

多智能体协同的空间感知图像恢复方法综述

2026 年 2 月 26 日

摘要

图像恢复旨在从由噪声、模糊、雾霾、雨纹及压缩失真导致的退化观测中重建高质量图像，近年来随深度学习技术的进步取得了显著突破。基于 Transformer 的单任务模型通过全局自注意力机制实现了性能飞跃；All-in-One (AiO) 统一框架借助降质感知机制在单一模型内处理多种退化类型；扩散生成模型利用强大的图像先验在严重退化场景下实现了卓越的感知质量。然而，上述方法均以全图均匀处理为前提，难以应对现实场景中退化在空间上异质分布这一核心挑战。近年兴起的基于大型语言模型 (LLM) 和视觉语言模型 (VLM) 的智能体框架，通过动态感知、规划与工具编排开辟了新范式。本文系统梳理上述演进历程，重点聚焦空间感知与多智能体协同，分析基础模型在退化感知中的支撑作用，探讨轨迹蒸馏在轻量化边缘部署中的潜力，并展望空间感知多智能体图像恢复的未来研究方向。

1 引言

图像恢复是计算机视觉的基础性问题，其目标是从退化观测 $\mathbf{y} = \mathcal{D}(\mathbf{x}) + \mathbf{n}$ 中恢复洁净图像 \mathbf{x} ，其中 $\mathcal{D}(\cdot)$ 为退化算子， \mathbf{n} 为噪声。在监控、遥感、医学成像等实际场景中，图像往往同时受到多类退化影响，恢复难度极大。

深度学习的兴起彻底改变了图像恢复领域的格局。从卷积神经网络 (CNN) 到基于 Transformer 的架构，单任务恢复模型性能持续突破。然而，为每种退化类型单独维护模型既不经济又不实用，由此催生了 AiO 统一框架：Li 等人 [1] 提出 AirNet，利用对比学习构建降质感知表征；Potlapalli 等人 [2] 引入可学习提示向量驱动统一恢复；Conde 等人 [3] 首次以自然语言指令引导恢复过程。基于扩散概率模型的生成式方法也在严重退化场景下展现出无与伦比的感知质量 [4, 5]。

然而，上述所有范式共享一个根本假设：对全图采用统一恢复策略。现实场景中，单张图像内不同区域的退化类型与程度往往存在显著差异——前景可能存在运动模糊，背景可能弥漫雾霾，阴影区域充斥噪声，场景中还可能横贯雨纹。这一观察催生了基于智能体的图像恢复范式：Zhu 等人 [6] 提出 AgenticIR，以 VLM 为核心实现感知 → 调度 → 执行 → 反思 → 重调度的五阶段动态恢复规划；Chen 等人 [7] 开发 RestoreAgent 以端到端微调多模态 LLM；Jiang 等人 [8] 的多智能体框架 MAIR 将推理效率提升 44%；Zhou 等人 [9] 提出质量驱动贪婪策略 Q-Agent，实现线性复杂度规划。

本文系统梳理从单任务 Transformer 到多智能体空间感知图像恢复的演进历程，归纳关键方法论创新，指出现有研究的局限，并展望未来研究方向。

2 基于 Transformer 的图像恢复

2.1 核心架构创新

标准自注意力机制的计算复杂度与像素数量的平方成正比，对高分辨率恢复代价高昂。SwinIR[10] 通过残差 Swin Transformer 块 (RSTB) 和移位窗口多头自注意力将计算复杂度降至线性，在图像去噪、超分辨率和 JPEG 压缩失真消除等任务上较此前最优方法提升 0.14–0.31 dB PSNR，成为 Transformer 图像恢复的基准方法。Restormer[11] 则沿通道维度而非空间维度计算注意力，提出多深度卷积头转置注意力 (MDTA) 与门控深度卷积前馈网络 (GDFN)，在保持全局感受野的同时实现线性空间复杂度，在 SIDD 去噪基准上较前一最优提升 0.13–0.52 dB。Uformer[12] 构建层次化 U 型 Transformer 并配合可学习多尺度恢复调制器，兼顾了去噪和去模糊任务的性能与效率。

频域方法 FFTformer[13] 将自注意力替换为傅里叶域操作，以更低计算代价实现全局感受野，尤其适用于图像去模糊。基于 Mamba 状态空间模型的 VmambaIR[14] 通过选择性状态空间扫描以线性复杂度建模长程依赖，为高效图像恢复提供了新思路。Wu 等人 [15] 在 DSwinIR 中指出 SwinIR 固定网格窗口引入边界伪影问题，提出内容自适应动态窗口注意力，改善了非均匀退化的处理能力。

