

<https://doi.org/10.3799/dqkx.2025.203>



# 多源融合的“岩石-矿物-光谱”知识图谱构建及应用

韩伟<sup>1</sup>, 张照坤<sup>1</sup>, 雷辛亚<sup>2</sup>, 王力哲<sup>1\*</sup>

1. 中国地质大学计算机学院, 湖北武汉 430078

2. 武汉网络安全大学筹建办公室人工智能安全学院, 湖北武汉 430040

**摘要:** 在全球资源竞争加剧的背景下, 地质资源勘查已成为多国战略布局的核心焦点, 我国要提升战略性矿产的国内保障水平, 强化地质资源勘查, 提高资源保障能力。随着大数据与人工智能技术的深度渗透, 岩石学、光谱学与矿物学领域已积累海量多源异构数据, 然而当前存在多源异构数据融合不足、知识关联薄弱、分类体系难以对齐等问题, 导致了现有知识难以高效服务于资源勘查等地质应用。因此, 提出“岩石-矿物-光谱”知识图谱 (rock mineral spectrum knowledge graph, RMS-KG) 的系统化构建方案。通过一体化编码遥感影像、光谱曲线、矿物特征及地质文献等多来源知识, 采用自顶向下与自底向上相结合的领域本体建模策略, 构建涵盖岩石分类、矿物特征和光谱特征等核心概念的知识图谱模式层。随后, 融合深度学习与语义解析技术, 从结构化数据库、半结构化报告及非结构化文献中抽取相关实体、属性及关系, 进而实现知识融合与语义关联, 并在图数据库中完成知识融合、可视化查询与动态推理。RMS-KG 包含各类实体、实体属性及关系超数十万个, 涵盖岩矿类型超 1000 种, 形成覆盖“岩石-矿物-光谱”的统一模式层与数据层, 支持光谱指纹到矿物、岩石的映射以及基于矿物组合的成矿类型推理, 并在“光谱导向矿物识别”和“成矿类型推理”两类场景中验证了有效性与可解释性。RMS-KG 为岩矿识别与找矿预测提供可复用的知识底座与推理能力, 提升地质知识的可检索、可计算与可复用性, 为地质大数据的知识化表达与智能应用提供可推广范式。

**关键词:** 知识图谱; 岩石学; 光谱学; 矿物学; 多源数据融合。

中图分类号: P583

文章编号: 1000-2383(2026)02-634-13

收稿日期: 2025-03-15

## Construction and Application of the “Rock-Mineral-Spectrum” Knowledge Graph Based on Multi-Source Fusion

Han Wei<sup>1</sup>, Zhang Zhaokun<sup>1</sup>, Lei Xinya<sup>2</sup>, Wang Lizhe<sup>1\*</sup>

1. School of Computer Science, China University of Geosciences, Wuhan 430078, China

2. The Preparatory Office for Wuhan University of Cyber Security, School of Artificial Intelligence Security, Wuhan 430040, China

**Abstract:** Amid intensifying global competition for mineral resources, raising domestic security of strategic minerals requires higher-precision, more explainable exploration. Although petrology, spectroscopy, and mineralogy have amassed large volumes of heterogeneous data, limited cross-source fusion, weak semantic linkage, and misaligned taxonomies hinder their use in exploration. This paper proposes a systematic “Rock-Mineral-Spectrum” Knowledge Graph (RMS-KG) to address these gaps. We integrate remote-sensing imagery, reflectance spectra, mineral characteristics, and geological literature using a hybrid ontology

**基金项目:** 国家自然科学基金项目 (Nos. U21A2013, 42201415).

**作者简介:** 韩伟 (1993-), 男, 副研究员, 硕士生导师, 主要从事地质遥感解译与地质遥感知识图谱构建。ORCID: 0000-0003-3882-1616. E-mail: weihan@cug.edu.cn

\* **通讯作者:** 王力哲, ORCID: 0000-0003-2766-0845. E-mail: lizhe.wang@foxmail.com

**引用格式:** 韩伟, 张照坤, 雷辛亚, 王力哲, 2026. 多源融合的“岩石-矿物-光谱”知识图谱构建及应用. 地球科学, 51(2): 634-646.

**Citation:** Han Wei, Zhang Zhaokun, Lei Xinya, Wang Lizhe, 2026. Construction and Application of the “Rock-Mineral-Spectrum” Knowledge Graph Based on Multi-Source Fusion. *Earth Science*, 51(2): 634-646.

