

面向空天地一体化的多模态知识增强 卫星智能计算方法

李 倩, 王尚广*

(北京邮电大学网络与交换技术国家重点实验室, 北京 100876)

摘 要: 随着空天地一体化体系的快速发展, 卫星端智能计算逐渐成为支撑遥感知、态势分析与应急决策的重要技术基础, 但其同时面临星载算力与存储受限、多模态异构数据噪声复杂以及任务语义动态变化等挑战。现有多模态融合方法多强调视觉-文本对齐或大模型驱动推理, 知识图谱方法强调结构化语义建模, 但二者大多面向地面高算力环境设计, 缺乏面向卫星边缘计算场景的统一部署机制。针对上述问题, 本文提出一种面向空天地一体化的知识引导多模态模块化卫星网络(Knowledge-guided Multi-Modal Modular Satellite Network, KM³-SatNet)。与直接引入全局知识图谱或采用固定融合结构的既有方法不同, 本文围绕“任务相关知识裁剪-知识约束鲁棒融合-预算感知动态推理-端云协同自适应部署”开展一体化设计: 首先构建卫星任务知识图谱, 并通过语义匹配、关系感知扩展与个性化 PageRank 实现动态知识子图抽取; 其次设计知识约束门控融合机制, 在视觉、文本与知识之间建立可解释的动态交互约束; 进一步提出模块化可裁剪推理结构与算力感知控制器, 实现星载预算约束下的自适应模块选择; 最后给出考虑链路波动与中断回退的端云协同优化策略。在多模态遥感任务数据集上的实验结果表明, 与典型多模态融合模型相比, 本文方法在目标识别任务上精度提高 6.3%, 在灾害场景识别任务上提高 5.8%, 同时相较于轻量级遥感多模态大语言模型(Remote Sensing Multimodal Large Language Model Lite, RS-MLLM-Lite) 参数量减少 50.0%, 相较于 Swin-L, 参数量减少 98.7%, 推理时延降低 67.7%。结果表明, KM³-SatNet 能够在保持较高精度的同时兼顾模型规模、时延、能耗与鲁棒性, 适用于卫星边缘环境下的多模态智能计算部署。

关键词: 空天地一体化; 卫星计算; 多模态融合; 知识增强; 任务推理; 端云协同

基金项目: 国家科技计划雄安新区科技创新专项(No.2024XAGG0025); 国家自然科学基金(No.62402054)

中图分类号: TP391

文献标识码: A

文章编号: 0372-2112(2026)06-2740-23

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

DOI: 10.12263/DZXB.20260290

Knowledge-Guided Multi-Modal Modular Satellite Network for Space-Air-Ground Integrated Intelligent Satellite Computing

LI Qian, WANG Shangguang*

(State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications,
Beijing 100876, China)

Abstract: With the rapid development of space-air-ground integrated systems, intelligent on-board satellite computing has become an important technical foundation for supporting remote sensing perception, situation analysis, and emergency decision making. However, it faces several challenges, including limited on-board computing power and storage, complex noise in multimodal heterogeneous data, and dynamic changes in task semantics. Existing multimodal fusion methods mainly emphasize vision-text alignment or large-model-driven reasoning, while knowledge graph methods focus on structured semantic modeling. Most of these methods are designed for ground-side high-performance computing environments and lack a unified deployment mechanism for satellite edge computing scenarios. To address these problems, this paper proposes a knowledge-guided multi-modal modular satellite network (KM³-SatNet) for space-air-ground integrated systems. Different from existing methods that directly introduce a global knowledge graph or adopt a fixed fusion structure, this paper develops an integrated design around “task-related knowledge pruning, knowledge-constrained robust fusion, budget-aware dynamic reasoning, and edge-cloud collaborative adaptive deployment”. First, a satellite task knowledge graph is constructed, and dynamic knowledge subgraph extraction is achieved through semantic matching, relation-aware expansion, and personalized PageRank. Second, a knowledge-constrained gated fusion mechanism is designed to establish interpretable dynamic interaction constraints among vision, text, and knowledge. Furthermore, a

modular and prunable inference structure and a computing-power-aware controller are proposed to realize adaptive module selection under on-board budget constraints. Finally, an edge-cloud collaborative optimization strategy considering link fluctuation and interruption fallback is presented. Experimental results on multimodal remote sensing task datasets show that, compared with typical multimodal fusion models, the proposed method improves accuracy by 6.3% on the object recognition task and by 5.8% on the disaster scene recognition task. Meanwhile, compared with the lightweight remote sensing multimodal large language model lite (RS-MLLM-Lite), the number of model parameters is reduced by 50.0%; compared with Swin-L, the number of parameters is reduced by 98.7%, and inference latency is reduced by 67.7%. The results demonstrate that KM³-SatNet can maintain high accuracy while balancing model scale, latency, energy consumption, and robustness, making it suitable for multimodal intelligent computing deployment in satellite edge environments.

Keywords: space-air-ground integration; satellite computing; multimodal fusion; knowledge enhancement; task reasoning; edge-cloud collaboration

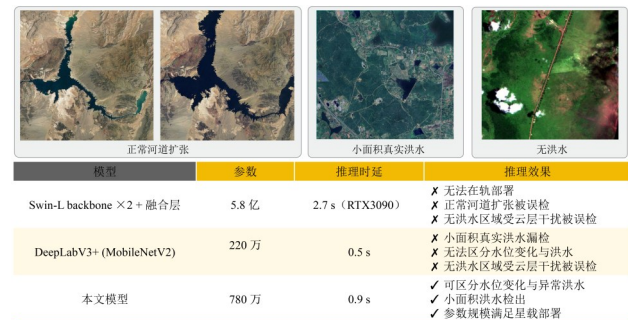
Foundation Item(s): Science and Technology Innovation Program of Xiongan New Area (No.2024XAGG0025); National Natural Science Foundation of China (No.62402054)

0 引言

空天地一体化信息系统的高速发展,推动卫星从传统数据采集平台向在轨智能处理节点演进,在洪水监测、灾害损失评估、应急指挥决策等重大民生与国防场景中,已成为不可或缺的技术支撑。其中,城市洪水精细分割作为灾害监测的核心任务,其结果直接服务于风险等级判定与应急资源调度,是保障人民生命财产安全的关键数据基础^[1]。与地面算力充足、链路稳定的计算环境不同,卫星端存在显著的资源约束特征,低轨卫星单节点算力受限,显著低于地面高端GPU (Graphics Processing Unit)的数十至数百TOPS (Tera Operations Per Second)算力水平,星地链路带宽有限且具有时变特性,同时遥感任务需求随监测场景动态变化。传统依赖大规模深度学习模型的智能分析方法,因计算开销高、实时性差等问题难以直接迁移至星载环境,因此,构建兼顾推理精度与星载算力、存储、时延约束的在轨智能计算框架,已成为空天地一体化智能计算领域的迫切需求^[2-3]。

城市洪水精细分割任务的核心难点在于异常水体的精准判别。在光学、合成孔径雷达 (Synthetic Aperture Radar, SAR) 遥感图像中,正常河道扩张、水库水位动态变化与小面积真实洪水区域的水体特征高度相似,这使得模型极易出现两类关键误差:一是将正常水体变化误判为灾害洪水区域,二是漏检小尺度真实淹没区域,最终导致分割结果失准并引发应急决策偏移。而现有方法难以融合永久水体地理属性、水体时序变化信息、地形空间结构语义等多维度知识开展精准推理,进一步加剧了上述问题。如图1所示,在洪水精细分割任务中,现有模型普遍面临“规模-性能”的两难冲突。以双分支Swin-L为代表的大模型(5.8亿参数)虽精度潜力高,但推理时延达2.7 s (RTX3090),完全无法在轨部署,且易将正常河道扩张误检为洪水、受云层干扰误检为无洪水区域;以Dee-

pLabV3+(MobileNetV2)为代表的轻量化模型(220万参数)虽推理时延仅0.5 s,却存在小面积真实洪水漏检、无法区分水位变化与洪水的问题。



模型	参数	推理时延	推理效果
Swin-L backbone ×2 + 融合层	5.8 亿	2.7 s (RTX3090)	✗ 无法在轨部署 ✗ 正常河道扩张被误检 ✗ 无洪水区域受云层干扰被误检
DeepLabV3+ (MobileNetV2)	220 万	0.5 s	✗ 小面积真实洪水漏检 ✗ 无法区分水位变化与洪水 ✗ 无洪水区域受云层干扰被误检
本文模型	780 万	0.9 s	✓ 可区分水位变化与异常洪水 ✓ 小面积洪水检出 ✓ 参数规模满足星载部署

图1 洪水精细分割任务中的模型规模-性能冲突示意图(RTX3090环境下推理)

Figure 1 Model size-performance conflict in fine-grained flood segmentation under the RTX3090 inference environment

近年来,多模态融合与大模型技术显著提升了遥感图像分割与场景理解能力,但现有方法大多面向地面高算力环境设计,在星载边缘计算场景中的适配性仍然不足。以双分支Swin-L为代表的多模态Transformer模型^[4],虽然能够取得较高分割精度,但模型规模达到5.8亿参数,推理时延在高端GPU环境下仍超过2 s,难以满足卫星端实时推理需求;以DeepLabV3+(MobileNetV2)为代表的轻量化模型^[5],虽然在参数规模和时延方面更适合星载部署,但由于缺乏领域知识约束,在云遮挡、SAR相干斑和水体语义混淆等复杂条件下,容易出现小尺度洪水漏检以及正常水位变化与异常洪水扩张难以区分的问题,模型鲁棒性和语义判别能力仍然受限。进一步分析现有研究可以发现,其在星载端应用中主要存在三方面不足。首先,现有多模态融合方法(如Cross-Attention、RS-MLLM等)^[6-7]主要关注视觉-文本特征对齐与语义

增强,但通常默认算力和存储资源相对充足,缺乏针对星载场景中云层遮挡、光照突变、SAR 噪声等复杂干扰的鲁棒机制,且模型参数量和推理开销普遍较大。其次,现有知识图谱方法^[8-9]虽能够提供结构化先验,但多采用全局图谱直接参与推理,缺少面向具体任务的知识筛选与裁剪机制,因而容易引入无关知识并增加额外计算负担,难以适应卫星端受限的算力与存储条件。最后,现有星地端云协同策略^[10-11]更多侧重任务卸载与资源分配,尚未充分考虑多模态知识增强推理过程中的模块依赖关系和链路时变特性,因而难以在保证推理连续性的同时实现精度、时延与能耗的统一优化。此外,已有星载智能计算方法大多采用固定推理结构^[12],缺乏根据实时预算、任务需求和输入数据质量动态调整推理路径的能力,难以实现性能与开销之间的自适应平衡。

基于上述分析,本文关注的核心问题是:在星载端算力、存储和时延均受到严格约束的条件下,如何实现任务语义约束、多模态鲁棒融合以及推理过程自适应调度的一体化设计。围绕这一问题,本文提出面向空天地一体化场景的知识引导多模态模块化卫星网络(Knowledge-guided Multi-modal Modular Satellite Network, KM³-SatNet)。与现有方法相比,本文的主要区别不在于简单叠加知识图谱、多模态融合和端云协同模块,而在于面向星载部署约束,构建了一种统一的任务驱动推理框架。具体而言:其一,本文提出任务驱动的动态知识子图构建机制,通过任务语义匹配、关系感知扩展与个性化 PageRank 压缩,将知识使用方式由全局图谱推理转化为任务相关子图推理,从而在降低推理开销的同时提高知识约束的有效性;其二,本文提出知识约束的双粒度门控融合机制,以任务知识为条件,同时对视觉 Token 保留和跨模态交互强度进行约束,从而抑制云遮挡、SAR 相干斑等噪声引起的错误关联,提升多模态语义融合的鲁棒性;其三,本文提出预算感知的模块化可裁剪推理与端云协同优化机制,通过轻量化控制器根据实时算力预算、任务需求和链路状态动态选择推理结构与执行位置,实现精度、时延与能耗的协同平衡。在多模态遥感任务数据集上的实验结果表明, KM³-SatNet 在多模态遥感目标识别任务中精度提升 6.3%,在灾害场景识别任务中提升 5.8%;同时,相较于轻量级遥感多模态大语言模型(Remote Sensing Multimodal Large Language Model Lite, RS-MLLM-Lite),模型参数量减少 50.0%;相较于 Swin-L,参数量减少 98.7%,推理时延降低 67.7%。在云遮挡、噪声干扰和光照变化等复杂条件下,模型平均精度衰减率仅为 4.5%,存储占用控制在 50 MB 以内。上述结果表明,本文方法能够在保持较

高推理精度的同时显著降低资源消耗,为受限星载环境中的多模态知识增强智能计算提供了一种可部署的技术路径。

1 相关工作

1.1 卫星智能计算与多模态融合方法

卫星智能计算的核心目标是在星载算力、存储和时延受限条件下,实现遥感任务的高精度与高鲁棒性推理。早期研究主要围绕轻量化网络、模型剪枝、量化和硬件加速展开,常借助 MobileNet、ShuffleNet 等骨干支撑土地覆盖分类和基础目标检测等在轨任务^[13-16]。随着空天地一体化体系的发展,单源遥感信息已难以满足灾害监测和复杂场景理解需求,多模态融合逐渐成为提升遥感智能计算性能的重要路径^[6]。现有多模态遥感方法主要包括特征级融合和决策级融合。其中特征拼接、跨注意力(Cross-Attention)等方法能够联合建模光学与 SAR 等多源信息,并在地面高算力环境下取得较好效果。然而,这类方法在星载边缘场景中仍面临明显局限:一是云层遮挡、SAR 相干斑和光照变化等噪声会破坏跨模态对齐质量,导致融合特征失真^[17-19];二是不同模态的成像机理差异会加剧语义鸿沟^[20];三是复杂融合结构通常伴随较高计算与存储开销,难以满足星载端实时部署需求。近年来,遥感多模态大语言模型(RS-MLLM)进一步推动了高层语义理解研究,并在洪水、地震等灾害监测任务中展现出潜力^[4]。但现有 RS-MLLM 多面向地面或云端环境,参数规模大、部署成本高,即使经过轻量化压缩,也普遍缺乏与遥感领域先验知识的深度结合,因而仍难有效处理“正常水体变化与异常洪水扩张混淆”以及“复杂噪声条件下语义失真”等问题。此外,现有研究虽在云去除、图像配准等预处理方面取得进展^[18-19, 21-22],并通过多模态数据集缓解了数据异构性问题^[23],部分语义分割方法也取得了较好性能^[24-25],但总体上仍缺乏面向星载约束的端到端一体化设计,尚未系统解决精度、鲁棒性与资源开销之间的平衡问题。综上,现有卫星智能计算与多模态融合研究仍主要服务于地面高资源环境,缺乏面向星载边缘场景的知识引导型轻量化融合框架,尚难在严格算力与时延约束下实现复杂噪声条件下的稳健语义推理。

1.2 知识图谱推理与端云协同方法

知识图谱推理通过显式建模实体、属性和关系提升遥感推理的可解释性与逻辑一致性,星地端云协同通过联合调度星载端与地面端资源突破单星算力上限。在知识图谱推理方面,现有研究已应用于卫星任务规划和遥感地物解译等问题^[3, 26-31]。传统方法多

