

结合频域物理感知与高阶语义融合的一体化遥感图像复原框架

董 喆, 孙钰哲, 刘天竹, 谷延锋*

(哈尔滨工业大学电子与信息工程学院, 黑龙江哈尔滨 150001)

摘要: 遥感图像在成像链路中常受大气散射、传感器噪声及极端光照等多种异质退化因素的严重影响, 现有一体化复原方法主要依赖隐式特征学习, 缺乏对退化物理频域特性的显式建模及高阶语义交互能力。为此, 本文提出一种结合频域物理感知与高阶语义融合的一体化遥感图像复原框架 Aether。该框架构建频域自适应退化解析器 (Harmonic-Adaptive Degradation Analyzer, HADA), 利用数据驱动的可学习谐波滤波器组替代固定基底变换, 实现对不同退化类型频谱指纹的自适应解析与精准提取。此外, 设计基于 Kolmogorov-Arnold 表示定理的高阶非线性交互融合模块 (Higher-Order Nonlinear Interaction Fusion, HONIF), 通过样条函数网络构建高维映射空间, 突破传统线性注意力的表达瓶颈, 实现退化先验与图像特征的深度语义对齐。在 MD-RSID、MD-RRSHID 及 MDRS-Landsat 三个基准数据集上的实验表明, Aether 在处理雾霾、噪声、模糊及低光等退化问题时均取得最优性能。特别是在 MDRS-Landsat 数据集上, 该方法在去雾任务中的峰值信噪比 (Peak Signal-to-Noise Ratio, PSNR) 较次优方法提升 3.63 dB, 在低光增强任务中 PSNR 提升 1.60 dB, 且可学习感知图像块相似度 (Learned Perceptual Image Patch Similarity, LPIPS) 改善了 75.2%, 有效解决了复杂遥感场景下的一体化通用复原难题。

关键词: 遥感图像复原; 一体化图像复原; 频域自适应; 高阶非线性交互

基金项目: 国家自然科学基金 (No.624B2051)

中图分类号: TP391.41; TP751

文献标识码: A

文章编号: 0372-2112(2026)02-0799-19

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.12263/DZXB.20251152

A Unified Remote Sensing Image Restoration Framework Integrating Frequency-Domain Physical Perception and Higher-Order Semantic Fusion

DONG Zhe, SUN Yuzhe, LIU Tianzhu, GU Yanfeng*

(School of Electronics and Information Engineering, Harbin Institute of Technology, Harbin, Heilongjiang 150001, China)

Abstract: Remote sensing imagery is often severely degraded along the imaging chain by heterogeneous factors such as atmospheric scattering, sensor noise, and extreme illumination conditions. Existing all-in-one restoration approaches predominantly rely on implicit feature learning, lacking explicit modeling of the physical frequency-domain properties of degradation and the ability to capture higher-order semantic interactions. To address these limitations, we present Aether, a unified remote sensing image restoration framework that integrates frequency-domain physical perception with higher-order semantic fusion. Aether introduces a harmonic-adaptive degradation analyzer (HADA), which replaces fixed-basis transforms with a data-driven learnable harmonic filter bank. This enables adaptive parsing and precise extraction of spectral fingerprints associated with diverse degradation types. In addition, we design a higher-order nonlinear interaction fusion module (HONIF) grounded in the Kolmogorov-Arnold representation theorem. HONIF constructs a high-dimensional mapping space via spline-based function networks, overcoming the representational bottleneck of conventional linear attention and facilitating deep semantic alignment between degradation priors and image features. Experiments on three benchmark datasets—MD-RSID, MD-RRSHID, and MDRS-Landsat—demonstrate that Aether achieves state-of-the-art performance across haze removal, denoising, deblurring, and low-light enhancement tasks. Notably, on the MDRS-Landsat dataset, Aether surpasses the second-best method by 3.63 dB in peak signal-to-noise ratio (PSNR) for dehazing and by 1.60 dB for low-light enhancement, while improving learned perceptual image patch similarity (LPIPS) by 75.2%, effectively addressing the long-standing challenge of unified and generalizable restoration in complex remote sensing scenarios.

Keywords: remote sensing image restoration; all-in-one image restoration; harmonic-adaptive; higher-order nonlinear interaction

Foundation Item(s): National Natural Science Foundation of China (No.624B2051)

0 引言

遥感成像作为国家对地观测体系的重要基础环节,在环境监测、灾害评估、城市精细化治理等众多领域发挥着不可替代的作用^[1-2]。随着多光谱、高分辨率和多时相卫星的持续组网,遥感图像在空间、时间与光谱维度的丰富性显著提升。然而,真实成像过程中不可避免地受到大气散射、传感器电子噪声、平台姿态扰动以及极端光照变化等多物理因素的干扰^[3-4]。这些退化机制在空间上呈现强烈的非均质特性,在频域上体现出明显的能量分布差异。其结果不仅表现为辐射精度下降与结构清晰度损失,而且直接影响到目标检测、场景理解与变化检测等下游视觉任务的性能边界^[5-6]。因此,构建能够同时应对多类退化并具备物理一致性的统一遥感图像复原方法,具有重要科学意义与应用需求。

传统图像复原方法多基于物理退化模型或先验正则化,通过变分优化实现逆过程求解^[7]。尽管其具有较好的可解释性,但无法充分刻画真实遥感成像链路中高度非线性、空间非平稳、退化耦合等复杂现象,因而在实际场景中容易出现性能瓶颈。深度学习方法凭借其强大的非线性拟合能力显著提升了复原效果^[8-9],但早期方法通常面向单一退化类型设计结构,一旦面对真实遥感场景中类型未知的“异质退化”,便需预先判定并分别调用模型,不仅导致计算开销线性累积,更难以有效提炼不同退化类型之间潜在的共享表征^[10]。近年来出现的“一体化”复原框架通过动态卷积、提示编码或专家混合等手段在单一网络中处理多种退化^[11-13],虽然在自然图像领域取得一定进展,但在遥感复原任务中仍存在三大核心局限,具体如图1所示。

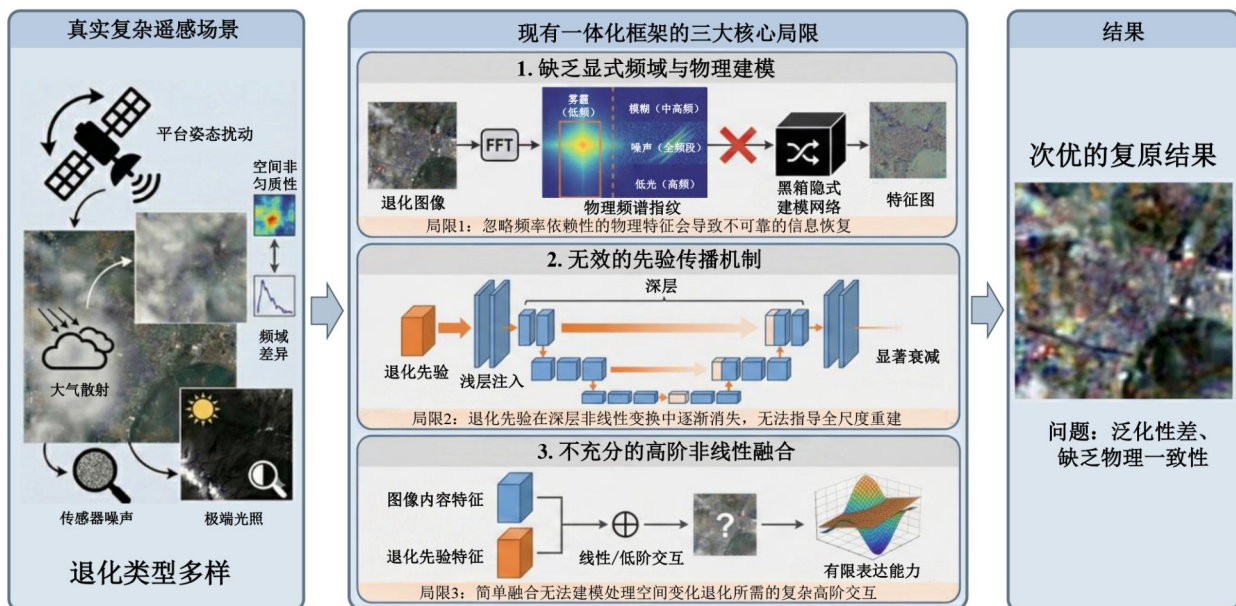


图1 真实复杂场景下现有一体化遥感图像复原框架面临的关键挑战

Figure 1 Key challenges faced by existing all-in-one remote sensing image restoration frameworks in real-world complex scenarios

首先,现有一体化框架大多依赖深度网络通过端到端学习隐式推断退化类型,几乎未对遥感成像的物理属性及其频域特征进行显式刻画^[14]。遥感图像退化具有显著的频率依赖性:雾霾主要压制低频能量;噪声在全频段产生均匀扰动;运动模糊沿特定方向削弱中高频细节;低光则因信噪比下降导致中高频显著衰减,而低频仅呈整体下移^[15-16]。如果缺乏频域物理规律的约束,仅凭隐式学习的特征表示容易偏离真实

的辐射与结构分布,导致复原结果在量化遥感应用中缺乏可靠性。其次,现有方法普遍采用单点式的先验注入策略,将退化信息在网络浅层输入后,任其在深层传播过程中逐渐衰减^[17]。由于深度网络的多层非线性变换会不断重构特征空间,退化先验往往无法在从全局结构重建到局部纹理生成的多尺度流程中持续发挥指导作用。最后,目前主流的融合方式基本基于线性或二阶交互,如通道拼接后卷积或标准点积

注意力^[18-19]。这些低阶变换在函数逼近空间中具有先天表达上界,难以建模退化先验与图像内容之间潜在的高阶非线性关系,也难以适应遥感场景中高度复杂的空间异质性。

上述限制表明,现有统一复原框架在物理一致性、先验传播机制以及非线性融合表达能力方面均存在结构性短板,尚不足以支撑真实遥感场景下的复杂退化建模^[20]。然而,从频域角度观察,各类退化在能量分布上呈现稳定而具有区分度的“频谱指纹”^[21-22];从空间角度观察,真实退化具有显著的局部差异化,要求复原网络在不同区域实施强度不同的校正策略。因此,一个能够结合频域物理建模、层次化先验传播、多尺度空间适应性和高阶非线性融合的新型一体化遥感图像复原框架,是实现高保真、大范围适应性复原的关键。

基于上述分析,本文提出一种结合频域物理感知与高阶语义融合的一体化遥感图像复原框架 Aether。该方法首次在统一复原任务中将可学习谐波频域分析与高阶函数逼近理论相结合,通过自适应频谱分解模块(Harmonic-Adaptive Degradation Analyzer, HADA)显式建模退化类型在频域的差异性特征,并通过高阶非线性交互融合模块(Higher-Order Nonlinear Interaction Fusion, HONIF)利用样条函数构建强表达力的非线性空间,实现退化先验与图像内容间深层语义对齐。进一步地,通过在解码器不同尺度中嵌入 HADA-HONIF 协同结构,构建跨层级的退化引导机制,使退化信息能够在结构复原、局部重建和纹理增强的全流程中持续发挥作用,从根本上克服了传统方法中先验衰减的问题。

本文的主要贡献体现在以下几个方面:

(1)提出一种融合频域物理建模、层次化先验传播与高阶非线性交互的新型一体化遥感图像复原范式,理论上突破现有统一框架的表达限制。

(2)设计可学习谐波滤波器组,实现面向不同退化类型的数据驱动频谱分解,有效提升退化表征的判别能力。

(3)构建基于 Kolmogorov-Arnold 表示定理的高阶非线性融合模块,从根本上增强退化先验与图像内容的深度融合能力。

(4)提出层级化多尺度复原架构,在整个重建流程中实现退化先验的持续有效引导,显著提升真实遥感场景中的复原稳定性与细节保真度。

1 相关工作

1.1 一体化图像复原

图像复原作为计算机视觉的基础任务,经历了从

任务特定到任务统一的重要范式转变。传统方法主要针对单一退化类型设计专门化模型,如去噪、去模糊、超分辨率等。然而,真实场景中的图像往往同时受到噪声、模糊、压缩伪影以及恶劣天气等多种退化因素的复合影响。这种现实需求催生了一体化图像复原研究范式的兴起,旨在构建统一框架使单一模型能够自适应处理多种未知退化类型。

早期探索性工作为一体化复原奠定了基础。TransWeather^[23]创新性地将 Transformer 引入复原任务,通过天气类型查询和可学习退化表征实现了无需任务标定的自适应处理。AirNet 进一步推进了盲复原技术,借助对比学习机制实现了无需退化类型先验的特征提取和重建。近年来,架构设计取得了系列突破。Restormer 提出高效 Transformer 架构,通过门控多头自注意力机制在保持全局感受野的同时显著降低了计算复杂度。MambaIR^[24]首次将选择性结构化状态空间模型引入图像复原,针对标准 Mamba 在低层视觉任务中的局部像素遗忘和通道冗余问题,设计了残差状态空间块融合局部增强和通道注意力机制,在保持全局感受野的同时充分利用图像的局部先验,在多项复原任务上均优于 Transformer 基线且具有更高计算效率。