approach that combines top-down domain modeling with bottom-up data construction. The schema covers core concepts in rock taxonomy, mineral attributes, and spectral features. Deep learning and semantic parsing extract entities, attributes, and relations from structured databases, semi-structured reports, and unstructured texts; knowledge is then fused in a graph database to enable semantic linkage, visual querying, and dynamic reasoning. RMS-KG contains on the order of tens of thousands of nodes and edges and includes more than 1,000 rock-mineral types. It unifies the “rock-mineral-spectrum” semantics, supports mapping spectral fingerprints to minerals and rocks, and enables metallogenic-type inference from mineral assemblages. Two application scenarios, “spectrum-guided mineral identification” and “metallogenic-type inference”, demonstrate its effectiveness and interpretability. RMS-KG provides a reusable knowledge substrate and reasoning capability for rock-mineral recognition and prospecting, improving the retrievability, computability, and reusability of geological knowledge and offering a generalizable paradigm for knowledge-centric geological AI.

**Key words:** knowledge graph; petrology; spectroscopy; mineralogy; multi-source data fusion.

## 0 引言

在全球资源竞争持续加剧的形势下,矿产资源的安全稳定供给关乎国家战略与可持续发展能力.我国正在推进新一轮找矿突破行动,旨在提升战略性矿产的国内保障水平与国际话语权(金晓斌等,2024).在矿产勘查与相关地质应用中,精准识别岩石与矿物、判定成矿类型并预测矿体空间位置,依赖对岩石物化性质、矿物光谱指纹与区域构造背景的综合分析.传统方法侧重专家经验与单一数据源,难以满足对精度、效率与深度的现实需求.因此,地质勘查正由“数据-知识”协同推动从经验驱动向智能化与精准化转型,成为地质科学的重要方向(翟明国等,2018).这一转型的关键在于以领域知识为纽带,实现多源信息的结构化表达与可解释推理,为知识表达与图谱建模奠定问题基础.

基于转型的诉求,领域知识被视为连接多源数据与勘查决策的关键纽带.以空间数据挖掘与知识发现为视角,结构化的知识表达能够将观测数据转化为可计算与可推理的语义网络,进而揭示隐含的空间关联与时空演化机理,为复杂地质过程建模提供逻辑支撑(李德仁等,2001).知识图谱作为语义知识库,可在实体与关系层面打通数据壁垒,提升多源信息的融合分析能力(官赛萍等,2018).虽然知识图谱技术效果显著,但当前仍面临三项核心挑战.其一,多源数据高度异构,高光谱影像、实验室光谱与地学文本在格式、语义与时空尺度上差异显著,融合分析难度大(Shen *et al.*, 2024).其二,矿物共生、蚀变特征与成矿环境之间关系复杂且不确定,传统表示方法难以支撑有效推理.其三,海量文献以非结构化文本为主,自动化知识抽取与标准化表达存在瓶颈,限制从数据到知识再到决策的转化

效率.

针对这些挑战,近年来地质领域知识图谱研究呈多维发展趋势,围绕本体构建、知识融合与智能应用形成了较为系统的技术谱系(周成虎等,2021; Ma, 2022; 牛凤桂等,2024).在模式层方面,基于基本形式化本体(basic formal ontology, BFO)开发了地质学核心本体(geo-core ontology, GCO)明确了地质对象、物质组成与地质过程等核心概念,为专业子域本体扩展提供了稳定中间层(Garcia *et al.*, 2020).地质科学本体(geo-science ontology, GSO)构建了“基础层-地质层-扩展模块”的3层架构并融合北美地质图数据模型(North American geologic map data model, NADM)与GeoSciML标准,实现跨尺度地质实体的语义关联与统一表达(Brodaric and Richard, 2021).面向大规模协同构建,深时数字地球计划(deep-time digital Earth, DDE)计划联合多学科工作组形成跨领域知识图谱与协同编辑框架,并在地质时间本体上实现了百万年尺度的时间分辨率,展示了面向地学大数据的群智化标注与冲突消解能力(Wang *et al.*, 2021; Hu *et al.*, 2023).在此框架支撑下,面向“岩石-矿物-光谱”的数据层与抽取方法逐步成熟,为数据获取与关系建模提供了技术抓手.

