

一种基于小波变换的两阶段低照度图像增强方法

孙 静¹⁾ 孙福奇¹⁾ 郝世杰²⁾ 孙福明¹⁾

¹⁾(大连民族大学信息与通信工程学院 辽宁 大连 116600)

²⁾(合肥工业大学计算机与信息学院 合肥 230009)

摘 要 低照度环境下采集的图像普遍存在亮度衰减、对比度弱化及细节模糊等退化现象,导致视觉质量显著降低,严重制约计算机视觉系统在目标检测、语义分割等高级视觉任务中的性能表现。针对上述问题,本文提出了一种基于小波变换的两阶段低照度图像增强网络 TSUNet (Two-Stage Wavelet Recovery U-Net)。本文创新性地构建了基于小波变换理论的 U 型网络架构,通过初级恢复与精细增强两阶段的渐进式处理,分别实现基础特征重建和细节特征优化。为提升网络的特征表达能力,本文设计了增强小波域特征融合模块,该模块集成离散小波变换与逆变换操作,并设计了由动态门控空间注意力与轻量融合曲线注意力组成的双重注意力机制,通过双重注意力机制与小波变换协同工作,以更精细化的方式实现噪声抑制与细节增强的平衡。在优化策略方面,本文提出了融合感知损失函数,通过综合考量像素级误差与视觉感知质量,引导模型生成具有自然的视觉效果的高质量图像。实验结果表明,本文提出的方法在多个公开低照度数据集的关键指标(如峰值信噪比、结构相似性指数)中展现出出色的性能。代码已开源在 <https://github.com/HibobacX/TSUNet>。

关键词 图像增强;U 型网络;小波变换;注意力机制;损失函数

中图法分类号 TP391

DOI 号 10.11897/SP.J.1016.2025.01188

Two-Stage Low-Light Image Enhancement Based on Wavelet Transform

SUN Jing¹⁾ SUN Fu-Qi¹⁾ HAO Shi-Jie²⁾ SUN Fu-Ming¹⁾

¹⁾(School of Information and Communication Engineering, Dalian Minzu University, Dalian, Liaoning 116600)

²⁾(School of Computer Science and Information Engineering, Hefei University of Technology, Hefei 230009)

Abstract In realistic scenarios characterized by complex and dynamic environmental conditions, image acquisition devices frequently encounter insufficient ambient illumination, leading to widespread quality degradation issues such as suboptimal luminance levels, inadequate contrast ratios, and compromised detail retention. This significant deterioration in image quality not only degrades subjective visual perception for human observers but, more critically, hinders the operational effectiveness of computer vision systems in performing advanced visual tasks-including object detection and semantic segmentation-resulting in notable declines in both accuracy and reliability when processing such degraded imagery. The performance implications of low-light imaging are profound and multifaceted. Such performance degradation can be quantitatively assessed through considerable reductions in key evaluation metrics, consequently impeding the deployment of these vision systems in demanding low-light applications across diverse domains such as autonomous driving, medical imaging, security surveillance, and scientific research. The challenges are particularly acute in scenarios where precise visual information is paramount for decision-making.

收稿日期:2024-07-19;在线发布日期:2025-03-07。本课题得到国家自然科学基金(No. 62172137)、辽宁省科技计划联合计划项目(No. 2024JH2/102600091)、大连市科技人才创新支持政策实施计划(No. 2022RQ086)、中央高校基本科研业务费(No. 04442024061)资助。
孙 静(通信作者),博士,讲师,中国计算机学会(CCF)会员,主要研究领域为计算机视觉。E-mail:jingsun@dlmu.edu.cn。孙福奇,硕士研究生,中国计算机学会(CCF)会员,主要研究领域为图像处理。郝世杰,博士,教授,主要研究领域为图像处理、计算机视觉。孙福明,博士,教授,主要研究领域为图像处理、计算机视觉。

ing and analytical processes. While conventional enhancement algorithms (e. g. , histogram equalization and Retinex theory) provide fundamental luminance correction, they remain susceptible to noise amplification and lack adaptive modeling capabilities for complex degradation patterns. Recent deep learning-based enhancement approaches have overcome limitations of traditional methods through data-driven paradigms; however, inherent multi-objective optimization conflicts persist, rendering these methods prone to over-enhancement of exposed regions and detail distortion when handling complex scenarios involving simultaneous uneven illumination distribution and noise interference. To address these limitations, this paper proposes TSUNet (Two-Stage Wavelet Recovery U-Net), a dual-phase wavelet domain enhancement framework for low-light image restoration. This methodology employs a wavelet transform-enhanced U-Net architecture that decomposes the complex low-light enhancement problem into manageable sub-problems through sequential phases of preliminary restoration and refined enhancement. This hierarchical processing mechanism emulates the progressive perception characteristics of human visual systems: first establishing scene structural integrity before progressively resolving high-frequency details, thereby effectively mitigating multi-objective optimization conflicts inherent in single-phase models. This dual-phase architecture enables TSUNet to achieve superior equilibrium between global and local image quality improvements. To further bolster feature representation capabilities, the authors introduce an augmented wavelet domain feature fusion module integrating discrete wavelet transforms, inverse wavelet transforms, and dual attention mechanisms. Within this module, a novel dynamic gated spatial attention and lightweight fusion curve attention mechanisms synergistically interact with wavelet transforms, empowering TSUNet to process image features with enhanced spectral-spatial precision. Moreover, the researchers devise a perception-aware composite loss function combining Charbonnier loss, VGG perceptual loss, and a redesigned lightweight edge loss specifically optimized for low-light enhancement tasks. This comprehensive loss function guides the model to generate visually natural enhancements with authentic detail preservation by concurrently addressing pixel-level errors and perceptual fidelity. Experimental results demonstrate that the proposed method achieves state-of-the-art performance across multiple public low-light datasets in key metrics including Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM), while producing restored images with superior visual perception that effectively suppress noise artifacts while preserving critical structural details. The implementation is publicly available at <https://github.com/HibobacX/TSUNet>, facilitating further research and practical applications in low-light image enhancement.

Keywords image enhancement; U-Net; wavelet transformation; attention mechanism; loss function

1 引言

高质量的低照度图像增强已成为计算机视觉领域的研究前沿。图像增强技术作为预处理手段,对图像分类与目标检测^[1]等高级视觉任务具有重要价值,同时也是一种有效的图像恢复方法。

该领域仍面临诸多技术挑战,主要包括:噪声抑制、细节重构、色彩保真度下降、动态范围受限与局部对比度退化等核心问题。为应对这些挑战,早期

研究形成了两类主要技术范式,即基于直方图处理的方法和基于 Retinex 理论的方法。前者以直方图均衡化(Histogram Equalization, HE)和自适应直方图均衡化(Adaptive Histogram Equalization, AHE)为代表,通过调节图像亮度分布和对比度来提升低照度图像的视觉质量。后者则通过模拟人类视觉系统的色彩恒常性实现增强。在 Retinex 方法体系中,单尺度 Retinex (Single-Scale Retinex, SSR)采用高斯滤波器来估计照明分量,而多尺度 Retinex (Multi-Scale Retinex, MSR)通过多尺度高斯核优

化光照估计,在保持细节的同时提升亮度。

随着深度神经网络的快速发展,传统方法的局限性逐渐显现。直方图均衡化虽能通过灰度调整改善细节,但易导致过度增强与噪声放大。自适应直方图均衡化通过局部处理缓解了过度增强问题,却仍难以有效抑制噪声。基于 Retinex 理论的方法存在光晕效应、色彩失真和噪声敏感性问题,且计算复杂度较高。相比之下,深度神经网络凭借其强大的非线性表征能力,通过自适应多层次特征学习,已成为该领域的主流方法。

基于深度学习的增强方法可分为两个方向:传统算法的深度学习改进与端到端深度学习方法。在传统算法的深度学习改进方向上^[2-4],RetinexNet^[2]基于 Retinex 理论构建光照与反射联合估计模块,并通过控制曝光构建含合成噪声的 LOL 数据集^[2]以解决数据匮乏问题。端到端深度学习方法主要采用两种架构:单尺度流水线^[5-7]和与编解码器结构^[8-11]。前者如递归增强网络 PERIN^[6],通过递归结构降低参数量,但受限于表达能力难以恢复极低亮度图像的完整细节。后者以 LLNet^[8]为例,其编解码设计虽能平衡对比度与噪声,却存在高分辨率场景下的光照敏感性缺陷与泛化能力不足问题。

近年来,U-Net 及其变体在图像恢复领域表现突出:Yu 等人^[9]提出面向艺术图像的 U-Net 模型,Chao 等人设计了 CUI-Net^[10]处理不均匀光照,但在处理具有高噪声的低照度图像时仍存在挑战。相关研究表明^[12-15],U-Net 的层级化结构能自适应学习局部/全局特征以适应复杂光照,其跳跃连接实现多尺度特征融合。然而,低照度图像的高噪声特性易导致 U-Net 增强过程中噪声放大,影响视觉质量并增加后处理难度。因此,如何提升 U-Net 架构在噪声抑制方面的鲁棒性仍是亟待解决的问题。

除 U-Net 架构外,小波变换在低照度图像增强领域亦展现出独特优势。尽管该技术近年来已被引入低照度增强任务,但其应用仍存在固有局限性。例如,Kannoth 等人^[16]通过曲波变换与迭代投影技术结合实现了优异的细节保持,但面临噪声放大、阴影保留不足、马赛克效应及细节丢失等问题。Jiang 等人^[17]利用小波变换降低扩散过程的空间维度,但其性能在极暗场景下显著下降,主要归因于噪声主导效应。Wang 等人^[18]提出的小波变换注意力机制虽能有效抑制低频噪声,却因高频细节处理不足导致细节丢失。

进一步分析表明,小波变换存在以下核心缺陷:

首先,计算复杂度高:大尺寸图像处理时显著降低增强效率。其次,低频信息忽视:多数算法过度关注高频分量,导致低频对比度恢复不足。此外,噪声敏感性:高频增强过程易放大噪声干扰。最后,方向选择性缺失:难以精确识别连续边缘,导致边缘粗糙化与清晰度下降。这些缺陷要求在小波变换应用中需权衡其优缺点,并通过多技术融合优化增强效果。

值得注意的是,针对上述技术的不足之处,一些研究者开始探索 U-Net 与小波变换的融合应用。U-Net 凭借其强大的特征提取与重建能力,可有效弥补小波变换在细节重建方面的不足,而小波变换的多尺度分解能力为 U-Net 提供了更精细的特征表征。通过深度学习网络对小波子带进行联合学习与重构,二者协同作用可解决传统方法中的噪声放大与细节损失问题,最终实现低照度图像的精准全局增强。基于此,近期研究开始探索二者融合的技术路径:例如,Batziau 等人提出的 U-Net 与 Haar 小波池化融合技术^[19]在增强图像结构和纹理方面表现突出,但其采用的 Haar 小波池化在处理边缘细节时容易引入伪影。需特别指出的是,小波变换通过多尺度分解能够有效捕捉不同尺度的细节信息,并通过频域中有选择地增强特定频率分量,在保留边缘纹理细节的同时抑制噪声。然而,该方法存在两方面的局限性:其一,小波变换对不同类型低照度图像缺乏自适应性,在强噪声和不均匀光照场景下难以实现有效的全局增强;其二,U-Net 网络结构可能进一步放大此类复杂场景下的噪声干扰。

此外,由于小波变换本质上依赖频域特征分析,其在强噪声和不均匀光照场景下的全局增强能力受限,易导致局部区域增强失衡,影响视觉效果的自然性和真实性。现有传统 U-Net 与小波变换结合的方法虽综合了两者优势,但仍存在以下局限:(1)边界伪影问题:小波变换频域转换过程中引入的边界伪影经 U-Net 处理后显著增强,降低视觉质量;(2)噪声误识别:U-Net 易将高频分量中的残余噪声误判为有效特征并强化;(3)特征提取受限:基于固定基函数的小波变换难以适应复杂场景的特征多样性;(4)特征融合不一致:频域与空间域特征融合过程中存在语义不匹配现象。这些问题表明,现有融合方法在平衡细节增强与噪声抑制方面仍需优化,以实现更符合人眼感知的图像增强效果。

综上所述,尽管现有低照度图像增强技术已取得显著进展,但在噪声抑制与细节保留的平衡方面

仍面临挑战。针对上述问题,本文提出以下创新性解决方案:首先,本文设计了一种两阶段的 U 型低照度图像增强网络。通过将第一阶段增强结果与辅助特征图在精细增强阶段融合,该网络在保留图像全局结构的同时显著提升局部细节恢复能力,使低照度场景的增强结果呈现更自然的光照分布与更清晰的细节层次。该架构实现了多尺度特征的高效整合与深层语义提取。其次,本文提出了一种新颖的增强小波域特征融合模块。通过集成双重注意力机制与离散小波变换,该模块可精准聚焦小波域高频信息,提升模型对边缘和纹理的精细化处理能力,在避免传统小波变换过度增强问题的同时有效抑制细节丢失。最后,本文构建了一种融合感知损失函数。通过融合 Charbonnier 损失、VGG 感知损失与改进的轻量边缘损失,该函数联合优化像素级重建误差与高层语义特征匹配,引导模型生成兼具高保真度与强真实感的增强图像。通过上述技术的协同优化,本文方法解决了 U-Net 与小波变换融合中的边界伪影、噪声放大和跨域特征失配等问题。

本文的主要贡献总结如下:

(1)提出了一种两阶段低照度图像增强框架。该框架包含级联式的初步恢复与精细增强阶段。在初步恢复阶段,通过重新设计的浅层特征处理等模块共同作用,实现噪声抑制与基础亮度校正;在精细增强阶段,通过跨阶段特征融合机制实现更高质量的图像重建,显著提升图像的结构完整性与纹理真实性。

(2)设计了一种增强小波域特征融合模块。该模块利用离散小波在小波域上捕捉不同尺度的图像特征的能力,结合本文提出的动态门控空间注意力和轻量融合曲线注意力来强化关键特征表达,保留更多细节和纹理信息,从而实现更优异的图像增强效果。

(3)构建了一种融合感知损失函数。该损失函数由 VGG 感知损失、Charbonnier 损失以及改进的轻量边缘损失三部分组成。其中,针对低照度图像增强任务的特点,我们提出了一种轻量化边缘损失函数。该函数通过引入多通道处理机制和更高效的拉普拉斯核计算,能够高效提取图像边缘特征信息,使算法在降低计算复杂度的同时,显著提升了边缘细节的检测精度。

2 相关工作

2.1 低照度图像增强方法

在近期的研究中,Chen 等人针对夜间成像问题

创建了 SID 的数据集^[20]。该数据集通过提供不同曝光水平的图像对,为相关研究奠定了重要数据基础。Jiang 等人创新性地将注意力机制引入生成对抗网络^[21],有效验证了该机制在此任务中的适用性。Zhang 等人基于自监督学习改进了 RetinexNet 模型^[22],通过分离光照与反射分量实现了更高效的图像增强。Jie 等人提出的 LLFlow 网络^[23]创新性地融入了光流信息,但该方法的效果仍受限于光流估计的精度约束。Ma 等人开发的 SCI 模型^[24]采用级联照明学习框架,通过权重共享机制降低计算复杂度,在提升模型效率的同时增强了实际应用价值。