直接基于全局知识图谱开展推理,虽然能够引入结构化先验,但迁移到星载端时存在明显不足:其一,全局图谱规模较大,额外计算与存储开销难以被星载平台承受;其二,缺乏任务驱动的知识筛选与压缩机制,易引入无关知识并干扰推理;其三,知识表示与多模态特征融合之间耦合不足,尚难有效缓解遥感数据中的噪声干扰与语义模糊问题。近年来,大语言模型与知识图谱融合为遥感推理提供了新思路,并在卫星集群任务规划等场景中表现出一定优势^[12],但这类方法同样依赖较强算力支持,尚未形成适用于星载环境的轻量化统一设计。在星地端云协同方面,现有研究主要聚焦任务卸载和资源分配。针对低轨卫星高速移动、链路时变和能源受限等特点,已有方法构建了时延-能耗联合模型,并提出卫星协作卸载、能量感知卸载和区域自适应卸载等策略,以缓解单星计算压力^[32-37]。在资源优化方面,双层博弈、资源聚类、联邦学习等方法被用于联合配置通信与计算资源,从而提升资源利用率^[38-41];数字孪生与星地融合网络为协同调度提供了动态优化手段^[5, 42-43],深度强化学习则被用于处理卫星移动性带来的资源管理问题^[44-45]。然而,现有端云协同方法大多面向通用计算卸载,尚未充分考虑多模态知识增强推理中的模块依赖关系、知识子图协同传输需求以及链路退化场景下的鲁棒执行问题,因此难以实现精度、时延和能耗的联合优化。总体来看,现有知识图谱推理与星地端云协同研究尚未形成面向星载场景的统一框架:前者缺少任务驱动轻量知识使用机制,后者缺少与多模态知识增强推理深度耦合协同设计。本文正是在这一基础上,进一步探索面向空天地一体化场景的知识引导多模态模块化卫星智能计算方法。

2 问题定义

卫星端多模态知识增强推理任务以空天地一体化卫星智能计算需求为导向,核心是基于多模态遥感数据与领域先验知识实现精准、高效的在轨语义推理,为遥感感知、灾害监测等核心任务提供技术支撑,其输入与输出可通过以下空间定义精准刻画。视觉 Token 索引记为 n , 文本 Token 索引记为 m , 知识实体索引记为 v , 关系类型索引记为 r , 模块索引记为 k ; 任务语义向量记为 $\mathbf{z}_t$, 实体嵌入与关系嵌入分别记为 $\mathbf{e}_v$ 和 $\mathbf{e}_r$, 模块配置向量记为 $\mathbf{c}$, 星载实时预算记为 B 。卫星端智能计算任务的输入空间为 $\mathbf{X} = \mathbf{X}_{\text{vis}} \times \mathbf{X}_{\text{txt}} \times \mathbf{G}_{\text{all}}$, 其中 $\mathbf{X}_{\text{vis}} = \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ 为多模态视觉遥感数据集, 包含光学图像、SAR 图像等, 单幅图像记为 $\mathbf{x} \in \mathbf{X}_{\text{vis}}$, 分辨率为 $H \times W$; $\mathbf{X}_{\text{txt}} = T_1, T_2, \dots, T_m$ 为自然语言任务指令集合, 单条指令记为 $T = w_1, w_2, \dots, w_k (w_i$

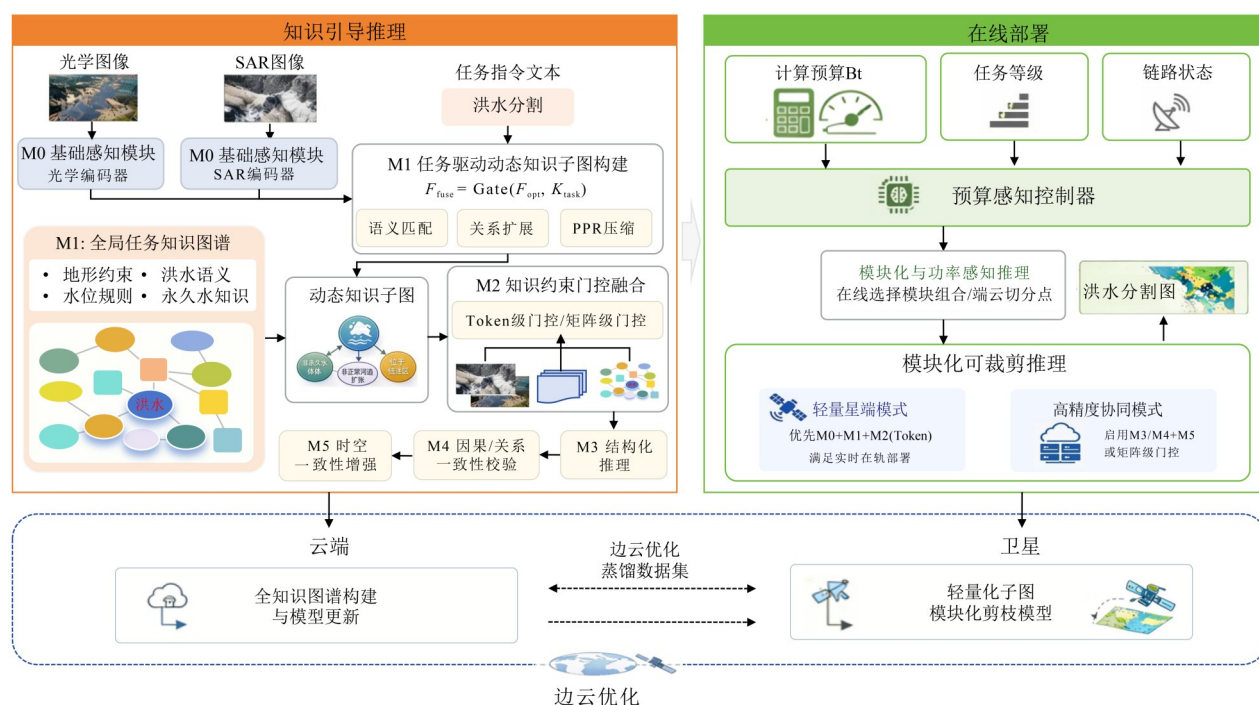
为分词词汇), 用于表征任务语义需求(如“洪水区域精细分割”“建筑目标检测”); $\mathbf{Q}_{\text{all}} = (\mathbf{Z}, \mathbf{R}, \mathbf{C})$ 为全局卫星任务知识图谱, $\mathbf{Z}$ 为实体集合(如洪水、永久水体、低洼区), $\mathbf{R}$ 为关系集合(如 *caused_by*、*related_to*), $\mathbf{C} \subseteq \mathbf{Z} \times \mathbf{R} \times \mathbf{Z}$ 为三元组边集合, 承载遥感领域先验知识。设任务输出空间为 $\mathbf{Y}$, 根据卫星遥感任务类型分为两类。若为目标识别任务, $\mathbf{y} \in \mathbf{Y}$ 为目标边界框与类别标签的集合, 即 $\mathbf{y} = (b_1, c_1), (b_2, c_2), \dots, (b_n, c_n)$, 其中 b_i 为目标边界框坐标, c_i 为目标类别。若为灾害场景识别/分割任务, $\mathbf{y} \in \mathbf{Y}$ 为像素级语义标签矩阵(分割)或场景类别标签(分类), 分割任务中 $\mathbf{y} \in \mathbf{R}^{H \times W}$, 每个像素值对应语义类别(如洪水、正常水体、建筑)。卫星端智能计算的核心映射关系为在资源约束下, 学习映射函数 $f: \mathbf{X} \rightarrow \mathbf{Y}$, 实现从多模态输入与全局知识图谱到任务推理结果的精准映射。

3 KM³-SatNet 方法

针对空天地一体化星载端的算力、存储、时延硬约束, 以及多模态遥感数据噪声干扰、知识利用效率低、推理结构无法动态适配等核心问题, 本文提出知识引导多模态模块化卫星网络(KM³-SatNet)。该网络整体架构如图 2 所示, 由输入层、感知层、知识层、推理层和部署层组成, 各层协同实现“多模态输入→知识增强→动态推理→端云协同”的端到端在轨智能推理, 在满足星载资源约束的前提下, 提升遥感任务推理精度与鲁棒性。其中, 输入层接收多模态遥感数据与任务文本指令; 感知层完成视觉与文本特征的轻量化提取; 知识层基于任务语义构建动态知识子图, 提供领域先验约束; 推理层通过知识约束门控融合与模块化推理单元完成精准任务决策; 部署层结合端云协同机制实现算力自适应优化, 确保模型适配星载边缘计算环境。

3.1 任务驱动的动态知识子图构建

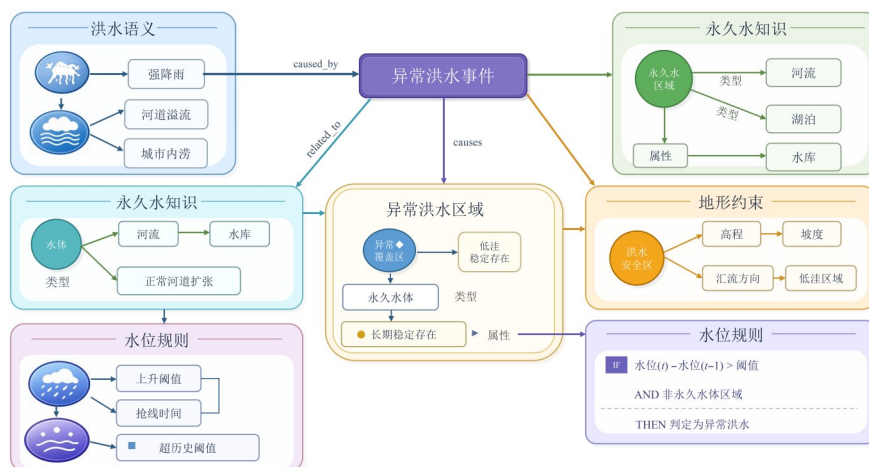
星载端算力与存储资源受限, 无法直接采用全局卫星任务知识图谱开展推理, 且不同卫星遥感任务(如洪水灾害监测、航道异常检测)的知识需求存在显著差异, 无关知识会增加计算开销并导致推理偏差。为此, 本文提出任务驱动的动态知识子图构建机制, 核心目标是从全局知识图谱中精准提取与当前任务强相关的局部知识单元, 构建“任务专用”的轻量化知识子图, 将知识推理由全局图谱推理转化为任务子图推理, 在保证知识支撑有效性的前提下, 降低推理复杂度, 适配星载资源约束。该机制围绕“语义锚定-关联检索-结构裁剪”三步展开, 整体流程如图 3 所示(以洪水场景为例)。



注:该架构由知识引导推理、在线部署与星地端云协同优化三大核心模块构成,通过任务驱动的动态知识子图构建、知识约束门控融合、模块化可裁剪推理及端云协同优化,实现星载端多模态知识增强推理的精度与算力平衡。M0表示基础感知模块,M1表示任务驱动动态知识子图构建,M2表示知识约束门控融合,M3表示结构化推理,M4表示因果/关系一致性校验,M5表示时空一致性增强。预算感知控制器根据实时算力预算、任务等级和链路状态,在“轻量星端模式”和“高精度协同模式”之间动态切换。

图2 KM³-SatNet整体架构示意图

Figure 2 Overall architecture of KM³-SatNet



注:图中构建洪水语义、永久水体知识、地形约束、水位规则四大知识域,定义核心实体与语义关系,设计异常洪水推理规则,为在轨洪水精细分割提供结构化知识约束,支撑动态知识子图的构建与后续推理。

图3 洪水场景下卫星任务知识图谱结构与推理规则示意图

Figure 3 Structure and reasoning rules of the satellite task knowledge graph (subgraph) in flood scenarios

3.1.1 任务语义向量建模

任务语义向量建模是动态知识子图构建的基础,其核心是将非结构化自然语言任务指令转化为结构化、可计算的低维语义向量,为全局知识图谱的语义匹配提供统一度量基准,同时适配星载端的计算效率

需求。给定自然语言任务文本指令 $T = w_1, w_2, \dots, w_k$, 其中 w_i 为任务指令中的分词词汇,为兼顾星载端的计算效率与语义表示的准确性,采用轻量化预训练文本编码器 $f_{\text{text}}(\cdot)$ 对任务文本进行上下文语义编码,生成初始任务语义向量:

$$\mathbf{z}_i = f_{\text{text}}(T) \in \mathbf{R}^{d_i} \quad (1)$$

其中, d_i 为语义向量的嵌入维度, 轻量化编码器的设计可有效降低星载端的编码计算开销, 同时保留任务指令的核心语义信息。卫星边缘环境下的遥感任务推理对语义表示的稳定性与鲁棒性要求较高, 初始语义向量易受文本表述形式、分词误差等因素影响, 出现向量分布离散的问题, 进而降低后续语义匹配的准确性。为此, 对初始任务语义向量进行 L_2 归一化与温度缩放联合处理, 得到标准化任务语义向量:

$$\tilde{\mathbf{z}}_i = \frac{\mathbf{z}_i}{\|\mathbf{z}_i\|_2} \cdot \frac{1}{\tau} \quad (2)$$

其中, τ 为温度系数 (取值范围为 0.1~1.0), 用于调节语义向量分布的“锐化”程度。当 τ 减小时, 向量间的语义区分度提升, 可增强后续语义匹配 τ 的精准性; 当 τ 增大时, 向量分布更平滑, 可提升匹配过程的抗干扰能力。该标准化处理既消除了向量尺度差异对语义匹配的影响, 又通过温度系数的灵活调节, 实现了语义表示区分性与鲁棒性的平衡, 使任务语义向量能够精准刻画任务的核心语义约束, 成为后续子图检索与裁剪的有效驱动信号。标准化后的任务语义向量将与全局知识图谱的实体/关系嵌入空间进行对齐, 为后续计算任务与知识图谱实体间的语义相似度、筛选任务相关实体奠定基础, 是实现“任务语义-知识图谱”关联的关键桥梁。

3.1.2 全局知识图谱语义匹配机制

设全局卫星任务知识图谱为 $\mathbf{Q} = (\mathbf{Z}, \mathbf{R}_{\text{rel}}, \mathbf{C})$, 其中 $\mathbf{Z}$ 为图谱中的实体集合, 涵盖卫星任务相关的目标、场景、操作、属性等各类知识实体, $\mathbf{R}_{\text{rel}}$ 为实体间的关系类型集合, 表征实体间的语义关联, $\mathbf{C} \subseteq \mathbf{Z} \times \mathbf{R}_{\text{rel}} \times \mathbf{Z}$ 为三元组边集合, 是知识图谱的基本语义单元。全局知识图谱语义匹配机制的核心是建立任务语义向量与知识图谱实体关系的语义关联, 通过语义相似度计算与结构化扩展, 筛选出与任务强相关的实体和关系集合, 为构建任务子图提供基础要素。为实现知识图谱实体与关系的数值化表示, 使其能够与任务语义向量进行定量的语义相似度计算, 对全局知识图谱中的每个实体 $\mathbf{v} \in \mathbf{Z}$ 与关系 $\mathbf{r} \in \mathbf{R}_{\text{rel}}$ 进行低维嵌入学习, 得到嵌入向量:

$$\mathbf{e}_v \in \mathbf{R}^{d_k}, \mathbf{e}_r \in \mathbf{R}^{d_k} \quad (3)$$

其中, d_k 为知识嵌入维度, 与后续任务语义向量的映射维度保持一致。考虑到卫星端的计算资源约束, 选用计算开销低、训练效率高的 TransE 模型作为知识嵌入方法, 该模型通过将实体与关系映射到同一低维空间, 利用向量的平移不变性刻画实体间的语义关系, 其能量函数为

$$\phi(h, r, t) = \|\mathbf{e}_h + \mathbf{e}_r - \mathbf{e}_t\|_2 \quad (4)$$

其中, $(h, r, t) \in \mathbf{C}$ 为知识图谱中的正采样三元组。为提升嵌入学习的效果, 采用基于间隔的排序损失对模型进行训练, 训练损失函数为

$$L_{\text{kg}} = \sum [\gamma + \phi(h, r, t) - \phi(h', r, t')]_{+(h, r, t) \in \mathbf{C}} \quad (5)$$

其中, (h', r, t') 为负采样三元组, 通过随机替换正三元组的头实体或尾实体生成; γ 为间隔超参数 (取值为 1.0~2.0), 用于区分正、负三元组的能量值; $[\cdot]_+ = \max(\cdot, 0)$ 为 ReLU 函数, 仅当负三元组的能量值小于正三元组时计算损失。该训练方式可使正三元组的向量平移关系更贴合实际语义, 负三元组则远离合理的语义空间, 从而保证实体关系嵌入向量的语义表征能力。

