量，本文采用预训练的 *VGG19* 网络作为特征提取器，用于计算真实图像与预测图像之间的差异，具体表示为：

$$L_{prec}=\frac{1}{N}\sum_{i=1}^{N}\frac{1}{C_iH_iW_i}\left\|VGG_i\left(y_{pred}\right)-VGG_i\left(y_{true}\right)\right\|_2^2 \tag{6}$$

式中，N 表示特征总数，VGG_i 表示预训练的 VGG 网络第 i 层的输出，C_i、H_i、W_i 为特征图的维数。

直方图损失（Histogram Loss）：在低照度图像的直方图中，像素点大多集中在低像素值区域，而正常光图像的直方图则呈现出更加均衡的分布。为了将真实图像和预测图像的像素强度分布对齐，本文引入直方图损失对全局亮度进行调整，具体表示为：

$$L_{hist}=\sum_{b=1}^{B}\left|\frac{hist\left(y_{true},b\right)}{\sum_{b=1}^{B}hist\left(y_{true},b\right)}-\frac{hist\left(y_{pred},b\right)}{\sum_{b=1}^{B}hist\left(y_{pred},b\right)}\right| \tag{7}$$

式中，$hist(y, b)$表示图像 y 的直方图第 b 个 bin 的像素数，B 表示直方图中 bin 的总数。

色彩损失（Color Loss）：为了确保预测图像和真实图像之间的色彩保真度，本文从 HSV 色彩空间（色相、饱和度、亮度）中提取图像的颜色信息，当两个像素满足式（8）时，它们具有相同的色相（H）和饱和度（S）[29]：

$$y_{true,ij}=\lambda\cdot y_{pred,ij} \tag{8}$$

式中，λ 表示任意正数，$y_{true,ij}$ 和 $y_{pred,ij}$ 都为三维向量，表示图像在像素点(i, j)的颜色信息向量。因此，可以采用余弦相似度来测量两个像素之间的色相（H）和饱和度（S）差[30]，具体表示为：

$$L_{color}=1-\sum_{i=1,j=1}^{m,n}<y_{true,ij},y_{pred,ij}> \tag{9}$$

式中，$<\cdot,\cdot>$ 表示两个向量的余弦相似度。

PSNR 损失：为了确保预测图像的细节清晰度，避免模糊，本文引入 PSNR 损失以调节图像中的噪声，L_{PSNR} 的定义如下：

$$L_{PSNR}=40.0-10\log_{10}\left(\frac{MAX_I^2}{MSE\left(y_{true},y_{pred}\right)}\right) \tag{10}$$

式中，MAX_I 表示输入图像的像素最大值（对于 8 位图像，MAX_I=255），$MSE()$表示对两个图像均方误差的计算。本文将 40.0dB 设定为高质量参考值，作为衡量低照度图像增强质量标准，具体设定参考文献[31]。

SSIM 损失：为了保留图像的结构和纹理，改善预测图像整体视觉效果，本文引入结构相似性损失 SSIM 以约束真实图像和预测图像之间的结构一致性，具体表示为：

$$L_{SSIM}=1-\frac{1}{N}\sum_{img}\left(\frac{2\mu_x\mu_y+C_1}{\mu_x^2+\mu_y^2+C_1}\cdot\frac{2\sigma_{xy}+C_2}{\sigma_x^2+\sigma_y^2+C_2}\right) \tag{11}$$

式中，x，y 分别表示预测图像和真实图像，μ_x，μ_y 表示 x 和 y 的像素平均值，σ_x^2，σ_y^2 表示像素值的方差，σ_{xy} 表示像素的协方差，C_1 和 C_2 为避免分母为 0 的常量。

为平衡各个损失函数约束效果，避免冗余，本文在文献[31]基础上根据各个损失在模型训练中的作用效果进行实验调节，进而确定最优参数组合，将权重贡献系数 $\alpha_1\sim\alpha_6$ 设定为 1.1，0.07，0.06，0.01，0.26，0.5。

2 实验结果分析

2.1 实验设计

本文框架在 NVIDIA Geforce RTX 3060 Laptop GPU 上使用 PyTorch 实现，实验设置中批大小固定为 171，训练轮次固定为 5000，选用 Adam 为优化器，将学习率固定为 0.0001。为评估所提方法的实际效果，本文选取真实数据集 LOL[17]进行训练与测试工作。LOL 数据集由 500 组低光图像及相应的正常光配对图像组成，

按照作者的设置规范，本文选取其中 485 组图像用于模型训练，余下 15 组图像用于测试算法性能，表示为 LOL_test。在此基础上，本文还选取了多个真实数据集对所提方法的泛化性能进行验证，包括 LOL_v2[32]、LSRW[33]和 SICE[34]。LOL_v2 数据集由 LOL_v2_real 和 LOL_v2_synthetic 两个子集组成，其中 LOL_v2_real 子集包含 689 组训练图像和 100 组测试图像，本文选取其测试集进行测试，表示为 LOL_v2_test；LSRW 数据集包含 5650 组由华为 P40 Pro 手机和尼康 D7500 相机拍摄的配对图像，涵盖了室内和室外场景下的低照度图像，本文选取华为手机拍摄的 30 组图像及尼康相机拍摄的 20 组图像进行测试，分别表示为 LSRW_Huawei_test 和 LSRW_Nikon_test；SICE 数据集专门用于单张图像对比度增强，包含多种场景和高分辨率的多曝光序列，本文选取 50 组室内外低光数据进行测试，表示为 SICE_test。这些多样化的测试图像涵盖了不同场景与环境，为全面评估本文方法在复杂场景下的增强效果提供了可靠依据。

