E-mail: jig@aircas.ac.cn
Website: www.cjig.cn
Tel: 010-58887035
中国图象图形学报
JOURNAL OF IMAGE AND GRAPHICS
©中国图象图形学报版权所有

中图法分类号:TP391.4;TP18　　**文献标识码**:A　　**文章编号**:1006-8961(2026)09-3352-16

论文引用格式:Li W G, Gui K, Xiao Z L, Jin H Y and Su H N. 2026. Low-light image enhancement with frequency feature refinement and truncated sampling in conditional diffusion models. Journal of Image and Graphics, 31(9):3352-3367(李文改, 桂柯, 肖照林, 金海燕, 苏浩楠. 2026. 频域特征优化与截断采样的条件扩散模型弱光图像增强方法. 中国图象图形学报, 31(9):3352-3367)[DOI:10.11834/jig.250342]

频域特征优化与截断采样的条件扩散模型弱光图像增强方法

李文改,桂柯,肖照林*,金海燕,苏浩楠

西安理工大学计算机科学与工程学院,西安　710048

摘　要:**目的**　现有基于扩散模型的弱光图像增强方法存在计算成本高、高频特征畸变以及色彩偏移等问题,限制了其在实际场景中的应用效果。为此,提出一种频域特征优化与截断采样的条件扩散模型弱光图像增强方法(low-light image enhancement with diagonal frequency feature refinement and truncated sampling in conditional diffusion models)。**方法**　该方法首先在小波低频子空间中实施扩散建模,有效降低了空间维度和计算复杂度,同时保留图像的主体结构与光照分布特征。针对高频子带中对角线方向信息较弱的问题,本文设计高频细化模块,利用水平方向与垂直方向的高频特征作为引导,通过双路径交叉注意力机制实现对角线高频纹理的补全,从而提升局部细节与纹理的清晰度。此外,提出一种高效的条件采样策略,采用特定步数截断技术,大幅减少了推理时间,同时保持增强结果的质量。为进一步提高对比度与色彩还原表现,在反向去噪过程中使用上一步采样的色彩与对比度信息调制当前采样步特征,从而实现对增强图像的对比度校正。**结果**　在LOLv1(low-light image enhancement dataset version 1)、LOLv2-Real和VE-LOL(vision enhancement in low-light conditions)3个配对数据集上,本文方法在定性与定量评估上均优于现有主流方法。在LOLv2-Real数据集上,峰值信噪比(peak signal-to-noise ratio,PSNR)和结构相似性(structural similarity index,SSIM)分别提升8.46%、1.22%,LPIPS(learned perceptual image patch similarity)和NIQE(naturalness image quality evaluator)分别下降10.19%、0.38%。在4个非配对数据集上亦展现良好的跨域泛化性能,相较于基于空域扩散建模的PyDiff(pyramid diffusion models for low-light image enhancement)方法,单幅图像的推理时间由7.617 s减少至0.473 s,显存占用降低48.53%。**结论**　大量定量及定性的实验结果表明,本文方法在提升图像亮度、细节和色彩保真的同时降低了推理开销,展现出良好的实用性与应用前景。代码已共享在Gitee上:https://gitee.com/cv_llie/llie_diff。

关键词:弱光图像增强(LLIE);扩散模型;小波包变换(WPT);截断采样策略;频域建模

Low-light image enhancement with diagonal frequency feature refinement and truncated sampling in conditional diffusion models

Li Wengai, Gui Ke, Xiao Zhaolin*, Jin Haiyan, Su Haonan

School of Computer Science and Engineering, Xi'an University of Technology, Xi'an　710048, China

Abstract: **Objective** Low-light image enhancement (LLIE) aims to recover visually pleasing images from severely under-

收稿日期:2025-08-03;**修回日期**:2026-01-30;**预印本日期**:2026-02-06

*通信作者:肖照林　xiaozhaolin@xaut.edu.cn

基金项目:国家自然科学基金项目(62371389,62031023)

Supported by:National Natural Science Foundation of China(62371389,62031023)

exposed inputs by improving brightness, contrast, and structural clarity while preserving natural colors and fine details. This task is fundamentally important for numerous downstream vision applications, including video surveillance, autonomous driving, remote sensing, and medical imaging, where low illumination conditions frequently degrade visual quality and impair algorithmic reliability. In recent years, diffusion probabilistic models have demonstrated remarkable generative capacity and robustness, making them a promising paradigm for LLIE. However, directly applying diffusion models to low-light enhancement remains challenging due to several intrinsic limitations. First, the iterative denoising mechanism of diffusion models introduces substantial computational and memory overhead, hindering their deployment in real-time or resource-constrained scenarios. Second, most existing diffusion-based LLIE approaches operate in the spatial domain, which often struggles to recover high-frequency textures and fine structural details faithfully under severe illumination degradation, leading to oversmoothed results or visually disturbing artifacts. Third, the accumulation of prediction errors during reverse diffusion may cause color shifts and temporal inconsistencies across denoising steps, thereby reducing color fidelity and perceptual naturalness. **Method** To address these challenges, this study proposes LLIE with diagonal frequency feature refinement and truncated sampling in conditional diffusion models. The core idea is to exploit jointly the complementary strengths of frequency-domain modeling and accelerated diffusion inference to achieve high-quality enhancement with considerably reduced computational cost. Specifically, instead of performing diffusion directly in the spatial domain, the proposed method conducts forward and reverse diffusion in the low-frequency subspace of the wavelet domain. This design choice is motivated by the observation that global illumination structure and coarse semantic content are predominantly concentrated in low-frequency components. By diffusing only low-frequency coefficients, the proposed framework effectively reduces spatial resolution and computational complexity while preserving the essential structural and illumination information required for faithful enhancement. However, operating in the low-frequency subspace inevitably leads to the loss of high-frequency details, particularly diagonal textures that are critical for visual sharpness. To compensate for this limitation, we introduce a dedicated high-frequency refinement module that explicitly models directional dependencies in the wavelet domain. In contrast to conventional approaches that treat high-frequency bands independently, the proposed module leverages horizontal and vertical high-frequency components to reconstruct missing diagonal information through a dual-path cross-directional attention mechanism. By enabling feature interaction across orthogonal directions, the refinement module effectively alleviates directional sparsity and enhances local texture fidelity, resulting in sharper edges and more natural fine details. To improve efficiency further, an efficient truncated sampling strategy is incorporated into the reverse diffusion process. Instead of executing the full sequence of denoising steps, the proposed approach adopts a step-interrupt mechanism that terminates diffusion at an earlier stage once sufficient reconstruction quality is achieved. This strategy considerably reduces inference time and memory consumption without introducing noticeable degradation in visual quality, making the method more suitable for practical applications. In addition, to mitigate color shifts and enhance stability during reverse diffusion, a contrast-aware correction module is integrated into the denoising stages. This module exploits contrast and color priors accumulated from previous diffusion steps to adjust brightness and chromatic consistency adaptively at each iteration. By explicitly enforcing contrast awareness throughout the sampling process, the proposed method effectively suppresses error accumulation and improves color fidelity in the final enhanced images. **Result** To validate the effectiveness of the proposed method, we conducted extensive experiments on three publicly available paired low-light datasets, comparing the proposed method with several state-of-the-art LLIE methods. Subjective visual comparisons and objective quantitative metrics demonstrate that the proposed method consistently outperforms existing approaches. On the LOLv2-Real dataset, the proposed method achieves improvements over the second-best method by 8.46% in peak signal-to-noise ratio, 1.22% in structural similarity index, a 10.19% reduction in learned perceptual image patch similarity (LPIPS), and a 0.38% reduction in naturalness image quality evaluator (NIQE), indicating improved perceptual quality and structural fidelity. To assess the cross-domain generalization ability, we further evaluated the model using the pretrained weights from LOLv2-Real datasets on four unpaired low-light datasets. The proposed method achieved lower NIQE scores than existing methods by 0.82 (LLIE via illumination map estimation), 0.39 (multiexposure image fusion) and 0.86 (naturalness photo enhancement), demonstrating strong generalization ability without target-domain fine-tuning. Additionally, ablation studies confirm the contribution of each core module. Compared with spatial-domain diffusion baselines, the proposed

method considerably reduced inference time from 7.617 s to 0.473 s per image, and memory usage decreased by 48.53%, highlighting the efficiency and practicality of the proposed design. **Conclusion** The proposed method integrates frequency-domain modeling with an efficient diffusion sampling strategy, introducing three key innovations: 1) performing the diffusion process in the low-frequency subspace of the wavelet domain considerably reduces computational complexity while preserving global structural and illumination information. 2) A dedicated high-frequency refinement module is developed to recover diagonal texture details by leveraging horizontal and vertical high-frequency information, thereby improving the sharpness and perceptual quality of the enhanced images. 3) The integration of an efficient truncated sampling strategy and a contrast-aware correction module improves sampling efficiency and color fidelity, effectively balancing enhancement quality and computational cost. Collectively, these contributions enable the proposed method to achieve superior enhancement performance with substantially reduced inference overhead, demonstrating strong potential for real-world low-light vision applications.

