

## 特色专题

# 基于大语言模型的考古遗址精确定位

陆宇翔<sup>1,2</sup>, 沈靖<sup>2,3\*</sup>, 姜侯<sup>2</sup>, 刘唐<sup>2</sup>

**摘要** 由于早期考古工作受限于技术手段,大量遗址仅以文本形式记载大致位置,缺乏精确坐标信息,给后续的考古调查、遗址保护与研究带来了诸多不便。传统的田野调查方法耗时费力,难以适应大规模、多批次的遗址定位需求。为解决这一问题,提出了一种基于大语言模型的考古遗址智能定位框架。该方法通过设计自然语言提示,引导大语言模型从互联网中的考古文献与资料中自动提取遗址的地理描述信息(如参照物、方位、距离等);随后结合遥感影像分析技术,获取目标区域的高分辨率卫星图像,综合推理并计算遗址的精确地理坐标。该流程实现了从文本信息抽取、遥感数据调用到空间位置反演的全流程自动化,显著提升了遗址定位的效率与智能化水平。在老虎山遗址、三星堆遗址和辽中京遗址的实证验证中,自动定位结果与实测坐标的偏差控制在千米以内,最优精度可达 10 m 级,满足考古学对定位精度的基本要求。定位误差主要来源于文献描述模糊和影像判读的不确定性。与传统人工方法相比,该方法不仅大幅减少了人力投入,还具备更强的可扩展性和应用潜力。研究成果为考古信息数字化、遗址保护与再研究提供了新的技术路径和工具支撑。

**关键词** 地理智能;考古遗址;精准定位;遥感信息提取;大语言模型

考古遗址的空间位置是遗址保护、文化遗产管理和历史复原的基础<sup>[1]</sup>。准确坐标不仅可绘制遗址分布,还可支撑遗址与周边环境关系的研究与文化景观再现<sup>[2]</sup>。然而,早期考古因测绘技术有限,报告多为口头或手绘描述,缺乏精确坐标,难以对应现代地图或遥感影像<sup>[3]</sup>。这种空间信息的缺失和偏差给遗址保护和后续考古研究带来诸多困难。因此,对历史遗址的精确定位成为亟需攻克的问题。

针对遗址空间定位的挑战,国内外学者提出了多

种解决思路。传统方法一般依赖对旧文献、手绘图纸等资料的人工研读,比对实地踏勘或遥感影像进行辅助验证,虽能改善定位精度,但效率低、主观性强,难以大规模推广应用<sup>[4-5]</sup>。

近年来,人工智能的发展带来了新机遇。大语言模型(large language model, LLM)在语义解析和信息抽取方面表现出色<sup>[6-8]</sup>。如基于提示增强的 LLM 信息抽取算法能够显著提升实体关系抽取的性能<sup>[9]</sup>。这种能力使得 LLM 可以自动识别考古文献中的复杂地理描述和空间关系。基于 LLM 的智能体技术可调用网络考古数据库、历史地图和文献资料,补充和校验遗址相关信息<sup>[10]</sup>。

此外,高分辨率遥感影像与开放地图应用程序接口(application program interface, API)的普及,为遗址定位提供了新的技术支撑<sup>[11-12]</sup>。例如,利用 WorldView-2

1. 中国地质大学(北京)人工智能学院,北京 100083

2. 中国科学院地理科学与资源研究所地理信息科学与技术  
全国重点实验室,北京 100101

3. 中国科学院大学,北京 100049

收稿日期: 2025-05-15; 修回日期: 2025-07-18

基金项目: 中国科学院战略性先导科技专项(XDB740200)

作者简介: 陆宇翔,博士研究生,研究方向为地理大模型,电子信箱: [luyuxiang21@mails.ucas.ac.cn](mailto:luyuxiang21@mails.ucas.ac.cn); 沈靖(通信作者),博士研究生,研究方向为遥感大数据,电子信箱: [shenjing6429@igsnrr.ac.cn](mailto:shenjing6429@igsnrr.ac.cn)

引用格式: 陆宇翔, 沈靖, 姜侯, 等. 基于大语言模型的考古遗址精确定位[J]. 科技导报, 2025, 43(18): 77-85; doi:10.3981/j.issn.1000-7857.2025.05.00086

卫星影像对突尼斯古罗马时期遗址进行勘测,成功识别并定位了 10 处此前未知的遗址<sup>[13]</sup>。这表明了结合历史资料与遥感影像能有效发现和定位遗址<sup>[14]</sup>。

基于此,本文提出一种融合 LLM 与遥感分析的多智能体遗址自动化定位方法。框架集成文本信息抽取、语义解析、影像获取与识别,实现端到端坐标计算,显著提升定位效率与规模化能力。以老虎山、三星堆和辽中京遗址为案例,结果显示定位偏差在十米至千米之间,满足考古研究精度需求。研究验证了大语言模型与遥感结合的可行性,并为考古遗址信息数字化与智能化提供了新路径。