2.2 评价指标

本文选用峰值信噪比（Peak Signal to Noise Ratio，PSNR）、结构相似性（Structural Similarity，SSIM）、学习感知图像相似度（Learned Perceptual Image Patch Similarity，LPIPS）和均方误差（Mean Square Error，MSE）作为模型评价指标。其中，PSNR 通过像素点间的误差计算来评估图像质量，其值越高表明图像质量也相应提高；SSIM 从亮度、对比度和结构三个维度对图像进行全面考量，其值大小能够反映两张图像的相似性；LPIPS 指标是一种更加贴合人类视觉感知的评价标准，用于量化两张图像之间的差异，其值越低表明两张图像越接近；MSE 通过度量两张图像间的差异来判断失真程度，越低的 MSE 值表明图像失真越小。

2.3 实验结果

为了评估算法的性能，将本文与近几年主流低照度增强算法 RetinexNet[17]、KinD[13]、EnlightenGAN[23]、Zero-DCE++[22]错误!未找到引用源.、PairLIE[35]、FLWNet[28]、HWMNet[36]、IAT[37]进行对比实验，其中，RetinexNet、KinD 为基于 Retinex 理论的监督学习方法，EnlightenGAN、PairLIE 为无监督学习方法，Zero-DCE++为零次学习方法，FLWNet、HWMNet、IAT 以及本文方法均为端到端增强方法。为保证实验的公正性，所有训练测试均在一致的环境下开展。

2.3.1 定性分析

为了检验所提方法在实际低光环境中对拍摄图像的增强能力，本文将其与 7 种主流方法进行对比研究。图 6 展示了不同低照度图像增强方法在 LOL_test 数据集上的主观视觉效果对比，如图所示，RetinexNet、PairLIE 增强后图像的亮度和对比度虽然有所提升，但存在明显噪声，图像失真严重。EnlightenGAN、Zero-DCE++、IAT 都存在图像整体亮度提升不足的现象，其中，EnlightenGAN 增强后黑暗处出现明显晕影，存在部分特征丢失现象；Zero-DCE++增强后的图像出现了显著的偏色（蓝）问题，IAT 增强后框图中的数字面板明亮区域局部曝光严重，产生了明显的光晕。KinD 对于数字面板的增强同样存在局部曝光严重的问题，导致背景墙面模糊，细节纹理丢失；FLWNet 能够有效保留图像的细节纹理，未出现明显的纹理失真，但仍存在光晕问题。实验结果表明，本文方法所增强的图像在亮度、对比度方面均有显著改善，且无颜色失真和纹理畸变，达到了良好的增强效果。

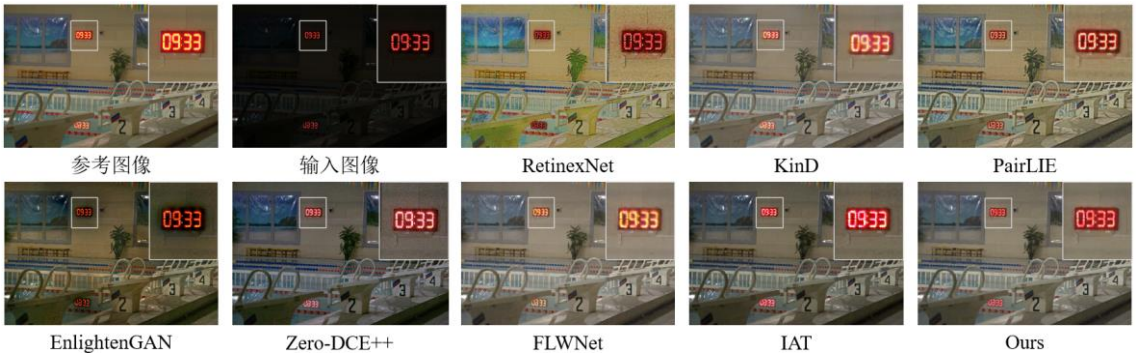


图 6　不同方法在 LOL_test 数据集上的定性对比结果

Fig. 6　Qualitative comparison with SOTA methods on LOL_test dataset

在实际应用中，能否有效处理真实环境下的复杂光照条件是一个巨大的挑战，因此，本文将基于 LOL 数据集训练的模型在 LOL_v2_test、LSRW_Huawei_test、LSRW_Nikon_test 和 SICE_test 多个数据集上进行了对比试验。如图 7 所示，RetinexNet、PairLIE、EnlightenGAN、Zero-DCE++增强后图像噪点严重，存在明显的色彩失真问题。IAT 虽然对整体亮度和对比度都有所提升，但其过高的饱和度致使纹理过于锐利，出现了图像失真现象。KinD、FLWNet 增强后的图像颜色保真度较好，但相较于本文所提方法，在图像右侧背景玻璃门的纹理细节保留上仍有所欠缺。本文方法能够在纹理细节、图像清晰度及色彩还原上都更加接近参考图像，在 LOL_v2_test 数据集上也获得了更优越的结果。