先进的学习策略成为提升一体化复原性能的关键。提示学习已成为该领域的主流范式, PromptIR 创新性地将可学习视觉提示作为轻量级自适应模块,通过多尺度退化感知引导实现对不同退化类型的自适应处理。多模态提示学习标志着向人机交互式复原的演进, DA-CLIP^[25]提出退化感知控制器,通过适配 CLIP 图像编码器预测高质量特征嵌入并集成到复原网络中,实现对多种退化类型的统一处理。Perceive-IR^[26]采用多层次质量驱动提示机制增强模型对不同程度退化的识别能力。多任务学习策略不断成熟,GRIDS^[27]基于退化相关性设计任务分组机制减少任务间干扰, TUR^[28]提出任务不确定性正则化方法动态调整不同任务的损失权重,显著提升了多任务训练的稳定性。

扩散模型作为强大的生成式先验模型为图像复原带来了新的技术路径。DiffIR^[29]提出高效扩散框架,通过在紧凑的图像复原先验空间而非完整图像空间进行扩散过程,显著降低了迭代次数并提高了采样效率。AutoDIR^[30]提出自动化一体化复原系统,结合盲图像质量评估阶段自动检测未知图像退化,以及基于潜在扩散的一体化复原阶段处理多种退化类型,展现出作为图像复原基础模型的潜力。大规模预训练模型的引入代表了前沿发展方向, RestoreAgent^[31]构建了首个具有自主退化分析和自适应模型选择能力的

智能体系统,通过多轮对话和工具调用实现了从固定流程向模拟人类认知过程的动态系统演进。近期综述性工作系统梳理了一体化图像复原的技术脉络,jiang等人^[32]建立了按架构设计、学习策略和应用场景分类的结构化分类体系,整合了广泛使用的数据集、评估协议和实现实践,为该领域的发展提供了重要参考。

1.2 一体化遥感图像复原

遥感图像复原研究面临着区别于自然图像的独特挑战。遥感图像在获取过程中受到大气散射、传感器噪声、运动模糊、光照变化以及气象条件等多重因素的复合影响,这些退化因素往往交织出现且空间分布非均匀。相比自然图像,遥感图像具有成像距离远导致大气影响显著、传感器特性差异大、场景复杂度高以及分辨率跨度大等特性。早期遥感图像复原研究主要针对特定退化类型设计专门方法,如去雾、去噪、去模糊等单一任务。然而,实际遥感场景中的复合退化需求推动了一体化复原框架的发展,使统一架构能够自适应处理多种退化类型,显著提升模型的实用性和部署效率。

模型驱动方法通过融合物理退化模型与深度学习架构,在提升性能的同时增强了可解释性,成为一体化遥感图像复原的重要研究方向。HDI-PRNet^[33]基于高阶退化成像理论框架,结合马尔可夫特性和最大后验估计提出渐进式复原网络,通过深度展开框架将优化算法的迭代过程转化为可学习的网络层。该方法包括基于近端映射先验的去噪模块、整合诺依曼级数展开与双域退化的去模糊模块以及超分辨率模块,突破了传统一阶模型的局限,实现了去噪、去模糊和超分辨率任务的渐进式处理。Ada4DIR^[34]提出自适应模型驱动的一体化网络,创新性地整合四种退化模型(去噪、去模糊、去雾和低光增强)与退化驱动融合Transformer模块。该方法运用提示学习机制根据退化类型引导复原进程,通过模型驱动复原模块在特征空间实现退化特征到清晰表征的转换。相比数据驱动方法,模型驱动方法能够利用物理先验减少对大规模训练数据的依赖,提供更好的可解释性和对分布外数据的泛化能力。

最新的研究探索了物理引导的专家混合系统用于一体化遥感复原。PhyDAE^[35]提出物理引导的退化自适应专家网络,采用两阶段级联架构将退化信息从隐式特征转换为显式决策信号,实现对雾霾、噪声、模糊和低光等多种异构退化的精确识别和差异化处理。该方法设计了频率感知退化分解器(FADD)和残差流形投影器(RMP)模块,前者通过多尺度频率分解策略捕获退化特定的频谱特征,后者通过流形学

习建立退化模式与复原策略之间的自适应映射。针对不同退化类型,PhyDAE构建了专门的物理模型引导的专家网络:针对雾霾采用大气散射模型引导的去雾专家;针对噪声采用空间噪声图估计引导的去噪专家;针对模糊采用参数化各向异性高斯模糊核引导的去模糊专家;针对低光采用Retinex理论引导的增强专家。这种物理先验与深度学习的深度融合,显著提升了模型的泛化能力和可解释性。

多退化建模与交互学习策略对于处理遥感图像复杂退化模式具有关键作用。SGIL^[36]针对双重退化(外部环境退化如雾霾、雨雪与内部传感器噪声)场景,提出基于显著性引导的交互学习统一框架,通过伪像素监督的显著性分析识别主要退化区域,并设计任务感知交互学习模块实现不同退化类型的协同建模。不同退化类型往往在空间分布和频谱特性上存在差异,通过显著性分析可以更有针对性地处理不同区域的退化,避免简单的全局处理可能带来的过度平滑或欠复原问题。此外,基于低秩约束的多视角高光谱图像融合方法^[37]通过挖掘不同角度间的互补信息与内在相关性,同时实现超分辨率重建和几何配准,展示了多任务协同处理的潜力。

尽管一体化遥感图像复原取得了重要进展,但构建真正意义上的统一框架仍面临诸多挑战。首先,多退化联合训练中的任务冲突问题尚需更有效的优化策略,特别是在不同退化类型的优化目标存在矛盾时。其次,退化类型的精确识别在复杂真实场景中仍存在困难,特别是面对未见过的退化组合时。再次,在多样化成像条件和传感器类型下保持光谱保真度的严苛要求,需要更精细的物理建模和约束设计。最后,计算效率与复原质量的平衡对于实际应用至关重要,特别是在边缘设备和实时处理场景中。未来研究需要在深化物理模型与深度学习融合、提升跨传感器跨场景泛化能力、构建大规模多退化遥感数据集等方面持续探索,以应对真实遥感应用中复杂多变的退化挑战。

2 方法

2.1 问题建模与模型架构

遥感图像在实际应用中面临多样化退化挑战,包括大气散射引起的雾霾效应、传感器系统噪声、低照度信号衰减以及平台运动模糊。各类退化具有独特的物理机制与频域特征:雾霾主要削弱低频并压低全局对比度;低光由于信噪比降低显著衰减中高频细节;噪声在全频段叠加扰动;而模糊则在特定方向抑制中高频能量。现有一体化复原方法面临三个核心难题:首先是退化类型的准确识别与特征表征问题,

由于只能从单张退化图像进行推断,网络难以将场景固有的视觉特征与退化效应分离;其次是退化先验与图像内容的自适应融合问题,不同退化类型和不同退化程度要求差异化的复原策略,网络需要在重度退化区域施加强力校正,同时在轻度退化或清晰区域保持原始信息不被过度处理;最后是退化引导信息在深度网络中的持续有效传播问题。现有方法多采用单点注入策略,退化先验信息在经过多层特征变换和非线性激活后逐渐衰减,无法在从粗粒度语义复原到细粒度纹理重建的多尺度复原过程中持续发挥作用。

针对上述挑战,如图 2 所示,本文提出 Aether 一体化复原框架,其在视觉上下文引导范式的基础上,创新性地从频域信号处理理论与高阶函数逼近理论相结合,构建双路协同的复原架构。Aether 通过引入配对上下文样本 $C=I_d, I_c$, 其中, $I_d \in \mathbb{R}^{3 \times H \times W}$ 为与待复原图像 I_q 具有相同退化类型的参考样本, $I_c \in \mathbb{R}^{3 \times H \times W}$ 为其对应的清晰真值,构建显式的退化特征解耦机

制。在训练阶段,配对上下文从同一退化类型的样本中随机选取;在测试阶段,若退化类型已知则直接选取对应类型样本,若未知则通过基于 CLIP 特征的轻量级分类器进行退化类型预判后选取。相比于直接注入基于物理模型的退化先验,配对上下文机制能够通过数据驱动方式捕捉真实场景中难以用统一物理参数建模的细粒度退化特性和空间非均匀性,为后续复原提供更精准的实例级退化表征。此外,提出了 HADA 和 HONIF 两大关键组件。HADA 模块突破传统固定频率分解的局限,通过数据驱动的可学习谐波滤波器实现自适应频谱分解,能够针对不同退化类型自动学习最优的频率划分策略,精确捕获退化在频域的差异化指纹特征。HONIF 模块基于 Kolmogorov-Arnold 表示定理,通过样条函数网络构建高阶非线性映射空间,突破传统线性融合和低阶多项式变换的表达局限,实现退化先验与图像特征的深度语义对齐和空间自适应调制。

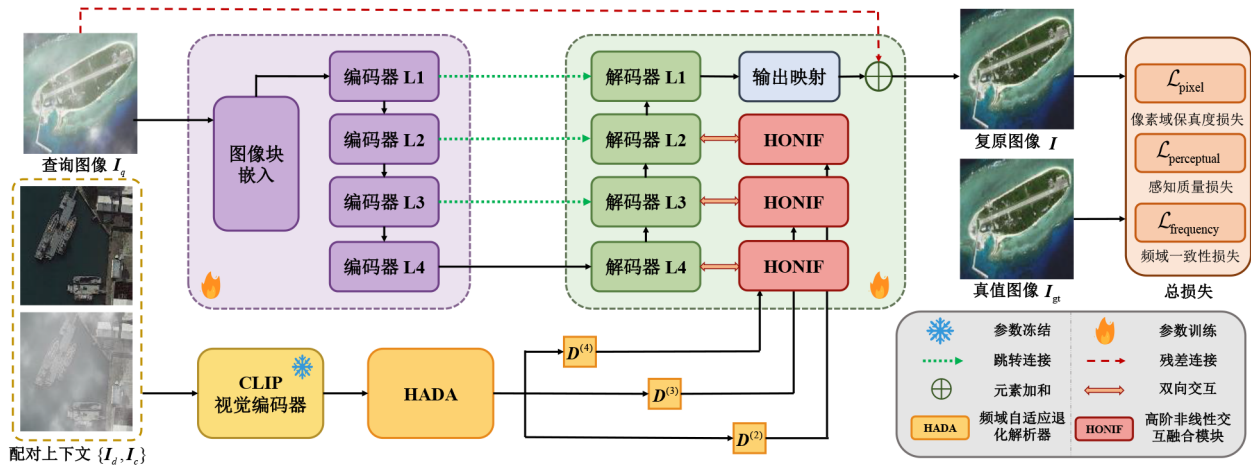


图 2 Aether 一体化遥感图像复原框架整体架构示意图

Figure 2 Overall architecture of the Aether all-in-one remote sensing image restoration framework

Aether 整体采用编码器-解码器架构,遵循由粗到细的渐进式复原策略,在不同空间分辨率下分别处理全局语义复原、中尺度结构修复和细粒度纹理重建任务。编码器通过重叠图像块嵌入层将输入查询图像 I_q 映射至初始特征空间 $F_0 \in \mathbb{R}^{48 \times H_0 \times W_0}$, 随后经过四个层级的特征提取,每个层级由多个 Transformer 块构成以捕获多尺度语义表征。第 i 层级 ($i \in [1, 2, 3, 4]$) 的特征维度设置为 $C_i = 48 \cdot 2^{i-1}$, 空间分辨率为 $(H_i, W_i) = (H_0/2^{i-1}, W_0/2^{i-1})$, 层级间通过下采样模块实现 $2 \times$ 空间下采样和通道数扩张。编码器最深层级的输出 $F_{enc}^{(4)} \in \mathbb{R}^{384 \times H/8 \times W/8}$ 进入潜在特征处理模块,该模块包含若干 Transformer 块以提取高层语义表征 F_{latent} , 这一潜在表征捕获了图像的全局结构和语义信息,为后续的解码重建提

供语义锚点。

解码器设计是 Aether 架构创新的核心所在。解码器包含四个对称层级 $l \in \{4, 3, 2, 1\}$, 每个层级首先通过上采样模块将低分辨率特征恢复至更高分辨率,上采样模块采用 1×1 卷积后接像素重排实现 $2 \times$ 空间上采样和 $2 \times$ 通道数压缩。上采样后的特征与编码器对应层级的跳跃连接特征拼接,经通道压缩卷积降维后送入 Transformer 块序列进行特征细化。解码器的每个层级 $l \in \{4, 3, 2\}$ 都集成了 HADA 和 HONIF 模块构成的退化引导分支,形成层次化的协同复原机制。具体而言, HADA 模块以配对上下文 C 为输入,输出层次化的退化表征 $D^{(l)} \in \mathbb{R}^{2L \times d_l}$, 其中, $L=50$ 为 CLIP 特征的 token 数量, $d_l \in \{320, 128, 64\}$ 分别对应解码器层

级 $l \in \{4, 3, 2\}$ 的退化表征维度。随后, HONIF 模块接收解码器特征 $F_{\text{dec}}^{(l)}$ 和退化表征 $D^{(l)}$, 通过高阶非线性交互生成退化感知的增强特征 $F_{\text{fused}}^{(l)}$ 。该增强特征与原始解码器特征在通道维度拼接后, 经过额外的 Transformer 块和通道压缩卷积处理, 生成融合退化先验的复原特征并传递至下一层级。