任务语义向量的嵌入维度 d_i 与知识图谱实体关系的嵌入维度 d_k 通常不一致, 无法直接进行语义相似度计算, 因此需引入可训练线性映射层, 实现任务语义空间到知识图谱语义空间的精准投影与维度对齐, 映射过程为

$$\mathbf{q}_i = \mathbf{W}_q \tilde{\mathbf{z}}_i, \mathbf{W}_q \in \mathbf{R}^{d_k \times d_i} \quad (6)$$

其中, $\mathbf{W}_q$ 为线性映射矩阵, 通过端云协同训练进行优化, 使映射后的任务语义向量 $\mathbf{q}_i \in \mathbf{R}^{d_k}$ 能够在知识图谱语 $O(|V|d_k)$, 难以满足卫星端的实时性要求。种子实体检索的核心是筛选与当前任务语义强相关的实体, 作为后续子图扩展的起点。采用余弦相似度计算映射后任务语义向量 $\mathbf{q}_i$ 与每个实体嵌入向量 $\mathbf{e}_v$ 的语义相似度, 相似度公式为

$$s(v|t) = \frac{(\mathbf{q}_i \cdot \mathbf{e}_v)}{(\|\mathbf{q}_i\|_2 \cdot \|\mathbf{e}_v\|_2)} \quad (7)$$

选取相似度最高的 Top- K 个实体作为种子实体集合 S_i (K 为超参数, 根据星载存储约束取值为 5~20)。若采用朴素计算方式, 该过程需遍历全局所有实体, 时间复杂度为 $O(|V|d_k)$, 难以满足星载端实时性要求。为此, 引入近似最近邻 (Approximate Nearest Neighbor, ANN) 索引算法 [星载端部署轻量化 FAISS (Facebook AI Similarity Search) 压缩索引] 对实体嵌入向量建立索引, 并配合常数级向量距离计算。星载端可部署轻量化的压缩索引, 或仅保留候选域索引, 在保证检索精度的前提下, 进一步降低星载端的内存占用与计算代价, 适配边缘计算的实时性要求。

仅依靠种子实体集合无法完整表征任务的知识关联结构, 易丢失实体间的关键语义关系, 导致任务子图的知识表达不完整。为此, 以种子实体集合 Z_i 为起点, 在全局知识图谱中进行 H -hop 邻域扩展, 挖掘种子实体的关联实体, 得到候选节点集合:

$$\mathbf{Z}_i^{\text{cand}} = N^{(H)}(\mathbf{Z}_i) \quad (8)$$

其中: $N^{(H)}(\cdot)$ 为 H -hop 邻域算子, 表示从种子实体出发, 遍历距离为 H 步的所有关联实体; H 为扩展步长超参数, 需根据任务的知识关联复杂度合理设置。

邻域扩展过程中易引入与任务无关的关系和边, 造成候选子图的规模冗余, 因此需引入关系门控机制对候选边进行筛选, 抑制无关关系的干扰。首先由任务语义向量预测每种关系的任务相关重要性, 得到关系门控权重:

$$a_r(t) = \sigma(\mathbf{w}_r^T \mathbf{q}_t) \in (0, 1) \quad (9)$$

其中: $\sigma(\cdot)$ 为 Sigmoid 激活函数, 将权重映射至 $(0, 1)$ 区间; $\mathbf{w}_r$ 为关系 r 的可训练参数, 用于刻画任务语义与关系间的关联程度。关系门控权重越大, 表示该关系与当前任务的语义相关性越强, 应予以保留; 反之则应弱化或删除。基于关系门控权重与实体语义相关性得分, 对候选边 $(u, r, v) \in \mathbf{C} \cap (\mathbf{Z}_t^{\text{cand}} \times \mathbf{R} \times \mathbf{Z}_t^{\text{cand}})$ 定义任务相关综合权重, 综合考量关系与实体的任务相关性:

$$w(u, r, v|t) = a_r(t) \cdot \frac{s(u|t) + s(v|t)}{2} \quad (10)$$

保留综合权重最高的 Top- M 条边 (M 为超参数, 用于控制初步任务子图的边规模, 取值为 20~50), 构成初步任务子图的边集合 $\mathbf{E}_t^0$ 。通过关系感知的结构扩展与边筛选, 既保证了任务子图的知识结构完整性, 又有效剔除了无关边, 实现了候选子图规模的初步约束。星端侧使用压缩 FAISS 索引完成实体 ANN 检索, Top- K 种子实体数 K 在 5~20 内搜索。候选子图通过 H -hop 扩展 ($H \in 1 \sim 4$) 与 (Personalized PageRank) Top- M 边筛选 ($M \in 20 \sim 60$) 得到, 随后采用 PPR 选择 Top- L 节点 ($L \in 10 \sim 30$, 重启概率 $\eta \in 0.2 \sim 0.4$), 保证了知识覆盖、计算代价与实时性之间的平衡。

3.1.3 动态子图压缩

通过上述步骤得到的候选子图 ($\mathbf{Z}_t^{\text{cand}}, \mathbf{Z}_t^0$) 仍可能存在部分冗余节点, 需进一步进行精细化裁剪, 将子图规模严格约束在卫星端可承受的范围内, 同时通过复杂度分析验证方法在星载端的可部署性。为实现对候选子图的精细化裁剪, 引入任务驱动的个性化 PageRank (PPR) 算法对候选节点进行二次筛选, 该算法以任务相关的种子实体为重启节点, 通过图上的随机游走挖掘与种子实体关联紧密的节点, 能够在保留任务核心知识的前提下, 实现子图规模的有效压缩。

设 $\mathbf{A}_t$ 为候选子图的行归一化邻接矩阵, $\mathbf{s}_t$ 为重启分布向量, 对种子实体集合 $\mathbf{Z}_t$ 中的节点赋予均匀的重启概率, 其余节点重启概率为 0; $\eta \in (0, 1)$ 为重启概率超参数, 表示随机游走过程中回到种子实体的概率。个性化 PageRank 的迭代公式为

$$\mathbf{p}_t = (1 - \eta) \mathbf{A}_t^T \mathbf{p}_t + \eta \mathbf{s}_t \quad (11)$$

其中, $\mathbf{p}_t$ 为节点的重要性得分向量, 迭代至收敛后, 得分越高的节点表示与种子实体的关联越紧密, 与当前任务的语义相关性越强。

选取 $\mathbf{p}_t$ 中值最大的 Top- L 个节点作为最终任务子图的实体集合 $\mathbf{Z}_t$, 并从初步边集合 $\mathbf{R}_t^0$ 中筛选出两个端点均属于 $\mathbf{Z}_t$ 的边, 构成最终任务子图的边集合 $\mathbf{R}_t$, 最终得到任务驱动的动态知识子图:

$$\mathbf{Q}_t = (\mathbf{Z}_t, \mathbf{C}_t), \mathbf{C}_t = \{(u, r, v) \in \mathbf{C}_t^0 \mid u, v \in \mathbf{Z}_t\} \quad (12)$$

3.2 知识约束门控的多模态融合

多模态数据 (光学图像、SAR 图像、任务文本、知识图谱) 的协同分析是卫星遥感任务精准推理的核心, 但真实遥感数据常伴随云层遮挡、SAR 相干斑、光照突变等复杂噪声, 易导致多模态特征融合失真, 影响推理精度。传统跨模态注意力主要依据内容相似性进行对齐, 缺乏“哪些视觉 Token 值得保留、哪些交互应被削弱”的外部任务先验。本文以任务语义和动态知识子图为条件, 分别在 Token 选择和交互矩阵两层施加知识约束, 其中, Token 级门控面向严格实时预算, 矩阵级门控面向算力相对充足的边缘/云端场景, 两者并非简单的替代关系, 而是预算自适应的互补设计。为此, 本文提出知识约束门控的多模态融合框架, 利用任务知识图谱的先验语义与结构约束, 构建动态门控机制, 主动抑制视觉噪声干扰, 强化学术特征与视觉特征的有效交互, 在星载资源受限环境下实现稳健、精准的多模态语义理解。该框架基于模块化设计, 分为多模态特征表示、知识引导门控矩阵生成、知识约束融合机制三个部分。

3.2.1 多模态特征表示

为适配星载端的轻量化部署需求, 框架采用模块化设计对输入的遥感图像、任务文本和动态生成的任务子图进行独立编码, 生成统一语义空间下的多模态特征向量, 为后续融合奠定基础。输入数据包括原始遥感图像 $\mathbf{x}$ 、自然语言任务指令 T 以及由 3.1 节生成的任务驱动知识子图 $\mathbf{Q}_t$ 。针对遥感图像的局部纹理与全局语义特征, 采用轻量化视觉编码器 $f_v(\cdot)$ 对输入图像 $\mathbf{x}$ 进行处理, 将其分割为固定数量的图像 Patch 并映射为高维语义向量, 输出视觉特征序列 $\mathbf{V} = f_v(\mathbf{x}) \in \mathbf{R}^{N \times d}$, 其中 N 为视觉 Token 的数量, d 为特征嵌入维度。同样, 针对任务文本指令的语义内涵, 利用轻量化文本编码器 $f_t(\cdot)$ 对文本 T 进行编码, 生成文本特征序列 $\mathbf{L} = f_t(T) \in \mathbf{R}^{M \times d}$, 其中 M 为文本 Token 的数量。视觉与文本编码器均选用轻量化预训练模型, 通过参数共享或知识蒸馏策略, 确保在推理阶段具备低延迟、高计算效率的特性。知识子图 $\mathbf{Q}_t$ 承载了任务相关的结构化语义信息, 为将其融入融合框架, 采

用关系图卷积网络(Relational Graph Convolutional Network, R-GCN)作为子图编码器。R-GCN能够显式建模知识图谱中不同类型关系的贡献,有效区分实体间的关联强弱,其核心传播规则为

$$\mathbf{H}^{(l+1)} = \phi \left(\sum_{r \in \mathbf{R}} \sum_{u \in N_r(v)} \frac{1}{c_{v,r}} \mathbf{W}_r^{(l)} \mathbf{h}_u^{(l)} \right) \quad (13)$$

其中, $\mathbf{H}^{(0)} = \{\mathbf{e}_v\}_{v \in V_i}$ 为知识子图中实体的初始嵌入向量, $\mathbf{R}$ 为关系类型集合, $N_r(v)$ 为实体 v 在关系 r 下的邻居实体, $c_{v,r}$ 为归一化常数以稳定训练, $\mathbf{W}_r^{(l)}$ 为关系 r 对应的可训练权重矩阵, $\phi(\cdot)$ 为非线性激活函数 GELU。通过 L_g 层信息传递,最终得到表征知识子图语义与结构的知识特征序列 $\mathbf{K} = \mathbf{H}^{(L_g)} \in \mathbf{R}^{P \times d}$, 其中 $P = |\mathbf{Z}_i|$ 为知识子图中的实体数量, $\mathbf{K}$ 中的每个向量对应一个知识实体的语义表示。为实现多模态特征的高效交互与门控控制,进一步计算各模态的全局摘要向量。通过对各模态特征序列取平均,得到能够表征整体语义倾向的聚合向量:

$$\bar{\mathbf{v}} = \frac{1}{N} \sum_{i=1}^N \mathbf{V}_i, \bar{\mathbf{l}} = \frac{1}{M} \sum_{j=1}^M \mathbf{L}_j, \bar{\mathbf{k}} = \frac{1}{P} \sum_{p=1}^P \mathbf{K}_p \quad (14)$$

视觉、文本、知识的全局摘要向量 $\bar{\mathbf{v}}_i, \bar{\mathbf{l}}, \bar{\mathbf{k}}$ 分别浓缩了对应模态的核心语义信息,将作为门控生成的关键引导信号,实现跨模态的语义对齐与交互控制。

3.2.2 知识引导门控矩阵生成

多模态融合的核心挑战是在抑制噪声的同时,强化有效特征交互。本文设计的知识引导门控机制,以任务知识(任务语义与图谱结构约束)为主导,动态决定视觉 Token 与文本/知识 Token 的交互强度,屏蔽噪声带来的无效特征关联。基于模态全局摘要构建知识引导查询向量,以此为基准生成动态门控权重,结合星载算力约束,设计两种不同复杂度的门控生成策略,适配不同部署场景。

门控生成的核心逻辑是基于模态全局摘要,构建知识引导查询向量,以此为基准对视觉 Token 或跨模态 Token 对生成动态门控权重。具体构造知识引导查询向量 $\mathbf{g}_i = \mathbf{W}_g [\bar{\mathbf{l}}; \bar{\mathbf{k}}] \in \mathbf{R}^d$, 其中 $\mathbf{W}_g$ 为可训练的投影矩阵, $[\bar{\mathbf{l}}; \bar{\mathbf{k}}]$ 表示将文本与知识全局摘要向量进行拼接。该向量融合了任务文本语义与知识图谱结构信息,能够精准反映当前任务的核心关注点,为特征交互提供权威的语义引导。基于 $\mathbf{g}_i$ 设计两种不同复杂度的门控生成策略,以适配不同的星载算力场景。

针对星载端算力严格受限的场景,采用 Token 级门控,仅对每个视觉 Token 的重要性进行独立评估,实现快速筛选。具体操作是,将每个视觉 Token $\mathbf{V}_i$ 与

知识引导查询向量 $\mathbf{g}_i$ 拼接,通过 Sigmoid 激活函数生成门控值,表征该视觉 Token 参与后续交互的必要性:

$$\mathbf{m}_i = \sigma(\mathbf{w}_m^T [\mathbf{V}_i; \mathbf{g}_i]) \in (0, 1) \quad (15)$$

所有视觉 Token 的门控值构成门控向量 $\mathbf{m} \in \mathbf{R}^N$ 。门控值接近 1 的 Token 被判定为与任务强相关,保留并参与后续交互;门控值接近 0 的 Token 则被视为噪声或无关信息,进行弱化或抑制。该策略计算开销极低,仅涉及向量拼接与单层感知机计算,能够在星载端实现实时推理。

对于地面站或具备较强算力的边缘节点,为实现更精细、更鲁棒的交互控制,采用交互矩阵门控。该策略不仅对视觉 Token 进行筛选,还对视觉与文本知识 Token 之间的两两交互关系进行显式建模与加权,从而构建全局的交互约束门控矩阵 $\mathbf{G}$:

$$\mathbf{G} = \sigma \left(\frac{(\mathbf{V}\mathbf{W}_1)(\mathbf{L}\mathbf{W}_2)^T}{\sqrt{d}} + (\mathbf{V}\mathbf{W}_3)\mathbf{g}_i^T \right) \in (0, 1)^{N \times M} \quad (16)$$

该公式由两部分组成,第一部分 $\frac{(\mathbf{V}\mathbf{W}_1)(\mathbf{L}\mathbf{W}_2)^T}{\sqrt{d}}$ 是标准的跨模态注意力分数,衡量视觉与文本 Token 间的固有相关性;第二部分 $(\mathbf{V}\mathbf{W}_3)\mathbf{g}_i^T$ 则引入了知识引导查询向量 $\mathbf{g}_i$ 的偏置调制,实现知识对交互关系的动态注入。通过 Sigmoid 激活,矩阵 $\mathbf{G}$ 中的每个元素 G_{ij} 均表示第 i 个视觉 Token 与第 j 个文本 Token 间的交互强度权重。矩阵门控能够精细地抑制噪声特征间的无效关联,同时强化关键语义特征间的关联,显著提升融合精度。星载终端默认采用轻量的 Token 级门控以保证推理速度,地面服务器或算力充足的边缘节点则启用交互矩阵门控,以换取更优的融合性能。

3.2.3 知识约束融合机制

基于生成的门控权重,构建知识约束的门控跨模态注意力融合机制。该机制以视觉特征为主体,契合遥感任务中图像信息为核心的特点,将文本指令与知识图谱作为语义约束条件,通过门控注意力实现多模态特征的深度融合,最终输出增强后的视觉特征,用于下游的任务预测(如目标检测、场景分类)。对视觉、文本与知识特征进行线性投影,生成查询、键、值向量。为简化计算并突出视觉主体,将文本与知识特征拼接作为条件信息,构建统一的上下文键值对 $[\mathbf{L}; \mathbf{K}]$ 。标准跨模态注意力的计算为

$$\begin{aligned} \mathbf{A} &= \text{softmax} \left(\frac{(\mathbf{V}\mathbf{W}_q)([\mathbf{L}; \mathbf{K}]\mathbf{W}_k)^T}{\sqrt{d}} \right) \\ \mathbf{O} &= \mathbf{A}([\mathbf{L}; \mathbf{K}]\mathbf{W}_v) \end{aligned} \quad (17)$$

