中图法分类号:TP391;TP18　　**文献标识码**:A　　**文章编号**:1006-8961(2026)09-3077-20

论文引用格式:Zheng M M, Guo C R, Zhang L and Wei W. 2026. Aerial image restoration driven by large models: a survey. Journal of Image and Graphics, 31(9):3077-3096(郑萌萌, 郭宸瑞, 张磊, 魏巍. 2026. 大模型赋能下的航空图像复原技术研究进展. 中国图象图形学报, 31(9):3077-3096)[DOI:10.11834/jig.260208]

大模型赋能下的航空图像复原技术研究进展

郑萌萌[1],郭宸瑞[2,3*],张磊[1],魏巍[1]

1. 西北工业大学计算机学院,西安　710129; 2. 宇航智能控制技术全国重点实验室,北京　100854;
3. 北京航天自动控制研究所,北京　100854

摘　要:航空图像复原作为低层视觉任务的重要研究方向,对于提升遥感监测与空天信息处理的精度具有重要意义。针对航空图像中普遍存在的噪声、模糊、低光及多退化耦合等问题,本文系统综述了近年来图像复原技术的研究进展。首先,从传统方法与深度学习方法两个层面,梳理了典型单一退化复原技术的发展脉络;进一步,围绕统一建模思想,总结了多退化联合建模方法,包括基于多退化建模、提示学习以及动态路由与专家模型的复原框架。在此基础上,重点分析了大模型驱动的图像复原方法,从视觉基础模型、多模态大模型以及基于智能体的复原方法3个方面展开讨论。视觉基础模型通过大规模预训练提供统一特征表示,多模态大模型通过跨模态语义建模提升复杂场景理解能力,而智能体方法则通过引入决策机制实现复原过程的动态优化。同时,本文进一步归纳了现有方法在多退化耦合建模、语义信息利用以及动态自适应机制等方面面临的关键挑战。最后,结合航空图像的应用特点,对未来发展方向进行了展望,包括多源数据构建、统一表征学习、多模态协同建模及智能化复原系统设计等。本文系统梳理了大模型驱动下航空图像复原的方法体系,并归纳出多退化耦合、语义—视觉协同、物理先验融合与动态决策优化等核心问题,可为后续航空图像智能复原研究提供系统性参考与发展思路。本文提及的数据集已汇总至https://github.com/MengmengZheng/Aerial-Image-Restoration-Datasets。

关键词:航空图像复原;计算机视觉(CV);提示学习;统一建模;多源协同;智能体

Aerial image restoration driven by large models: a survey

Zheng Mengmeng[1], Guo Chenrui[2,3*], Zhang Lei[1], Wei Wei[1]

1. School of Computer Science, Northwestern Polytechnical University, Xi'an　710129, China; 2. State Key Laboratory of Aerospace Intelligent Control Technology, Beijing　100854, China; 3. Beijing Aerospace Automatic Control Institute, Beijing　100854, China

Abstract: Aerial image restoration is an important low-level vision task for improving the quality and reliability of remote sensing and aerospace imagery. Compared with general natural image restoration, aerial image restoration faces more complex imaging conditions caused by several factors, including long-distance observation, platform motion, atmospheric interference, illumination variation, and sensor limitations. However, these factors often lead to diverse and coupled degradations, which not only reduce visual quality but also affect downstream tasks, such as object detection, semantic segmentation, change detection (CD), scene understanding, and autonomous perception. Therefore, robust aerial image restoration has become a key problem in both computer vision (CV) and aerospace information processing. In recent years, image restoration has evolved from task-specific models for individual degradations to unified restoration frameworks and large

收稿日期:2026-04-16;**修回日期**:2026-06-05;**预印本日期**:2026-06-12

*通信作者:郭宸瑞　g-ch-r@163.com

基金项目:国家自然科学基金项目(62372379, 62472359)

Supported by: National Natural Science Foundation of China(62372379, 62472359)

model-driven paradigms. To date, existing surveys mainly focus on general natural image restoration, specific low-level vision tasks, or individual degradation problems. In contrast, the present survey focuses on aerial image restoration from the perspective of large model empowerment. It also emphasizes the unique imaging characteristics and practical requirements of aerial scenarios, as well as systematically reviews the technical evolution from single-degradation restoration to unified modeling, and further to large model-driven intelligent restoration. In particular, this paper highlights the contributions of large-scale pretraining, multimodal representation learning, physical priors, and agent-based decision mechanisms to robust aerial image restoration under complex degradation conditions. First, this paper reviews representative restoration methods for typical single-degradation tasks and summarizes their basic modeling ideas, advantages, and limitations. Traditional methods usually rely on handcrafted priors, optimization models, and physical assumptions, all of which provide clear interpretability but are often limited in complex real-world scenarios. While deep learning-based methods improve restoration performance by learning nonlinear mappings from degraded images to clean images, many of them are still designed for specific degradation types and may lack generalization ability when facing unknown or mixed degradations in aerial scenes. To solve this problem, this paper discusses unified image restoration methods for complex aerial imaging scenarios. In practical applications, aerial image degradation is rarely caused by a single factor. Different degradation factors may coexist and interact with each other, resulting in complex coupled degradation and spatially non-uniform degradation distributions. As such, unified restoration methods aim to handle multiple degradations within a single framework. To date, existing studies have mainly explored multidegradation modeling, prompt-based restoration, and dynamic routing or expert-based frameworks. These methods improve the adaptability of restoration models by learning shared representations, introducing degradation-aware prompts, or dynamically selecting suitable restoration paths. However, challenges remain in terms of explicitly modeling degradation mechanisms, distinguishing different degradation patterns, and dealing with spatially varying degradations. On this basis, this survey focuses on large model-driven aerial image restoration. Vision foundation models provide transferable visual representations and structural priors through large-scale pretraining, which can improve restoration robustness under data-limited and cross-scene conditions. Multimodal large models introduce semantic and cross-modal information into the restoration process, thereby enabling models to better understand complex scenes and provide high-level guidance for structure recovery. The combination of physical models and large models provides a promising hybrid paradigm in which physical priors derived from imaging mechanisms constrain the restoration process, while large models provide strong representation and generation capabilities. Meanwhile, agent-based restoration methods extend image restoration from a static mapping process to a dynamic decision-making process, in which degradation analysis, strategy planning, model selection, feedback evaluation, and iterative optimization can be integrated into a unified restoration pipeline. Furthermore, this paper analyzes the main challenges faced by existing aerial image restoration methods. First, the modeling of coupled degradations remains insufficient, especially when multiple degradation factors interact in complex imaging environments. Second, existing collaborations between semantic priors and low-level visual features remain limited, making it difficult to achieve reliable structure recovery under severe degradation. Third, the integration of physical knowledge and data-driven large models remains underexplored, and balancing physical consistency with generative flexibility is still an open problem. Fourth, most existing methods lack dynamic adaptive mechanisms and are unable to adjust restoration strategies in accordance with image content, degradation state, or task requirements. Finally, the high computational cost of large model-driven methods restricts their deployment on airborne platforms, edge devices, and real-time remote sensing systems. Overall, aerial image restoration is moving from degradation-specific and pixel-level processing toward unified modeling, multimodal collaboration, physical prior-guided restoration, and intelligent decision-making. In summary, this survey aims to provide a systematic reference for researchers by clarifying the relationship between traditional restoration methods, unified restoration frameworks, and large model-driven paradigms. It also discusses future research directions, including multisource data construction, unified representation learning, multimodal and physical prior-guided restoration, adaptive decision-making, and efficient deployment, all of which are driven by the goal of promoting robust, interpretable, and intelligent aerial image restoration systems.

Key words: aerial image restoration; computer vision (CV); prompt-based learning; unified modeling; multi-source fusion; agent-based framework

0 引　言

在国家持续推进航空领域战略布局并大力发展空天信息产业的背景下，航空图像已逐渐成为空天信息系统中的重要数据来源之一，并在遥感监测、灾害评估、生态环境保护、城市规划及自主飞行等多个领域发挥着关键支撑作用。在这些应用中，图像质量是影响后续视觉感知任务准确性与可靠性的关键因素。然而，如图 1 所示，受远距离成像及复杂气象条件的影响，航空图像在采集过程中不可避免地受到多种退化因素的干扰，包括器件辐射引入的噪声干扰、复杂光照条件下的高动态范围失真、平台运动造成的运动模糊，以及弱光场景下的信噪比显著降低等。这些因素往往相互耦合，使退化过程更加复杂且具有较强不确定性。在其综合影响下，图像质量显著下降，且在细节表达、结构清晰度和视觉一致性等方面存在明显不足，进而严重制约了目标检测、语义分割及三维重建等关键高层视觉任务的性能。因此，如何有效提升航空图像质量，增强视觉信息的可辨性与稳定性，是航空视觉感知领域亟需解决的关键问题之一。

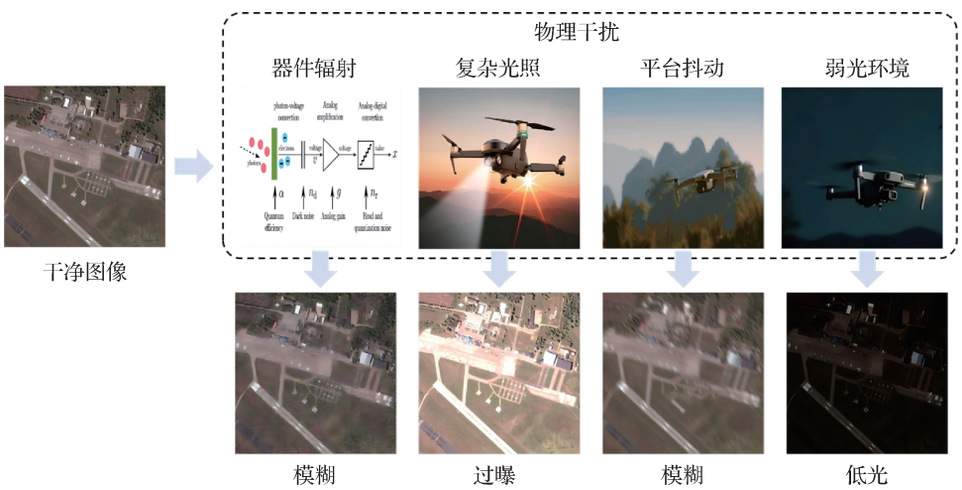


图 1　航空图像退化现象示意图

Fig. 1　Illustration of various degradation effects in aerial images

图像复原（image restoration，IR）作为计算机视觉领域的重要研究方向，其核心目标在于从退化观测中恢复图像的结构信息与视觉质量，从而为后续高层视觉感知任务提供可靠的数据基础。随着深度学习的快速发展，图像复原技术逐步由依赖人工设计先验的传统方法，演进为以数据驱动为主的深度学习方法。在这一发展过程中，基于卷积神经网络（convolutional neural network，CNN）和 Transformer 的模型逐渐成为图像复原领域的主流技术，这类方法通过从大规模数据中学习图像退化与恢复之间的映射关系，从而实现对退化图像的有效复原，并在多种复原任务中表现出优越性能。针对不同退化类型的图像复原任务，研究者对特定的退化进行建模，提出了大量的复原方法，并在对应任务中展现了良好的性能。然而，在真实场景中，图像退化的表现常呈现多重叠加的状态，在面对多种退化因素耦合的复杂场景时，面向单一退化设计的图像复原方法往往难以准确建模不同退化之间的耦合关系，从而导致图像细节重建产生偏差与结构失真。因此，一些研究开始探索基于统一建模的图像复原方法，以提升模型面对复杂退化场景的适应能力。