在数据层与信息抽取方面,面向“岩石-矿物-光谱”的关键要素识别与关系构建不断完善.矿物与岩石领域的实体识别、关系抽取及三元组构建逐步成熟,典型做法结合深度序列标注与句法依存以提升术语识别与关系准确性,为专业知识图谱的高质量填充奠定了基础(张春菊等,2023).同时,行业侧在测井与油气识别中引入图嵌入与神经网络,实现知识图谱驱动的智能识别与检索,验证了“图谱+学习”的可行性与可迁移性(刘国强等,

2022). 在此基础上,知识图谱逐步服务于资源勘查应用场景,进一步检验其在找矿预测与证据耦合中的实际价值.

在资源勘查与矿业应用方向,知识图谱方法逐步用于矿物组合到成矿环境再到靶区推荐的链式推理. 铀矿成矿知识图谱通过大规模勘探报告的结构化关联,成功定位潜在靶区,体现了文本证据到空间预测的可追溯链路(Enkhsaikhan *et al.*, 2021). 面向斑岩铜等类型矿床,区域尺度图谱将地质实体、成矿关系与演化过程纳入统一框架,显著提升了找矿预测性能(周永章等, 2021). 在数据融合层面,地质图知识与地球化学数据的联动建模表明,知识嵌入与深度网络的协同可提升矿产预测精度(Chen *et al.*, 2024). 面向本体扩展,矿床多层次本体与时空本体库的构建为复杂成矿系统的层级化表达与知识推理提供了可操作路径(张前龙等, 2024). 此外,面向地层与矿体的跨源证据耦合,语义网络与向量化表征支持隐伏矿体规律识别与大规模地层关联分析(Qiu *et al.*, 2023). 这些实践表明,面向复杂地质过程的知识表达与推理已具备跨尺度与跨数据源的集成能力,并为其他地学场景的迁移提供了方法基础.

在工程地质与灾害知识方向,滑坡领域本体与基于 Neo4j 的可视化与智能检索验证了任务导向的知识组织与服务能力(许强等, 2023),半监督命名实体识别框架 Semi-GDNER 与 Geo-Disaster-NER 数据集实现了灾害位置、类型与原因等关键实体的高精度抽取,体现了领域文本到结构化知识的可迁移路径,地质灾害链知识图谱框架融合了灾害事件、地质环境与应急处置本体,基于时空推理实现 72 小时滑坡预警,并以汶川地震案例验证实用性(邱芹军等, 2023). 这些成果表明图谱在多源证据耦合与可解释推理方面具备潜力,同时也凸显了面向“岩石—矿物—光谱”的系统化建模仍有待深化.

基于上述问题与进展,本研究提出了一种面向多源数据融合驱动的“岩石—矿物—光谱”知识图谱构建方法. 针对地质领域“岩石—矿物—光谱”多源异构数据模态差异显著、跨模态语义关联薄弱所导致的知识图谱构建困难,提出了多模态语义对齐与不确定性融合建模方法. 具体而言,通过整合岩石遥感图像、光谱反射曲线、矿物特征等多源异构数据,设计了涵盖“岩石分类—矿物组成—光谱特

征”3个核心维度的地质知识本体模型,并结合深度学习与概率图模型构建不确定性实体关系抽取框架,能够有效缓解数据异构性与语义歧义问题,实现跨模态知识的高效融合与精准语义对齐. 此外,针对岩石—矿物演化关系具有隐含性强、属性不确定性高,传统知识表示方法难以有效支撑动态推理的问题,进一步提出了不确定性嵌入表示与隐式关系挖掘方法. 具体方法结合图神经网络与概率软逻辑,将矿物共生规律、岩石成因模式等专家知识以可量化置信度的形式编码为推理规则,并通过自监督学习机制优化知识图谱中隐含关系的嵌入表示,显著提升了知识图谱在成矿预测与矿床成因分析方面的动态推理能力.