在解码器深层 ($l=4$), 特征图空间分辨率为 $(H/8) \times (W/8)$, 网络感受野覆盖大范围区域, 此时 HADA-HONIF 协同模块主要负责粗粒度的语义复原和全局退化模式识别。在解码器中层 ($l=3$), 特征图空间分辨率提升至 $(H/4) \times (W/4)$, 协同模块转向中粒度的局部结构调整, 关注区域性的退化校正和结构细节复原。在解码器浅层 ($l=2$), 特征图空间分辨率达到 $(H/2) \times (W/2)$, 协同模块专注于细粒度的纹理重建和边缘增强, 网络具备足够的空间精度复原精细的图像细节和低频信息。最终, 最浅层解码器输出 $F_{\text{dec}}^{(0)} \in \mathbb{R}^{96 \times H \times W}$ 经过若干精细化 Transformer 块进行最终的纹理细节优化, 得到精细化特征 F_{refine} 。输出层采用 3×3 卷积将其映射回 RGB 空间, 与输入图像进

行残差连接生成最终复原结果:

$$I_{\text{clean}} = I_q + \text{Conv}_{3 \times 3}(F_{\text{refine}}) \quad (1)$$

综上所述, Aether 这种从宏观到微观的层次化复原流程, 通过 HADA 和 HONIF 模块在多个解码层级的协同作用, 确保了不同粒度的退化先验信息能够在相应的特征抽象层次持续发挥作用, 系统性地解决了退化识别、特征解耦和自适应融合三大核心难题, 避免了传统单点注入方法中退化信息在深度传播过程中的逐渐衰减问题。

HADA 是 Aether 框架中退化特征提取的核心组件, 如图 3 所示, 其设计理念源于不同退化类型在频域空间呈现出显著差异化的能量分布模式这一物理观测。传统方法通常采用固定的频率分解策略, 如离散余弦变换或小波变换, 将信号分解为预定义的频带。然而, 这种固定分解范式存在根本性局限, 各类退化具有独特的物理机制与频域特征: 雾霾主要削弱低频并压低全局对比度; 低光由于信噪比降低显著衰减中高频细节; 噪声在全频段叠加扰动; 而模糊则在特定方向抑制中高频能量。固定的频率划分无法自适应不同退化的频谱特征, 导致退化表征的判别性不足。

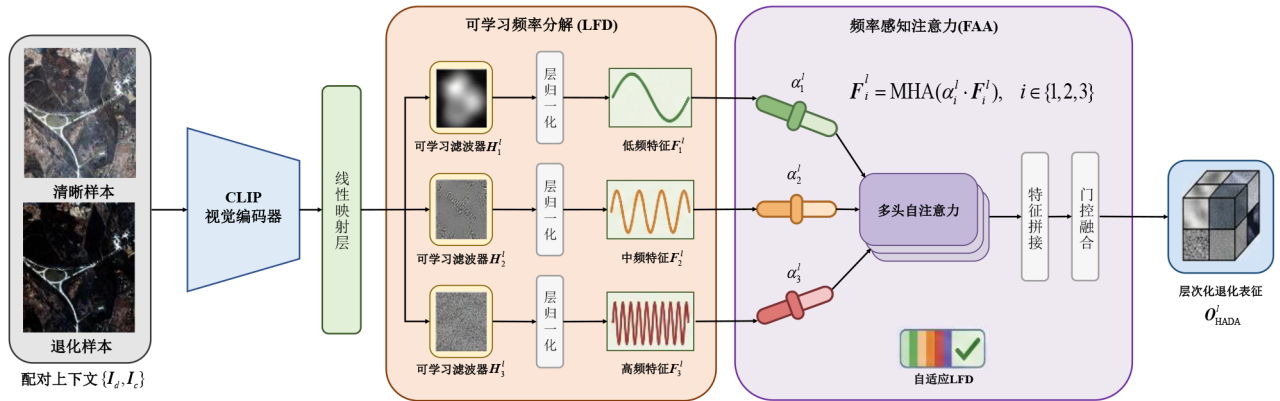


图 3 HADA 模块结构图

Figure 3 Architecture of the HADA module

2.2 HADA

HADA 模块通过引入可学习的谐波分解机制突破上述限制, 实现数据驱动的自适应频谱分析。该模块的输入为经 CLIP 视觉编码器提取的上下文特征 $E_C \in \mathbb{R}^{2L \times D}$, 其中, $L=50$ 为 ViT-B/32 架构产生的空间 token 数量, $D=768$ 为 CLIP 嵌入维度。配对上下文 $C=\{I_d, I_c\}$ 中的退化样本 I_d 和清晰样本 I_c 经 CLIP 编码后在 token 维度拼接, 形成联合表征 E_C 。为适配后续频域分解模块的计算需求并降低复杂度, 首先通过线性投影层将高维 CLIP 特征压缩至退化表征空间:

$$P_C^l = \text{GELU}(W_{\text{proj}}^l E_C + b_{\text{proj}}^l), P_C^l \in \mathbb{R}^{2L \times C_l} \quad (2)$$

其中: $W_{\text{proj}}^l \in \mathbb{R}^{C_l \times D}$ 为第 l 层级的投影权重矩阵;

$C_l \in \{64, 128, 320\}$ 为对应解码器层级 $l \in \{2, 3, 4\}$ 的退化表征维度; GELU 激活函数引入非线性变换能力。

HADA 的核心创新在于可学习频率分解 (Learnable Frequency Decomposition, LFD) 子模块。不同于传统基于正交基的固定分解, LFD 通过参数化的谐波滤波器组实现自适应频谱划分。具体而言, LFD 维护 $N_b=3$ 个可学习的频带滤波器 $\{H_i^l\} (i=1, 2, \dots, N_b)$, 每个滤波器 $H_i^l \in \mathbb{R}^{C_l \times C_l}$ 在退化表征空间中执行线性变换, 有

$$F_i^l = \text{LayerNorm}(P_C^l H_i^l), i=1, 2, \dots, N_b \quad (3)$$

值得注意的是, 这里的滤波器矩阵 H_i^l 不是预先定义的正交基 (如傅里叶基或小波基), 而是通过端到端学习自动发现最优的频谱分解策略。在训练过

程中, H_1^l 倾向于捕获低频全局亮度和大尺度结构变化; H_2^l 响应中频纹理和边缘信息衰减(对模糊退化敏感); H_3^l 专注于高频细节扰动检测。LayerNorm 的引入确保各频带分量在相同的数值尺度下进行后续处理, 避免梯度传播的不稳定性。

为进一步提升频带表征的判别能力, HADA 引入频率感知注意力(Frequency-Aware Attention, FAA)机制, 对分解得到的各频带特征进行加权融合。FAA 为每个频带分配可学习的重要性权重 $\{\alpha_i^l\} (i=1, 2, \dots, N_b)$, 并通过多头自注意力机制捕获频带间的相互依赖关系:

$$\tilde{F}_i^l = \text{MHA}(\alpha_i^l \cdot F_i^l, \alpha_i^l \cdot F_i^l, \alpha_i^l \cdot F_i^l), \alpha_i^l \in \mathbb{R}^+ \quad (4)$$

$$O_{\text{freq}}^l = \text{Concat}(\tilde{F}_1^l, \tilde{F}_2^l, \tilde{F}_3^l) \in \mathbb{R}^{2L \times (N_b \cdot C_l)} \quad (5)$$

其中, MHA 表示多头自注意力操作, 每个频带的加权表征 \tilde{F}_i^l 通过自注意力机制内部建模 token 间的空间依赖关系。频带权重 α_i^l 初始化为 $1/N_b$ 以确保训练初期各频带贡献均衡, 随训练进程自适应调整以突出对当前退化类型最具判别力的频带。将各频带的加权表征在通道维度拼接后, 经过门控融合网络进一步整合:

$$O_{\text{HADA}}^l = \text{GELU}(\mathbf{W}_{\text{fuse}}^l O_{\text{freq}}^l + \mathbf{b}_{\text{fuse}}^l) \in \mathbb{R}^{2L \times C_l} \quad (6)$$

其中, $\mathbf{W}_{\text{fuse}}^l \in \mathbb{R}^{C_l \times (N_b \cdot C_l)}$ 为跨频带融合的权重矩阵。最终输出的退化表征 O_{HADA}^l 包含了从配对上下文中提取的、经频域自适应分解和注意力加权后的高判别性退化特征。

通过数据驱动的方式学习最优频谱分解策略, HADA 能够针对不同退化类型自动发现其频域指纹

特征, 显著提升退化识别和特征解耦的准确性。

2.3 HONIF

HONIF 负责将 HADA 提取的退化先验 O_{HADA}^l 与解码器特征图 F^l 进行深度语义对齐和空间自适应调制。现有融合策略主要包括简单的通道拼接后卷积、双线性池化或低阶注意力机制(如点积注意力), 这些方法在函数逼近能力上存在本质局限。线性融合无法捕获退化先验与图像特征之间的复杂非线性交互关系, 而传统多层感知机(MultiLayer Perceptron, MLP)虽然具备万能逼近能力, 但其将非线性变换固定在节点激活函数中, 对输入特征维度间的高阶交互建模能力有限。

启发于 Kolmogorov-Arnold 表示定理中“高维连续函数可由有限个单变量函数的叠加与复合构成”的思想, 本文重新审视了交叉注意力中 $Q/K/V$ 生成过程的表达能力。传统线性投影本质上仅提供低阶线性变换, 其非线性建模能力主要依赖后续注意力矩阵的加权过程, 难以充分刻画解码器特征与退化先验之间固有的高阶耦合关系。为此, 本文提出了 HONIF 模块, 如图 4 所示, 通过在注意力机制的输入端引入可学习的单变量函数族, 以构建更具表达力的跨模态交互结构。具体而言, HONIF 采用 KAN 增强的交叉注意力架构(KAN-enhanced Cross-Attention), 将 Q, K, V 的线性映射替换为由 B-spline 参数化的可学习单变量函数, 实现基于“边函数”的高阶非线性变换, 从而在保持注意力框架整体结构不变的前提下显著提升特征重参数化能力。

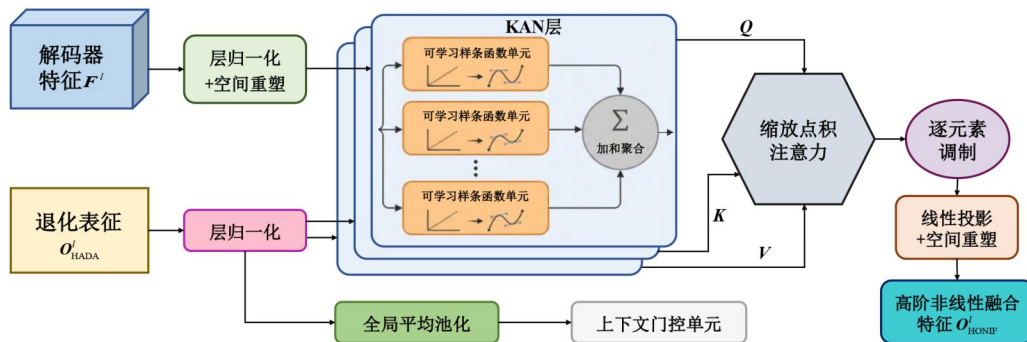


图 4 HONIF 模块结构图

Figure 4 Architecture of the HONIF module

具体而言, 给定解码器特征 $F^l \in \mathbb{R}^{C_l \times H_l \times W_l}$ 和退化表征 $O_{\text{HADA}}^l \in \mathbb{R}^{2L \times C_l}$, 首先对 F^l 进行空间展平和归一化:

$$F_{\text{flat}}^l = \text{LayerNorm}(\text{Reshape}(F^l)) \in \mathbb{R}^{(H_l \cdot W_l) \times C_l} \quad (7)$$

$$O_{\text{norm}}^l = \text{LayerNorm}(O_{\text{HADA}}^l) \in \mathbb{R}^{2L \times C_l} \quad (8)$$

随后, 通过 KAN 层生成注意力三元组。以查询生成为例, KAN 层首先对输入特征 F_{flat}^l 的每个维度 c

进行归一化并映射到 $[-1, 1]$ 区间:

$$\mathbf{x}_{\text{norm}} = \tanh(F_{\text{flat}}^l) \in \mathbb{R}^{(H_l \cdot W_l) \times C_l} \quad (9)$$

接着, 对每个输入维度 c 和输出维度 o , KAN 维护一组网格点 $\{g_j\} (j=1, 2, \dots, G)$, 其中 $G = N_g + N_s$ 为网格维度 ($N_g = 5$ 为网格大小, $N_s = 3$ 为样条阶数)。定义径向基函数(RBF)作为样条基函数的替代, 以高斯核形式表示:

$$\phi_j(x) = \exp\left(-\frac{(x-g_j)^2}{2\sigma^2}\right), \sigma = \frac{2}{G-1} \quad (10)$$