为应对多重退化耦合问题，研究者引入统一图像复原（all-in-one image restoration，AiOIR）框架，使单一模型能够在统一建模下处理多种退化类型。AiOIR 框架可同时应对多种单一或混合退化任务（如去噪、去模糊、低光复原及高动态范围重建等），且无需针对特定任务对模型进行单独训练。该类方法通过引入条件控制、退化感知建模及共享表示学习等机制，以增强模型对复杂退化场景的适应能力。在此基础上，研究者逐步将相关方法拓展至航空图像场景中，以应对其更为复杂的退化特性与应用需求。

然而，现有统一建模方法仍存在一定局限。一方面，其通常依赖共享特征空间，缺乏对不同退化机理的显式刻画，易导致特征混淆，从而影响复原效果；另一方面，这类方法多基于像素级映射，缺乏对高层语义信息的有效利用，在严重退化或结构缺失条件下难以实现稳定的结构恢复。此外，大多数方法仍基于单次前向推理的静态处理方式，缺乏对复杂退化场景的动态适应能力，限制了其在真实环境中的应用效果。

不同于传统方法直接在像素空间学习退化到清晰图像的映射，大模型能够通过预训练获得具有结构与语义信息的高层表征，从而在复原过程中对图像结构进行有效约束，缓解复杂退化条件下细节恢复不稳定的问题。在此基础上，多模态大模型通过融合图像、文本及其他传感器信息（如红外与深度数据），能够为退化区域提供补充信息，提升在弱光、遮挡及大气干扰等场景中的复原能力。此外，基于语义信息的引导机制使模型能够对目标结构进行合理约束，在细节重建过程中避免结构失真。基于此，航空图像复原逐步突破传统像素级恢复的范式，向融合多源信息与语义引导的方向发展。

总体而言，航空图像复原在计算机视觉与空天信息处理领域具有重要的研究价值与广阔的应用前景。围绕复杂退化建模、统一复原框架以及大模型驱动方法的深入研究，有助于进一步提升航空图像复原的性能与适用范围。未来研究可在多退化协同建模、多源信息融合及语义引导等方向持续探索，以充分挖掘航空图像数据的潜在价值，为智能感知与智慧航空等应用提供更加可靠的技术支撑。

本文首先介绍面向单一退化的航空图像复原方法，重点分析去噪、去模糊、低光复原及高动态范围重建等典型任务的建模思路与方法特点；接着，详细阐述面向统一建模的航空图像复原方法，重点讨论统一图像复原框架在多退化耦合场景中的应用优势及其关键技术；随后，进一步探讨大模型驱动的航空图像复原方法，从视觉基础模型、多模态大模型以及语义引导等角度对相关研究进行系统梳理；在此基础上，对现有方法的关键挑战与局限性进行归纳分析；最后，总结航空图像复原领域的研究进展，并展望未来发展方向和潜在研究热点。与现有图像复原综述相比，本文聚焦于航空图像复原应用场景及大模型相关技术，系统梳理单一退化复原、统一建模复原以及大模型驱动复原方法的发展脉络，并进一步分析视觉基础模型、多模态大模型、物理先验和智能体机制在航空图像复原中的作用，旨在为大模型驱动的航空图像复原研究提供参考与借鉴，推动相关技术的进一步发展与应用。

1 面向单一退化的航空图像复原方法

受复杂成像环境与飞行平台运动变化的影响，航空图像在采集过程中易产生噪声、模糊及光照不足等多种退化，从而降低图像质量与可用性。由于不同退化在形成机理与视觉表现上存在差异，研究者通常针对具体退化类型构建相应的复原模型，通过建模退化过程实现图像质量的有效恢复。如表1所示，本小节围绕典型单一退化问题，对航空图像去噪、去模糊、低光复原及高动态范围重建等方法进行介绍。与此同时，为进一步梳理相关研究的实验支撑，表2汇总了面向航空/遥感图像复原的数据集，并从主要任务、配对关系和数据规模等方面进行归纳。

1.1 图像去噪

在航空成像过程中，受传感器噪声、远距离传输及复杂环境干扰的影响，图像往往受到高斯噪声、泊松噪声及空间非均匀噪声的共同作用。这类噪声不仅会引入随机强度扰动，降低图像的信噪比，还会破坏局部纹理结构与边缘信息，导致细节模糊及结构不连续，从而削弱特征提取的稳定性与判别能力，并进一步影响目标检测与语义理解等高层视觉任务的性能。因此，如何在有效抑制噪声的同时保持图像结构与细节信息，成为航空图像去噪的核心问题。

从退化建模角度看，图像去噪通常可表示为

$$\boldsymbol{y}=\boldsymbol{x}+\boldsymbol{n} \tag{1}$$

式中，$\boldsymbol{x}$为清晰图像，$\boldsymbol{y}$为观测图像，$\boldsymbol{n}$为成像过程中引入的噪声。该模型刻画了噪声对原始图像的叠加干扰关系，是图像去噪任务中常用的基础退化表达。

传统方法通常基于噪声与信号独立且服从固定分布的假设对图像进行建模。例如，块匹配与三维滤波（block-matching and 3D Filtering，BM3D）方法（Dabov等，2007）利用图像的非局部自相似性，在变换域中对相似块进行协同滤波，从而实现高效去噪；K奇异值分解（K-singular value decomposition，K-SVD）方法（Aharon等，2006）基于稀疏表示理论，通

表 1　面向单一退化的航空图像复原方法

Table 1　Representative methods for single-degradation aerial image restoration

退化任务	年份	代表方法	核心思想
图像去噪	2017	DnCNN	残差学习显式建模噪声分布，将去噪任务转化为噪声估计问题
	2018	FFDNet	引入噪声水平图作为条件输入，将噪声强度显式嵌入网络，实现对不同强度噪声的自适应建模
图像去模糊	2018	DeblurGAN	利用生成对抗学习约束复原图像分布，使模型在恢复结构清晰度的同时增强视觉真实感
	2018	SRN-DeblurNet	通过多尺度递归结构逐层恢复不同尺度的模糊信息，实现从粗到细的渐进式结构重建
图像低光增强	2018	deep RetinexNet	将图像分解为照明与反射分量，并通过联合优化实现亮度提升与结构保持的协同建模
	2020	Zero-DCE	通过学习像素级非线性映射曲线对亮度进行自适应调整，在无需成对监督数据的情况下实现稳定增强
高动态范围重建	2017	HDRCNN	通过学习低动态范围到高动态范围的非线性映射，对过曝与欠曝区域进行分区域重建
	2021	DeepHDR	引入注意力机制对不同曝光区域进行自适应加权融合，从而提升亮度与细节的协调恢复能力

表 2　面向航空/遥感图像复原的数据集

Table 2　Datasets for aerial/remote sensing image restoration

数据集	主要任务	是否配对	数据规模
MDRS-Landsat	去噪、去模糊、去雾、低光增强	是，合成配对	基于 5 500 幅高质量 Landsat-8 图像构建，包含模糊、噪声、雾、低光照等退化，划分为 5 130 个训练、100 个验证、270 个测试样本
UAVIR-5D	去噪、去模糊、去雾、低光增强、去雨滴	是，合成或构建配对	面向 UAV 图像复原的多退化数据集，覆盖城市道路、建筑群、校园等航拍场景，包含模糊、噪声、雾、低光照和雨滴 5 类退化
iSAID-dark	低光增强	是，合成配对	由 iSAID 测试集中 751 幅高分辨率航空图像裁剪生成，包含 3 755 对训练图像和 66 对验证图像
U3D	超高分辨率、UAV 低光增强	否	1 000 幅 UHR (3 840 × 2 160 像素) 低光航拍图像和 1 000 幅非成对的正常光日间航拍图像

过字典学习实现图像重构；全变分（total variation，TV）模型（Rudin 等，1992）通过对图像梯度进行正则化约束，在抑制噪声的同时保留边缘结构。尽管上述方法在理想噪声条件下表现良好，但由于其依赖固定先验，难以刻画真实航空图像中信号相关与空间非均匀的噪声特性，在复杂航空场景中的适应能力有限。

为克服传统方法在复杂噪声建模方面的不足，研究者逐渐转向数据驱动的去噪方法。基于卷积神经网络的方法能够利用局部感受野与参数共享机制，有效建模图像的空间结构特征，从而在抑制噪声的同时保留边缘与纹理信息。例如，去噪卷积神经网络（denoising convolutional neural network，DnCNN）（Zhang 等，2017）通过残差学习策略显式建模噪声分量，使模型专注于噪声特征的学习；快速灵活去噪网络（fast and flexible denoising network，FFDNet）（Zhang 等，2018）通过引入噪声水平图，将噪声强度作为条件输入，实现对不同噪声水平的自适应处理。这类方法在不同类型噪声条件下均表现出较好的性能，因此，基于卷积神经网络的方法成为图像去噪领域的主流技术路线。在此基础上，研究者进一步引入生成式建模方法以提升复原图像的细节表达能力。例如，生成对抗网络（generative adversarial network，GAN）（Goodfellow 等，2014）通过对抗学习机制

约束生成结果的分布，从而缓解传统方法中的过平滑问题；扩散模型（diffusion model，DM）（Ho等，2020）通过逐步去噪的生成过程，对复杂噪声分布进行建模，使模型能够在多个尺度上恢复图像细节，在高质量图像复原任务中展现出较强能力。

针对航空与遥感图像中不同传感器所引入的差异性噪声特性，研究者在通用模型基础上结合具体成像机制进行改进。例如，Zhang等人（2017）通过构建基于深度残差网络的遥感图像去噪方法，增强模型对复杂背景纹理的表达能力，从而提升了模型在复杂遥感场景中的去噪鲁棒性；Chierchia等人（2017）针对合成孔径雷达（synthetic aperture radar，SAR）图像中的乘性斑点噪声，设计基于卷积神经网络的去斑方法，实现了对该类噪声的有效抑制。此外，通过融合光学与红外等多模态信息，可以利用不同传感器之间的互补特性辅助结构恢复，从而在复杂场景中进一步提升去噪效果。

1.2 图像去模糊

航空图像中的模糊主要来源于飞行平台运动、成像抖动以及大气扰动等因素，通常表现为空间结构的扩散与边缘信息的退化。从成像机理上看，图像模糊由成像系统响应与平台运动共同作用产生，其中模糊核反映了相机运动轨迹及能量扩散过程。在航空场景中，模糊核通常具有明显的空间变化特性，从而增加了模糊核估计的复杂性。因此，如何在复杂动态条件下准确建模模糊过程并恢复图像结构，是图像去模糊研究的核心问题。

经典去模糊退化模型可写为

$$\boldsymbol{y}=\boldsymbol{k}\otimes\boldsymbol{x}+\boldsymbol{n} \tag{2}$$

式中，$\boldsymbol{x}$表示清晰图像，$\boldsymbol{y}$表示模糊观测图像，$\boldsymbol{k}$表示模糊核，$\otimes$表示卷积运算，$\boldsymbol{n}$表示噪声项。图像去模糊的目标是在模糊核未知或不准确的情况下，从观测图像$\boldsymbol{y}$中恢复清晰图像$\boldsymbol{x}$。