其中: W_Q, W_K, W_V 为可训练的投影矩阵; A 为注意力权重矩阵, 表征视觉特征对上下文信息的关注程度; O 为融合得到的上下文语义特征。

结合两种门控策略, 设计两种门控注入方案, 适配不同星载算力约束。首先是 Token 门控。直接将 Token 级门控向量 m 作用于融合后的上下文特征 O , 对每个视觉 Token 的融合结果进行加权增强。增强后的视觉特征 $V'_i = V_i + m_i \times O_i$, 通过门控值 m_i 动态调节每个视觉 Token 对上下文信息的吸收量, 对高相关性 Token 增强其语义表达, 对噪声 Token 则抑制其融合贡献, 实现轻量化的知识约束。其次是矩阵门控。针对交互矩阵门控场景, 将门控矩阵 G 作为注意力权重的约束因子, 进行逐元素加权。由于知识子图实体数量 P 与文本 Token 数量 M 可能不一致, 需对门控矩阵进行维度扩展或采用关系权重 $a_r(t)$ 进行适配。门控注意力权重的计算为

$$A' = \text{Normalize}(A \odot G) \quad (18)$$

其中: $\odot$ 为哈达玛积 (Hadamard Product), 实现逐元素的精准约束; Normalize 为行归一化操作, 确保调整后的注意力权重仍为合法的概率分布。最终, 利用门控注意力权重 A' 进行特征融合, 得到增强视觉特征:

$$V' = V + A'([L; K]W_V) \quad (19)$$

该方式通过全局交互矩阵的约束, 从整体上优化了跨模态特征的关联模式, 有效屏蔽了噪声干扰, 确保特征融合完全遵循任务知识的引导。经过知识约束融合得到的增强视觉特征 V' , 融合了视觉表象、任务语义与知识结构的多层信息, 具备更强的抗干扰性与语义表征能力。

将 V' 输入至特定的任务头 (分割任务采用轻量化 U-Net 头, 检测任务采用 YOLOv8-tiny 头), 即可完成下游遥感任务的预测, 输出最终的推理结果。

3.3 面向卫星边缘环境下的模块化可裁剪推理

卫星边缘计算场景具有算力资源极度受限 (低轨卫星单节点算力通常不超过 10 TOPS)、任务负载动态波动、应用场景需求异构的核心特征, 存储资源需控制在百兆级规模, 无法支撑完整大规模多模态模型的在轨推理。传统多模态融合模型采用固定化网络结构, 不具备根据算力资源与任务需求动态调整的能力, 难以适配卫星边缘环境的实际工程要求。为此, 本文提出面向卫星边缘环境的模块化可裁剪推理结构, 将端到端推理流程解耦为功能独立、可插拔、可配置的子模块, 结合算力感知的动态模块选择机制, 实现推理性能与计算开销的动态平衡, 适配星载端动态资源约束与异构任务需求。模块化可裁剪结构采用“离线标定+在线动态选择”的混合机制。离线阶段对各模块不同配置下的 FLOPs、显存占用和推理时

延进行标定, 形成可调用的模块配置库, 在线阶段由轻量化控制器依据实时预算和任务等级选择执行路径, 因此预编译的是模块库, 运行时动态变化的是模块组合、门控粒度与端云切分点。

3.3.1 模块划分策略

模块划分遵循功能独立、低耦合、高内聚、可裁剪的原则, 结合卫星遥感多模态推理的任务流程, 将整体推理结构拆分为一组标准化的可插拔模块集合, 各模块承担单一核心功能, 且模块间仅通过标准化的特征接口进行交互, 既支持单模块内部的精细化配置, 也支持多模块的灵活组合, 适配不同任务场景与算力条件。具体模块划分及功能定义如下:

(1) M0 基础感知模块。作为推理结构的基础模块, 包含轻量化视觉编码器与轻量化文本编码器, 核心功能为对遥感图像与任务文本指令进行基础特征提取, 输出统一维度的视觉 Token 与文本 Token 特征序列, 是所有推理任务的必选模块, 其内部可通过调整编码器的层数、通道数、Token 分辨率实现轻量化配置。

(2) M1 知识子图构建模块。承接 3.1 节的核心逻辑, 包含 Top-K 种子实体检索、H-hop 关系感知扩展及可选的个性化 PPR 子图压缩功能, 核心功能为根据任务语义从全局知识图谱中抽取动态知识子图, 输出任务相关的知识实体嵌入特征, 支持通过调整 Top-K、H-hop、Top-M 等超参数控制子图规模与计算开销。

(3) M2 知识约束门控融合模块。承接 3.2 节的核心逻辑, 包含 Token 级门控与矩阵级门控两种融合模式, 核心功能为实现视觉、文本、知识多模态特征的知识约束融合, 输出增强后的视觉特征序列, 支持两种门控模式的二选一, 且可通过调整门控层的维度、激活方式实现轻量化。

(4) M3 结构化推理模块。基于图神经网络 R-GCN 实现知识特征的结构化推理, 核心功能为挖掘知识子图中实体间的深层语义关联, 强化知识特征对多模态推理的引导作用, 其内部可通过调整 R-GCN 的层数、隐藏层维度、邻居采样数实现裁剪, 是对知识推理精度有要求任务的可选模块。

(5) M4 因果/关系一致性模块。针对洪水、地震等灾害场景的强语义约束需求设计, 核心功能为基于知识图谱的因果关系规则, 对多模态推理结果进行一致性校验, 剔除不符合因果逻辑的错误预测。

(6) M5 时空一致性模块。针对多时相遥感监测任务设计, 核心功能为利用时序遥感数据的空间相关性与时间连续性, 对单帧推理结果进行时空平滑与补全, 提升动态监测任务的推理稳定性。

3.3.2 计算预算建模

计算预算建模的核心目标是对各模块不同配置

下的计算、存储与时延代价进行量化表征,构建统一的代价评估体系,为动态模块选择提供可量化的算力约束依据。该建模过程基于离线标定+在线适配的思路,先通过离线实验测得各模块不同配置下的硬件代价指标,再通过加权融合构建统一的代价函数,最后结合星载端实时算力状态设定全局预算约束。对每个模块的任意配置,通过星载硬件平台的离线实测与数值估计,得到三个核心代价指标。计算代价 $FLOP_m(\mathbf{w}_m)$ 模块在该配置下完成一次推理所需的浮点运算数,表征计算量大小。存储代价 $Mem_m(\mathbf{w}_m)$ 模块在该配置下推理过程中所需的实时存储容量(含参数存储与特征存储),表征对星载存储资源的占用。

时延代价 $Latency_m(\mathbf{w}_m)$ 模块在该配置下完成一次推理的实际耗时,直接表征对实时性的影响。由于计算、存储、时延三个代价指标的量纲不同,且不同星载硬件平台对各指标的敏感程度存在差异(如部分平台算力紧缺,部分平台存储紧缺),引入可由星载平台离线标定的权重系数 $\alpha + \beta + \gamma = 1$,对三个指标进行加权融合,构建模块统一代价函数:

$$C_m(\mathbf{w}_m) = \alpha \cdot FLOP_m(\mathbf{w}_m) + \beta \cdot Mem_m(\mathbf{w}_m) + \gamma \cdot Latency_m(\mathbf{w}_m) \quad (20)$$

其中,权重系数根据星载平台的资源瓶颈进行标定:若平台算力受限,增大 α ;若存储受限,增大 β ;若对实时性要求极高,增大 γ 。该函数将多维度的硬件代价转化为单一标量值,实现不同模块、不同配置代价的统一量化比较。任务推理的模块配置向量为 $\mathbf{w} = \{\mathbf{w}_m\}_{m \in \mathbb{N}}$,则整体推理结构的总计算代价为各模块代价之和 $\sum_{m \in \mathbb{N}} C_m(\mathbf{w}_m)$,令 B 为星载端当前可用算力预算,该预算并非固定值,而是由星载平台的实时状态动态决定,包括电源剩余电量、设备温控状态、任务优先级等:电源电量低时降低 B ,高优先级任务时适当提高 B ,温控超标时限制 B 。为保证模型在星载端的可部署性,满足全局算力预算约束 $\sum_{m \in \mathbb{N}} C_m(\mathbf{w}_m) \leq B$,该约束为后续动态模块选择设定了明确的算力边界,确保选择的模块组合及配置不超出星载端实时资源承载能力。

3.3.3 动态模块选择机制

为在算力预算约束下,自适应选择最优的模块配置向量 $\mathbf{w}$,实现推理精度与计算开销的折中,本文引入轻量化控制器(Controllor)作为决策核心。该控制器以任务语义、输入数据质量、链路状态与实时算力预算为输入,通过端到端的训练学习动态选择策略,自动输出各模块的启用概率与配置分布,无需人工干预,适配卫星边缘环境的动态变化特性。控制器的决

策依据为融合多维度信息的状态特征向量 $\mathbf{s} = [\mathbf{z}_t; q_{img}; R_{link}; B]$,该向量整合了任务、数据、链路、资源四大核心维度的信息,全面表征当前推理任务的整体场景,其中 $\mathbf{z}_t$ 为 3.1.1 节中标准化后的任务语义向量,表征任务的核心语义需求与精度要求, q_{img} 为遥感图像的质量指标向量,包含云遮挡比例、图像对比度、目标清晰度等量化指标,表征输入数据的质量与推理难度。 R_{link} 为当前星地链路的速率估计值,表征星地协同的通信条件, B 为当前星载端的可用算力预算,表征资源约束边界。状态特征向量通过标准化处理后输入控制器,确保各维度信息的量纲统一,为控制器的精准决策提供基础。控制器采用轻量化的多层感知机(MLP)实现,避免因控制器自身计算开销过大影响星载推理效率,其核心为参数化的决策函数 $f_\theta(\cdot)$ (θ 为控制器的可训练参数)。控制器以状态特征向量 $\mathbf{s}$ 为输入,通过 Softmax 激活函数输出各模块不同配置的概率分布:

$$\pi_\theta(\mathbf{w}_m | \mathbf{s}) = \text{softmax}(f_\theta(\mathbf{s})) \quad (21)$$

其中, $\pi_\theta(\mathbf{w}_m | \mathbf{s})$ 表示在当前状态特征 $\mathbf{s}$ 下,选择模块 m 的配置 $\mathbf{w}_m$ 的条件概率。该概率分布既表征了各模块配置的选择优先级,也为后续的随机采样与端到端训练提供了概率基础。控制器的训练目标为在算力预算约束下,实现推理任务损失最小化与计算代价不超限的双重目标,通过构建联合损失函数,实现控制器与推理模块的端到端协同训练。具体训练目标为最小化以下期望损失:

$$\min_{\theta, \text{Net}} E_{\omega \sim \pi_\theta} [L_{\text{task}}(\omega) + \lambda \cdot \max(0, \sum C_m(\omega_m) - B_m)] \quad (22)$$

其中, $L_{\text{task}}(\omega)$ 为推理任务损失(目标检测的交并比损失、场景分类的交叉熵损失),由模块配置向量 ω 决定,表征推理精度。第二项为预算违约惩罚项, λ 为惩罚系数,当总计算代价超出算力预算时,该项输出正的惩罚值,且超出越多惩罚越大,当总计算代价未超限时,该项为 0,确保模型严格遵守算力约束。 $E_{\omega \sim \pi_\theta}$ 表示对基于控制器概率分布 π_θ 采样的所有模块配置向量 ω 求期望,保证训练的鲁棒性。由于模块配置为离散值,无法直接通过梯度下降进行优化,本文采用 Gumbel-Softmax 重参数化技巧对离散配置进行连续近似,将离散的随机采样过程转化为可微的计算过程,实现控制器与推理模块的端到端梯度回传与联合优化。在星载端的实际推理阶段,控制器无需进行概率采样,直接根据状态特征向量 $\mathbf{s}$ 输出概率最大的模块配置向量 ω 作为当前任务的最优推理结构,该过程计算开销极低,可在星载端实时完成,满足卫星边缘计算的实时性要求。

3.4 端云协同优化策略

卫星边缘环境存在算力稀缺、星地链路带宽有限且时变的问题,单一星载轻量化推理难以满足高复杂度遥感任务的精度要求,纯云端推理则因星地通信时延过高而损失实时性并增加星载发射能耗。为此,本文提出端云协同优化策略,基于模块依赖关系进行任务合理分解,通过星载端与云端的算力协同、数据交互,在遵循星载算力预算约束的前提下,联合优化推理精度、时延与能耗三大目标,实现星载与云端资源的优势互补。该策略通过任务分解、通信开销建模、两阶段联合优化三步实现,完成星载端与云端的推理流程协同。端云协同是性能增强机制,而非星端推理的前提条件。星端保留轻量化本地模型、压缩索引和任务子图生成能力,当链路波动或中断时,系统立即回退到仅星端执行的轻量链路,待链路恢复后再重新评估是否触发协同推理。决策权划分遵循星端实时决策为主、地面离线更新为辅的原则。

3.4.1 任务分解策略

端云协同的前提是对整体推理任务进行合理分解与卸载,而KM³-SatNet的推理模块遵循严格的功能依赖顺序(基础感知→知识子图构建→多模态融合→结构化推理→任务头预测),前序模块的输出为后序模块的输入,若随意卸载模块会导致中间特征传输断裂、推理流程失效。为此,本文提出基于模块依赖的切分点任务分解策略,核心原则为仅在推理流程的关键特征输出节点设置可切分点,将推理任务划分为星载前段子任务与云端后段子任务,保证星地协同推理的流程连续性与特征可传输性。

$$y_m = \begin{cases} 0, & \text{星载执行} \\ 1, & \text{云端执行} \end{cases} \quad (23)$$

卸载变量的取值并非独立,而是由选定的可切分点唯一确定。若选择切分点 p ,则切分点之前的所有模块卸载变量为0(星载执行),切分点之后的所有模块卸载变量为1(云端执行)。通过该切分点-卸载变量的映射关系,将离散的模块卸载决策转化为对可切分点的选择,大幅降低了决策复杂度,同时从根本上避免了因模块乱序卸载导致的推理流程失效问题。

3.4.2 通信开销建模

星地之间的中间特征传输是端云协同的核心开销来源,其传输时延与发射能耗直接决定了端云协同的实际效果,若通信开销过高,会抵消云端算力带来的性能提升。为此,本文针对可切分点的特征传输过程,构建星地通信开销量化模型,分别对传输时延与发射能耗进行精准建模,并引入特征压缩策略降低传输代价,适配星地链路的低带宽特性。

对于选定的可切分点 p ,设其需上传的中间特征

为 F_p ,特征的二进制存储大小为 $|F_p|$ (单位:bit),当前星地链路的实际传输速率为 v_R (单位:bps,由卫星轨道、大气环境动态决定),星地通信的固定往返时延为 T_{rtt} (单位:s,包含信号传播、地面站处理等固定耗时)。则该切分点下的星地通信时延为

$$T_{\text{comm}}(p) = \frac{|F_p|}{v_R} + T_{rtt} \quad (24)$$

该模型表明,通信时延由可变传输时延(与特征大小正相关、与链路速率负相关)和固定往返时延两部分组成,其中可变传输时延是端云协同优化的主要调控对象。卫星的星上能源由太阳能电池与储能电池提供,星地通信的射频发射能耗是星载能耗的重要组成部分,需进行严格管控。针对切分点 p 的特征传输过程,设卫星射频发射功率为 P_{tx} ,则该切分点下的星地传输发射能耗为

$$E_{tx}(p) = P_{tx} \cdot \frac{|F_p|}{v_R} \quad (25)$$

发射能耗与特征传输时间正相关,降低特征传输量或提升链路速率均可有效减少发射能耗,契合卫星的低能耗设计要求。为进一步降低星地通信的时延与能耗代价,在特征传输前引入轻量特征压缩策略,通过量化、稀疏化等无监督/轻监督压缩方法,在尽量保留特征语义信息的前提下,对中间特征 F_p 压缩处理,压缩后特征大小为

$$|F_p| \leftarrow \rho \cdot |F_p|, 0 < \rho \leq 1 \quad (26)$$

其中, ρ 为压缩比, ρ 越小表示压缩程度越高,传输代价越低,但可能带来一定的特征信息损失。工程实现中,可根据任务对精度的要求动态调整压缩比:对精度要求高的任务采用高压缩比(ρ 接近1),对实时性要求高的任务采用低压缩比(ρ 偏小),实现特征传输代价与推理精度的平衡。星载端采用轻量化压缩算法,避免压缩过程引入额外的计算开销,云端则对应采用解压算法恢复特征,保证推理流程的连续性。