在图像自然性恢复方面,Li 等人提出的 ZeroDCE 网络^[25]通过像素级曲线估计和零参考损失函数实现了突破,后续研究团队进一步推出了加速版本^[26],在保持性能优势的同时显著提升了运算效率。Jiang 等人基于 MDARNet 的增强方法^[14]在色彩还原方面取得显著突破,而 Zhang 等人提出的 KinD 网络^[27]通过引入多约束损失函数,增强了 Retinex 架构的学习能力。值得关注的是,Cai 等人基于 Retinex 理论设计的新变体 Retinexformer^[3],作为首个基于 Transformer 的低照度增强算法,充分展现了该架构在视觉任务中的潜力。这些创新成果共同推动了低照度图像处理领域的技术发展。

2.2 U 型网络

U-Net 作为一种专为医学图像分割设计的卷积神经网络^[28],其架构灵感源于 Long 等人提出的全卷积网络框架^[29]。凭借强大的多尺度特征提取能力与高效的上下文信息整合机制,U-Net 已在计算机视觉领域得到广泛应用。

然而,当直接迁移 U-Net 至图像增强任务时,其网络结构易导致纹理细节与边缘信息的丢失。针对这一局限性,研究者提出了多种改进方案:Niu 等人通过任务解耦策略设计专用网络^[12],部分研究采用分阶段递归架构实现渐进式图像优化^[6,13,15,30-31],另有方法在编解码结构中引入多层次特征融合策略以增强细节保留^[32-33]。近年来,Shi 等人基于 UNet++ 算法^[34]提出的改进方案,通过嵌套跳跃连接结构与实例归一化残差模块的协同设计,显著提升了低照度增强性能。尽管这些方法在不同程度上改善了图像质量,但由于缺乏全局上下文建模能力,增强图像易出现噪声放大、阴影残留及细节丢失等问题,导致其在复杂光照场景下的泛化性能受限。

2.3 小波分析

小波变换通过将复杂信号分解为多尺度子带信

号,能够有效捕捉信号的多层次局部特征。针对低照度图像普遍存在的噪声显著、对比度低下等问题,其处理复杂性主要源于信号特征的耦合性。通过小波域分解获得的频率子带,可精准区分图像细节与噪声成分。基于子带系数调节的技术路线,在抑制噪声的同时能有效保留关键特征。然而,由于多尺度分解特性,其在处理大规模图像时面临计算资源需求高的问题,且模型泛化性能受制于训练数据的多样性与质量。更关键的是,小波变换过程中易引发噪声放大与阴影残留,进而导致细节特征退化。

针对上述瓶颈,研究者提出基于小波-CNN 的融合框架^[35],通过子带自适应机制提升低光场景适应性。进一步研究表明,离散小波变换与注意力驱动去噪网络的结合^[36],可在降噪过程中有效保持低照度细节。为增强结构特征与亮度分布,双分支网络与密集小波变换的协同架构^[37]展现出显著优势,但其高计算复杂度导致伪影与噪声的平衡控制依然

具有挑战性。为此,学界开始探索 U-Net 架构与小波变换的深度耦合^[38-40],利用编解码结构在浅层特征提取上的优势实现精细建模。值得关注的是,小波变换与扩散网络的创新性结合^[17,41],通过扩散模型的强生成能力产出高质量增强图像,但此类方法对训练数据质量敏感度显著增加,且计算开销远超传统小波方法。特别需要指出的是,扩散模型的引入虽能提升生成质量,但需严格控制其潜在的噪声干扰与输出稳定性。

3 方法实现

3.1 TSUNet 网络框架

针对低照度图像增强中存在的细节丢失、模糊及噪点问题,本文提出了一种名为 TSUNet (Two-Stage Wavelet Recovery U-Net) 的双阶段 U 型低照度图像增强网络,其整体结构如图 1 所示。

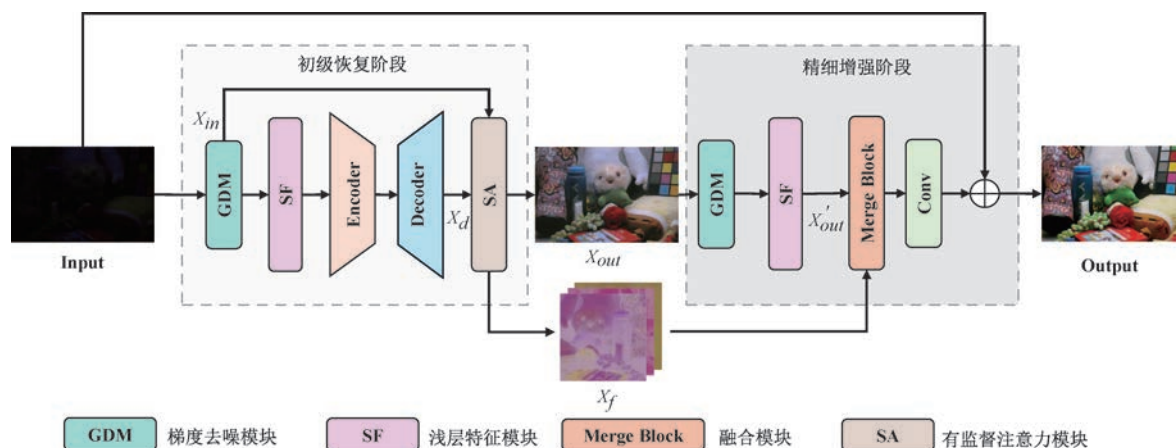


图 1 TSUNet 网络结构图

低照度图像增强网络的设计需要在增强效果和计算效率之间进行权衡。尽管增加网络阶段数可以拓展特征提取与处理能力,但也带来了一系列技术挑战。首先,网络参数量与计算复杂度呈近似线性增长,显著提高了硬件需求与处理时间。其次,随着网络深度增加,误差累积效应与梯度消失风险加剧,降低了训练的稳定性。此外,过多的处理阶段可能导致细节过度平滑或过度增强,从而引发整体亮度分布失衡,以及高光过曝和暗区欠曝等问题。

通过深入分析低照度图像增强任务,我们发现其核心挑战主要体现在两个相互关联但优化目标不同的方面:噪声抑制与细节增强。基于此,本研究创新性地提出了一种双阶段网络结构。其中,初级恢复阶段专注于噪声抑制和基础照度提升,而精细增

强阶段侧重于图像细节重建与整体质量优化。这种设计不仅确保了处理的完整性,同时有效避免了冗余阶段可能导致的性能损失。实验结果充分表明,该双阶段结构在控制计算复杂度的同时,确保了训练的稳定性,实现了增强效果与计算效率的最佳平衡。值得注意的是,进一步增加网络阶段数所带来的性能提升有限,反而会显著增加模型复杂度与训练难度。

TSUNet 主要由初级恢复阶段与精细增强阶段组成。在初级恢复阶段,首先利用梯度去噪模块 (Gradient Denoising Module, GDM) 和浅层特征模块 (Shallow Features, SF) 提取输入低照度图像的特征,并进行去噪与初步增强。随后,结合编解码器结构进一步处理图像特征,以完成初步增强和亮度

提升。最后,将编解码器处理后的特征图像 X_d 与提取并去噪后的特征 X_{in} 共同输入有监督的注意力模块(Supervised Attention Module, SA),以实现特征的有效分离。其中,经过分离处理后的特征一部分用于生成初步增强的特征图像 X_{out} ,并经由梯度去噪模块和浅层特征模块处理,得到特征图像 X'_{out} ;另一部分则用于生成辅助特征图 X_f 。

经过梯度去噪和特征提取的特征图像 X'_{out} 与辅助特征图 X_f 随后被输入第二阶段的融合模块(Merge Block),以提供细节恢复所需的特征信息。精细增强阶段主要用于提升图像细节,并进一步精细化重建第一阶段生成的特征。在该阶段,第一阶段恢复的辅助特征图 X_f 与初步增强的图像 X_{out} 进行融合,以进一步提炼和增强特征。最终,将精细增强阶段的输出与输入图像逐像素相加,以生成最终的增强图像。本文通过重新设计 U 型网络结构,并结合双阶段策略,实现了去噪、去模糊和细节增强的目标,从而有效提升了低照度图像的质量。

3.1.1 初级恢复阶段

本阶段的目标是实现初步的低照度增强,并为后续的精细增强阶段提供丰富的图像特征。该阶段主要由以下模块组成:梯度去噪模块(GDM)^[42]、浅层特征模块(SF)、编码器(Encoder)、解码器(Decoder)以及有监督的注意力模块(SA)。

(1)梯度去噪模块:为有效抑制低照度图像中的噪声与模糊,本文引入了梯度去噪模块(GDM)^[42],其结构如图 2 所示。该模块采用双残差分支处理输入图像,并通过比例参数调控原始图像,从而实现初步的去噪和去模糊。

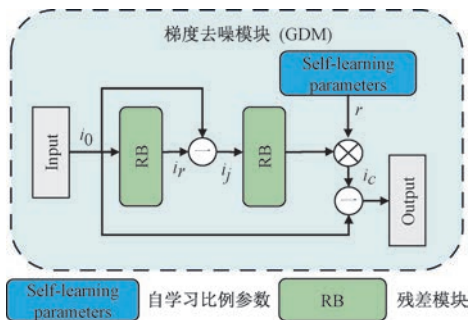


图 2 梯度去噪模块结构图

输入图像 i_0 先通过第一个残差分支(ResBlock, RB)进行特征计算,随后,将 i_0 减去 RB 处理后的特征 i_r ,获得模糊噪声纹理信息 i_j 。该信息随后经过第二个残差分支处理,并与自学习比例参数(Self-learning parameters)相乘,生成噪声纹理信息 i_c 。

通过迭代处理残差信息,可有效降低噪声纹理中的误差。自学习比例参数 r 作为步长调节因子,其作用类似于梯度下降中的学习率。最终,通过从输入图像 i_0 中减去噪声纹理信息 i_c ,实现降噪目标。

(2)浅层特征模块:为了有效捕获经过 GDM 模块去模糊与去噪处理后的低照度图像中的空间细节信息,本文设计了一个浅层特征模块(SF),其结构如图 3 所示。

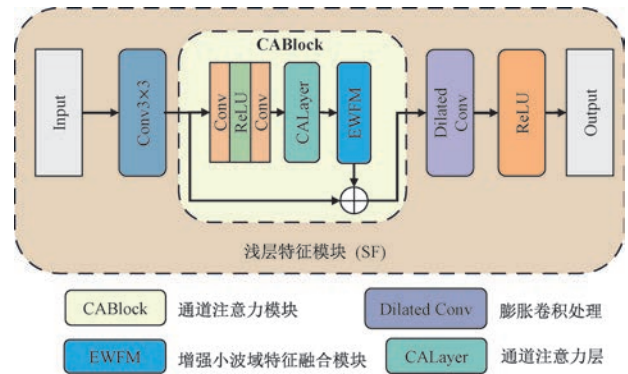


图 3 浅层特征模块结构图

SF 模块由卷积层($\text{Conv}3 \times 3$)、通道注意力块(Channel Attention Block, CABlock)和膨胀卷积块(Dilated Conv)^[43]三部分组成。

具体而言, SF 模块采用序列化结构,首先通过 3×3 卷积层($\text{Conv}3 \times 3$)提取基础特征,随后将特征输入 CABlock 进行局部特征优化。CABlock 的设计目标是通过以下三阶段处理生成高质量输入特征:

(1)通过双层卷积与 ReLU 激活函数实现特征精确提取;(2)引入通道注意力层(CALayer)筛选关键特征通道;(3)利用增强小波域特征融合模块(Enhanced Wavelet Domain Feature Fusion Module, EWFM)优化特征表示。

最终,处理后的特征图通过逐元素相加方式与原始输入特征图融合,生成的特征信息图将输入膨胀卷积块。通过膨胀卷积扩展感受野,该模块实现了对空间上下文信息的高效捕获。

SF 模块通过结合通道注意力机制与膨胀卷积,实现了特征提取的双重优化:前者提升网络对关键特征的敏感性,后者在保持计算效率的同时增强上下文感知能力。实验表明,该设计显著提升了低照度图像细节提取能力,进而改善整体增强效果。

(3)编解码器:作为图像增强的核心组件,编解

码器集成了增强小波域特征融合模块(EWFM)、混合实例标准化(Hybrid Instance Normalization,

HIN)以及残差结构(Residual Block, RB),其结构如图4所示。

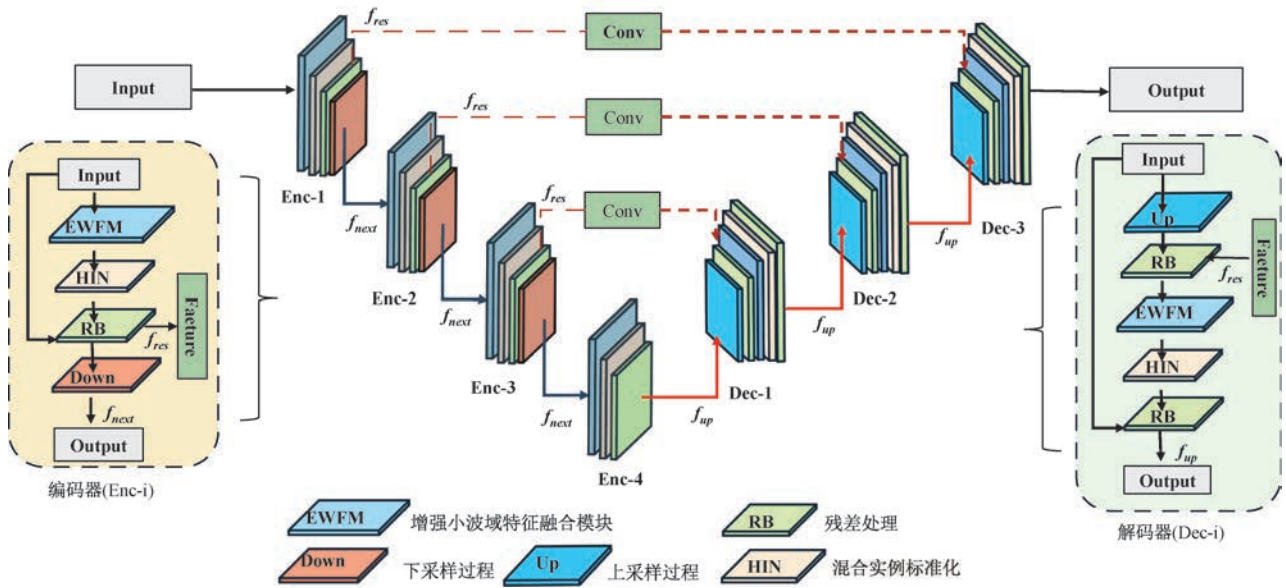


图4 编解码器结构图

编码器处理流程如下:EWFM 首先对输入图像执行特征去噪与优化,随后 HIN 模块提取多尺度特征信息,通过残差连接(RB)将处理后特征与原始输入特征进行融合。在下采样操作前,增强特征分两路处理:一路通过跳跃连接传输至解码器作为融合特征 f_{res} ;另一路经下采样后输出特征 f_{next} 作为下一阶段输入。