早期研究主要围绕模糊核估计与反卷积求解展开。Fergus等人（2006）从概率建模角度出发，构建盲去模糊的贝叶斯推断框架，并利用自然图像梯度的稀疏先验，对清晰图像与模糊核进行联合估计，从而实现单幅图像的模糊恢复。在此基础上，Levin等人（2007）进一步分析了模糊核估计中的不适定性问题，即仅依赖最大后验估计容易产生退化解，并通过引入稀疏梯度先验约束图像结构，从而提高反卷积过程的稳定性与恢复结果的可靠性。这类方法在均匀模糊条件下具有一定效果，但由于依赖显式模糊核的建模与估计，在空间变化的模糊场景中往往难以获得稳定的恢复结果。

针对显式模糊核建模在复杂场景中的局限性，研究者逐渐采用端到端建模策略，对模糊退化过程进行隐式表达，使模型能够直接恢复退化图像中的结构与细节信息（吴迪 等，2020）。Kupyn等人（2018）通过构建DeblurGAN模型，引入对抗学习机制约束复原图像分布，使恢复结果在提升清晰度的同时保持较好的视觉真实感；Tao等人（2018）提出尺度递归网络（SRN-DeblurNet），通过多尺度递归结构逐步恢复不同尺度的模糊细节，从而提升模型对复杂运动模糊的处理能力。与传统方法相比，这类方法避免了对模糊核的显式估计，在复杂动态场景中表现出更强的适应能力。

进一步地，针对动态模糊中普遍存在的方向性与空间变化特性，研究者开始引入基于自注意力机制的模型结构。例如，Tsai等人（2022）提出Stripformer模型，通过构建条带式注意力机制，在不同方向上对模糊特征进行建模，从而有效捕获具有方向性的运动模糊特征。

在航空成像场景中，模糊通常与平台运动轨迹及成像几何关系密切相关，并呈现明显的非均匀特性。针对这一特点，Pan等人（2016）通过引入暗通道先验约束图像结构，并结合优化策略实现去模糊，从而在复杂场景下获得更优恢复效果；Su等人（2017）通过构建多帧融合模型，利用视频序列中的时序信息进行联合恢复，从而在动态模糊条件下显著提升去模糊性能。此外，一些研究进一步结合运动估计与物理成像模型，对模糊形成过程进行显式建模，以提升复原精度与稳定性。

1.3 图像低光增强

低光环境会导致图像亮度不足、对比度下降以及细节信息丢失，是航空夜间成像及弱光条件下的典型退化问题。在航空场景中，由于成像距离远、光照条件复杂以及大气散射效应的影响，低光退化往往与噪声干扰及成像模糊等因素相互耦合，从而进一步增加了图像复原的难度。因此，如何在提升图像亮度的同时保持结构信息与视觉真实性，成为低光增强研究的核心问题。

低光增强常借助Retinex模型描述图像形成过程，即

$$\boldsymbol{I}=\boldsymbol{R}\odot\boldsymbol{L}+\boldsymbol{n} \tag{3}$$

式中，$\boldsymbol{I}$为观测图像，$\boldsymbol{R}$为反射分量，$\boldsymbol{L}$为照明分量，$\odot$表示逐像素乘法，$\boldsymbol{n}$为噪声项。低光增强的目标是在提升$\boldsymbol{L}$的同时保持$\boldsymbol{R}$的结构稳定，并抑制亮度提升过程中被放大的噪声。

传统低光增强方法多基于Retinex理论，将图像显式分解为反射分量与照明分量，并通过对照明分量进行调节实现亮度提升；此外，直方图均衡与伽马校正等方法通过对像素分布进行全局或非线性变换，在一定程度上改善图像亮度与对比度。然而，这类方法通常依赖简单的光照假设，难以刻画复杂场景中的非均匀光照变化，容易引入过度增强、细节丢失及色彩失真等问题。

针对上述不足，研究者引入深度学习方法对低光退化过程进行建模。例如，Wei等人（2018）提出基于Retinex的深度分解网络（deep RetinexNet），通过联合优化分解与增强过程，使模型在提升亮度的同时保持结构信息；在此基础上，一些方法（Zhang等，2019；Zhang等，2021）进一步通过改进分解策略或引入约束机制，提高对复杂光照条件的适应能力。

除显式分解策略外，研究者还探索直接对光照调整过程进行建模的方法。Guo等人（2020）提出零参考深度曲线估计方法（zero-reference deep curve estimation，Zero-DCE），通过构建深度曲线估计网络，对像素亮度进行逐点自适应调整，在无需成对数据的条件下实现低光增强；该方法通过设计可微分的曲线映射函数，使模型能够在保证稳定性的同时实现灵活的亮度调节。此外，针对低光增强任务中普遍存在的细节恢复困难问题，一些研究引入多尺度建模机制，通过在不同尺度上建模光照分布与结构信息，从而增强对暗区细节的恢复能力。

针对低光增强过程中普遍存在的细节失真与分布偏移问题，研究者进一步设计了更具针对性的模型结构。Wang等人（2019）提出基于归一化流（normalizing flow）的低光增强模型（low-light flow，LLFlow），对低光图像与正常光图像之间的分布关系进行建模，使模型能够从概率分布角度恢复图像细节，从而缓解传统方法中细节重建不真实的问题。与此同时，Lyu等人（2021）从特征选择角度出发，提出基于注意力引导的低光增强模型，在亮度调整过程中引入空间注意力机制，使模型能够自适应聚焦于结构信息较为重要的区域，从而避免对无效区域的过度增强，并提升结构一致性。

在航空与遥感场景中，低光退化不仅表现为整体亮度不足，还常伴随显著的光照非均匀分布，使得暗区细节难以有效恢复。与一般图像增强问题不同，低光增强需要在提升亮度的同时避免对已有结构信息的破坏，因此对光照分布的精确建模尤为关键。此外，由于暗区域信号较弱，亮度提升过程容易放大潜在噪声并引入伪影，使得增强结果在视觉上出现不自然现象。因此，如何在复杂光照条件下实现对照明分量的合理估计，并在增强过程中保持结构一致性与视觉自然性，成为航空低光图像增强中的核心挑战。

1.4 高动态范围重建

由于成像设备动态范围有限，在高对比度场景中难以同时记录暗区与亮区的有效信息，航空图像常出现过曝与欠曝并存的现象，导致细节信息丢失与结构退化。高动态范围（high dynamic range，HDR）重建的目标在于恢复更完整的亮度分布，使图像能够同时呈现高光与阴影区域的细节信息，从而提升视觉质量与信息可用性。

根据建模形式的不同，HDR重建方法主要分为多曝光融合方法与单幅图像重建方法两类。传统方法以多曝光融合为主，通过对不同曝光图像进行加权组合实现动态范围扩展。例如，Mertens等人（2007）提出曝光融合方法，通过对局部对比度、饱和度及曝光度进行加权，该方法在无需显式构建HDR图像的情况下实现视觉增强。然而，该类方法依赖多帧输入，在航空动态场景中由于平台运动及场景变化，图像之间难以实现精确对齐，容易引入融合伪影。

在深度学习的单幅图像HDR重建方法中，Eilertsen等人（2017）提出高动态范围重建卷积神经网络（high dynamic range convolutional neural network，HDRCNN），通过对过曝区域与欠曝区域进行分区域建模，恢复缺失的亮度信息，从而在仅输入单幅低动态范围图像的条件下实现动态范围扩展，有效避免了多帧方法中的对齐问题。多曝光HDR重建方法中，Kalantari和Ramamoorthi（2017）提出动态场景高动态范围重建网络，通过在特征层引入对齐模块，对不同曝光图像进行显式对齐，并结合融合网络实现亮度信息的统一重建，从而有效缓解动态场景中的运动伪影问题。

在航空与遥感场景中，HDR重建面临更加复杂的成像条件。一方面，平台运动与视角变化使得多曝光图像之间的空间对应关系更加复杂；另一方面，大气散射与光照变化会进一步加剧亮度分布的不稳定性。Wei等人(2020)针对遥感图像中亮度分布复杂及动态范围差异显著的问题，设计了基于深度学习的HDR重建方法DeepHDR，通过在特征层引入自适应融合机制，对不同曝光信息进行加权整合，从而提升复杂场景下的亮度恢复能力与细节表现；Zheng等人(2021)针对无人机(unmanned aerial vehicle, UAV)图像设计HDR重建模型，通过改进对齐与融合机制，有效减少动态场景中的伪影。此外，基于RAW数据的方法由于保留了更高比特深度与线性响应特性，使模型能够更加准确地刻画亮度信息，在高动态范围重建中展现出显著优势。

2 面向统一建模的航空图像复原方法

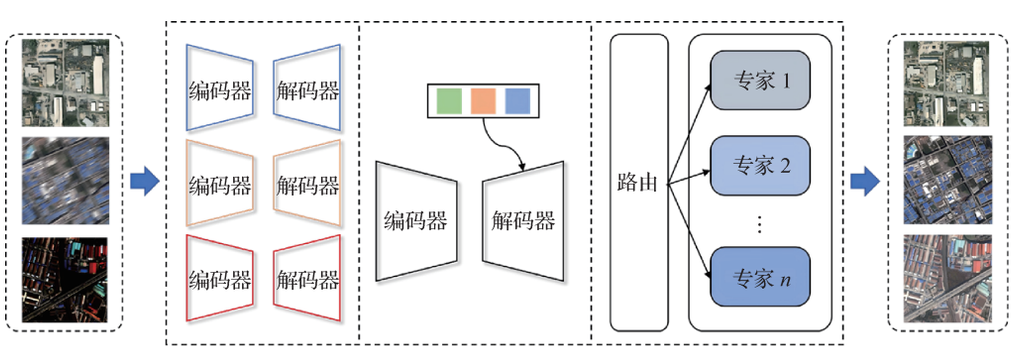
在实际航空成像过程中，图像退化往往并非由单一因素引起，而是由噪声、模糊以及光照变化等多种退化共同作用并相互耦合形成。这类复合退化不仅使图像质量显著下降，也使得不同退化之间呈现出复杂的交互关系，从而增加了复原建模的难度。针对单一退化设计的图像复原方法通常依赖特定退化假设，难以适应复杂场景中多退化叠加的情况，容易在实际应用中出现退化类型不匹配或恢复结果不稳定的问题。

为应对多退化耦合带来的挑战，研究者逐渐从“针对单一退化建模”转向“统一建模多退化”的研究范式，提出统一图像复原方法。该类方法通过构建单一模型对多种退化类型进行联合建模，使模型能够在不同退化条件下自适应调整复原策略，从而提升在复杂航空场景中的稳定性与适应能力。

统一图像复原的关键在于如何在同一框架下同时刻画不同退化的共性特征与差异特性，使模型既具备跨退化的共享表达能力，又能够针对特定退化进行有效区分与处理。围绕这一目标，现有研究从不同角度提出了多种建模策略。根据实现机制的不同，如图2所示，相关方法大致可分为3类：基于多退化建模的方法、基于提示学习的方法，以及基于动态路由与专家模型的方法。具体而言，图2(a)表示基于多退化建模的统一复原方法，主要通过共享主干网络、退化表示学习或条件编码等方式，在同一模型中学习不同退化之间的共性与差异。图2(b)表示基于提示学习的方法，通过可学习提示、任务提示或退化提示对复原过程进行条件调制，使模型能够根据不同退化状态调整复原策略。图2(c)表示基于动态路由与专家模型的方法，通过路由机制为不同退化类型、退化程度或空间区域动态选择合适的专家模块，从而提升模型对复杂退化和空间非均匀退化的适应能力。表3展示了面向统一建模下航空图像复原的代表方法。