3.4.3 两阶段启发式联合优化机制

端云协同的优化目标为在满足星载算力预算、切分点选择规则的前提下,实现推理精度最大化、端云总时延最小化、星载总能耗最小化。由于优化变量包含离散的切分点选择、模块配置向量与连续的特征压缩比,属于混合整数优化问题,传统凸优化方法难以求解。为此,提出两阶段启发式联合优化机制,将复杂多目标优化分解为两个串行单目标优化子问题,在保证优化效果的同时降低计算复杂度,适配星载端实时决策需求。首先构建端云协同的多目标优化函数,将精度、时延、能耗三大目标统一为一个全局最小化目标,具体形式为

$$\min L_{\text{task}}(\omega) + \mu \cdot T(\omega, y) + v \cdot E(\omega, y) \quad (27)$$

其中, $L_{\text{task}}(\omega)$ 为推理任务损失, 表征推理精度, 损失越小精度越高, $T(\omega, y) = \sum_m [(1 - y_m) T_m^{\text{edge}}(\omega_m) + y_m T_m^{\text{cloud}}(\omega_m)] + T_{\text{comm}}(p)$ 为端云总时延, 包含星载模块推理时延、云端模块推理时延与星地通信时延。
 $-E(\omega, y) = \sum_m [(1 - y_m) E_m^{\text{edge}}(\omega_m) + y_m E_m^{\text{tx}}(\omega_m)]$ 为星载总能耗, 包含星载模块推理能耗与星地特征传输的发射能耗。

针对上述混合整数优化问题, 本文采用两阶段启发式算法进行求解, 将模块配置优化与切分点选择进行解耦, 分阶段完成全局优化任务。

第一阶段: 星载算力约束下的模块配置优化。在星载实时算力预算 B 的严格约束下, 控制器综合任务语义向量、图像质量指标、星地链路状态及实时资源剩余等多维环境特征, 自动完成对所有推理模块的最优裁剪配置决策。通过端到端训练的轻量化控制器, 能够在毫秒级时间内快速生成满足算力约束的最优模块配置向量 ω^* 。该阶段的优化目标在于星载推理精度与计算代价之间的动态平衡, 确保模型在资源受限的星载边缘环境下仍能保持高效的推理响应速度。

第二阶段: 切分点的时延最优选择。在获取满足算力约束的最优模块配置向量 ω^* 后, 本阶段将优化目标聚焦于端云总时延的最小化。时延是提高端云协同推理效率的核心瓶颈, 因此通过比较不同卸载策略下的总时延, 选择最优实施方案。首先, 针对每一

个可选切分点 $p \in P$, 分别计算两种不同部署模式下的总推理时延。纯星载推理时延为 $T_{\text{edge_all}}(\omega^*)$, 该模式表示所有推理模块均在星载端执行, 无需进行星地数据传输, 总时延即为星载端所有模块的推理耗时之和。端云协同推理时延为 $T_{\text{edge_part}}(\omega^*, p) + T_{\text{comm}}(p) + T_{\text{cloud}}(p)$, 该模式包含三部分时延, $T_{\text{edge_part}}(\omega^*, p)$ 为切分点 p 之前星载模块的推理时延, $T_{\text{comm}}(p)$ 为星载与云端之间的中间特征传输通信时延, $T_{\text{cloud}}(p)$ 为切分点 p 之后云端模块的推理时延。随后, 遍历所有可切分点, 计算对应的端云协同推理时延, 并与纯星载推理时延进行对比。最终选择总时延最小的部署方案作为最优决策。若纯星载推理时延最小, 则选择不进行任务卸载, 所有模块全部在星载端执行。若某一切分点对应的端云协同推理时延最小, 则将该切分点确定为最优切分点 P^* , 并根据切分点与卸载变量的映射关系, 确定所有模块的最优卸载决策 y_m^* 。

4 实验

4.1 数据集和评价指标

本文采用公开数据集+自建数据集的组合方式, 构建两大实验数据集, 覆盖卫星遥感典型任务场景, 所有遥感图像均为 512×512 分辨率, 包含光学图像与 SAR 图像 (Sentinel-1/2 卫星数据), 并配套自然语言任务指令与人工标注标签, 具体信息如表 1 所示, 自建灾害场景识别数据集统计信息如表 2 所示。

表 1 实验数据集详细信息

Table 1 Details of the experimental datasets

数据集名称	数据类型	样本规模	任务类型	标注方式
多模态遥感目标识别数据集	光学+SAR+任务文本	10 万张图像、5 万条指令	目标检测(建筑、道路、水体、船舶)	边界框标注 (mAP 评价)
灾害场景识别数据集 (自建)	光学+SAR+任务文本	5 万张图像、2 万条指令	二分类(正常场景/灾害场景)、洪水精细分割	像素级标注 (IoU/Dice 评价)

表 2 自建灾害场景识别数据集统计信息

Table 2 Statistics of the self-built disaster scene recognition dataset

统计项	数值	说明
灾害事件数	1 248	按事件级去重后的独立灾害事件数量
总图像数	50 000	通过光学/SAR 配对与样本清洗后保留的有效图像样本
文本指令数	20 000	由检测/分类/分割模板生成并经人工抽检校验的指令
分类样本分布	正常场景 26 800、灾害场景 23 200	场景级二分类样本统计
分割样本分布	小面积洪水 7 400、大面积洪水 6 100、河道扩张易混淆 5 500、无有效淹没 3 200	参与像素级掩膜标注的主要样本类型
标注一致率	分类 96.8%、边界框 91.2%、分割掩膜 Dice 93.4%	双人交叉标注后的一致率与专家复核结果
训练/验证/测试事件数	874/249/125	按 7:2:1 划分, 并结合地区与季节分层抽样

自建灾害场景识别数据集基于 Sentinel-1/2 卫星的实际遥感数据构建, 覆盖我国长江流域、黄河流域

等洪水高发区域, 包含正常水体、河道扩张、小面积洪水、大面积洪水等典型场景, 同时加入不同程度的

云遮挡、光照变化、SAR 相干斑噪声等干扰,贴近卫星遥感的实际数据特征,为鲁棒性实验提供数据支撑。所有数据集按 7:2:1 的比例划分为训练集、验证集与测试集,训练过程采用数据增强(随机裁剪、翻转、亮度调整、噪声添加)提升模型的泛化能力。针对不同的遥感任务,选取精度、模型规模、推理性能、鲁棒性四大类核心评价指标,全面量化模型的性能。目标识别任务指标为平均精度均值(mAP@0.5),表示 IoU 阈值为 0.5 时所有类别的平均检测精度;灾害场景识别任务指标为分类准确率(Acc)、洪水分割交并比(IoU)、Dice 相似系数(Dice),均为越高越好;模型规模指标为参数量(Params)、浮点运算数(FLOPs),均为越低越好;推理性能指标为推理时延(Latency,单

位:ms)、星载端能耗(Energy,单位:J),均为越低越好;鲁棒性指标为不同干扰条件下(云遮挡、噪声、光照变化)的精度衰减率(Accuracy Decay Rate,ADR),ADR 越低表示鲁棒性越强, $ADR = (\text{精度无干扰} - \text{精度有干扰}) / \text{精度无干扰} \times 100\%$ 。自建灾害场景识别数据集以 Sentinel-1/2 时空配对样本为基础,结合灾害事件清单和区域水文信息构建候选事件池,随后按照“光学/SAR 配对-任务文本生成-双人交叉标注-专家复核-质量控制-事件级划分”的流程形成最终数据集。标注阶段采用“双人独立标注+冲突复判+专家抽检”的策略,像素级分割样本的一致率达到 93.4%,并通过地区/季节分层划分避免空间泄漏。自建灾害场景数据集构建与质量控制流程见图 4。



注:该流程展示了数据来源、模态配对、任务文本生成、双人交叉标注、专家复核与事件级划分等关键环节。

图4 自建灾害场景数据集构建与质量控制流程

Figure 4 Construction and quality control process of the self-built disaster scene dataset

如图 4 所示,自建灾害场景识别数据集采用“多源获取-模态配对-文本生成-标注质检-事件级划分”的流程构建。首先,在数据来源方面,综合利用 Sentinel-1/2 遥感影像、灾害事件清单以及区域水文与地形信息,获取覆盖典型灾害场景的多源样本。其次,在模态配对阶段,按照统一的观测时间窗和地理范围对光学与 SAR 样本进行时空对齐,并剔除严重云污染、几何失配或质量异常的样本,以保证多模态数据的一致性与可用性。在任务文本构建方面,依据目标检测、场景分类和洪水分割等不同任务模板自动生成文本指令,并通过人工抽检与语义一致性校核,确保文本描述与样本内容匹配。标注阶段采用“双人交叉标注+专家复核”的质量控制机制,其中边界框、场景标签和像素级掩膜均由两名标注员独立完成,对存在分

歧的样本交由领域专家进行复判。最后,在数据划分上采用事件级训练集、验证集和测试集划分策略,按照 7:2:1 的比例,并结合地区和季节进行分层抽样,以避免同一灾害事件泄漏到不同子集,从而提高实验评估的真实性与可靠性。

4.2 卫星遥感任务知识图谱

卫星遥感任务知识图谱(Satellite Remote Sensing task Knowledge Graph, SRS-KG)为面向空天地一体化遥感任务的领域专用知识图谱,整体包含实体类型 12 类、实体总数 1 863 个、关系类型 8 类、有效三元组 7 528 条,核心覆盖洪水监测、建筑/船舶/道路目标识别、植被监测、地震灾害评估四大典型卫星遥感任务场景,典型遥感场景的知识域扩展示例如表 3 所示。

表3 典型遥感场景的知识域扩展示例

Table 3 Examples of knowledge-domain expansion for typical remote sensing scenarios

场景类型	核心实体	关键关系	迁移注意事项
洪水监测	洪水、永久水体、低洼区、降雨事件	caused_by、has_attr、related_to	重点约束异常水体与正常水位变化的区分
建筑/道路识别	建筑、道路、桥梁、阴影区域	adjacent_to、support_by、occluded_by	需控制高密度城市纹理导致的误检扩散
植被监测	植被覆盖、旱情、土壤湿度、作物类型	correlated_with、affected_by	需引入季节性和地区差异,避免静态规则失效
地震灾害评估	建筑损毁、废墟、道路阻断、次生灾害	caused_by、part_of、related_to	应强化事件级因果链与灾后时序变化约束

构建数据源采用权威领域知识库+标注遥感数据集+行业标准规范多源融合方式,具体包括:(1)遥感领域知识库(如 Remote Sensing Knowledge Base、地理空间知识图谱 GeoKG);(2)公开遥感标注数据集(Sentinel-1/2 标注集、高分卫星数据集、xView2 灾害数据集^[46-47]);(3)行业标准(《卫星遥感洪涝灾害监测技术规范》^[45]、《航天遥感图像解译规范》^[48]);(4)领域核心文献(近 5 年遥感多模态融合、灾害监测相关顶刊/顶会论文);(5)自建灾害场景标注数据集集中的语义关联信息。SRS-KG 并非仅服务于洪水场景,还

覆盖建筑/船舶/道路目标识别、植被监测和地震灾害评估等典型任务。对其他任务,读者可按照任务本体定义-实体/关系 schema 设计-多源知识采集-规则与专家校核-任务驱动子图抽取的流程扩展知识图谱。实体抽取采用规则+深度学习联合抽取方法。首先基于遥感领域关键词表、正则表达式完成规则化粗抽取,再采用轻量命名实体识别模型对非结构化文本(任务指令、文献、行业规范)进行精细化抽取,最后经领域专家人工校验去重,最终确定 12 类核心实体,各类别实体数量与定义如表 4 所示。

表 4 SRS-KG 实体类型统计信息
Table 4 Statistics of SRS-KG entity types

实体类型	数量/个	核心定义	典型示例
遥感目标	426	卫星遥感任务中需识别/分割的核心地理要素	洪水、建筑、船舶、道路、植被
场景类型	158	遥感任务对应的地理场景/灾害场景	城市洪涝、河道扩张、地震废墟
数据类型	63	遥感任务涉及的多模态数据类型	光学图像、SAR 图像、时序数据
地理属性	327	遥感目标/场景的空间/地理特征	低洼区、永久水体、高坡度区
语义属性	289	遥感目标/场景的语义特征/状态特征	异常扩张、淹没、正常水位
任务类型	45	卫星端核心智能计算任务	精细分割、目标检测、场景分类
推理规则	56	遥感任务的领域先验推理逻辑	水位超阈值判定异常、低洼区易积水
噪声类型	39	遥感数据中的典型干扰噪声	云遮挡、SAR 相干斑、光照变化
传感器类型	78	采集遥感数据的卫星传感器	Sentinel-1 SAR、高分二号光学传感器
地理区域	198	典型遥感监测地理区域/灾害高发区域	长江流域、黄河流域、华北平原
处理方法	126	遥感数据/任务的基础处理方法	云去除、特征融合、语义分割
评价指标	58	遥感任务推理结果的核心评价指标	IoU、mAP、分类准确率

SRS-KG 的关系为实体间的语义关联与逻辑关联,基于遥感任务推理逻辑划分为 8 类核心关系,所有关系均经领域专家定义语义约束规则,确保三元组

的逻辑合理性,关系类型、定义及三元组数量如表 5 所示,且所有关系均为有向关系,明确头实体与尾实体的语义指向。

表 5 SRS-KG 关系类型统计信息
Table 5 Statistics of SRS-KG relation types

关系类型	三元组数量/条	核心定义	典型三元组示例
属于(belong_to)	892	子实体归属于父实体的类别关联	洪水→属于→遥感目标;城市洪涝→属于→场景类型
具有属性(has_attr)	1 658	实体具有的地理/语义属性关联	低洼区→具有属性→易积水;洪水→具有属性→异常水体
因果关系(caused_by)	986	实体间的因果逻辑关联(结果→因果关系→原因)	城市洪涝→caused_by→强降雨;水体异常→caused_by→河道扩张
相关关联(related_to)	1 257	实体间的非因果语义关联,相互关联但无明确因果	SAR 图像→related_to→洪水监测;云遮挡→related_to→光学图像
数据支撑(data_support)	763	数据类型为某任务/目标/场景提供数据支撑	光学图像→data_support→建筑检测;时序数据→data_support→水位变化监测
适用方法(apply_method)	685	某任务/目标适用的遥感数据处理方法	洪水分割→apply_method→语义分割;云遮挡→apply_method→云去除
存在噪声(have_noise)	526	某数据类型/场景存在的典型干扰噪声	光学图像→have_noise→光照变化;SAR 图像→have_noise→相干斑
监测区域(monitor_area)	771	某任务/目标的典型监测地理区域	洪水监测→monitor_area→长江流域;植被监测→monitor_area→华北平原