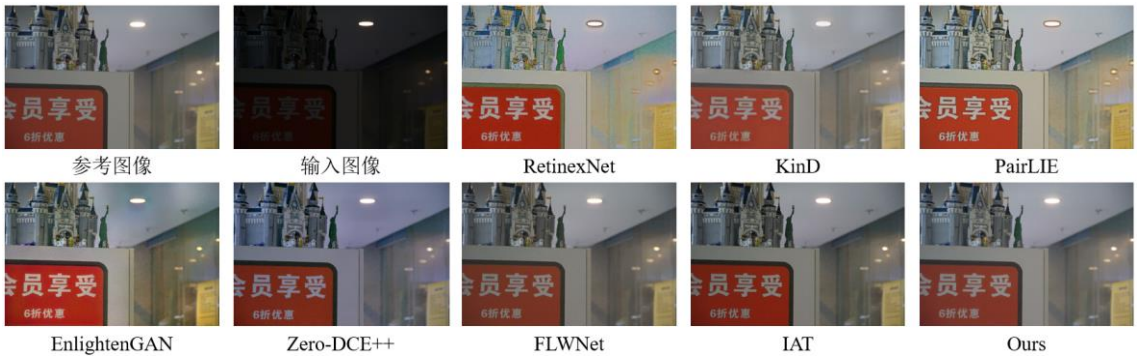


图 7　不同方法在 LOL_v2_test 数据集上的定性对比结果

Fig. 7　Qualitative comparison with SOTA methods on LOL_v2_test dataset

如图 8~10 所示，不同方法在低照度图像增强任务中表现出各自的优缺点。RetinexNet 在三个数据集上的增强结果均存在明显的颜色失真、噪点以及晕影问题；KinD 尽管在颜色还原上有所改善，但存在颜色不均的现象，且产生了与预测图像不一致的阴影；PairLIE 能够较准确地表征颜色和亮度，但增强后的图像噪点显著，影响视觉体验；Zero-DCE++在不同数据集上均出现图像偏色（蓝调）问题，同时亮度增强不足，难以满足图像增强的实际需求；IAT 方法在保留图像颜色及细节信息上表现较好，但整体亮度提升不足，增强效果有限；HWMNet 与 FLWNet 在提升图像亮度和对比度方面效果明显，但相较于本文所提方法，其增强结果在颜色保真度方面仍显不足，难以精确还原真实场景的色彩信息。相比之下，本文方法在多个真实数据集上的表现均达到了令人满意的效果。所提方法不仅能够显著提升低照度图像的亮度，同时在纹理细节保留、图像清晰度以及色彩还原方面均优于现有方法，充分展现了其在低照度图像增强任务中的鲁棒性和泛化能力。