解码器的处理流程与编码器相似,其核心差异在于执行小波域特征融合前需先进行上采样(Up),并将上采样结果与对应编码层特征进行残差融合。

编解码器的协同运作最终生成融合多尺度信息的解码特征 f_{up} ,该特征将作为有监督注意力模块(SA)的输入。

(4) 有监督注意力模块:为深度挖掘增强图像信息,本文引入有监督注意力模块(SA)^[44],其结构如图5所示。

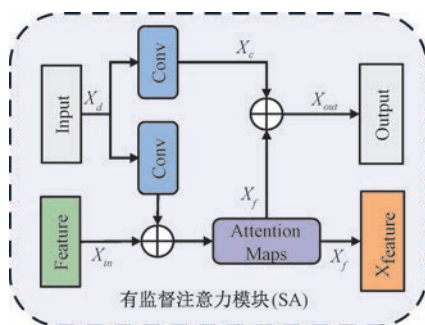


图5 有监督注意力模块结构图

该模块接收解码器输出特征 X_d 和经 GDM 去噪的输入特征 X_{in} ,通过双分支处理实现特征优化:

分支一:对 X_d 执行卷积操作后与 X_{in} 相加,生成空间注意力增强特征表示 X_f ;

分支二:对 X_d 执行卷积操作获得特征 X_c ,将其与 X_f 相加生成初步处理后的特征 X'_{out} 。两路输出 X_f 和 X_{out} 将被传入精细增强阶段进行后续处理。

3.1.2 精细增强阶段

精细增强阶段由梯度去噪模块(GDM)、浅层特征模块(SF)和融合模块(Merge Block)构成,其核心目标在于实现高质量图像重建。该阶段通过深度融合初级恢复阶段输出的特征图与图像细节纹理特征,最终输出经过增强的高品质图像。

在图像重建领域,多层次特征信息的有效融合始终面临严峻挑战,特别是在处理具有丰富细节纹理的图像时,传统融合方法往往存在信息冗余与特征遗漏的双重缺陷。为此,我们在精细增强阶段引入了一种新型融合模块,其结构如图6所示。该模块通过联合处理保留细节纹理的特征 X_f 与经 GDM 和 SF 二次优化的特征图 X'_{out} ,实现更精细的特征整合。

融合模块的处理流程包含四个关键步骤:

(1) 特征拼接:将输入特征 X'_{out} 与初级恢复阶段输出的细节纹理特征 X_f 沿特征维度进行拼接,经卷积处理后得到复合特征 X_{cat} ;

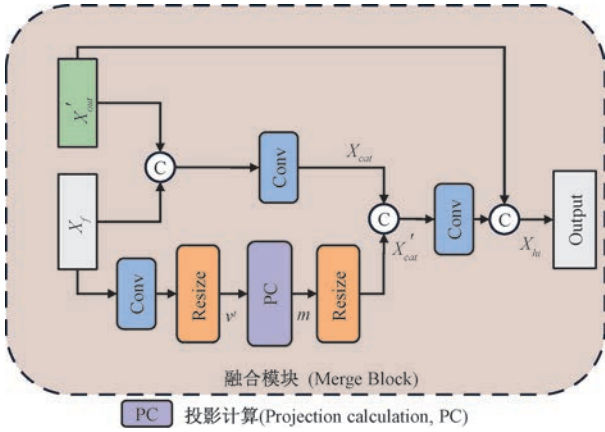


图6 融合模块结构图

(2)特征投影:通过投影计算^[45] (Projection calculation, PC),将经过变形(Resize)处理的细节特征 v' 压缩至低维空间,生成特征投影 m ;

(3)特征融合:将降维后的特征投影 m 与复合特征 X_{cat} 进行拼接,形成增强型融合特征 X'_{cat} ;

(4)最终整合:对融合特征 X'_{cat} 实施卷积处理(Conv),并与原始输入特征 X_{out} 进行最终拼接,输出重建后的高分辨率图像 X_{hi} ,在保障细节保真的同时显著提升图像质量。

3.2 增强小波域特征融合模块

深度神经网络凭借其强大的非线性拟合能力,在计算机视觉领域取得了突破性进展。理论研究表明,三层及以上的神经网络即可逼近任意非线性函数。然而,在低照度图像增强任务中,单纯依赖端到端的神经网络学习存在以下固有局限:

数据依赖性:需要大规模训练数据和计算资源,且训练收敛性难以保证。

可解释性缺失:网络的“黑盒”特性导致处理结果缺乏透明性,难以针对性优化。

过拟合风险:易对噪声特征产生过拟合,导致增强图像出现伪影或细节失真。

针对传统方法的固有缺陷,本文引入小波变换理论以提升低照度图像增强效果。需要指出的是,传统小波变换方法仍存在以下固有不足:第一,参数固化:预设变换基函数缺乏自适应性,需人工调整参数应对不同图像退化类型;第二,失真风险:简单阈值处理易引发马赛克效应和边缘模糊;第三,信息利用不足:复杂低照度场景下难以有效利用图像上下文信息。

为克服传统小波变换方法的局限性,本文提出一种新型增强小波域特征融合模块(EWFM),其核心创新在于将小波变换与深度学习注意力机制深度融合。该模块的技术方案包含:(1)引入双重注意力机制,实现特征权重的自适应学习和调整;(2)采用深度可分离卷积(Depthwise Separable Convolution, DSC)对高频分量进行编码,强化复杂纹理表征能力;(3)运用特征分离与融合技术,保持原始图像低频信息,避免过度增强导致的噪声放大与细节丢失。

在具体实现中,EWFM 模块通过离散小波变换(Discrete Wavelet Transform, DWT)与逆小波变换(Inverse Wavelet Transform, IWT)的协同处理,结合双重注意力机制和特征融合技术实现细节增强与噪声抑制。通过双重注意力机制对小波域高频信息的精细化处理,该模块显著提升了网络的去噪能力、边缘锐化精度及纹理复原效果。该模块结构如图7所示。

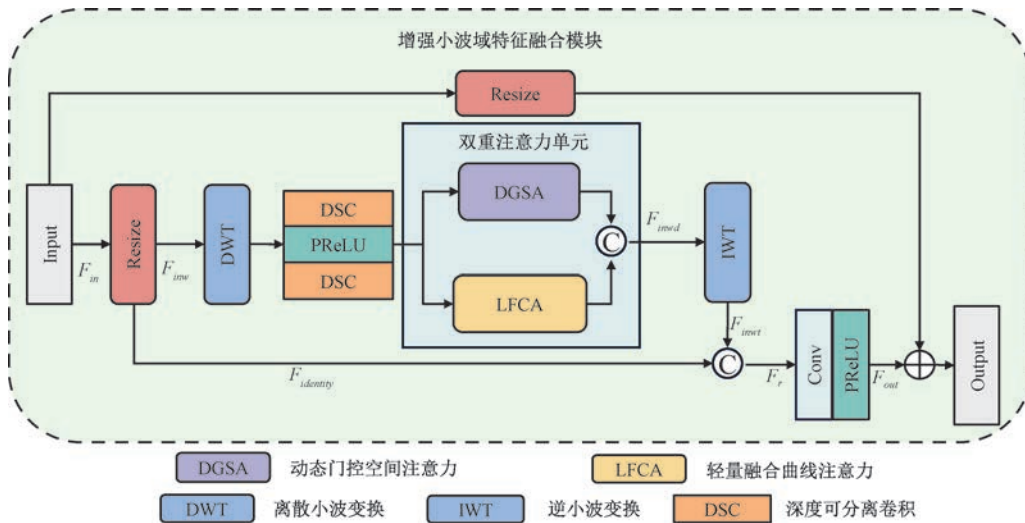


图7 增强小波域特征融合模块结构图

EWFM 模块的结构流程如下:

(1) 特征预处理: 将输入特征 F_{in} 进行维度变换后分为两部分: 一部分是保留低频信息的分离融合特征 $F_{identity}$, 用于 DWT 处理的待分解特征 F_{inw} ;

(2) 小波变换处理: 对 F_{inw} 进行 DWT 分解, 获取高频分量 (包含噪声、纹理信息) 与低频分量, 采用深度可分离卷积^[46] (DSC) 和参数化修正线性单元^[47] (Parametric Rectified Linear Unit, PReLU) 对高频分量编码, 经双重注意力单元处理后获得深度特征 F_{inwd} ;

(3) 特征融合重建: 通过 IWT 将特征信息 F_{inwd} 重建为原始的空间分辨率的重建特征图 F_{inwt} , 与 $F_{identity}$ 拼接生成拼接特征 F_r ; 经卷积层 ($\text{Conv}1 \times 1$) 和 PReLU 激活函数处理产生残差特征 F_{out} , 最终通过将 F_{out} 与原始输入残差连接输出增强特征。

为优化高频信息处理, 本文提出包含动态门控空间注意力 (Dynamic Gated Spatial Attention, DGSA) 和轻量融合曲线注意力 (Lightweight Fusion Curve Attention, LFCA) 的双重注意力机制, 其协同优势体现在:

(1) 动态范围扩展: 在不同维度增强模型处理能力, 提升暗区细节的同时抑制亮区过曝;

(2) 局部特征调控: 实现细节与纹理的高质量恢

复, 显著提升图像整体质量。

动态门控空间注意力 (DGSA): 通过自适应权重调节聚焦关键空间区域, 采用门控结构控制信息流向, 根据图像内容动态调整注意力分布, 实现对空间细节的精准处理。

轻量融合曲线注意力 (LFCA): 通过像素级特征调节优化细节表现与对比度, 结合轻量化网络与曲线融合策略, 在保证计算效率的同时提升色彩恢复效果。

DGSA 与 LFCA 协同作用显著增强了小波域特征融合模块的图像恢复性能, 具体实现细节将在后续章节详细阐述。

3.2.1 动态门控空间注意力

在现有低照度图像增强研究中, 注意力机制通常聚焦于图像显著区域以增强特征表达能力。然而, 在处理具有复杂光照和纹理的图像时, 传统方法往往难以全面捕捉关键细节。因此, 本文提出一种集成门控卷积的新型注意力模块: 动态门控空间注意力模块 (DGSA), 其结构如图 8 所示。该模块通过引入门控调制机制能够增强注意力的表达能力, 实现注意力分配的精细化控制, 从而显著提升模型对低照度图像细节的恢复性能。

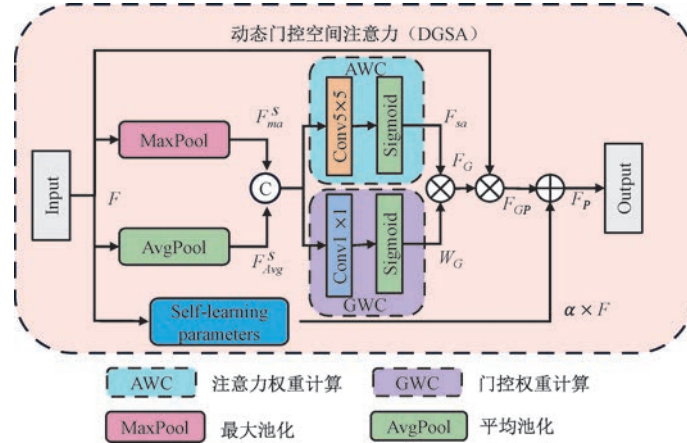


图 8 动态门控空间注意力结构图

DGSA 模块由空间注意力计算与门控调制两个核心单元构成, 具体实现流程如下:

(1) 空间特征提取: 采用最大池化 (Maximum Pooling, MaxPool) 和平均池化 (Average Pooling, AvgPool) 并行提取空间特征。其中, MaxPool 能够捕获图像的极端特征, 而 AvgPool 能够获取图像平均响应。双路径特征互补融合, 为模型提供全局内容理解与局部细节感知能力。

(2) 空间注意力计算 (AWC): 将 MaxPool 和

AvgPool 两种池化结果 (F_{Avg}^s, F_{ma}^s) 沿通道维度拼接。经注意力权重计算 (Attention Weight Calculation, AWC) 处理, 使用 5×5 卷积层和 Sigmoid 激活函数生成初始空间注意力图 F_{sa} 。具体计算过程如公式 (1) 所示:

$$F_{sa} = \sigma(\text{Conv}_5(\text{Cat}([F_{Avg}^s; F_{ma}^s]))) \quad (1)$$

其中, σ 表示 Sigmoid 激活函数, F_{Avg}^s 表示经过平均池化操作得到的特征图, F_{ma}^s 表示经过最大池化操作得到的特征图, Cat 表示特征拼接操作, Conv_5

表示 5×5 的卷积操作。

(3)门控权重生成(GWC):门控卷积层用于门控权重计算(Gate Weight Calculation, GWC)。GWC采用 1×1 卷积核来调节特征响应,通过 Sigmoid 函数处理生成门控权重 W_G 。具体计算过程如公式(2)所示:

$$W_G = \sigma(\text{Conv}_1(\text{Cat}([F_{\text{avg}}^s; F_{\text{ma}}^s]))) \quad (2)$$

其中, Conv_1 表示 1×1 的卷积操作。

(4)注意力权重生成:将门控权重 W_G 与空间注意力图 F_{sa} 逐元素相乘,生成精细化注意力分布 F_G 。具体计算过程如公式(3)所示:

$$F_G = W_G \times F_{sa} \quad (3)$$

(5)特征增强处理:将调制后的注意力权重 F_G 与输入特征 F 进行相乘,获得增强特征图 F_{GP} ,该过程如公式(4)所示:

$$F_{GP} = F_G \times F \quad (4)$$

通过门控调制机制,模型能够自适应地调节空间位置的权重系数,精准增强图像特征,从而显著提升对图像细节的处理能力。

(6)特征融合计算:为了增强恢复图像细节及视觉效果,本文设计了一个带有自学习参数(Self-learning parameters)的残差连接机制。将增强特征图 F_{GP} 与原始输入 F 加权融合,动态门控机制通过可学习参数 α 自动调节融合强度,生成处理后的特征图 F_P 。动态门控空间注意力如式(5)所示:

$$F_P = F_{GP} + \alpha \times F \quad (5)$$

DGSA 不仅能够保留原始图像信息,还避免了因过度调制而导致的信息损失,确保模型在不同光照条件下的图像处理效果。

3.2.2 轻量融合曲线注意力

动态门控空间注意力(DGSA)机制通过增强模型对全局结构的感知能力,有效缓解了局部光照不足导致的细节丢失问题,同时显著抑制了噪声背景对图像质量的干扰。然而,仅使用 DGSA 难以有效理解局部照明。因此,本文提出一种结合曲线估计方法的轻量融合曲线注意力机制(LFCA),旨在提升关键特征提取能力并实现高效噪声抑制。DGSA 模块结构如图 9 所示。

LFCA 模块的核心架构包含三个关键组件:曲线参数计算(Curve Parameter Calculation, CPC)、梯度放大模块(Gradient Amplifier, GA)和多尺度特征融合(Multi Scale Feature Fusion, MSF)。该机制通过协同优化多层次特征提取与融合,实现图像细节恢复、色彩校正和对比度增强,同时保持计算