Key words: low-light image enhancement (LLIE); diffusion model; wavelet packet transform (WPT); truncated sampling strategy; frequency domain modeling

0 引 言

弱光图像增强(low-light image enhancement, LLIE)任务旨在提升低照度场景下图像的视觉感知质量,使图像能够传达更丰富的结构与语义信息。该任务在自动驾驶(Li 等,2024;余烨 等,2023)、医疗影像(Ma 等,2021)和安防监控(Liang 等,2023;毕诗帆 等,2026)等下游视觉任务中具有重要的应用价值。传统的 LLIE 方法主要基于图像的统计规律或人类视觉感知进行模型设计。直方图均衡化方法(Zhang 等,2024)通过拉伸图像的灰度范围提升图像亮度,但该方法缺乏对图像内容与噪声的区分能力,导致噪声过度放大以及局部对比度失衡。Retinex 理论(Wei 等,2018)将图像分解为光照和反射分量分别处理,能够实现自然的增强效果,但容易导致伪影和色彩失真问题。随着深度学习技术的发展,LLIE 逐渐转向以数据驱动的端到端方法,通过光照估计(Jiang 等,2021;Zhou 等,2023)、正则化流(黄颖 等,2024;Wang 等,2022)与可学习分解(Wu 等,2025)等策略显著提升了增强效果。然而,这些方法仍受限于光照建模不稳定、高频纹理恢复不足和计算开销较大的问题。

现有 LLIE 方法虽然在亮度提升与结构还原上取得了进展,但存在增强亮度时放大图像噪声的问题,导致细节模糊、伪影失真。为解决这一问题,去噪扩散概率模型(denoising diffusion probabilistic model, DDPM)(Ho 等,2020)为 LLIE 提供了新的解决方案。DDPM 通过在正向过程中逐步添加噪声,并在反向过程中逐步去噪以恢复图像,从而能够生成高保真且细节丰富的结果。在此基础上,大量基于空域扩散的 LLIE 方法在像素空间中执行迭代重建,以同时增强亮度并抑制噪声。LLDiffusion(Wang 等,2025)提出两阶段扩散框架,结合物理先验与扩散模型的生成能力,提升了结构还原和光照建模的准确性。然而,该方法训练过程较为复杂,且易受到先验误差的影响。Diff-Retinex(Yi 等,2023)则分别对光照图和反射图进行优化,并通过联合训练保持两者的物理一致性,从而避免了传统方法中容易出现的过曝光或伪影问题。为了缓解扩散模型采样速度慢的问题,PyDiff(Zhou 等,2023)通过构建金字塔结构,从低分辨率到高分辨率逐级引导扩散过程,但该方法在高频细节保留方面仍存在不足,且金字塔重建过程中存在结构失真累积的风险。与空域扩散方法不同,WCDM(Jiang 等,2023)将扩散建模迁移到频域,通过对小波低频子带进行扩散重建来降低计算量,但其缺乏高低频信息交互与色彩对比度约束,仍难以在复杂光照条件下准确恢复细节与保持图像自然度。此外,已有研究还尝试通过加速策略以提高模型效率,如确定性隐式采样(Lan 等,2025;Song 等,2020)、知识蒸馏(Li 等,2023;Zhang 等,2026)和重构扩散常微分方程(Liu 等,2022)等方式降低训练开销。然而,这些方法的计算复杂度依然随着图像尺寸的增加呈现指数级的增加。例如,在智能监控场景中,对于一幅分辨率为 2 176 × 1 536 像素的夜间监控图像,在 RTX 3090D GPU 上采用扩散模型执行 25 步采样,处理时间约需 650 s,严重限制了其在实际场景中的应用。

综上所述，尽管扩散模型在弱光增强任务中展现出巨大潜力，但仍面临以下问题：1）反向去噪的迭代机制具有高昂的计算成本，严重限制了模型在实际场景的部署效率；2）空域的扩散方法难以重建弱光图像中的高频纹理与结构细节，导致增强图像仍存在模糊或伪影问题；3）反向扩散过程中偏差的逐步积累，导致生成图像出现色偏，影响增强结果的自然度与一致性。

针对上述问题，本文提出一种频域特征优化与截断采样的条件扩散模型弱光图像增强方法，旨在降低计算复杂度，提升增强图像的质量。该方法通过在小波低频子空间中实施扩散建模，有效保留了图像的主体结构与光照信息。并针对高频子带设计高频细化模块，利用水平与垂直方向的高频特征重建对角线方向的高频系数，实现局部纹理的高质量还原。此外，本文利用高频细化模块生成的高频先验作为条件，引导反向去噪过程在较少的采样步骤内完成收敛，并通过步数截断进一步减少计算量，从而显著降低推理时间与显存消耗。为进一步提升增强结果的稳定性与色彩还原能力，本文在反向去噪过程中设计了对比度校正模块，通过引入上一步的颜色与对比度信息，修正当前步的色彩和亮度表现，且几乎不增加额外计算负担。本文方法的创新性主要包括以下 3 个方面：1）在小波低频子空间中实施扩散过程，相较于空域建模显著降低了计算复杂度，并保留了全局结构和光照信息；2）提出高频细化模块，利用水平与垂直方向的高频信息补全对角线方向的纹理细节，从而提升纹理细节恢复；3）采用高效截断采样策略和对比度校正模块，提升采样效率与色彩保真性，兼顾模型实用性与增强质量。

1 方　法

针对现有基于扩散的 LLIE 方法在推理过程中高度依赖编码输入，且迭代去噪机制导致计算耗时较长的问题，本文提出一种频域特征优化与截断采样的条件扩散模型弱光图像增强方法。该方法流程如图 1 所示。

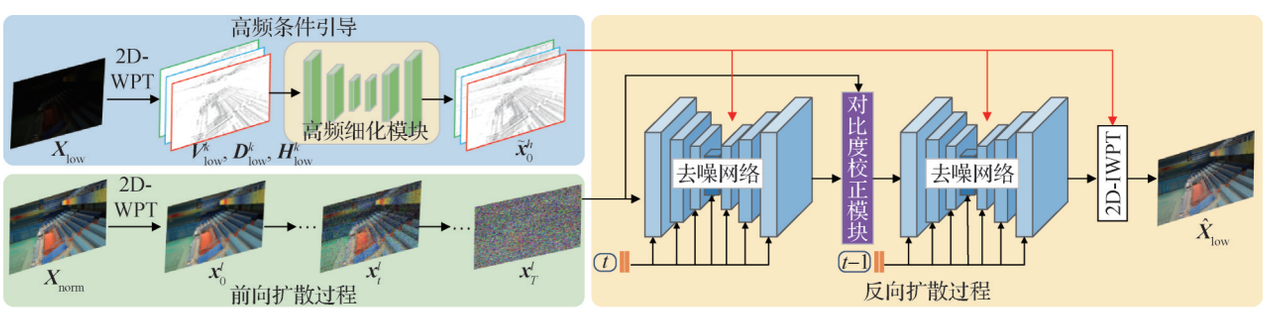


图 1　频域特征优化与截断采样的条件扩散模型弱光图像扩散增强方法流程

Fig. 1　Pipeline of low-light image enhancement with diagonal frequency feature refinement and truncated sampling in conditional diffusion models

首先，本文利用小波包变换分离图像的光照与细节信息，将弱光图像 $\boldsymbol{X}_{\text{low}} \in \mathbf{R}^{H\times W\times 3}$ 和正常光图像 $\boldsymbol{X}_{\text{norm}} \in \mathbf{R}^{H\times W\times 3}$ 分别分解为多个频率子带。低频子带包含图像的主要结构和整体亮度分布；高频子带则包含边缘、纹理等细节信息。高频子带根据方向不同分为水平、垂直及对角线。鉴于低频子带的信息密度更高，且包含主要的光照与结构特征，本文选择在小波低频子带 x^l_0 的子空间中实施前向扩散过程。相较于在像素空间中进行扩散建模的方法，该策略有效降低了计算复杂度，同时完整地保留了图像的主体结构和光照特征。在反向去噪过程中，本文引入高频子带 $\left(\tilde{\boldsymbol{V}}^k_{\text{low}}, \tilde{\boldsymbol{D}}^k_{\text{low}}, \tilde{\boldsymbol{H}}^k_{\text{low}}\right)$ 作为条件引导信息，以充分利用其所包含的边缘、纹理等细节特征，从而提高生成图像的清晰度。然而，小波变换的高频子带更倾向于捕捉水平方向 H^k_{low} 与垂直方向 V^k_{low} 的局部特征，导致对角线方向 D^k_{low} 的信息易被忽略。因此，本文设计高频细化模块来重建高频子带 $\tilde{x}^h_0$，通过垂直和水平方向的细节特征补充对角线方向的缺失信息，实现图像纹理细节的高质量重建；同时，本文采用高效的截断采样策略，从而降低计算资源开销，并抑制输出内容的随机性。与传统扩散模型采样方法相比，该策略在不影响增强图像质量的前提下，将

采样步数从20步降低至5步，有效简化了扩散模型的采样过程，减轻了大量迭代步骤带来的计算资源负担。此外，传统的扩散反向过程不仅耗时长，还存在图像对比度不足的问题，导致模型推理时间长且出现严重的色偏问题。为此，本文设计了轻量化的对比度校正模块，从而提高了增强图像的颜色与对比度的视觉质量。最终，本文通过逆小波包变换将频域重构结果映射至图像空间，生成增强图像$\hat{\boldsymbol{X}}_{\text{low}}$。

1.1 基于WPT的频率分解方法

在弱光条件下，图像呈现出光照不足、细节模糊等问题，且光照信息主要集中在低频成分，而边缘与纹理等细节信息则分布在高频区域。因此，频域建模被认为是LLIE的一种有效方法。相较于快速傅里叶变换（fast Fourier transform，FFT）等传统频率分析方法，离散小波变换（discrete wavelet transform，DWT）具有良好的时频局部化特性，能够在不同尺度上对图像内容进行分解，兼顾局部特征与全局结构，从而更适配弱光图像中信息分布稀疏、频域跨域性强的特点。进一步，本文采用二维小波包变换（wavelet packet transform，WPT）代替二维DWT。与DWT仅对每层的低频子带进行递归分解不同，WPT能够在每一层中对所有子带进行进一步分解，从而更充分地捕捉高频成分中的多尺度细节变化，提升图像的细节清晰度。此外，WPT通过最小化代价函数计算出最优的信号分解路径，在不损失有效信息的前提下将每次变换后的空间维度减半，降低后续处理过程中的计算资源开销。具体而言，对于输入图像$\boldsymbol{X}_{\text{low}}$，二维WPT将图像分解为4个子带，具体表示为

$$\left(\left(\boldsymbol{V}_{\text{low}}^{k},\boldsymbol{D}_{\text{low}}^{k},\boldsymbol{H}_{\text{low}}^{k}\right),\boldsymbol{x}_{0}^{l}\right)=WPT_{\text{2D}}\left(\boldsymbol{X}_{\text{low}}\right) \tag{1}$$