## 1 多模态“岩石—矿物—光谱”知识图谱构建

知识图谱的构建方法通常包括自底向上、自顶向下及两者结合的方式(Zhao *et al.*, 2018). 自底向上的构建方法以数据为核心驱动,从海量多源异构数据中自动或半自动抽取知识,逐步累积形成知识图谱. 自顶向下的构建方法以领域知识为顶层框架,由专家团队先定义知识图谱的模式层,再基于模式层填充数据层,形成“概念→实例”的层次化知识体系. 由于岩石学、矿物学、光谱学所覆盖的知识面非常广泛、数据量大、语义关系多,故本文构建知识图谱采用自顶向下与自底向上结合的混合模式,通过领域本体的顶层设计与多源数据的底层知识抽取形成双向驱动,确保知识体系的完整性与时效性,具体构建方法流程图如图 1 所示.

本文基于已有的专家知识和国际岩石分类标准等领域知识(Le Bas and Streckeisen, 1991). 先定义知识图谱的模式层,包括概念层次、属性约束与推理规则等;再构建数据层,从遥感图像、光谱库、地质文献中抽取实体与关系,形成三元组;然后通过实体对齐与冲突消解,消除多源数据的语义歧义进行知识融合;最后基于 Neo4j 可视化存储图谱,嵌入领域规则实现动态推理. 在本章节中,着重讲述本文所构建知识图谱的模式层和数据层,在下一章节详细描述知识融合与知识存储.