1 基于 LLM 的考古遗址定位框架

本研究提出一种融合 LLM 与遥感解析的多智能体协同框架,通过层级化任务分解实现考古遗址的端到端自动化定位(图 1)。框架依次包括:(1) 文本阶段。由“遗址名称搜索智能体”和“遗址描述搜索智能体”对非结构化文本进行语义建模,提取遗址信息。(2) 粗定位阶段。“遗址文本定位智能体”将文本与空间坐标建立对应关系,获取遗址大致位置,并调用地理信息服务平台 API 获取不大于 0.6 m 分辨率的遥感影像。(3) 精定位阶段。“遗址影像定位智能体”利用多模态大模型识别地表特征并进行空间配准,最终输出 WGS84 坐标系下的精确坐标。



图 1 基于 LLM 的考古遗址定位框架

1.1 PEER 多智能体协作

本研究构建了基于计划、执行、表达、审查(plan-execute-express-review, PEER)架构的多智能体协同定位框架,旨在实现考古遗址的自动化精确定位(图 2)。框架采用分层递进机制:在规划阶段(plan),规划智能体通过语义解析生成包含文献解析、检索与影像获取的调度策略;执行阶段(execute),执行智能体并发驱动多模块运行;表达阶段(express),整合各模块输出,生成含空间要素、异常检测及溯源路径的

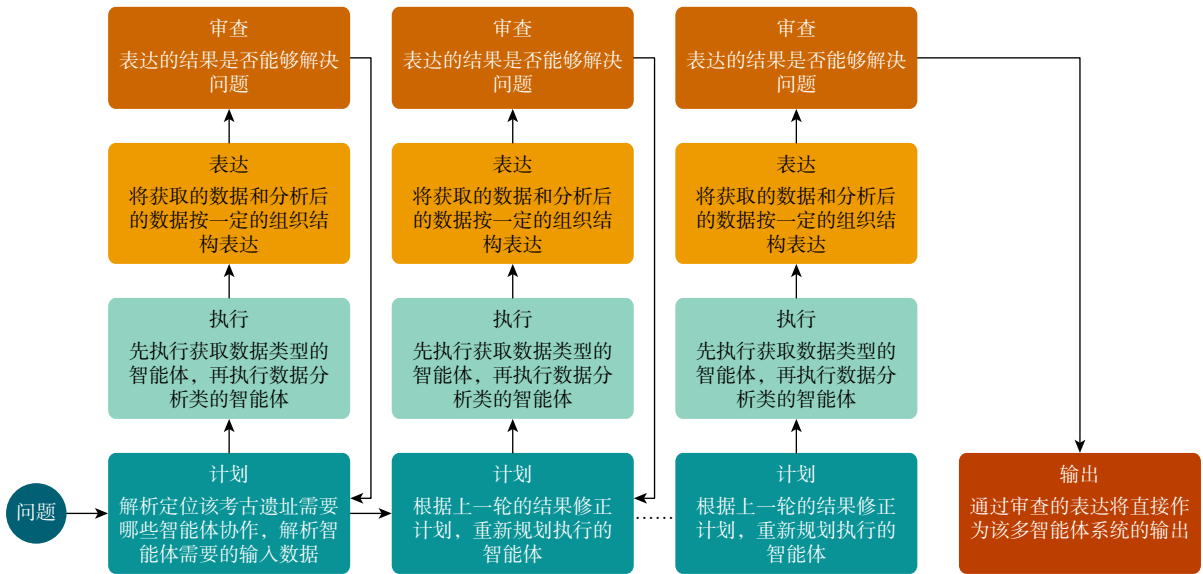


图 2 基于 LLM 的考古遗址精确定位遗址多智能体协作流程

结构化报告;审查阶段(review),基于置信度矩阵对定位结果进行多维验证,若未达精度阈值,则触发回溯机制迭代优化。

执行阶段是框架的核心,包含5个功能异构的智能体模块。数据获取层由遗址名称搜索、遗址描述搜索和影像获取智能体组成,分别完成文本清洗、地理实体识别和高分辨率遥感数据采集;计算层由文本定位和影像定位智能体协同工作,前者基于空间拓扑关系生成候选区域,后者利用多模态模型实现目标识别。两层级智能体通过数据总线交互,输出符合

WGS84坐标系的遗址坐标集合。系统结合动态闭环优化机制,提升了复杂地理环境下遗址定位的鲁棒性与可重复性。

1.2 遗址名称搜索智能体

考古遗址常因时代、学术流派或地方志差异而存在多种异名,若未校准将导致定位偏差。本研究设计“遗址名称校验智能体”,对输入名称进行多源核实与标准化(图3)。该智能体构建“名称-时代-地域”多维关键词,对考古文献、地方志、测绘报告及在线数据库进行并行检索,汇总名称变体及其使用频次;随后