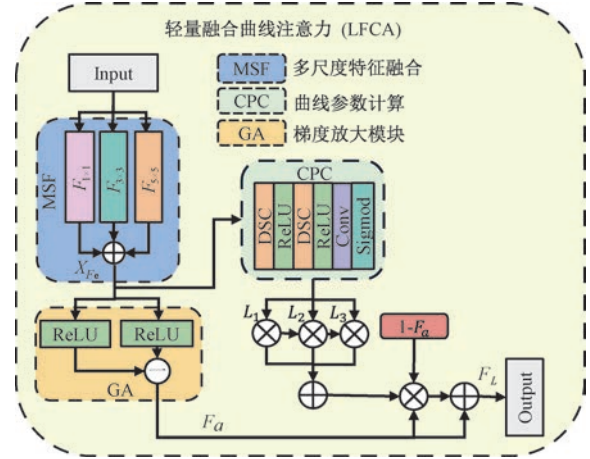


图9 轻量融合曲线注意力结构图

效率。具体实现流程如下:

(1)多尺度特征融合(MSF):采用不同尺寸卷积核($F_{s \times s}$ ($s = 1, 3, 5$))从输入图像中提取多尺度特征,多尺度特征通过分支内局部通道压缩与相加融合生成融合特征 X_{Fe} 。该特征将并行输入至曲线参数计算(CPC)和梯度放大模块(GA)进行后续处理,其计算过程如公式(6)所示:

$$X_{Fe} = \sum_{s \in \{1, 3, 5\}} F_s(x) \quad (6)$$

其中, s 表示多尺度特征融合中 $s \times s$ 大小的卷积核; F_s 表示使用 $s \times s$ 大小的卷积核提取的特征; x 表示输入的特征图像。

(2)曲线参数计算(CPC):为了优化计算, CPC 模块引入了深度这种方法将标准卷积分解为深度卷积和逐点卷积,从而在降低模型参数和计算复杂度的同时,确保了特征提取能力。融合特征 X_{Fe} 通过 CPC 生成逐像素的曲线参数 L_i ($i = 1, 2, 3$)。其表达式如公式(7)所示:

$$L_i = \sigma(\text{Conv}_1(\theta(\text{DSC}(\theta(\text{DSC}(X_{Fe})))))) \quad (7)$$

其中, DSC 表示深度可分离卷积; θ 表示 ReLU 激活函数; σ 表示 Sigmoid 激活函数。该设计增强了轻量融合曲线注意力在低照度图像增强任务中的效率,更适用于实时或资源受限的应用场景。

(3)梯度放大模块(GA):采用双路 ReLU 激活策略对融合特征信息 X_{Fe} 进行特征选择:第一路保持原始特征传递,第二路执行零值截断,通过计算两路输出的差值生成窗口函数形式的特征图 F_a 。在保留关键特征信息,抑制冗余特征。具体计算过程如公式(8)所示:

$$F_a = \theta(X_{Fe}) - \theta(X_{Fe} - 1) \quad (8)$$

该模块通过在 $[0, 1]$ 区间的非线性窗口函数来

突出关键特征。更具体地说,GA 使用双路 ReLU 激活函数处理融合特征,以确保特征值被限制在 0 到 1 之间,超出该范围的值则被截断。随后,将这个窗口函数与自适应的曲线参数结合,动态优化特征权重的分配,最终增强重要区域的特征表达并抑制非关键区域。相较于固定的注意力权重机制,该方法具有更强的自适应性和灵活性,能够根据不同场景动态调整特征权重,从而提高模型性能。

(4) 整合处理与特征调整:将梯度放大模块输出的特征 F_a 和曲线参数 L_i 通过非线性变换 $(1-F_a)$ 进行融合:首先对融合的特征 F_a 经过非线性变换,随后与各曲线参数逐元素相乘并累加,最终将优化结果与融合特征 F_a 相加得到增强特征 F_L 。轻量融合曲线注意力如式(9)所示:

$$F_L = F_a + \sum_{i=1}^3 L_i \times F_a \times (1 - F_a) \quad (9)$$

其中, i 表示第 i 个像素集曲线参数。

在上述公式中,非线性项 $F_a \times (1 - F_a)$ 确保了当特征值 F_a 接近 1 时抑制增长,而接近 0 时促进增长。通过这种自适应的动态平衡机制可精准增强关键区域特征表达,显著提升模型在低照度场景下的特征判别能力。

3.3 融合感知损失函数

传统的低照度图像增强方法通常使用 L_1 损失函数作为训练的损失函数,然而,由于 L_1 损失函数在度量图像结构信息和高层语义相似度方面存在局限性,会导致生成的图像缺乏真实感,制约了图像质量提升。针对此局限性,本文提出一种融合感知损失函数,该损失函数旨在利用 Charbonnier 损失函数优化局部像素级重建精度,利用 VGG 感知损失函数提取并匹配高层语义特征。该损失函数兼具 Charbonnier 损失函数^[48]和 VGG 感知损失函数的优点,综合考虑了全局特征和局部细节,同时辅以本文改进的边缘损失函数,实现了图像细节信息和整体视觉感知之间的平衡。

3.3.1 Charbonnier 损失

Charbonnier 损失^[48]作为 L_1 损失函数的改进形式,具有连续光滑特性。Charbonnier 损失被用于计算生成图像像素矩阵 \mathbf{X} 与真实图像像素矩阵 \mathbf{Y} 的单像素相似性,相较于标准 L_1 损失,其在单像素误差 $\mathbf{X}-\mathbf{Y}$ 等于 0 处一阶连续可导,避免了 L_1 损失在零点不可导的问题,从而确保优化过程的数值稳定性。通过调节参数 ϵ ,该损失函数在误差接近 0 时具有类似 L_2 损失的平滑性,能够提供更精细的梯

度信息,而在误差较大时表现出类似 L_1 损失的鲁棒性。其公式如公式(10)所示:

$$L_{Char} = \sqrt{\|\mathbf{X}-\mathbf{Y}\|^2 + \epsilon^2} \quad (10)$$

其中, ϵ 是一个常数,用于保证数值稳定性。该损失函数通常在总损失函数中占主导地位。

3.3.2 VGG 感知损失

为提升重建图像的视觉真实性,本文引入 VGG 感知损失函数^[49],其通过预训练 VGG 网络提取多层级特征,模拟人类视觉系统的特征感知机制。通过匹配中间层特征,提升图像纹理细节保真度与全局结构一致性。VGG 感知损失的数学定义为

$$L_{VGG} = \frac{1}{C_j H_j W_j} \|\phi_j(\hat{\mathbf{y}}) - \phi_j(\mathbf{y})\|_2^2 \quad (11)$$

其中, j 表示网络的第 j 层, $\phi(\cdot)$ 表示 VGG 网络第 j 层的特征映射函数, $C_j H_j W_j$ 表示第 j 层特征图的通道数、高度和宽度, $\hat{\mathbf{y}}$ 表示生成图像像素矩阵, \mathbf{y} 表示真实图像像素矩阵。

3.3.3 轻量边缘损失

在低照度图像增强任务中,细节恢复是影响增强效果的关键因素,特别是对高频边缘特征和纹理结构的精确重建。尽管当前的低照度增强方法逐步提升了整体亮度恢复能力,但在边缘细节保真度方面仍存在显著局限。本文从图像超分辨率领域的边缘损失^[50]获得启发:在图像超分辨率任务中,边缘损失函数借助高斯核来提取图像边缘信息,通过比较原始图像和增强图像的边缘差异来计算损失。然而,这一过程涉及复杂的下采样和上采样操作,不仅带来了较高的计算复杂度,还会因多通道信息交叉混合影响边缘检测的准确性。

针对上述技术瓶颈,本研究提出了一种面向低照度图像增强的轻量化边缘损失函数。该创新性设计摒弃了传统方法依赖高斯模糊的间接特征提取范式,通过端到端的边缘特征提取机制,在降低计算复杂度的同时显著提升了边缘检测精度。具体而言,本方案采用了两种关键技术改进:

(1) 采用标准 3×3 拉普拉斯卷积核对图像进行直接特征提取,省却了传统流程中的高斯预处理和下采样操作。

(2) 设计多通道独立处理机制,允许 RGB 各颜色通道独立执行边缘检测,有效避免跨通道信息耦合导致的特征混淆。该损失函数的具体公式如下:

$$L_{edge} = \frac{1}{N} \sum_{i=1}^N \sqrt{(\Delta_{LO}(\mathbf{X}) - \Delta_{LO}(\mathbf{Y}))^2 + \epsilon^2} \quad (12)$$

其中, \mathbf{X} 和 \mathbf{Y} 分别代表生成图像像素矩阵和真实图

像像素矩阵, ϵ 为调整参数, Δ_{LO} 表示拉普拉斯核处理, N 表示像素数量乘以通道数得到的像素总数。

3.3.4 融合感知损失

为实现主观视觉质量与客观指标的综合优化, 本文构建了融合多维度约束的复合感知损失函数。该函数创新性地整合了 VGG 感知损失、Charbonnier 平滑损失以及新型轻量边缘损失, 通过联合优化局部细节重建质量与全局视觉一致性, 有效提升了模型对低照度场景的适应能力。融合感知损失定义为

$$L_{total} = \alpha_{VGG} L_{VGG} + \alpha_{Char} L_{Char} + \alpha_{Edge} L_{Edge} \quad (13)$$

其中, α_{VGG} 、 α_{Char} 和 α_{Edge} 为权重系数。通过对不同权重系数的对比实验和消融研究, 发现当权重系数设置为 $\alpha_{VGG} = 1.3$ 、 $\alpha_{Char} = 1.7$ 、 $\alpha_{Edge} = 0.05$ 时, 模型在保留图像结构信息和抑制噪声方面能够达到最佳平衡。这种权重配置使模型在保留图像结构信息的同时, 有效抑制了噪声产生, 避免了细节丢失和噪声过度增强等问题。

4 实验结果及分析

4.1 实验细节

本文使用 NVIDIA GeForce RTX 4090 GPU, 基于 PyTorch 深度学习框架, 在 LOLv1 基准数据集上进行了 300 轮次的训练。在训练过程中, 将低照度图像随机裁剪为 256×256 大小的图像块, 作为模型的输入, 批量大小设置为 2。实验中采用 Adam 优化器, 初始学习率设置为 $1e-4$, 随后通过余弦退火算法将其衰减到 $1e-6$ 。在测试阶段, 使用 256×256 大小的测试集图像进行评估。

4.2 数据集

本文分别在真实数据集 LOLv1^[2]、LSRW^[51]、LOLv2-real^[52]、LOLv2-synthetic^[52] 和无参考数据集 DICM^[53]、LIME^[54]、MEF^[55] 上进行测试, 以验证本文方法的有效性。其中, LOLv1 包含 500 对图像, 本文将其中 485 对图像作为训练图像, 剩余的 15 对图像作为测试图像。LSRW 数据集包含两个子数据集: 一个是使用华为手机拍摄的 2480 对图像, 其中 2450 对用于训练, 30 对用于测试; 另一个是使用尼康相机拍摄的 3170 对图像, 其中 3140 对用于训练, 30 对用于测试。LOLv2 数据集包括两个子集, 即 LOLv2-real 和 LOLv2-synthetic。LOLv2-real 子集是在真实场景中通过调整 ISO 和曝光时间捕获得到的, 本文使用其中 689 对图像用

于训练, 100 对用于测试。LOLv2-synthetic 子集则通过分析低光照图像的照明分布, 从 RAW 图像合成低光照图像。该子集包含 1000 对低光照/正常图像, 其中 900 对用于训练, 100 对用于测试。无监督数据集 DICM、LIME 和 MEF 分别包含 69 张、10 张和 44 张商用数码相机采集的低照度图像, 均用于测试。

4.3 方法比较

4.3.1 客观评价

将本文提出的方法与当前先进方法进行比较, 包括无监督学习方法: EnlightenGAN^[21] 和 SCI^[24]; 零样本学习方法: ZeroDCE^[26]; 监督学习方法: IAT^[56]、LLFormer^[23]、CF-Uformer^[11]、R2RNet^[51]、MIRNet^[57]、Instruct-IR^[58]、PPFormer^[59] 和 Entropy-SDE^[60]; 基于小波变换的方法: CFWD^[41] 和 HWMNet^[61]。此外, 实验中采用峰值信噪比 (Peak Signal to Noise Ratio, PSNR)、结构相似性 (Structural Similarity Index Measure, SSIM)^[62] 以及学习感知图像相似度 (Learned Perceptual Image Patch Similarity, LPIPS)^[63] 作为客观评估指标。

表 1 展示了本文方法与现有先进方法在 LOLv1 和 LSRW 数据集上的对比结果, 最优和次优结果分别用加粗和下划线表示 (后面表格同此)。实验结果表明, 本文提出的方法在 PSNR、SSIM、LPIPS 三个指标上均优于所有对比方法。其中, 在 LSRW 数据集中, SSIM 相比于次优方法提升较大, 这一优势源于: (1) 小波域高频分量的精细化处理: 增强小波域特征融合模块 (EWMF) 通过动态门控空间注意力 (DGSA) 强化边缘梯度信息, DGSA 更关注图像的边缘和梯度等结构信息, 并精准地增强这些关键信息。这种精细化处理直接优化了 SSIM 指标, 有效提升了图像的结构相似度; (2) 噪声抑制与细节解耦: 双阶段设计中的梯度去噪模块 (GDM) 有效抑制了噪声对梯度信息的影响, 避免噪声干扰网络学习图像的真实结构。这种设计让网络更专注于学习和恢复图像的本质结构, 这对提升结构信息敏感的 SSIM 指标至关重要。

为了进一步验证本文所提方法的性能优势具有统计学意义, 我们对实验结果进行了显著性检验。在 LOLv1 和 LSRW 数据集的所有评价指标上, 为确保 t 检验方法的适用性, 我们首先验证了数据满足独立性和正态性假设 (通过 Shapiro-Wilk 检验, $p > 0.05$), 且方差齐性检验 (Levene 检验) 显示组间方差无显著差异 ($p > 0.05$)。在满足参数检验前提

表 1 在 LOLv1 数据集和 LSRW 数据集上的结果比较

方法	LOLv1			LSRW		
	PSNR \uparrow	SSIM \uparrow	LPIPS \downarrow	PSNR \uparrow	SSIM \uparrow	LPIPS \downarrow
MIRNet 2020 ^[57]	24.140	0.842	0.131	16.470	0.477	0.430
EnlightenGAN 2021 ^[21]	14.311	0.575	0.315	15.932	0.482	0.315
ZeroDCE 2021 ^[26]	14.861	0.562	0.256	16.105	0.467	0.291
SCI 2022 ^[24]	13.806	0.528	0.279	15.087	0.413	0.288
IAT 2022 ^[56]	23.380	0.809	0.216	16.211	0.492	0.271
HWMNet 2022 ^[61]	22.920	0.830	0.114	19.221	<u>0.592</u>	0.208
LLFormer 2022 ^[23]	22.529	0.792	0.173	19.077	0.542	0.305
R2RNet 2023 ^[51]	18.132	0.739	0.207	17.256	0.541	0.306
CFWD 2024 ^[41]	22.730	<u>0.853</u>	0.087	<u>19.566</u>	0.572	0.374
CF-UFormer 2024 ^[11]	23.820	0.850	<u>0.079</u>	16.217	0.430	0.215
PPFormer 2024 ^[59]	23.910	0.845	0.217	19.281	0.552	0.350
Instruct-IR 2024 ^[58]	22.830	0.836	0.116	19.043	0.532	0.311
Entropy-SDE 2024 ^[60]	24.050	0.848	0.081	19.241	0.548	<u>0.195</u>
本文方法	24.336	0.861	0.056	21.175	0.642	0.158
p -value	4.40e-4	1.35e-4	5.51e-5	2.31e-6	1.33e-6	1.05e-8