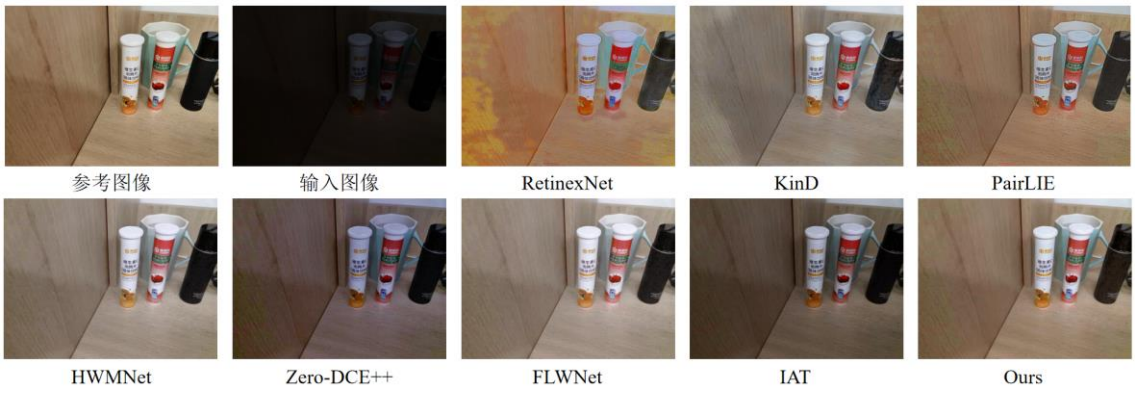


图 8　不同方法在 LSRW_Huawei_test 数据集上的定性对比结果

Fig. 8　Qualitative comparison with SOTA methods on LSRW_Huawei_test dataset

图 9　不同方法在 LSRW_Nikon_test 数据集上的定性对比结果

Fig. 9　Qualitative comparison with SOTA methods on LSRW_Nikon_test dataset

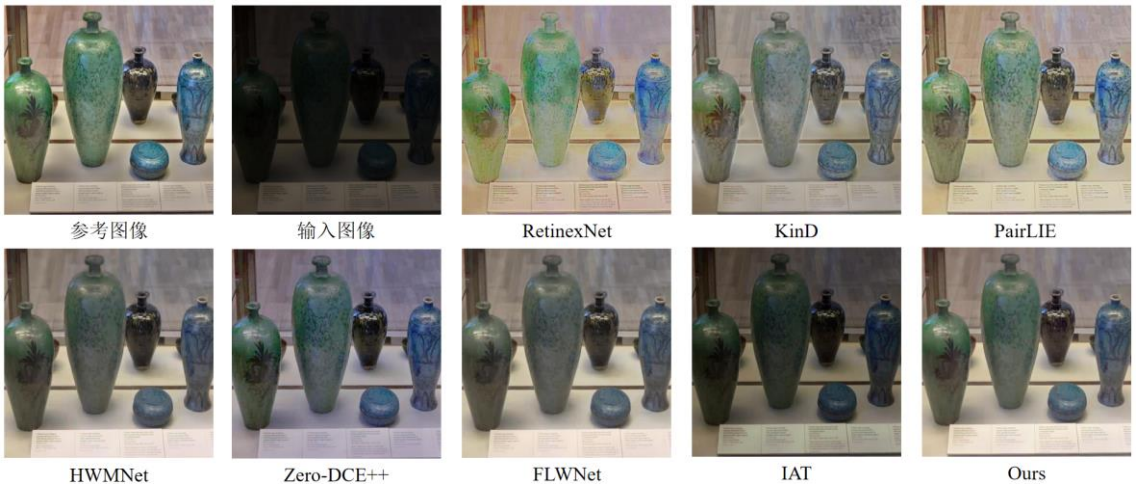
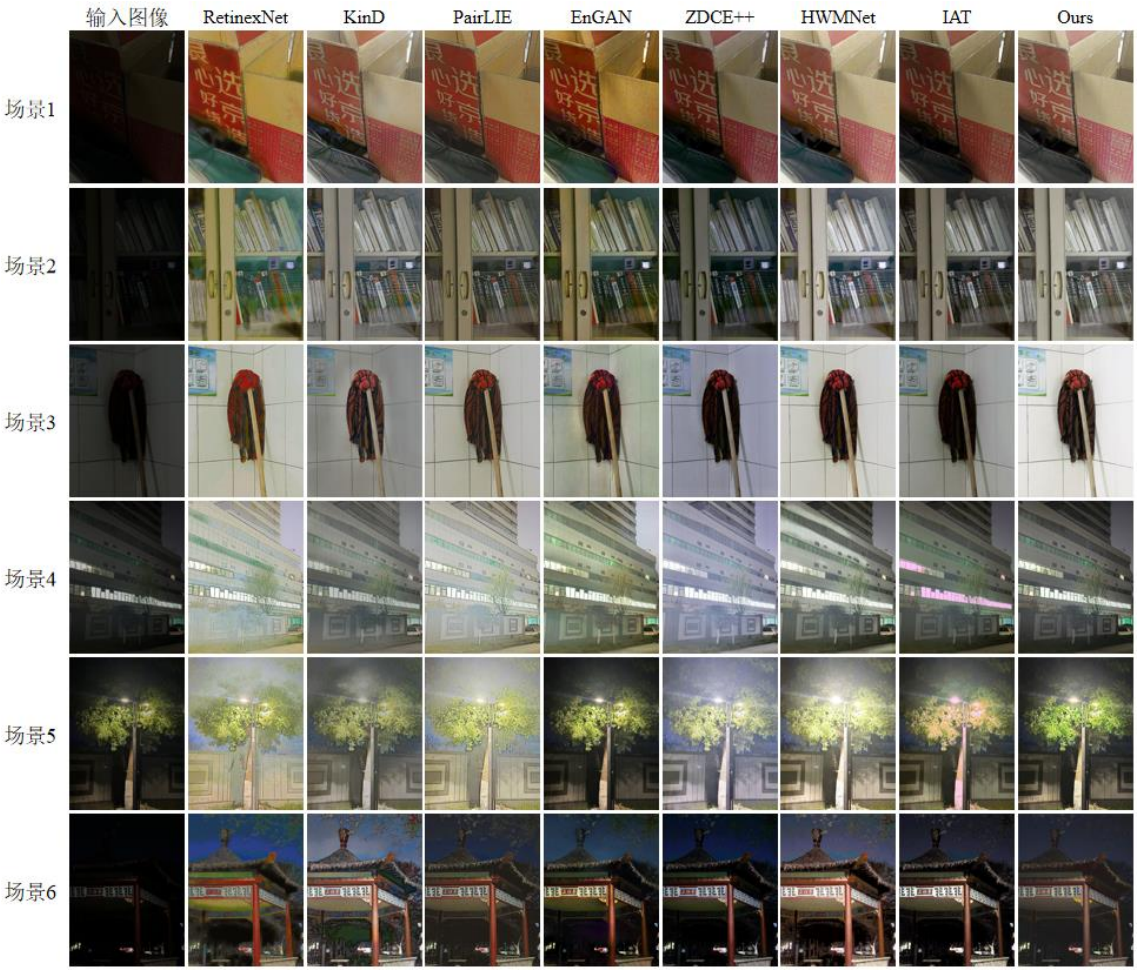


图 10　不同方法在 SICE_test 数据集上的定性对比结果

Fig. 10　Qualitative comparison with SOTA methods on SICE_test dataset

为验证本文所提方法在真实低照度条件下对室内外实拍图像的增强效果，本文选取南京邮电大学（三牌楼校区）校园内 6 处实拍场景（室内外各 3 处）进行对比实验，拍摄设备为 iPhone14 Pro，图像分辨率均为 3024×4032 像素。测试场景涵盖室内反光玻璃、高饱和物体以及室外非均匀光照、夜空等复杂条件，所提方法与其他 7 种低照度图像增强方法的对比结果如图 11 所示。从图中可以看出，RetinexNet 虽能显著提升图像整体亮度与对比度，但室内场景增强后颜色饱和度过高，导致纹理突出，图像失真严重，室外场景增强后出现明显光晕与噪声扩散；KinD 虽在色彩与亮度还原方面表现较优，但存在黑色晕影现象（场景 3 墙面与拖把边缘产生断层），纹理细节模糊。PairLIE 对室内场景的处理结果整体色调偏暗黄，且暗部噪点显著，同时高光区域存在晕影扩散（场景 4 和场景 5 的灯光光晕覆盖场景细节）。EnlightenGAN 在室内外场景均产生泛黄晕影，且边缘细节模糊。Zero-DCE++存在整体亮度不足与偏色（蓝）问题，其灯光处理亦出现晕染效应（场景 4 和场景 5 灯光光晕过度扩散导致图像整体泛白）。HWMNet 虽在亮度提升与暗部细节保留方面表现突出，但局部色彩失衡问题显著（场景 1 的黑色皮质座椅泛绿，场景 4 的灰色墙体呈现不规则白斑）。IAT 全局亮度提升不足导致视觉观感昏暗，且在强光源区域出现严重色偏（场景 4 和场景 5 的灯光粉色异常渲染）。相比之下，本文方法在保持色彩自然性（场景 1 纸箱和场景 3 拖把本色还原准确）与纹理保真度（场景 2 的柜门玻璃反射纹理连贯）的前提下，实现了均衡的亮度对比度调整，有效抑制了光晕伪影（场景 4 和场景 5 的灯光周围场景清晰）、噪声增殖（场景 6 的暗部噪点平滑）与局部色偏（场景 3 墙面和场景 6 夜空色调纯净），在复杂室内外低照度条件下均展现出更符合人类视觉感知的增强效果。