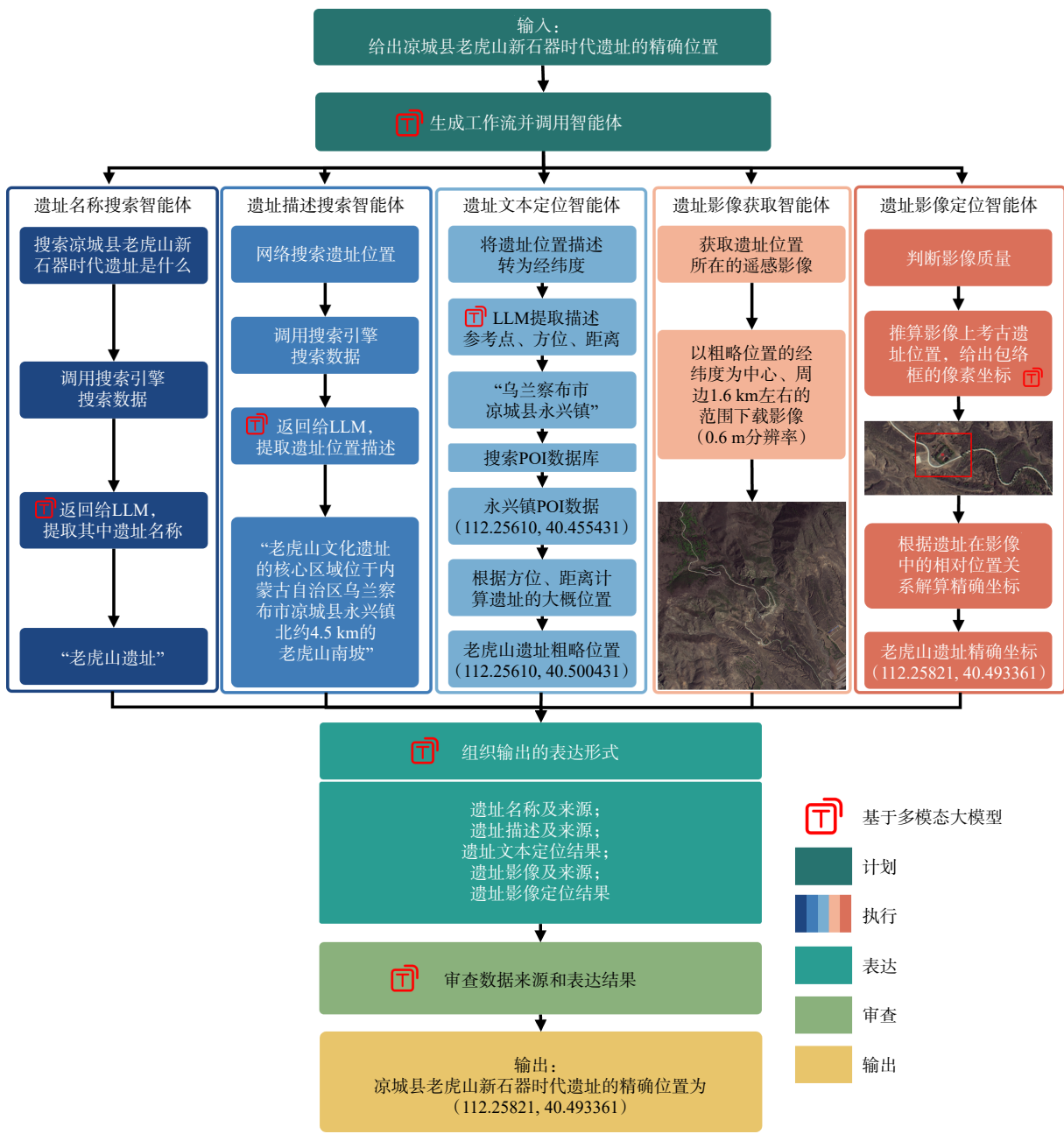


图3 以老虎山遗址为例的考古遗址精确定位

利用 LLM 进行语义聚类与权重排序, 筛选出官方发布且学界公认的标准名称。为提升准确率, 采用多轮思维链(chain-of-thought, COT)提示<sup>[15]</sup>, 先解析名称间的语义关联和使用场景, 再生成最终的标准化结果及其可追溯文献。

### 1.3 遗址描述搜索智能体

考古文献中常包含对遗址位置的文字描述, 是实现精确定位的重要线索。本研究设计“遗址描述搜索智能体”, 以名称搜索结果输入(图 3), 自动检索并解析相关文本, 提取空间信息要素。其主要任务是识别目标对象(遗址区、发掘区等)、参照对象(村镇、山丘、河流等)及其空间关系(方向、距离、拓扑等), 从而形成可用于坐标推理的描述单元。

本研究利用 LLM 结合提示策略, 对考古报告、地方志等材料进行处理, 自动提取与定位相关的语句, 包括参照物、方向和距离等空间特征。由于自然语言存在模糊性且 LLM 可能产生“幻觉”, 抽取结果多为自然语言而非结构化数据。例如“凉城县永兴镇北约 4.5 km 的老虎山南坡”这一表述, 仍需进一步由“遗址文本定位智能体”判断参照物的有效性。通过多智能体协同机制, 系统能够提升空间描述解析的准确性与鲁棒性, 为遗址数字化定位奠定基础。

### 1.4 遗址文本定位智能体

文本定位智能体基于结构化空间描述, 将参照物映射到兴趣点(point of interest, POI)数据库中。LLM 的提示语设计上, 优先匹配公共 POI 数据库中的精确地物; 若缺失或模糊, 则回退至行政区中心点或历史地图数据。例如, 当“老虎山”缺乏精确坐标时, 系统可调用“凉城县永新镇”POI, 并将“北约 4.5 km”转换为经纬度偏移量, 推算遗址坐标。具体计算中, 系统将文本中的方向与距离转化为笛卡尔投影偏移量, 结合参照物坐标生成遗址的初步位置。为降低 LLM “幻觉”风险, 系统在每一步均生成溯源链路, 记录参照物来源及推算公式, 以保证结果的可追溯性。