搭建星载端模拟平台与云端平台,模拟空天地一体化的实际计算环境,硬件参数如表 6 所示。采用 AdamW 优化器,学习率初始值为 1×10^{-4} ,学习率衰减策略为余弦退火,批次大小 (Batch Size) 星载端为 8,云端为 32,训练轮数为 100,权重衰减为 1×10^{-5} ,早停策略 (Patience=10) 防止过拟合。所有模型均采用 8 位量化进行轻量化处理,星载端推理为单张图像推理,云端推理为批量推理 (Batch Size=16);端云协同的星地链路速率按 1/5/10 Mbps 分档测试,特征压缩比按 0.2/0.5/0.8 动态调整。温度系数 τ 的取值范围为 $(0, 1]$,目标识别任务最优值为 0.3,灾害场景识别(洪水分割)任务最优值为 0.2,设置依据为遥感任务语义区分度要求高, τ 可增强语义匹配精准性,预实验显示 $\tau > 0.2$ 时匹配鲁棒性下降, $\tau > 0.4$ 时精准性降低;知识子图扩展步长 H 为正整数,目标识别任务最优值为 2,洪水分割任务最优值为 3,因目标识别知识关联简单,2-hop 即可覆盖核心知识,而洪水分割涉及多维度因果关联,需 3-hop 实现完整知识覆盖,且 $H \geq 4$ 时子图规模会超出星载端存储约束;PPR 重启概率 η 的取值范围为 $(0, 1)$,目标识别任务最优值为 0.2,洪水分割任务最优值为 0.3,重启概率越大,子图越聚焦种子实体,洪水分割需兼顾种子实体与关联属性实体,故 η 略大,预实验显示 η 在 0.2~0.3 范围内时子图性价比最高;惩罚系数 λ 的取值范围为 1~10,目标识别任务最优值为 5,洪水分割任务最优值为 6,主要用于保证算力预算约束的严格性,灾害任务对精度要求更高。因此, λ 略大以避免超预算,且实验表明 $\lambda > 8$ 时模型易过度裁剪导致精度

骤降;计算代价权重 α 的取值范围为 0~1,目标识别任务最优值为 0.4,洪水分割任务最优值为 0.35,目标识别对推理速度要求更高,故 α 略大,且需满足 $\alpha + \beta + \gamma = 1$ 的约束,权重分配遵循星载端“算力>存储>时延”的资源敏感程度;存储代价权重 β 的取值范围为 0~1,两大任务最优值均为 0.35,因两类任务对存储约束均为百兆级,该取值可适配星载端存储 ≤ 512 MB 的硬约束;时延代价权重 γ 的取值范围为 0~1,目标识别任务最优值为 0.25,洪水分割任务最优值为 0.3,洪水分割作为应急任务,对时延实时性要求更高,故 γ 略大,可满足星载端推理时延 ≤ 100 ms 的工程要求;种子实体数量 K 为正整数,目标识别任务最优值为 5,洪水分割任务最优值为 8,目标识别核心种子实体较少,5 个即可覆盖核心需求,而洪水分割涉及洪水、永久水体、低洼区等多个核心实体,需 8 个种子实体才能保证知识覆盖的完整性;子图边保留数 M 为正整数,目标识别任务最优值为 20,洪水分割任务最优值为 30,该取值与扩展步长相匹配,洪水分割子图关联更复杂,需保留更多边,且所有取值均能保证子图存储 ≤ 10 MB,适配星载端存储约束;子图节点保留数 L 为正整数,目标识别任务最优值为 15,洪水分割任务最优值为 25,结合实验可知,该取值在保证知识完整性的前提下,可使子图推理计算开销 ≤ 0.05 TOPS,符合星载端算力约束要求。默认视觉编码器采用 4-stage 轻量骨干,文本编码器采用 DistilBERT-base 的前 6 层并将输出统一投影到 256 维共享空间,知识子图编码器使用 2 层。部署测试中统一采用 8-bit 量化。

表 6 实验硬件平台参数
Table 6 Experimental hardware platform parameters

平台类型	硬件配置	算力/性能指标	适用场景
星载端模拟平台	ARM Cortex-A72 4 核、4 GB LPDDR4、Mali-G52 GPU	算力 0.5 TOPS、存储 ≤ 512 MB、功耗 5 W	星载端推理、模块裁剪、端云协同决策
云端平台	Intel i9-13900K、64 GB DDR5、NVIDIA RTX4090	算力 1 000+ TOPS、存储 1 TB、带宽 10 Gbps	模型训练、云端推理、端云协同验证

4.3 对比模型

选取当前卫星遥感、多模态融合、知识增强及近期代表性方法共 8 种模型作为对比基线,并按输入模态、模型规模与部署目标分为三类:(1)轻量星端模型:DeepLabV3+ (MobileNetV2)^[5]、YOLOv8-nano^[49];(2)多模态融合/遥感大模型:Swin-L+CrossAttention^[50]、RS-MLLM-Lite、EarthGPT-X^[7]、LAMF-Ne^[51];(3)知识增强模型:ViT-B + KGAT^[9]。其中,EarthGPT-X^[7] 为 2025 年代表性遥感空间 MLLM,LAMF-Net^[51] 为 2026 年轻量自适应多模态融合方法。为保证比较公平性,主比较对象优先选择输入模态、部署目标和资源约束条件与本文接近的方法;对于参

数规模显著更大、主要服务于地面高算力环境的方法,仅作为参考比较,不将其作为唯一主结论依据。

4.4 主实验

多模态遥感目标识别任务的对比实验结果如表 7 所示,本文提出的 KM³-SatNet 在精度、模型规模、推理性能上均取得最优综合表现。KM³-SatNet 的 mAP@0.5 达到 92.6%,相较于最优对比模型 RS-MLLM-Lite 提升 6.3%,相较于轻量化模型 YOLOv8-nano 提升 10.2%,验证了知识增强与多模态门控融合的有效性。KM³-SatNet 参数量仅为 7.8×10^6 ,FLOPs 为 1.2×10^9 ,相较于大规模模型 Swin-L+CrossAttention (参数量 5.80×10^8 ,FLOPs 2.56×10^{11}),参数量减少 98.7%,FLOPs 减少

99.5%。KM³-SatNet 星载端推理时延为 9.2 ms, 能耗为 0.046 J, 相较于 RS-MLLM-Lite (时延 28.5 ms, 能耗 0.142 J), 时延降低 67.7%, 能耗降低 67.6%。灾害场景识别任务的对比实验结果如表 8 所示, 包含分类准确率 (Acc)、洪水分割 IoU/Dice、推理性能三大维度, 本文方法同样表现最优。KM³-SatNet 灾害场景分类 Acc 达到 94.5%, 相较于 RS-MLLM-Lite 提升 5.8%; 洪水分割 IoU 达到 89.2%, Dice 达到 93.8%, 分别提升 6.1% 与 5.7%, 有效解决了正常河道扩张误检、小面积洪水漏检的问题。KM³-SatNet 星载端推理时延为 8.8 ms, 能耗为 0.044 J, 相较于轻量化模型 DeepLabV3+ (MobileNetV2), 精度大幅提升的同时, 时延与能耗仅小幅增加, 完全满足星载端的实时性与低能耗要求。KM³-SatNet 模型参数量 7.8×10^6 , 远低于星载端存储约束 (≤ 512 MB), 推理时延 ≤ 10 ms, 满足星载

端实时推理需求 (≤ 100 ms), 验证了本文方法的工程可部署性。与 EarthGPT-X 和 LAMF-Net 相比, KM³-SatNet 在保持更高精度的同时显著降低参数量、星载端时延与能耗, 更符合受限星载环境的部署需求。实验结果验证了本文提出的 KM³-SatNet 方法通过任务驱动的动态知识子图构建与知识约束门控融合, 引入遥感领域先验知识, 有效抑制了视觉噪声干扰, 解决了传统模型的误检、漏检问题, 大幅提升了推理精度。通过模块化可裁剪结构、轻量化编码器、特征量化压缩等策略, 在保证精度的前提下, 将模型参数量与计算开销控制在星载端可承受范围内, 实现了精度与算力的平衡。模型在星载端模拟平台上的推理时延 ≤ 10 ms, 能耗 ≤ 0.05 J, 参数量 $\leq 8 \times 10^6$, 完全满足低轨卫星的算力、存储、能耗约束, 具备实际工程部署价值。

表 7 多模态遥感目标识别任务对比实验结果

Table 7 Comparison results on multimodal remote sensing object recognition

模型	mAP@0.5/%	参数量	FLOPs	星载端时延/ms	星载端能耗/J
DeepLabV3+(MobileNetV2)	78.5	2.2×10^6	5×10^8	5.1	0.026
YOLOv8-nano	82.4	3.6×10^6	8×10^8	6.8	0.034
Swin-L + CrossAttention	88.9	5.8×10^8	2.56×10^{11}	278.3	1.391
ViT-B + KGAT	87.2	1.285×10^8	3.26×10^{10}	45.2	0.226
RS-MLLM-Lite	86.3	1.56×10^7	2.8×10^4	28.5	0.142
EarthGPT-X	85.7	4.12×10^8	1.68×10^{11}	198.6	0.993
LAMF-Net	85.9	1.24×10^7	3.1×10^9	12.6	0.063
KM ³ -SatNet(无知识增强)	89.5	6.5×10^6	1.0×10^9	7.5	0.038
KM ³ -SatNet(本文方法)	92.6	7.8×10^6	1.2×10^9	9.2	0.046

表 8 灾害场景识别任务对比实验结果

Table 8 Comparison results on disaster scene recognition

模型	分类 Acc/%	分割 IoU/%	分割 Dice/%	星载端时延/ms	星载端能耗/J
DeepLabV3+(MobileNetV2)	82.3	76.5	81.2	5.0	0.025
YOLOv8-nano	84.6	78.2	83.5	6.5	0.033
Swin-L + CrossAttention	90.2	85.6	90.1	265.4	1.327
ViT-B + KGAT	89.5	84.3	89.2	42.6	0.213
RS-MLLM-Lite	88.7	83.1	88.1	27.8	0.139
EarthGPT-X	87.5	82.4	87.2	191.8	0.959
LAMF-Net	88.1	82.8	87.9	11.9	0.060
KM ³ -SatNet(无知识增强)	91.2	86.5	91.0	7.2	0.036
KM ³ -SatNet(本文方法)	94.5	89.2	93.8	8.8	0.044

4.5 消融实验

为验证 KM³-SatNet 各核心模块的有效性, 设计消融实验, 依次移除/关闭各核心模块, 在灾害场景识别任务 (洪水分割 IoU) 上进行测试, 结果如表 9 所示。

由表 9 可知, 各核心模块均对模型性能有显著提升。知识子图构建模块 (M1) IoU 提升 6.5%, 证明任务驱动的动态知识子图能够为推理提供有效的领域

先验知识, 减少无关知识干扰。知识约束门控融合模块 (M2) IoU 进一步提升 5.3%, 验证了知识约束门控机制能够有效实现多模态特征的鲁棒融合, 抑制视觉噪声, 强化关键语义特征的交互。结构化推理模块 (M3) IoU 提升 1.6%, 小幅提升精度的同时仅增加 1.0 ms 时延, 证明结构化推理能够挖掘知识子图的深层语义关联, 进一步提升推理精度。因果/时空一致性模块

表9 KM³-SatNet 消融实验结果(洪水分割 IoU)Table 9 Ablation results of KM³-SatNet on flood segmentation IoU

模型变体	核心模块配置	IoU/%	精度变化/%	星载端时延/ms	时延变化/ms
基础模型	仅 M0(基础感知模块)	75.8	—	4.5	—
+M1	基础模型+知识子图构建	82.3	+6.5	6.2	+1.7
+M1+M2	基础模型+M1+知识约束门控融合	87.6	+5.3	7.8	+1.6
+M1+M2+M3	基础模型+M1+M2+结构化推理	89.2	+1.6	8.8	+1.0
+M1+M2+M3+M4/M5	基础模型+全模块	89.2	0	8.8	0
KM ³ -SatNet(本文方法)	全模块+端云协同优化	89.2	0	8.8	0

(M4/M5)对洪水分割 IoU 无明显提升,但能有效降低异常值比例,提升推理结果的稳定性,适用于灾害监测的实际场景。端云协同优化模块对星载端推理精度与时延无影响,但其核心价值在于为高复杂度任务提供云端算力支撑,实现精度与实时性的动态平衡。

消融实验结果表明, KM³-SatNet 的各核心模块设计具有必要性与合理性,模块间的协同作用实现了精度与算力的最优平衡。除模块级消融外,进一步补充了知识使用方式与门控策略的细粒度替换实验,结果如表 10 所示。

表 10 知识子图与门控策略的细粒度消融/替换结果(洪水分割任务)

Table 10 Fine-grained ablation and replacement results for knowledge subgraphs and gating strategies on flood segmentation

模型变体	知识使用方式	门控策略	IoU/%	星载端时延/ms
A1	不使用知识图谱	标准跨模态注意力	81.7	7.1
A2	静态全局图谱	标准跨模态注意力	86.4	12.8
A3	动态任务子图	无门控,仅标准注意力	85.9	8.9
A4	动态任务子图	仅 Token 门控	88.6	7.8
A5	动态任务子图	仅矩阵门控	89.0	9.6
A6(完整模型)	动态任务子图	预算自适应双粒度门控	89.2	8.8

动态任务子图优于静态全局图谱,说明任务相关知识裁剪能够在降低冗余知识干扰的同时保持更好的精度—时延平衡。知识条件化门控优于标准跨模态注意力,说明引入任务语义与知识先验后能够有效抑制错误的跨模态关联。Token 门控和矩阵门控分别适用于严格实时预算与算力相对充足的场景。完整模型在精度—时延折中上最优。

4.6 复杂度分析

为进一步验证 KM³-SatNet 在星载端的可部署性,对模型的时间复杂度与空间复杂度进行定量分析,并与对比模型进行对比,结果如图 5 所示。

本文方法的时间复杂度为 $O(N+K)$, 远低于传统 Transformer 模型的 $O(N^2 \log N)$ 与 $O(N^2)$, 因为通过动态知识子图构建将知识推理复杂度从全局图谱的 $O(K_{\text{total}})$ 降至任务子图的 $O(K)$ ($K \leq 20$), 同时采用轻量化编码器降低视觉特征提取复杂度, 实现了线性时间复杂度。本文方法的空间复杂度为 $O(N+K)$, 仅需存储视觉 Token、知识子图实体特征, 无需存储大规模注意力矩阵, 星载端存储占用仅为 45.2 MB, 远低于星载端存储约束 (≤ 512 MB)。本文方法的星载端推理吞吐量达到 113.6 FPS (每秒处理图像数), 远高于大规模模型与知识增强模型, 满足星载端实时推

理的吞吐量要求。复杂度分析结果进一步验证了 KM³-SatNet 的轻量化设计能够有效降低模型的计算复杂度。

4.7 鲁棒性实验

为验证 KM³-SatNet 在卫星遥感实际复杂环境中的鲁棒性, 设计鲁棒性实验, 在灾害场景识别任务中加入“云遮挡 (0%~80%)、SAR 相干斑噪声 (0~0.5)、光照变化 (0.2~1.8)”三种典型干扰, 测试不同模型的精度衰减率 (ADR), 结果如图 6 所示。

本文方法在三种典型干扰条件下的平均精度衰减率仅为 4.5%, 远低于所有对比模型, 相较于最优对比模型 RS-MLLM-Lite (平均 ADR 10.6%), 鲁棒性提升 57.5%, 验证了知识约束门控融合机制能够有效抑制遥感数据的噪声干扰, 提升模型的环境适应性。在 SAR 相干斑噪声干扰下, 本文方法的 ADR 仅为 5.2%, 表现最优, 证明知识图谱中的 SAR 数据特征知识能够为模型提供有效的噪声抑制先验; 在云遮挡干扰下, ADR 为 4.5%, 证明知识约束能够引导模型聚焦于核心语义特征, 减少云遮挡带来的特征缺失影响。对比 KM³-SatNet (无知识增强), 本文方法的平均 ADR 从 8.2% 降至 4.5%, 证明知识增强是提升模型鲁棒性的关键因素, 领域先验知识能够有效弥补遥感数据的噪声与缺失。