### 1.1 模式层构建

知识图谱模式层是概念、属性、关系等集合,其核心作用是为多源地质数据提供统一的语义框架,确保知识结构的系统性与可扩展性. 本研究构建的

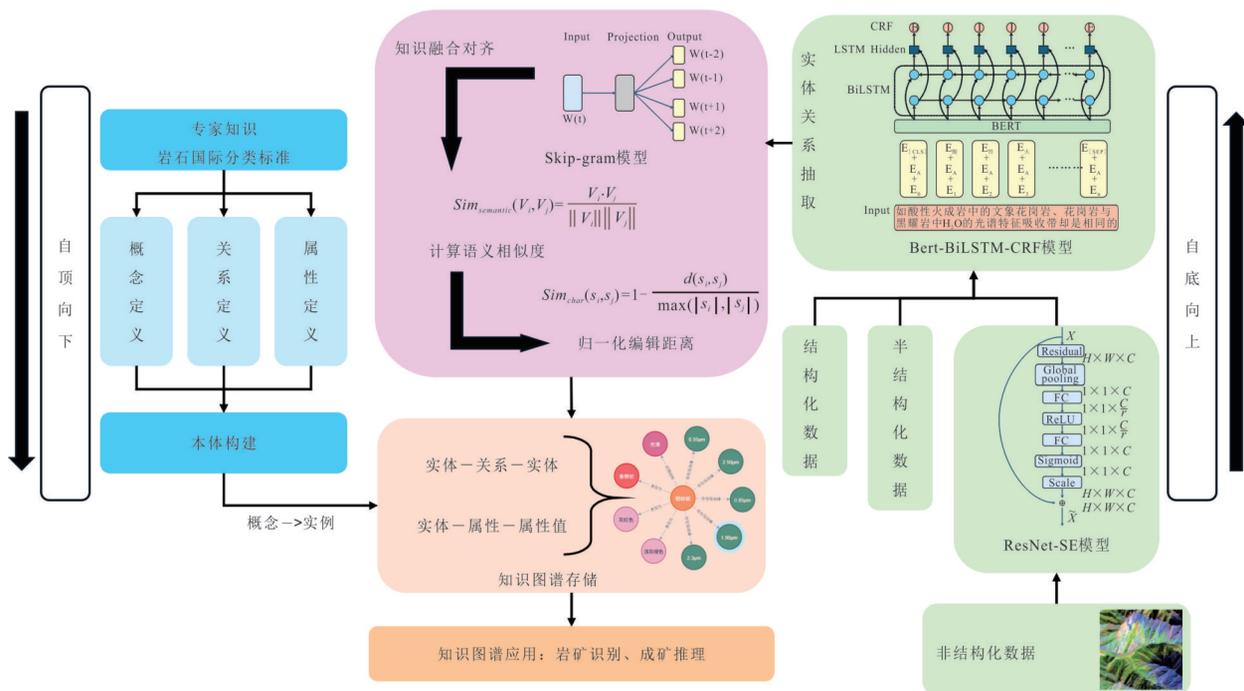


图1 “岩石-矿物-光谱”知识图谱构建流程

Fig. 1 Construction process of rock-mineral-spectral knowledge graph

表 1 模式层概念实体、关系及属性

Table 1 Concept of entities, relationships and attributes in the pattern layer

模式层	一级实体	关系	属性
岩石类型	岩浆岩	HAS_COLOR	颜色
		HAS_FORM	形态
		HAS_DRAINAGE	水系特征
	沉积岩	HAS_TEXTURE	纹理特征
		HAS_TERRAIN	地形地貌
	变质岩	HAS_LITHOLOGY	岩性分类
矿物类型	硅酸盐矿物	HAS_MINERAL	矿物特征
	硫酸盐矿物	HAS_ABS_PEAK	吸收峰波长
	硫化物矿物	HAS_REF_PEAK	反射峰波长
	氧化物矿物		
光谱特征	波长	ABS_PEAK_DTH	吸收峰深度
		ABS_PEAK_WTH	吸收峰宽度
成矿环境	斑岩型	MINERAL_CHAR	矿物特征
	热液型		
	岩浆型	ALTER_CHAR	蚀变特征
	沉积型	SPEC_MINERAL	标型矿物
	变质型		

“岩石-矿物-光谱”知识图谱 (rock mineral spectrum knowledge graph, RMS-KG) 模式层采用自顶向下建模方法, 通过领域专家知识定义核心概念与

关系, 展示部分实体、关系及属性, 如下表 1 所示。

RMS-KG 模式层以岩石类型、矿物类型、光谱特征与成矿环境为核心实体, 通过各自延伸的次级实体及属性关系, 构成层次清晰、逻辑严谨的知识网络结构。其中, 岩石类型实体包括岩浆岩、沉积岩和变质岩等一级实体分类, 岩浆岩进一步细分形成花岗岩、玄武岩、安山岩、英云闪长岩等次级实体, 而沉积岩细分形成砂岩、页岩、石灰岩等次级实体, 以及变质岩细分形成大理岩、片麻岩、石英岩等次级实体。岩石类型的属性涵盖颜色、纹理特征、形态及地形地貌特征等, 实体间的关联关系包括“颜色”、“形态”、“水系特征”、“纹理特征”、“地形地貌”、“岩性分类”以及“矿物特征”等。矿物类型实体包含硅酸盐矿物、碳酸盐矿物、硫化物矿物等一级分类, 其次级实体分别为石英、长石、辉石, 白云石、孔雀石、方解石以及绿矾、芒硝、重晶石等。矿物类型实体间的关系涵盖“颜色”、“存在吸收峰”、“存在反射峰”等特征。光谱特征实体主要包括波长实体, 具备“吸收峰深度”、“吸收峰宽度”等属性。此外, 成矿环境实体包括斑岩型、热液型与岩浆型等一级分类, 并包含“矿物特征”、“蚀变特征”及“标型矿物”等关联关系。

### 1.2 数据层构建

数据层构建是知识图谱的关键环节, 其主要目

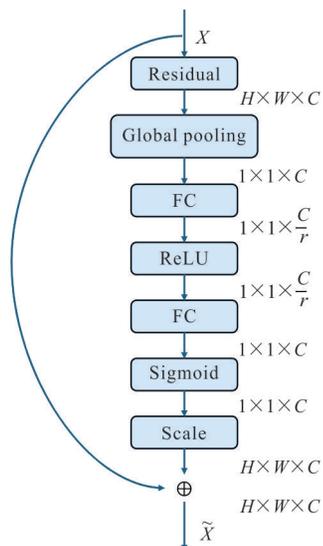


图2 ResNet-SE 模型结构

Fig. 2 Model structure of ResNet-SE

标是从多源异构数据中准确地提取实体、属性及关系,进而形成结构化的知识单元(刘峤等,2016).为实现这一目标,本研究综合利用岩石遥感影像、光谱反射曲线及地质文献文本3类数据,采用深度学习与半自动化方法进行跨模态知识抽取.针对地质数据类型的异构性与复杂性,设计了“图像—文本—光谱”多模态协同处理框架,通过跨模态特征级融合与语义关联策略,有效克服了传统单一方法在处理地质领域异构数据时的局限性,提升了知识抽取的准确性与可靠性.