条件的基础上,采用配对 t 检验将四次训练得到的指标与基准指标进行对比分析。统计结果显示,我们将四次训练得到的指标和基准指标进行了统计显著性检验,在所有指标上均满足 p -value <0.05 。本文方法均具有统计显著性,表明本文方法的性能提

升显著且非随机。

同时,为展现本文模型的鲁棒性,在 LOLv2-real 和 LOLv2-synthetic 数据集上也进行了对比实验,实验结果如表 2 所示。

表 2 在 LOLv2-real 数据集和 LOLv2-synthetic 数据集上的结果比较

方法	LOLv2-real			LOLv2-synthetic		
	PSNR \uparrow	SSIM \uparrow	LPIPS \downarrow	PSNR \uparrow	SSIM \uparrow	LPIPS \downarrow
MIRNet 2020 ^[57]	20.020	0.820	0.233	21.940	0.846	0.249
EnlightenGAN 2021 ^[21]	18.640	0.677	0.309	16.570	0.770	0.212
ZeroDCE 2021 ^[26]	16.059	0.580	0.352	17.577	0.815	0.187
SCI 2022 ^[24]	17.304	0.540	0.345	16.695	0.723	0.242
IAT 2022 ^[56]	20.310	0.789	0.220	15.956	0.743	0.259
HWMNet 2022 ^[61]	22.831	0.827	0.193	15.150	0.706	0.268
LLFormer 2022 ^[23]	20.056	0.792	0.260	24.038	0.909	0.067
R2RNet 2023 ^[51]	20.267	0.803	0.395	16.052	0.736	0.409
CFWD 2024 ^[41]	23.321	0.825	0.193	23.872	0.903	0.078
CF-UFormer 2024 ^[11]	22.273	0.837	<u>0.092</u>	22.231	0.892	0.107
PPFormer 2024 ^[59]	<u>23.721</u>	0.830	0.187	<u>24.807</u>	<u>0.919</u>	<u>0.053</u>
Instruct-IR 2024 ^[58]	22.912	<u>0.841</u>	0.271	23.310	0.891	0.116
Entropy-SDE 2024 ^[60]	21.310	0.832	0.120	22.145	0.872	0.173
本文方法	24.604	0.877	0.053	25.296	0.943	0.029
p -value	3.61e-4	8.19e-5	1.04e-6	3.26e-5	1.75e-6	2.52e-5

从表 2 可以看出,在 PSNR、SSIM、LPIPS 三个指标上,本文提出的方法较于其他方法,均取得了最佳性能表现。LOLv2 数据集上的实验结果进一步验证了本文提出方法的鲁棒性,无论是在真实场景的 LOLv2-real 数据集上还是合成场景的 LOLv2-synthetic 数据集上,本文模型在 PSNR、SSIM、LPIPS 三个评价指标中均表现最优,这表明本文提出的模型能够适应不同的环境和条件,在真实和合成场景下均能表现出稳定的性能。同样,在表 2 中

对 LOLv2-real 和 LOLv2-synthetic 数据集的实验结果进行了统计显著性检验,为确保 t 检验方法的适用性,我们首先验证了数据满足独立性和正态性假设(通过 Shapiro-Wilk 检验, $p>0.05$),且方差齐性检验(Levene 检验)显示组间方差无显著差异($p>0.05$)。在满足参数检验前提条件的基础上,采用配对 t 检验将四次训练得到的指标与基准指标进行对比分析。统计结果显示,我们将四次训练得到的指标和基准指标进行了统计显著性检验,在所

有指标上均满足 $p\text{-value} < 0.05$, 再次表明本文方法的性能提升具有统计学显著性且非随机。

为了深入理解本文方法在低照度图像增强方面的优越性能, 我们对实验结果进行了详细分析。首先, 得益于本文设计的增强小波域特征融合模块和两阶段网络中的初步恢复阶段, 本文方法在 LOLv1、LSRW、LOLv2-real 和 LOL-synthetic 数据集中的 PSNR 和 SSIM 两个失真评价指标上均取得了领先。这表明, 采用本文提出的方法增强的图像具有更多的高频细节和结构完整性, 为图像恢复结果提供了更高质量的输出。

此外, LPIPS 是一种用于衡量图像感知质量的指标。通过观察对比实验结果, 发现本文方法在 LPIPS 指标上表现显著优于现有技术, 这种提升得益于本文模型中融合感知损失函数和两阶段网络中精细恢复阶段的设计。

最后, 为进一步评估本文模型的有效性与泛化能力, 对本文和主流方法使用无参考数据集 DICM、LIME 和 MEF 进行了对比实验。由于这三个数据集无参考图像, 所以选用无参考评价指标 (Natural Image Quality Evaluator, NIQE) 来客观评价模型的恢复效果, 实验结果如表 3 所示。

表 3 在 DICM、LIME、MEF 数据集上的 NIQE 分数对比

方法	DICM	LIME	MEF	Avg
原始输入	7.623	7.624	7.146	7.464
EnlightenGAN ^[21]	7.297	<u>6.914</u>	6.738	6.983
ZeroDCE ^[26]	7.612	7.356	6.920	7.296
SCI ^[24]	6.678	7.341	6.557	6.859
HWMNet ^[61]	7.230	7.011	6.605	6.949
R2RNet ^[51]	<u>6.771</u>	7.797	6.757	7.108
CFWD ^[41]	6.963	6.982	<u>6.535</u>	<u>6.827</u>
本文方法	7.070	6.899	6.279	6.750

实验结果表明, 本文方法在 LIME 和 MEF 数据集中取得了最优结果, 且在三个数据集的平均表

现上优于其他对比方法, 这显示了该方法在低照度图像增强任务中的有效性和泛化能力。

以上实验结果表明, 本文提出的方法不仅能够显著增强图像质量, 还能在恢复图像的结构和感知质量方面表现出色, 并在不同场景和条件下表现出良好的鲁棒性。本文方法在算法设计和模型优化上的创新是实现这些优势的关键, 使其能够在不同的数据集和场景下均保持稳定的性能。

表 4 给出了本文方法和具有代表性的低照度增强方法以及基于小波变换的低照度增强方法在参数数量和计算复杂度方面的对比结果。结果显示, 尽管本文方法在参数数量和计算复杂度上与其他方法相近, 但在性能表现上更为优越。

表 4 低照度增强方法的复杂度和参数对比

方法	参数量 (M)	Flops (G)
MIRNet ^[57]	31.79	785.00
HWMNet ^[61]	29.59	86.50
CFWD ^[41]	24.67	87.72
本文方法	24.29	84.07

4.3.2 主观评价

如图 10 所示, 本研究对各方法的体育场图像增强效果进行了可视化对比分析, 实验结果表明: ZeroDCE 和 SCI 方法虽能提升全局亮度, 但在恢复座椅等关键细节方面表现欠佳, 不仅出现了严重的细节丢失, 还伴有明显的噪声和色彩偏差。这些问题无疑会对后续的高级计算机视觉任务产生负面影响。EnlightenGAN 方法虽然在整体亮度方面有所提升, 但存在色彩偏黄、细节丢失和噪声等问题, 其性能瓶颈会限制后续视觉任务的进行。反观 MIRNet 方法在亮度增强方面表现过度, 导致细节丢失和局部模糊。IAT 方法在色彩恢复方面表现较佳, 座椅细节较为清晰, 但存在较多噪声和点状模糊。HWMNet 和 CFWD 方法在座椅细节和色彩恢复方

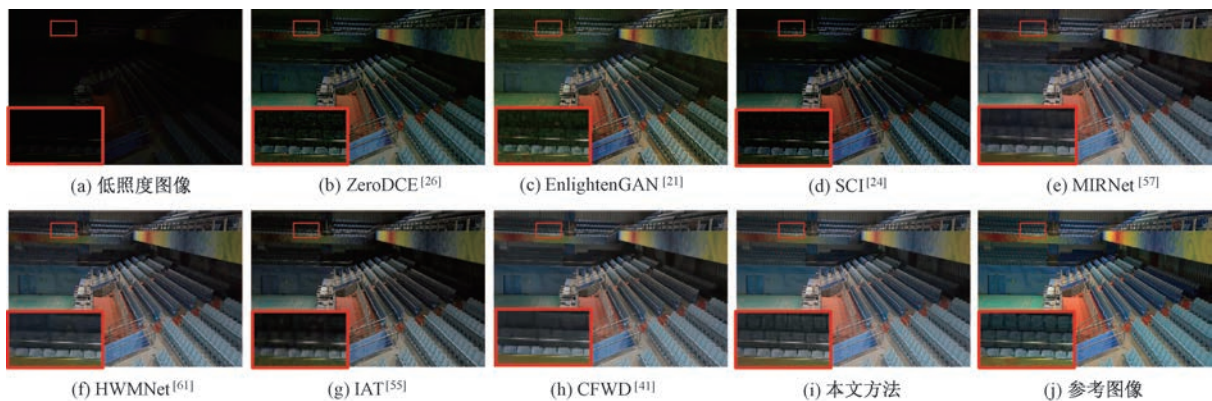


图 10 不同方法在 LOLv1 数据集的体育场图像上增强结果的视觉对比

面表现较优,但在深层细节处理时出现棋盘效应,细节边缘模糊,色彩饱和度不足。相比之下,本文提出的方法在细节保持方面表现突出,准确恢复了座椅的细节与轮廓,同时有效保留了原始图像的色彩信息,显著抑制了模糊现象。实验结果表明,本文方法在图像质量、细节信息以及色彩真实性等方面均展现出卓越的综合性能。

图 11 展示了各方法在储物柜图像上的增强效果。可以观察到,ZeroDCE、EnlightenGAN 和 SCI 方法的亮度提升效果有限,具体表现为纸盒边缘消失且

盒子细节丢失;此外,EnlightenGAN 方法存在严重的色偏问题。虽然 MIRNet 和 IAT 方法提升了亮度,但存在边缘噪点较多、细节不清晰的情况,导致盒子花纹因噪点变得模糊不清。相较于前述方法,HWMNet 和 CFWD 方法相对于其他方法恢复效果更佳,但仍存在花纹和盒子不清晰、色偏的问题。而本文提出的方法不仅成功恢复了左侧盒子的边缘细节,有效抑制了噪点和模糊,还清晰地恢复了右侧盒子的花纹,且更准确地还原了盒子的颜色。这些结果进一步证明了本文方法在细节保留和色彩真实性方面的优势。

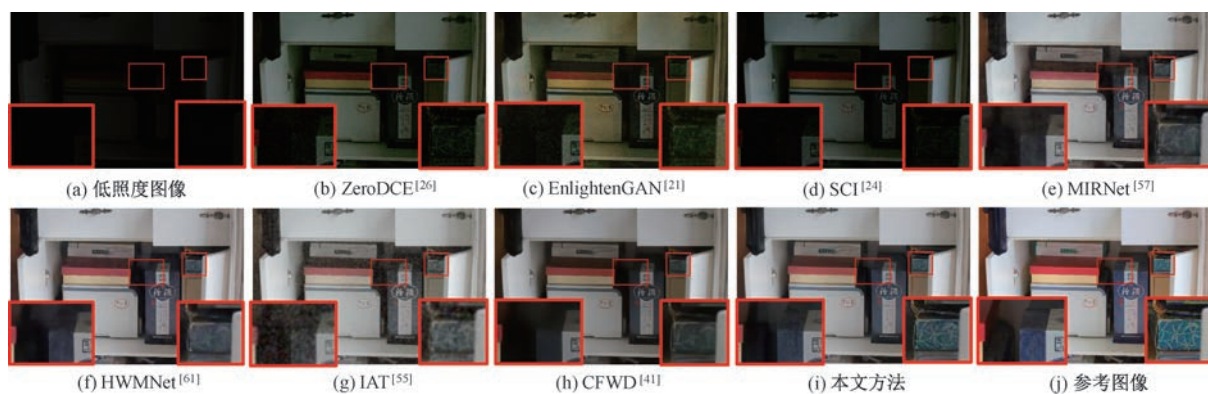


图 11 不同方法在 LOLv1 数据集的储物柜图像上增强结果的视觉对比

最后,为了更全面地评估本文所提方法的鲁棒性,我们又选取了三个不同数据集的图像进行测试:从 LOLv2-synthetic 数据集中随机选取了 3 张图像(图 12),从 LIME 和 MEF 数据集中各选取了 1 张图像(图 13、图 14),增强效果如图 12 至图 14 所示。

实验结果表明,相比于其他方法,本文方法在亮度提升、细节恢复和颜色还原方面均表现出优越性,有效抑制了噪声,恢复了细节信息,生成的结果更加真实自然。以上结果进一步证明了本文提出的 TSUNet 网络在低照度图像恢复方面的优越性能。

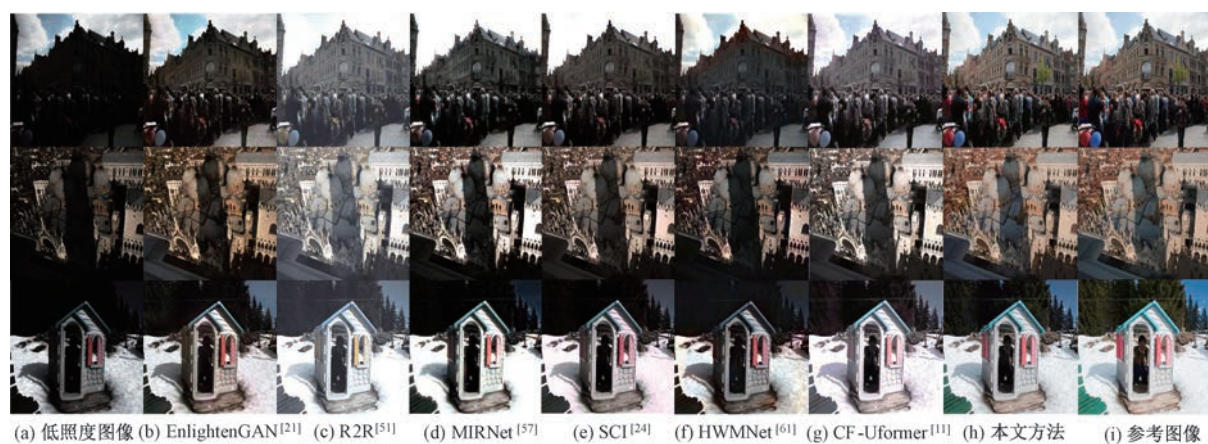


图 12 不同方法在 LOLv2-synthetic 数据集上的主观视觉对比效果

4.3.3 复杂场景对比分析

为了全面评估本文所提方法在极限场景下的图像恢复能力,我们设计了三组具有挑战性的低照度测试场景:(1)真实环境下的多光源高动态范围场景,用于验证算法在复杂光照条件下的细节保持能

