

基于即插即用分裂算法的盲图像修复扩散模型^①



侯晓伦, 藺素珍, 王彦博, 王志茹

(中北大学 计算机科学与技术学院, 太原 030051)

通信作者: 藺素珍, E-mail: lsz@nuc.edu.cn

摘要: 针对现行即插即用图像修复方法无法处理弱光图像增强等盲图像修复任务中图像退化模型难以精确建模的问题, 本文构造了一种将即插即用分裂算法与引导扩散模型相结合的求解方案. 该方案巧妙地避免了直接求解由复杂退化模型导致的复杂数据子问题, 转而使用真实退化图像来解决数据子问题, 并利用数据子问题的解作为“锚点”间接约束并优化先验子问题的求解过程. 从而确保了图像修复结果能够更紧密地贴近真实的图像恢复目标. 本方法在多个公开数据集上进行验证, 结果表明, 所提算法与现行具有代表性的方法相比, PSNR 和 SSIM 分别平均提升 4.89% 和 9.48%. 实验证明, 所提方法在修复指标上表现更好, 验证了本方法的有效性.

关键词: 图像修复; 即插即用; 扩散模型; 非线性修复; 深度学习

引用格式: 侯晓伦, 藺素珍, 王彦博, 王志茹. 基于即插即用分裂算法的盲图像修复扩散模型. 计算机系统应用, 2025, 34(3): 105-114. <http://www.c-s-a.org.cn/1003-3254/9787.html>

Blind Image Restoration Diffusion Model Based on Plug-and-play Splitting Algorithm

HOU Xiao-Lun, LIN Su-Zhen, WANG Yan-Bo, WANG Zhi-Ru

(School of Computer Science and Technology, North University of China, Taiyuan 030051, China)

Abstract: In response to the problem that current plug-and-play image restoration methods cannot accurately model image degradation models in blind image restoration tasks such as low-light image enhancement, this study constructs a solution that combines a plug-and-play splitting algorithm with a guided diffusion model. This solution cleverly avoids directly solving complex data sub-problems caused by complex degradation models. Instead, it uses real degraded images to solve data sub-problems and takes the solutions of data sub-problems as “anchor points” to indirectly constrain and optimize the solving process of prior sub-problems. This ensures that the image restoration results can be more closely approximated to the real image restoration target. This method is validated on multiple public datasets. The results show that the proposed algorithm achieves an average improvement of 4.89% in PSNR and 9.48% in SSIM compared to current representative methods. Experiments prove that the proposed method performs better in repair metrics, validating its effectiveness.

Key words: image restoration; plug-and-play; diffusion model; nonlinear repair; deep learning

由于受到噪声干扰、信号衰减等多重影响, 图像在采集、储存以及传输过程中质量降低是不可避免的现象. 通常, 一副降质图像 y 可以看作是理想图像 x 通过一个与噪声无关的退化模型 $T(\cdot)$ 并加上一个加性高斯白

噪声 n 所产生的^[1]. 而图像恢复的目的就是反转这种退化过程, 即求解由数据拟合项 $L(y, T(x))$ 和正则化项 $R(x)$ 所构成的优化问题, 以改善图像质量^[2].

作为一种不适定的逆问题, 传统的图像修复方法

① 基金项目: 国家自然科学基金 (62271453); 国家自然科学基金联合基金 (U21A20524); 山西省自然科学基金 (202303021211147); 山西省基础研究计划 (20210302123025)

收稿时间: 2024-08-06; 修改时间: 2024-09-19; 采用时间: 2024-09-30; csa 在线出版时间: 2025-01-17

CNKI 网络首发时间: 2025-01-17

通常依赖特定的退化模型和正则化项去约束求解过程,从而获得合理的修复结果^[3].例如在传统的机器学习算法中,通常会选择L1范数、L2范数、全变分等作为正则化项.但是正则化项的设计通常依赖于先验知识^[4],并且控制正则化项强度的参数也十分敏感,难以选择.

随着深度学习的发展,即插即用图像修复方法允许用户使用包含扩散模型在内的训练好的神经网络代替先验知识嵌入到先验项 $R(x)$ 中,从而使得修复过程无需额外训练^[5].对于退化模型 $T(\cdot)$ 较为简单的情况,例如线性模型,即插即用图像修复方法可以借助半二次分裂法(half quadratic splitting, HQS)^[6]等变量分离优化算法,将退化模型与正则化项分裂交替优化^[7],对于退化模型的优化称为数据子问题,对于正则化项的优化称为先验子问题.由于线性退化模型通常具有近似封闭数值解,因此上述方法能够有效处理多种线性修复任务.

但是对于不存在近似封闭数值解的盲图像修复任务等复杂图像恢复问题,由于其退化模型 $T(\cdot)$ 未知且很可能具有非线性和非凸性质,因此直接寻找全局最优解会变得异常艰难.这就导致现有的即插即用图像修复方法在处理非线性或盲任务^[8]时无法使用数值优化方法去约束修复结果的产生,从而生成质量较差或与原始图像毫不相关的修复结果.

为此,本文提出了一种基于即插即用分裂算法的盲图像修复扩散模型.具体来说,该方法可以分为以下3个主要内容:(1)为了将即插即用图像修复方法拓展到非线性或盲任务领域,以弱光图像增强任务为例,本文方法抛弃了现行的数值优化方法转而采用真实退化图像去解决数据子问题.(2)本文采用一种间接引导的方法,利用数据子问题的解作为“锚点”,以此来约束先验子问题解的生成过程,确保修复后的图像更接近真实的理想图像.(3)面对非线性及盲修复任务中退化机理不明确的挑战,本文采用了一种简洁高效的近似模型,该模型可以在修复过程中动态地进行同步优化,旨在更准确地模拟真实退化模型.实验结果表明,所提算法不仅可以解决现行即插即用图像修复方法无法处理非线性或盲任务的问题,还具有更优越的性能.

1 相关工作

即插即用图像修复算法的数学表达如式(1)所示:

$$\hat{x} = \arg \min_x \frac{1}{2\sigma^2} \|y - T(x)\|^2 + \lambda R(x) \quad (1)$$

其中, \hat{x} 表示修复后的图像, $\frac{1}{2\sigma^2} \|y - T(x)\|^2$ 表示数据项, σ 表示给定常量, $R(x)$ 表示先验项, λ 用于平衡数据项和先验项.