2.1 基于多退化建模的统一图像复原方法

基于多退化建模的统一图像复原方法旨在通过同一模型对多种退化类型进行联合建模，使模型能够在统一框架下处理不同类型的图像退化问题。这类方法的核心在于刻画不同退化之间的共性特征与差异特性，使模型在共享表示的基础上实现对特定



(a) 基于多退化建模的架构　(b) 基于提示学习的架构　(c) 基于专家模型的架构

图2　面向统一建模的航空图像复原方法

Fig. 2　Overview of unified modeling framework for aerial image restoration

((a) multi-degradation modeling-based architecture; (b) prompt learning-based architecture; (c) expert-based architecture)

表3　面向统一建模的航空图像复原代表性方法

Table 3　Representative methods for unified modeling-based aerial image restoration

方法	年份	核心思想	适用退化类型	模型复杂度/推理效率
AirNet	2022	通过退化表示学习实现未知退化条件下的统一复原	去噪、去雨、去雾	中等复杂度，推理效率较高
TransWeather	2022	采用统一Transformer框架和天气查询建模多种恶劣天气退化	去雨、去雾、去雪	中等复杂度，推理效率一般
PromptIR	2023	通过可学习提示对不同退化类型进行条件调制	去噪、去雨、去雾	中等复杂度，推理效率较高
DA-CLIP	2024	通过退化感知视觉语言表征统一描述内容与退化信息	去噪、去雨、去雾、去模糊、低光增强	复杂度较高，推理效率一般
InstructIR	2024	利用自然语言指令统一表达不同复原任务需求，引导模型完成多任务复原	去噪、去雨、去模糊、去雾、低光增强	中等复杂度，推理效率一般
MoCE-IR	2025	通过专家路由机制为不同退化类型或不同退化程度动态选择复原路径	去雾、去雨、去噪、去模糊、低光增强	复杂度较高，推理效率一般

退化的自适应处理。根据退化信息的建模方式不同，相关方法主要可分为基于混合数据驱动的隐式建模、基于退化感知的条件建模以及基于结构设计的显式建模3类。

在隐式多退化建模方面，早期统一复原方法通常通过混合多种退化数据进行联合训练，使模型在统一特征空间中学习跨退化的通用表示。例如，Zamir等人(2022)提出高分辨率图像复原模型，通过在去噪、去模糊、去雨等多种任务上进行联合训练，使模型具备较强的跨退化复原能力；Wang等人(2022)提出U形Transformer图像复原网络(U-shaped Transformer network，Uformer)，通过多尺度编码—解码结构与局部自注意力机制，在统一框架下建模不同退化条件下的图像特征。这类方法通过共享模型参数实现多任务统一处理，但由于缺乏显式退化建模，在退化差异较大时容易出现特征混淆问题。

为增强模型对不同退化类型的区分能力，研究者进一步引入退化感知机制，通过显式建模退化信息引导复原过程。例如，Li等人(2022)提出统一图像复原网络，通过学习退化表示并将其编码为条件特征输入网络，使模型能够根据不同退化类型自适应调整复原策略。此外，一些研究通过引入退化标签或退化编码向量，将不同退化类型映射到特征空间，从而在统一模型中实现多任务协同建模。这类方法通过显式条件约束，有效增强了模型对多退化的区分能力，并缓解了特征混淆问题。

另外，为同时建模退化的共性与差异性，部分研究从网络结构设计角度出发，引入多分支或模块化结构进行显式建模。例如，Zamir等人(2021)提出多阶段渐进式图像复原网络(multi-stage progressive image restoration network，MPRNet)，通过分阶段结构逐步恢复图像细节，并在多退化任务中实现性能提升；一些方法进一步通过设计任务特定分支以提取差异特征，同时共享主干网络以学习通用表示，并通过特征融合实现统一复原，从而在模型表达能力与参数效率之间取得平衡。这类结构化方法通过解耦不同退化特征，提高了模型对复杂退化的适应能力。

在航空与遥感图像场景中，多退化问题往往更加复杂，例如大气散射、传感器噪声与运动模糊之间的耦合作用，使得退化过程呈现出更强的不确定性。针对这一特点，研究者在统一建模框架中进一步引入遥感先验或成像约束，使模型能够更好地刻画真实退化过程。例如，一些研究通过结合大气散射模型或成像几何关系，对退化过程进行约束，从而提升模型在复杂环境下的稳定性与泛化能力。

2.2　基于提示学习的统一图像复原方法

提示学习(prompt learning)作为统一建模的重要技术路径，其核心思想在于通过引入可学习或预定义的提示信息，引导模型感知不同任务或退化类型，从而实现统一框架下的多任务适配。在图像复

原任务中，提示通常作为条件信息嵌入网络结构，通过调制特征表示或控制模型行为，使模型能够在共享参数的基础上实现对不同退化的自适应处理。

从提示信息的表达形式来看，现有方法主要可分为视觉提示、文本提示以及多模态提示3类。

在视觉提示方法中，提示通常以可学习的向量、特征块或退化编码的形式嵌入网络，并与图像特征在不同层级进行融合。这类方法通过在特征空间中引入退化相关信息，使模型能够根据不同退化条件自适应调整特征表示，从而实现对多种退化的统一建模。例如，Potlapalli等人（2023）提出PromptIR，通过引入可学习的退化提示向量，并在多尺度特征建模过程中对特征进行调制，使模型能够在统一框架下处理多种图像复原任务。视觉提示方法具有结构简单、参数开销小的特点，是当前统一复原中应用最广泛的提示形式。

在文本提示方法中，退化类型或任务需求通过离散标签或语义描述进行表示，并作为条件信息输入模型。与视觉提示相比，文本提示能够提供更加抽象的语义信息，使模型能够在更高层次上理解不同退化之间的关系。在不依赖大规模预训练模型的情况下，这类方法通常采用简单的嵌入方式将文本标签映射到特征空间，并与图像特征进行融合，从而实现对不同任务的条件控制。

在多模态提示方法中，视觉提示与文本提示得到进一步结合，通过构建统一的提示表示，实现对复杂退化场景的协同建模。这类方法通过融合多源信息，使模型不仅能够感知低层退化特征，还能够利用语义信息辅助复原，从而在复杂场景中提升模型的适应能力与鲁棒性。

在航空图像复原场景中，提示学习具有较大的应用潜力。例如，可以将成像条件（如传感器类型、拍摄时间或环境状态）作为提示信息引入模型，引导其进行针对性复原；同时，提示机制还可以用于描述不同退化组合关系，从而增强模型对复合退化的适应能力。

2.3 基于动态路由与专家模型的统一复原方法

为进一步提升模型在多退化场景中的适应能力，研究者提出基于动态路由与专家模型（mixture-of-experts，MoE）的统一图像复原方法。该类方法通过构建多个功能差异化的专家子网络，并结合路由机制根据输入特征动态选择或组合合适的专家进行处理，从而实现对不同退化类型及其组合形式的自适应建模。与基于提示学习的方法通过条件信息调控模型行为不同，MoE方法通过引入结构化分工与动态路径选择机制，使模型能够在统一框架下实现对复杂退化的有效建模。

动态路由专家模型通过联合建模专家选择与专家分工机制，实现对不同退化特征的自适应处理。其中，路由机制用于根据输入特征分配专家权重，从而选择合适的处理路径；差异化专家结构则用于刻画不同退化模式，使各专家能够专注于特定特征或处理策略，从而提升模型对复杂场景的表达能力。

在专家选择机制方面，模型通过引入路由函数对输入特征进行分析，并据此生成专家权重，实现样本级或任务级的动态专家分配。Fedus等人（2022）提出的稀疏专家Transformer（switch Transformer）通过限制参与计算的专家数量，在控制计算开销的同时显著提升模型容量，为后续图像复原模型引入MoE结构提供了重要基础。在此基础上，Zamfir等人（2025）提出MoCE-IR，通过构建具有不同计算复杂度与感受野的专家模块，并设计偏向低复杂度的路由策略，使模型能够根据任务难度自适应分配计算资源，从而在保证复原性能的同时减少冗余计算。

在专家分工机制方面，研究者通过设计具有差异化处理能力的专家模块，使不同专家能够针对不同退化模式或特征类型进行建模，从而实现更有效的任务解耦。Cho等人（2021）通过动态权重分配机制，使不同分支能够根据输入特征自适应调整贡献，从而提升多任务复原性能；Li等人（2023）通过引入退化感知的特征调制策略，使模型能够在多退化场景中动态选择更合适的处理路径，从而缓解不同任务之间的干扰问题。

进一步地，为提升模型对空间非均匀退化的建模能力，研究者将动态路由机制扩展至特征空间，使专家选择能够在空间维度上进行自适应调整。通过在同一图像中对不同区域分配不同处理路径，模型能够更加精细地恢复局部结构，从而提升复杂场景下的复原效果。

在航空与遥感图像场景中，由于退化通常具有多因素耦合与空间分布不均的特点，同一图像中不同区域可能受到噪声、模糊、光照不足及大气扰动等多种因素影响。基于动态路由与专家模型的方法能够在更细粒度上为不同退化模式分配差异化处理路

径，从而显著提升复杂环境下的复原性能。

3 大模型驱动的航空图像复原方法

随着视觉基础模型（vision foundation model）与多模态学习技术的发展，图像复原方法正从以任务为中心的专用建模范式，逐步演进为以大规模预训练为基础的统一表征学习范式。该类方法通过在海量无标注或弱标注数据上进行预训练，学习具有良好迁移能力的通用视觉表示，并在下游任务中通过特征复用、参数微调或结构适配实现性能提升。相较于传统深度学习方法依赖任务特定数据与结构设计，大模型方法能够在复杂退化、跨场景及小样本条件下表现出更强的泛化能力与鲁棒性。

根据技术路径可以将大模型驱动的图像复原方法划分为以下 3 类：1）以自监督学习为基础的视觉基础模型逐渐成为统一的特征提取框架，为多退化场景下的图像复原提供稳定且具有良好泛化能力的表征；2）多模态学习通过联合建模图像、文本以及多源传感器信息，实现对复杂成像环境的协同理解与信息互补，从而提升复原效果的鲁棒性与适应性；3）随着决策机制与推理能力的发展，基于智能体的图像复原方法将复原过程建模为动态决策问题，通过引入策略选择与反馈优化机制，使模型能够在复杂退化条件下自适应地调整处理流程，进一步拓展了大模型在图像复原任务中的应用形式与能力边界。表 4 展示了大模型驱动的航空图像复原的代表方法。

在航空图像复原场景中，由于退化类型复杂、数据获取困难以及成像条件多变，大模型所具备的跨任务迁移能力与跨模态建模能力具有更加突出的优势。因此，从视觉基础模型、多模态建模及语义引导等角度对相关方法进行系统梳理，对于理解当前研究进展及未来发展方向具有重要意义。

表 4　大模型驱动的航空图像复原代表方法

Table 4　Representative large model-driven methods related to aerial image restoration