力;(2)接近全黑的极限暗光场景,用于测试算法对极弱光信号的提取能力;(3)高度退化的高噪声场景,用于评估算法的噪声抑制及细节重建效果。这三组测试场景分别对应低照度图像增强中的关键难点:高动态范围压缩、极限弱光细节恢复以及噪声

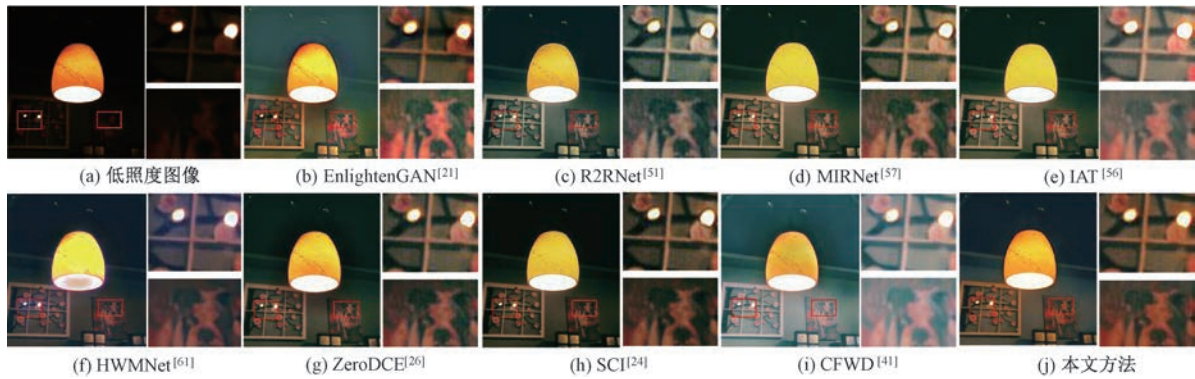


图 13 不同方法在 LIME 数据集上的主观视觉对比效果

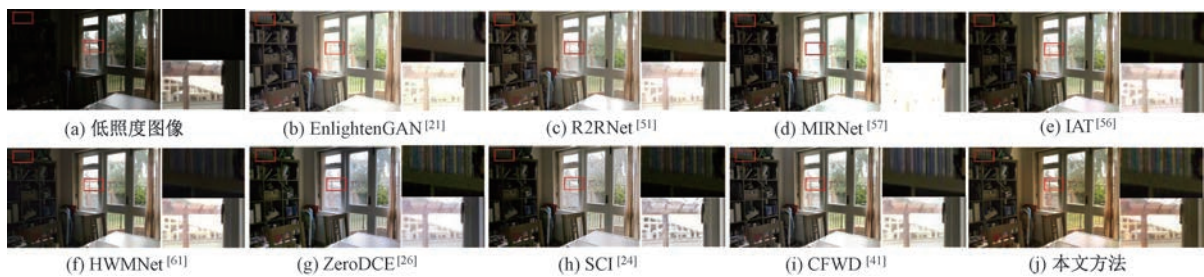


图 14 不同方法在 MEF 数据集上的主观视觉对比效果

抑制。这些挑战性场景将有助于全面评估所提算法在极端情况下的表现,并揭示其在处理不同类型低照度图像时的优势和局限性。

在多光源高动态范围场景测试中,我们选择了一个具有代表性的室内建筑场景,对比结果如图 15 所示。该场景包含水晶吊灯、壁灯等多种不同亮度的光源,并呈现拱形穹顶、墙面浮雕等精细的建筑细节,对算法处理能力构成显著挑战。从对比实验结果中可观察到,现有方法在处理此类复杂场景时均存在不同

程度的颜色失真问题,其中 MIRNet (图 e)、IAT (图 g)、CFWD (图 h) 均出现不同程度的灯光过曝现象。尽管其他方法提升了图像亮度,但仍然存在不同程度的细节丢失和恢复不足。相比之下,本文所提出的方法不仅保持了场景整体的亮度层次,而且准确还原了建筑细节,尤其在吊灯周围的穹顶纹理和墙面装饰等区域表现突出;此外,本文所提方法还对参考图像中的暗部场景也进行了增强,展现出 TSUNet 优异的细节恢复能力,使得增强后的图像更接近真实场景。

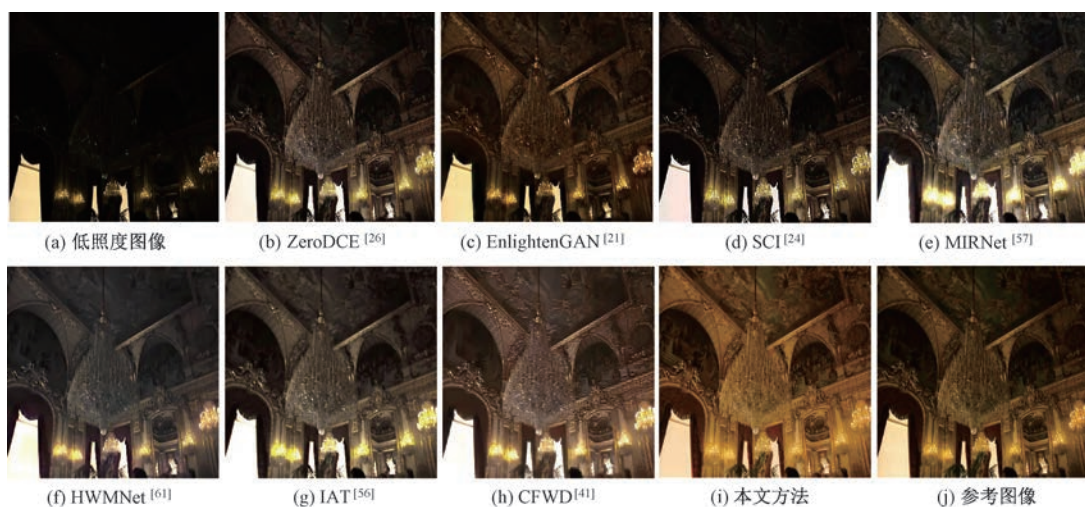


图 15 不同方法在极限多光源高对比度场景上的主观视觉对比效果

在接近全黑的极限暗光场景测试中,我们选取了一个具有挑战性的户外山地低照度场景作为测试对象。该场景整体亮度极低,且包含从近景松树到远景山体的不同景深层次,这对算法在暗光条件下提取弱光信号和恢复细节能力构成了严峻考验,尤其是在极限暗光场景下,针对树木、树叶以及水面细节的恢复难度极大。如图 16 所示,对比实验结果显示,现有方法在处理极限暗光场景时存在明显局限:ZeroDCE (图 b)、SCI (图 d)、MIRNet (图 e)和 IAT (图 g)方法虽然增强了亮度并恢复了部分山体轮廓,但在树木和

树叶细节的恢复方面表现出明显局限,恢复程度有限。EnlightenGAN (图 c)方法相较其他方法较好地恢复了树叶细节,但相对于参考图像出现了严重的色偏问题。HWMNet (图 f)方法在图像恢复过程中出现了不自然的夕阳色调,并完全丢失了树木和树叶的细节。CFWD (图 h)方法则表现出大量细节丢失和马赛克效应。相比之下,本文方法不仅成功提取出极弱光条件下的图像细节,并且还还原了场景的自然色彩,尤其是在树木和树叶的细节以及水面反光等方面,展现了 TSUNet 在极限暗光场景下的恢复能力。

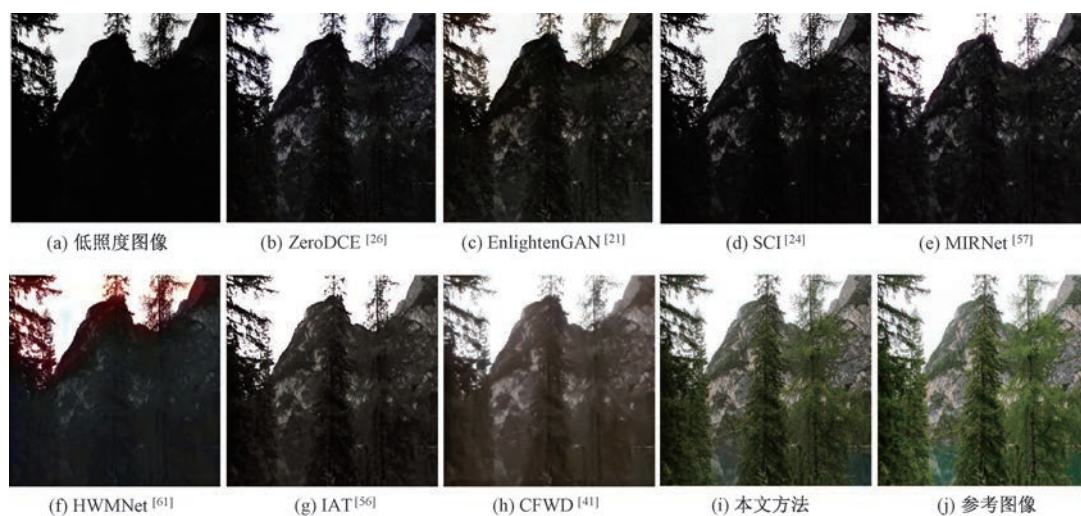


图 16 不同方法在极限暗光条件下恢复细节的主观视觉对比效果

为进一步验证本文方法在高噪声场景下的处理能力,我们选取了一张具有显著噪声干扰的室内图像作为测试图像,如图 17 所示。该图像包含标准色卡、玩偶等目标物体。从对比实验结果来看,现有方法在处理高噪声图像时存在一定的局限性,均无法有效去除噪声:EnlightenGAN (图 c)方法的恢复颜色呈现偏色,MIRNet (图 e)方法整体呈现蓝色。其他方法在图像恢复和噪声抑制方

面虽有一定效果,但均未能完全消除噪声,且图像整体清晰度和细节恢复方面存在明显不足。相比之下,本文方法在噪声抑制与细节恢复之间取得了良好平衡,不仅有效抑制了大量图像噪声,还同时保留并恢复了诸如牛奶盒和图书文字等细节。此外,通过色卡观察,本文方法在颜色恢复方面也取得了良好效果,展现了本文方法在高噪声极限环境下的恢复能力。

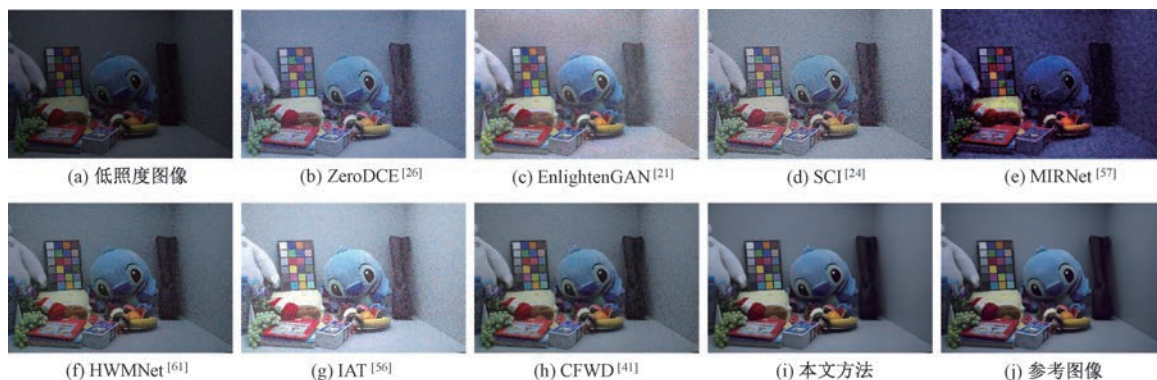


图 17 不同方法在高噪声场景上的主观视觉对比效果

最后,为了有效分析不同光照对本文方法的影响,我们在不同曝光条件下对同一图像进行恢复并进行对比。如图 18 所示,从左向右(A 到 D),原始图像的亮度逐渐降低,代表了从轻微昏暗到极度不足的光照条件。从恢复效果来看,在所有亮度等级下,算法成功将暗部细节提升至相近的视觉水平,展现出优异的稳定性。特别是书架上的书籍颜色和纹

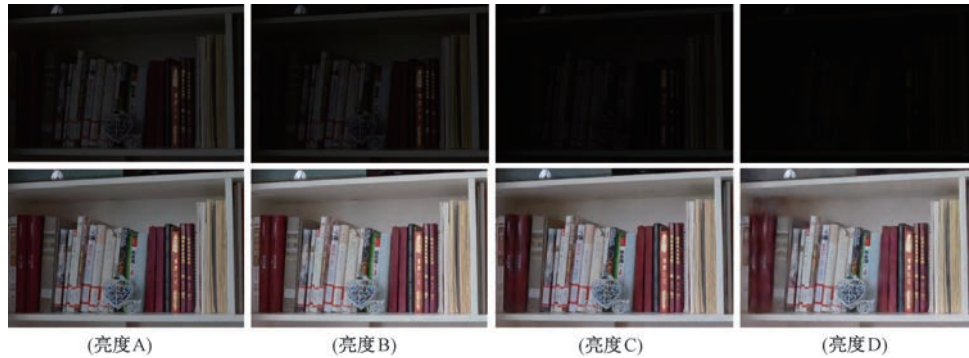


图 18 不同光照场景下本文方法恢复效果的主观视觉对比效果

4.4 消融实验

为了验证子模块的有效性,针对增强小波域特征融合模块、融合感知损失函数以及两阶段低照度图像增强网络进行了消融实验。通过逐一测试这些关键子模块能够验证它们对所提出的低照度图像增强网络的有效性。

4.4.1 增强小波域特征融合模块

为了展示 EWFM、双重注意力单元(DGSA + LFCA)以及其中的动态门控空间注意力(DGSA)和轻量融合曲线注意力(LFCA)对图像恢复的影响,我们设计了消融实验。这些消融实验清晰展示了各组成部分在低照度图像恢复中的作用,为整体网络优化提供了实验依据。

(1)从整体模块的影响来看:将增强小波域特征融合模块(EWFM)从 TSUNet 网络中移除,保持其他模块不变,并重新训练模型,实验结果如表 5 所示。

表 5 增强小波域特征融合模块对模型性能的影响

测试集	PSNR	SSIM
w/o EWFM	22.92	0.839
TSUNet	24.34	0.861

移除 EWFM 后,模型的 PSNR 从 24.34 下降至 22.92,SSIM 从 0.861 下降至 0.839。这一显著的性能下降表明 EWFM 在低照度图像增强中起到了关键作用。EWFM 通过 DWT 将图像分解为不同频率子带,使模型能够捕捉多尺度特征。移除 EWFM 意味着失去了这种多尺度分析能力,导致模

型,在四种不同程度的暗光条件下,恢复后都呈现出高度一致的视觉效果。即使在最暗的亮度 D 下,算法仍然保持了与轻微昏暗条件 A 下相当的恢复能力,展示了该方法出色的鲁棒性,即不会因为输入图像亮度的显著降低而导致恢复质量的明显下降,这表明该算法具有稳定的输出一致性,这对于实际应用场景来说是极为重要的特性。

型无法有效地区分和处理图像中的噪声和细节信息。EWFM 中的双重注意力机制(DGSA 和 LFCA)为模型提供了精细化的特征选择和融合机制。缺少这些机制会削弱模型的多尺度特征整合能力,影响图像增强效果。EWFM 专为提升低照度图像质量而设计,尤其在噪声抑制和细节保留方面。移除 EWFM 后,模型难以有效抑制噪声并增强细节,使图像质量下降。