为了简化式(1)的求解过程,通常的做法是将数据项与先验项分开处理.为此,引入一个辅助变量 z ,从而构建出形如式(2)的约束优化问题,并通过式(3)的方法进行求解.

$$\hat{x} = \arg \min_x \frac{1}{2\sigma^2} \|y - T(x)\|^2 + \lambda R(z) \quad (2)$$

$$L_\mu(x, z) = \frac{1}{2\sigma^2} \|y - T(x)\|^2 + \lambda R(z) + \frac{\mu}{2} \|z - x\|^2 \quad (3)$$

其中, $z = x$, μ 是一个惩罚因子.

而式(3)可以分为以下两个子问题来迭代求解.

$$z_k = \arg \min_z \frac{1}{2(\sqrt{\lambda/\mu})^2} \|z - x_k\|^2 + R(z) \quad (4)$$

$$x_k = \arg \min_x \|y - T(x)\|^2 + \mu\sigma^2 \|x - z_{k-1}\|^2 \quad (5)$$

具体而言,式(4)是先验子问题,从贝叶斯的角度来看,其对应于高斯水平为 $\sqrt{\lambda/\mu}$ 的 x_k 上的高斯去噪;式(5)是数据子问题,其旨在利用依赖于 $T(\cdot)$ 的封闭数值解来约束修复图像的生成.因此,任何高斯去噪器都可以插入到式(4)中去解决先验子问题,并通过式(4)与式(5)的交替迭代来求解式(1).

随着深度学习的不断发展,Zhang等人^[9]提出了一种基于卷积神经网络(convolutional neural network, CNN)的强大去噪器,并将其作为隐式正则化器整合到即插即用图像修复算法中.这种方法有效地解决了传统即插即用图像修复算法需要指定显式和手工制作的图像先验的问题.Wang等人^[10]针对传统深度降噪器在即插即用框架下可能无法产生令人满意的实际图像修复结果这一挑战,提出了一种基于深度强化学习的即插即用框架,从而实现了更稳健的修复结果.此外,Li等人^[11]为了解决即插即用算法在处理缺失球面图像修复时的局限性,设计了一种新的面向球面图像的CNN降噪器并将其插入到即插即用算法中,使其可以恢复具有缺失损伤的球面图像.

2023年,Zhu等人^[12]针对图像信息严重缺失时即插即用算法恢复效果不佳的问题,受到深度生成模型具有卓越建模复杂分布能力的启发,将扩散模型与即插即用算法相结合,使得即插即用算法可以处理大规模掩码等一系列具有挑战性逆问题.同年,He等人^[13]

在利用空间先验的基础上,首次将频域先验引入到即插即用算法中,通过这种结合多域特征约束的方式使得即插即用算法在热红外图像去噪任务中取得了卓越的成果.2024年,Hurault等人^[14]针对现有即插即用算法在处理由泊松噪声引起的图像去噪和去模糊任务时所面临的局限性,提出了一种基于Bregman评分去噪器的即插即用算法,使得即插即用修复算法能够处理泊松逆问题.

从当前的研究趋势来看,现有的即插即用图像修复算法虽然能够取得良好的修复效果,但其研究重点主要集中在先验项上的优化,而对数据项的改进则相对较少.此外,这些研究大多针对的是图像补绘、去模糊以及超分辨率等退化模型已知的线性任务.相比之下,有关弱光图像增强等退化模型未知的非线性或盲任务的即插即用修复算法的研究较为缺乏.而在实际应用中,许多图像修复任务对应的退化模型往往都是未知的.因此,开发适用于非线性或盲任务的即插即用图像修复方法变得尤为重要.为此,本文从约束条件和引导策略两个角度进行研究,提出了一种基于即插即

用分裂算法的盲图像修复扩散模型(blind image restoration diffusion model based on plug-and-play splitting algorithm, SplitPIR).

2 SplitPIR

现有的即插即用图像修复方法可以通过使用不同的依赖于 $T(\cdot)$ 的封闭数值解来完成不同的线性修复任务.但是,这些方法往往也会受到封闭数值解的约束,即它们仅适用于存在真实或近似封闭数值解的情况.对于封闭数值解不存在的非线性或盲图像修复任务,这些方法则无能为力.为此,本文提出了一种名为基于即插即用分裂算法的盲图像修复扩散模型的方法.

本文方法首先利用HQS算法将退化模型和正则化项分解为形如式(4)和式(5)的独立数据子问题和先验子问题.之后,本文借用先验子问题的解求解出一个中间变量,并用真实退化图像 y 对其进行约束以解决数据子问题.最后,利用数据子问题的解间接引导下一步先验子问题解的生成,并通过迭代过程逐步逼近最终的解决方案.本文方法的整体思路如图1所示.