注：图中 EnGAN、ZDCE++分别表示 EnlightenGAN、Zero-DCE++

图 11　不同方法在实拍图像上的定性对比结果

Fig. 11　Qualitative comparison with SOTA methods on real images

2.3.2　定量分析

本文采用 PSNR、SSIM、LPIPS 和 MSE 作为评价指标，将所提方法与其他算法在多个真实数据集上进行定量对比分析。选取计算量 Flops 和计算量 Params 在 400×600 像素的图像上进行效率对比，结果如表 1~2 所示，其中，最优结果以加粗形式展示，次优结果则通过下划线加以标注。

如表 1~2 所示，本文所提方法在 LOL_test 数据集上的 PSNR 和 MSE 指标均为最优，与次优算法相比 PSNR 提升了 3.66%，说明本文算法在改善亮度、抑制噪声等性能上有显著优势；在 SSIM 和 LPIPS 指标均为次优，但相较于最优算法 HWMNet 的 943.3857G 计算量和 66.5647M 参数量，本文仅需要 3.0494G 计算量和 0.0149M 参数量即可达到相近的结果，显示出更优的综合性能。在 LOL_v2_test、LSRW_Huawei_test、LSRW_Nikon_test 和 SICE_test 数据集上本文方法在各个指标上都达到了最优或次优的结果，并表现出稳定的增强效果。这表明，相较于对比方法，本文所提方法在非训练数据的真实场景下同样能够提供更优的增强效果，从定量分析的角度验证了其较强的鲁棒性和泛化能力。

表 1　不同方法在 LOL_test 和 LOL_v2_test 上的定量对比结果

Table 1 Quantitative comparison with SOTA methods on LOL_test and LOL_v2_test

Method	Complexity		LOL_test (15)				LOL_v2_test (100)			
	FLOPS(G)	Params(M)	PSNR↑	SSIM↑	LPIPS↓	MSE↓	PSNR↑	SSIM↑	LPIPS↓	MSE↓

RetinexNet	79.5587	0.5552	17.8037	0.6126	0.3822	0.0193	17.4371	0.5950	0.4126	0.0204
KinD	236.7537	15.7969	19.6633	0.8160	0.1978	0.0146	20.7775	0.8482	0.1941	0.0105
EnlightenGAN	48.7149	8.6367	17.4829	0.6515	0.3268	0.0307	18.6397	0.6767	0.3213	0.0186
Zero-DCE++	**0.0245**	**0.0106**	15.5527	0.5730	0.3253	0.0447	18.4032	0.5855	0.2972	0.0241
PairLIE	81.8381	0.3418	19.4205	0.7342	0.2572	0.0166	18.7751	0.7487	0.2578	0.0186
HWMNet	943.3857	66.5647	24.2424	**0.8532**	**0.1138**	0.0081	**30.2964**	0.8914	**0.0800**	0.0021
IAT	5.2564	0.0869	23.3834	0.8090	0.2356	0.0056	26.4603	0.8448	0.2226	0.0034
FLWNet	3.8016	0.0174	24.1028	0.8380	0.2258	0.0054	27.8379	0.8741	0.2214	0.0029
Ours	3.0494	0.0149	**24.9848**	0.8522	0.1864	**0.0040**	28.7353	**0.9197**	0.1663	**0.0017**

表 2　不同方法在 LSRW_Huawei_test、LSRW_Nikon_test 和 SICE_test 上的定量对比结果

Table 2 Quantitative comparison with SOTA methods on LSRW_Huawei_test，LSRW_Nikon_test and SICE_test

Method	LSRW_Huawei_test (30)				LSRW_Nikon_test (20)				SICE_test (50)			
	PSNR↑	SSIM↑	LPIPS↓	MSE↓	PSNR↑	SSIM↑	LPIPS↓	MSE↓	PSNR↑	SSIM↑	LPIPS↓	MSE↓
RetinexNet	16.7783	0.4433	0.4411	0.0237	13.746	0.3189	0.3737	0.0456	19.2904	0.7790	0.2864	0.0129
KinD	17.1794	0.5429	0.3806	0.0213	14.7765	0.4342	0.3137	0.035	20.0311	0.8277	0.2314	0.0129
Zero-DCE++	16.8730	0.4773	0.3843	0.0260	15.8629	0.4289	0.3055	0.0341	19.7917	0.8414	0.2330	0.0135
PairLIE	18.6654	0.5406	0.3778	0.0165	14.5293	0.4088	0.2676	0.0403	19.8464	0.8447	0.2321	**0.0097**
HWMNet	18.4997	0.5887	0.3841	0.0177	15.7409	0.4711	0.2797	0.0358	18.6594	0.8369	0.2106	0.0177
IAT	17.1434	0.5366	0.3634	0.0253	15.4107	0.4844	0.3009	0.0456	14.9152	0.7447	0.2290	0.0385
FLWNet	21.2603	0.6231	0.3699	**0.0085**	**18.3256**	0.4975	0.2887	**0.0174**	16.2964	0.5361	0.3492	0.027
Ours	**22.1853**	**0.6388**	**0.3600**	0.0088	18.2040	**0.4999**	**0.2369**	0.0179	**20.7548**	**0.8615**	**0.1772**	0.0109

2.4　消融实验

为验证所提网络框架中简单门控注意力模块（SGABlock）和多注意力特征融合模块（MAFFusion）的有效性，本文设置了两组消融实验以进行分析。一是评估 SGABlock、MAFFusion 两个模块对于整体网络的必要性，对模块整体进行消融实验；二是在不改变整体网络架构的基础上，单独对各模块内部组件进行消融实验，以便深入理解各个部分的功能与贡献。

