DOI:10.13203/j.whugis20260243 文章编号:1671-8860(2026)07-1508-01

遥感智能解译基础模型关键技术研究

王迪 武汉大学计算机学院,湖北 武汉,430072

Key Technologies of Foundation Models for Remote Sensing Intelligent Interpretation

WANG Di School of Computer Science, Wuhan University, Wuhan 430072, China

中图分类号:P237 文献标识码:D

随着对地观测体系与遥感传感器技术的快速发展,遥感影像在数据规模快速增长的同时,其地理覆盖、空间分辨率、光谱范围、时间跨度与数据模态等也日益多样,对遥感解译技术提出了更高的要求。如何充分利用大规模遥感数据资源,构建具备通用表征能力、跨任务迁移能力以及高层语义理解与推理能力的人工智能基础模型,已成为遥感智能解译领域亟需解决的关键科学问题。

围绕遥感智能解译基础模型构建这一核心目标,该文针对遥感数据标注成本高、模态差异显著、场景结构复杂等问题,以及现有模型在跨任务迁移与推理能力方面存在的不足,以提升遥感基础模型的通用表征能力与跨任务适应能力为主线,从数据集构建、模型结构设计与预训练范式等方面展开研究,构建面向遥感场景的基础模型方法体系。在此基础上,进一步拓展至多模态语义理解与地理空间推理,并提出不依赖显式思维链监督的推理增强范式。

通过上述研究,该文在数据构建、结构设计、预训练方法与多模态推理能力提升等方面取得了重要进展。实验结果表明,所构建的遥感基础模型在多类下游任务中优于现有主流方法,表现出良好的计算效率、鲁棒性与跨任务迁移能力。与此同时,在缺乏显式推理路径监督的条件下,模型仍能实现稳定的跨场景地理空间推理,为构建通用、可扩展且具备自主思考能力的遥感人工智能基础模型提供了新的技术路径,对推动遥感智能解译向更高层次的通用化与自主化发展具有重要的理论意义与应用价值。主要研究内容如下:

1)在数据层面,针对遥感影像像素级标注成本高、难以高效扩增及大规模预训练数据不足的问题,构建了支撑遥感基础模型训练的数据资源体系。通过自动化标注策略扩展像素级语义分割监督信号规模,并整合多源传感器观测数据构建大规模高光谱预训练数据资源,为基础模型的表征学习与跨场景泛化提供了稳定的数据支撑。

2)在模型结构层面,针对遥感影像中普遍存在的地物尺度变化显著、目标朝向任意以及空间几何关系复杂等问题,同时考虑高光谱数据中空谱冗余严重的特性,设计了多种基础模型结构与关键组件,有效增强了模型对复杂空间结构与光谱特征的建模能力,为构建高效且可扩展的遥感基础模型提供了新的结构设计思路。

3)在预训练方法层面,针对现有遥感基础模型预训练范式以单一任务或单一架构为主、跨任务迁移能力受限的问题,系统研究了遥感监督预训练与多任务预训练两种典型范式,深入评估了视觉Transformer架构在大规模遥感监督预训练条件下的表征能力与迁移性能,并通过多任务协同预训练有效缓解了预训练与下游任务之间的差异,为提升遥感基础模型的通用性与跨任务适应性提供了新的方法路径。

4)在多模态推理层面,针对现有遥感多模态大语言模型高度依赖显式思维链监督、推理成本高且易引入人为偏差的问题,探索了一种不依赖预定义思维链标注的多模态推理增强范式,通过结合监督微调与基于强化学习的后训练策略,在获得基本遥感理解的同时,激活模型的深度推理能力,实现了跨任务、跨场景的地理空间推理能力涌现,为构建具备自主思考能力的遥感多模态基础模型指明了新的研究方向。

作者简介:王迪,2026年6月毕业于武汉大学(指导教师:陶大程教授、杜博教授、张良培教授、张敬教授),研究方向为遥感基础模型。
d_wang@whu.edu.cn

引文格式:WANG Di. Key Technologies of Foundation Models for Remote Sensing Intelligent Interpretation[J]. Geomatics and Information Science of Wuhan University, 2026, 51(7): 1508. DOI:10.13203/j.whugis20260243.(王迪. 遥感智能解译基础模型关键技术研究[J]. 武汉大学学报(信息科学版), 2026, 51(7): 1508. DOI:10.13203/j.whugis20260243.)