对于岩石遥感图像特征提取,本文采用残差网络(ResNet)架构和引入注意力机制的SE模块,来提取岩石遥感图像中的特征(Kong *et al.*, 2021). ResNet通过跳跃连接解决梯度消失问题,其核心公式为:

$$H(x) = F(x) + x, \quad (1)$$

其中: $H(x)$ 为输出特征图; $F(x)$ 为卷积层输出; $x$ 为输入特征图. ResNet-50包含50层卷积层,通过瓶颈结构减少参数量,提升特征提取效率.引入挤压—激励(Squeeze-and-Excitation, SE)模块,通过通道注意力增强关键特征表达,模型结构如图2所示.

首先,对初始输入 $X$ 执行全局平均池化(Global Average Pooling)操作,将 $H \times W \times C$ 的三维特征图压缩为 $1 \times 1 \times C$ 的通道维度向量,实现对空间维度信息的全局聚合,获取每个通道的整体响应值:

$$z_c = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W x_{c,i,j}, \quad (2)$$

其中: $H$ 和 $W$ 为特征图尺寸; $x_{c,i,j}$ 为第 $C$ 个通道的像素值.接下来,压缩后的数据进入通道注意力计算阶段,通过两个全连接层(FC)构成的瓶颈结构生成通道权重.首先使用降维全连接层降低计算复杂度,通道数由 $C$ 降至 $C/r$ , $r$ 为压缩比,通常取16,引入ReLU激活函数增加非线性,然后通过升维全连接层恢复至 $C$ 个通道,最后经Sigmoid函数生成范围在 $[0,1]$ 的通道注意力权重向量:

$$s = \sigma[W_2 \delta(W_1 z)], \quad (3)$$

其中: $W_1$ 和 $W_2$ 为权重矩阵; $\delta$ 为ReLU激活函数; $\sigma$ 为Sigmoid函数.最终的特征重标定阶段,将生成的通道权重与原始特征图进行逐通道相乘,实现对各通道特征的自适应调整增强高响应通道的特征表达,抑制无关通道的冗余信息,从而提升模型对关键特征的聚焦能力:

$$\tilde{x}_{c,i,j} = x_{c,i,j} \times s_c. \quad (4)$$

针对地质文献语义的解析,本文主要是采用Bert-BiLSTM-CRF模型对文献中的实体和属性进行提取(谢腾等,2020). Bert预训练模型采用双向Transformer编码器,通过遮蔽语言建模(MLM)和下一句预测(NSP)任务预训练,目标函数为:

$$L = \mathcal{L}_{\text{MLM}} + \lambda \mathcal{L}_{\text{NSP}}, \quad (5)$$

其中: $\mathcal{L}_{\text{MLM}}$ 为遮蔽词预测损失, $\mathcal{L}_{\text{NSP}}$ 为句子关系预测损失, $\lambda$ 为平衡系数. BiLSTM层的双向长短期记忆网络捕捉上下文语义,隐藏层状态计算为:

$$\begin{cases} \vec{h}_t = \text{LSTM}(\vec{h}_{t-1}, x_t) \\ \overleftarrow{h}_t = \text{LSTM}(\overleftarrow{h}_{t+1}, x_t) \\ h_t = \vec{h}_t \oplus \overleftarrow{h}_t \end{cases} \quad (6)$$

其中: $\vec{h}_t$ 表示前向LSTM在时刻 $t$ 的隐藏层状态; $x_t$ 表示当前时刻的输入; $\overleftarrow{h}_t$ 表示后向LSTM在时刻 $t$ 的隐藏层状态. CRF层通过条件随机场优化标注序列,转移矩阵 $T$ 定义状态间转移概率,最终损失函数为:

$$\mathcal{L}_{\text{CRF}} = -\log \frac{\exp[\text{score}(y, x)]}{\sum_{y' \in Y} \exp[\text{score}(y', x)]}, \quad (7)$$

其中: $\text{score}(y, x)$ 表示在给定输入 $x$ 的情况下,预测标注序列 $y$ 的得分,得分越高表示该标注序列越符合真实情况. $\sum_{y' \in Y} \exp[\text{score}(y', x)]$ 是对所有可能标注序列 $y'$ 的得分进行指数化求和,用于归一化处