式中，$WPT_{\text{2D}}(\cdot)$表示二维小波包变换，k表示分解层数。每次分解后得到1个低频子带$\boldsymbol{x}_{0}^{l}$，3个高频子带$\boldsymbol{V}_{\text{low}}^{k}$，$\boldsymbol{D}_{\text{low}}^{k}$，$\boldsymbol{H}_{\text{low}}^{k}$，其分别表示在垂直、对角线、水平方向经过高通滤波器卷积后的高频子带。WPT的显著优势在于，其在各个分解层级中所有子带具有一致的空间维度，使得各频率分量能够以结构化的方式进行拼接与分析，从而捕捉图像的多尺度频率细节。这对弱光图像中光照和纹理信息的提取至关重要。在完成对低频子带和高频子带的增强处理后，通过二维逆小波包变换（inverse wavelet packet transform，IWPT）将图像从频域转换回空域，以实现图像的重建，其具体过程为

$$\hat{\boldsymbol{X}}_{\text{low}}=IWPT_{\text{2D}}\left(\tilde{\boldsymbol{x}}_{0}^{h},\hat{\boldsymbol{x}}_{0}^{l}\right) \tag{2}$$

式中，$\tilde{\boldsymbol{x}}_{0}^{h}$表示增强后的高频子带$\left(\tilde{\boldsymbol{V}}_{\text{low}}^{k},\tilde{\boldsymbol{D}}_{\text{low}}^{k},\tilde{\boldsymbol{H}}_{\text{low}}^{k}\right)$，$\hat{x}_{0}^{l}$表示扩散模型重建后的低频子带。IWPT重建过程从最深的分解层开始，逐步向上至第1层，最终将所有子带合并恢复出增强图像$\hat{\boldsymbol{X}}_{\text{low}}$。该过程不仅保证了频域增强信息的有效传递，还完整保留了图像的结构和细节纹理。

1.2 高频细化模块

尽管WPT在细节分解方面具有优势，但其高频子带的对角线方向细节信息在多层分解后往往变得稀疏甚至被削弱，影响了图像纹理信息的完整重建。为此，本文设计了高频细化模块（high-frequency refinement module，HFRM），通过融合水平方向与垂直方向的细节特征，引导对角线高频特征的重建与补全，从而实现更精确的纹理还原，提升增强图像的细节质量与清晰度。其结构如图2所示。该过程表示为

$$\left(\tilde{\boldsymbol{V}}_{\text{low}}^{k},\tilde{\boldsymbol{D}}_{\text{low}}^{k},\tilde{\boldsymbol{H}}_{\text{low}}^{k}\right)=HFRM\left(\boldsymbol{V}_{\text{low}}^{k},\boldsymbol{D}_{\text{low}}^{k},\boldsymbol{H}_{\text{low}}^{k}\right) \tag{3}$$

在特征提取阶段，高频细化模块采用深度可分离卷积结构，以实现高频系数的多尺度、低冗余特征表示。在此基础上，本文构建了双路径交叉注意机制分别通过水平方向$\boldsymbol{H}_{\text{low}}^{k}$与垂直方向$\boldsymbol{V}_{\text{low}}^{k}$的特征补偿对角线方向$\boldsymbol{D}_{\text{low}}^{k}$的细节信息。随后，利用膨胀卷积组恢复局部细节，其中第一个和最后一个卷积用于提取局部信息，中间的膨胀卷积用于改善感受野，以捕获长距离依赖关系。通过膨胀率的渐进式调节策略（先递增后递减），设置膨胀率分别为$d=\{1,2,3,2,1\}$。该策略不仅提升了对特征全局依赖和局部连续的建模能力，同时避免了常规膨胀卷积易产生的网格效应。最后，再次采用深度可分离卷积对融合后的特征进行处理，输出重建后的高频子带$\tilde{\boldsymbol{V}}_{\text{low}}^{k}$，$\tilde{\boldsymbol{D}}_{\text{low}}^{k}$，$\tilde{\boldsymbol{H}}_{\text{low}}^{k}$，从而实现高质量的细节恢复与频率补全。

1.3 高效的条件采样策略

基于空域的LLIE扩散方法计算复杂度高、推理过程耗时长，限制了其在实际应用中的部署效率。因此，本文将扩散建模从空域迁移至频域，选择在WPT的低频子带上实施前向扩散过程，即向初始低频子带$\boldsymbol{x}_{0}^{l}$逐步添加随机高斯噪声，该过程表示为

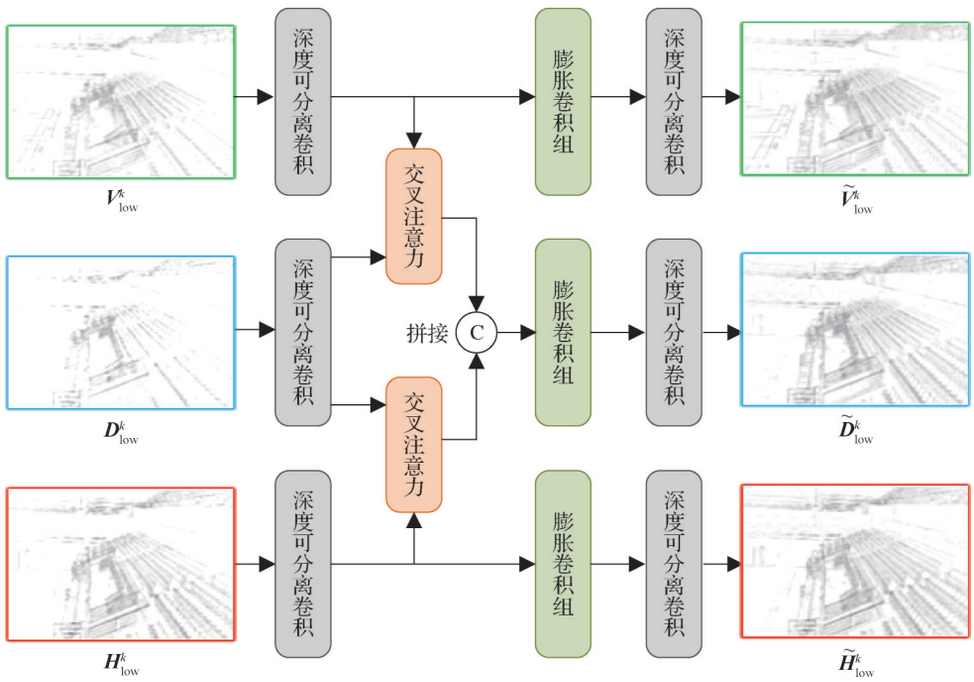


图 2　高频细化模块结构示意图

Fig. 2　Illustration of the high-frequency refinement module structure

$$q\left(\boldsymbol{x}_t^l \middle| \boldsymbol{x}_0^l\right)=\mathcal{N}\left(\boldsymbol{x}_t^l;\sqrt{\bar{\alpha}_t}\,\boldsymbol{x}_0^l,\left(1-\bar{\alpha}_t\right)\boldsymbol{I}\right) \tag{4}$$

式中，$t=1,2,\cdots,T$，$\bar{\alpha}_t$ 为控制加噪强度的累积量，$\boldsymbol{I}$ 表示协方差。前向过程的步数与标准扩散模型一致，以确保训练阶段的稳定性。本文在反向扩散过程采用截断采样机制，将反向去噪步骤压缩至 5 步。与 DDIM（denoising diffusion implicit model）的跳跃加速不同，本文在低秩频率空间中进行扩散建模，使得反向过程在更少的采样步数下快速收敛至高质量的增强结果。在每个反向采样步骤中，高频细化模块提供的高频先验 $\tilde{x}_0^h$ 作为条件，以引导低频子带逐步去噪重构。该条件引导强化了高低频之间的协同作用，使低频重建方向更加明确，从而提升收敛效率，在显著减少采样步数的同时获得更优的视觉质量。整个操作流程如图 3 所示。

具体而言，输入的弱光图像 $\boldsymbol{X}_{\text{low}}\in\mathbf{R}^{H\times W\times 3}$ 首先通过 WPT 从空域映射至小波域。随后，将高频子带 $\boldsymbol{V}_{\text{low}}^k,\boldsymbol{D}_{\text{low}}^k,\boldsymbol{H}_{\text{low}}^k$ 输入到高频细化模块中，获得增强后的高频分量 $\tilde{x}_0^h$。整个反向去噪采样过程从 $\boldsymbol{x}_T^l\in\mathbf{R}^{\frac{H}{4}\times\frac{W}{4}\times 3}\sim\mathcal{N}(0,\boldsymbol{I})$ 开始，将 $\tilde{\boldsymbol{x}}_0^h$ 作为去噪过程的条件引导，通过噪声估计网络 $\epsilon_\theta(\boldsymbol{x}_t^l,\tilde{\boldsymbol{x}}_0^h,t)$ 估计当前低频子带 x_t^l 中的噪声成分。随后将噪声从 $\boldsymbol{x}_t^l$ 逐步去除，以生成高质量的增强图像。该条件扩散采样过程表示为

$$\boldsymbol{x}_{t-1}^l=\sqrt{\bar{\alpha}_{t-1}}\left(\frac{\boldsymbol{x}_t^l-\sqrt{1-\bar{\alpha}_t}\cdot\epsilon_\theta(\boldsymbol{x}_t^l,\tilde{\boldsymbol{x}}_0^h,t)}{\sqrt{\bar{\alpha}_t}}\right)+\sqrt{1-\bar{\alpha}_{t-1}}\cdot\epsilon_\theta(\boldsymbol{x}_t^l,\tilde{\boldsymbol{x}}_0^h,t) \tag{5}$$

式中，$\epsilon_\theta(\boldsymbol{x}_t^l,\tilde{\boldsymbol{x}}_0^h,t)$ 为噪声估计网络，其主体结构为 U-Net 网络。为此，本文提出了一种高效的截断采样策略，即在采样过程给定步数 P 以完成去噪过程。具体而言，当采样步数达到 P 时，模型不再继续执行剩余的采样步骤，而是基于当前低频子带 $\boldsymbol{x}_P^l$ 中所包含的全局结构与光照特征，直接预测最终的重建结果 $\boldsymbol{x}_0^l$。该策略突破了传统扩散模型严格依赖逐步迭代的采样范式，在保证重建质量的同时减少推理步骤与计算成本。本文将式(5)优化为

$$\begin{cases}\boldsymbol{x}_{t-1}^l=\sqrt{\bar{\alpha}_{t-1}}\left(\dfrac{\boldsymbol{x}_t^l-\sqrt{1-\bar{\alpha}_t}\cdot\epsilon_\theta\left(\boldsymbol{x}_t^l,\tilde{\boldsymbol{x}}_0^h,t\right)}{\sqrt{\bar{\alpha}_t}}\right)+\sqrt{1-\bar{\alpha}_{t-1}}\cdot\epsilon_\theta\left(\boldsymbol{x}_P^l,\tilde{\boldsymbol{x}}_0^h,P\right),t=T,\cdots,P\\ \hat{\boldsymbol{x}}_0^l=\dfrac{\boldsymbol{x}_P^l-\sqrt{1-\bar{\alpha}_P}\cdot\epsilon_\theta\left(\boldsymbol{x}_P^l,\tilde{\boldsymbol{x}}_0^h,P\right)}{\sqrt{\bar{\alpha}_P}}\end{cases} \tag{6}$$