2.2 混合架构、自监督训练与局限

结合 CNN 局部纹理提取能力与 Transformer 全局建模能力的混合架构在多任务中表现优异。Chen 等人 [16] 提出 CNN-Transformer 混合特征融合网络，通过交叉注意力融合局部与全局特征，专门用于单图去雨。在训练范式方面，Zhang 和 Zhou[17] 提出基于上下文感知 Transformer 的自监督去噪方法，无需配对洁净-噪声训练数据，性能可与全监督方法相媲美。

尽管性能优异，单任务 Transformer 的根本局限在于任务专一性。Ali 等人 [18] 的综述指出，大量任务专用模型的维护给实际部署带来了显著挑战，这直接驱动了统一恢复方法的发展。

3 统一化 All-in-One 图像恢复

3.1 对比学习与提示驱动方法

AiO 恢复的核心挑战在于使单一模型实现对多种退化类型的判别性表征。AirNet[1] 通过对比式降质编码器 (CBDE) 学习降质判别表征，实现了在去噪、去雾、去雨任务上无需显式降质标签的自适应处理。Hu 等人 [19] 进一步在多级别 (图像、特征、块) 对比框架中提升了降质判别精度。

提示学习范式的引入是 AiO 恢复的重大进展。PromptIR[2] 通过可学习提示向量对 Transformer 主干进行软条件化，在去噪、去雾、去雨、去模糊四任务内实现统一处理，展示了连续提示空

间能够表征不同退化之间的中间状态。随后多项工作扩展了提示示范式：Ma 等人 [20] 的 ProRes 从降质先验构建可解释提示；Wu 等人 [21] 提出输入自适应的动态提示生成；Wu 等人 [22] 在 FrePrompter 中直接从输入频谱提取物理有据的降质提示；Wu 等人 [23] 以对比提示学习解决不同降质类型间的提示表征冗余问题。

3.2 指令引导、混合专家与局限

InstructIR[3] 通过自然语言指令控制恢复过程，是首个将图像恢复与自然语言理解相结合的方法，可同时处理七类恢复任务。Tang 等人 [24] 提出 RamIR，将 Mamba 状态空间架构与链式思维推理相结合，通过显式推理-行动分解处理复杂混合退化。Dudhane 等人 [25] 采用动态预训练策略，以课程学习方式渐进暴露于更复杂退化，提升了训练效率和最终性能。Zamfir 等人 [26] 提出 MoCE-IR 混合专家架构，通过稀疏专家激活将计算资源按恢复难度分配，在多个基准上达到最优性能。Wu 等人 [27] 的“多样中的和谐”方法通过动态生成的降质感知卷积核使单一网络能为每个输入定制处理，无需显式任务切换。

尽管 AiO 方法进展显著，Jiang 等人 [28] 的综述指出其三项持续局限：性能在遇到训练集外的新型退化时显著下降；全图单一流程处理忽视了退化的空间异质性；单次前向推理无法基于中间质量评估进行迭代细化。这些局限性直接驱动了更灵活的智能体化方法的发展。

4 基于扩散模型的生成式图像恢复

4.1 基础框架与条件生成

扩散概率模型通过迭代去噪过程建模复杂数据分布，为图像恢复提供了强大的生成先验。Palette[4] 将条件扩散模型应用于图像间转换任务，展示了扩散方法能够生成多样化高质量输出的能力，这是确定性方法无法实现的。RDDM[29] 将扩散过程重新定义为在洁净图像与退化图像残差上操作，显著减少了所需采样步数，在保持高恢复质量的同时大幅降低了推理成本。Zheng 等人 [30] 提出选择性漏斗映射策略，仅对最严重退化的频率成分应用扩散，对保存良好的部分直接直通。Ding 等人 [31] 引入显式约束使扩散采样在保持对退化输入忠实度的同时引入生成先验，在保真度与感知质量之间寻求平衡。