表 2 光谱数据库介绍

Table 2 Overview of spectral databases

光谱数据库	样品与波段	关键内容
GreenPEG 欧洲伟晶岩光谱库	岩矿反射光谱 350~2 500 nm	原始与连续谱扣除、主吸收峰参数、样品照片、光谱矿物解释
Fregeneda-Almendra 锂专题库	锂矿物与主要露头岩性	野外与实验室光谱,附坐标、蚀变、测面、设备,实验室样本含照片、CR 光谱与吸收峰细节
1981—2016 中国岩矿反射率数据集	多类岩矿,全国 28 省外业观测	光谱反射率与观测时空、仪器、波段参数等元数据,共 9 348 记录
PDS/RELAB 矿物光谱库	矿物与 Mars 模拟材料	绝对反射率,两列制表符分隔,干燥环境 FTIR 测量

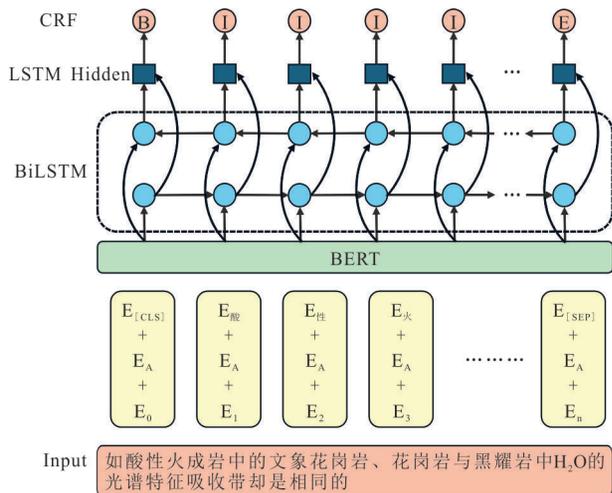


图 3 Bert-BiLSTM-CRF 模型结构

Fig. 3 Model structure of Bert-BiLSTM-CRF

理, Y 为所有可能标注序列的集合. 模型结构如图 3 所示.

本文主要采用半自动化特征峰定位方法来提取光谱曲线的特征,使用一阶导数法计算光谱曲线斜率,二阶导数用来检测曲率变化,判定一阶导数过零点且二阶导数极值点为特征峰. 公式分别为:

$$\begin{cases} \frac{dA}{d\lambda} = \frac{A(\lambda + \Delta\lambda) - A(\lambda - \Delta\lambda)}{2\Delta\lambda} \\ \frac{d^2A}{d\lambda^2} = \frac{\frac{dA}{d\lambda}(\lambda + \Delta\lambda) - \frac{dA}{d\lambda}(\lambda - \Delta\lambda)}{2\Delta\lambda} \end{cases}, \quad (8)$$

其中: A(λ) 为波长 λ 的吸光度; Δλ 为波长间隔.

## 2 知识融合与图谱存储

数据层抽取的多源异构知识存在实体歧义、关系冲突及格式差异等问题,如同一矿物在不同数据源中可能被标注为“高岭石”与“Kaolinite”,或同一岩石的岩石类型存在不同记载. 知识融合通过实体对齐、冲突消解等技术实现跨源知识的语义统一,是构建“岩石—矿物—光谱”知识图谱的关键环节.

本章提出基于语义相似度与字符相似度融合的实体对齐方法,并采用 Neo4j 图数据库实现知识的高效存储与查询,为后续地质应用提供统一的数据基底.

### 2.1 研究数据

为满足“光谱—矿物—岩石—成矿环境”链式推理的需求,本文优先选用覆盖可见光到短波红外约 350~2 500 nm 且元数据完整、可追溯的实验室光谱库,并以 GreenPEG 伟晶岩光谱库为主,辅以 Fregeneda-Almendra 锂专题数据以强化锂端元识别,采用 PDS/RELAB 等权威矿物库进行机理对照与异常峰位核验,同时引入 1981—2016 年中国岩矿反射率数据集作为区域先验与本地化基线,如下表 2 所示具体光谱数据库信息. 上述库具有标准化采集与处理流程,包含原始与连续谱扣除曲线、吸收峰参数和样品照片,部分还提供按 Landsat 8 与 Sentinel 2 等传感器重采样的版本,便于与遥感 workflow 直接衔接. 在文献资料方面,本文系统检索知网与万方的中文研究,覆盖知识图谱、地质、光谱与遥感方向,用于术语标准化、规则归纳与证据分级,并将结论与不确定性显式写入知识图谱,形成从实验室到影像再到决策的可解释证据链. 综合使用这些数据与文献,可以降低人工经验的主观性与误差传播,提升光谱诊断与矿物识别的准确度,支撑矿物组合到成矿类型的稳健推断,加速找矿区优选与工作量配置.