实验表明，本文所提出的高效条件采样策略可将扩散模型的采样步数由传统的 20 步显著减少至 5 步，极大简化了模型的推理流程，能够保持生成图像的清晰度与保真度。

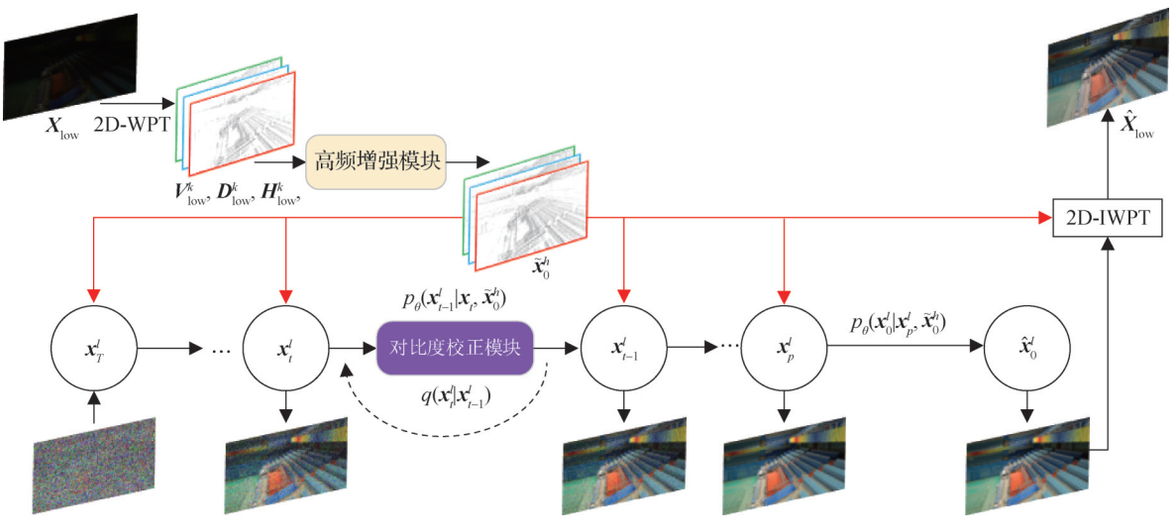


图3　反向扩散过程的截断采样策略

Fig. 3　The truncated sampling strategy in the reverse diffusion process

1.4　对比度校正模块

此外，反向扩散过程中预测的噪声与真实噪声存在偏差，该误差在多步迭代中会逐步积累并放大，进而导致生成图像出现色偏及对比度不足等问题。为此，本文设计了对比度校正模块（contrast-aware correction module，CCM），采用上一步的色彩和对比度信息调制当前步特征，实现对比度校正。该模块的具体结构如图4所示。

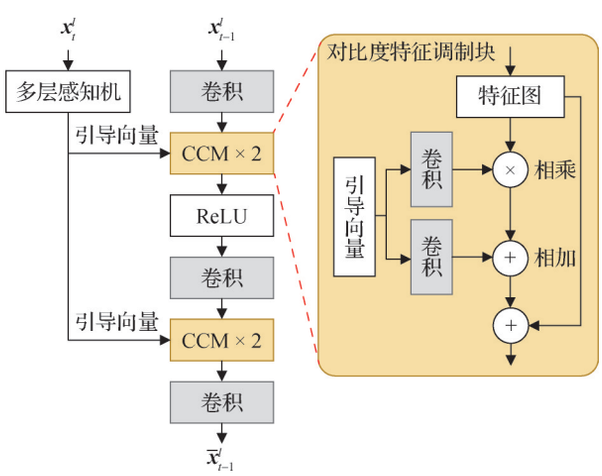


图4　对比度校正模块的结构示意图

Fig. 4　The structure of the contrast-aware correction module

在反向去噪阶段，对比度校正模块首先通过多层感知机（multilayer perceptron，MLP）独立处理上一步的全局特征$\boldsymbol{x}^l_t$，以提取颜色偏移与亮度等全局信息，将其编码为引导向量。同时，通过卷积操作从当前特征$\boldsymbol{x}^l_{t-1}$中提取包含边缘和纹理的局部特征图。随后，该引导向量和局部特征图输入对比度特征调制块，以实现对比度与细节的自适应增强。具体而言，CCM通过逐点调制、加权融合与残差补偿操作对当前特征进行调制，并通过残差连接机制保留原始结构信息。随后在融合特征上施加修正线性单元（rectified linear unit，ReLU），以增强非线性表达能力。最终，对比度校正模块输出$\bar{\boldsymbol{x}}^l_{t-1}$作为下一步迭代的初始特征。该模块在反向去噪的过程中显式地引入上一步的颜色与对比度信息，从而提升扩散模型在色彩保真度与视觉对比度，且几乎不会引入额外的计算负担。

1.5　损失函数

扩散模型的训练本质是对多步马尔可夫链进行逐步优化，每一步都需要计算损失函数。为兼顾模型训练效率与增强图像的整体质量，本文设计了组合损失函数，用于联合优化小波域中的低频扩散建模与高频子带增强过程。该总损失函数定义为

$$L_{total} = L_{noise} + L_{high} + L_{col} + L_{ce} \tag{7}$$

式中，L_{noise}约束噪声估计网络的训练，具体为

$$L_{noise} = E_{x^l_0, t, \epsilon_t \sim \mathcal{N}(0,I)} \left\| \epsilon_t - \epsilon_\theta(\tilde{\boldsymbol{x}}^h_0, \boldsymbol{x}^l_t, t) \right\|^2 \tag{8}$$

式中，$\boldsymbol{x}^l_t$为t步的图像的低频子带。L_{high}用于约束高频细化模块的训练，具体为

$$L_{high} = \left\| \tilde{\boldsymbol{x}}^h_0 - \boldsymbol{x}^h_0 \right\|_1 \tag{9}$$

式中，$\|\cdot\|_1$表示计算L1-范数。$\boldsymbol{x}^h_0 = \left(\boldsymbol{V}^k_{low}, \boldsymbol{D}^k_{low}, \boldsymbol{H}^k_{low}\right)$为正常光图像经过WPT的高频子带，$\tilde{\boldsymbol{x}}^h_0 = \left(\tilde{V}^k_{low}, \tilde{D}^k_{low}, \tilde{H}^k_{low}\right)$为高频细化模块增强的高频子带。

此外，基于自然图像统计的 Gray-World 颜色恒定性假设，本文采用了无监督的颜色损失 L_{col}，以确保增强图像的颜色自然，从而提高整体感知质量。L_{col} 计算为

$$L_{\mathrm{col}}=\sum_{c\in\{R,G,B\}}\left\|\mu\left(\hat{\boldsymbol{X}}_{\mathrm{low}}^{c}\right)-0.5\right\|_{2}^{2} \tag{10}$$

式中，$\mu\left(\hat{\boldsymbol{X}}_{\mathrm{low}}^{c}\right)$ 表示增强图像 $\hat{\boldsymbol{X}}_{\mathrm{low}}$ 颜色通道 c 的平均强度值。

本文引入对比度增强损失 L_{ce} 对增强图像中的对比度进行约束，以提高细节和纹理的可见性。该损失计算为

$$L_{\mathrm{ce}}=\left\|V\left(\hat{\boldsymbol{X}}_{\mathrm{low}}\right)-S\left(\hat{\boldsymbol{X}}_{\mathrm{low}}\right)\right\|_{2}^{2} \tag{11}$$

式中，$V\left(\hat{\boldsymbol{X}}_{\mathrm{low}}\right)$ 表示增强图像 $\hat{\boldsymbol{X}}_{\mathrm{low}}$ 的亮度，$S\left(\hat{\boldsymbol{X}}_{\mathrm{low}}\right)$ 表示增强图像 $\hat{\boldsymbol{X}}_{\mathrm{low}}$ 的饱和度。

2 实验与结果分析

2.1 实验环境与数据集

本文使用深度学习框架为 PyTorch，在单个 NVIDIA RTX 3090 GPU 上完成训练，并使用 Adam 优化器进行损失函数最小化，固定学习率为 4×10^{-4}。小波变换尺度为 2，时间嵌入使用正弦位置编码，以及权重使用 0.999 的指数移动平均策略。训练阶段和推理阶段步长均为 5，批处理大小为 12。

为了全面评估本文方法在 LLIE 任务上的性能，在配对数据集 LOLv1（low-light image enhancement dataset version1）、LOLv2-Real 以及 VE-LOL（vision enhancement in low-light conditions）数据集上进行训练，并在该测试集和非配对数据集上进行性能测试。LOLv1 数据集包含 485 对弱光/正常光图像用于训练和 15 对用于测试的图像。LOLv2-Real 数据集中 689 对图像用于训练，100 对图像用于测试。VE-LOL 包括 2 500 对低光照/正常光照的图像，场景更加多样化。

2.2 评价指标

为了更客观地评估不同算法的图像增强效果，本文使用有参考图像指标和无参考图像指标对实验结果进行对比分析。有参考的图像评价指标包括：峰值信噪比（peak signal-to-noise ratio，PSNR）、结构相似性（structural similarity，SSIM）和学习感知图像块相似度（learned perceptual image patch similarity，LPIPS）；无参考图像评价指标包括：自然图像质量评估（naturalness image quality evaluator，NIQE）（Mittal 等，2013）和彩色图像质量评价（color image quality evaluation，CIQE）。NIQE 用于衡量图像的自然程度；CIQE 通过色度、饱和度和对比度的线性组合来量化图像质量。