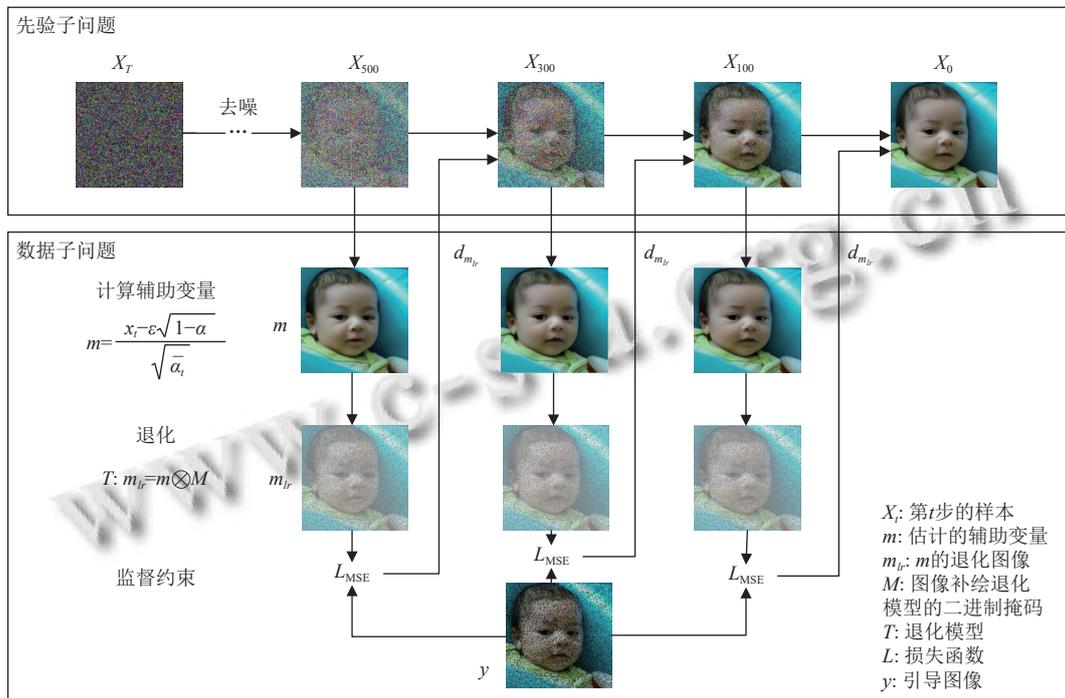


图1 本文算法整体架构图

2.1 数据子问题

对于线性任务,现行即插即用图像修复方法在解决数据子问题时依赖于封闭数值解的计算,这一数值解是基于先验子问题的解推导而来.然而,先验子问题

的解在扩散模型的影响下具有一定的随机性,这导致封闭数值解与真实解之间不可避免地存在偏差.这种偏差会随着数值解的传递而持续累积,最终对修复结果产生影响,使其与原始图像之间产生明显差异.相比

之下,本文方法摒弃了这种数值优化方法,转而采用真实退化图像来解决数据子问题并利用数据子问题的解间接地引导先验子问题解的生成,从而完成图像修复任务.这种方法能够有效减少偏差的累计.

因此,对于数据子问题,本文摒弃现行数值优化方法,转而采用真实退化图像来解决数据子问题并将数据子问题的解用于引导先验子问题解生成的做法,不仅可以极大地减少封闭数值解计算过程中产生的偏差对最终结果产生的负面影响,还可以将现行即插即用图像修复方法的应用领域拓展到非线性或盲任务中,以弱光图像增强任务为例.其具体的实现过程和原理在后续章节中介绍.

2.1.1 引导策略

假设我们有一个退化图像 y ,它是由原始的干净图像 x 通过一个退化过程 $T(\cdot)$ 产生的.那么,在不了解退化模型具体细节的情况下,图像修复任务可以被描述为从一组可能的原始图像 x 中寻找与退化图像 y 最为匹配的那一个.而本文方法抛弃了现行的数值优化方法,转而采用真实退化图像 y 来解决数据子问题.那么,在这个过程中,我们就可以模拟“锚定效应”来处理图像修复任务,即以退化图像 y 作为参照点,来定位和重建目标干净图像 x .

研究表明,去噪扩散概率模型(denoising diffusion probabilistic model, DDPM)中的逆向去噪过程能够以退化图像 y 作为条件^[15,16].这意味着,在DDPM框架下,原始的逆向去噪分布 $p_\theta(x_{t-1}|x_t)$ 可以被转化为一个条件分布 $p_\theta(x_{t-1}|x_t, y)$,从而使DDPM能够在去噪过程中利用特定的条件信息.这一特性与本文提出的方法不谋而合,即通过退化图像 y 来指导并重建对应的干净图像 x ,从而解决数据子问题.其数学推导如下^[17].

$$\begin{aligned} \log p_\theta(x_{t-1}|x_t, y) &= \log(p_\theta(x_{t-1}|x_t)p(y|x_t)) + K_1 \\ &\approx \log p(u) + K_2 \end{aligned} \quad (6)$$

为简洁起见, u 使用如下分布:

$$u \sim N(u; \mu_\theta(x_t, t) + \Sigma g, \Sigma) \quad (7)$$

其中, $g = \nabla_{x_t} \log p(y|x_t)$, $\Sigma = \Sigma_\theta(x_t)$, K_1 和 K_2 均为一个常数, $p_\theta(x_{t-1}|x_t)$ 表示DDPM中的反向去噪分布, $p(y|x_t)$ 可以理解为将 x_t 去噪后所得到的高质量图像与 y 一致的概率.

$p(y|x_t)$ 的一个近似解如下所示:

$$p(y|x_t) = \frac{1}{Z} \exp(-sE(T(x_t), y)) \quad (8)$$

其中, E 是一个图像的距离度量, Z 是一个归一化因子, s 是控制引导幅度的比例因子.

也就是说,式(8)鼓励通过退化模型 $T(\cdot)$ 后的 x_t 与真实退化图像 y 一致.对式(8)两边进行梯度计算:

$$\log p(y|x_t) = -\log Z - sE(T(x_t), y) \quad (9)$$

$$\nabla_{x_t} \log p(y|x_t) = -s \nabla_{x_t} E(T(x_t), y) \quad (10)$$

其中,距离度量 E 采用均方误差(mean square error, MSE)来衡量.

综上所述,通过向无条件逆向去噪分布 $p_\theta(x_{t-1}|x_t)$ 的均值中加入 $-s \nabla_{x_t} E(T(x_t), y)$ 就可以将其转化为条件分布 $p_\theta(x_{t-1}|x_t, y)$,从而实现借助真实退化图像 y 来解决数据子问题的目的.

此外,有研究表明,在扩散模型的逆向去噪过程中,在每一步生成图像前,对辅助变量 m 进行引导会获得更好的修复效果^[8].实际上,在DDPM的具体求解过程中,逆向去噪过程通常会先从 x_t 中预测出一个辅助变量,然后利用这个辅助变量 m 和 x_t 一起去采样下一步的图像^[18].辅助变量的求解方法如下所示:

$$m = \frac{x_t - \sqrt{1 - \bar{\alpha}_t} \varepsilon_\theta(x_t, t)}{\sqrt{\bar{\alpha}_t}} \quad (11)$$

因此,为了使本文方法能够获得更佳的修复效果,本文将数据子问题分裂成式(10)和式(11),并对这两部分分别进行处理,即首先利用数据子问题的输入求解出对应的中间变量 m ,其次利用退化模型 $T(\cdot)$ 将 m 退化为 m_t ,最后使用真实的退化图像 y 和 m_t 计算距离损失,用于指导下一步先验子问题解的生成.通过这种分裂策略,我们期望能够实现更稳定的图像修复效果^[19].