4.2 预训练先验利用与任务适应

DiffBIR[5] 开创了利用预训练大型扩散模型先验的方法：第一阶段消除主要退化，第二阶段借助预训练 Stable Diffusion 的生成先验细化结果，在真实世界盲图像超分辨率和人脸恢复上取得了卓越感知质量。Yang 等人 [32] 提出像素感知交叉注意力 (PASD)，使扩散模型在合成细节的同时保持对输入的高保真度。扩散模型被广泛适应于特定任务：Guo 等人 [33] 提出 ShadowDiffusion 用于去阴影；Yi 等人 [34] 将 Retinex 理论与扩散模型结合提出 Diff-Retinex 用于低光增强，并在 Diff-Retinex++[35] 中引入强化学习优化采样轨迹；Luo 等人 [36] 提出 Refusion 解决大尺寸图像的分辨率限制；Wang 等人 [37] 设计块级扩散策略，使扩散恢复能够处理任意分辨率图像。

在 AiO 扩散方向，Tu 等人 [38] 利用扩散桥接对不同退化分布转变建模；Luo 等人 [39] 以视觉指令引导扩散用于 AiO 恢复；Yue 等人 [40] 同时以多降质描述符为条件处理混合退化。尽管感

知质量出众，迭代采样的高计算成本及可能引入的幻觉细节（在医学图像等高保真度场景尤为关键）仍是扩散方法的持续挑战。

5 基于智能体的图像恢复系统

5.1 单智能体框架

基于智能体的图像恢复从根本上将恢复过程重新定义为智能决策过程。AgenticIR[6] 建立了该范式的基础框架：以 VLM 通过五阶段流程（感知 → 调度 → 执行 → 反思 → 重调度）分析退化、规划工具序列，并通过深度优先搜索进行迭代细化。该框架能够动态组合任意专用工具序列，并通过自我评估和迭代细化改善结果，与先前所有方法存在根本性差异。然而，其深度优先搜索策略在最坏情况下导致指数级计算代价，对实时部署构成挑战。

Zhu 等人 [41] 进一步提出的智能体系统按复杂度将恢复问题分为三级（简单、中等、复杂），并自适应调整推理深度：对简单问题执行单步工具选择，对复杂问题激活完整审议流程。该系统还引入经验记忆模块，使系统能够在部署过程中积累和复用成功的恢复策略，提供了一条随部署经验持续改进而无需周期性重训练的路径。

RestoreAgent[7] 通过在恢复任务-计划对上端到端微调多模态 LLM，将多步推理压缩为单次前向传播，实现更快推理速度。Q-Agent[9] 从搜索策略角度出发，以图像质量评估模型预测的预期质量提升贪婪选择下一步工具，实现了相对工具数量的线性计算复杂度，显著提升了实用部署的可行性。

5.2 多智能体架构与空间感知扩展

MAIR[8] 引入三阶段降质先验（压缩失真 → 成像退化 → 场景退化）的多智能体框架，通过限制工具选择空间和并行化处理，在保持与 AgenticIR 相当恢复质量的同时将推理效率提升 44%。Tripathi 等人 [42] 分析指出，多智能体层次化设计——全局协调智能体管理整体策略，专用子智能体处理特定方面——能够随退化类型和工具数量的增加实现自然扩展。

当前基于智能体系统的关键局限在于将整幅图像作为单一实体处理，忽视了退化在空间上的显著异质性。实现空间感知所需的工具链已基本就绪：SAM[43] 提供零样本语义分割能力，可将图像分解为具有语义意义的空间区域；RobustSAM[44] 专门针对退化图像的分割质量下降问题进行优化；Grounded SAM[45] 结合文字提示目标检测，支持通过自然语言定位关注区域（如“雾霾天空区域”或“噪声阴影区域”）；DepictQA[46] 等 VLM 可对各区域独立进行退化类型识别和严重程度的自然语言描述。协调智能体、分割工具与区域级恢复策略的空间感知多智能体框架，代表了应对现实图像异质退化的最有前途的范式。

6 基础模型与图像质量评估

6.1 视觉语言模型用于退化分析

VLM 为基于智能体的恢复系统提供了关键的感知基础。DepictQA[46] 借助 VLM 提供对图像质量属性的详细自然语言描述，包括具体退化类型、严重程度及受影响区域，远超简单分类的表达能力。You 等人 [47] 扩展 DepictQA 以支持图像对之间的比较质量评估，直接对应智能体系统中恢复前后效果的自我评估需求。Wu 等人 [48] 构建了面向低级视觉感知任务的大规模指令遵循数据集 Q-Instruct，在该数据集上微调的 VLM 在降质识别和描述方面显著优于通用 VLM。Wu 等人 [49] 建立的 Q-Bench 基准及其扩展 Q-Bench+[50]，提供了覆盖感知、描述和比较任务的 VLM 质量评估标准化评估协议。