其中, g_j 为均匀分布在 $[-1, 1]$ 区间的网格点。对归一化后的输入 $x_{n,c}$ (第 n 个 token 的第 c 个维度), 计算其在所有网格点上的基函数响应并归一化:

$$\begin{aligned} \mathbf{b}_{n,c} &= [\phi_1(x_{n,c}), \phi_2(x_{n,c}), \dots, \phi_G(x_{n,c})], \\ \tilde{\mathbf{b}}_{n,c} &= \frac{\mathbf{b}_{n,c}}{\sum_{j=1}^G \phi_j(x_{n,c}) + \epsilon} \end{aligned} \quad (11)$$

其中, $\epsilon = 10^{-8}$ 为数值稳定性常数。KAN 层维护可学习的样条权重张量 $\mathbf{S}_Q^l \in \mathbb{R}^{d_{\text{inner}} \times C_l \times G}$, 其中, $d_{\text{inner}} = n_h \cdot d_h$ 为注意力机制的内部维度 (n_h 为头数, d_h 为每头维度)。样条输出通过爱因斯坦求和约定计算:

$$\mathbf{q}_{\text{spline}} = \sum_{c=1}^{C_l} \sum_{j=1}^G \tilde{\mathbf{b}}_{:,c,j} \cdot \mathbf{S}_{Q:,c,j}^l \in \mathbb{R}^{(H_l \cdot W_l) \times d_{\text{inner}}} \quad (12)$$

此外, KAN 层包含一个标准线性基底分支 $\mathbf{q}_{\text{base}} = \mathbf{F}_{\text{flat}}^l \mathbf{W}_Q^{\text{base}}$, 最终查询为两者之和:

$$\mathbf{Q}^l = \mathbf{q}_{\text{spline}} + \mathbf{q}_{\text{base}} \in \mathbb{R}^{(H_l \cdot W_l) \times d_{\text{inner}}} \quad (13)$$

类似地, 键和值通过 KAN 层从退化表征生成:

$$\begin{aligned} \mathbf{K}^l &= \text{KAN}_K(\mathbf{O}_{\text{norm}}^l) \in \mathbb{R}^{2L \times d_{\text{inner}}}, \\ \mathbf{V}^l &= \text{KAN}_V(\mathbf{O}_{\text{norm}}^l) \in \mathbb{R}^{2L \times d_{\text{inner}}} \end{aligned} \quad (14)$$

标准的缩放点积注意力机制计算如下:

$$\text{Attn}^l = \text{Softmax}\left(\frac{\mathbf{Q}^l (\mathbf{K}^l)^T}{\sqrt{d_h}}\right) \in \mathbb{R}^{(H_l \cdot W_l) \times 2L} \quad (15)$$

$$\mathbf{O}_{\text{attn}}^l = \text{Attn}^l \mathbf{V}^l \in \mathbb{R}^{(H_l \cdot W_l) \times d_{\text{inner}}} \quad (16)$$

为实现空间自适应调制, HONIF 引入上下文门控机制 (Context Gating)。该机制通过小型 MLP 网络学习全局上下文强度, 动态控制退化先验对特征图的影响程度:

$$\mathbf{g}^l = \sigma\left(\mathbf{W}_2^l \cdot \text{GELU}(\mathbf{W}_1^l \bar{\mathbf{O}}_{\text{norm}}^l)\right) \in \mathbb{R}^1 \quad (17)$$

其中: $\bar{\mathbf{O}}_{\text{norm}}^l = \frac{1}{2L} \sum_{i=1}^{2L} \mathbf{O}_{\text{norm},i}^l$ 为退化表征的全局平均池化特征; σ 为 Sigmoid 激活函数; $\mathbf{W}_1^l \in \mathbb{R}^{(C_l/4) \times C_l}$ 和 $\mathbf{W}_2^l \in \mathbb{R}^{1 \times (C_l/4)}$ 构成两层瓶颈 MLP。门控系数 \mathbf{g}^l 对注意力输出进行逐 token 调制:

$$\tilde{\mathbf{O}}_{\text{attn}}^l = \mathbf{O}_{\text{attn}}^l \odot \mathbf{g}^l \in \mathbb{R}^{(H_l \cdot W_l) \times d_{\text{inner}}} \quad (18)$$

最后, 调制后的特征通过线性投影映射回原始通道维度, 并重塑为空间形式:

$$\mathbf{O}_{\text{HONIF}}^l = \text{Reshape}(\tilde{\mathbf{O}}_{\text{attn}}^l \mathbf{W}_O^l) \in \mathbb{R}^{C_l \times H_l \times W_l} \quad (19)$$

其中, $\mathbf{W}_O^l \in \mathbb{R}^{d_{\text{inner}} \times C_l}$ 为输出投影矩阵。

HONIF 模块通过将 KAN 的高阶函数逼近能力与注意力机制的长程依赖建模能力相结合, 从而能够在更低的参数开销下学习输入特征维度间的复杂非线性交互模式。

性交互模式。

2.4 多维度联合优化策略

Aether 的训练优化策略从像素域保真度、感知质量和频域一致性三个维度协同约束网络学习过程, 以实现高保真度的全方位复原效果。

像素域保真度损失构成基础约束项, 采用 L_1 范数衡量复原图像 $\hat{\mathbf{I}}$ 与真值图像 \mathbf{I}_{gt} 之间的逐像素差异:

$$\mathcal{L}_{\text{pixel}} = \frac{1}{CHW} \sum_{c=1}^C \sum_{h=1}^H \sum_{w=1}^W |\hat{\mathbf{I}}_{c,h,w} - \mathbf{I}_{\text{gt},c,h,w}| \quad (20)$$

L_1 损失相较于 L_2 损失对离群值的处理更为鲁棒, 在边缘和纹理区域提供更清晰的重建。

感知质量损失引入预训练 VGG-19 网络的深层特征表征, 在感知空间而非像素空间约束复原结果, 提升主观视觉质量。具体地, 提取 VGG 网络中 relu3_3 层的特征映射, 计算特征域的 L_1 距离:

$$\mathcal{L}_{\text{perceptual}} = \frac{1}{C} \sum_{c=1}^C \left\| \phi_c(\hat{\mathbf{I}}) - \phi_c(\mathbf{I}_{\text{gt}}) \right\|_1 \quad (21)$$

其中, $\phi(\cdot)$ 表示 VGG-19 特征提取器。感知损失能够有效缓解过度平滑问题, 促进网络复原更丰富的纹理细节和语义一致的结构信息。

频域一致性损失从频谱分析角度约束复原图像的频率成分分布, 确保高频细节的准确复原。通过二维离散余弦变换 (2D-DCT) 将图像变换至频域, 计算频谱幅度的 L_1 差异:

$$\mathcal{L}_{\text{frequency}} = \frac{1}{C} \left\| \text{DCT}_2(\hat{\mathbf{I}}_c) - \text{DCT}_2(\mathbf{I}_{\text{gt},c}) \right\|_1 \quad (22)$$

其中, $\text{DCT}_2(\cdot)$ 表示二维离散余弦变换, $C=3$ 为 RGB 通道数。频域损失显式约束网络在频率空间的保真度, 对模糊和噪声退化尤为重要, 因为模糊主要破坏中高频信息, 而噪声在全频段引入虚假能量, 频域约束能够引导网络更精确地复原正确的频谱分布。

总体优化目标为三项损失的加权组合:

$$\mathcal{L}_{\text{total}} = \lambda_{\text{pixel}} \mathcal{L}_{\text{pixel}} + \lambda_{\text{perceptual}} \mathcal{L}_{\text{perceptual}} + \lambda_{\text{frequency}} \mathcal{L}_{\text{frequency}} \quad (23)$$

其中, 权重系数设置为 $\lambda_{\text{pixel}} = 1.0$, $\lambda_{\text{perceptual}} = 0.1$, $\lambda_{\text{frequency}} = 0.05$, 以平衡像素精度、感知质量和频域保真度。这些超参数通过在验证集上进行网格搜索确定, 确保三种损失项在训练过程中的贡献相对均衡。

3 实验结果与分析

3.1 实验数据集

为客观验证所提一体化遥感图像复原方法在复合退化条件下的有效性和鲁棒性, 本文选取了 MD-RSID、MD-RRSHID 以及 MDRS-Landsat 三个具有代表性的公开多退化基准数据集进行全面评估。这三个数据集均针对去雾、去噪、去模糊以及低光增强四类典型退化任务提供了配对的退化-清晰图像样本, 覆

盖了从合成到真实世界、从均匀退化到高度非均匀退化的多种场景,能够充分考察算法在不同成像条件、传感器特性以及退化强度下的泛化性能。

(1) MD-RSID 数据集源于 RSID^[38]去雾基准的扩展。该数据集以高质量军事应用遥感图像为基础,通过物理指导的退化注入策略系统性地引入噪声(零均值高斯噪声模拟传感器与传输干扰)、模糊(各向异性高斯核结合随机尺寸与方向模拟大气湍流及运动效应)以及低光(伽马变换模拟夜间或遮蔽条件)三种附加退化。最终形成约 4 000 对多退化-清晰图像对,训练、验证、测试划分比例为 8:1:1(3 200/400/400)。该数据集的显著优势在于退化参数分布高度可控,同时保留了原始图像丰富的纹理与细节,为评估算法在结构化退化场景下的复原精度提供了理想平台。

(2) MD-RRSHID 数据集基于 RRSID^[39]真实世界遥感雾霾基准扩展而成。RRSHID 包含 3 053 对真实采集的雾霾-清晰遥感图像对,覆盖中国多区域、多气候与多地表类型(城市、农田、山地、水体等),雾霾浓度从轻度到重度不等,伴随显著的真实颜色偏移与光谱漂移现象。需要特别说明的是,MD-RRSHID 数据集的去雾测试样本直接使用了这 3 053 对真实雾霾图像,这些雾霾退化完全来自真实大气条件而非任何人工合成。在其清晰真值图像上,本文采用与 MD-RSID 一致的物理一致性退化生成流程,进一步叠加噪声、模糊与低光退化,形成 3 053 对复合退化样本,训练/验证/测试划分为 2 442/305/306。该数据集的最大价值在于其完全真实的雾霾基础,能够有效检验算法面对分布外、非均匀大气干扰时的适应能力。

(3) MDRS-Landsat 数据集以 RSHaze^[40]提供的 5 500 张高分辨率遥感清晰图像(512 × 512)为核心,通过严谨的物理建模方式同步生成四类退化。其中去雾过程特别引入 Landsat 8 多光谱波段信息,实现波段相关的非均匀雾霾分布模拟,更贴近真实卫星成像链路。该数据集包含 5 130 对训练样本、100 对验证样本以及 270 对测试样本,为评估算法在多光谱保真度与物理一致性方面的表现提供了极具挑战性的基准。

3.2 评估指标

为全面、客观地衡量 Aether 在一体化遥感图像复原任务中的性能,本文选用峰值信噪比(Peak Signal-to-Noise Ratio, PSNR)、结构相似性(Structural Similarity Index Measure, SSIM)以及可学习感知图像块相似度(Learned Perceptual Image Patch Similarity, LPIPS)三种主流指标,从像素级精度、结构保真度以及人类感知质量三个维度对复原结果进行量化评估。

PSNR 是图像复原领域应用最广泛的客观指标,

主要反映像素级重建误差,其计算建立在均方误差(Mean Squared Error, MSE)基础上,定义为

$$\text{PSNR} = 10 \cdot \log_{10} \left(\frac{\text{MAX}^2}{\text{MSE}} \right) \quad (24)$$

其中,MAX 为图像像素的最大值,MSE 为复原图像与清晰参考图像之间的均方误差。PSNR 值越高,表明像素级失真越小,重建精度越高。

SSIM 从人类视觉系统(Human Visual System, HVS)的角度出发,同时考量亮度(luminance)、对比度(contrast)和结构(structure)三方面相似性,更贴近主观感知质量。其经典表达式为

$$\text{SSIM}(x, y) = [l(x, y)]^\alpha \cdot [c(x, y)]^\beta \cdot [s(x, y)]^\gamma \quad (25)$$

其中, $l(x, y)$ 、 $c(x, y)$ 、 $s(x, y)$ 分别表示亮度、对比度和结构比较函数,通常取 $\alpha = \beta = \gamma = 1$ 并加入小常数以保证数值稳定性。SSIM 的取值范围为 $[0, 1]$, 其值越接近 1, 表示两幅图像的结构信息越一致。

LPIPS 通过预训练深度网络(如 VGG-16)提取的多层感知特征来度量图像间的视觉差异,能够更好地模拟人类对图像质量的真实判断。其核心计算公式为

$$\text{LPIPS}(x, y) = \sum_l \frac{1}{H_l W_l} \left\| \mathbf{w}_l \odot (\hat{\phi}_l(x) - \hat{\phi}_l(y)) \right\|_2^2 \quad (26)$$

其中: $\hat{\phi}_l(\cdot)$ 表示第 l 层经单位归一化的激活特征图; H_l 、 W_l 为对应特征图的空间维度; $\mathbf{w}_l \in \mathbb{R}^{C_l}$ 为可学习的通道加权向量; \odot 表示逐通道乘法。LPIPS 值越小,说明两幅图像在人类感知层面越相似。

3.3 实验设置

本文的所有实验在配备 NVIDIA A800 GPU 的分布式计算集群上执行,训练框架采用 PyTorch Lightning, 基于分布式数据并行策略实现多 GPU 协同训练。