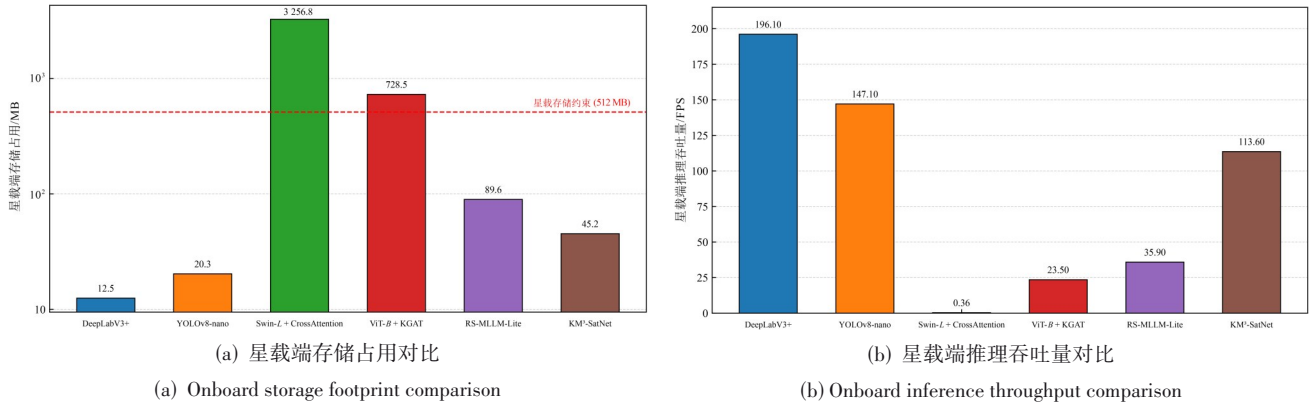


图5 星载端存储占用与推理吞吐量对比

Figure 5 Comparison of onboard storage footprint and inference throughput

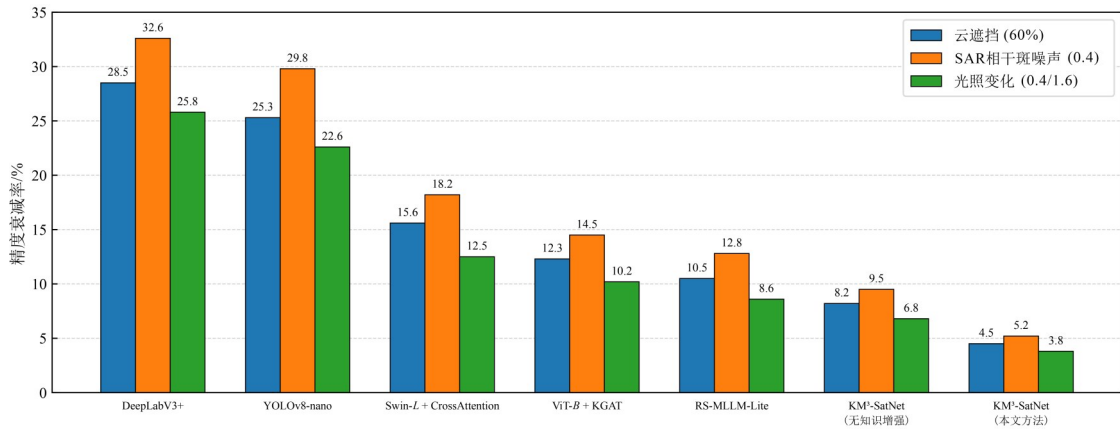


图6 不同干扰条件下模型的鲁棒性对比

Figure 6 Robustness comparison of models under different interference conditions

4.8 端云协同优化实验结果与分析

为验证端云协同优化策略的有效性,在不同星地链路速率下,测试KM³-SatNet在灾害场景识别任务中的端云总时延、星载总能耗与推理精度,结果如图7所示。

在不同链路速率下,本文的端云协同优化策略(动态特征压缩)相较于无优化策略,端云总时延降低,星载总能耗降低,验证了特征压缩策略能够有效降低星地通信开销。在链路速率 ≥ 1 Mbps时,精度损失均 $\leq 0.9\%$,在链路速率 ≥ 5 Mbps时,精度损失 $\leq 0.5\%$,实现了通信代价与推理精度的平衡,证明轻量特征压缩策略能够在降低传输代价的同时,最大程度保留特征语义信息。通过动态调整压缩比 ρ ,模型能够适应不同的星地链路状态,在高带宽时追求高精度,在低带宽时保证实时性,验证了端云协同优化策略的动态适配性。端云协同实验结果表明,本文提出的端云协同优化策略能够有效实现星载端与云端的算力协同,在保证推理精度的前提下,大幅降低端云总时延与星

载总能耗,提升模型在空天地一体化场景中的适应性。

4.9 关键超参数敏感性实验与分析

选取温度系数 τ 、扩展步长 H 、PPR重启概率 η 、惩罚系数 λ 等对模型性能影响最显著的超参数,开展敏感性实验,实验结果如图8所示。

以灾害场景识别(洪水分割)任务为基准,固定其他超参数为最优值,单独调整目标超参数的取值,测试洪水分割IoU(精度指标)、星载端推理时延(实时性指标)、子图存储占用(存储指标)三大核心指标的变化规律。

(1) 温度系数 τ 敏感性分析

温度系数 τ 取值范围 $[0.1, 0.2, 0.3, 0.4, 0.5, 0.6]$ 。当 $\tau=0.1\sim 0.2$ 时,洪水分割IoU维持在88.5%~89.2%的高值,语义匹配精准性高,几乎无无关实体被抽取。当 $\tau=0.3\sim 0.4$ 时,IoU降至86.2%~88.0%,因语义区分度下降,少量无关实体(如正常湖泊)被纳入子图。当 $\tau\geq 0.5$ 时,IoU骤降至80%以下,语义匹配失效,大

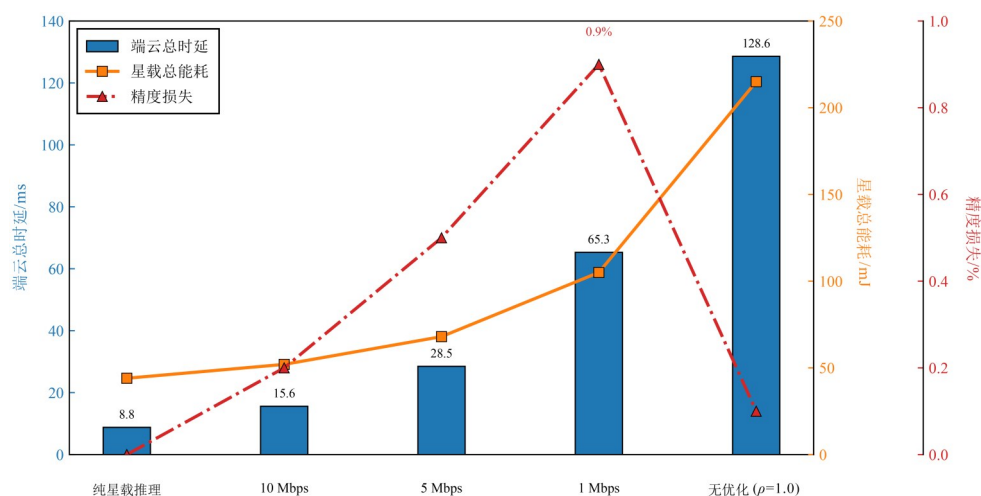


图7 不同链路速率下端云协同优化实验结果

Figure 7 Edge-cloud collaborative optimization results under different satellite-ground link rates

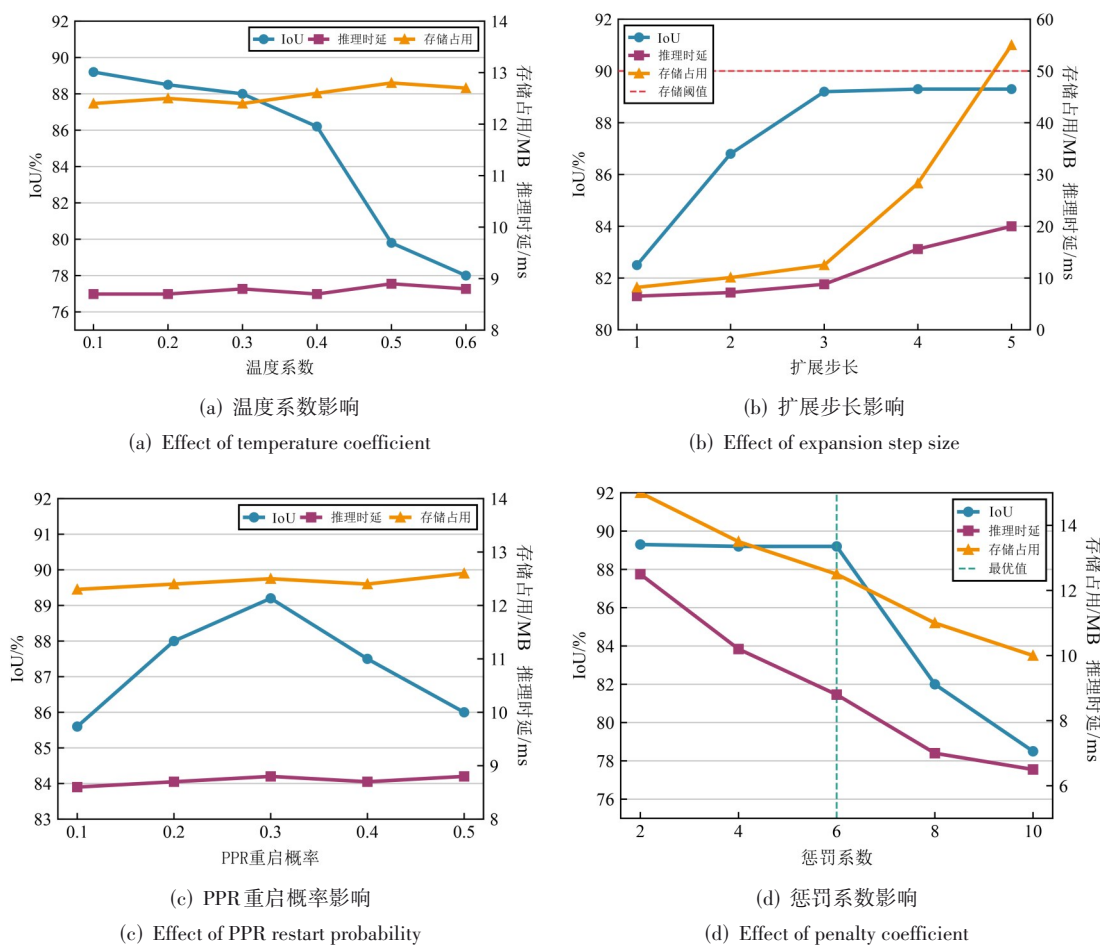


图8 关键超参数对洪水分割IoU、推理时延及子图存储占用的敏感性分析

Figure 8 Sensitivity analysis of key hyperparameters with respect to flood segmentation IoU, inference latency, and subgraph storage footprint

量无关知识引入导致推理偏差。时延与存储指标随 τ 变化无明显波动(波动 ≤ 0.2 ms、 ≤ 0.5 MB)。遥感任务需设置小 τ (0.2~0.3), 优先保证语义匹配精准性, 对

实时性与存储无显著影响。

(2) 扩展步长 H 敏感性分析

扩展步长 H 取值范围 $[1, 2, 3, 4, 5]$ 。当 $H=1$ 时,

IoU 仅为 82.5%, 知识子图过于精简, 缺失洪水分割的核心因果关联(如强降雨→洪水)。当 $H=3$ 时, IoU 从 86.8% 提升至 89.2%, 知识子图覆盖完整的洪水分割领域知识, 时延从 6.5 ms 升至 8.8 ms, 存储从 8.2 MB 升至 12.5 MB, 均在星载端约束内。当 $H=4$ 时, IoU 微升至 89.3% (提升 0.1%), 但时延骤升至 15.6 ms, 存储升至 28.3 MB, 性价比极低。当 $H=5$ 时, 子图规模超限(存储 > 50 MB), 星载端无法部署。根据任务知识关联复杂度选择 H , 简单任务(目标识别)取 2, 复杂任务(洪水分割)取 3, $H \geq 4$ 时存在明显的精度-算力失衡。

(3) PPR 重启概率 η 敏感性分析

PPR 重启概率 η 取值范围 [0.1, 0.2, 0.3, 0.4, 0.5]。当 $\eta=0.1$ 时, IoU 为 85.6%, 随机游走过远, 子图引入大量无关实体(如地理区域下的非洪水区域)。当 $\eta=0.2 \sim 0.3$ 时, IoU 升至 88.0%~89.2%, 子图聚焦种子实体且保留核心关联实体, 时延与存储稳定。当 $\eta \geq 0.4$ 时, IoU 逐渐下降(87.5%~86.0%), 因随机游走过度聚焦种子实体, 缺失重要的关联属性实体(如低洼区、水位阈值)。 η 的最优取值为 0.2~0.3, 实现种子实体聚焦与关联知识完整的平衡, 过大或过小均会导致精度下降。

(4) 惩罚系数 λ 敏感性分析

惩罚系数 λ 取值范围 [2, 4, 6, 8, 10]。当 $\lambda=2 \sim 4$ 时, 算力预算约束宽松, 模型未进行有效裁剪, IoU 为 89.2%~89.3%, 但时延升至 12.5~10.2 ms, 超出部分星载端的时延要求。当 $\lambda=6$ 时, IoU 保持 89.2%, 时延 8.8 ms, 存储 12.5 MB, 实现精度与算力的最优平衡。当 $\lambda=8 \sim 10$ 时, 算力预算约束过严, 模型过度裁剪, IoU 骤降至 82.0%~78.5%, 虽时延降至 7.0~6.5 ms, 但精度损失不可接受。惩罚系数需适中(5~6), 兼顾算力约束的严格性与模型精度, 避免过松或过严。

5 结束语

本文针对空天地一体化背景下卫星端多模态智能计算面临的算力受限、多模态数据噪声复杂、任务语义动态变化、端云协同开销高等核心问题, 提出了知识引导的多模态模块化卫星网络 KM^3 -SatNet。提出任务驱动的动态知识子图构建机制, 构建了卫星遥感任务知识图谱(SRS-KG), 通过语义匹配、结构扩展与 PPR 压缩, 实现与当前任务强相关的知识子图抽取, 将知识推理复杂度从全局图谱降至任务子图, 有效降低了计算开销与无关知识干扰。设计知识约束门控的多模态融合框架, 引入知识图谱的先验语义约束, 构建 Token 级/矩阵级双模式门控机制, 实现多模态特征的鲁棒融合, 有效抑制了遥感数据的噪声干

扰, 提升了模型的语义判别能力与鲁棒性。提出模块化可裁剪推理结构, 将推理流程解耦为功能独立的可插拔模块, 设计算力感知的轻量化控制器, 实现模块配置的自适应选择, 严格遵循星载端的算力与存储约束, 实现了推理性能与计算开销的动态平衡。提出端云协同优化策略, 基于模块依赖关系设置可切分点, 构建星地通信开销模型与两阶段启发式联合优化机制, 实现星载端与云端的算力协同, 在保证精度的前提下, 大幅降低端云总时延与星载总能耗, 提升了模型在空天地一体化场景中的适应性。大量实验结果表明, KM^3 -SatNet 在多模态遥感目标识别任务中精度提升 6.3%, 在灾害场景识别任务中精度提升 5.8%, 同时模型参数数量减少, 推理时延降低, 在复杂干扰条件下的平均精度衰减率仅为 4.5%。