(2)从双重注意力单元的影响来看:将双重注意力单元(DGSA + LFCA)、动态门控空间注意力(DGSA)和轻量融合曲线注意力(LFCA)分别从 TSUNet 网络中移除,保持其他模块不变,并重新训练模型,实验结果如表 6 所示。

表 6 双重注意力单元对模型性能的影响

测试集	PSNR	SSIM
w/o DGSA	23.90	0.853
w/o LFCA	23.67	0.846
w/o DGSA+LFCA	23.13	0.844
TSUNet	24.34	0.861

根据表 6 的结果显示,移除 DGSA 时,PSNR 和 SSIM 分别下降至 23.90 和 0.853。这是因为 DGSA 通过门控机制和空间注意力强化特征图中的关键空间信息,有助于模型更好地理解并恢复图像结构和纹理。DGSA 的移除导致模型无法充分捕捉这些关键信息,从而影响增强效果。此外,移除 LFCA 后,PSNR 和 SSIM 分别下降至 23.67 和 0.846。

LFCA 通过调整像素间的权重,优化了图像的色彩和对比度,这对于低照度图像中的暗区域细节恢复尤为重要。没有 LFCA,模型无法有效区分和增强这些细节,导致视觉质量下降。当同时移除 DG-SA 和 LFCA 时,PSNR 和 SSIM 进一步下降至 23.13 和 0.844。这一结果强调了两种注意力机制在提升图像质量方面的互补性。DGSA 能够提供空间维度的关注点,LFCA 则可以在逐像素上进行优化,两者结合使模型能更全面地理解并增强图像特征。

4.4.2 融合感知损失函数

为验证本文设计的融合感知损失 TotalLoss,包含 VGG 感知损失 VGGLoss、Charbonnier 损失 CharLoss 和轻量边缘损失 EdgeLoss 对模型性能的影响,进行了消融实验,结果如表 7 所示。实验结果表明,融合感知损失函数在提升模型性能方面发挥了显著的积极作用。

表 7 不同损失函数对模型性能的影响

测试集	PSNR	SSIM
w/o VGGLoss	23.11	0.818
w/o EdgeLoss	<u>23.71</u>	<u>0.851</u>
w/o CharLoss	22.83	0.847
w/o CharLoss+VGGLoss	22.39	0.822
w/o EdgeLoss+VGGLoss	22.93	0.821
w/o CharLoss+EdgeLoss	22.81	0.848
w/o TotalLoss	21.92	0.805
TSUNet	24.34	0.861

当仅使用 VGGLoss (w/o CharLoss+EdgeLoss)时,SSIM 指标相比于单独使用 CharLoss 和 EdgeLoss 时得到了显著提升,但 PSNR 指标明显下降。这表明 VGGLoss 有效捕捉了图像的高级语义和全局结构特征,增强了图像的结构相似性。然而,由于缺乏对像素级细节和边缘信息的约束,导致图像的还原不足,细节纹理有所损失。在移除 VGGLoss (w/o VGGLoss)后,PSNR 指标轻微下降,SSIM 指标显著下降。这意味着 VGG 感知损失虽然在像素级细节上有一定作用,但在全局结构和语义一致性方面具有更重要的作用。VGGLoss 利用预训练的 VGG 网络提取高级特征,对全局结构进行约束,有助于保持图像的结构相似性。

当仅使用 CharLoss (w/o EdgeLoss+VGGLoss)时,PSNR 指标相比于单独使用 VGGLoss 和 EdgeLoss 时得到了明显提升,这说明模型在像素级细节恢复方面表现良好。但相对较低的 SSIM 指标表明,图像的整体结构和纹理一致性有所欠缺。这种

现象是因为 CharLoss 专注于最小化像素级误差,而忽视了对全局结构和高级特征的约束。当移除 CharLoss (w/o CharLoss)后,模型无法有效减少像素级误差,导致 PSNR 和 SSIM 指标均有下降。这强调了 CharLoss 在细节恢复中的关键作用,通过直接约束像素差异,提升了图像的细节质量。

当仅使用 EdgeLoss (w/o CharLoss+VGGLoss)时,相比于单独使用 CharLoss 和 VGGLoss 时 PSNR 和 SSIM 指标均出现略微下降。虽然 EdgeLoss 强化了边缘特征的恢复,但由于缺乏对整体像素级细节和高级语义信息的约束,导致图像的整体质量不佳。这表明虽然 EdgeLoss 能够指导模型在图像恢复时生成更丰富的细节信息和更完整的结构,但仅依赖 EdgeLoss 无法满足图像恢复的全面需求。当移除 EdgeLoss (w/o EdgeLoss)后,模型的 PSNR 和 SSIM 指标都有所下降,这说明 EdgeLoss 在增强图像边缘细节和清晰度方面发挥了重要作用。缺少 EdgeLoss 后,图像的边缘细节变得模糊,影响了视觉效果中的细节表现。

当移除融合感知损失函数(w/o TotalLoss)时,模型的 PSNR 和 SSIM 指标均显著下降。这表明各个损失函数在模型训练中协同作用,共同提升了图像恢复的质量。融合感知损失函数通过综合 CharLoss、VGGLoss 和 EdgeLoss 的优点,平衡了细节恢复、结构保持和边缘增强,显著提升了模型性能。各损失函数的协同作用使模型在细节、结构和边缘等方面取得了良好的平衡。消融实验结果验证了这些损失函数在提升模型性能中的关键作用,缺少任何一个函数都会导致图像恢复质量的下降。因此,融合感知损失函数能够更有效地满足图像恢复过程中对细节、结构和边缘信息的综合需求。

4.4.3 两阶段低照度图像增强网络

为验证两阶段低照度图像增强网络中各部分对模型性能的影响,本文对浅层特征模块(SF)、精细增强阶段、改进的编一解码器以及有监督注意力模块(SA)进行了消融实验,实验结果如表 8 所示。实验结果表明,这些改进部分对模型性能都产生了显著的正面影响。

具体而言,浅层特征模块(SF)负责提取图像的初级特征,为后续增强过程奠定基础。消融实验结果显示,移除 SF 后,模型的 PSNR 降至 23.22,SSIM 降至 0.841,相较于 TSUNet 模型存在显著的性能下降。这一变化表明,SF 在保留图像基本信息方面发挥了重要作用,为后续处理提供了高质量输

入。SF 能有效捕获低层次特征,使模型更好地保留图像的整体结构,从而在初步恢复阶段中避免信息丢失,为细节增强奠定基础。

表 8 不同模块对模型性能的影响

测试集	PSNR	SSIM
w/o SF	23.22	0.841
w/o 精细增强阶段	23.46	0.849
w/o 改进的编-解码器	22.92	0.839
w/o GDM	24.06	0.855
w/o Mergeblock	23.63	0.851
w/o SA	23.75	0.852
w/o SA+GDM	23.01	0.848
w/o SF+GDM	23.04	0.845
w/o SA+SF	22.75	0.836
w/o SF+GDM+SA	22.42	0.835
TSUNet	24.34	0.861

精细增强阶段在进一步细化图像的细节和纹理方面起到了重要作用。移除该阶段后,PSNR 降至 23.46,SSIM 降至 0.849。尽管其对性能的影响略小于 SF 等模块,但依然表明了其对图像增强的重要贡献。精细增强阶段通过进一步提取局部特征,能够提升图像的边缘清晰度和纹理细节,使增强后的图像呈现更为细腻和丰富的视觉效果。该阶段有效地补充了初级恢复阶段难以捕捉的细节特征,特别是在复杂光照场景下,通过提升纹理和对比度质量,显著改善了整体的视觉感知效果。

改进的编-解码器模块是 TSUNet 架构的核心,负责多尺度特征的提取与融合,并通过增强小波域特征融合模块(EWFM)优化特征表示。消融实验中移除编解码器导致 PSNR 降至 22.92,SSIM 降至 0.839,表现出最显著的性能下降。结果表明,编-解码器模块通过有效的上下采样和多尺度特征融合,增强了模型对图像全局信息的捕捉能力,从而更好地恢复图像的对比度和结构。特别是 EWFM 模块的引入,极大地提升了图像的细节保真度,使得模型能够在保留细节和抑制噪声方面表现优异。因此,编-解码器模块对于 TSUNet 的整体性能具有重要的意义,是实现高质量图像恢复的关键之一。

梯度去噪模块(GDM)在去除图像噪声和增强清晰度方面发挥了作用。移除 GDM 导致 PSNR 降至 24.06,SSIM 降至 0.855,性能略有下降。GDM 模块通过双重残差机制有针对性地抑制噪声信息,使得图像增强过程更加稳定和清晰。实验结果表明,尽管去除 GDM 对模型性能的影响较小,但其在细节保留和噪声抑制方面提供了重要支持,特别是在高噪声的低照度环境中,通过降低干扰和抑制噪

声放大,提升了图像的整体质量。

融合模块(Mergeblock)的消融实验显示,在移除该模块后,PSNR 和 SSIM 分别降至 23.63 和 0.851,性能略有下降。Mergeblock 负责将不同阶段的特征进行融合,促进上下文信息的交互。这种特征融合能够增强图像的整体一致性和细节表现力,特别是在多层次信息的整合上,Mergeblock 发挥了关键作用。其作用在于确保图像的结构信息保持连贯性,并进一步提升细节层次感。

有监督注意力模块(SA)在聚焦图像的关键区域和提升细节表现上起到了辅助作用。移除 SA 后,PSNR 降至 23.75,SSIM 降至 0.852,性能略微下降。该模块通过注意力机制显著聚焦于关键特征区域,使得模型在复杂场景下能够更好地捕捉到重要细节。尽管 SA 对模型整体性能的影响小于其他核心模块,但其引入确保了对目标区域的聚焦,从而在细节增强和对比度提升方面发挥了积极作用。

联合消融实验进一步分析了各模块之间对增强效果的协同作用。例如,移除 SA 和 GDM 导致 PSNR 和 SSIM 分别降至 23.01 和 0.848,而移除 SA、SF 和 GDM 后,PSNR 和 SSIM 分别降至 22.42 和 0.835,表现出显著的性能下降。这表明,多模块的协同作用对于最终的增强效果具有至关重要的意义。尤其是 SA 和 GDM 的结合,能够同时提升关键区域的噪声抑制和细节保真度,而 SF 与 GDM 的组合则在初步恢复阶段提供了稳健的基础。联合消融实验揭示了 TSUNet 各模块的互补性,表明其协同作用对实现高质量图像增强效果不可或缺。

通过以上消融实验的深入分析,我们可以明确各模块在 TSUNet 中的独特贡献和重要性。这些模块在增强图像细节、降低噪声以及保持图像结构的一致性方面密切配合,共同促进了模型的整体性能提升,从而显著增强了低照度图像的视觉质量。

5 讨 论

本研究提出的双阶段小波增强网络 TSUNet,已在多个公开标准数据集上展现出其卓越的性能。深入分析表明,其性能优势主要源于以下几项核心创新:

(1)增强型小波域特征融合模块(EWFM)的创新设计:相较于 CFWD 等传统小波方法,EWFM 通过小波多尺度分析与双重注意力机制的嵌入式融合提升性能优势。具体而言,EWFM 模块采用离散小

波变换实现图像特征的频域解耦,并通过动态门控空间注意力(DGSA)和轻量融合曲线注意力(LF-CA)的协同作用实现了高频细节的定向强化。实验结果表明,相比同类方法,TSUNet 不仅能够在保证高质量增强效果的同时,还能够降低计算资源的消耗,进而在模型的复杂度和性能之间实现了良好平衡,显著提升了增强后图像的视觉质量和客观评价指标。

(2)双阶段级联网络架构的优化策略:相比主流 U-Net 类方法(如 R2RNet、CF-UFormer 等),TSUNet 采用创新的双阶段架构:初级恢复阶段旨在对输入的图像进行去噪和进行基础亮度增强,主要通过梯度去噪模块和浅层特征模块的协同作用来抑制噪声、提取基础特征;而精细增强阶段则进一步重建和细化图像的细节特征。通过双阶段处理任务,TSUNet 能够更为细致地逐步提升图像质量。不仅如此,TSUNet 在初级恢复阶段还整合了 EWFm 模块来增强模型的特征提取能力,使模型能够精确处理不同光照条件下的图像特征差异,从而提高图像增强的稳定性和鲁棒性。实验结果表明,与其他同类型方法相比,TSUNet 在多个数据集的图像质量评价指标中均方面处于领先地位。

尽管 TSUNet 在多个方面取得了显著进展,但该方法仍存在一些局限性:

(1)尽管 TSUNet 在静态图像增强任务中表现出色,但在处理动态视频的低照度增强时,动态场景中的复杂变化会对其性能产生不利影响,需进一步研究以解决动态场景中的帧间一致性问题。

(2)该方法在计算资源方面的消耗依然相对较高,尤其是在处理高分辨率图像时,这限制了其在实时或资源受限的应用场景中的适用性。因此,未来的工作将致力于优化网络结构,减少计算复杂度,并提高运行效率。

6 结 论

本文提出一种基于小波变换的双阶段低照度图像增强方法,该方法由初级恢复阶段和精细增强阶段构成。在初级恢复阶段,通过重新设计的浅层特征提取模块等组件的协同作用,有效提取图像信息与特征,实现初步的低照度增强,并为后续阶段输送丰富的恢复特征。精细增强阶段则专注于细节重建,最终生成高质量的增强图像。

为进一步提升性能,初级恢复阶段引入增强小波域特征融合模块,并结合本文重新设计的双重注

意力机制(包含轻量融合曲线注意力与动态门控空间注意力),实现对图像特征的精细化处理。同时,设计融合感知损失函数以优化训练过程,改善增强效果。实验表明,该方法在性能上优于现有先进方法。未来工作将重点优化模型结构以提升效率,并验证各模块在其他模型中的迁移能力,同时探索该方法在低照度视频增强中的应用潜力。

参 考 文 献

- [1] Jiang Ze-Tao, Li Hui, Lei Xiao-Chun, Zhu Ling-Hong, Shi Dao-Quan, Zhai Feng-Shuo. A low-light object detection method based on SAM-MSFF network. *Acta Electronica Sinica*, 2024, 52(1): 81-93(in Chinese)
(江泽涛,李慧,雷晓春,朱玲红,施道权,翟丰硕.一种基于 SAM-MSFF 网络的低照度目标检测方法. *电子学报*, 2024, 52(1): 81-93)
- [2] Wei C, Wang W, et al. Deep retinex decomposition for low-light enhancement. *arXiv preprint arXiv:1808.04560*, 2018
- [3] Cai Y, Bian H, Lin J, et al. Retinexformer: One-stage retinex-based transformer for low-light image enhancement// *Proceedings of the 2023 IEEE/CVF International Conference on Computer Vision*. Paris, France, 2023: 12504-12513
- [4] Wu W, Weng J, Zhang P, Wang X, Yang W, Jiang J. Ur-retinex-net: Retinex-based deep unfolding network for low-light image enhancement// *Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. New Orleans, USA, 2022: 5901-5910
- [5] Lee S, Jang D, Kim D. Temporally averaged regression for semi-supervised low-light image enhancement// *Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*. Vancouver, Canada, 2023: 4208-4217
- [6] Li J, Feng X, Hua Z. Low-light image enhancement via progressive-recursive network. *IEEE Transactions on Circuits and Systems for Video Technology*, 2021, 31(11): 4227-4240
- [7] Özcan M, Ergezer H, Ayazoglu M. Flight mode on: A feather-light network for low-light image enhancement// *Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*. Vancouver, Canada, 2023: 4226-4235
- [8] Lore K, Akintayo A, Sarkar S. LLNet: A deep autoencoder approach to natural low-light image enhancement. *Pattern Recognition*, 2017, 61: 650-662
- [9] Yu J, Zhao L. A novel U-Net model for low-light image enhancement and its application in art design. *Journal of Applied Science and Engineering*, 2023, 27(12): 3613-3622
- [10] Chao K, Song W, Shao S, Liu D, Liu X, Zhao X. CUI-Net: A correcting uneven illumination net for low-light image enhancement. *Scientific Reports*, 2023, 13(1): 12894