2.1.2 未知退化模型预测

通过上述推导过程可以发现,在将一个无条件逆向去噪分布 $p_\theta(x_{t-1}|x_t)$ 转化为一个条件分布 $p_\theta(x_{t-1}|x_t, y)$ 的过程中会涉及退化模型 $T(\cdot)$,而在非线性或盲任务中,退化模型通常都是未知的.因此,为了保证上述引导策略可以顺利进行,针对非线性或盲任务中退化模型未知的问题,本文引入了一个简单而有效的退化模型来模拟复杂的退化过程,引入模型的具体定义为:

$$y = f x + B \quad (12)$$

其中, f 表示光因子,是一个标量; B 表示光掩模,是一个与 x 维数相同的向量.

当 x 和 y 具有相同的大小时,任意一个损伤图像与

其对应的高质量图像之间的变换都可以被 f 和 B 所捕获. 因此我们可以使用这个模型来模拟复杂的未知退化模型.

此外, 由于 f 和 B 是退化模型的未知参数, 所以本文方法需要同时估计原始图像和退化模型的参数. 具体操作为: 针对每个单独的损坏图像, 首先随机初始化 f 和 B , 之后在 DDPM 的反向采样过程中对其进行同步优化, 优化方式可参见第 2.1.3 节算法 1 的第 6 步.

2.1.3 采样

本文所提出的方法在采样过程中遵循 DDPM 的采样框架, 但是在 DDPM 每一步迭代过程中均需要在解决了先验子问题和数据子问题后才能生成下一步迭代的初始图像, 具体流程如算法 1 所示. 其中, Σ 表示扩散模型估计出来的方差.

算法 1. SplitPIR

- 1) 从标准正态分布中随机采样一副噪声图像 x_T ;
- 2) for $t = T$ to 1 do
- 3) $x_0^{(t)} = \arg \min_z \frac{1}{2\sigma_t^2} \|z - x_t\|^2 + R(z)$
- 4) $m^{(t)} = \frac{x_0^{(t)} - \sqrt{1 - \alpha_t} \epsilon_{\theta}(x_0^{(t)}, t)}{\sqrt{\alpha_t}}$
- 5) $L_{\phi, m^{(t)}} = L(y, T_{\theta}(m^{(t)}))$
- 6) $\phi \leftarrow \phi - \eta \nabla_{\phi} E_{\phi, m^{(t)}}$
- 7) 从分布 $N(\mu + s \nabla_{\mu} E_{\phi, m^{(t)}}, \Sigma)$ 中随机采样一副图像 x_{t-1}
- 8) 结束循环
- 9) 返回

2.2 先验子问题

从概率的角度来看, 在大规模图像数据集 ImageNet 256×256 上训练好的扩散模型可以学习到自然图像的统计特性, 包括像素间的相关性和图像的结构特征, 这些特性可以被视为图像像素值的联合分布. 此时, 这个训练好的扩散模型就可以看作是一个学习到大量先验知识的“智者”. 因此, 本文选用在 ImageNet 256×256 数据集上预训练好的扩散模型作为正则化项去解决先验子问题, 使得恢复后的图像更趋近于真实图像.

3 实验分析

3.1 实验设计

由于本文使用了文献[17]所提供的在 ImageNet 数据集上预训练好的单个无条件的 DDPM, 所以本文方法无需训练. 为了有效评估实验结果, 对于线性任务, 本文采用了 FFHQ 256×256 和 ImageNet 256×256 数据集中的随机样本进行测试; 对于非线性或盲任务 (以弱光图像增强任务为例), 则采用 LOL 和 VE_LOL_L 数

据集中的随机样本进行测试, 以便对比分析本文方法与其他方法的性能.

本文的实验平台为 Linux 系统、Python 3.9 和 CUDA 11.7 结合的编程环境. 硬件配置环境为 Intel(R) Xeon(R) Silver 4210 CPU@2.20 GHz. 在测试过程中, 为了优化近似模型, 本文选择了 Adam 优化器, 其参数设置为 $\beta_1 = 0.9$, $\beta_2 = 0.999$, 初始学习率设定为 0.001. 同时, 本文将模型整体修复过程中的扩散次数设定为 1000 次, 引导尺度设置为 4000.

为了更加客观地评估本文方法的性能, 本文首先采用峰值信噪比 (peak signal-to-noise ratio, PSNR) 和结构相似性 (structure similarity index measure, SSIM) 来衡量修复图像的质量; 其次使用均方根误差 (root mean square error, RMSE) 来量化原始高质量图像与修复后的图像之间的差异.

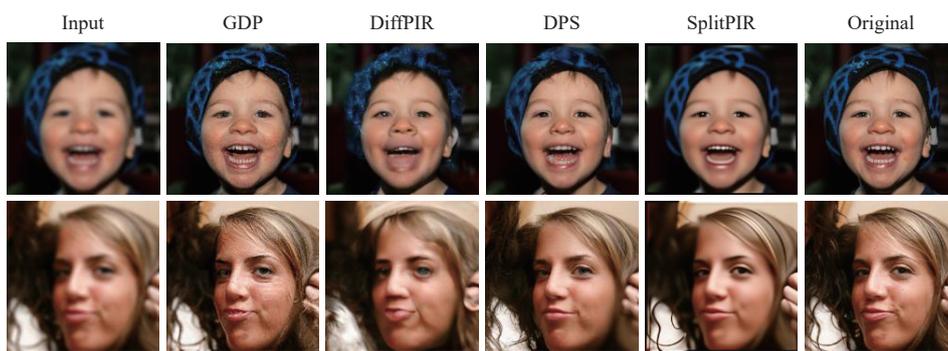
3.2 定性分析

为了可视化比较本文所提出的方法和对比方法, 本文针对不同的任务分别在对应的测试结果中随机抽取了两幅图像用于结果展示.