### 1.5 遗址影像获取智能体

获得遗址的初步坐标后, “影像获取智能体”以其为中心构建约 1.6 km 半径缓冲区, 并通过“天地图”API 获取该范围的 18 级瓦片影像, 天地图仅支持瓦片编号调用, 本研究利用经纬度与瓦片编号的转换公式完成编号计算:

$$\begin{cases} x_{\text{tile}} = \left\lfloor \frac{\lambda + 180}{360} \cdot 2^z \right\rfloor \\ y_{\text{tile}} = \left\lfloor \left( 1 - \ln \left[ \tan \left( \frac{\pi}{180} \cdot \phi \right) + \frac{1}{\cos \left( \frac{\pi}{180} \cdot \phi \right)} \right] / \pi \right) \cdot 2^{z-1} \right\rfloor \end{cases} \quad (1)$$

式中,  $\lambda$  为十进制表示的经度;  $\phi$  为十进制表示的纬度;  $z$  为瓦片等级, 在本实验中固定为 18。

完成转换后, 通过编号将下载的切片拼接为完整影像。由于天地图影像多为 2023 年后的数据, 能够保证考古遗址定位对影像现势性的需求。

### 1.6 遗址影像定位智能体

“遗址影像定位智能体”以多模态大模型为核心, 流程包括影像质量评估与遗址识别 2 阶段。输入数据涵盖拼接后的遥感影像、遗址名称及由 LLM 抽取的空间描述信息(如“永兴镇北约 4.5 km 的老虎山南坡”)。在质量评估阶段, 系统自动检测云层、色彩异常或对比度不足等问题, 以剔除或修正低质量影像。随后进入识别阶段, 智能体在结构化提示引导下判断影像中是否存在目标遗址, 并输出置信度、空间位置描述及候选区域的像素级包围框(bounding box, BBOX)。当存在多个疑似区域时, 模型会逐一评估可信度并选出最优结果。最终, 系统基于遥感影像的地理参考, 将 BBOX 中心像素转换为地理坐标, 实现对遗址的精准定位。该过程融合文本与影像双重线索, 提升了定位的精度、稳定性与可解释性, 具体过程如表 1 所示。

## 2 实验过程

为验证所提多智能体协同框架的适用性与有效性, 本文选取老虎山、三星堆和辽中京遗址为实验对象, 并依据表 2 所示伪代码逻辑, 开展基于真实考古数据的自动化定位试验。

### 2.1 实验对象与数据准备

所选 3 处遗址分别代表新石器、青铜和辽代的重要考古遗存, 既具历史分期差异, 又在空间分布上差异显著, 可用于全面评估框架的泛化性能。首先, 以真实遗址名称为检索词, 通过“遗址名称搜索智能体”和“遗址描述搜索智能体”调用必应 API 获取网页文本, 用于空间描述提取。其次, “遗址文本定位智能



表 1 考古遗址遥感识别任务的结构化提示词模板

|                                                                                                                |
|----------------------------------------------------------------------------------------------------------------|
| 外部信息: 遗址名称/遥感影像/位置描述信息                                                                                         |
| 期望输出: 置信度/位置描述/候选区域包围框/判断依据                                                                                    |
| 提示词 遥感影像目标定位:                                                                                                  |
| 1: 角色定义:                                                                                                       |
| 你是一名资深考古专家,长期从事遥感考古工作,熟悉中国不同类型考古遗址在遥感影像中的表现特征(如夯土结构、环壕痕迹、异常色差、规则扰动、地貌突起、土壤压实等)。你能够结合影像特征与已知遗址的空间描述信息进行专业分析与比对。 |
| 2: 信息输入: 现在请根据以下信息判断遥感影像中是否存在指定考古遗址                                                                            |
| 4: 遗址名称: 辽中京遗址大明塔基座                                                                                            |
| 5: 遥感影像: 请查看图像输入                                                                                               |
| 位置描述信息: 该遗址位于“内蒙古自治区乌兰察布市凉城县永兴镇北约4.5 km的老虎山南坡”。                                                                |
| 6: 任务要求:                                                                                                       |
| 请判断影像中是否可能存在该遗址相关迹象,并从以下等级中选择其一:                                                                               |
| ‘不可能/几乎不可能/有一定可能/高概率存在/一定存在’                                                                                   |
| 若可能性为“有一定可能”及以上,请继续输出以下内容:                                                                                     |
| 空间位置描述: 用自然语言描述疑似遗址区域在图像中的位置(例如:“影像中部偏北有一片规则的高反射区域”。                                                           |
| 9: 像素坐标包围框(BBOX): 请以[xmin, ymin, xmax, ymax]格式标出该区域的外接矩形框(基于影像像素坐标)。                                           |
| 判断依据: 简要说明你判断该区域为遗址的专业理由(例如:“该区域呈现典型夯土台地特征,位置与描述高度一致”。                                                         |
| 如未发现任何可疑迹象,请直接返回:“未在影像上发现考古遗址”。                                                                                |
| 10: 若存在多个候选区域,请编号并逐一列出,并根据遗址描述给出你认为最可能的遗址位置。                                                                   |
| 11:                                                                                                            |