模型总计训练 150 个 epoch, 优化器为 AdamW ($\beta_1 = 0.9, \beta_2 = 0.999$), 初始学习率为 2×10^{-4} , 权重衰减系数为 0.01。学习率调度采用 15 个 epoch 线性 warm-up 后接余弦退火策略, 最终衰减至 1×10^{-6} 。每张 GPU 批大小为 4, 在 8 卡 A800 上全局批大小为 32。训练时输入图像随机裁剪至 128×128 像素, 测试与验证阶段采用原始分辨率。数据增强包括随机水平/垂直翻转以及 $\pm 15^\circ$ 随机旋转, 以增强模型对遥感场景多视角变化的鲁棒性。

网络骨干采用 Restormer 的四阶段 Transformer 编码-解码结构。初始嵌入维度为 48, 并在每个下采样阶段通道数翻倍(48 → 96 → 192 → 384 → 768), 编码器各阶段 Transformer 块数量分别为 4、6、6、8, 多头注意力头数相应配置为 1、2、4、8, 前馈网络扩展因子为

2.66,层归一化采用 WithBias 形式。

提出的 HADA 模块通过可学习的频率选择性分解与频率感知多头注意力机制,从成对退化-干净上下文的 CLIP 特征中提取多尺度退化感知表示,投影通道维度统一为 128。HONIF 模块则采用 8 头跨注意力(头维度 64),并引入基于上下文显著性的门控单元,实现融合强度的动态自适应调节。三级上下文注入模块分别作用于解码器第 1、2、3 层,处理特征维度依次为 64、128、320,与对应阶段解码特征通道数精确对齐,从而确保退化感知信息在整个上采样路径中的有效传递。CLIP 视觉编码器采用冻结的 ViT-B/32 模型,从倒数第二层提取上下文嵌入。

3.4 对比实验结果

为全面验证所提 Aether 的有效性,本文在三个公开基准数据集上进行系统评估,并与 Restormer、NAF-Net、PromptIR、AirNet、IDR、TransWeather、Ada4DIR、PhyDAE、MPMF-Net^[41]、MWFormer^[42]、AdaIR^[43]、OneRestore^[44]、GridFormer^[45]、DA-RCOT^[46]、AutoDIR、MoCE-

IR^[47]、DGUNet^[48]、SrResNet-AP^[49]、Restormer-AP、Uformer-AP 等主流方法进行对比。所有对比方法均使用官方权重或在完全相同训练条件下复现,以确保对比的公平性。

3.4.1 MD-RSID 数据集结果分析

表 1 展示了各方法在 MD-RSID 数据集上的定量结果。该数据集通过物理指导策略在高质量遥感图像上注入噪声、模糊和低光退化,形成约 4 000 对样本,退化参数高度可控。

Aether 在四类任务中均取得最优性能。去雾任务中达到 28.21 dB PSNR、0.966 8 SSIM 和 0.051 4 LPIPS,较次优方法 PhyDAE (26.86 dB、0.961 3、0.058 6) 提升 1.35 dB、0.005 5 和 12.3%。传统方法如 AirNet (19.79 dB) 和 MPMF-Net (19.12 dB) 与 Aether 差距超过 8 dB,表明简单端到端学习难以有效建模大气散射物理过程。去模糊任务中,Aether 的 29.86 dB PSNR 较 PhyDAE (27.73 dB) 提升 2.13 dB, LPIPS 改善 20.0%,显著优于 MWFormer (25.44 dB) 和 GridFormer (26.13 dB)。

表 1 MD-RSID 数据集各类退化情况的定量比较结果表

Table 1 Quantitative comparison results for various degradation types on the MD-RSID dataset

方法	来源	去雾			去模糊			去噪			低光增强		
		PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓
AirNet	22'CVPR	19.79	0.890 0	0.142 0	21.40	0.652 6	0.438 0	31.28	0.871 4	0.237 0	24.25	0.918 3	0.075 5
PromptIR	23'NeurIPS	20.00	0.899 2	0.121 2	22.39	0.651 0	0.437 4	<u>32.71</u>	<u>0.885 8</u>	0.225 6	24.04	0.917 5	0.070 4
MPMF-Net	25'AAAI	19.12	0.886 8	0.158 7	20.97	0.600 2	0.458 8	28.56	0.830 8	0.289 5	24.35	0.908 2	0.060 9
MWFormer	24'TIP	22.81	0.865 7	0.160 3	25.44	0.676 3	0.426 4	29.87	0.792 9	0.295 5	28.54	0.952 8	0.068 2
AdaIR	25'ICLR	22.21	0.898 9	0.143 1	26.02	0.730 9	0.386 1	30.76	0.837 3	0.271 9	28.01	0.957 2	0.055 3
Ada4DIR	25'IF	21.95	0.899 2	0.127 7	24.19	0.709 9	0.388 9	28.78	0.779 6	0.295 7	24.61	0.932 6	0.069 5
OneRestore	24'ECCV	25.54	0.906 1	0.124 9	25.89	0.689 6	0.416 2	30.48	0.819 5	0.280 2	29.17	0.969 9	0.044 1
GridFormer	24'IJCV	20.12	0.876 2	0.129 6	26.13	0.718 5	0.330 0	30.39	0.822 1	0.249 3	27.22	0.959 3	0.028 4
DA-RCOT	25'TPAMI	23.29	0.921 9	0.118 5	26.33	0.736 9	0.377 1	31.08	0.843 6	0.269 6	27.56	0.964 2	0.046 9
AutoDIR	24'ECCV	24.53	0.938 6	0.098 3	26.28	0.723 3	0.390 1	29.49	0.768 7	0.300 2	27.24	0.961 5	0.049 9
PhyDAE	25'arXiv	<u>26.86</u>	<u>0.961 3</u>	<u>0.058 6</u>	<u>27.73</u>	<u>0.777 2</u>	<u>0.322 1</u>	<u>32.77</u>	<u>0.886 2</u>	<u>0.213 6</u>	<u>31.96</u>	<u>0.985 5</u>	<u>0.021 1</u>
Aether	—	28.21	0.966 8	0.051 4	29.86	0.832 8	0.257 6	32.60	0.882 0	<u>0.216 3</u>	33.61	0.987 9	0.013 2

注:最优结果以加粗标注,次优结果以下划线标注。

低光增强任务展现最显著优势,Aether 取得 33.61 dB PSNR、0.987 9 SSIM 和 0.013 2 LPIPS,较 PhyDAE (31.96 dB、0.985 5、0.021 1) 提升 1.65 dB、0.002 4 和 37.4%。LPIPS 的大幅改善表明 HONIF 模块的空间自适应门控机制能够对重度欠曝区域施加强力校正,同时保护相对明亮区域,避免过度增强。去噪任务中,Aether (32.60 dB、0.882 0) 与 PhyDAE (32.77 dB、0.886 2) 性能接近,表明在相对简单的全频段噪声退化中,现有方法已具备较强处理能力。如图 5 所示,Aether 在四类退化任务中均能复原更清晰的纹理细节和更自然的色彩,特别是在低光增强任务中

有效抑制了噪声放大现象,同时保持了边缘锐度。

3.4.2 MD-RRSHID 数据集结果分析

表 2 展示了 MD-RRSHID 数据集结果。该数据集包含 3 053 对真实采集的雾霾-清晰图像对,覆盖多区域、多气候与多地表类型,伴随显著的颜色偏移与光谱漂移,为算法的分布外泛化能力提供严峻考验。

Aether 在真实退化场景下保持全面领先。去雾任务达到 24.13 dB PSNR 和 0.680 3 SSIM,较 PhyDAE (22.96 dB、0.651 1) 提升 1.17 dB 和 0.029 2。与 MD-RSID 相比,所有方法 PSNR 普遍下降 3~5 dB,但 Aether 下降幅度 (4.08 dB) 相对较小,验证了配对上下文退化解耦机制的鲁棒性。

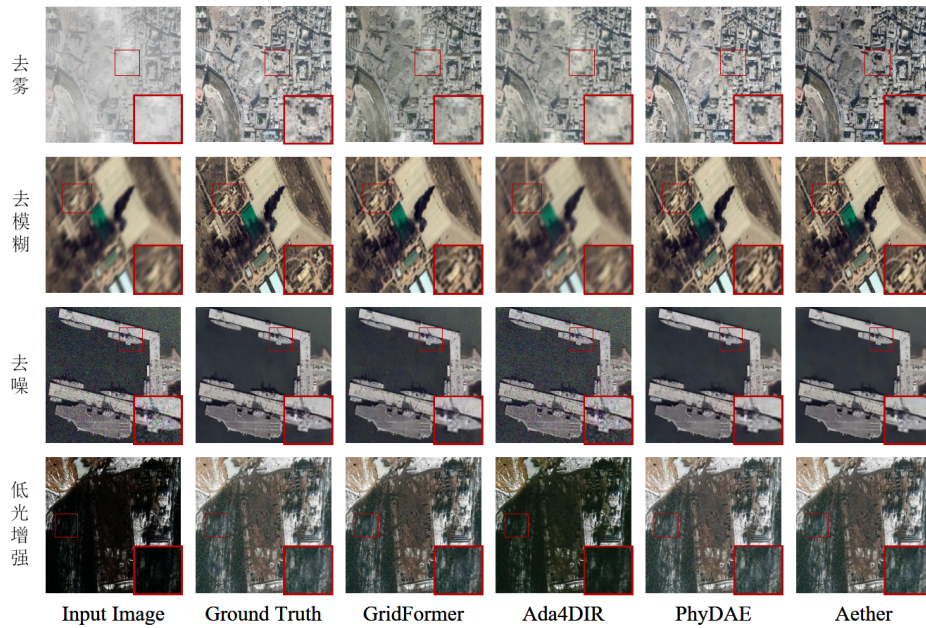


图5 MD-RSID数据集上不同一体化图像复原算法的可视化结果对比图

Figure 5 Visual comparison of different unified image restoration methods on the MD-RSID dataset

表 2 MD-RRSHID数据集各类退化情况的定量比较结果表

Table 2 Quantitative comparison results for various degradation types on the MD-RRSHID dataset

方法	来源	去雾			去模糊			去噪			低光增强		
		PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓
AirNet	22'CVPR	20.74	0.518 7	0.580 9	22.69	0.660 6	0.368 8	33.83	0.883 4	0.289 1	27.01	0.842 2	0.302 0
PromptIR	23'NeurIPS	19.72	0.504 9	0.558 5	23.83	0.704 9	0.364 0	<u>35.46</u>	<u>0.918 2</u>	0.260 7	36.52	0.969 9	0.084 2
MPMF-Net	25'AAAI	18.70	0.516 8	0.596 6	27.50	0.755 2	0.299 9	25.60	0.816 3	0.310 5	34.62	0.924 7	0.085 1
MoCE-IR	25'CVPR	22.86	0.647 9	0.468 5	33.12	0.879 5	0.250 2	35.04	0.908 9	0.239 6	35.31	0.979 7	0.059 1
MWFormer	24'TIP	18.63	0.509 3	0.592 0	25.29	0.711 3	0.452 7	25.80	0.534 5	0.492 6	28.16	0.864 4	0.173 3
AdaIR	25'ICLR	22.58	0.636 9	0.469 0	32.51	0.867 6	0.259 9	34.80	0.901 6	0.248 5	34.24	0.972 9	0.063 6
Ada4DIR	25'IF	19.97	0.514 8	0.555 8	27.76	0.767 9	0.307 1	25.22	0.536 1	0.425 8	24.19	0.820 0	0.278 4
OneRestore	24'ECCV	20.04	0.587 0	0.574 4	23.62	0.721 9	0.369 3	31.40	0.825 1	0.301 5	28.26	0.643 5	0.204 9
GridFormer	24'IJCV	21.46	0.575 7	0.458 4	31.94	0.846 5	0.224 8	33.28	0.850 9	0.254 3	33.69	0.966 5	0.048 3
DA-RCOT	25'TPAMI	22.80	0.643 9	<u>0.456 9</u>	<u>33.86</u>	<u>0.893 8</u>	0.227 5	33.09	0.897 0	0.284 5	35.91	0.978 5	0.061 2
PhyDAE	25'arXiv	<u>22.96</u>	<u>0.651 1</u>	0.442 3	33.73	0.878 7	<u>0.212 2</u>	35.17	0.914 1	<u>0.228 9</u>	<u>37.35</u>	<u>0.985 2</u>	<u>0.044 2</u>
Aether	—	24.13	0.680 3	0.459 3	36.26	0.925 7	0.175 2	35.68	0.920 4	0.228 8	45.46	0.990 9	0.035 6

注:最优结果以加粗标注,次优结果以下划线标注。

去模糊任务中,36.26 dB PSNR 较 DA-RCOT(33.86 dB)提升 2.40 dB,LPIPS改善 23.0%(0.175 2 vs 0.227 5)。

最显著的是低光增强任务,Aether 达到 45.46 dB PSNR、0.990 9 SSIM 和 0.035 6 LPIPS,较 PhyDAE(37.35 dB、0.985 2、0.044 2)提升 8.11 dB、0.005 7 和 19.5%。这是所有实验中最显著的性能优势,充分验证了 HADA 模块精确解析真实雾霾与低光退化频域差异的能力。GridFormer (33.69 dB) 和 AdaIR (34.24 dB) 与 Aether 差距超过 11 dB,表明传统方法在真实雾霾背景下的低光退化处理存在严重瓶颈。去噪任务中,