方法	年份	核心思想	模型复杂度	推理效率
DiffPIR	2023	将即插即用复原思想融入扩散采样过程，利用退化算子约束观测一致性	高	较慢
DPS	2023	将显式观测模型与扩散后验采样结合，处理含噪声图像逆问题	高	较慢
DDNM	2023	基于线性退化算子和零空间约束实现零样本图像复原	高	较慢
Diff-Retinex	2023	将 Retinex 分解与扩散生成结合，增强低光图像复原的物理解释性	高	较慢
LightenDiffusion	2024	将 Retinex 理论融入潜空间扩散模型，实现无监督低光增强	较高	中等
LLMRA	2024	利用多模态大语言模型生成退化描述和语义上下文，引导图像复原	高	较慢
RestoreAgent	2024	构建复原智能体，自动判断退化类型并规划复原流程	高	较慢

3.1 大模型基本范式

随着视觉基础模型与大规模预训练技术的发展，图像复原方法逐渐由面向单一退化设计的任务专用建模范式，演进为以统一表征学习为核心的大模型驱动范式。该类方法通过在海量无标注或弱标注数据上进行预训练，使模型能够学习具有良好泛化能力的视觉表示，并通过特征迁移、参数微调或结构适配等方式应用于具体复原任务，从而在复杂退化与跨场景条件下获得更加稳定的性能表现。与传统依赖显式退化建模的方法相比，该范式更加关注通用表征能力的提升，其本质体现为由“退化建模驱动”向“表征学习驱动”的转变。

在模型体系上，如图 3 所示，大模型基本范式可划分为视觉基础模型、视觉语言模型、多模态大模型和大语言模型 4 类。其中，视觉基础模型侧重于通用视觉表征学习和结构先验提取；视觉语言模型通过图文对齐增强语义理解能力；多模态大模型进一步融合图像、文本及多源传感器信息，实现跨模态协同建模；大语言模型则主要提供语言推理、任务规划和策略生成能力。上述模型共同构成大模型驱动图像复原的重要技术基础，为航空图像复杂退化建模、语义引导复原和动态决策优化提供支撑。

视觉基础模型通过大规模预训练获得通用视觉表征，是大模型体系中支撑图像复原任务的重要基础。例如，DINO 模型通过自蒸馏机制实现特征一致性学习，使模型能够在不同视角下获得稳定的语义

表示(Caron等,2021);掩码自编码模型通过随机遮挡输入并预测缺失区域,使模型能够捕获图像的全局上下文关系(He等,2022)。相关后续研究(Oquab等,2024)在特征对齐与跨任务迁移能力方面持续改进,使视觉基础模型在复杂场景中的表现更加稳定。视觉基础模型在图像复原中的核心作用并不在于直接完成复原任务,而在于提供具有良好泛化能力的特征表示,为后续复原模型的设计与优化提供基础支撑。

视觉语言模型通过引入文本信息,将视觉特征映射至语义空间,从而增强模型对图像内容的理解能力。CLIP(contrastive language-image pre-training)模型(Radford等,2021)通过对比学习实现图像与文本特征的对齐,使模型能够在跨模态空间中建立语义一致表示。在此基础上,视觉指令学习模型(Liu等,2023)通过结合视觉编码器与语言模型,使模型具备图像理解与语义推理能力。该类方法能够在图像复原过程中引入语义约束,从而提升结构合理性与视觉一致性。

多模态大模型在视觉语言模型基础上进一步扩展至多源信息融合,通过统一建模图像、文本以及其他传感器数据,实现更加全面的场景理解能力。例如,Segment Anything模型(Kirillov等,2023)通过大规模分割数据训练获得通用视觉表示,在结构建模方面具有重要意义。此外,多模态模型通过跨模态特征交互,使不同信息源能够协同作用,从而在复杂退化条件下获得更加鲁棒的特征表达能力。在航空图像复原中,该类方法能够有效融合红外与可见光等多源信息,从而提升模型对复杂环境的适应能力。

大语言模型虽然不直接参与像素级建模,但在视觉任务中的辅助作用逐渐增强。通过提供任务描述、语义推理或结构先验,大语言模型能够对图像复原过程进行间接指导。例如,可利用语言模型生成复原策略或提供语义解释,从而提升模型在复杂场景中的决策能力。这种“语言引导视觉”的建模方式,为图像复原提供了新的研究思路。

总体而言,视觉基础模型提供结构表征能力,视觉语言模型增强语义理解,多模态大模型实现多源信息融合,而大语言模型提供高层推理能力,这4类模型共同构成大模型驱动图像复原方法的核心技术体系。在航空图像复原场景中,该范式能够有效应对复杂退化与数据稀缺问题,但仍面临跨域迁移、多模态对齐及计算效率等方面的挑战。

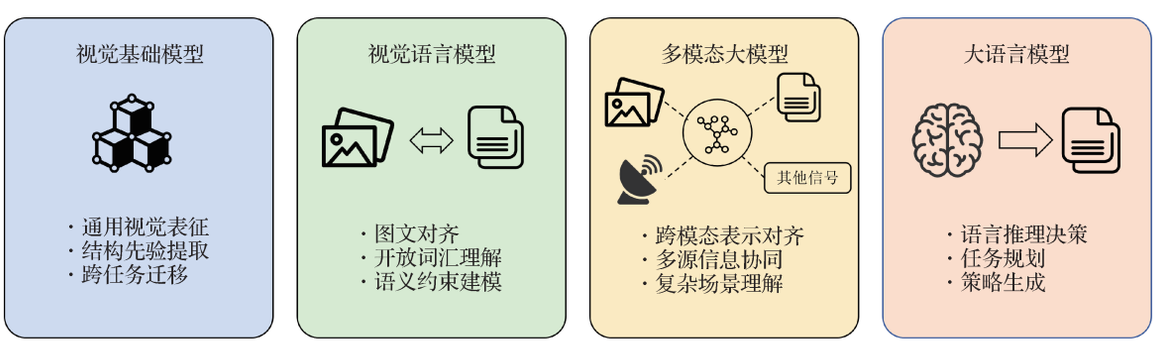


图3　大模型基本范式分类示意图

Fig. 3　Classification diagram of basic large model paradigms

3.2　基于视觉基础模型的图像复原方法

基于视觉基础模型的图像复原方法通过引入大规模预训练模型所学习到的通用视觉表征,对复原网络的特征提取与结构建模过程进行增强(韦炎炎等,2025)。与传统从头训练的复原模型不同,该类方法利用预训练阶段所获得的结构与语义信息,使模型在复杂退化条件下获得更加稳定的特征表达能力。当前主流研究主要围绕预训练先验引导、特征融合增强以及轻量化适配3类技术路径展开。

在预训练先验引导方面,研究者利用预训练模型中隐含的自然图像分布或结构信息,对复原过程进行约束。一类方法通过引入生成模型先验,将图像复原问题转化为受约束的生成过程,使模型在恢复细节的同时保持整体结构一致性。例如,Saharia等人(2023)提出基于扩散模型的图像恢复方法,通过逐步去噪过程实现高质量图像重建;Kawar等人(2022)进一步将预训练扩散模型作为先验约束,用于解决图像逆问题,从而在去模糊与去噪任务中取得良好效果。这类方法的核心在于利用预训练模型所学习到的自然图像分布,对复原过程进行全局约

束，从而减少结构失真。

进一步地，近年来部分研究开始将物理退化模型与大模型先验相结合，以增强复原过程的可解释性和观测一致性。与仅依赖生成先验的方法不同，该类方法通常将退化算子、观测模型或物理成像约束显式引入扩散采样过程，使复原结果既符合预训练模型学习到的图像分布，又满足退化成像过程的约束。例如，Zhu 等人（2023）提出 DiffPIR（denoising diffusion models for plug-and-play image restoration）方法，将即插即用图像复原思想融入扩散模型采样框架，通过数据一致性项约束复原图像与退化观测之间的关系，并利用预训练扩散模型提供图像先验，可用于超分辨率、去模糊和图像修复等任务。Chung 等人（2023）提出 DPS（diffusion posterior sampling）方法，从后验采样角度将显式观测模型与扩散生成过程结合，使扩散模型能够处理含噪声的线性或非线性图像逆问题。Wang 等人（2023）提出 DDNM 方法，基于线性退化算子和零空间约束，在无需额外训练的条件下利用预训练扩散模型实现零样本图像复原。

此外，针对特定退化任务，也有研究将物理成像模型作为条件信息或结构约束嵌入生成式复原框架中。例如，Yi 等人（2023）提出 Diff-Retinex，将 Retinex 理论中的照明—反射分解与扩散生成过程相结合，使低光增强过程在具备生成能力的同时保持一定物理解释性；Jiang 等人（2024）提出 LightenDiffusion，将 Retinex 先验引入潜空间扩散模型，在无监督条件下实现低光图像增强。这类方法表明，物理模型不仅可以作为传统优化方法中的显式约束，也可以转化为退化条件、控制特征或采样引导信号，与大模型的生成先验协同作用。对于航空图像复原而言，平台运动、大气散射、复杂光照和传感器噪声等退化往往具有明确的成像来源，因此将物理先验与大模型结合，有助于提升复杂退化条件下复原结果的稳定性、可解释性和跨场景泛化能力。

在特征融合增强方面，研究者将预训练模型提取的多层特征与复原网络中的特征进行深度融合，以提升模型的表达能力。一类方法通过跨尺度特征融合，使高层语义信息与低层细节信息协同作用，从而在保持细节的同时增强结构一致性。例如，Liang 等人（2021）提出基于层次化 Transformer 的图像复原方法，通过引入滑动窗口注意力机制实现跨尺度特征交互，从而提升模型对复杂退化的建模能力；Chen 等人（2022）提出简化非线性激活的高效复原网络，通过增强特征表达能力实现更高效的细节恢复。此外，一些研究通过特征对齐或跨注意力机制实现不同特征之间的动态融合，使模型能够根据退化程度自适应调整特征权重，从而在复杂退化条件下获得更加稳定的性能表现。该类方法强调多层次特征协同与结构建模，在复杂场景中具有较强的鲁棒性。

在轻量化适配方面，研究者通过引入参数高效的适配机制，使预训练视觉模型能够以较低成本迁移到图像复原任务中。一类方法通过在预训练模型中插入轻量化适配模块，对特征进行局部调制，从而在不改变主体结构的情况下实现任务适配。例如，Chen 等人（2023）针对视觉 Transformer 提出适配模块结构，使模型能够通过少量参数实现对下游任务的有效迁移；另一类方法基于低秩适配思想，通过引入少量可学习参数对模型权重进行调整，在保证性能的同时显著降低训练开销（Hu 等，2022）。

此外，基于提示的适配方法通过在输入或特征空间中引入可学习提示向量，对模型行为进行调制，使模型能够针对不同退化条件自适应调整复原策略（Jia 等，2022）。同时，一些研究通过特征调制机制（如通道缩放与偏移）对预训练特征进行动态调整，从而实现对不同退化类型的灵活适配。这类方法在多退化场景中具有良好的扩展性与效率优势。

在航空图像复原场景中，基于视觉基础模型的方法具有显著优势。一方面，航空图像标注数据有限，而预训练模型能够提供稳定的特征表示，从而缓解数据不足问题；另一方面，航空图像中的退化通常具有复杂结构特性，如运动模糊、大气扰动及传感器噪声等，预训练先验能够为复原过程提供有效约束，从而提升模型对复杂退化的建模能力。例如，在去模糊任务中引入生成先验可以改善边缘恢复效果，在去噪任务中融合多层特征能够减少过平滑现象。