参考文献

- [1] 孟瑜, 陈静波, 张正, 等. 知识与数据驱动的遥感图像智能解译: 进展与展望[J]. 遥感学报, 2024, 28(11): 2698-2718.
Meng Yu, Chen Jingbo, Zhang Zheng, et al. Knowledge and data driven remote sensing image interpretation: Recent developments and prospects[J]. National Remote Sensing Bulletin, 2024, 28(11): 2698-2718. (in Chinese)
- [2] 蒋长林, 李清, 王羽, 等. 天地一体化网络关键技术研究综述[J]. 软件学报, 2024, 35(1): 266-287.
Jiang Changlin, Li Qing, Wang Yu, et al. Survey on key technologies in space-ground integrated network[J]. Journal of Software, 2024, 35(1): 266-287. (in Chinese)
- [3] 康利鸿, 田菁, 江碧涛. 巨星座时代遥感卫星应用技术挑战与思考[J]. 遥感学报, 2024, 28(7): 1658-1666.
Kang Lihong, Tian Jing, Jiang Bitao. Challenges and research on remote sensing satellite application technology in the Giant Constellation Era[J]. National Remote Sensing Bulletin, 2024, 28(7): 1658-1666. (in Chinese)
- [4] 刘昇, 张寅捷, 敖洋, 等. Transformer 架构下跨多尺度信息融合的遥感影像建筑提取[J]. 遥感学报, 2024, 28(12): 3173-3183.
Liu Yi, Zhang Yinjie, Ao Yang, et al. Building extraction from remote sensing images via the multiscale information fusion method under the Transformer architecture[J]. National Remote Sensing Bulletin, 2024, 28(12): 3173-3183. (in Chinese)
- [5] Chen L C, Zhu Yukun, Papandreou G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[C]//Computer Vision - ECCV 2018. Cham: Springer International Publishing, 2018: 833-851.
- [6] 李树涛, 李聪好, 康旭东. 多源遥感图像融合发展现状与

- 未来展望[J]. 遥感学报, 2021, 25(1): 148-166.
- Li Shutao, Li Congyu, Kang Xudong. Development status and future prospects of multi-source remote sensing image fusion[J]. National Remote Sensing Bulletin, 2021, 25(1): 148-166. (in Chinese)
- [7] Zhang Wei, Cai Miaoxin, Ning Yaqian, et al. EarthGPT-X: A spatial MLLM for multilevel multisource remote sensing imagery understanding with visual prompting[J]. IEEE Transactions on Geoscience and Remote Sensing, 2025, 63: 1-21.
- [8] 张雪英, 张春菊, 吴明光, 等. 顾及时空特征的地理知识图谱构建方法[J]. 中国科学(信息科学), 2020, 50(7): 1019-1032.
- Zhang Xueying, Zhang Chunju, Wu Mingguang, et al. Spatiotemporal features based geographical knowledge graph construction[J]. Scientia Sinica (Informationis), 2020, 50(7): 1019-1032. (in Chinese)
- [9] Wang Xiang, He Xiangnan, Cao Yixin, et al. KGAT: Knowledge graph attention network for recommendation[C]// Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. New York: ACM, 2019: 950-958.
- [10] 夏玮玮, 胡静, 宋铁成. 低地球轨道卫星边缘计算场景中任务卸载与资源分配联合优化算法[J]. 通信学报, 2024, 45(7): 48-60.
- Xia Weiwei, Hu Jing, Song Tiecheng. Joint optimization algorithm for task offloading and resource allocation in low earth orbit satellites edge computing scenario[J]. Journal on Communications, 2024, 45(7): 48-60. (in Chinese)
- [11] 徐晓斌, 王琪, 范存群, 等. 面向空天地一体化信息网络的边缘计算资源融合管理方法[J]. 计算机学报, 2023, 46(4): 690-710.
- Xu Xiaobin, Wang Qi, Fan Cunqun, et al. An aggregated edge computing resource management method for space-air-ground integrated information networks[J]. Chinese Journal of Computers, 2023, 46(4): 690-710. (in Chinese)
- [12] 朱立东, 张勇, 贾高一. 卫星互联网路由技术现状及展望[J]. 通信学报, 2021, 42(8): 33-42.
- Zhu Lidong, Zhang Yong, Jia Gaoyi. Current status and future prospects of routing technologies for satellite Internet[J]. Journal on Communications, 2021, 42(8): 33-42. (in Chinese)
- [13] 付琨, 王佩瑾, 冯瑛超, 等. 遥感跨模态智能解译: 模型、数据与应用[J]. 中国科学(信息科学), 2023, 53(8): 1529-1559.
- Fu Kun, Wang Peijin, Feng Yingchao, et al. Cross-modal remote sensing intelligent interpretation: Method, data, and application[J]. Scientia Sinica Informationis, 2023, 53(8): 1529-1559. (in Chinese)
- [14] Howard A G, Zhu Menglong, Chen Bo, et al. MobileNets: Efficient convolutional neural networks for mobile vision applications[PP/OL]. arXiv (2017-04-17)[2026-04-20]. <https://doi.org/10.48550/arXiv.1704.04861>.
- [15] Ma Ningning, Zhang Xiangyu, Zheng Haitao, et al. ShuffleNet V2: Practical guidelines for efficient CNN architecture design[C]//Computer Vision - ECCV 2018. Cham: Springer International Publishing, 2018: 122-138.
- [16] Redmon J, Farhadi A. YOLOv3: An incremental improvement[PP/OL]. arXiv (2018-04-08)[2026-04-20]. <https://doi.org/10.48550/arXiv.1804.02767>.
- [17] 徐萌, 王思涵, 郭仁忠, 等. 遥感影像云检测和云去除方法综述[J]. 计算机研究与发展, 2024, 61(6): 1585-1607.
- Xu Meng, Wang Sihan, Guo Renzhong, et al. Review of cloud detection and cloud removal methods for remote sensing images[J]. Journal of Computer Research and Development, 2024, 61(6): 1585-1607. (in Chinese)
- [18] 鲁自立, 贾鑫, 朱卫纲, 等. SAR图像相干斑抑制方法综述[J]. 兵器装备工程学报, 2017, 38(6): 104-108.
- Lu Zili, Jia Xin, Zhu Weigang, et al. Study on SAR image despeckling algorithm[J]. Journal of Ordnance Equipment Engineering, 2017, 38(6): 104-108. (in Chinese)
- [19] 丁一帆, 尤红建, 张浩, 等. 面向高分辨率遥感影像的地形辐射校正方法[J]. 北京航空航天大学学报, 2018, 44(1): 27-35.
- Ding Yifan, You Hongjian, Zhang Hao, et al. Topographic correction method for high-resolution remote sensing images[J]. Journal of Beijing University of Aeronautics and Astronautics, 2018, 44(1): 27-35. (in Chinese)
- [20] Hong Danfeng, Gao Lianru, Yao Jing, et al. Graph convolutional networks for hyperspectral image classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, 59(7): 5966-5978.
- [21] 张继贤, 李国胜, 曾钰. 多源遥感影像高精度自动配准的方法研究[J]. 遥感学报, 2005, 9(1): 73-77.
- Zhang Jixian, Li Guosheng, Zeng Yu. The study on automatic and high-precision rectification and registration of multi-source remote sensing imagery[J]. National Remote Sensing Bulletin, 2005, 9(1): 73-77. (in Chinese)
- [22] 张良培, 武辰. 多时相遥感影像变化检测的现状与展望[J]. 测绘学报, 2017, 46(10): 1447-1459.
- Zhang Liangpei, Wu Chen. Advance and future development of change detection for multi-temporal remote sensing imagery[J]. Acta Geodaetica et Cartographica Sinica, 2017, 46(10): 1447-1459. (in Chinese)
- [23] Xia Guisong, Bai Xiang, Ding Jian, et al. DOTA: A large-

- scale dataset for object detection in aerial images[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 3974-3983.
- [24] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation[C]//Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015. Cham: Springer International Publishing, 2015: 234-241.
- [25] 徐丹青, 吴一全. 光学遥感图像目标检测的深度学习算法研究进展[J]. 遥感学报, 2024, 28(12): 3045-3073.
Xu Danqing, Wu Yiquan. Progress of research on deep learning algorithms for object detection in optical remote sensing images[J]. National Remote Sensing Bulletin, 2024, 28(12): 3045-3073. (in Chinese)
- [26] Bordes A, Usunier N, Garcia-Durán A, et al. Translating embeddings for modeling multi-relational data[C]//Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2. New York: ACM, 2013: 2787-2795.
- [27] Schlichtkrull M, Kipf T N, Bloem P, et al. Modeling relational data with graph convolutional networks[C]//The Semantic Web. Cham: Springer International Publishing, 2018: 593-607.
- [28] Johnson J, Douze M, Jégou H. Billion-scale similarity search with GPUs[J]. IEEE Transactions on Big Data, 2021, 7(3): 535-547.
- [29] Haveliwala T H. Topic-sensitive PageRank: A context-sensitive ranking algorithm for Web search[J]. IEEE Transactions on Knowledge and Data Engineering, 2003, 15(4): 784-796.
- [30] Jang E, Gu Shixiang, Poole B. Categorical reparameterization with Gumbel-Softmax[C]//Proceedings of the 5th International Conference on Learning Representations. Toulon: OpenReview.net, 2017: 24-26.
- [31] Loshchilov I, Hutter F. Decoupled weight decay regularization[C]//Proceedings of the 7th International Conference on Learning Representations. New Orleans: OpenReview.net, 2019: 1-12.
- [32] 施建锋, 陈忻阳, 李宝龙. 面向物联网的云边端协同计算中任务卸载与资源分配算法研究[J]. 电子与信息学报, 2025, 47(2): 458-469.
Shi Jianfeng, Chen Xinyang, Li Baolong. Research on task offloading and resource allocation algorithms in cloud-edge-end collaborative computing for the Internet of Things[J]. Journal of Electronics & Information Technology, 2025, 47(2): 458-469. (in Chinese)
- [33] Song Yaohui, Li Xi, Ji Hong, et al. Energy-aware task offloading and resource allocation in the intelligent LEO satellite network[C]//2022 IEEE 33rd Annual International Symposium on Personal, Indoor and Mobile Radio Communications. Piscataway: IEEE, 2022: 481-486.
- [34] Lan Wenjun, Chen Kongyang, Cao Jiannong, et al. Security-sensitive task offloading in integrated satellite-terrestrial networks[J]. IEEE Transactions on Mobile Computing, 2025, 24(3): 2220-2233.
- [35] 张其阳, 邢若琳, 李元哲, 等. 卫星边缘计算智能化技术研究进展[J]. 软件学报, 2026, 37(1): 326-343.
Zhang Qiyang, Xing Ruolin, Li Yuanzhe, et al. Research progress on intelligent technologies for satellite edge computing[J]. Journal of Software, 2026, 37(1): 326-343. (in Chinese)
- [36] 沈学民, 承楠, 周海波, 等. 空天地一体化网络技术: 探索与展望[J]. 物联网学报, 2020, 4(3): 3-19.
Shen Xuemin, Cheng Nan, Zhou Haibo, et al. Space-air-ground integrated networks: Review and prospect[J]. Chinese Journal on Internet of Things, 2020, 4(3): 3-19. (in Chinese)
- [37] 赵亚军, 郁光辉, 徐汉青. 6G 移动通信网络: 愿景、挑战与关键技术[J]. 中国科学(信息科学), 2019, 49(8): 963-987.
Zhao Yajun, Yu Guanghui, Xu Hanqing. 6G mobile communication networks: Vision, challenges, and key technologies[J]. Scientia Sinica (Informationis), 2019, 49(8): 963-987. (in Chinese)
- [38] 倪少杰, 岳洋, 左勇, 等. 卫星网络路由技术现状及展望[J]. 电子与信息学报, 2023, 45(2): 383-395.
Ni Shaojie, Yue Yang, Zuo Yong, et al. The status quo and prospect of satellite network routing technology[J]. Journal of Electronics & Information Technology, 2023, 45(2): 383-395. (in Chinese)
- [39] 付一阳, 胡博, 刘人鹏, 等. 星地融合网络中基于异质图表征的多智能体协作切换方法[J]. 计算机学报, 2026, 49(2): 347-364.
Fu Yiyang, Hu Bo, Liu Renpeng, et al. A multi-agent heterogeneous graph representation-based handover method in satellite-terrestrial integrated network[J]. Chinese Journal of Computers, 2026, 49(2): 347-364. (in Chinese)
- [40] 梁广俊, 王群, 辛建芳, 等. 移动边缘计算资源分配综述[J]. 信息安全学报, 2021, 6(3): 227-256.
Liang Guangjun, Wang Qun, Xin Jianfang, et al. Survey of mobile edge computing resource allocation[J]. Journal of Cyber Security, 2021, 6(3): 227-256. (in Chinese)
- [41] Konečný J, McMahan H B, Yu F X, et al. Federated learning: Strategies for improving communication efficiency[PP/OL]. V2. arXiv (2017-10-30)[2026-04-20]. <https://doi.org/10.48550/arXiv.1610.05492>.

- [42] 张新长, 华淑贞, 齐霁, 等. 新型智慧城市建设与展望: 基于 AI 的大数据、大模型与大算力[J]. 地球信息科学学报, 2024, 26(4): 779-789.
Zhang Xinchang, Hua Shuzhen, Qi Ji, et al. Progress and prospects of new smart city construction: AI-based big data, big models and big computing power[J]. Journal of Geo-Information Science, 2024, 26(4): 779-789. (in Chinese)
- [43] 徐常志, 靳一, 李立, 等. 面向 6G 的星地融合无线传输技术[J]. 电子与信息学报, 2021, 43(1): 28-36.
Xu Changzhi, Jin Yi, Li Li, et al. Wireless transmission technology of satellite-terrestrial integration for 6G mobile communication[J]. Journal of Electronics & Information Technology, 2021, 43(1): 28-36. (in Chinese)
- [44] 刘雪芳, 毛伟灏, 杨清海. 基于深度强化学习的空天地一体化网络资源分配算法[J]. 电子与信息学报, 2024, 46(7): 2831-2841.
Liu Xuefang, Mao Weihao, Yang Qinghai. A resource allocation algorithm for space-air-ground integrated network based on deep reinforcement learning[J]. Journal of Electronics & Information Technology, 2024, 46(7): 2831-2841. (in Chinese)
- [45] 冯炼, 皮雪晖. 地表水体范围遥感监测方法、应用与挑战[J]. 遥感学报, 2025, 29(6): 2139-2161.
Feng Lian, Pi Xuehui. Remote sensing monitoring methods, applications, and challenges for surface water bodies[J]. National Remote Sensing Bulletin, 2025, 29(6): 2139-2161. (in Chinese)
- [46] Gupta R, Hosfelt R, Sajeve S, et al. xBD: A dataset for assessing building damage from satellite imagery[PP/OL]. arXiv (2019-11-21) [2026-04-20]. <https://doi.org/10.48550/arXiv.1911.09296>.
- [47] 吴立新, 龚健雅. 中国遥感学科的形成过程、现状与挑战[J]. 遥感学报, 2025, 29(6): 1314-1330.
Wu Lixin, Gong Jianya. The formation process, construction status and challenging issues of remote sensing discipline in China[J]. National Remote Sensing Bulletin, 2025, 29(6): 1314-1330. (in Chinese)
- [48] 王志华, 杨晓梅, 张俊瑶, 等. 面向复杂自然场景的遥感地学分区智能解译框架及初探[J]. 地球信息科学学报, 2025, 27(2): 305-330.
Wang Zhihua, Yang Xiaomei, Zhang Junyao, et al. A remote sensing intelligent interpretation framework through geo-science zoning for complex nature scenes and its preliminary experiments[J]. Journal of Geo-Information Science, 2025, 27(2): 305-330. (in Chinese)
- [49] Varghese R, Sambath M. YOLOv8: A novel object detection algorithm with enhanced performance and robustness[C]//2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems. Piscataway: IEEE, 2024: 1-6.
- [50] Liu Ze, Lin Yutong, Cao Yue, et al. Swin Transformer: Hierarchical vision Transformer using shifted windows[C]//2021 IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE, 2021: 9992-10002.
- [51] Shen Cheng, Ding Bo, Zhang Jing, et al. Lightweight and adaptive multimodal fusion network for multisource remote sensing[C]//International Conference on Remote Sensing and Digital Earth (RSDE 2025). Bellingham: SPIE, 2025: 1405408.

作者简介



李倩 女, 1994年5月生, 甘肃会宁人。博士, 北京邮电大学网络与交换技术国家重点实验室助理教授。主要研究方向为多模态大数据、卫星网络大数据分析。
E-mail: li.qian@bupt.edu.cn



王尚广 男, 1982年2月生, 河南周口人。博士, 北京邮电大学计算机学院教授、博士生导师。主要研究方向为计算机视觉、卫星计算、服务计算与边缘智能等。中国电子学会会员编号: E190027924S。
E-mail: sgwang@bupt.edu.cn