6.2 基础模型先验与无参考质量评估

大规模基础模型编码的视觉语义先验为真实世界图像恢复提供了强大的泛化能力。Ai 等人 [51] 提出 DreamClear，结合大容量扩散模型与隐私安全数据集构建流程，在多个真实世界恢复基准上实现最优性能，表明大模型容量与大规模精选数据的结合对处理真实降质多样性至关重要。Luo 等人 [52] 开发了通过受控 VLM 实现真实场景照片级图像恢复的方法，其中 VLM 不仅用于感知还主动参与恢复过程，通过生成对目标结果的自然语言描述来指导条件扩散模型，形成了“语义指导”的新恢复模式。

经典无参考图像质量评估 (NR-IQA) 方法为恢复流程提供高效质量分数，指导工具选择和终止决策。NIMA[53] 预测人类质量评分的分布而非单一分值，同时提供质量估计和不确定性度量。LPIPS[54] 通过预训练网络特征计算与人类判断高度相关的感知相似度指标。Zheng 等人 [55] 提出专门用于评估图像恢复质量的条件知识蒸馏方法，联合利用退化输入和恢复结果进行质量评估；Li 等人 [56] 将质量评估反馈信号直接整合到去模糊恢复过程中，建立了质量驱动恢复的概念——这正是 Q-Agent 等方法的核心理念。

7 模型压缩与轻量化部署

Transformer、扩散模型和基于智能体方法的高计算需求，对资源受限设备的部署构成了重大障碍。知识蒸馏是最重要的压缩策略之一。Zhang 等人 [57] 提出软知识蒸馏结合多维交叉网络注意力 (SKD) 的压缩方法，学生模型不仅复现教师输出还学习其跨空间和通道维度的中间注意力模式，以 2-4 倍参数压缩实现极小质量损失。Yang 等人 [58] 提出面向 Mamba 的异构知识蒸馏，将 Transformer 教师的知识迁移至 Mamba 架构学生，实现 Transformer 质量水平的高效边缘部署。Wang 等人 [59] 提出无数据蒸馏方法，利用扩散模型生成合成退化图像，解决了原始训练数据不可用时的蒸馏挑战。

针对基于智能体的恢复系统，**轨迹蒸馏**代表着最具潜力的轻量化路径：离线运行完整智能体系统，记录包含区域分割、降质识别、工具选择、参数设置与质量评分的完整决策轨迹，随后训练轻量紧凑网络直接从输入图像预测这些决策，在部署时无需昂贵的 LLM/VLM 推理，可实现实时处理。如何确保轻量模型忠实复现智能体对边缘情况和未见退化组合的细粒度决策，是该方

向的核心挑战。

8 讨论

8.1 质量与效率的三难困境

贯穿所有恢复范式的持续张力在于：实现更高恢复质量的方法往往需要显著更多的计算资源。单任务 Transformer（如 Restormer[11]）在特定任务上实现了良好的质量-效率平衡但无法处理多种退化；AiO 模型以牺牲部分单任务质量换取通用性；扩散模型实现最佳感知质量但推理成本高出数个数量级；基于智能体的系统最灵活，潜在恢复质量最高，但当前需要多轮 LLM/VLM 推理。轨迹蒸馏是解决这一三难困境最有前途的路径，其有效性依赖于轨迹数据的覆盖质量与轻量模型的泛化能力。

8.2 合成退化与真实退化的差距

大多数恢复方法在合成退化图像上开发和评估，而真实世界退化显著不同：它是空间非均匀的、涉及多种退化类型的复杂交互，且无法用精确数学模型描述。Zhai 等人 [60] 的综述揭示了合成与真实基准之间的显著性能差距；Guan 等人 [61] 贡献了 WeatherBench 基准，提供了更真实的评估条件。基于智能体和空间感知的方法天然适于弥合这一差距，因为它们能够自适应响应每幅图像的特定退化特征，而非依赖对退化模型的先验假设。

8.3 空间感知的缺失

本文最重要的发现之一是空间感知在所有现有范式中均未得到充分发展：单任务与 AiO 方法均匀处理整幅图像；扩散模型对全图应用同一生成过程；即使是当前最先进的基于智能体系统，在制定恢复决策时也将图像视为整体实体。真正的空间感知需要将图像分解为有意义区域，独立分析各区域退化特征，制定区域特定的恢复策略，并无缝融合结果，这将从根本上提升对真实世界异质退化图像的恢复质量。