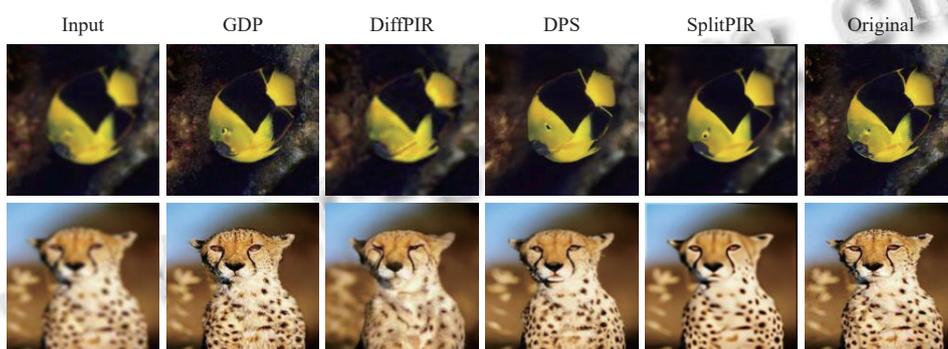
3.2.1 线性任务

对于线性任务, 本文方法与现行即插即用图像修复方法 DiffPIR 以及现有的非即插即用方法的统一恢复方法 GDP^[8]和 DPS^[20]方法进行了比较. 实验结果的部分展示如图 2、图 3 和图 4 所示. 从视觉效果来看, 采用本文方法得到的修复结果比基线方法的修复结果更贴近原始的真实图像. 这一优势在图像去模糊和超分辨率任务中表现得尤为突出.

观察图 2 和图 4 可以发现, GDP 产生的人脸修复结果在整体上较为接近原始的真实图像, 但其在细节上表现出来的效果不尽人意, 特别是在眼睛和嘴巴的邻近区域, 会出现一些真实图像中并不存在的细纹或瑕疵. DiffPIR 在两个数据集上的修复结果从整体上来看, 存在图像模糊、质量较低的问题; 从细节上来看, 也存在生成一些真实图像中不存在的细纹或噪声的问题. DPS 所产生的修复结果的整体质量较高, 细节上也不存在生成一些多余的纹理或噪声的问题, 但是通过这种方法所产生的修复结果可能会与真实图像产生较大的差异, 这个问题在图像超分辨率任务上表现尤其严重. 而本文方法所产生的修复结果既可以保证整体的图像质量, 在细节上, 也不会产生多余的信息, 从而确保了修复结果与真实图像的高度近似.