2.3 对比实验

2.3.1 定量比较

为了定量比较各方法的增强效果，本文对比了 RUAS（retinex-inspired unrolling with architecture search）（Liu 等，2021）、SCI（self-calibrated illumination）（Ma 等，2022）、PairLIE（learning a simple low-light image enhancer from paired low-light instances）（Fu 等，2023）、LLFlow（a flow-based low-light image enhancement）（Wang 等，2022）、WCDM（a conditional diffusion network based on wavelet transformation）（Jiang 等，2023）、MIRNet（Sumathi 等，2024）、Retinexformer（Cai 等，2023）、SNRNet（SNR-aware low-light image enhancement）（Xu 等，2022）、HVI（a new color space for low-light image enhancement）（Yan 等，2025）和 URetinex-Net++（Wu 等，2025）方法，分别计算了 3 个数据集上的 PSNR、SSIM、LPIPS 和 NIQE，量化指标的对比结果如表 1 所示。

从表 1 可以看出，本文方法在 PSNR、LPIPS 和 NIQE 评价指标上均取得最优表现，整体性能优于当前大多数主流 LLIE 方法。在 LOLv1 数据集上，本文方法增强结果 PSNR 指标为 25.81 dB，比次优的 LLFlow 提升了 1.65%；LPIPS 指标为 0.101，比次优的 URetinex-Net++ 降低了 4.72%；NIQE 指标为 4.480，比最优的 URetinex-Net++仅高 0.026。SSIM 指标没有表现出最优的效果，这主要归因于本文方法在设计上更强调全局结构一致性与感知质量优化。

在 LOLv2-Real 数据集上，本文同样展现出优越性能，PSNR 和 SSIM 指标分别为 26.68 dB、0.910，比次优的 LLFlow 分别提高了 8.46%、1.22%。LPIPS 指标达到了 0.097，比次优的 HVI 降低了 10.19%；NIQE 指标达到 4.443，比次优的 URetinex-Net++降低了 0.38%。此外，在包含合成图像与更多场景的 VE-LOL 数据集上，本文方法相较于次优的 PyDiff 方法，PSNR 和 SSIM 指标分别提高了 8.26%、1.93%。

表1　在LOLv1、LOLv2-Real以及VE-LOL数据集上的定量评价

Table 1　Quantitative evaluation on the LOLv1, LOLv2-Real and VE-LOL datasets

方法	LOLv1				LOLv2-Real				VE-LOL			
	PSNR/dB	SSIM	LPIPS	NIQE	PSNR/dB	SSIM	LPIPS	NIQE	PSNR/dB	SSIM	LPIPS	NIQE
RUAS(Liu等,2021)	16.40	0.701	0.270	4.528	15.14	0.614	0.322	4.472	15.11	0.609	0.401	4.472
SCI(Ma等,2022)	14.17	0.600	0.346	4.554	15.75	0.597	0.315	4.460	15.05	0.592	0.353	4.460
PairLIE(Fu等,2023)	18.47	0.807	0.243	4.491	19.44	0.791	0.302	4.484	19.10	0.795	0.324	4.508
LLFlow(Wang等,2022)	25.39	**0.926**	0.118	4.513	24.60	0.899	0.143	4.507	24.17	**0.899**	0.143	4.512
WCDM(Jiang等,2023)	21.90	0.849	0.213	4.541	17.95	0.769	0.293	4.561	16.45	0.760	0.293	4.571
PyDiff(Zhou等,2023)	23.18	0.898	0.107	4.499	21.76	0.843	0.233	4.481	24.22	0.882	0.162	4.502
WF-Diff(Zhao等,2024)	21.52	0.877	0.129	4.505	18.45	0.793	0.355	4.731	18.21	0.706	0.412	4.789
MIRNet(Sumathi等,2024)	24.14	0.893	0.134	4.510	24.53	0.887	0.157	4.504	23.92	0.881	0.203	4.516
Retinexformer(Cai等,2023)	25.15	0.895	0.131	4.510	22.79	0.859	0.171	4.495	21.47	0.721	0.342	4.376
SNRNet(Xu等,2022)	24.61	0.897	0.151	4.522	21.48	0.866	0.157	4.502	22.93	0.767	0.323	4.458
HVI(Yan等,2025)	23.81	0.885	0.186	4.490	23.23	0.882	0.108	4.482	20.27	0.825	0.160	4.479
URetinex-Net++(Wu等,2025)	23.83	0.885	0.106	**4.454**	20.72	0.859	0.127	4.460	21.59	0.775	0.307	4.460
本文	**25.81**	0.894	**0.101**	4.480	**26.68**	**0.910**	**0.097**	**4.443**	**26.22**	**0.899**	**0.124**	**4.367**

注:加粗、下划线字体分别表示各列最优、次优结果。

LPIPS和NIQE指标相较于次优方法分别降低了13.29%、0.21%。由于VE-LOL数据集加入了合成数据,其他方法增强质量均有下降,而本文方法质量反而有所提升,证明本文方法具有良好的泛化性与鲁棒性。值得注意的是,本文方法在所有测试数据集上均显著优于现有的扩散模型增强方法WCDM,这表明所提出的频域对角线特征优化与截断采样的弱光图像扩散增强方法不仅提升了增强图像的质量与细节还原的能力,还具备较强的应用前景。

为了更全面地评估本文方法在跨域场景中的泛化能力与鲁棒性,本文在多个未参与训练的非配对弱光图像数据集上进行了测试。具体而言,使用LOLv2-Real数据集训练得到的模型权重,在4个经典非配对数据集LIME、DICM、NPE和MEF上进行测试,使用无参考图像质量评估指标NIQE进行评估,其测试结果如表2所示。可以看出,本文方法在3个数据集上均获得了最优的NIQE得分,较其他方法的平均值分别下降了0.82(LIME)、0.39(MEF)和0.86(NPE),显著优于当前主流方法。在DICM数据集上,本文方法比最优方法EnlightenGAN的NIQE指标仅高0.02。这表明本文方法在跨域数据集测试中依然展现出良好的稳定性与泛化能力,能够在非配对目标域中生成具有自然视觉效果的增强图像。

表2　非配对弱光数据集上不同方法的NIQE结果

Table 2　NIQE results of different methods on unpaired low-light datasets

方法	LIME	DICM	MEF	NPE
RUAS(Liu等,2021)	5.39	4.43	5.45	7.09
SCI(Ma等,2022)	4.68	4.13	4.60	4.52
RetinexNet(Wei等,2018)	6.53	4.76	4.69	4.69
EnlightenGAN(Jiang等,2021)	4.54	**4.06**	4.95	4.45
Diff-retinex(Yi等,2023)	5.13	4.34	4.42	5.19
WCDM(Jiang等,2023)	4.55	4.97	4.71	5.01
本文	**4.32**	4.08	**4.41**	**4.30**

注:加粗字体表示各列最优结果。

2.3.2　定性比较

如图5所示,本文对比了多种典型方法在LOLv1、LOLv2-Real以及VE-LOL数据集上的增强效果,从主观视觉质量层面对不同方法的性能进行了评估。图中第1—3行分别是LOLv1、LOLv2-Real以及VE-LOL数据集上可视化结果。可以观察到,SCI

与RUAS方法在色彩真实感与光照一致性方面表现较差，例如SCI方法整体亮度提升较低，RUAS在第3行对于“灯具”区域的恢复出现过曝光现象。其原因是上述方法未能充分建模图像的色彩分布，导致增强图像普遍存在颜色失真、曝光过度问题。URetinex-Net++在光照分量估计中未能准确捕捉边缘结构信息，造成边缘区域出现明显的伪影与光晕效应，影响图像的整体感知质量。此外，基于流场建模的LLFlow和基于扩散建模的WCDM方法在一定程度上恢复了图像结构，但对局部细节区域的建模能力有限，导致增强图像细节丢失严重，整体画面显得模糊。例如WCDM在第1行中对于“衣领的红色边缘”和第2行“座椅靠背”区域的细节无法恢复。本文方法得益于反向扩散过程中对比度增强模块与高频细节增强信息的协同机制，不仅保持了图像整体结构清晰，还抑制了色偏与边缘伪影问题，使得增强结果在视觉观感上更真实自然。无论在复杂照明条件的真实场景，还是具有合成特征的低照度图像，本文方法均能生成具有高保真度与感知一致性的图像。

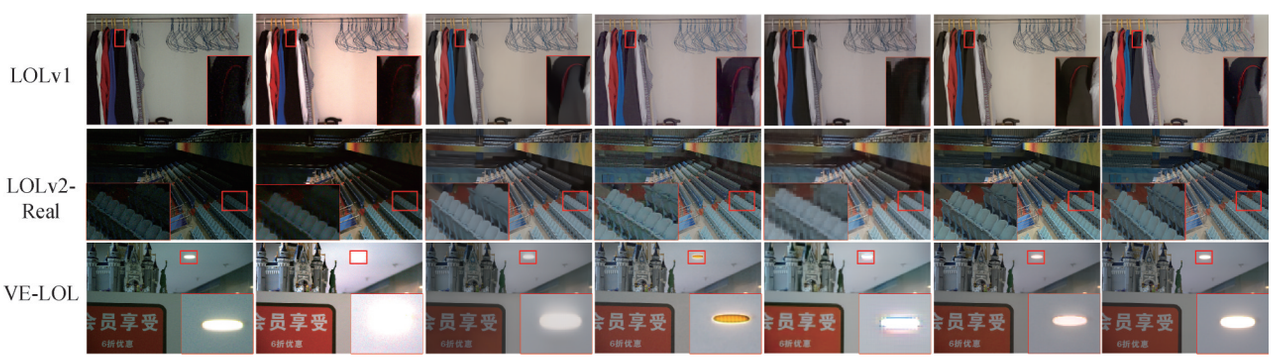


图5　在LOLv1、LOLv2-Real以及VE-LOL数据集上的可视化结果比较

Fig. 5　Visualization comparison of results on the LOLv1、LOLv2-Real, and VE-LOL datasets

((a) SCI; (b) RUAS; (c) LLFlow; (d) URetinex-Net++; (e) WCDM; (f)ours; (g) reference images)