9 展望

基于本文分析，以下方向有望推动图像恢复领域的进一步突破。

层次化多智能体空间感知恢复架构：全局协调智能体利用 SAM 分解图像，区域专家智能体独立处理各空间区域，质量评估智能体监督整体效果并触发迭代细化，构建自然支持退化空间异质性的层次化多智能体系统。

连续参数优化：当前智能体主要做离散工具选择，而连续参数优化——智能体不仅选择工具还调整其操作参数（如不同区域使用不同噪声估计）——将大幅提升恢复的灵活性与精度，可结合强化学习与质量驱动反馈实现。

面向边缘部署的轨迹蒸馏：关键挑战包括：设计同时预测区域分割、降质类型、工具路由和连续参数的多任务学生架构；开发优先处理多样化和挑战性样本的轨迹选择策略；以及建立同时评估逐图恢复质量、决策一致性和鲁棒性的评估协议。

跨域泛化与标准化基准：领域亟需涵盖具有空间异质退化真实图像、支持区域级质量评估指标的标准化基准，以及跨医学、遥感、工业成像的泛化评估体系。

10 结论

本文系统梳理了图像恢复从单任务 Transformer 架构，经 AiO 统一模型和扩散式生成方法，到基于智能体空间感知系统的演进历程。每一范式均针对其前驱的特定局限：Transformer 引入超越 CNN 的全局上下文建模；AiO 模型消除了对特定任务网络的需求；扩散模型通过生成先验实现了前所未有的感知质量；基于智能体的系统引入了动态推理、工具编排和迭代细化。

仍然存在的关键缺口是空间感知。空间分解基础模型（SAM）、区域级退化分析视觉语言模型（DepictQA 等）与多智能体协同恢复框架的汇聚，为弥合这一差距创造了独特机遇。结合轨迹蒸馏实现轻量化边缘部署，空间感知多智能体恢复有望建立同时更精确、更自适应、更易部署的新范式，在感知、推理与行动的融合中推动图像恢复迈入智能化新阶段。

参考文献

- [1] Li B, Liu X, Hu P, Wu Z, Lv J, Peng X. AirNet: All-in-One Image Restoration Network via Contrastive Learning[C/OL] // IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2022. <https://doi.org/10.1109/CVPR52688.2022.02014>.
- [2] Potlapalli V, Zamir S W, Khan S, Khan F S. PromptIR: Prompting for All-in-One Blind Image Restoration[J/OL]. arXiv (Cornell University), 2023. <https://doi.org/10.48550/arxiv.2306.13090>.
- [3] Conde M V, Geigle G, Timofte R. InstructIR: High-Quality Image Restoration Following Human Instructions[C] // European Conference on Computer Vision (ECCV). 2024.
- [4] Saharia C, Chan W, Chang H, Lee C, Ho J, Salimans T, Fleet D J, Norouzi M. Palette: Image-to-Image Diffusion Models[J/OL], 2022. <https://doi.org/10.1145/3528233.3530757>.
- [5] Lin X, He J, Chen Z, Lyu Z, Dai B, Yu F, Qiao Y, Ouyang W, Dong C. DiffBIR: Toward Blind Image Restoration with Generative Diffusion Prior[J/OL]. Lecture notes in computer science, 2024. https://doi.org/10.1007/978-3-031-73202-7_25.
- [6] Zhu H, Xu K, Wang R. AgenticIR: Agentic Image Restoration with Visual Language Models[J]. arXiv preprint, 2025.
- [7] Chen H, Chen S, Gu J, Li W, Pei R, Ren J, Song F, Tian Y, Zhou K, Zhu L. RestoreAgent: Autonomous Image Restoration Agent via Multimodal Large Language Models[J/OL], 2024. <https://doi.org/10.52202/079017-3512>.
- [8] Jiang X, Li G, Chen B, Zhang J. Multi-Agent Image Restoration[J/OL]. arXiv (Cornell University), 2025. <https://doi.org/10.48550/arxiv.2503.09403>.
- [9] Zhou J, Others. Q-Agent: Quality-Driven Agent for Image Restoration[J]. arXiv preprint, 2025.
- [10] Liang J, Cao J, Sun G, Zhang K, Gool L V, Timofte R. SwinIR: Image Restoration Using Swin Transformer[J/OL], 2021. <https://doi.org/10.1109/iccvw54120.2021.00210>.
- [11] Zamir S W, Arora A, Khan S, Hayat M, Khan F S, Yang M. Restormer: Efficient Transformer for High-Resolution Image Restoration[C/OL] // 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2022. <https://doi.org/10.1109/cvpr52688.2022.00564>.