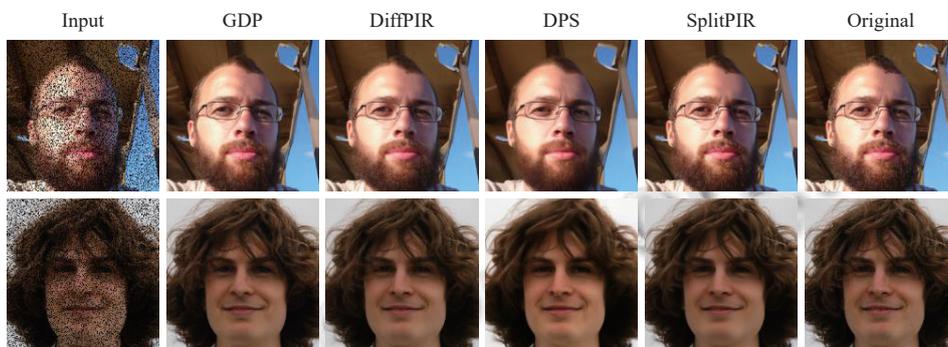


(a) FFHQ 256×256数据集上修复对比定性结果

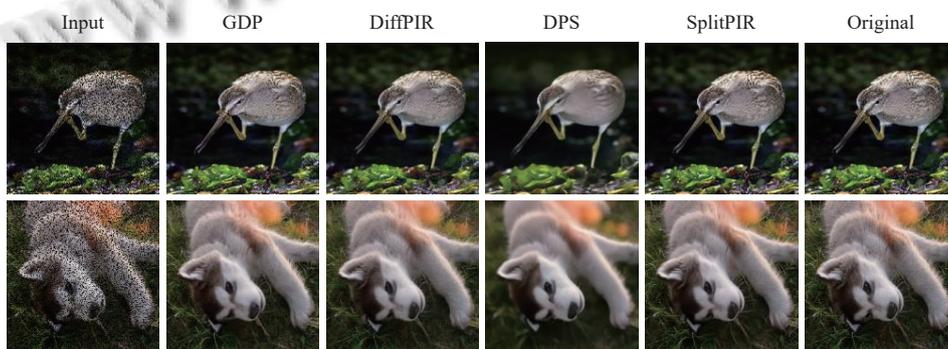


(b) ImageNet 256×256数据集上修复对比定性结果

图2 图像去模糊任务的定性比较结果

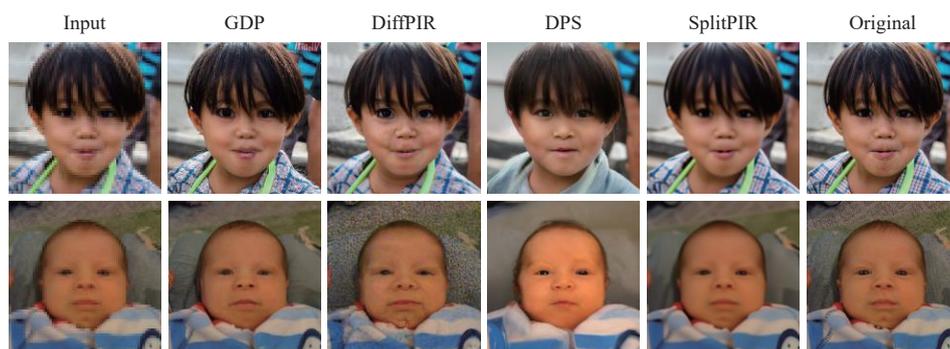


(a) FFHQ 256×256数据集上修复对比定性结果

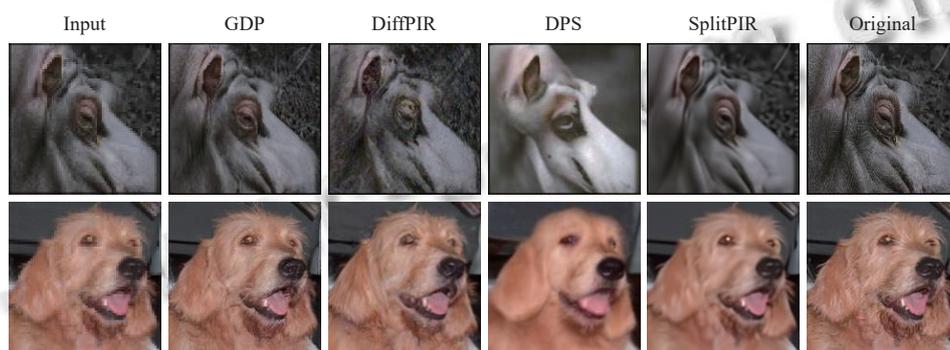


(b) ImageNet 256×256数据集上修复对比定性结果

图3 图像补绘任务的定性比较结果



(a) FFHQ 256×256数据集上修复对比定性结果



(b) ImageNet 256×256数据集上修复对比定性结果

图4 图像超分辨率任务的定性比较结果

3.2.2 非线性或盲任务

对于非线性或盲任务(以弱光图像增强任务为例),由于 DiffPIR 和 DPS 这两种方法无法完成类似于弱光图像增强的非线性任务,所以本文采用了一种既可以完成线性任务又可以完成非线性任务的非即插即用图像修复框架的 GDP 方法和两个专门针对弱光图像增强任务而设计的 RRDNet^[21]和 Zero-DCE++^[22]方法作为弱光图像增强任务的基线方法.实验结果的部分展示如图5所示.通过观察图5可以发现,由 RRDNet 和 Zero-DCE++两种方法所恢复出来的结果图像整体亮度较低,与原图有些差距,且结果图中存在大量真实图像中不存在的噪点.而 GDP 方法所修复的图像的亮度虽然比较接近原图,但是其结果图像中也存在大量噪点.而本文方法所产生的结果图像在保证整体亮度足够接近原始图像的同时,也可以很好地缓解对比方法生成多余噪点的问题.

3.3 定量分析

本文针对不同模型在相同数据集上进行的多种修复任务的修复结果进行了定量比较.定量结果中,最优值用加粗表示,次优值用下划线表示.

3.3.1 线性任务

对于线性任务,本文方法和基线方法在 FFHQ 和 ImageNet 数据集上的定量结果如表1和表2所示.通过对比分析实验结果,可以明显看出,无论是针对哪一种线性任务,本文方法在 PSNR 和 SSIM 这两个性能指标上的结果均优于基线方法.在 ImageNet 数据集上的3种线性任务测试过程中,PSNR 指标的值提高了1.12%–8.48%,SSIM 指标的值提高了4.49%–17.24%.在 FFHQ 数据集上的3种线性任务测试过程中,PSNR 指标的值提高了0.76%–9.31%,SSIM 指标的值提高了3.26%–12.16%.在 ImageNet 数据集上, RMSE 指标在图像补绘和图像超分辨率任务上均达到了最优值,虽然在图像去模糊任务上未能取得最优成绩,但仍然获得了次优的结果且与最优值相差不大,这充分验证了本文方法的有效性.在 FFHQ 数据集上, RMSE 指标的表现与其在 ImageNet 数据集上的情况类似.这些定量结果从科学的角度有效地证实了本文方法相较于其他基线方法在线性图像修复任务上具有显著优势.

3.3.2 非线性或盲任务

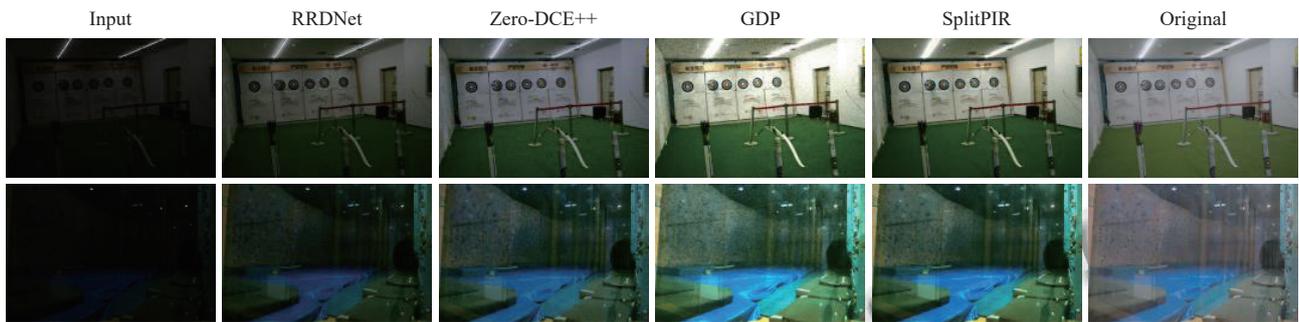
对于非线性或盲任务(以弱光图像增强任务为例),

本文方法和基线方法在 LOL 和 VE_LOL_L 数据集上的定量结果如表 3 所示。通过分析实验结果, 我们可以清晰地看到本文方法展现出了较强的竞争性。在 LOL 数据集上, 本文方法在 PSNR 和 SSIM 两个指标上均取得了最优结果, 相比于基线方法分别提升了 5.11% 和 3.17%, 在 RMSE 指标上本文方法虽然未能取得最佳成绩, 但是获得的次优结果与最优结果相差无几, 这也充分证明了本文方法的有效性。在 VE_LOL_L 数据集上, 本文方法在 SSIM 指标上达到了最优水平, 相比于基线

方法提升了 12.07%, 虽然 PSNR 和 RMSE 指标未能取得最佳成绩, 但获得的次优结果也充分体现了本文方法的有效性。上述这些定量结果分析充分证明了本文所提出的 SplitPIR 方法不仅有效克服了 DiffPIR 方法在处理非线性或盲任务时的局限性, 并且可以取得比现有的既可以完成线性任务又可以完成非线性或盲任务的统一框架更好的效果, 此外, 在弱光图像增强方面, 本文方法表现出了堪比专门设计的方法, 有时甚至超越了它们。这些分析综合印证了本文方法的有效性和实用性。



(a) LOL数据集上修复对比定性结果



(b) VE_LOL_L数据集上修复对比定性结果

图 5 弱光图像增强任务的定性比较结果

表 1 ImageNet 数据集上线性任务的定量比较结果

方法	图像补绘			图像超分辨率			图像去模糊		
	PSNR (dB)	SSIM	RMSE	PSNR (dB)	SSIM	RMSE	PSNR (dB)	SSIM	RMSE
GDP	28.44	0.83	6.610	22.30	0.55	8.054	22.13	<u>0.53</u>	8.053
DiffPIR	<u>30.31</u>	<u>0.89</u>	<u>4.899</u>	<u>23.60</u>	<u>0.58</u>	8.215	<u>22.30</u>	0.50	8.191
DPS	26.23	0.72	6.449	21.33	0.51	<u>7.543</u>	21.99	<u>0.53</u>	7.103
SplitPIR	32.88	0.93	4.887	24.77	0.68	7.150	22.55	0.61	<u>7.627</u>