Aether(35.68 dB、0.920 4)超越 PromptIR (35.46 dB、0.918 2),LPIPS 改善 12.2%,说明高阶非线性交互能力可有效处理真实与合成退化的混合模式。如图 6 所示,面对真实雾霾的非均匀分布和颜色偏移,Aether 能够更有效地复原场景对比度和色彩保真度,在低光增强任务中避免了其他方法出现的过饱和和色彩失真现象。

3.4.3 MDRS-Landsat数据集结果分析

表 3 展示了 MDRS-Landsat 数据集结果。该数据集基于 5 500 张高分辨率图像,通过物理建模生成退

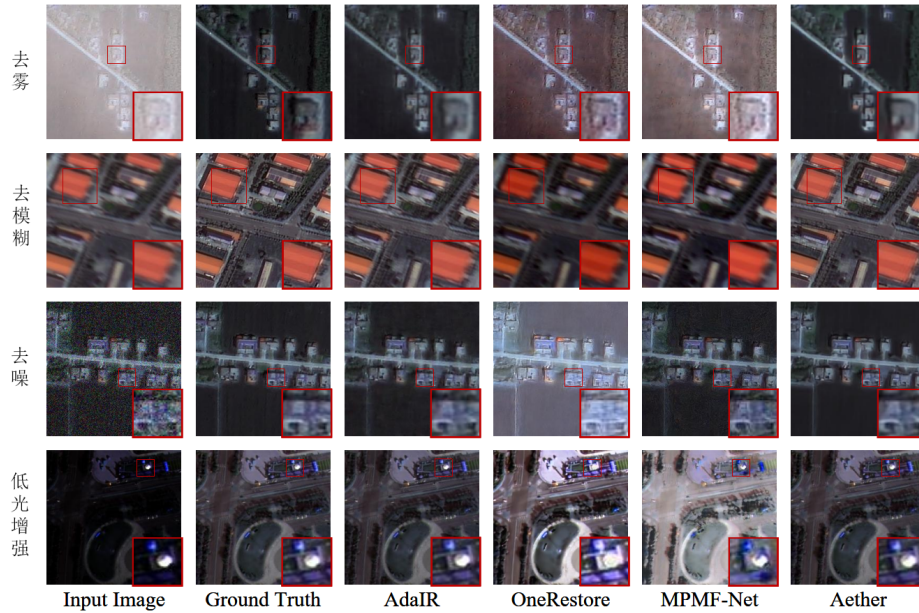


图6 MD-RRSHID数据集上不同一体化图像复原算法的可视化结果对比图

Figure 6 Visual comparison of different unified image restoration methods on the MD-RRSHID dataset

表3 MDRS-Landsat数据集各类退化情况的定量比较结果表

Table 3 Quantitative comparison results for various degradation types on the MDRS-Landsat dataset

方法	来源	去雾			去模糊			去噪			低光增强		
		PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓
NAFNet	22'ECCV	31.56	0.964 2	0.041 7	33.10	0.812 0	0.319 4	33.08	0.826 3	0.187 2	30.40	0.951 6	0.054 2
Restormer	22'CVPR	36.18	0.986 7	0.012 4	35.23	0.855 9	0.209 9	34.53	0.858 9	0.135 6	37.86	0.987 2	0.011 0
DGUNet	22'CVPR	27.45	0.933 8	0.071 3	29.64	0.782 2	0.340 5	30.31	0.731 4	0.249 1	27.15	0.901 0	0.133 9
TransWeather	22'CVPR	35.02	0.968 9	0.024 1	33.45	0.815 9	0.286 8	33.69	0.842 8	0.153 0	36.33	0.970 5	0.019 3
AirNet	22'CVPR	24.39	0.933 1	0.064 1	28.27	0.788 7	0.324 4	30.30	0.744 6	0.191 8	28.38	0.947 2	0.056 9
PromptIR	23'NeurIPS	37.61	0.989 7	0.008 4	36.41	0.886 1	0.155 7	<u>34.99</u>	<u>0.872 9</u>	<u>0.102 9</u>	39.09	0.990 0	0.008 4
IDR	23'CVPR	36.99	0.989 2	0.008 7	36.57	<u>0.890 2</u>	<u>0.149 8</u>	34.88	0.868 1	0.109 1	35.19	0.986 5	0.009 6
SrResNet-AP	24'CVPR	34.78	0.982 5	0.018 6	34.63	0.847 9	0.227 3	34.70	0.862 0	0.137 5	33.87	0.982 3	0.016 5
Restormer-AP	24'CVPR	37.36	0.988 8	0.009 8	35.75	0.873 2	0.180 0	34.96	0.869 7	0.123 9	37.27	0.988 5	0.010 2
Uformer-AP	24'CVPR	36.06	0.987 7	0.010 7	34.64	0.848 8	0.234 2	34.39	0.853 3	0.133 4	36.58	0.989 9	0.006 9
Ada4DIR	25'IF	38.79	0.991 1	0.006 5	36.40	0.887 4	0.156 7	34.97	0.868 8	0.124 1	40.69	0.992 3	<u>0.005 0</u>
PhyDAE	25'arXiv	<u>39.12</u>	<u>0.992 8</u>	0.022 7	36.88	0.882 4	0.148 7	34.53	0.865 1	0.099 1	42.24	0.994 9	0.012 1
Aether	—	42.75	0.994 2	<u>0.007 1</u>	37.55	0.907 8	0.187 1	35.15	0.874 5	0.172 0	43.84	0.996 3	0.003 0

注:最优结果以加粗标注,次优结果以下划线标注。

化,特别引入 Landsat 8 多光谱波段信息实现波段相关的非均匀雾霾模拟,贴近真实卫星成像链路。

Aether刷新所有任务最优纪录。去雾任务达到42.75 dB PSNR和0.994 2 SSIM,较PhyDAE(39.12 dB、0.992 8)提升3.63 dB和0.001 4, LPIPS为0.007 1,较PhyDAE(0.022 7)改善68.7%。所有方法PSNR普遍显著高于前两个数据集,反映了严格物理建模的退化更加规则可预测。去模糊任务中,37.55 dB PSNR超

越PhyDAE(36.88 dB)0.67 dB,但LPIPS(0.187 1)略高于IDR(0.149 8),表明在高分辨率场景下存在像素精度与感知质量的权衡。

低光增强任务最为显著,43.84 dB PSNR、0.996 3 SSIM和0.003 0 LPIPS,较PhyDAE(42.24 dB、0.994 9、0.012 1)提升1.60 dB、0.001 4和75.2%。0.003 0的LPIPS值接近感知质量理论下限,表明复原图像在人类视觉层面几乎无法与真值区分。去噪任务中,

35.15 dB PSNR 超越 PromptIR (34.99 dB) 0.16 dB, 在竞争激烈的任务中保持领先。如图 7 所示, 在大尺度图像复原中, Aether 展现出卓越的细节保持能

力和多光谱一致性, 特别是在去雾任务中准确复原了不同波段的光谱信息, 避免了色彩漂移和伪影。

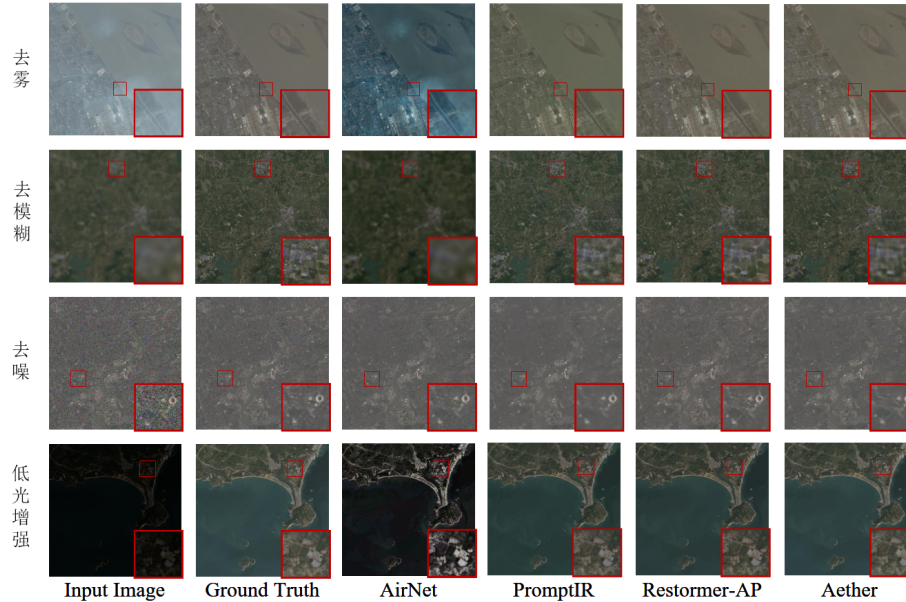


图 7 MDRS-Landsat 数据集上不同一体化图像复原算法的可视化结果对比图

Figure 7 Visual comparison of different unified image restoration methods on the MDRS-Landsat dataset

3.4.4 计算效率分析

表 4 展示了 Aether 与主流一体化图像复原方法的效率对比结果。从计算复杂度角度, Aether 的 FLOPs 为 441.98 G, 显著低于高性能方法如 DGUNet (1 050.57 G)、IDR (887.12 G) 和 TransWeather (782.45 G), 分别降低了 57.9%、50.2% 和 43.5% 的浮点运算量。虽然 Aether 的 FLOPs 比最轻量的 Ada4DIR (366.67 G) 高 20.5%, 但性能提升极为显著, 在 MD-RSID、MD-RRSHID 和 MDRS-Landsat 数据集上分别领先 6.19 dB、11.09 dB 和 2.11 dB。推理延迟方面, Aether 达到 421.46 ms 的次优表现, 相比 Ada4DIR (380.79 ms) 仅慢 10.7%, 但后者的显存占用高达 5.02 GB, 而 Aether 仅需 2.74 GB, 降低了 45.4%。更为关键的是, Aether 相比性能接近的高计算量方法展现出显著的速度优势: 相比 IDR 快 47.6%, 相比 TransWeather 快 41.3%, 相比 DGUNet 快 35.4%。显存占用方面, Aether 的 2.74 GB 在所有方法中排名第二, 仅比 PromptIR (2.38 GB) 高 15.1%, 但性能在三个数据集上平均领先约 6.28 dB、6.50 dB 和 2.79 dB。

综合三项效率指标与复原性能, Aether 在精度-效率帕累托前沿上占据最优位置。虽然 HONIF 模块引入了 B-spline 样条函数计算, 但通过选择性应用策略、紧凑的维度设计以及高效的向量化实现, 额外计算开销得到有效控制。实验结果表明, 相比计算量相近的 PromptIR (632.56 G) 和 AirNet (518.24 G), Aether 在保

表 4 不同方法在 512×512 分辨率遥感图像上的计算效率对比

Table 4 Comparison of computational efficiency of different methods on 512×512 resolution remote sensing images

方法	FLOPs/G	推理延迟/ms	显存占用/GB
AirNet	518.24	576.31	3.62
Ada4DIR	366.67	380.79	5.02
DGUNet	1 050.57	652.84	3.85
PromptIR	632.56	1 010.96	2.38
Transweather	782.45	718.63	4.51
IDR	887.12	805.28	4.76
Aether	441.98	421.46	2.74

持更低 FLOPs 的同时分别实现了 58.3% 和 26.9% 的延迟降低, 且性能大幅领先。这一优异的性能-效率权衡充分验证了 HADA 和 HONIF 模块设计的有效性, 使 Aether 不仅适用于高性能计算平台, 也能在显存受限的边缘设备或嵌入式遥感处理系统中高效部署, 展现出良好的实际应用价值。

3.5 消融实验结果

3.5.1 核心组件消融

为了系统性验证 Aether 提出的两大核心组件——HADA 和 HONIF 的有效性, 本文在 MD-RSID、MD-RRSHID 和 MDRS-Landsat 三个数据集上进行了渐进式消融实验。为全面评估各组件对模型整体性能的

影响,本文采用四类退化任务的平均指标(4D-AVG,即去雾、去模糊、去噪和低光增强的平均PSNR、SSIM和LPIPS)进行量化分析。基线模型(Baseline)构建于Restormer骨干网络之上,在移除HADA和HONIF模块的同时,仅采用标准的交叉注意力机制将CLIP视觉编码器提取的退化先验特征注入解码器,以实现基础的退化引导复原。

如表5的消融实验结果所示,相比于基线模型,单独引入HADA模块后,模型在三个数据集上的PSNR分别显著提升了1.50 dB(MD-RSID)、1.71 dB(MD-RRSHID)和2.24 dB(MDRS-Landsat)。这一显著增益表明,遥感图像中的退化模式在频域具有强烈的区分度。HADA通过数据驱动的自适应频谱分解,能够精准捕获这些频域指纹特征,从而为后续的复原过程提供了关键的退化先验信息,证明了显式频域建模在一体化复原任务中的必要性。

表5 Aether核心组件消融实验结果

Table 5 Ablation study results of the core components of Aether

方法	组件配置		性能表现(4D-AVG, PSNR ↑)		
	HADA	HONIF	MD-RSID	MD-RRSHID	MDRS-Landsat
Baseline	×	×	28.65	32.71	36.17
+ HADA	√	×	30.15	34.42	38.41
+ HONIF	×	√	29.43	33.89	37.76
Aether	√	√	31.07	35.38	39.82