2.4.1　模块消融实验

为评估简单门控注意力模块和多注意力特征融合模块对图像增强效果的影响，本文共设计以下 3 种实验组合：①M1：将 SGABlock 模块替换为 2 个 3×3 卷积；②M2：将 MAFFusion 模块替换为连接操作（Concat）；③M1+M2：结合 SGABlock、MAFFusion 模块进行网络训练。**不同组合的定性对比结果如图 12 所示，结果表明，替换 SGABlock 模块导致图像细节退化，M1 增强结果中柜门图案边缘模糊，暗部噪点显著增多；替换 MAFFusion 模块导致增强图像在色彩还原（木柜、文件夹等饱和度较低）和暗处纹理细节保留方面均有所降低，完整模型在纹理保留、噪声抑制和色彩一致性上表现最优。**不同组合的定量对比结果如表 3 所示，其中最优结果已加粗展示。结果表明，将 SGABlock 模块替换后，峰值信噪比显著下降，表明该模块在减少图像噪声方面起到了关键作用；将 MAFFusion 模块替换后，各项评价指标均低于原算法，**表明该模块能有效协调全局与局部特征，**增强模型表征能力。**完整模型在各项指标上均达到最优，且参数量与计算量最低。**

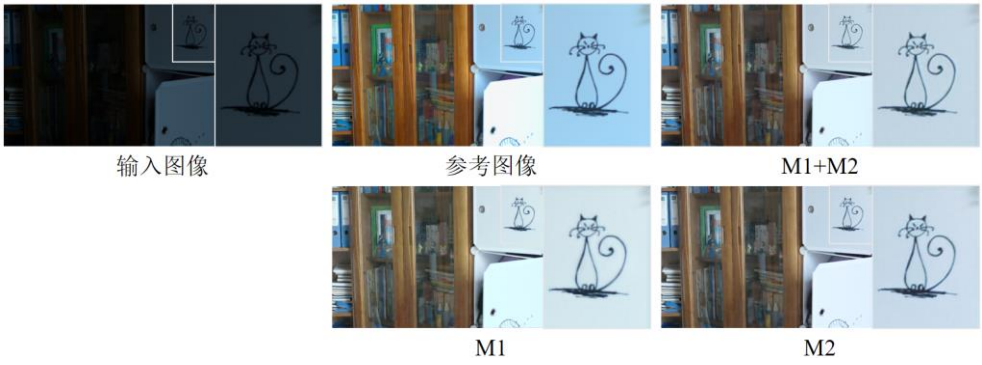


图 12　模块消融实验对比图

Fig. 12　Comparison of module ablation experiments

表 3　模块消融实验对比结果

Table 3 Comparative results of module ablation experiments

Combination	Module		Evaluation Metric				Complexity	
	SGABlock	MAFFusion	PSNR↑	SSIM↑	LPIPS↓	MSE↓	FLOPS(G)	Params(M)
M1		√	24.2744	0.8320	0.2083	0.0046	3.1608	0.0150
M2	√		24.4359	0.8287	0.2161	0.0044	3.6902	0.0173
M1+M2	√	√	**24.9848**	**0.8522**	**0.1864**	**0.0040**	**3.0494**	**0.0149**

2.4.2　模块内部组件消融实验

为验证简单门控注意力模块和多注意力特征融合模块中各个组件对图像增强性能的影响，对 2 个模块中的不同组件进行消融实验。

针对简单门控注意力模块中层归一化（LayerNorm）、简单门控（SimpleGate）和简单通道注意力（SCA）3 个组件进行消融实验，实验组合设置如下：①M1：仅删除 LayerNorm 层；②M2：将 SimpleGate 层替换为激活函数 LeakyReLU；③M3：将 SCA 组件替换为 CA；④M1+M2+M3：结合 LayerNorm、SimpleGate、SCA 组件构建简单门控注意力模块进行网络训练。**不同组合的定性对比结果如图 13 所示，结果显示，移除任一组件均会导致性能下降，三者组合的增强效果在色彩保真度与细节清晰度上最接近参考图像，说明各个组件均能够一定程度上帮助网络更好地还原颜色细节。**不同组合的定量对比结果如表 4 所示，其中最优结果已加粗展示。**从表中可以看出，定量对比结果与定性分析结果一致，**移除三个组件中的任何一个，均会导致模型性能在不同程度上出现下降，三者组合指标最优，说明简单门控注意力模块中设置的各个组件对于模型性能都有增益效果。

表 4　SGABlock 模块组件消融实验对比结果

Table 4 Comparative results of ablation experiments of SGABlock module components

Combination	Component			Evaluation Metric			
	LayerNorm	SimpleGate	SCA	PSNR↑	SSIM↑	LPIPS↓	MSE↓
M1		√	√	24.4985	0.8436	0.2044	0.0045
M2	√		√	24.5942	0.8473	0.1967	0.0041
M3	√	√		24.4267	0.8448	0.2042	0.0044
M1+M2+M3	√	√	√	**24.9848**	**0.8522**	**0.1864**	**0.0040**

图 13　SGABlock 模块组件消融实验对比图

Fig. 13　Comparison of ablation experiments of SGABlock module components

针对多注意力特征融合模块中简单通道注意力（SCA）和像素注意力（PA）2 个组件设计消融实验，共有以下 3 种实验组合：①M1：将 SCA 组件替换为 CA；②M2：仅删除 PA 组件；③M1+M2：结合 SCA 和 PA 组件构建多注意力特征融合模块进行网络训练。**不同组合的定性对比结果如图 14 所示，替换 SCA 组件导致增强后图像数字表盘区域出现了色彩不均和晕影的现象；删除 PA 组件后的增强结果整体模糊，数字表盘区域出现了明显的光晕，噪点显著；完整模型在抑制噪声与恢复细节上表现最佳。**不同组合的定量对比结果如表 5 所示，其中最优结果已加粗展示。结果表明，替换 SCA 组件后，各项评价指标均有所降低，说明该组件能够在简化网络模型的同时提高增强效果；删除 PA 组件后，PSNR 指标明显降低，说明该组件在融合注意力特征，抑制噪声方面有着至关重要的作用。