表 2 FFHQ 数据集上线性任务的定量比较结果

方法	图像补绘			图像超分辨率			图像去模糊		
	PSNR (dB)	SSIM	RMSE	PSNR (dB)	SSIM	RMSE	PSNR (dB)	SSIM	RMSE
GDP	31.48	0.90	5.418	25.54	<u>0.74</u>	<u>6.929</u>	<u>25.07</u>	<u>0.72</u>	6.885
DiffPIR	<u>34.08</u>	<u>0.92</u>	3.769	<u>25.90</u>	<u>0.74</u>	7.355	24.41	0.68	7.324
DPS	29.83	0.87	5.248	23.61	0.68	6.960	24.97	<u>0.72</u>	6.182
SplitPIR	35.60	0.95	<u>3.914</u>	28.31	0.83	5.836	25.26	0.78	<u>6.433</u>

3.4 偏差分析

为了验证本文所提方法能有效减少现行即插即用

图像修复方法在解决数据子问题时产生的偏差, 我们进行了对比实验, 将本文方法与现行具有代表性的即

插即用图像修复方法 DiffPIR 进行了比较. 鉴于这两种方法都需要通过多次迭代来求解最终图像, 因此我们计算了每一轮迭代生成的图像与原始完好图像之间的均方误差 (mean squared error, MSE), 以此作为衡量最终修复结果与原始图像之间偏差的指标. 实验结果表明, 本文方法能够显著降低偏差的累积, 从而提升了图像修复的准确性, 具体实验结果如图 6 所示. 正如第 3.3.1 节中的定量分析所示, 本文方法相较于 DiffPIR 表现出更佳的效果, 这进一步证实了本文方法能够有效地抑制偏差的累积.

表 3 弱光图像增强任务的定量比较结果

方法	LOL			VE LOL L		
	PSNR (dB)	SSIM	RMSE	PSNR (dB)	SSIM	RMSE
RRDNet	11.37	0.53	5.325	13.46	<u>0.58</u>	5.275
Zero-DCE++	<u>14.86</u>	0.62	5.220	16.12	0.45	5.080
GDP	13.93	<u>0.63</u>	5.114	13.04	0.55	5.186
SplitPIR	15.62	0.65	<u>5.180</u>	<u>13.48</u>	0.65	<u>5.185</u>

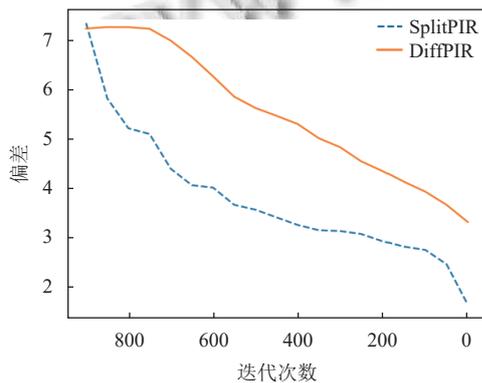


图 6 偏差分析结果图

3.5 消融实验

为了验证本文提出的引导策略和未知退化模型预测模块对最终修复结果的影响, 我们对这两个关键组件分别进行了消融实验. 考虑到去除这两个模块后模型将无法完成修复任务, 因此我们采用了替代方案来进行消融实验: 即用其他策略替换本文引导策略及未知退化模型预测模块. 通过这种方式, 我们能够更清晰地分析每个组件在整体性能中的贡献及其必要性.

3.5.1 引导策略

对于引导策略, 本文采用了两种常见的方案来替代本文方法: 一种是广泛应用于即插即用方法中的数值引导策略, 另一种是引导扩散模型中常用的直接引导策略^[23]. 通过对比分析实验结果 (如表 4 所示), 我们可以发现, 相较于其他两种引导策略, 本文提出的间接引导策略能够获得最优的结果.

表 4 引导策略消融实验结果

策略	PSNR (dB)	SSIM	MSE
数值引导策略	34.39	0.93	0.41
直接引导策略	<u>34.63</u>	<u>0.94</u>	<u>0.40</u>
间接引导策略 (ours)	35.78	0.95	0.39

3.5.2 未知退化模型预测模块

对于未知退化模型预测模块, 其核心在于近似退化模型的选择, 因此对该模块进行消融实验实际上就是对所选近似退化模型进行消融实验. 所以, 本文通过消融实验对比了两种替代方案: 基于 CNN 的 DLN^[24]模型和基于 Transformer 的 LLFormer^[25]模型. 实验结果如表 5 所示. 通过分析实验结果, 我们可以观察到, 虽然使用 DLN 和 LLFormer 作为预测模型能够获得较好的修复效果, 但相对于显著增加的参数量, 性能提升并不明显. 因此, 在综合考虑模型复杂度与性能之间的平衡后, 本文最终选择了线性模型作为未知退化模型的预测模块, 以简化了模型结构、降低计算资源的需求, 同时保证了令人满意的修复性能.

表 5 未知退化模型预测模块消融实验结果

模型	PSNR (dB)	SSIM	MSE	参数量
线性模型 (ours)	13.25	0.68	0.31	2
DLN	<u>13.53</u>	<u>0.69</u>	0.31	<u>700 586</u>
LLFormer	13.95	0.71	0.31	24549014

4 结论与展望

本文针对现有的即插即用图像修复方法在非线性或盲复原任务中的局限性, 提出了一种适用范围更广的即插即用图像修复方法. 该方法利用真实退化图像来解决数据子问题, 并利用数据子问题的解间接引导先验子问题解的生成. 通过这种间接引导的策略可以使最终的修复结果更接近真实理想图像. 此外, 面对非线性或盲复原任务中未知退化模型的挑战, 本文方法嵌入了一种简单高效的近似模型, 该模型能伴随着修复过程同步优化估计, 以此来逼近真实退化模型. 大量实验表明本文方法不仅可以提升即插即用图像修复方法的修复质量, 还扩展了其处理未知复杂退化情况的能力, 具有重要的实用价值和广泛的应用前景.