- [12] Wang Z, Cun X, Bao J, Zhou W, Liu J, Li H. Uformer: A General U-Shaped Transformer for Image Restoration[C/OL] // 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2022. <https://doi.org/10.1109/cvpr52688.2022.01716>.
- [13] Kong L, Dong J, Ge J, Li M, Pan J. Efficient Frequency Domain-based Transformers for High-Quality Image Deblurring[J/OL], 2023. <https://doi.org/10.1109/cvpr52729.2023.00570>.
- [14] Shi Y, Xia B, Jin X, Wang X, Zhao T, Xia X, Xiao X, Yang W. VmambaIR: Visual State Space Model for Image Restoration[J/OL]. IEEE Transactions on Circuits and Systems for Video Technology, 2025. <https://doi.org/10.1109/tcsvt.2025.3530090>.
- [15] Wu G, Jiang J, Jiang K, Liu X, Nie L. DSwinIR: Rethinking Window-based Attention for Image Restoration[J/OL]. arXiv (Cornell University), 2025. <https://doi.org/10.48550/arxiv.2504.04869>.
- [16] Chen X, Pan J, Lu J, Fan Z, Li H. Hybrid CNN-Transformer Feature Fusion for Single Image Deraining[C/OL] // Proceedings of the AAAI Conference on Artificial Intelligence. 2023. <https://doi.org/10.1609/aaai.v37i1.25111>.
- [17] Zhang D, Zhou F. Self-Supervised Image Denoising for Real-World Images With Context-Aware Transformer[J/OL]. IEEE Access, 2023. <https://doi.org/10.1109/access.2023.3243829>.
- [18] Ali A M, Benjdira B, Koubâa A, El-Shafai W, Khan Z, Boulila W. Vision Transformers in Image Restoration: A Survey[J/OL]. Sensors, 2023. <https://doi.org/10.3390/s23052385>.
- [19] Hu B, Yang S, Liu F, Ding W. Collaborative Semantic Contrastive for All-in-one Image Restoration[J/OL]. Engineering Applications of Artificial Intelligence, 2025. <https://doi.org/10.1016/j.engappai.2025.110017>.
- [20] Ma J, Cheng T, Wang G, Zhang Q, Wang X, Zhang L. ProRes: Exploring Degradation-aware Visual Prompt for Universal Image Restoration[J/OL]. arXiv (Cornell University), 2023. <https://doi.org/10.48550/arxiv.2306.13653>.
- [21] Wu G, Jiang J, Jiang K, Liu X, Nie L. Learning Dynamic Prompts for All-in-One Image Restoration[J/OL]. IEEE Transactions on Image Processing, 2025. <https://doi.org/10.1109/tip.2025.3567205>.
- [22] Wu Z, Liu W, Wang J, Li J, Huang D. FrePrompter: Frequency self-prompt for all-in-one image restoration[J/OL]. Pattern Recognition, 2024. <https://doi.org/10.1016/j.patcog.2024.111223>.
- [23] Wu G, Jiang J, Jiang K, Liu X, Nie L. Beyond Degradation Redundancy: Contrastive Prompt Learning for All-in-One Image Restoration[J/OL]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2025. <https://doi.org/10.1109/tpami.2025.3642852>.
- [24] Tang A H, Wu Y, Zhang Y. RamIR: Reasoning and action prompting with Mamba for all-in-one image restoration[J/OL]. Applied Intelligence, 2025. <https://doi.org/10.1007/s10489-024-06226-y>.
- [25] Dudhane A, Thawakar O, Zamir S W, Khan S, Khan F S, Yang M-H. Dynamic Pre-training: Towards Efficient and Scalable All-in-One Image Restoration[J/OL]. arXiv (Cornell University), 2024. <https://doi.org/10.48550/arxiv.2404.02154>.
- [26] Zamfir E, Wu Z, Mehta N, Tan Y, Paudel D P, Zhang Y, Timofte R. Complexity Experts are Task-Discriminative Learners for Any Image Restoration[J/OL], 2025. <https://doi.org/10.1109/cvpr52734.2025.01190>.
- [27] Wu G, Jiang J, Jiang K, Liu X. Harmony in Diversity: Improving All-in-One Image Restoration via Multi-Task Collaboration[J/OL], 2024. <https://doi.org/10.1145/3664647.3680762>.