- [11] Sun Y, Sun J, Sun F, Wang F, Li H. Low-light image enhancement using transformer with color fusion and channel attention. *The Journal of Supercomputing*, 2024, 80: 18365-18391
- [12] Niu Yu-Zhen, Chen Ming-Ming, Li Yue-Zhou, et al. Task decoupling guided low-light image enhancement. *Acta Electronica Sinica*, 2024, 52(1): 34-45 (in Chinese)
(牛玉贞, 陈铭铭, 李悦洲, 赵铁松. 基于任务解耦的低照度图像增强方法. *电子学报*, 2024, 52(1): 34-45)
- [13] Zhou F, Sun X, Dong J, Zhu X. SurroundNet: Towards effective low-light image enhancement. *Pattern Recognition*, 2023, 141: 109602
- [14] Jiang Ze-Tao, Qin Lu-Lu, Qin Jia-Qi, et al. Low-light image enhancement method based on MDARNet. *Journal of Software*, 2021, 32(12): 81-93 (in Chinese)
(江泽涛, 覃露露, 秦嘉奇, 张少钦. 一种基于 MDARNet 的低照度图像增强方法. *软件学报*, 2021, 32(12): 81-93)
- [15] Yu N, Shi H, Han Y. Joint correcting and refinement for balanced low-light image enhancement. *IEEE Transactions on Multimedia*, 2024, 26(26): 6310-6324
- [16] Kannoth S, Kumar S, Raja K. Low light image enhancement using curvelet transform and iterative back projection. *Scientific Reports*, 2023, 13(1): 872
- [17] Jiang H, Luo A, Fan H, Han S, Liu S. Low-light image enhancement with wavelet-based diffusion models. *ACM Transactions on Graphics*, 2023, 42(6): 238
- [18] Wang Q, Huang P, Li L, Han L. WATNet: A wavelet-aware lightweight hybrid model for fast low-light enhancement//*Proceedings of the 7th Asian Conference on Artificial Intelligence Technology*. Jiaxing, China, 2023: 1556-1564
- [19] Batziou E, Ioannidis K, Patras I, et al. Low-light image enhancement based on U-Net and haar wavelet pooling//*Proceedings of the 2023 International Conference on Multimedia Modeling*. Bergen, Norway, 2023: 510-522
- [20] Chen C, Chen Q, Xu J, Koltun V. Learning to see in the dark//*Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City, USA, 2018: 3291-3300
- [21] Jiang Y, Gong X, Liu D, Cheng Y, Fang C, Shen X, et al. Enlightengan: Deep light enhancement without paired supervision. *IEEE Transactions on Image Processing*, 2021, 30: 2340-2349
- [22] Zhang Y, Di X, Zhang B, Wang C. Self-supervised image enhancement network: Training with low light images only. *arXiv preprint arXiv:2002.11300*, 2020
- [23] Jie H, Zuo X, Gao J, et al. LLformer: An efficient and real-time lidar lane detection method based on transformer//*Proceedings of the 2023 International Conference on Pattern Recognition and Intelligent Systems*. Shenyang, China, 2023: 18-23
- [24] Ma L, Ma T, Liu R, Fan X, Luo Z. Toward fast, flexible, and robust low-light image enhancement//*Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Vancouver, Canada, 2023: 5627-5636
- [25] Guo C, Li C, Guo J, et al. Zero-reference deep curve estimation for low-light image enhancement//*Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle, USA, 2020: 1780-1789
- [26] Li C, Guo C, Loy C. Learning to enhance low-light image via zero-reference deep curve estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021, 44(8): 4225-4238
- [27] Zhang Y, Zhang J, Guo X. Kindling the darkness: A practical low-light image enhancer//*Proceedings of the 27th ACM International Conference on Multimedia*. Nice, France, 2019: 1632-1640
- [28] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation//*Proceedings of the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention*. Munich, Germany, 2015: 234-241
- [29] Long J, Evan Shelhamer, et al. Fully convolutional networks for semantic segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(4): 640-651
- [30] Jiang Ze-Tao, Wu Xu, Zhang Shao-Qin. Low-illumination image enhancement based on MR-VAE. *Chinese Journal of Computers*, 2020, 43(7): 1328-1339 (in Chinese)
(江泽涛, 伍旭, 张少钦. 一种基于 MR-VAE 的低照度图像增强方法. *计算机学报*, 2020, 43(7): 1328-1339)
- [31] Huang Shu-Ying, Hu Wei, Yang Yong, Li Hong-Xia, Wang Bin. A low-exposure enhancement based on progressive dual network model. *Chinese Journal of Computers*, 2021, 44(2): 384-394 (in Chinese)
(黄淑英, 胡威, 杨勇, 李红霞, 汪斌. 基于渐进式双网络模型的低曝光图像增强方法. *计算机学报*, 2021, 44(2): 384-394)
- [32] Singh A, Chougule A, Narang P, Chamola V, Yu F. Low-light image enhancement for UAVs with multi-feature fusion deep neural networks. *IEEE Geoscience and Remote Sensing Letters*, 2022, 19: 1-5
- [33] Lim C, Loh Y, Wong L. LAU-Net: A low light image enhancer with attention and resizing mechanisms. *Signal Processing: Image Communication*, 2023, 115: 116971
- [34] Shi P, Xu X, Fan X, et al. LL-UNet++: UNet++ based nested skip connections network for low-light image enhancement. *IEEE Transactions on Computational Imaging*, 2024, 10: 510-521
- [35] Ji Z, Jung C. Subband adaptive enhancement of low light images using wavelet-based convolutional neural networks//*Proceedings of the 2021 IEEE International Conference on Image Processing*. Anchorage, USA, 2021: 1669-1673
- [36] Karthikeyan V, Raja E, Gurumoorthy K. Denoising convolutional neural network with energy-based attention for image enhancement. *Journal of Applied Analysis and Computation*, 2024, 14(4): 1893-1914
- [37] Chen Jun-Jie, Zhou Yong-Xia, Zu Jia-Zhen, Shen Wei, Zhao Ping. Dual-branch low-light image enhancement combined

- with dense wavelet transform. *Computer Engineering and Applications*, 2024, 60(4): 200-210 (in Chinese)
(陈俊杰, 周永霞, 祖佳贞, 盛威, 赵平. 结合稠密小波变换的双分支低照度图像增强. *计算机工程与应用*, 2024, 60(4): 200-210)
- [38] Zhou Q, Yang K, Ke Z, Wang G, Zhang Y, Jia Y, et al. Deep learning strategies for addressing anomalous exposure in image processing: the FARDBUNet approach//*Proceedings of the 5th International Conference on High Performance Big Data and Intelligent Systems*. Macao, China, 2023: 39-46
- [39] Song Y, Dai Y, Liu W, Liu Y, Liu X, Yu Q, et al. DesTrans: A medical image fusion method based on transformer and improved DenseNet. *Computers in Biology and Medicine*, 2024, 174: 108463
- [40] Zou W, Gao H, Yang W, Liu T. Wave-Mamba: Wavelet state space model for ultra-high-definition low-light image enhancement//*Proceedings of the 32nd ACM International Conference on Multimedia*. New York, USA, 2024: 1534-1543
- [41] Xue M, He J, He Y, Liu Z, Wang W, Zhou M. Low-light image enhancement via CLIP-fourier guided wavelet diffusion. *arXiv preprint arXiv:2401.03788*, 2024
- [42] Mou C, Wang Q, Zhang J. Deep generalized unfolding networks for image restoration//*Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. New Orleans, USA, 2022: 17399-17410
- [43] Wang P, Chen P, Yuan Y, Liu D, Huang Z, Hou X, Cottrell G. Understanding convolution for semantic segmentation//*Proceedings of the 2018 IEEE Winter Conference on Applications of Computer Vision*. Lake Tahoe, USA, 2018: 1451-1460
- [44] Mehri A, Ardakani PB, Sappa AD. MPRNet: Multi-path residual network for lightweight image super resolution//*Proceedings of the 2021 IEEE/CVF Winter Conference on Applications of Computer Vision*. Waikoloa, USA, 2021: 2704-2713
- [45] Li L, Zhang Y, Huang A. Learnable subspace orthogonal projection for semi-supervised image classification//*Proceedings of the 16th Asian Conference on Computer Vision*. Macao, China, 2023: 477-490
- [46] Howard A. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861*, 2017
- [47] He K, Zhang X, Ren S, Sun J. Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification//*Proceedings of the 2015 IEEE International Conference on Computer Vision*. Santiago, Chile, 2015: 1026-1034
- [48] Lai W, Huang J, et al. Deep laplacian pyramid networks for fast and accurate super-resolution//*Proceedings of the 2017 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Honolulu, USA, 2017: 624-632
- [49] Justin J, Alahi A, Li F. Perceptual losses for real-time style transfer and super-resolution//*Proceedings of the 14th European Conference on Computer Vision*. Amsterdam, The Netherlands, 2016: 694-711
- [50] Seif G, Androutsos D. Edge-based loss function for single image super-resolution//*Proceedings of the 2018 IEEE International Conference on Acoustics, Speech and Signal Processing*. Calgary, Canada, 2018: 1468-1472
- [51] Hai J, Xuan Z, Yang R, Hao Y, Zou F, Lin F, Han S. R2rnet: Low-light image enhancement via real-low to real-normal network. *Journal of Visual Communication and Image Representation*, 2023, 90: 103712
- [52] Yang W, Wang W, Huang H, Wang S, Liu J. Sparse gradient regularized deep retinex network for robust low-light image enhancement. *IEEE Transactions on Image Processing*, 2021, 30: 2072-2086
- [53] Lee C, Lee C, Kim CS. Contrast enhancement based on layered difference representation//*Proceedings of the 2012 IEEE International Conference on Image Processing*. Orlando, USA, 2012: 965-968
- [54] Guo X, Li Y, Ling H. LIME: Low-light image enhancement via illumination map estimation. *IEEE Transactions on Image Processing*, 2016, 26(2): 982-993
- [55] Ma K, Zeng K, Wang Z. Perceptual quality assessment for multi-exposure image fusion. *IEEE Transactions on Image Processing*, 2015, 24(11): 3345-3356
- [56] Cui Z, Li K, et al. You only need 90k parameters to adapt light: a light weight transformer for image enhancement and exposure correction//*Proceedings of the 33rd British Machine Vision Conference*. London, UK, 2022: 21-24
- [57] Zamir S, Arora A, Khan S, Hayat M, Khan F, Yang M, Shao L. Learning enriched features for real image restoration and enhancement//*Proceedings of the 16th European Conference on Computer Vision*. Glasgow, UK, 2020: 492-511
- [58] Oh H, Lee H, Ye S, et al. Instructir: A benchmark for instruction following of information retrieval models. *arXiv preprint arXiv:2402.14334*, 2024
- [59] Dang J, Zhong Y, Qin X. PPformer: Using pixel-wise and patch-wise cross-attention for low-light image enhancement. *Computer Vision and Image Understanding*, 2024, 241: 103930
- [60] Lian W, Lian W, Luo Z. Equipping diffusion models with differentiable spatial entropy for low-light image enhancement. *arXiv preprint arXiv:2404.09735*, 2024
- [61] Fan C, Liu T, Liu K. Half wavelet attention on M-Net+ for low-light image enhancement//*Proceedings of the 2022 IEEE International Conference on Image Processing*. Bordeaux, France, 2022: 3878-3882
- [62] Wang Z, Bovik A, Sheikh H, Simoncelli E. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 2004, 13(4): 600-612
- [63] Zhang R, Isola P, Efros A, Shechtman E, Wang O. The unreasonable effectiveness of deep features as a perceptual metric//*Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City, USA, 2018: 586-595



SUN Jing, Ph. D. , lecturer. Her main research interest is computer vision.

SUN Fu-Qi, M. S. candidate. His main research interest is image processing.

HAO Shi-Jie, Ph. D. , professor.

His main research interests include image processing and computer vision.

SUN Fu-Ming, Ph. D. , professor. His main research interests are image processing and computer vision.

Background

Images acquired in low-light environments often exhibit issues such as insufficient luminance, reduced contrast, and obscured details. These degradations not only compromise visual perception but also induce feature extraction biases, substantially impairing the performance of advanced computer vision tasks including object detection and semantic segmentation.

While conventional enhancement techniques based on histogram equalization and Retinex theory achieve elementary luminance correction, they exhibit marked limitations in multi-objective optimization involving noise suppression, detail enhancement, and dynamic range expansion, and are prone to chromatic aberrations and artifacts. Deep learning-based approaches, despite leveraging data-driven feature representation for notable progress, remain constrained by multi-objective optimization conflicts in complex scenarios, struggling to holistically resolve detail loss, noise suppression, and balanced enhancement across all dimensions.

To address these challenges, this paper presents TSU-Net—a dual-stage low-light enhancement network integrating wavelet transforms with U-Net architecture. First, we construct a cascaded two-phase network employing progressive preliminary restoration and refined enhancement. By decoupling the process into distinct stages, this architecture enables targeted optimization at each phase, mitigating risks of

over-correction or under-correction inherent in simultaneous multi-degradation resolution attempts. Second, we design an enhanced wavelet-domain feature fusion module incorporating discrete wavelet transforms, inverse wavelet transforms, and dual-attention mechanisms (dynamic gated spatial attention and lightweight fusion curves attention). This configuration empowers the network to adaptively prioritize spatially critical regions in feature maps, thereby accentuating and processing essential details. The lightweight fusion curve attention (LFCA) further facilitates efficient multi-branch and multi-scale feature integration, ensuring comprehensive yet refined feature representations for subsequent enhancement. Finally, a perceptual fusion loss function balancing pixel-level accuracy with visual quality metrics guides the model to produce enhanced images with naturalistic aesthetics and authentic detail preservation. Experimental validation demonstrates state-of-the-art performance across LOL and LSRW benchmarks, achieving superior metrics in PSNR, SSIM, and LPIPS evaluations compared to mainstream methods.

This work was supported by the National Natural Science Foundation of China (No. 62172137), Liaoning Science and Technology Joint Program (No. 2024JH2/102600091), Dalian Science and Technology Talent Innovation Policy Implementation Project (No. 2022RQ086), and Fundamental Research Funds for Central Universities (No. 04442024061).