本文进一步在4个经典非配对弱光图像数据集上，使用LOLv2-Real数据集上训练得到的模型权重进行测试，部分增强结果如图6所示。其中，第1列为输入图像，第2—5列为WCDM、DarkIR(Feijoo等，2025)、EnlightenGAN以及本文方法的增强结果。从可视化结果可以观察到，基于扩散的WCDM方法在提升亮度和细节恢复方面都有所欠缺，增强图像仍存在暗区域残留与纹理信息缺失的问题；而基于频域建模的DarkIR方法虽然具备一定的全局光照调节能力，但在多个样本中出现了明显的过曝光和细节缺失问题。EnlightenGAN方法在亮度和细节恢复上有所提升，但同时引入了大量的噪声。例如，MEF数据集上的增强图像“地面”区域存在较明显的噪声和模糊。相比之下，本文方法在未使用目标域监督的情况下，仍展现出较优的跨域增强性能、首先，得益于频域建模与高频细化机制的协同设计，本文方法能够有效抑制噪声干扰，显著提升纹理细节。其次，在反向扩散过程中引入的对比度校正模块，通过显式建模颜色偏移与亮度趋势，不仅增强了图像的对比度，还保持了自然真实的色彩视觉体验。然而，对于极暗场景，频域能量分布过于微弱，仍可能导致特征提取效果受限，从而影响增强质量。

2.3.3　模型效率对比分析

为了探究在频域实施扩散能否有效降低模型的计算复杂度并提升推理速度，本文在保持扩散模型参数不变的前提下，与基于空域扩散LLIE方法进行了对比分析，其对比结果如表3所示。在推理时间方面，本文将每幅图像的处理时间从PyDiff的7.617 s降至0.473 s，推理速度提升约16倍。其根本原因在于本文将扩散模型从空域迁移至小波低频子空间中进行，并结合高效的条件采样策略，显著减少了扩散迭代的计算量，简化了模型的推理流程。在显存占用方面，本文方法相较于PyDiff减少了48.5%的显存消耗，表明在硬件资源受限的设备上具有更好的适配性和更高的运行效率。在参数量方面，本文方法相比于WCDM减少了36.5%。相比于空域扩散方法，本文模型在保持相同扩散框架能力的前提下，通过对低频子空间的高效建模，将网络的

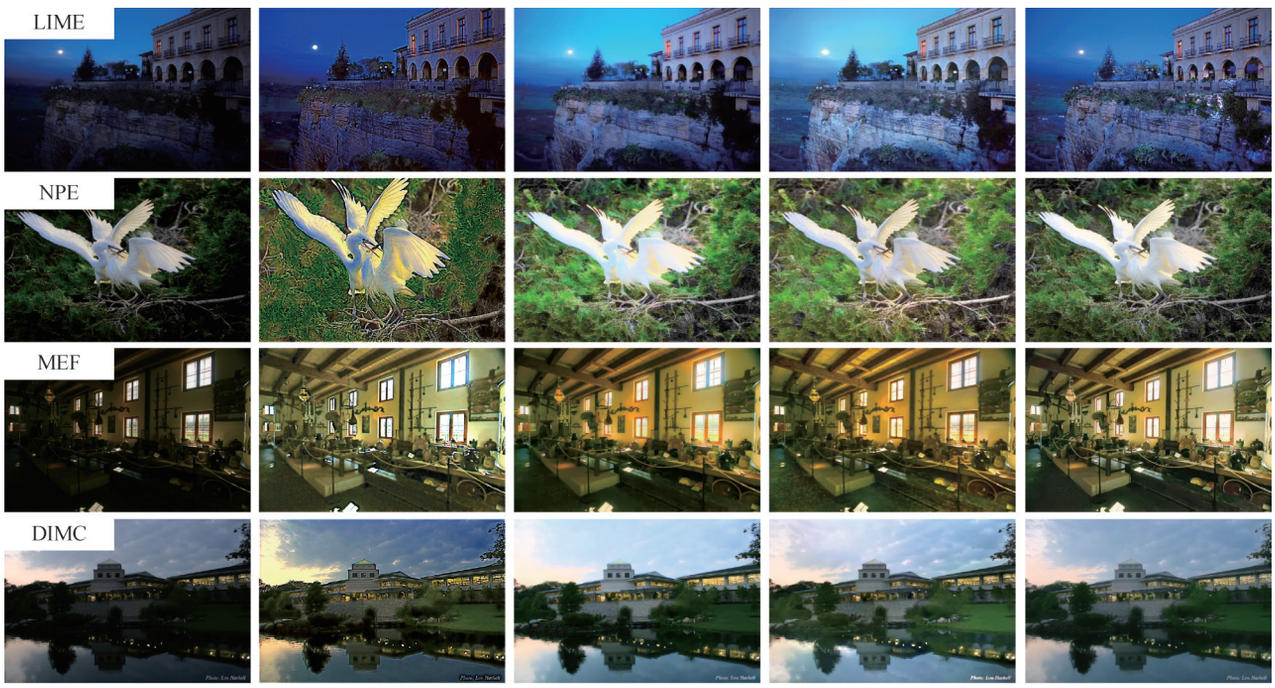


图6　在经典非配对弱光数据集上增强结果的可视化比较

Fig. 6　Visual comparison of enhancement results on classic unpaired low-light datasets

((a) input images; (b) WCDM; (c) DarkIR; (d) EnlightenGAN; (e) ours)

卷积通道数与特征映射规模进一步压缩，使得整体参数量显著降低。另一方面，模型无需引入大量的结构先验或深层卷积堆叠即可捕获全局结构与局部纹理信息，从根本上减少了内存占用的峰值。综上，本文能够在保持增强质量的同时，在资源受限场景中具有更高应用潜力。

表3　本文与空间域扩散LLIE方法的推理效率和模型复杂度比较

Table 3　Comparison of inference efficiency and model complexity with spatial-domain diffusion LLIE methods

方法	推理时间/s	显存占用/G	参数量/M
DDIM(Song等，2020)	12.138	5.881	115.78
PyDiff(Zhou等，2023)	7.617	3.594	97.89
WF-Diff(Zhao等，2024)	13.489	6.351	56.88
WCDM(Jiang等，2023)	15.123	5.837	48.32
本文	**0.473**	**1.850**	**30.68**

注：加粗字体表示各列最优结果。

2.4　消融实验

2.4.1　核心组件的消融分析

为了深入分析各组件对本文方法性能的影响，本文分别移除高频细化模块、截断采样策略和对比度校正模块，实验结果如表4所示。首先，将WPT替换为DWT后，推理时间增加了0.421 s，表明WPT在保持频域表达能力的同时，显著提升了推理效率。其次，去除高频细化模块后，PSNR和SSIM指标分别下降了1.70 dB和0.042，表明高频细化模块在补偿高频细节退化方面发挥着关键作用。当采用DDPM的完整采样替代截断采样机制后，推理时间显著增加至11.32 s，表明该机制能够显著减少推理时间。此外，去除对比度校正模块后，PSNR和SSIM指标分别下降了1.12 dB和0.020，表明该模块能够进一步提升恢复质量。各组件的协同作用使得本文方法在LLIE实现了高效且高质量的图像恢复。

2.4.2　高频细化模块的有效性实验

为了验证高频细化模块对恢复边缘、纹理等细节信息更为有效，本文对比了该方法使用高频细化模块和不使用高频细化模块的增强结果，可视化结果见图7。在未添加高频细化模块时(图7(b))，乒乓球桌面边缘、桌脚边缘及桌上LOGO均模糊不清，视觉效果极差。相比之下，添加高频细化模块的可视化结果(图7(c))结构清晰度大幅提升，尤其在乒乓球桌边缘区域、桌面反射和LOGO标识等细节恢

表 4　各组件对本文方法增强性能影响分析

Table 4　Analysis of the impact of each module on the performance enhancement of our method

方法	PSNR /dB	SSIM	参数量/M	推理时间/s
w/o WPT(DWT)	25.07	0.891	30.68	0.894
w/o 高频细化模块	24.11	0.852	**26.08**	0.590
w/o 截断采样机制	24.93	0.890	30.68	11.320
w/o 对比度校正模块	24.69	0.874	28.23	**0.436**
本文	**25.81**	**0.894**	30.68	0.473

注：加粗字体表示各列最优结果。“w/o”表示移除该组件。

复效果提升显著。从定量指标来看，添加高频细化模块后，增强图像的 PSNR 从 24.11 dB 提升至 26.35 dB，提升了 2.24 dB，而 SSIM 也从 0.874 提升至 0.917，提升了 0.043，说明图像在像素级重建与结构一致性方面均获得提升。这得益于高频细化模块通过引入水平、垂直方向与对角线方向的特征交互机制，构建跨方向的注意力引导，增强了对角线方向的细节恢复能力，从而实现了对局部纹理和边缘结构的精准重建。值得注意的是，该模块的参数量较少，仅增加不到 15% 的模型参数量，即可带来明显的性能增益。



(a) 弱光图像　(b) 不使用高频增强模块

(c) 使用高频增强模块　(d) 参考图像

图 7　高频细化模块对 LLIE 的影响

Fig. 7　Effect of high frequency refinement module on LLIE ((a) low-light image; (b) without HFRM; (c) with HFRM; (d) reference image)