体”接入高德地图 API<sup>[16]</sup>,在 POI 库中检索参照物,实现文本向地理实体的初步映射;同时,“遗址影像获取智能体”调用天地图 API<sup>[17]</sup>获取研究区 18 级遥感影像。根据式(2),在 Web Mercator 下该级别分辨率约为 0.6 m,可清晰识别遗址常见地表标志,并因影像多为 2023 年后获取,保证了时间现势性:

$$GSD = 2 \times R \times \pi / (px \times 2^z) \quad (2)$$

式中,GSD(ground sampling distance)为地面采样间距,即空间分辨率; $R=6378137$ ,为以米为单位的地球半径(基于 WGS84 参考椭球体); $z$  为瓦片等级,本研究中固定为 18; $px$  为在指定瓦片等级下单个瓦片的像素大小,在天地图中,18 级瓦片大小为 256 像素。

在定位阶段,“遗址影像定位智能体”调用多模态

表 2 多智能体驱动的考古遗址精确定位算法

|                                                |
|------------------------------------------------|
| 输入: 遗址名称或初步描述信息InputText                       |
| 输出: 遗址精确坐标LocationResult                       |
| 算法 遗址精确定位(InputText):                          |
| 1: 计划智能体(InputText)→TaskPlan                   |
| 2: 遗址名称搜索智能体(TaskPlan)→SiteName                |
| 3: 遗址描述搜索智能体(SiteName)→Description             |
| 4: 遗址文本定位智能体(Description)→POIList              |
| 5: 遗址影像获取智能体(POIList)→Images                   |
| 6: 遗址影像定位智能体(Images)→Location                  |
| 7: 表达智能体(Location, Description)→LocationResult |
| 8: 审查智能体(LocationResult)→Validity              |
| 9: 若Validity为True,则:                           |
| 10: 返回 LocationResult                          |
| 11: 否则:                                        |
| 12: 返回“定位失败”                                   |
| 结束算法                                           |

大模型 GPT-4o<sup>[18]</sup>,结合天地图真彩色合成影像、遗址名称与空间描述进行识别,输出候选区域的包围框与置信度,并通过排序选出最优结果。所有文本解析任务由 ChatGPT<sup>[18]</sup>完成,包括空间信息抽取与语义判断。整体流程形成“名称检索-描述提取-参照物定位-影像分析”的多智能体协同路径,实现文本与影像信息融合,支撑遗址的自动化精准定位。

2.2 多智能体框架部署

为实现考古遗址从非结构化文本到精确坐标的全自动定位,本文构建了基于 PEER 架构的多智能体协同系统。各智能体具备自主工具调用能力,可按任务目标动态调用搜索引擎、地图 API、遥感影像服务和多模态模型,实现异步调度与跨模态数据融合。

系统流程分为“任务规划-联合执行-结构化表达-质量审查”4 阶段,包含规划智能体、5 个执行类智能体、表达智能体和审查智能体。规划智能体为核心调度模块,解析输入任务并生成 JSON 格式的执行计划,明确子任务顺序、所需工具与数据依赖。例如,对于“永兴镇北约 4.5 km 的老虎山南坡”,系统自动调用“名称搜索-描述抽取-文本定位-影像获取-影像识别”等模块,逐步完成语义解析与坐标推算。

5 个执行类智能体负责遗址信息抽取、文本空间推理与影像识别等关键子任务,并依执行计划顺序调用外部工具。表达智能体则将结果整合为结构化对

象,内容涵盖遗址名称、文本描述、参照物坐标、影像信息与影像识别结果,确保输出的规范化与可溯源性。

最后,审查智能体对结果进行结构化与逻辑一致性校验。其检查要点包括:文本与影像描述是否一致,坐标是否在影像范围内,BBOX 是否与地貌匹配,文本与影像定位的欧氏距离是否在 2 km 阈值内,以及输出结构是否完整。为提升稳健性,系统采用“双轮采样+多数投票”机制,如结果不一致,则回溯至表达阶段重新生成候选区域。审查日志与判断结果同步记录,保证定位流程的可追溯性与可解释性。

2.3 考古遗址精确定位流程执行

在完成框架部署与调试后,系统可通过输入遗址名称或简要描述自动启动定位流程。首先,“遗址名称搜索智能体”和“遗址描述搜索智能体”协同确认遗址名称并提取空间描述;随后,“遗址文本定位智能体”解析参照物与空间关系,并利用高德地图 POI 库推理候选区域;接着,“遗址影像获取智能体”调用天地图 API 下载对应区域的高分影像;最后,“遗址影像定位智能体”分析影像并结合空间先验输出最高置信度的坐标。该流程实现了文本抽取、空间推理与影像识别的跨模态联动,有效提升了自动定位的准确性与

智能化水平。

2.4 成果记录与误差评估

实验将系统输出的遗址坐标与权威考古资料中的标注比对,采用欧氏距离量化定位偏差,并统计不同遗址样本的误差表现,以评估整体精度与稳定性。同时,对各智能体调用的数据源(如网页检索量、POI 候选数、影像下载量)及平均处理时间进行记录与分析,用于识别影响误差的关键环节并指导性能优化,确保系统在精度与效率之间实现平衡。