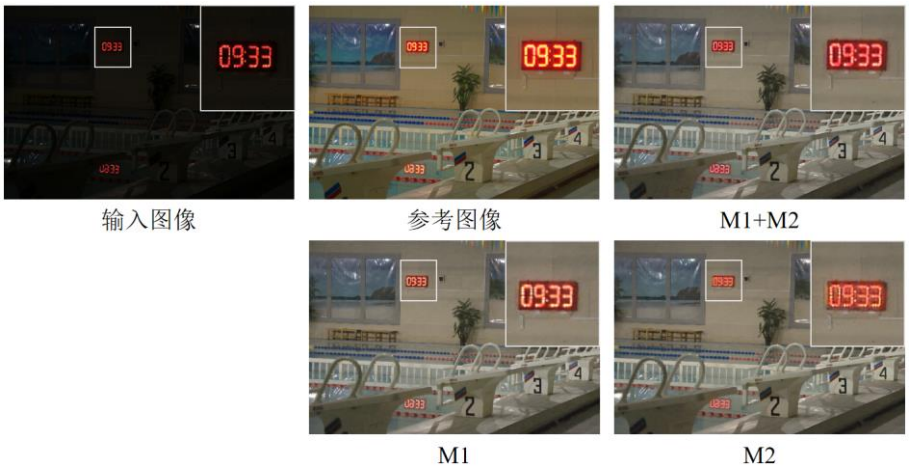


图 14　MAFFusion 模块组件消融实验对比

Fig. 14　Comparison of ablation experiments of MAFFusion module components

表 5　MAFFusion 模块组件消融实验对比结果

Table 5 Comparative results of ablation experiments of MAFFusion module components

Combination	Component		Evaluation Metric			
	SCA	PA	PSNR↑	SSIM↑	LPIPS↓	MSE↓
M1		√	24.6462	0.8383	0.2233	0.0044

M2	√		24.5139	0.8414	0.2124	0.0044
M1+M2	√	√	**24.9848**	**0.8522**	**0.1864**	**0.0040**

3 结束语

现有低照度图像增强方法常面临计算效率低下、噪声过多、色彩失真、细节丢失等问题，为此，本文基于多注意力特征融合提出了一种轻量级方法，该方法分为全局特征提取模块和特征融合增强模块两个部分。全局特征提取模块通过简单门控注意力模块在通道信息指导下提取输入图像全局细节信息，并简化注意力及门控单元组件以简化网络；通过从低照度输入 V 通道的直方图中提取亮度信息，并通过高阶曲线调整方法生成全局亮度特征图。特征融合增强模块通过多注意力融合模块对全局特征及局部接收场提取的局部特征进行信息整合，借助像素注意力进一步强化通道及空间注意力对于全局及局部特征的表征，提高模型的色彩还原和噪声抑制能力。在联合损失函数约束下，所提方法能够有效地提升图像的亮度、色彩表现与细节纹理，抑制噪声，具有良好的视觉效果。将本文方法与最先进方法在多个真实数据集上进行主客观对比，皆达到了更优的综合性能。现实世界的光照场景非常丰富，本文也将在未来进一步研究如何提升模型的泛化能力和鲁棒性，以适应更加复杂的光照场景。

参考文献：

[1] 孙涵, 刘译善, 林昱涵. 基于深度学习的显著性目标检测综述[J]. 数据采集与处理, 2023, 38(1): 21-50.
Sun Han, Liu Yishan, Lin Yuhan. Deep Learning Based Salient Object Detection: A Survey[J]. Journal of Data Acquisition and Processing, 2023, 38(1): 21-50.

[2] 吴鹏, 张孙杰, 王永雄, 等. 基于感知推理和外部空间先验特征的图像修复[J]. 数据采集与处理, 2024, (4): 933-943.
Wu Peng, Zhang Sunjie, Wang Yongxiong, et al. Image Inpainting Based on Perceptual Inference and External Spatial Prior Features[J]. Journal of Data Acquisition and Processing, 2024, (4): 933-943.

[3] Yeganeh H, Ziaei A, Rezaie A. A novel approach for contrast enhancement based on histogram equalization[C]//2008 International Conference on Computer and Communication Engineering. IEEE, 2008: 256-260.

[4] Huang S C, Cheng F C, Chiu Y S. Efficient contrast enhancement using adaptive gamma correction with weighting distribution[J]. IEEE transactions on image processing, 2012, 22(3): 1032-1041.

[5] Dorothy R, Joany R M, Rathish R J, et al. Image enhancement by histogram equalization[J]. International Journal of Nano Corrosion Science and Engineering, 2015, 2(4): 21-30.

[6] Pizer S M, Amburn E P, Austin J D, et al. Adaptive histogram equalization and its variations[J]. Computer vision, graphics, and image processing, 1987, 39(3): 355-368.

[7] Hines G, Rahman Z, Jobson D, et al. Single-scale retinex using digital signal processors[C]//Global signal processing conference. 2005: 1324.

[8] Rahman Z, Jobson D J, Woodell G A. Multi-scale retinex for color image enhancement[C]//Proceedings of 3rd IEEE international conference on image processing. IEEE, 1996, 3: 1003-1006.