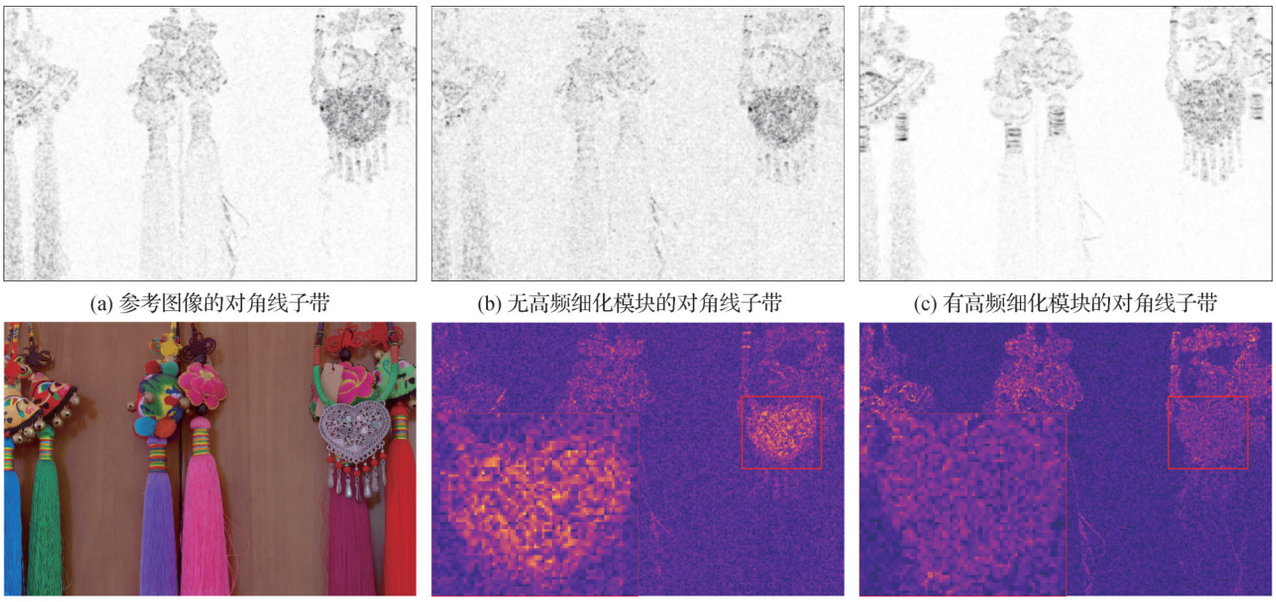
图 8 展示了高频细化模块增强对角线子带的有效性，其中图 8(e)和图 8(f)越亮的区域表示误差越大。未使用高频细化模块时，对角线子带结构模糊且噪声较强，其增强结果与参考图像的对角线子带存在明显的差异(图 8(e))。引入高频细化模块后，利用水平与垂直子带中的方向梯度信息，有效补全了对角线子带缺失的纹理细节，使高频边缘更加连续，并显著降低对角线子带误差(图 8(f))。这表明对角线方向的特征对细节恢复至关重要，而高频细化模块正是利用这种跨方向相关性实现了对角线子带的增强，从而提升了图像的细节与纹理表现。

2.5　对比度校正模块的泛化性实验

为进一步验证对比度校正模块在反向去噪过程中对隐空间特征的色彩与对比度修正效果，并评估其在扩散框架中的泛化性，本文将该模块集成至 PyDiff 方法的反向扩散过程，实验结果如图 9 所示。从图 9(a)可以看出，PyDiff 增强图像的整体对比度明显不足，细节区域缺乏层次感，色彩表现也较为暗淡。由于传统反向扩散过程主要依赖像素级重构损失，使生成图像的像素值趋于均值，亮度范围压缩，图像灰度分布集中，缺乏对暗部与亮部的有效区分。此外，模型可能因训练数据中某一主色通道占比过高而出现色彩偏向，导致生成图像的色调不自然。图 9(b)引入对比度校正模块后，增强图像不仅在整体亮度分布上更均衡，同时色彩更加自然，明显提升了图像的视觉质量。CIQE 指标从 0.68 提升至 0.82，提升 20.6%。核心原因在于，对比度校正模块能够利用前一扩散步骤的亮度与色彩信息，对当前隐空间特征进行动态调控，从而抑制色彩偏移。实验结果表明，对比度校正模块具备良好的跨框架泛化能力，为扩散模型在 LLIE 任务中提供了一种通用且轻量的方案。

2.6　截断采样策略的高效性实验

为验证截断采样策略在降低采样开销方面的有效性，本文在 LOLv2-Real 和 VE-LOL 数据集上进行了定量分析，并采用 PSNR 作为评价指标，对比结果如图 10 所示。本文与 DDIM 的采样过程均采用等间隔迭代。“★”表示本文在采样过程中所达到的最优 PSNR 值，横轴 t 表示反向扩散采样步。从图 10 可以观察到，DDIM 的整体曲线始终低于本文方法，且随着采样步数增加，其性能提升趋于缓慢。尤其 t = 1 000～800 之间时，本文的性能提升明显优于 DDIM。当 t = 600 时，本文性能趋于稳定，而 DDIM 则需要完整执行 1 000 步迭代才能达到相对稳定的性能。这得益于本文在低秩的频域进行扩散建模，



(a) 参考图像的对角线子带　(b) 无高频细化模块的对角线子带　(c) 有高频细化模块的对角线子带

(d) 本文方法的增强图像　(e) 无高频细化模块的差异图　(f) 有高频细化模块的差异图

图8　高频细化模块增强对角线子带的有效性

Fig. 8　Effectiveness of diagonal features enhanced by HFRM ((a) diagonal features of the reference image; (b) diagonal features without HFRM; (c) diagonal features with HFRM; (d) the enhanced result of our method; (e) diagonal feature difference map without HFRM; (f) diagonal feature difference map with HFRM)



(a) PyDiff　(b) PyDiff + 对比度校正模块　(c) 参考图像

图9　PyDiff + 对比度校正模块的影响

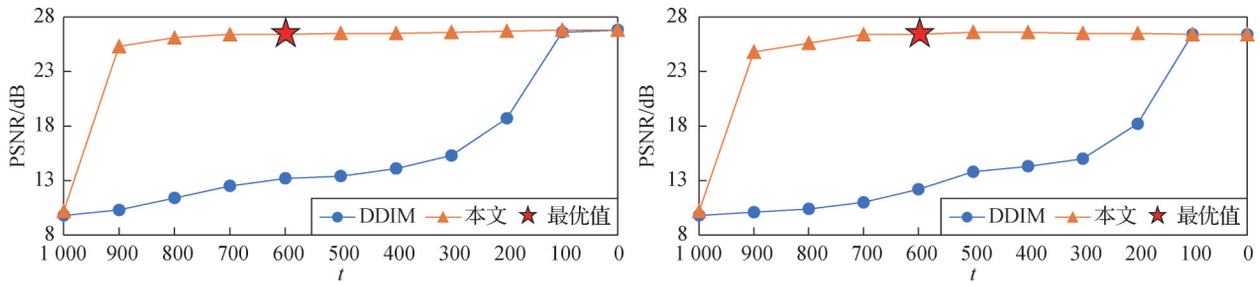
Fig. 9　Effect of the contrast-aware correction module in PyDiff

((a) PyDiff; (b) PyDiff + contrast-aware correction module; (c) reference images)

使得反向扩散过程仅需较少的采样步数，从而显著降低了采样开销。

图11展示了本文与DDIM在不同总采样点P设置下的性能对比，P可以理解为截断采样的时间划分点数量。在实验设置中，两种方法的总步数均为1 000步，在此基础上采用相同的截断采样策略。例如，当总步数为1 000时，仅在5个等间隔的时间点进行采样($P = 5$)。从图11可以观察到，DDIM的图像质量随总采样点P的增加逐步提升，并在25左右趋于收敛。相比之下，本文仅需5个采样节点到达收敛且优于DDIM的增强性能。

此外，表5展示了$P = 4$和$P = 5$的采样轨迹，并用4个评估指标进行了量化评价。$P = 5$时，本文在性能和计算成本之间达到了平衡。因此，本文有效

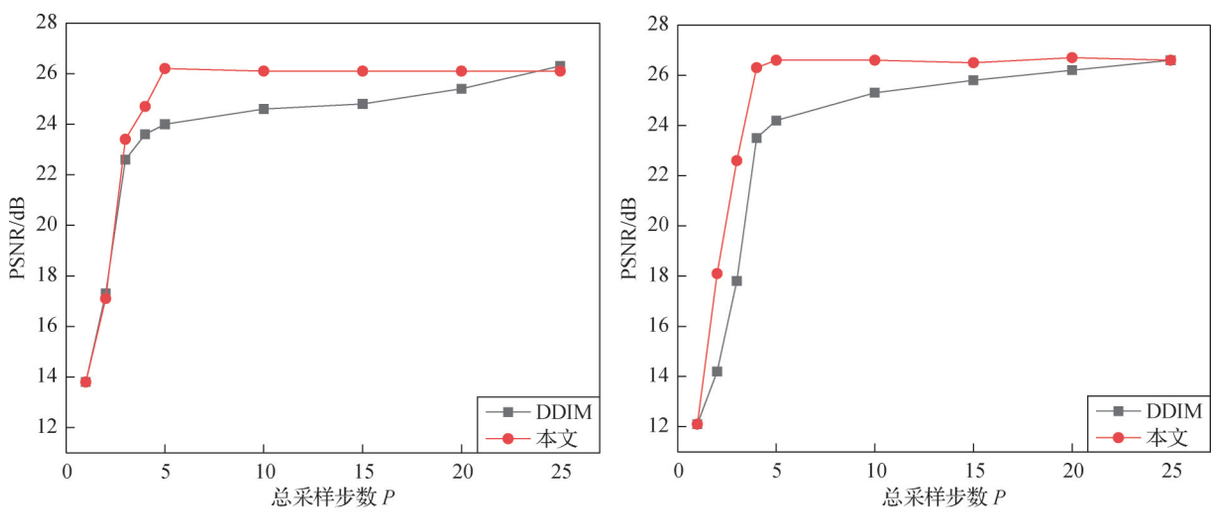


(a) LOLv 2-Real 数据集的采样过程　　(b) VE-LOL 数据集的采样过程

图 10　本文与 DDIM 采样过程的收敛速度比较

Fig. 10　Comparison of convergence speed between our method and DDIM

((a) sampling process of the LOLv2-Real dataset; (b) sampling process of the VE-LOL dataset)



(a) LOLv1 数据集上不同总采样步数的 PSNR　　(b) LOLv2-Real 数据集上不同总采样步数的 PSNR

图 11　本文与 DDIM 不同总采样点 P 的性能比较

Fig. 11　Performance comparison between our method and DDIM under different total sampling steps P ((a) PSNR with different total sampling steps on the LOLv1 dataset; (b) PSNR with different total sampling steps on the LOLv2-Real dataset)

表 5　本文与 DDIM 在 $P = 4$ 和 $P = 5$ 设置下的性能比较

Table 5　Performance comparison on $P = 4$ and $P = 5$ settings between our method and DDIM

方法	P	采样轨迹	PSNR/dB	SSIM	LPIPS	NIQE
DDIM	4	1 000→750→500→250→0	22.98	0.817	0.166	4.621
本文		1 000→870→730→600→0	25.50	0.876	0.123	4.502
DDIM	5	1 000→800→600→400→200→0	23.56	0.824	0.151	4.553
本文		1 000→900→800→700→600→0	**25.81**	**0.894**	**0.101**	**4.480**