注:4D-AVG表示四类退化任务(去雾、去模糊、去噪、低光增强)的平均指标,最优结果以加粗标注。

此外,在基线模型基础上单独引入HONIF模块,PSNR分别提升了0.78 dB、1.18 dB和1.59 dB。这证实了传统的线性融合或简单的注意力机制难以充分处理退化特征与图像内容之间复杂的耦合关系。HONIF利用基于样条函数的高阶非线性映射,成功建立了退化先验与图像语义特征之间的深度对齐,有效提升了特征融合的表达力。

当同时集成HADA和HONIF构成完整的Aether模型时,性能达到了最佳水平,在MD-RSID、MD-RRSHID和MDRS-Landsat上分别取得了31.07 dB、35.38 dB和39.82 dB的最优成绩。值得注意的是,完整模型的性能提升并非两个组件增益的简单线性叠加,而是展现出显著的协同增强效应。这表明HADA提供的精准退化表征为HONIF的高阶交互提供了高质量的输入基础,而HONIF则将这些先验信息有效地转化为对图像特征的空间自适应调制。两者互为补充,共同确立了Aether在处理复杂多变遥感退化场景下的鲁棒性与优越性。

3.5.2 HADA内部机制消融

为深入探究HADA中核心组件——LFD的有效性,并验证其相对于传统固定基底频率变换的优越性,本文在真实场景MD-RRSHID数据集上进行了对比消融实验。实验设置包括:(1)无频率分解,直接利用CLIP特征进行后续处理;(2)固定DCT分解,采用标准离散余弦变换划分频带;(3)固定小波分解,采用Haar小波变换划分频带;(4)LFD,本文提出的数据驱动谐波滤波器组。详细量化结果如表6所示。

表6 MD-RRSHID数据集上HADA模块不同频率分解策略的消融实验结果

Table 6 Ablation study results of different frequency decomposition strategies in the HADA module on the MD-RRSHID dataset

频率分解策略	机制描述	性能表现(4D-AVG)		
		PSNR ↑	SSIM ↑	LPIPS ↓
无频率分解	直接使用CLIP特征	34.32	0.911 7	0.305 2
固定DCT分解	使用标准离散余弦变换分频	34.89	0.914 8	0.249 1
固定小波分解	使用Haar小波变换	34.85	0.915 2	0.248 7
LFD	数据驱动的谐波滤波器组	35.38	0.917 4	0.224 7

注:最优结果以加粗标注。

相比于“无频率分解”的基准配置,引入任何形式的显式频率分解(DCT或小波)均带来了性能提升。这有力地佐证了前文所述的物理观察,即遥感图像中的不同退化类型在频域具有强烈的依赖性和区分度,显式的频域建模是提升复原上限的关键。此外,传统的固定分解策略虽然优于无分解策略,但二者性能接近且遭遇瓶颈。这是因为预定义的正交基函数(如余弦波或方波)形态固定,难以灵活适应真实遥感场景中高度非平稳、非均匀的混合退化频谱特征。相比之下,本文提出的LFD策略取得了决定性的性能优势,PSNR达到35.38 dB,SSIM提升至0.917 4,且感知指标LPIPS显著降低至0.224 7。具体而言,LFD较次优的DCT策略在PSNR上提升了0.49 dB,在LPIPS上改善了约9.8%。

LFD的优越性可归因于其“数据驱动”的本质。通过端到端的学习,LFD的谐波滤波器组能够自动发现并锁定对当前退化类型最具判别力的频段,而非被动地接受固定的频率划分。这种自适应机制使得HADA模块能够更精准地将退化特征与图像内容分离,从而为后续的复原网络提供更纯净、更具指导意义的先验信息。

此外,为了直观验证所提HADA模块在特征流空间中退化模式的判别能力,本文利用t-SNE算法对不同处理阶段的特征分布进行了降维可视化,结果

如图 8 所示。从图 8(a) 的 CLIP 输出特征和图 8(b) 的 HADA 输入特征中可以观察到, 四种不同类型的退化(雾霾、模糊、噪声、低光)在特征空间中缺乏清晰的判别边界。这表明, 虽然 CLIP 提取的视觉特征富含通用语义信息, 但在未经特定适配的情况下, 难以有效区分和提取遥感领域复杂的异质退化模式, 导致不同退化类型的样本点相互交织, 缺乏可分性。

相比之下, 如图 8(c) 所示, 经过 HADA 模块的频

域自适应解析处理后, 特征分布发生了质的飞跃。四类退化特征被成功映射到了界限清晰、类内紧凑的独立流形簇中, 实现了从模糊难分到高度可判别的转变。这种显著的聚类效应有力地证明了 HADA 模块能够精准地捕捉不同退化类型独特的频谱指纹, 成功将隐式的退化信息从通用特征中剥离并转化为高判别性的显式表征, 从而为后续网络针对特定退化执行差异化的复原策略提供了可靠的先验基础。

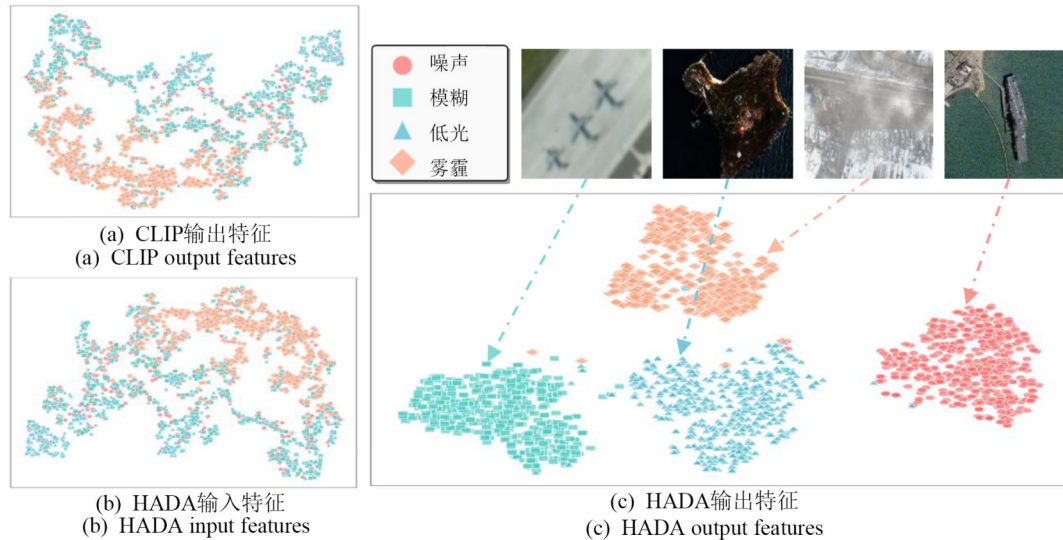


图 8 退化特征在不同处理阶段的 t-SNE 可视化结果

Figure 8 t-SNE visualization of degradation features at different processing stages

3.5.3 HONIF 内部机制消融

为深入验证 HONIF 在注入退化先验与特征调制方面的有效性, 本研究在 MDRS-Landsat 数据集上对通道拼接卷积、标准缩放点积注意力、交叉注意力后接 MLP 以及本文提出的 HONIF 四种递进式特征交互策略进行了系统性对比分析。如表 7 所示, 实验结果有力地揭示了特征融合深度与非线性程度对复原性能的决定性影响。仅采用浅层的“通道拼接后卷积”策略时, 模型性能受限明显 (PSNR 仅为 36.32 dB, LPIPS 高达 0.182 7), 表明简单的线性局部融合难以有效捕捉退化先验与解码器特征间复杂的语义关联。相比之下, 引入“标准缩放点积注意力”通过建立长期依赖关系将 PSNR 显著提升至 38.89 dB, 而进一步增加 MLP 引入非线性变换后, PSNR 继续攀升至 39.34 dB, 验证了非线性映射在拟合复杂退化特征中的价值。

然而, 上述方法仍受限于标准注意力机制中固有的线性投影瓶颈, 而本文提出的 HONIF 策略通过创新性地引入基于样条函数的可学习高阶变换替代传统线性投影, 突破了这一限制, 取得了最优的性能表现 (PSNR 达到 39.82 dB, SSIM 提升至 0.943 2, LPIPS 降

表 7 MDRS-Landsat 数据集上 HONIF 模块不同交互融合策略的消融实验结果

Table 7 Ablation study results of different interaction fusion strategies in the HONIF module on the MDRS-Landsat dataset

交互融合方式	性能表现 (4D-AVG)		
	PSNR \uparrow	SSIM \uparrow	LPIPS \downarrow
通道拼接后卷积	36.32	0.905 5	0.182 7
标准缩放点积注意力	38.89	0.928 9	0.149 1
交叉注意力后接 MLP	39.34	0.931 8	0.125 4
HONIF	39.82	0.943 2	0.092 3

注: 最优结果以加粗标注。

至 0.092 3)。这一近 1 dB 的 PSNR 提升与 38.1% 的感知指标优化充分证明, HONIF 强大的函数逼近能力能够实现退化先验对空间特征的自适应精准调制, 在确保高频细节复原的同时维持了极高的感知自然度。

此外, 为了直观揭示 HONIF 模块在特征空间中的调制机制与作用效能, 本文对高阶非线性交互前后的空间特征激活图进行了可视化分析, 结果如图 9 所示。观察 HONIF 输入的预融合特征图可以看出, 其空间响应分布呈现出显著的离散性与混沌状态, 高激活区域(暖色调)缺乏明确的语义指向, 表明仅经由

线性路径注入的退化先验难以有效区分场景固有结构与退化干扰。相比之下,经 HONIF 模块处理后的增强特征图发生了质的转变,呈现出高度相干且结构化的激活模式。高阶非线性交互机制成功抑制了非结构化的背景噪声,同时显著增强了纹理丰富区域及

边缘轮廓的特征响应强度。这种从无序到有序的特征演变有力地证实了 HONIF 模块中基于 KAN 的高阶函数逼近能力,其能够实现退化先验与图像内容在深层语义上的精准对齐与空间自适应重校准,从而为后续的高保真复原提供了具备高判别力的特征基础。

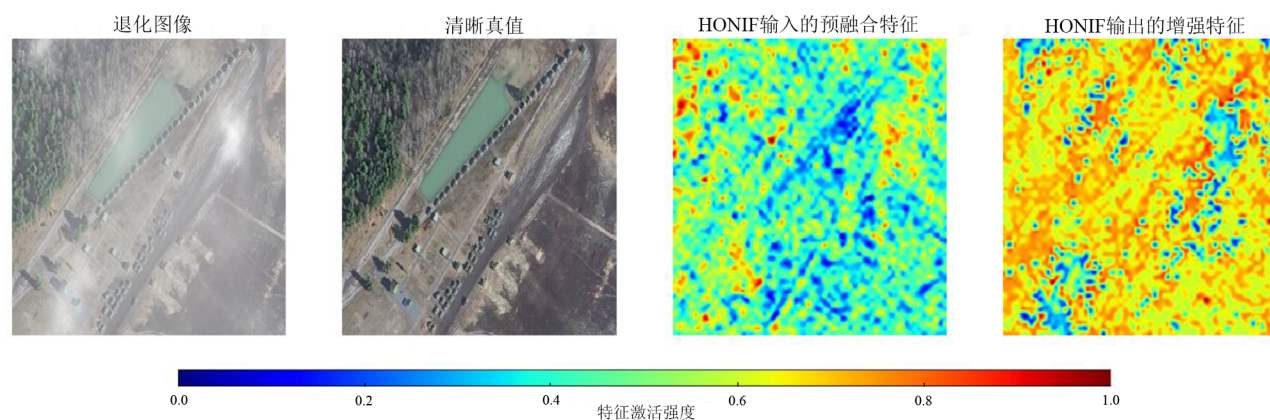


图9 HONIF模块处理前后的特征激活图可视化对比

Figure 9 Visual comparison of feature activation maps before and after processing by the HONIF module

3.5.4 损失函数消融

为了验证 Aether 提出的多维度联合优化策略在提升复原精度与感知质量方面的有效性,本研究对像素域保真度损失、感知质量损失以及频域一致性损失进行了逐步叠加的消融实验。实验结果如表 8 所示。

表8 损失函数消融实验结果

Table 8 Ablation study results of different loss functions

损失函数配置			性能指标 (4D-AVG)		
$\mathcal{L}_{\text{pixel}}$	$\mathcal{L}_{\text{perceptual}}$	$\mathcal{L}_{\text{frequency}}$	PSNR \uparrow	SSIM \uparrow	LPIPS \downarrow
✓	×	×	34.81	0.886 4	0.172 7
✓	✓	×	35.29	0.895 2	0.166 7
✓	✓	✓	35.43	0.913 3	0.150 6

注:最优结果以加粗标注。

仅采用像素域保真度损失作为基准时,模型虽然能确保基础的信号保真度,获得 34.81 dB 的 PSNR,但在结构相似性(SSIM 为 0.886 4)和感知指标(LPIPS 为 0.172 7)上表现一般。这印证了单一像素级约束容易导致复原结果产生过度平滑现象,难以从统计意义上复原符合人类视觉认知的精细纹理。