3.3 基于多模态大模型的图像复原方法

基于多模态大模型的图像复原方法以大规模跨模态预训练为基础，通过统一建模图像与文本等多模态信息，使模型能够在共享表征空间中融合语义与视觉特征，从而提升对复杂退化场景的建模能力。与传统多模态融合方法主要关注不同传感器数据之间的特征拼接不同，如图 4 所示，基于多模态大模型

的图像复原方法通常将退化图像作为主要输入，而文本描述、红外图像、SAR图像或语义分割结果等辅助信息可为复原过程提供额外的场景语义与结构线索。多模态大模型通过对不同模态信息进行联合编码，提取图像内容、退化状态和高层语义信息，并将其转化为语义嵌入或条件特征，用于引导后续复原网络。该类方法更加依赖多模态大模型所学习到的跨模态对齐能力与语义理解能力，使复原过程能够引入高层语义约束，从而在结构恢复与视觉一致性方面取得更优表现。当前研究主要围绕语义先验生成、语义反馈优化以及统一多模态建模3类技术路径展开。

在语义先验生成方面，多模态大模型用于解析退化图像并生成对应的语义描述，从而为复原过程提供高层先验信息。例如，Jin等人(2024)提出多模态大语言模型辅助的图像复原(multi-modal large language model based restoration assistant，LLMRA)方法，利用多模态大语言模型对低质量图像进行语义分析，生成退化描述，并通过CLIP文本编码器将其映射为语义嵌入，引导复原网络进行结构恢复。该方法的核心在于利用多模态大模型的跨模态理解能力，将复杂退化过程转化为可解释的语义信息，从而增强模型对退化类型的感知能力。

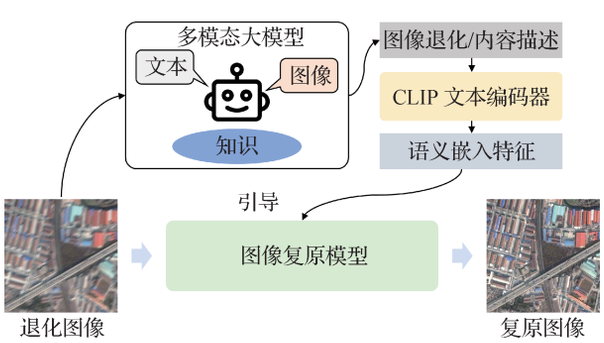


图4　基于多模态大模型的图像复原方法框架图

Fig. 4　Framework of multimodal large model-based image restoration methods

在语义反馈优化方面，多模态大模型用于对复原结果进行评价与迭代优化，而不仅仅作为特征提取器。例如，Lai等人(2025)提出SnowMaster方法，在真实图像去雪任务中引入多种多模态大模型，包括Q-Instruct-13B、LLaVA-1.5、LLaVA-1.6以及MiniCPM-V-2.6，构建多模型反馈机制，对复原结果进行偏好评估与伪标签筛选。通过利用多模态大模型对图像质量与语义一致性的综合判断能力，该方法能够在缺乏精确标注的情况下实现更稳定的模型优化。该类方法表明，多模态大模型不仅能够提供先验信息，还可以作为“评价器”参与复原过程的闭环优化。

在统一多模态建模方面，研究者进一步探索将多模态大模型作为统一理解与生成框架的核心，从而实现端到端的复原建模。例如，Wei等人(2025)提出PURE(perceive, understand and restore)方法，采用Lumina-mGPT作为多模态自回归生成模型，并利用LLaVA生成高质量图像的语义描述，将图像理解与图像复原过程统一到同一框架中。此外，Luo等人(2025)提出基于视觉指令驱动的扩散复原方法，通过引入多模态语义提示调制扩散模型的生成过程，实现多退化统一复原。该类方法的特点在于将多模态大模型作为核心建模单元，使模型能够同时具备语义理解与图像生成能力，从而在复杂场景中获得更强的适应性。

在航空图像复原场景中，基于多模态大模型的方法具有显著优势。一方面，航空图像退化通常具有复杂性与不确定性，单一模态难以充分刻画退化特征，而多模态大模型能够通过语义理解增强模型对场景的认知能力；另一方面，复杂环境(如夜间、雾霾或大气扰动)往往导致图像结构信息缺失，而语义先验可以在一定程度上弥补低层信息不足，从而提升复原结果的结构一致性。例如，在低光或极端天气条件下，通过语义描述引导复原过程可以有效减少结构失真问题。

3.4　基于智能体的图像复原方法

基于智能体的图像复原方法通过引入决策机制，将图像复原过程由传统的端到端映射转化为多步骤的动态优化过程。如图5所示，在该范式下，基于智能体的图像复原方法通常由退化感知、策略规划、模型调用、反馈评价和迭代优化等模块组成，模型不再一次性输出复原结果，而是通过策略选择与状态更新，在多个处理步骤中逐步优化图像质量。首先，退化感知模块对输入图像的退化类型、退化程度及空间分布进行分析，为后续复原策略提供状态信息；随后，策略规划模块根据退化分析结果确定处理顺序和模型调用方案，并选择合适的复原模型或工具执行具体操作；最后，反馈评价模块对中间复原结果进行质量评估，并将评价结果用于调整后续复

原策略。随着大模型在感知与推理能力上的提升，智能体方法逐渐与视觉基础模型及多模态大模型相结合，使图像复原系统具备一定的自适应决策能力与过程控制能力。根据实现方式的不同，现有研究主要可分为基于序列决策的复原方法、基于反馈优化的复原方法以及基于多智能体协同的复原方法3类。

在基于序列决策的复原方法中，研究者将图像复原过程建模为马尔可夫决策过程，通过策略网络在每一步选择合适的操作，从而逐步改善图像质量。例如，Yu等人(2018)提出基于强化学习的图像增强方法，将去噪、锐化等操作视为动作，通过策略学习实现逐步优化；Hu等人(2018)进一步引入深度强化学习机制，使模型能够根据当前图像状态自适应选择处理策略，从而在复杂退化条件下获得更优的复原路径。该类方法的核心在于通过决策序列替代固定映射，使模型具备一定的过程自适应能力。

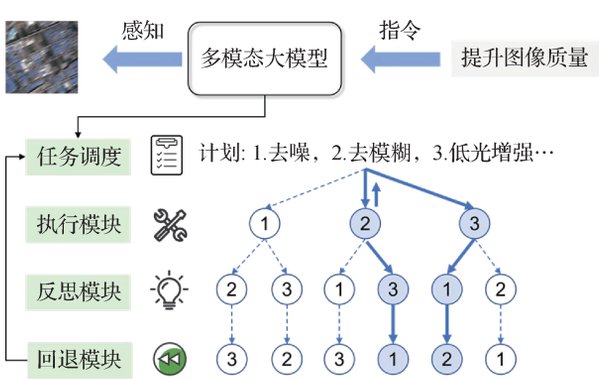


图5　面向复杂退化的航空图像复原智能体框架

Fig. 5　An agent-based framework for aerial image restoration under complex degradations

在基于反馈优化的复原方法中，智能体通过引入评价与反馈机制，对复原过程进行动态调整，从而实现逐步优化。与传统一次性输出结果的复原模型不同，该类方法通过构建“评估—决策—更新”的闭环过程，使模型能够根据当前复原效果不断修正后续操作。例如，Chen等人(2024)提出RestoreAgent框架，将图像复原过程划分为退化识别、任务规划与模型执行3个阶段，并通过智能体实现处理流程的动态调度，从而在多退化场景中获得更具适应性的复原策略。在此基础上，Li等人(2025)提出HybridAgent方法，通过构建“快速响应—精细优化”的双路径结构，使智能体能够在初始阶段快速生成粗略结果，并在后续阶段结合反馈信息进行逐步优化，从而在保证效率的同时提升复原质量。此外，一些研究通过引入质量评估模块或奖励函数，将感知质量与结构一致性等指标纳入决策过程，使智能体能够在优化过程中不断逼近更优复原结果，从而增强模型在真实复杂场景中的鲁棒性。

在基于多智能体协同的复原方法中，研究者通过构建多个功能分化的智能体，实现对复杂复原任务的分解与协同处理。例如，一些方法将复原过程划分为退化分析、策略规划与图像重建等子任务，由不同智能体分别负责，并通过信息交互实现整体优化；此外，还有研究通过构建“规划智能体—执行智能体”的协同框架，使系统能够先生成全局复原策略，再逐步执行具体操作，从而在多退化耦合场景中获得更加稳定的处理效果。该类方法通过结构化分解复杂任务，有效提升了模型在复杂环境下的适应能力与扩展能力。

近年来，智能体驱动的图像复原研究进一步朝着更强的探索能力与更精细的优化目标发展。Zhang等人(2026)提出TIR-Agent方法，通过构建探索型复原智能体，在训练过程中增强策略搜索能力，使模型能够在复杂退化条件下学习更加高效的复原路径；Wang等人(2026)提出PaAgent方法，通过引入主观—客观联合强化学习机制，将感知质量与客观评价指标共同纳入优化目标，从而使智能体在保证视觉效果的同时兼顾定量性能。这表明，智能体方法正由早期的“操作序列学习”逐步扩展为“探索增强与多目标优化”并重的研究方向。

在航空图像复原场景中，基于智能体的方法具有较强的应用潜力。一方面，航空图像退化通常具有多因素耦合与动态变化的特点，单一映射模型难以兼顾不同退化条件，而智能体方法能够根据图像状态动态调整处理策略；另一方面，结合视觉基础模型与多模态大模型，智能体可以利用语义信息与结构信息对复原过程进行全局规划，从而在复杂环境中获得更加稳定的复原效果。例如，在多退化场景中，智能体可以优先处理主导退化，再逐步细化细节恢复，从而提高整体复原效率。

4　挑战与局限

尽管近年来航空图像复原方法在单一退化建

模、统一建模以及大模型驱动等方面取得了显著进展，但在应对复杂实际场景时仍存在一定局限。为更直观地比较不同类型方法在关键能力维度上的差异，本文从多退化建模能力、语义协同能力、物理先验利用、动态决策能力、计算效率及主要局限等方面，对现有航空图像复原方法进行结构化归纳，如表5所示。总体来看，当前方法仍面临以下主要瓶颈：1）多退化耦合关系难以有效建模，不同退化特征在共享表示空间中易发生混淆；2）语义信息与低层视觉特征缺乏有效协同，导致在严重退化或结构信息缺失情况下难以实现稳定的结构恢复；3）现有方法多基于单次前向推理的静态处理方式，缺乏对复原过程的动态调整与自适应优化能力。上述问题在一定程度上制约了模型在真实复杂环境中的复原性能与泛化能力。围绕这些关键瓶颈，本文进一步对现有航空图像复原方法面临的挑战与局限展开分析。

1）多退化联合建模能力不足。当前图像复原方法虽然已由面向单一退化的任务驱动范式逐步演进至统一建模范式，但在多退化联合建模方面仍存在明显不足。一方面，基于混合数据训练的隐式建模方法通常依赖共享特征空间对不同退化进行统一表示，缺乏对退化形成机理的显式刻画，容易导致不同退化特征在特征空间中发生混淆，从而影响模型对复杂退化的判别能力与复原精度；另一方面，即使引入退化感知或条件建模机制，现有方法仍多将不同退化视为相对独立的因素，难以准确刻画真实航空场景中多种退化之间复杂的关联关系与相互作用。