3 结果与分析

为验证多智能体协同定位系统的效果,本文以老虎山、三星堆和辽中京遗址为样本开展自动化试验,并将系统输出坐标与权威资料实测位置对比,从定量分析和定性分析两方面进行评估。

3.1 定量分析

表 3 展示了 3 处遗址的系统定位结果与实测坐标的误差。结果显示,大部分偏差控制在 0.1~2 km,整体精度可满足遗址复查、保护区划定与重定位等应用需求。

表 3 遗址定位结果与实测位置对比

| 遗址名称  | 实测坐标(WGS84)                          | 系统推算坐标(WGS84)         | 误差距离/km |
|-------|--------------------------------------|-----------------------|---------|
| 老虎山遗址 | 112.25831, 40.495455 <sup>[19]</sup> | 112.25821, 40.493361  | 0.23    |
| 三星堆遗址 | 104.199791, 30.99365 <sup>[20]</sup> | 104.194793, 31.007628 | 1.31    |
| 辽中京遗址 | 119.155434, 41.569743 <sup>[4]</sup> | 119.155995, 41.570031 | 0.06    |

其中,辽中京遗址以大明塔为标志,范围集中,误差仅 60 m;老虎山遗址因遗迹点分布广,系统定位落在城址中心而非文化遗址中心,偏差约 230 m;三星堆遗址误差最大(1.31 km),主要因网络资料中地名混用及公开街道地图(openstreetmap, OSM)默认位置<sup>[21]</sup>与发掘区不一致所致。

3.2 定性分析

老虎山遗址发掘报告提供了明确的空间描述(如“永兴乡政府南 4.5 km”“老虎山南坡”)及地形图。智能体成功提取“永兴乡—4.5 km—老虎山南坡”等要素,在文本定位中因“老虎山”缺失 POI 而回退至“永兴乡”,并自动识别其为历史地名,LLM 更新“永兴乡”为“永兴镇”。结合 POI 坐标与方位信息,系统生

成缓冲区并在高分辨率影像中识别出疑似遗址土台,最终与实测位置偏差约 230 m,具体如图 4 所示。

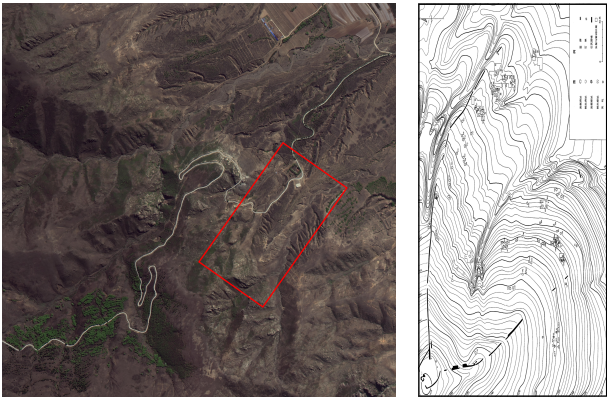


图 4 智能体搜索结果影像与发掘报告中的遗址地形图对比



图4显示智能体结果与发掘报告地形图高度一致,尤其在城址遗迹聚集区表现出显著匹配性。辽中京遗址(图5(a))中,系统定位与大明塔基座偏差仅60 m,表明影像定位在小尺度遗址上具有高精度。三星堆案例(图5(b))中,系统因OSM数据默认位置与发掘区不同而出现约1.3 km偏移,但结果仍落在遗址保护范围内,验证了框架在复杂遗址群中的适应性。



(a) 实测坐标为辽中京遗址大明塔基座 (b) 实测坐标为三星堆遗址一号坑

图5 遗址实测坐标与推算坐标对比

总体而言,该框架在多源数据支撑下实现了较高精度与稳定性,空间误差均在考古调查允许阈值内,同时有效控制了计算复杂度。

## 4 结果与讨论

本文提出的多智能体框架在老虎山、三星堆和辽中京遗址实验中表现出较高的自动化定位能力,定位误差多在十米至千米之间,基本满足重定位、复查和规划需求。框架通过任务分解与智能体协同提升了效率与稳定性,LLM在非结构化文本解析中发挥了核心作用,并与遥感影像分析结合形成“语言—视觉—空间”闭环,提高了结果的可靠性和解释性。然而,系统仍受文本模糊、地名变迁、地图与影像时效性差异等因素影响,需进一步优化。

### 4.1 误差来源分析

多智能体考古遗址定位系统在实验中表现出较强的自动识别与推理能力,但与权威资料对比仍存在多类系统性误差。首先,历史文献缺乏标准化空间描述,常无明确坐标或方向,或所依赖的地理参照物已难以对应现代实体,致使空间语义难以准确转换。模糊表述如“西侧”“东南约200 m”在缺乏参照点时易

致方向歧义,扩大不确定性。其次,地图数据库POI存在更新滞后与标注偏差,例如“永兴乡”已更新为“永兴镇”,若系统未能识别地名变迁,将直接影响推算精度。尽管大模型具备语义纠错能力,但稳定性有限,仍可能生成错误信息。此外,系统缺乏对网络坐标信息的有效筛查,易直接采用来源不明或代表性不足的数据。例如,三星堆遗址坐标若取自维基百科的博物馆位置,而非实际发掘区,便会造成显著偏差。