[9] Jobson D J, Rahman Z, Woodell G A. A multiscale retinex for bridging the gap between color images and the human observation of scenes[J]. IEEE Transactions on Image processing, 1997, 6(7): 965-976.

[10] Lore K G, Akintayo A, Sarkar S. LLNet: A deep autoencoder approach to natural low-light image enhancement[J]. Pattern Recognition, 2017, 61: 650-662.

[11] Lv F, Lu F, Wu J, et al. MBLLEN: Low-light image/video enhancement using cnns[C]//BMVC. 2018, 220(1): 4.

[12] Zamir S W, Arora A, Khan S, et al. Learning enriched features for real image restoration and enhancement[C]//Computer Vision-ECCV 2020: 16th European Conference, Glasgow, UK, August 23-28, 2020, Proceedings, Part XXV 16. Springer International Publishing, 2020: 492-511.

[13] Zhang Y, Zhang J, Guo X. Kindling the darkness: A practical low-light image enhancer[C]//Proceedings of the 27th ACM

international conference on multimedia. 2019: 1632-1640.

[14]Liu R, Ma L, Zhang J, et al. Retinex-inspired unrolling with cooperative prior architecture search for low-light image enhancement[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2021: 10561-10570.

[15]Zhao Z, Xiong B, Wang L, et al. RetinexDIP: A unified deep framework for low-light image enhancement[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2021, 32(3): 1076-1088.

[16]Bai J, Yin Y, He Q. Retinexmamba: Retinex-based Mamba for Low-light Image Enhancement[J]. arXiv preprint arXiv:2405.03349, 2024.

[17]Wei C, Wang W, Yang W, et al. Deep retinex decomposition for low-light enhancement. arXiv 2018[J]. arXiv preprint arXiv:1808.04560.

[18]Wu W, Weng J, Zhang P, et al. Uretinex-net: Retinex-based deep unfolding network for low-light image enhancement[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2022: 5901-5910.

[19]Cai Y, 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: 12504-12513.

[20]Ma L, Ma T, Liu R, 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: 5637-5646.

[21]Guo C, Li C, Guo J, et al. Zero-reference deep curve estimation for low-light image enhancement[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2020: 1780-1789.

[22]Li C, Guo C, Loy C C. Learning to enhance low-light image via zero-reference deep curve estimation[J]. IEEE transactions on pattern analysis and machine intelligence, 2021, 44(8): 4225-4238.

[23]Jiang Y, Gong X, Liu D, et al. Enlightengan: Deep light enhancement without paired supervision[J]. IEEE transactions on image processing, 2021, 30: 2340-2349.

[24]Jiang H, Luo A, Fan H, et al. Low-light image enhancement with wavelet-based diffusion models[J]. ACM Transactions on Graphics (TOG), 2023, 42(6): 1-14.

[25]Jiang H, Luo A, Liu X, et al. Lightendiffusion: Unsupervised low-light image enhancement with latent-retinex diffusion models[C]//European Conference on Computer Vision. Cham: Springer Nature Switzerland, 2024: 161-179.

[26]Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 7132-7141.

[27]Chen L, Chu X, Zhang X, et al. Simple baselines for image restoration[C]//European conference on computer vision. Cham: Springer Nature Switzerland, 2022: 17-33.

[28]Zhang Y, Di X, Wu J, et al. Simplifying low-light image enhancement networks with relative loss functions[J]. arXiv preprint arXiv:2304.02978, 2023.

[29]Zhang Y, Di X, Zhang B, et al. Better than reference in low-light image enhancement: Conditional re-enhancement network[J]. IEEE Transactions on Image Processing, 2021, 31: 759-772.

[30]Liu C, Wu F, Wang X. EFINet: Restoration for low-light images via enhancement-fusion iterative network[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2022, 32(12): 8486-8499.

[31]Brateanu A, Balmez R, Avram A, et al. Lyt-net: Lightweight yuv transformer-based network for low-light image enhancement[J]. arXiv preprint arXiv:2401.15204, 2024.

[32]Jiang H, Luo A, Fan H, et al. Low-light image enhancement with wavelet-based diffusion models[J]. ACM Transactions on Graphics (TOG), 2023, 42(6): 1-14.

[33]Hai J, Xuan Z, Yang R, et al. R2rnet: Low-light image enhancement via real-low to real-normal network[J]. Journal of Visual Communication and Image Representation, 2023, 90: 103712.

[34]Cai J, Gu S, Zhang L. Learning a deep single image contrast enhancer from multi-exposure images[J]. IEEE Transactions on Image Processing, 2018, 27(4): 2049-2062.

[35]Fu Z, Yang Y, Tu X, et al. Learning a simple low-light image enhancer from paired low-light instances[C]//Proceedings of the

IEEE/CVF conference on computer vision and pattern recognition. 2023: 22252-22261.

[36]Fan C M, Liu T J, Liu K H. Half wavelet attention on M-Net+ for low-light image enhancement[C]//2022 IEEE International Conference on Image Processing (ICIP). IEEE, 2022: 3878-3882.

[37]Cui Z, Li K, Gu L, et al. You only need 90k parameters to adapt light: a light weight transformer for image enhancement and exposure correction[J]. arXiv preprint arXiv:2205.14871, 2022.

作者简介：

刘艺(2000−)，女，硕士研究生，研究方向：图像处理、计算机视觉、深度学习等，E-mail: ly1622525373@163.com。

朱佳慧(2000−)，女，硕士研究生，研究方向：目标检测、图像增强、深度学习等。

郑涤尘(2000−)，男，硕士研究生，研究方向：图像处理、计算机视觉等。

张登银(1964−)，**通信作者**，男，博士，研究员，博士生导师，CCF 会员，研究方向：智能信号与信息处理、IP 网络技术（物联网）、信息安全等，E-mail: zhangdy@njupt.edu.cn。