- [28] Jiang J, Zuo Z, Wu G, Jiang K, Liu X. A Survey on All-in-One Image Restoration: Taxonomy, Evaluation and Future Trends[J/OL]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2025. <https://doi.org/10.1109/tpami.2025.3598132>.
- [29] Liu J, Wang Q, Fan H, Wang Y, Tang Y, Qu L. Residual Denoising Diffusion Models[J/OL], 2024. <https://doi.org/10.1109/cvpr52733.2024.00268>.
- [30] Zheng D, Wu X, Yang S, Zhang J, Hu J-F, Zheng W-S. Selective Hourglass Mapping for Universal Image Restoration Based on Diffusion Model[J/OL]. arXiv (Cornell University), 2024. <https://doi.org/10.48550/arxiv.2403.11157>.
- [31] Ding Z, Zhang X, Tu Z, Xia Z. Restoration by Generation with Constrained Priors[J/OL], 2024. <https://doi.org/10.1109/cvpr52733.2024.00248>.
- [32] Yang T, Wu R, Ren P, Xie X, Zhang L. Pixel-Aware Stable Diffusion for Realistic Image Super-Resolution and Personalized Stylization[J/OL]. Lecture notes in computer science, 2024. https://doi.org/10.1007/978-3-031-73247-8_5.
- [33] Guo L, Wang C, Yang W, Huang S, Wang Y, Pfister H, Wen B. ShadowDiffusion: When Degradation Prior Meets Diffusion Model for Shadow Removal[J/OL], 2023. <https://doi.org/10.1109/cvpr52729.2023.01350>.
- [34] Yi X, Xu H, Zhang H, Tang L, Ma J. Diff-Retinex: Rethinking Low-light Image Enhancement with A Generative Diffusion Model[J/OL], 2023. <https://doi.org/10.1109/iccv51070.2023.01130>.
- [35] Yi X, Xu H, Zhang H, Tang L, Ma J. Diff-Retinex++: Retinex-Driven Reinforced Diffusion Model for Low-Light Image Enhancement[J/OL]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2025. <https://doi.org/10.1109/tpami.2025.3563612>.
- [36] Luo Z, Gustafsson F, Zhao Z, Sjölund J, Schön T B. Refusion: Enabling Large-Size Realistic Image Restoration with Latent-Space Diffusion Models[J/OL]. arXiv (Cornell University), 2023. <https://doi.org/10.48550/arxiv.2304.08291>.
- [37] Wang Y, Yu J, Yu R, Zhang J. Unlimited-Size Diffusion Restoration[J/OL], 2023. <https://doi.org/10.1109/cvprw59228.2023.00123>.
- [38] Tu L, Wu J, Luo X, Jin Z. Unifying Heterogeneous Degradations: Uncertainty-Aware Diffusion Bridge Model for All-in-One Image Restoration[J/OL]. arXiv (Cornell University), 2026. <https://doi.org/10.48550/arxiv.2601.21592>.
- [39] Luo W, Qin H, Chen Z, Wang L, Zheng D, Li Y, Liu Y, Li B, Hu W. Visual-Instructed Degradation Diffusion for All-in-One Image Restoration[J/OL], 2025. <https://doi.org/10.1109/cvpr52734.2025.01191>.
- [40] Yue Y, Meng Y, Yang L, Yang Y. Joint Conditional Diffusion Model for Image Restoration with Mixed Degradations[J/OL]. arXiv (Cornell University), 2024. <https://doi.org/10.48550/arxiv.2404.07770>.
- [41] Zhu K, Gu J, You Z, Qiao Y, Dong C. An Intelligent Agentic System for Complex Image Restoration Problems[J/OL]. arXiv (Cornell University), 2024. <https://doi.org/10.48550/arxiv.2410.17809>.
- [42] Tripathi M, Kongprawechnon W, Yahia S B. Advancements in Agentic Ai Architecture and Prompting Strategies for Low-Light Enhancement[J/OL]. SSRN Electronic Journal, 2025. <https://doi.org/10.2139/ssrn.5351735>.
- [43] Kirillov A M, Mintun E, Ravi N, Mao H, Rolland C, Gustafson L, Xiao T, Whitehead S, Berg A C, Lo W-Y. Segment Anything[J/OL], 2023. <https://doi.org/10.1109/iccv51070.2023.00371>.