参考文献

- 1 Pan JS, Dong JX, Liu Y, et al. Physics-based generative adversarial models for image restoration and beyond. IEEE Transactions on Pattern Analysis and Machine Intelligence,

- 2021, 43(7): 2449–2462. [doi: [10.1109/TPAMI.2020.2969348](https://doi.org/10.1109/TPAMI.2020.2969348)]
- 2 黄健, 王虎, 赵小飞. 基于边缘图与多尺度特征融合的图像修复. 计算机系统应用, 2024, 33(4): 215–225. [doi: [10.15888/j.cnki.csa.009456](https://doi.org/10.15888/j.cnki.csa.009456)]
- 3 Belhasin O, Romano Y, Freedman D, *et al.* Principal uncertainty quantification with spatial correlation for image restoration problems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2024, 46(5): 3321–3333. [doi: [10.1109/TPAMI.2023.3343031](https://doi.org/10.1109/TPAMI.2023.3343031)]
- 4 Kamilov US, Bouman CA, Buzzard GT, *et al.* Plug-and-play methods for integrating physical and learned models in computational imaging: Theory, algorithms, and applications. *IEEE Signal Processing Magazine*, 2023, 40(1): 85–97. [doi: [10.1109/MSP.2022.3199595](https://doi.org/10.1109/MSP.2022.3199595)]
- 5 Bacca J, Monroy B, Arguello H. Deep plug-and-play algorithm for unsaturated imaging. *Proceedings of the 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Seoul: IEEE, 2024. 2460–2464.
- 6 Sun YB, Yang Y, Liu QS, *et al.* Learning non-locally regularized compressed sensing network with half-quadratic splitting. *IEEE Transactions on Multimedia*, 2020, 22(12): 3236–3248. [doi: [10.1109/TMM.2020.2973862](https://doi.org/10.1109/TMM.2020.2973862)]
- 7 Teodoro AM, Bioucas-Dias JM, Figueiredo MAT. Image restoration and reconstruction using targeted plug-and-play priors. *IEEE Transactions on Computational Imaging*, 2019, 5(4): 675–686. [doi: [10.1109/TCI.2019.2914773](https://doi.org/10.1109/TCI.2019.2914773)]
- 8 Fei B, Lyu ZY, Pan L, *et al.* Generative diffusion prior for unified image restoration and enhancement. *Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Vancouver: IEEE, 2023. 9935–9946.
- 9 Zhang K, Li YW, Zuo WM, *et al.* Plug-and-play image restoration with deep denoiser prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, 44(10): 6360–6376. [doi: [10.1109/TPAMI.2021.3088914](https://doi.org/10.1109/TPAMI.2021.3088914)]
- 10 Wang C, Zhang RK, Ravishanker S, *et al.* REPnP: Plug-and-play with deep reinforcement learning prior for robust image restoration. *Proceedings of the 2022 IEEE International Conference on Image Processing (ICIP)*. Bordeaux: IEEE, 2022. 2886–2890.
- 11 Li JF, Huang CY, Chan R, *et al.* Spherical image inpainting with frame transformation and data-driven prior deep networks. *SIAM Journal on Imaging Sciences*, 2023, 16(3): 1177–1194. [doi: [10.1137/22M152462X](https://doi.org/10.1137/22M152462X)]
- 12 Zhu YZ, Zhang K, Liang JY, *et al.* Denoising diffusion models for plug-and-play image restoration. *Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Vancouver: IEEE, 2023. 1219–1229.
- 13 He Y, Zhang C, Zhang B, *et al.* FSPnP: Plug-and-play frequency-spatial-domain hybrid denoiser for thermal infrared image. *IEEE Transactions on Geoscience and Remote Sensing*, 2023, 62: 5000416.
- 14 Hurault S, Kamilov U, Leclaire A, *et al.* Convergent bregman plug-and-play image restoration for poisson inverse problems. *Proceedings of the 37th International Conference on Neural Information Processing Systems*. New Orleans: ACM, 2024. 1187.
- 15 Saharia C, Ho J, Chan W, *et al.* Image super-resolution via iterative refinement. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, 45(4): 4713–4726.
- 16 Baldassari L, Siahkoobi A, Garnier J, *et al.* Conditional score-based diffusion models for Bayesian inference in infinite dimensions. *Proceedings of the 37th International Conference on Neural Information Processing Systems*. New Orleans: ACM, 2024. 1055.
- 17 Dhariwal P, Nichol A. Diffusion models beat GANs on image synthesis. *Proceedings of the 35th International Conference on Neural Information Processing Systems*. ACM, 2021. 672.
- 18 Ho J, Jain A, Abbeel P. Denoising diffusion probabilistic models. *Proceedings of the 34th International Conference on Neural Information Processing Systems*. Vancouver: ACM, 2020. 574.
- 19 Condat L, Kitahara D, Contreras A, *et al.* Proximal splitting algorithms for convex optimization: A tour of recent advances, with new twists. *SIAM Review*, 2023, 65(2): 375–435. [doi: [10.1137/20M1379344](https://doi.org/10.1137/20M1379344)]
- 20 Chung H, Kim J, McCann MT, *et al.* Diffusion posterior sampling for general noisy inverse problems. *Proceedings of the 11th International Conference on Learning Representations*. Kigali: ICLR, 2023.
- 21 Zhu AQ, Zhang L, Shen Y, *et al.* Zero-shot restoration of underexposed images via robust retinex decomposition. *Proceedings of the 2020 IEEE International Conference on Multimedia and Expo (ICME)*. London: IEEE, 2020. 1–6.
- 22 Li CY, Guo CL, Loy CC. Learning to enhance low-light image via zero-reference deep curve estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, 44(8): 4225–4238.
- 23 Liu XH, Park DH, Azadi S, *et al.* More control for free! image synthesis with semantic diffusion guidance. *Proceedings of the 2023 IEEE/CVF Winter Conference on Applications of Computer Vision*. Waikoloa: IEEE, 2023. 289–299.
- 24 Wang LW, Liu ZS, Siu WC, *et al.* Lightning network for low-light image enhancement. *IEEE Transactions on Image Processing*, 2020, 29: 7984–7996. [doi: [10.1109/TIP.2020.3008396](https://doi.org/10.1109/TIP.2020.3008396)]
- 25 Wang T, Zhang KH, Shen TR, *et al.* Ultra-high-definition low-light image enhancement: A benchmark and Transformer-based method. *Proceedings of the 37th AAAI Conference on Artificial Intelligence*. Washington: AAAI, 2023. 2654–2662.

(校对责编: 张重毅)