遥感影像的不确定性亦是重要因素。部分遗址因农田覆盖或城市扩张,地貌特征不再明显,导致智能体难以识别或结果置信度不足,影像定位智能体主要依赖表层纹理判别,缺乏对遗址功能与类型的语义理解。当地物特征不突出或存在多目标干扰时,难以准确选择候选点。本研究虽通过设置1.6 km缓冲区部分缓解不确定性,但在遗址密集区仍可能误判。

综上所述,系统误差主要源自文本描述模糊、POI数据时效性不足、遥感图像纹理弱化、网络信息未经验证及智能体判别能力受限等多个方面。

### 4.2 模型输出一致性与幻觉控制策略

在遗址精确定位任务中,LLM虽能理解并抽取复杂空间语义,但在处理高自由度文本时仍会出现“幻觉”,表现为与上下文不符、结构缺失或语义冲突,若不加控制将削弱系统的稳定性与可用性。为此,本研究构建了多维度输出控制机制,以提升结果一致性与可验证性。实验表明,该机制能有效抑制非法坐标生成、空间描述偏差和结果缺项。具体而言,系统在输入中联合引入遥感影像、遗址名称与自动抽取的空间描述,强化了推理过程的上下文约束,显著提升了区域定位的聚焦能力。

在输出环节,审查智能体对结果进行结构化校验,是保证稳定性的关键。其依据预设语义逻辑和字段规则,自动识别位置描述与图斑不符、字段缺失或坐标偏移过大的问题,并可在无人工干预条件下完成异常剔除与迭代重构,确保结果可信与结构完整。

需要强调的是,尽管系统通过结构化提示、多模态输入、审查机制与回溯流程实现了幻觉的初步控制,但在文本极度模糊、影像干扰严重或存在多目标竞争时,仍可能产生误判。因此,提升模型对地貌特征的专业理解,以及引入基于证据链的可验证推理逻辑,将是未来增强系统可靠性的重点方向。

### 4.3 改进方向

为克服多智能体框架在空间语义理解、地名时效性、影像识别与结果可解释性方面的不足,未来可通过多源多模态数据与增强推理机制提升性能。在空间推断层面,可在文本分析基础上融合历史地图与现代影像的配准对比,丰富空间参照信息,增强复杂地理关系理解;同时利用数字高程模型的地形信息,在参照物缺失或描述不完整时辅助实现基于地貌结构的定位。

在遥感影像方面,当前系统主要依赖 RGB 图像,未充分发挥多光谱优势。未来可引入近红外、短波红外等多波段数据,增强对夯土遗迹、植被扰动、土壤湿度异常等考古相关特征的识别。当大模型无法直接处理非 RGB 数据时,可通过特征合成、伪彩色渲染或辅助监督将其显性化,从而提升在复杂背景下的判读精度与稳定性<sup>[22-23]</sup>。地名匹配方面,可构建时态地理知识库,整合历史地名、行政区划演化与现代地物的动态对应,实现基于时间标签的实体检索,以减少地名变更、语义漂移和边界调整造成的偏差。

此外,引入不确定性量化与决策支持模块,可通过对文本、POI 分布与影像识别结果的模糊性进行概率建模,输出多候选区域及置信区间,为考古人员提供分层次、可验证的定位参考,并推动系统实现主动反馈与自我优化,朝高可靠性与可控性方向演进。

## 5 结论

本文提出了一种基于 LLM 的多智能体协同定位方法,构建了涵盖资料检索、信息抽取、位置搜索、影像识别与坐标推算的自动化流程。通过老虎山、三星堆和辽中京遗址实证表明,该系统可在自动考古遗址的精确推理与定位,误差普遍控制在十米至千米,满足考古应用需求。

研究结果显示,结合 LLM 的语义理解、多智能体的任务协作与遥感影像分析,可显著提升考古遗址重定位的自动化与规模化能力,减少对传统田野调查和人工解读的依赖,推动考古信息处理智能化。同时也发现,在文献缺失、参照物变迁或地表改造严重的情境下,仍会出现误差累积。未来可在多源数据融合、古今地名演变、深度影像识别与不确定性建模等方面