- [44] Chen W, Vong Y-J, Kuo S, Ma S, Wang J. RobustSAM: Segment Anything Robustly on Degraded Images[J/OL]. arXiv (Cornell University), 2024. <https://doi.org/10.48550/arxiv.2406.09627>.
- [45] Ren T, Liu S, Zeng A. Grounded SAM: Assembling Open-World Models for Diverse Visual Tasks[J/OL]. arXiv (Cornell University), 2024. <https://doi.org/10.48550/arxiv.2401.14159>.
- [46] You Z, Gu J, Li Z, Kong X, Dong C. DepictQA: Depicted Image Quality Assessment with Vision Language Models[J]. arXiv preprint, 2023.
- [47] You Z, Li Z, Gu J, Yin Z, Xue T, Dong C. Depicting Beyond Scores: Advancing Image Quality Assessment Through Multi-modal Language Models[J/OL]. Lecture notes in computer science, 2024. https://doi.org/10.1007/978-3-031-72970-6_15.
- [48] Wu H, Zhang Z. Q-Instruct: Improving Low-Level Visual Abilities for Multi-Modality Foundation Models[J/OL], 2024. <https://doi.org/10.1109/cvpr52733.2024.02408>.
- [49] Wu H, Zhang Z, Zhang E. Q-Bench: A Benchmark for General-Purpose Foundation Models on Low-level Vision[J/OL]. arXiv (Cornell University), 2023. <https://doi.org/10.48550/arxiv.2309.14181>.
- [50] Zhang Z, Wu H, Zhang E, Zhai G, Lin W. Q-Bench+: A Benchmark for Multi-Modal Foundation Models on Low-Level Vision From Single Images to Pairs[J/OL]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2024. <https://doi.org/10.1109/tpami.2024.3445770>.
- [51] Ai Y, Zhou X, Huang H, Han X, Chen Z, You Q, Yang H. DreamClear: High-Capacity Real-World Image Restoration with Privacy-Safe Dataset Curation[J/OL]. arXiv (Cornell University), 2024. <https://doi.org/10.48550/arxiv.2410.18666>.
- [52] Luo Z, Gustafsson F, Zhao Z, Sjolund J, Schon T B. Photo-Realistic Image Restoration in the Wild with Controlled Vision-Language Models[J/OL], 2024. <https://doi.org/10.1109/cvprw63382.2024.00658>.
- [53] Talebi H, Milanfar P. NIMA: Neural Image Assessment[J/OL]. IEEE Transactions on Image Processing, 2018. <https://doi.org/10.1109/tip.2018.2831899>.
- [54] Zhang R, Isola P, Efros A A, Shechtman E, Wang O. The Unreasonable Effectiveness of Deep Features as a Perceptual Metric[J/OL]. arXiv (Cornell University), 2018. <https://doi.org/10.48550/arxiv.1801.03924>.
- [55] Zheng H, Yang H, Fu J, Zha Z-J, Luo J. Learning Conditional Knowledge Distillation for Degraded-Reference Image Quality Assessment[C/OL] // 2021 IEEE/CVF International Conference on Computer Vision (ICCV). 2021. <https://doi.org/10.1109/iccv48922.2021.01008>.
- [56] Li J, Yan B, Lin Q, Li A, Ma C. Motion Blur Removal With Quality Assessment Guidance[J/OL]. IEEE Transactions on Multimedia, 2021. <https://doi.org/10.1109/tmm.2021.3068561>.
- [57] Zhang Y, Yan D. Soft Knowledge Distillation with Multi-Dimensional Cross-Net Attention for Image Restoration Models Compression[J/OL], 2025. <https://doi.org/10.1109/icassp49660.2025.10887587>.
- [58] Yang S, Hu B, Wu X, Liu F, Zhou J. Image Restoration Model Compression via Mamba-oriented Heterogeneous Knowledge Distillation[J/OL]. SSRN Electronic Journal, 2025. <https://doi.org/10.2139/ssrn.5407564>.
- [59] Wang P, Luo X, Xie Y, Qu Y. Data-free Distillation with Degradation-prompt Diffusion for Multi-weather Image Restoration[J/OL]. arXiv (Cornell University), 2024. <https://doi.org/10.48550/arxiv.2409.03455>.
- [60] Zhai L, Wang Y, Cui S, Zhou Y. A Comprehensive Review of Deep Learning-Based Real-World Image Restoration[J/OL]. IEEE Access, 2023. <https://doi.org/10.1109/access.2023.3250616>.
- [61] Guan Q, Yang Q, Chen X, Song T, Jin G, Jin J. WeatherBench: A Real-World Benchmark Dataset for All-in-One Adverse Weather Image Restoration[J/OL], 2025. <https://doi.org/10.1145/3746027.3758196>.