此外，在实际航空图像中，退化通常具有明显的空间非均匀性，不同区域可能同时受到噪声、模糊及光照变化等多种因素的差异性影响。然而，现有统一建模方法大多采用全局一致的处理策略，缺乏对空间变异退化的细粒度建模能力，导致局部区域复原效果不稳定。因此，如何构建能够同时表征退化共性、差异性及其内在关联的统一建模框架，仍是当前图像复原领域需要重点解决的问题之一。

2）语义信息与低层视觉建模耦合不足。随着视觉基础模型与多模态大模型的发展，语义信息已逐渐被引入图像复原任务，用以弥补低层视觉信息在严重退化条件下的不足。然而，从现有研究来看，语义信息的利用仍主要停留在浅层融合阶段，其对复原过程的指导作用尚未得到充分发挥。

具体而言，当前多模态复原方法多通过特征拼接或注意力机制将语义特征与视觉特征进行融合，但两者之间缺乏统一的表示约束与协同优化机制，导致语义信息难以有效参与细节重建过程。在严重退化场景中，低层结构信息严重缺失，理论上应由语义先验主导结构恢复，但现有方法仍以像素级重建

表5　现有航空图像复原方法的性能和复杂性对比

Table 5　Performance and complexity comparison of existing aerial image restoration methods

方法类型	多退化建模能力	语义协同能力	动态决策能力	计算效率	主要局限
面向单一退化的航空图像复原	弱	弱	弱	高	任务针对性强，但难以处理多退化耦合和未知退化场景
基于多退化建模的统一图像复原	较强	中	中	中	可处理多种退化，但对退化机理和空间非均匀性的刻画不足
基于提示学习的统一图像复原	较强	中	中	中	依赖提示表达能力，复杂退化条件下提示与退化状态可能不完全匹配
基于动态路由与专家模型的统一复原	较强	中	较强	中	具有自适应路径选择能力，但路由稳定性和专家分工需优化
基于视觉基础模型的图像复原	中	中	弱	中	具备较强视觉表征能力，但跨域适应和航空场景迁移存在挑战
基于多模态大模型的图像复原	较强	强	中	低	语义理解能力强，但跨模态对齐、计算开销和语义偏差问题较突出
基于智能体的图像复原	较强	强	强	低	可动态规划复原流程，但反馈可靠性、多步骤推理效率和部署成本仍需改进

为主导，语义信息仅作为辅助调制信号，难以实现“语义驱动”的结构生成。此外，多模态对齐误差及语义表达不确定性也可能引入结构偏差，甚至导致复原结果出现不符合真实场景的伪结构。

因此，如何实现语义信息与低层视觉信息的深度耦合，使语义信息从“辅助信息”转变为“主导约束”，并在统一框架下实现结构与细节的协同优化，是提升复杂退化场景复原性能的关键挑战。

3）缺乏统一的动态决策与自适应优化机制。现有大多数图像复原方法仍基于端到端映射范式，通过单次前向推理直接输出复原结果，缺乏对复杂退化条件下复原过程的动态调控能力。这类方法在面对多退化耦合及退化程度变化较大的场景时，往往难以兼顾不同退化之间的处理优先级，从而限制了复原性能的进一步提升。

尽管基于智能体的图像复原方法通过引入序列决策与反馈优化机制，在一定程度上实现了复原过程的动态调整，但当前研究仍处于探索阶段。一方面，现有方法多依赖预定义的操作集合或有限策略空间，缺乏对未知退化场景的自主探索能力；另一方面，反馈机制通常依赖图像质量评估模型或简单奖励函数，其稳定性与泛化能力仍有待提升。此外，多步骤决策过程也带来了较高的计算开销，限制了其在实际场景中的应用。

因此，如何构建具备感知、决策与反馈能力的一体化复原框架，使模型能够根据输入图像状态自适应调整复原策略，并在保证效率的同时实现高质量复原，仍是当前大模型驱动图像复原研究的重要发展方向。

5　未来展望

随着视觉基础模型、多模态大模型和智能体的发展，航空图像复原正由面向单一退化的任务处理范式，逐步转向以统一表征学习、跨模态协同建模和动态决策优化为特征的智能复原体系。总体来看，现有研究已由“针对特定退化设计专用模型”扩展至“依托大模型实现多退化统一建模”，并在结构恢复、场景理解及复杂条件适应等方面取得了重要进展。

面向未来，航空图像复原可从数据、方法和应用3个层面继续发展。在数据层面，构建覆盖多源传感器、多尺度分辨率和复杂动态场景的高质量数据集，是提升模型泛化能力的重要基础。红外、SAR、激光雷达（light detection and ranging，LiDAR）和高光谱等多源数据能够提供互补的结构、辐射和物理信息，有助于增强模型在复杂退化条件下的稳健性。同时，结合无人机平台、自动采集系统和弱监督标注机制，可进一步缓解航空图像数据获取困难和标注成本高的问题。

在方法层面，如何发挥大模型的统一表征能力，实现多退化协同建模、跨域适应和结构一致性恢复，仍是重要研究方向。参数高效微调、迁移学习、无监督与半监督学习等策略，可降低大模型在航空图像复原中的适配成本；多模态大模型、扩散模型与物理成像模型的结合，则为复杂退化过程的精细建模提供了新的可能。此外，智能体机制能够引入策略规划和反馈优化能力，但仍需进一步提升其稳定性、可解释性和计算效率。

在应用层面，航空图像复原应更加重视与下游任务的协同。高质量复原不仅服务于视觉质量提升，也直接影响目标检测、变化检测、场景理解和决策分析等任务。因此，面向任务一致性的评价体系、面向应用闭环的联合建模方法，以及适用于边缘计算和机载平台的轻量化部署策略，将成为推动航空图像复原走向实际应用的关键。

总体而言，航空图像复原正在向统一建模、多模态融合、物理先验约束和智能决策协同的方向发展。未来研究不仅需要提升复原质量，还应构建兼顾表征能力、物理一致性、任务适配性和部署可行性的系统化复原框架，从而推动航空图像复原在复杂真实场景中的应用。

参考文献（References）

Aharon M, Elad M and Bruckstein A. 2006. K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation. IEEE Transactions on Signal Processing, 54(11): 4311-4322 [DOI: 10.1109/TSP.2006.881199]

Caron M, Touvron H, Misra I, Jegou H, Mairal J, Bojanowski P, et al. 2021. Emerging properties in self-supervised vision transformers// Proceedings of 2021 IEEE/CVF International Conference on Computer Vision. Montreal, Canada: IEEE: 9630-9640 [DOI: 10.1109/ICCV48922.2021.00951]

Chen L Y, Chu X J, Zhang X Y and Sun J. 2022. Simple baselines for image restoration//Proceedings of the 17th European Conference on

Computer Vision. Tel Aviv, Israel: Springer: 17-33 [DOI: 10.1007/978-3-031-20071-7_2]

Chen H, Li W, Gu J, Ren J, Chen S, Ye T, et al. 2024. RestoreAgent: autonomous image restoration agent via multimodal large language models. Advances in Neural Information Processing Systems, 37: 110643-110666 [DOI: 10.52202/079017-3512]

Chen Z, Duan Y, Wang W, He J, Lu T, Dai J, et al. 2023. Vision Transformer adapter for dense predictions//Proceedings of the 11th International Conference on Learning Representations. OpenReview.net

Chierchia G, Cozzolino D, Poggi G and Verdoliva L. 2017. SAR image despeckling through convolutional neural networks//Proceedings of 2017 IEEE International Geoscience and Remote Sensing Symposium. Fort Worth, USA: IEEE: 5438-5441 [DOI: 10.1109/IGARSS.2017.8128234]

Cho S J, Ji S W, Hong J P, Jung S W and Ko S J. 2021. Rethinking coarse-to-fine approach in single image deblurring//Proceedings of 2021 IEEE/CVF International Conference on Computer Vision. Montreal, Canada: IEEE: 4621-4630 [DOI: 10.1109/ICCV48922.2021.00460]

Chung H, Kim J, Mccann M T, Klasky M L and Ye J C. 2023. Diffusion posterior sampling for general noisy inverse problems//Proceedings of the 11th International Conference on Learning Representations. OpenReview.net

Dabov K, Foi A, Katkovnik V and Egiazarian K. 2007. Image denoising by sparse 3-D transform-domain collaborative filtering. IEEE Transactions on Image Processing, 16(8): 2080-2095 [DOI: 10.1109/TIP.2007.901238]

Eilertsen G, Kronander J, Denes G, Mantiuk R K and Unger J. 2017. HDR image reconstruction from a single exposure using deep CNNs. ACM Transactions on Graphics, 36(6): #178 [DOI: 10.1145/3130800.3130816]

Fedus W, Zoph B and Shazeer N. 2022. Switch transformers: scaling to trillion parameter models with simple and efficient sparsity. The Journal of Machine Learning Research, 23(1): #120

Fergus R, Singh B, Hertzmann A, Roweis S T and Freeman W T. 2006. Removing camera shake from a single photograph//Proceedings of the ACM SIGGRAPH 2006 Papers. Boston, USA: ACM: 787-794 [DOI: 10.1145/1179352.1141956]

Goodfellow I J, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. 2014. Generative adversarial nets//Proceedings of the 28th International Conference on Neural Information Processing Systems. Montreal, Canada: MIT Press: 2672-2680

Guo C L, Li C Y, Guo J C, Loy C C, Hou J H, Kwong S, et al. 2020. Zero-reference deep curve estimation for low-light image enhancement//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE: 1777-1786 [doi: 10.1109/CVPR42600.2020.00185]

He K M, Chen X L, Xie S N, Li Y H, Dollár P and Girshick R. 2022. Masked autoencoders are scalable vision learners//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, USA: IEEE: 15979-15988 [DOI: 10.1109/CVPR52688.2022.01553]

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

Hu E J, Shen Y L, Wallis P, Allen-Zhu Z, Li Y Z, Wang S A, et al. 2022. LoRA: low-rank adaptation of large language models//Proceedings of the 10th International Conference on Learning Representations. OpenReview.net

Hu Y, He H, Xu C, Wang B and Lin S. 2018. Exposure: a white-box photo post-processing framework. ACM Transactions on Graphics, 37(2): 1-17 [DOI: 10.1145/3181974]

Jia M L, Tang L M, Chen B C, Cardie C, Belongie S, Hariharan B, et al. 2022. Visual prompt tuning//Proceedings of the 17th European Conference on Computer Vision. Tel Aviv, Israel: Springer: 709-727 [DOI: 10.1007/978-3-031-19827-4_41]

Jiang H, Luo A, Liu X H, Han S C and Liu S C. 2024. LightenDiffusion: unsupervised low-light image enhancement with latent-Retinex diffusion models//Proceedings of the 18th European Conference on Computer Vision. Milan, Italy: Springer: 161-179 [DOI: 10.1007/978-3-031-73195-2_10]

Jin X Y, Shi Y, Xia B and Yang W M. 2024. LLMRA: multi-modal large language model based restoration assistant [EB/OL]. [2026-04-16]. https://arxiv.org/pdf/2401.11401.pdf

Kalantari N K and Ramamoorthi R. 2017. Deep high dynamic range imaging of dynamic scenes. ACM Transactions on Graphics, 36(4): #144 [DOI: 10.1145/3072959.3073609]

Kawar B, Elad M, Ermon S and Song J M. 2022. Denoising diffusion restoration models//Proceedings of the 36th International Conference on Neural Information Processing Systems. New Orleans, USA: Curran Associates Inc.: #1714