进一步优化,以增强系统的准确性、适应性和可解释性。

**致谢:**感谢“坤元”地理科学大模型团队对本论文资料整理工作的贡献与支持。

### 参考文献(References)

- [1] 耿同, 杨瑞霞, 杨树刚. 早期考古发掘遗址重定位方法研究[J]. [中国科学院大学学报](#), 2020, 37(3): 317-323.
- [2] 刘方, 田苗, 吕杨. 考古遗址位置图的制图规范[J]. [草原文物](#), 2019(1): 95-102.
- [3] 晁春浩, 李毅. 元上都遗址田野考古测绘[J]. [北京测绘](#), 2010, 24(4): 30-32,71.
- [4] 任冠. 辽中京道城址的考古学观察[J]. [故宫博物院院刊](#), 2016(2): 66-76,161.
- [5] 刘建国. 地理信息系统在考古研究中的应用[J]. [社会科学管理与评论](#), 2006(1): 79-86.
- [6] 李国杰. DeepSeek引发的AI发展路径思考[J]. [科技导报](#), 2025, 43(3): 14-19.
- [7] 任福继, 张彦如. 通用大模型演进路线[J]. [科技导报](#), 2024, 42(12): 44-50.
- [8] 吴文峻, 廖星创, 赵金琨. DeepSeek技术创新与通用人工智能发展趋势[J]. [科技导报](#), 2025, 43(6): 14-20.
- [9] Zhang T, Ma L B, Cheng S, et al. Automatic prompt design via particle swarm optimization driven LLM for efficient medical information extraction[J]. [Swarm and Evolutionary Computation](#), 2025, 95: 101922.
- [10] 邓佳文, 任福继. 2023年生成式AI大模型发展热点回眸[J]. [科技导报](#), 2024, 42(1): 266-285.
- [11] Luo H, Feng X B, Du B, et al. A multimodal feature fusion network for building extraction with very high-resolution remote sensing image and LiDAR data[J]. [IEEE Transactions on Geoscience and Remote Sensing](#), 2024, 62: 1-19.
- [12] Chen Z, Liu K, Wang Q, et al. Agent-flan: Designing data and methods of effective agent tuning for large language models[J]. 2024, arXiv preprint arXiv:2403.12881.
- [13] Bachagha N, Wang X Y, Luo L, et al. Remote sensing and GIS techniques for reconstructing the military fort system on the Roman boundary (Tunisian section) and identifying archaeological sites[J]. [Remote Sensing of Environment](#), 2020, 236: 111418.
- [14] Argyrou A, Agapiou A. A review of artificial intelligence and remote sensing for archaeological research[J]. [Remote Sensing](#), 2022, 14(23): 6000.
- [15] Wei J, Wang X, Schuurmans D, et al. Chain-of-thought



- prompting elicits reasoning in large language models[J]. Advances in Neural Information Processing Systems, 2022, 35: 24824–24837.
- [16] GaoDe. Agreements and Declarations–GaoDe Map[EB/OL]. [2025–05–01]. <https://map.amap.com/doc/serviceitem.html>.
- [17] 黄蔚. 国家地理信息公共服务平台天地图[J]. 卫星应用, 2014(1): 41–44.
- [18] Achiam J, Adler S, Agarwal S, et al. Gpt-4 technical report[J]. 2023, arXiv preprint arXiv:2303.08774.
- [19] 田雨. 岱海考古(一)——老虎山文化遗址发掘报告集[M]. 北京: 科学出版社, 2000: 198.
- [20] 陈显丹. 广汉三星堆遗址发掘概况, 初步分期——兼论“早蜀文化”的特征及其发展[J]. 南方民族考古, 1990, 2: 213–231.
- [21] Foundation O. Terms of use–OpenStreetMap foundation [EB/OL]. [2025–05–01]. [https://osmfoundation.org/w/index.php?title=Terms\\_of\\_Use](https://osmfoundation.org/w/index.php?title=Terms_of_Use).
- [22] Zhang C H, Wang S. Good at captioning, bad at counting: Benchmarking GPT-4V on Earth observation data[C]// Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). IEEE, 2024: 7839–7849.
- [23] Kuckreja K, Danish M S, Naseer M, et al. GeoChat: Grounded large vision–language model for remote sensing [C]//Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2024: 27831–27840.

## Accurate archaeological site localization using large language models

LU Yuxiang<sup>1,2</sup>, SHEN Jing<sup>2,3\*</sup>, JIANG Hou<sup>2</sup>, LIU Tang<sup>2</sup>

1. School of Artificial Intelligence, China University of Geosciences (Beijing), Beijing 100083, China

2. State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China

3. University of Chinese Academy of Sciences, Beijing 100049, China

**Abstract** Due to technological limitations in early archaeological work, a large number of sites were documented only with vague textual descriptions, lacking precise geographic coordinates. This has posed significant challenges for subsequent archaeological investigations, site protection, and research. Traditional field survey methods are time-consuming and labor-intensive, making them unsuitable for large-scale or high-throughput site localization tasks. To address this issue, this study proposes an intelligent localization framework for archaeological sites based on large language models (LLMs). By designing tailored natural language prompts, the framework guides LLMs to automatically extract geographic descriptions—such as landmarks, directions, and distances—from online archaeological literature and records. It then integrates this information with high-resolution satellite imagery analysis to infer and calculate the precise geographic coordinates of the target site. This pipeline achieves full-process automation, covering text-based information extraction, remote sensing data retrieval, and spatial reasoning, thereby significantly improving the efficiency and intelligence level of site localization. Validation experiments conducted on the Laohushan Site, Sanxingdui Site, and Liao Zhongjing Site show that the deviation between the automatically inferred locations and actual surveyed coordinates is within one kilometer, with the best accuracy reaching approximately 10 meters—meeting the basic precision requirements of archaeological applications. Localization errors primarily stem from ambiguities in textual descriptions and uncertainties in image interpretation. Compared with traditional manual methods, this approach greatly reduces labor costs and offers enhanced scalability and application potential. The proposed method provides a novel technical pathway and tool support for the digitalization of archaeological information, site conservation, and further research.

**Keywords** geographic intelligence; archaeological sites; precise localization; remote sensing image information extraction; large language model ●



(责任编辑 赵庆圆)