在引入基于 VGG-19 深层特征的感知质量损失后,模型性能获得明显改善,PSNR 提升至 35.29 dB,同时 LPIPS 降低至 0.166 7。这一增益表明,在语义特征空间施加约束能够有效引导网络关注图像的高层结构信息,缓解了像素级损失带来的模糊效应。

进一步引入频域一致性损失后,模型达到了最优性能配置,PSNR 攀升至 35.43 dB,SSIM 显著跃升至

0.913 3,且 LPIPS 降至最低的 0.150 6。上述结果说明,由于退化在频域具有特定的能量分布特征,显式的频域一致性约束强制网络在复原空间细节的同时校正频谱幅度的分布偏差。这不仅强化了边缘和纹理的高频重建,还确保了全频段能量分布的物理真实性,从而实现了像素精度、感知质量与物理一致性的最佳平衡。

4 结束语

本文针对现有一体化复原框架在物理一致性建模与特征交互深度上的局限,提出了一种结合频域物理感知与高阶语义融合的一体化遥感图像复原框架 Aether。通过引入可学习频率分解机制,HADA 模块成功克服了传统固定频率变换在处理非平稳遥感退化时的适应性瓶颈,实现了对异质退化频谱特征的精准捕获。同时,HONIF 模块创新性地将样条函数逼近理论引入注意力机制,构建了退化先验与图像内容间的高阶非线性交互通道,显著增强了模型在复杂场景下的特征调制能力。在多个高分辨率与真实场景数据集上的广泛实验证实,Aether 不仅在 PSNR 与 SSIM 等客观指标上全面超越现有最先进方法,更在光谱保真度与纹理细节复原等主观感知层面展现出显著优势。该研究验证了显式频域物理建模与高阶神经网络结合的有效性,为未来构建更具物理解释性和泛化能力的智能遥感处理系统提供了新的理论视角与技术支撑。

参考文献

- [1] Wang Peijuan, Bayram B, Sertel E. A comprehensive review on deep learning based remote sensing image super-resolution methods[J]. *Earth-Science Reviews*, 2022, 232: 104110.
- [2] Liu Baodi, Zhao Lifei, Shao Shuai, et al. RAN: Region-aware network for remote sensing image super-resolution[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2023, 61: 5408113.
- [3] Lei Sen, Shi Zhenwei, Mo Wenjie. Transformer-based multistage enhancement for remote sensing image super-resolution[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 5615611.
- [4] Dong Runmin, Mou Lichao, Zhang Lexing, et al. Real-world remote sensing image super-resolution via a practical degradation model and a kernel-aware network[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2022, 191: 155-170.
- [5] Chen Hao, Yokoya N, Chini M. Fourier domain structural relationship analysis for unsupervised multimodal change detection[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2023, 198: 99-114.
- [6] Roy S K, Deria A, Hong Danfeng, et al. Multimodal fusion transformer for remote sensing image classification[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2023, 61: 5515620.
- [7] Kar A, Dhara S K, Sen D, et al. Zero-shot single image restoration through controlled perturbation of koschmieder's model[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2021: 16200-16210.
- [8] Zamir S W, Arora A, Khan S, et al. Restormer: Efficient transformer for high-resolution image restoration[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2022: 5718-5729.
- [9] Chen Liangyu, Chu Xiaojian, Zhang Xiangyu, et al. Simple baselines for image restoration[C]//Computer Vision-ECCV 2022. Cham: Springer, 2022: 17-33.
- [10] Li Yawei, Fan Yuchen, Xiang Xiaoyu, et al. Efficient and explicit modelling of image hierarchies for image restoration[C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2023: 18278-18289.
- [11] Khan S, Potlapalli V, Khan F S, et al. PromptIR: Prompting for all-in-one image restoration[C]//Advances in Neural Information Processing Systems 36. Neural Information Processing Systems Foundation, Inc. (NeurIPS), 2023: 71275-71293.
- [12] Li Boyun, Liu Xiao, Hu Peng, et al. All-In-one image restoration for unknown corruption[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2022: 17431-17441.
- [13] Zhang Jianhai, Huang Jian, Yao Mingda, et al. Ingredient-oriented multi-degradation learning for image restoration[C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2023: 5825-5835.
- [14] Zhao Chen, Cai Wenli, Dong Chengyu, et al. Wavelet-based Fourier information interaction with frequency diffusion adjustment for underwater image restoration[C]//2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2024: 8281-8291.
- [15] Cui Yuning, Ren Wenqi, Cao Xiaochun, et al. Image restoration via frequency selection[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2024, 46(2): 1093-1108.
- [16] Cui Yuning, Tao Yi, Bing Zhenshan, et al. Selective frequency network for image restoration[C]//The 11th International Conference on Learning Representations. ICLR, 2023. <https://openreview.net/forum?id=tyZ1ChG-ZIKO>.
- [17] Tu Zhengzhong, Talebi H, Zhang Han, et al. MAXIM: Multi-axis MLP for image processing[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2022: 5759-5770.
- [18] Wang Zhendong, Cun Xiaodong, Bao Jianmin, et al. Uformer: A general U-shaped transformer for image restoration[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2022: 17662-17672.
- [19] Liang Jingyun, Cao Jiezhong, Sun Guolin, et al. SwinIR: Image restoration using swin transformer[C]//2021 IEEE/CVF International Conference on Computer Vision Workshops. Piscataway: IEEE, 2021: 1833-1844.
- [20] Song Tianyu, Fan Suming, Jin Jianyu, et al. Exploring an efficient frequency-guidance transformer for single image deraining[J]. *Signal, Image and Video Processing*, 2024, 18(3): 2429-2438.
- [21] Fuoli D, Van Gool L, Timofte R. Fourier space losses for efficient perceptual image super-resolution[C]//2021 IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE, 2021: 2340-2349.

- [22] Huang Jing, Liu Yanjun, Zhao Feng, et al. Deep Fourier-based exposure correction network with spatial-frequency interaction[C]//Computer Vision-ECCV 2022. Cham: Springer, 2022: 163-180.
- [23] Valanarasu J M J, Yasarla R, Patel V M. TransWeather: Transformer-based restoration of images degraded by adverse weather conditions[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2022: 2343-2353.
- [24] Guo Hang, Li Jianmin, Dai Tao, et al. MambaIR: A simple baseline for image restoration with state-space model[C]//Computer Vision - ECCV 2024. Cham: Springer, 2025: 222-241.
- [25] Luo Ziwei, Gustafsson F K, Zhao Zheng, et al. Controlling vision-language models for multi-task image restoration[PP/OL]. V2. arXiv (2024-02-28) [2025-12-29]. <https://doi.org/10.48550/arXiv.2310.01018>.
- [26] Zhang Xin, Ma Jianqi, Wang Guoli, et al. Perceive-IR: Learning to perceive degradation better for all-in-one image restoration[J]. IEEE Transactions on Image Processing, 2026, 35: 2018-2033.
- [27] Cao Shen, Liu Yuhuan, Zhang Weilong, et al. GRIDS: Grouped multiple-degradation restoration with image degradation similarity[C]//Computer Vision - ECCV 2024. Cham: Springer, 2025: 70-87.
- [28] Wu Gang, Jiang Junjun, Wang Yujie, et al. Debaised all-in-one image restoration with task uncertainty regularization[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2025, 39(8): 8386-8394.
- [29] Xia Bin, Zhang Yulun, Wang Shiyin, et al. DiffIR: Efficient diffusion model for image restoration[C]//2023 IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE, 2023: 13049-13059.
- [30] Jiang Yuntian, Zhang Zeyu, Xue Tianfan, et al. AutoDIR: Automatic all-in-one image restoration with latent diffusion[C]//Computer Vision - ECCV 2024. Cham: Springer, 2025: 340-359.
- [31] Chen Haoyu, Li Wenbo, Gu Jinjin, et al. RestoreAgent: Autonomous image restoration agent via multimodal large language models[C]//Advances in Neural Information Processing Systems 37. Neural Information Processing Systems Foundation, Inc., 2024: 110643-110666.
- [32] Jiang Junjun, Zuo Zhenyu, Wu Gang, et al. A survey on all-in-one image restoration: Taxonomy, evaluation and future trends[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2025, 47(12): 11892-11911.
- [33] Feng Yujie, Yang Yu, Fan Xiaohui, et al. A progressive image restoration network for high-order degradation imaging in remote sensing[J]. IEEE Transactions on Geoscience and Remote Sensing, 2025, 63: 5633816.
- [34] Li Hezhe, Yuan Qiangqiang, He Jiang, et al. Ada4DIR: An adaptive model-driven all-in-one image restoration network for remote sensing images[J]. Information Fusion, 2025, 118: 102930.
- [35] Dong Zhe, Sun Yuzhe, Jiang Haochen, et al. PhyDAE: Physics-guided degradation-adaptive experts for all-in-one remote sensing image restoration[PP/OL]. V1. arXiv (2025-10-09) [2025-12-29]. <https://doi.org/10.48550/arXiv.2510.08653>.
- [36] Wang Shuo, Zhang Libo. A unified framework for double-degradation remote sensing image restoration through saliency-guided interaction learning[J]. IEEE Transactions on Geoscience and Remote Sensing, 2024, 62: 5607719.
- [37] Dian Renwei, Li Shutao. Hyperspectral image super-resolution via subspace-based low tensor multi-rank regularization[J]. IEEE Transactions on Image Processing, 2019, 28(10): 5135-5146.
- [38] Chi Kancheng, Yuan Yuan, Wang Qi. Trinity-net: Gradient-guided swin transformer-based remote sensing image dehazing and beyond[J]. IEEE Transactions on Geoscience and Remote Sensing, 2023, 61: 4702914.
- [39] Zhu Zhihui, Lu Wei, Chen Shibin, et al. Real-world remote sensing image dehazing: Benchmark and baseline [J]. IEEE Transactions on Geoscience and Remote Sensing, 2025, 63: 4705014.
- [40] Song Yudong, He Zhiqiang, Qian Hao, et al. Vision transformers for single image dehazing[J]. IEEE Transactions on Image Processing, 2023, 32: 1927-1941.
- [41] Wen Yubo, Gao Tian, Zhang Jian, et al. Multi-axis prompt and multi-dimension fusion network for all-in-one weather-degraded image restoration[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2025, 39(8): 8323-8331.
- [42] Zhu Ruixing, Tu Zhengzhong, Liu Jianmin, et al. MWFormer: Multi-weather image restoration using degradation-aware transformers[J]. IEEE Transactions on Image Processing, 2024, 33: 6790-6805.
- [43] Cui Yuning, Zamir S W, Khan S, et al. AdaIR: Adaptive all-in-one image restoration via frequency mining and modulation[C]//13th International Conference on Learning Representations. ICLR, 2025: 57335-57356.
- [44] Guo Yan, Gao Yu, Lu Yuxin, et al. OneRestore: A universal restoration framework for composite degradation[C]//

Computer Vision - ECCV 2024. Cham: Springer Nature Switzerland, 2024: 255-272.

- [45] Wang Ting, Zhang Kehai, Shao Zhiqiang, et al. GridFormer: Residual dense transformer with grid structure for image restoration in adverse weather conditions[J]. International Journal of Computer Vision, 2024, 132(10): 4541-4563.
- [46] Tang Xiaolin, Gu Xin, He Xiaoyun, et al. Degradation-aware residual-conditioned optimal transport for unified image restoration[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2025, 47(8): 6764-6779.
- [47] Zamfir E, Wu Ziwei, Mehta N, et al. Complexity experts

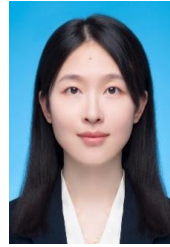
are task-discriminative learners for any image restoration[C]//2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2025: 12753-12763.

- [48] Mou Chong, Wang Qi, Zhang Jian. Deep generalized unfolding networks for image restoration[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2022: 17378-17389.
- [49] Kong Xiangtao, Dong Chao, Zhang Lei. Towards effective multiple-in-one image restoration: A sequential and prompt learning strategy[PP/OL]. V3. arXiv (2024-03-20) [2025-12-29]. <https://doi.org/10.48550/arXiv.2401.03379>.

作者简介



董 喆 男, 1997年出生于黑龙江省哈尔滨市。现为哈尔滨工业大学电子与信息工程学院副研究员。主要研究方向为多模态遥感大模型。中国电子学会会员编号:E190186330A。
E-mail: 1151874850@qq.com



刘天竹 女, 1990年出生于黑龙江省哈尔滨市。现为哈尔滨工业大学电子与信息工程学院研究员。主要研究方向为遥感信息智能处理。中国电子学会会员编号:E190025027M。
E-mail: tzliu@hit.edu.cn



孙钰哲 男, 2001年出生于黑龙江省北安市。现为哈尔滨工业大学电子与信息工程学院博士研究生。主要研究方向为跨模态遥感数据精细解译。中国电子学会会员编号:E190186333A。
E-mail: hityzsun@gmail.com



谷延锋 男, 1977年出生于黑龙江省佳木斯市。现为哈尔滨工业大学电子与信息工程学院教授。主要研究方向为高光谱遥感图像处理。中国电子学会会员编号:E190030341S。
E-mail: guyf@hit.edu.cn