Kirillov A, Mintun E, Ravi N, Mao H Z, Rolland C, Gustafson L, et al. 2023. Segment anything//Proceedings of 2023 IEEE/CVF International Conference on Computer Vision. Paris, France: IEEE: 3992-4003 [DOI: 10.1109/ICCV51070.2023.00371]

Kupyn O, Budzan V, Mykhailych M, Mishkin D and Matas J. 2018. DeblurGAN: blind motion deblurring using conditional adversarial networks//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE: 8183-8192 [DOI: 10.1109/CVPR.2018.00854]

Lai J Y, Chen S X, Lin Y L, Ye T, Liu Y, Fei S, et al. 2025. SnowMaster: comprehensive real-world image desnowing via MLLM with multi-model feedback optimization//Proceedings of 2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, USA: IEEE: 4302-4312 [doi: 10.1109/CVPR52734.2025.00406]

Levin A, Fergus R, Durand F and Freeman W T. 2007. Image and depth from a conventional camera with a coded aperture//Proceedings of the ACM SIGGRAPH 2007 Papers. San Diego, USA: ACM: #70 [DOI: 10.1145/1275808.1276464]

Li B C, Li X, Lu Y T and Chen Z B. 2025. Hybrid agents for image restoration [EB/OL]. [2026-04-16]. https://arxiv.org/pdf/2503.10120.pdf

Li B Y, Liu X, Hu P, Wu Z Q, Lyu J C and Peng X. 2022. All-in-one image restoration for unknown corruption//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, USA: IEEE: 17431-17441 [DOI: 10.1109/CVPR52688.2022.01693]

Li R, Zhang K, Van Gool L, Timofte R, Gu S and Liu J. 2023. All-in-one image restoration via degradation-aware modeling//Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver, Canada: IEEE: 17452-17462 [DOI: 10.1109/CVPR52729.2023.01673]

Liang J Y, Cao J Z, Sun G L, Zhang K, Van Gool L and Timofte R. 2021. SwinIR: image restoration using swin transformer//Proceedings of 2021 IEEE/CVF International Conference on Computer Vision Workshops. Montreal, Canada: IEEE: 1833-1844 [DOI: 10.1109/ICCVW54120.2021.00210]

Liu H T, Li C Y, Wu Q Y and Lee Y J. 2023. Visual instruction tuning//Proceedings of the 37th International Conference on Neural Information Processing Systems. New Orleans, USA: Curran Associates Inc.: #1516

Luo W Y, Qin H N, Chen Z W, Wang L B, Zheng D D, Li Y M, et al. 2025. Visual-instructed degradation diffusion for all-in-one image restoration//Proceedings of 2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, USA: IEEE: 12764-12777 [DOI: 10.1109/CVPR52734.2025.01191]

Lyu F, Li Y and Lu F. 2021. Attention guided low-light image enhancement with a large scale low-light simulation dataset. International Journal of Computer Vision, 129(7): 2175-2193 [DOI: 10.1007/s11263-021-01466-8]

Mertens T, Kautz J and Van Reeth F. 2007. Exposure fusion//Proceedings of 2007 Pacific Conference on Computer Graphics and Applications. Maui, USA: IEEE: 382-390 [DOI: 10.1109/PG.2007.17]

Oquab M, Darcet T, Moutakanni T, Vo H V, Szafraniec M, Khalidov V, et al. 2024. DINOv2: learning robust visual features without supervision. Transactions on Machine Learning Research: 1-31 [DOI: 10.48550/arXiv.2304.07193]

Pan J S, Sun D Q, Pfister H and Yang M H. 2016. Blind image deblurring using dark channel prior//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE: 1628-1636 [doi: 10.1109/CVPR.2016.180]

Potlapalli V, Zamir S W, Khan S and Khan F S. 2023. PromptIR: prompting for all-in-one blind image restoration//Proceedings of the 37th International Conference on Neural Information Processing Systems. New Orleans, USA: Curran Associates Inc.: #3121

Radford A, Kim J W, Hallacy C, Ramesh A, Goh G, Agarwal S, et al. 2021. Learning transferable visual models from natural language supervision//Proceedings of the 38th International Conference on Machine Learning. PMLR: 8748-8763

Rudin L I, Osher S and Fatemi E. 1992. Nonlinear total variation based noise removal algorithms. Physica D: Nonlinear Phenomena, 60(1/4): 259-268 [DOI: 10.1016/0167-2789(92)90242-F]

Saharia C, Ho J, Chan W, Salimans T, Fleet D J and Norouzi M. 2023. Image super-resolution via iterative refinement. IEEE Transactions on Pattern Analysis and Machine Intelligence, 45(4): 4713-4726 [DOI: 10.1109/TPAMI.2022.3204461]

Su S C, Delbracio M, Wang J, Sapiro G, Heidrich W and Wang O. 2017. Deep video deblurring for hand-held cameras//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE: 237-246 [DOI: 10.1109/CVPR.2017.33]

Tao X, Gao H Y, Shen X Y, Wang J and Jia J Y. 2018. Scale-recurrent network for deep image deblurring//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE: 8174-8182 [DOI: 10.1109/CVPR.2018.00853]

Tsai F J, Peng Y T, Lin Y Y, Tsai C C and Lin C W. 2022. Stripformer: strip transformer for fast image deblurring//Proceedings of the 17th European Conference on Computer Vision. Tel Aviv, Israel: Springer: 146-162 [DOI: 10.1007/978-3-031-19800-7_9]

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

Wang Y H, Yu J W and Zhang J. 2023. Zero-shot image restoration using denoising diffusion null-space model//Proceedings of the 11th International Conference on Learning Representations. OpenReview.net

Wang Y J, Yan Q S, Zhou J T, Dai D W and Dong W. 2026. PaAgent: portrait-aware image restoration agent via subjective-objective reinforcement learning [EB/OL]. [2026-04-16]. https://arxiv.org/pdf/2603.17055.pdf

Wang Z D, Cun X D, Bao J M, Zhou W G, Liu J Z, and Li H Q. 2022. Uformer: a general U-shaped transformer for image restoration//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, USA: IEEE: 17662-17672 [DOI: 10.1109/CVPR52688.2022.01716]

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

Wei H Y, Liu S Z, Yuan C and Zhang L. 2025. Perceive, understand and restore: real-world image super-resolution with autoregressive multimodal generative models//Proceedings of 2025 IEEE/CVF International Conference on Computer Vision. Honolulu, USA:

IEEE: 18640-18650 [DOI: 10.1109/ICCV51701.2025.01732]

Wei S, Wu Y, Liang J, Liu S, Wang M and Zhang X. 2020. FDBP-InSAR: an efficient algorithm for InSAR imaging via frequency domain back projection. Remote Sensing, 12(21): #3527 [DOI: 10.3390/rs12213527]

Wei Y Y, Mao T Y, Li B A, Wang F, Li F, Zhang Z, et al. 2025. Visual and large multimodal models promote image restoration and enhancement: research progress. Journal of Image and Graphics, 30(5): 1197-1219 (韦炎炎, 毛天一, 李柏昂, 王飞, 李锋, 张召, 等. 2025. 视觉模型及多模态大模型推进图像复原增强研究进展. 中国图象图形学报, 30(5): 1197-1219) [DOI: 10.11834/jig.240436]

Wu D, Zhao H T and Zheng S B. 2020. Motion deblurring method based on DenseNets. Journal of Image and Graphics, 25(5): 890-899 (吴迪, 赵洪田, 郑世宝. 2020. 密集连接卷积网络图像去模糊. 中国图象图形学报, 25(5): 890-899) [DOI: 10.11834/jig.190400]

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 K, Dong C, Lin L and Loy C C. 2018. Crafting a toolchain for image restoration by deep reinforcement learning//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE: 2443-2452 [DOI: 10.1109/CVPR.2018.00259]

Zamfir E, Wu Z, Mehta N, Tan Y, Paudel D P, Zhang Y, et al. 2025. Complexity experts are task-discriminative learners for any image restoration//Proceedings of 2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, USA: IEEE: 12753-12763 [DOI: 10.1109/CVPR52734.2025.01190]

Zamir S W, Arora A, Khan S, Hayat M, Khan F S, Yang M H, et al. 2021. Multi-stage progressive image restoration//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, USA: IEEE: 14816-14826 [DOI: 10.1109/CVPR46437.2021.01458]

Zamir S W, Arora A, Khan S, Hayat M, Khan F S and Yang M H. 2022. Restormer: efficient transformer for high-resolution image restoration//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, USA: IEEE: 5718-5729 [DOI: 10.1109/CVPR52688.2022.00564]

Zhang K, Zuo W M, Chen Y J, Meng D Y and Zhang L. 2017. Beyond a Gaussian denoiser: residual learning of deep CNN for image denoising. IEEE Transactions on Image Processing, 26(7): 3142-3155 [DOI: 10.1109/TIP.2017.2662206]

Zhang K, Zuo W M and Zhang L. 2018. FFDNet: toward a fast and flexible solution for CNN-based image denoising. IEEE Transactions on Image Processing, 27(9): 4608-4622 [DOI: 10.1109/TIP.2018.2839891]

Zhang Y, Guo X, Ma J, Liu W and Zhang J. 2021. Beyond brightening low-light images. International Journal of Computer Vision, 129(4): 1013-1037 [DOI: 10.1007/s11263-020-01407-x]

Zhang Y H, Zhang J W and Guo X J. 2019. Kindling the darkness: a practical low-light image enhancer//Proceedings of the 27th ACM International Conference on Multimedia. Nice, France: ACM: 1632-1640 [DOI: 10.1145/3343031.3350926]

Zhang Y L, Tian Y P, Kong Y, Zhong B N and Fu Y. 2021. Residual dense network for image restoration. IEEE Transactions on Pattern Analysis and Machine Intelligence, 43(7): 2480-2495 [DOI: 10.1109/TPAMI.2020.2968521]

Zhang Y S, Jia G L, Hu H T, Zhao S X, Zhao K K, Sun L, et al. 2026. TIR-Agent: training an explorative and efficient agent for image restoration [EB/OL]. [2026-05-08]. https://arxiv.org/pdf/2603.27742.pdf

Zheng Z, Ren W, Cao X, Wang T and Jia X. 2021. Ultra-high-definition image HDR reconstruction via collaborative bilateral learning//Proceedings of 2021 IEEE/CVF International Conference on Computer Vision. Montreal, Canada: IEEE

Zhu Y Z, Zhang K, Liang J Y, Cao J Z, Wen B H, Timofte R, et al. 2023. Denoising diffusion models for plug-and-play image restoration//Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Vancouver, USA: IEEE: 1219-1229 [DOI: 10.1109/CVPRW59228.2023.00129]

作者简介

郑萌萌，女，博士研究生，主要研究方向为计算机视觉、图像重建和高光谱图像处理。
E-mail：zhengmengmeng@mail.nwpu.edu.cn
郭宸瑞，通信作者，男，工程师，主要研究方向为目标检测、多模态大模型和雷达光学系统设计。E-mail：g-ch-r@163.com
张磊，男，教授，主要研究方向为计算机视觉、光谱重建、图像超分重建和压缩感知。E-mail：nwpuzhanglei@nwpu.edu.cn
魏巍，男，教授，主要研究方向为计算机视觉、遥感图像处理、机器学习和人工智能。E-mail：weiweinwpu@nwpu.edu.cn