注：加粗字体表示各列最优结果。

简化了扩散模型的反向采样过程，避免了大量冗余迭代步骤带来的计算开销。值得强调的是，图 10 和图 11 的实验具有互补性。图 10 侧重于展示模型在完整采样过程中的图像质量演化趋势；图 11 则侧重于对比不同总采样点 P 的性能表现。两组实验分别从“收敛速度最优化”与“采样点最少化”两个维度，验证了本文在提升采样效率与增强质量方面的综合优势。

3　结　论

针对现有扩散模型在 LLIE 任务中存在的高计算成本、细节重建不足以及生成不稳定等问题，本文提出了一种频域特征优化与截断采样的扩散弱光图

像增强方法。该方法在小波低频子空间中进行扩散建模,在有效保留图像整体结构与光照信息的同时显著降低了计算复杂度。并且引入高频细化模块,融合水平方向与垂直方向的高频特征以补全对角线细节,增强图像的纹理表现与局部清晰度。为解决扩散过程中存在的色偏与采样效率低的问题,进一步设计了高效截断采样策略与对比度校正模块,有效提升了采样效率、生成结果的稳定性与色彩保真度。在多个标准配对与非配对弱光图像数据集上的定量与定性实验结果均表明,本文方法不仅在生成质量、采样速度与资源消耗方面表现优异,同时具备良好的跨域泛化能力与鲁棒性。未来的研究将针对极暗场景下频域信息不足的问题,引入更具鲁棒性的多尺度融合策略与自适应频率引导机制,以增强模型在复杂光照条件下的泛化性能,推动其在实际智能视觉系统中的落地应用。

参考文献(References)

Bi S F, Ye L, Wang Z X, Zhang Z Y, Hong H Y and Sang N. 2026. Research progress on object detection and tracking methods for nighttime UAV aerial imagery. Journal of Image and Graphics, 2026-01-06. https://doi.org/10.11834/jig.250459 (毕诗帆, 叶亮, 王志祥, 张梓阳, 洪汉玉, 桑农. 2026. 夜间无人机航拍图像目标检测与跟踪方法研究进展. 中国图象图形学报, 2026-01-06. https://doi.org/10.11834/jig.250459 [DOI: 10.11834/jig.250459]

Cai Y H, Bian H, Lin J, Wang H Q, Timofte R and Zhang Y L. 2023. Retinexformer: one-stage retinex-based transformer for low-light image enhancement//Proceedings of 2023 IEEE/CVF International Conference on Computer Vision. Paris, France: IEEE: 12470-12479 [DOI: 10.1109/ICCV51070.2023.01149]

Feijoo D, Benito J C, Garcia A and Conde M V. 2025. DarkIR: robust low-light image restoration//Proceedings of 2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, USA: IEEE: 10879-10889 [DOI: 10.1109/CVPR52734.2025.01016]

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

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

Huang Y, Peng H, Li C S, Gao S M and Chen F. 2024. LLFlowGAN: a low-light image enhancement method for constraining invertible flow in a generative adversarial manner. Journal of Image and Graphics, 29(1): 65-79 (黄颖, 彭慧, 李昌盛, 高胜美, 陈奉. 2024. LLFlowGAN: 以生成对抗方式约束可逆流的低照度图像增强. 中国图象图形学报, 29(1): 65-79) [DOI: 10.11834/jig.230063]

Jiang H, Luo A, Fan H Q, Han S C and Liu S C. 2023. Low-light image enhancement with wavelet-based diffusion models. ACM Transactions on Graphics, 42(6): #238 [DOI: 10.1145/3618373]

Jiang Y F, Gong X Y, Liu D, Cheng Y, Fang C, Shen X H, et al. 2021. EnlightenGAN: deep light enhancement without paired supervision. IEEE Transactions on Image Processing, 30: 2340-2349 [DOI: 10.1109/TIP.2021.3051462]

Lan G Z, Ma Q L, Yang Y Q, Wang Z G, Wang D, Li X L, et al. 2025. Efficient diffusion as low light enhancer//Proceedings of 2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, USA: IEEE: 21277-21286 [DOI: 10.1109/CVPR52734.2025.01982]

Li J L, Li B L, Tu Z Z, Liu X Y, Guo Q, Juefei-Xu F, et al. 2024. Light the night: a multi-condition diffusion framework for unpaired low-light enhancement in autonomous driving//Proceedings of 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE: 15205-15215 [DOI: 10.1109/CVPR52733.2024.01440]

Li Y Y, Wang H, Jin Q, Hu J, Chemerys P, Fu Y, et al. 2023. SnapFusion: text-to-image diffusion model on mobile devices within two seconds//Proceedings of the 37th International Conference on Neural Information Processing System. New Orleans, USA: Curran Associates Inc.: #907

Liang J X, Yang Y X, Li B Y, Duan P Q, Xu Y and Shi B X. 2023. Coherent event guided low-light video enhancement//Proceedings of 2023 IEEE/CVF International Conference on Computer Vision. Paris, France: IEEE: 10581-10591 [DOI: 10.1109/ICCV51070.2023.00974]

Liu R S, Ma L, Zhang J A, Fan X and Luo Z X. 2021. Retinex-inspired unrolling with cooperative prior architecture search for low-light image enhancement//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, USA: IEEE: 10556-10565 [DOI: 10.1109/CVPR46437.2021.01042]

Liu X C, Gong C Y and Liu Q. 2022. Flow straight and fast: learning to generate and transfer data with rectified flow//Proceedings of the 8th International Conference on Learning Representations. Kigali, Rwanda: OpenReview.net

Ma L, Ma T Y, Liu R S, Fan X and Luo Z X. 2022. Toward fast, flexible, and robust low-light image enhancement//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, USA: IEEE: 5627-5636 [DOI: 10.1109/CVPR52688.2022.00555]

Ma Y H, Liu J, Liu Y H, Fu H Z, Hu Y, Cheng J, et al. 2021. Struc-

ture and illumination constrained GAN for medical image enhancement. IEEE Transactions on Medical Imaging, 40(12): 3955-3967 [DOI: 10.1109/TMI.2021.3101937]

Mittal A, Soundararajan R and Bovik A C. 2013. Making a "completely blind" image quality analyzer. IEEE Signal Processing Letters, 20(3): 209-212 [DOI: 10.1109/LSP.2012.2227726]

Song J M, Meng C L and Ermon S. 2020. Denoising diffusion implicit models//Proceedings of the 9th International Conference on Learning Representations. [s.l.]: OpenReview.net: 1423-1426

Sumathi M, Jayasudha F V, Rajeswari D and Sanju M S. 2024. Performance study of enhancing low light images using deep vision networks//Proceedings of 2024 International Conference on IoT Based Control Networks and Intelligent Systems. Bengaluru, India: IEEE: 1664-1669 [DOI: 10.1109/ICICNIS64247.2024.10823375]

Wang T, Zhang K H, Zhang Y, Luo W H, Stenger B, Lu T, et al. 2025. LLDiffusion: learning degradation representations in diffusion models for low-light image enhancement. Pattern Recognition, 166: #111628 [DOI: 10.1016/j.patcog.2025.111628]

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

Wei C, Wang W J, Yang W H and Liu J Y. 2018. Deep retinex decomposition for low-light enhancement//Proceedings of 2018 British Machine Vision Conference. Newcastle, UK: BMVA Press: 1-14

Wu W H, Weng J, Zhang P P, Wang X, Yang W H and Jiang J M. 2025. Interpretable optimization-inspired unfolding network for low-light image enhancement. IEEE Transactions on Pattern Analysis and Machine Intelligence, 47(4): 2545-2562 [DOI: 10.1109/TPAMI.2024.3524538]

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

Yan Q S, Feng Y X, Zhang C, Pang G S, Shi K B, Wu P, et al. 2025. HVI: a new color space for low-light image enhancement//Proceedings of 2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, USA: IEEE: 5678-5687 [DOI: 10.1109/CVPR52734.2025.00533]

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

Yu Y, Chen W X and Chen F X. 2023. RIC-NVNet: night-time vehicle enhancement network for vehicle model recognition. Journal of Image and Graphics, 28(7): 2054-2067 (余烨, 陈维笑, 陈凤欣. 2023. 面向车型识别的夜间车辆图像增强网络RIC-NVNet. 中国图象图形学报, 28(7): 2054-2067) [DOI: 10.11834/jig.220122]

Zhang F, Shao Y J, Sun Y S, Gao C X and Sang N. 2024. Self-supervised low-light image enhancement via histogram equalization prior//Proceedings of the 6th Chinese Conference on Pattern Recognition and Computer Vision. Xiamen, China: Springer: 63-75 [DOI: 10.1007/978-981-99-8552-4_6]

Zhang L, Li Z Y, Cheng L, Zhang Q, Liu Z, Zhang X L, et al. 2026. DLIENet: a lightweight low-light image enhancement network via knowledge distillation. Pattern Recognition, 169: #111777 [DOI: 10.1016/j.patcog.2025.111777]

Zhao C, Cai W L, Dong C Y and Hu C W. 2024. Wavelet-based fourier information interaction with frequency diffusion adjustment for underwater image restoration//Proceedings of 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE: 8281-8291 [DOI: 10.1109/CVPR52733.2024.00791]

Zhou D W, Yang Z X and Yang Y. 2023. Pyramid diffusion models for low-light image enhancement//Proceedings of the 32nd International Joint Conference on Artificial Intelligence. Macao, China: International Joint Conference on Artificial Intelligence: #199 [DOI: 10.24963/ijcai.2023/199]

作者简介

李文改，女，博士研究生，主要研究方向为计算机视觉与图像处理。E-mail: 1241211015@stu.xaut.edu.cn

肖照林，通信作者，男，教授，博士生导师，主要研究方向为计算摄像和计算机视觉。E-mail: xiaozhaolin@xaut.edu.cn

桂柯，男，硕士研究生，主要研究方向为计算机视觉与图像处理。E-mail: 2241221139@stu.xaut.edu.cn

金海燕，女，教授，主要研究方向为计算机视觉、图像处理和智能信息处理。E-mail: jinhaiyan@xaut.edu.cn

苏浩楠，男，讲师，主要研究方向为计算机视觉和图像处理。E-mail: suhaonan@xaut.edu.cn