### 2.2 实体融合对齐

实体融合对齐旨在识别不同数据源中指向同一真实对象的实体,核心步骤包括语义相似度计算、字符相似度计算及综合评分决策(张富等, 2022). Word2Vec 通过浅层神经网络将文本转化为低维稠密向量,使语义相近的词汇在向量空间中距离相近(Ye *et al.*, 2016). 本研究采用 Skip-Gram 模型训练地质领域词向量. 如图 4 所示为 Skip-Gram 模型结构,输入为地质实体名称,输出为 100 维的实

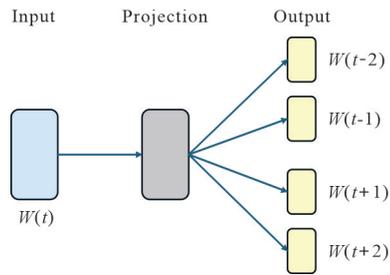


图 4 Skip-Gram 模型结构

Fig. 4 Model structure of skip-gram

表 3 阈值判定规则

Table 3 Threshold determination rule

触发条件	判定结果
$0.6 \leq \text{Sim}_{\text{final}}$	判定为同一实体
$0.4 \leq \text{Sim}_{\text{final}} < 0.6$	触发人工审核
$\text{Sim}_{\text{final}} < 0.4$	判定为不同实体

体向量  $V_i$  (Mikolov *et al.*, 2013).

首先整合光谱库的数据和多篇地质文献内容, 构建一个领域语料库, 使用 jieba 分词工具进行分词, 去除停用词等预处理 (Ding *et al.*, 2021). 然后进行模型训练, 目标函数为最大化上下文词汇的预测概率:

$$L = \sum_{\omega \in C} \sum_{-c \leq j \leq c, j \neq 0} \log p(\omega_{t+j} | \omega_t), \quad (9)$$

其中:  $C$  为语料库,  $c$  为上下窗口大小,  $p(\omega_{t+j} | \omega_t)$  为给定中心词  $\omega_t$  时上下文  $\omega_{t+j}$  的条件概率. 训练完成后, 每个实体获得向量表示  $V_i \in \mathbb{R}^{100}$ , 如“石英”向量为  $V_{\text{石英}}$ , “Quartz”向量为  $V_{\text{Quartz}}$ . 采用余弦相似度衡量实体向量间的语义距离:

$$\text{Sim}_{\text{semantic}}(V_i, V_j) = \frac{V_i \cdot V_j}{\|V_i\| \|V_j\|}. \quad (10)$$

取值范围为  $[0, 1]$ , 值越大表示语义越相近. 编辑距离指将一个字符串转换为另一个字符串所需的最少编辑操作次数. 为消除字符串长度影响, 采用归一化编辑距离:

$$\begin{cases} \text{Sim}_{\text{char}}(s_i, s_j) = 1 - \frac{d(s_i, s_j)}{\max(|s_i|, |s_j|)}, \\ \text{Sim}_{\text{final}} = \alpha \cdot \text{Sim}_{\text{semantic}} + (1 - \alpha) \cdot \text{Sim}_{\text{char}} \end{cases}, \quad (11)$$

其中:  $d(s_i, s_j)$  为编辑距离;  $|s_i|$  和  $|s_j|$  为字符串长度. 通过网格搜索确定权重  $\alpha = 0.7$ , 平衡语义关联与字符匹配, 综合语义相似度与字符相似度, 采用线性加权公式得到最终的  $\text{Sim}_{\text{final}}$ , 然后根据阈值判定规则, 如下表 3 所示, 显示判定的结果.

### 2.3 知识存储

知识存储环节是将抽取的实体、关系及属性转化为结构化知识的重要步骤. 模式层通过对岩石类型、矿物名称与光谱特征的概念、属性和关系进行规范化、形式化的定义, 实现领域知识的高层抽象与理论框架构建. 数据层则以三元组的形式对知识图谱模式进行实例填充, 完成从理论模型到实际地质数据的映射过程, 具体采用“实体—关系—实体”与“实体—属性—属性值”两类知识表示模型. 针对岩石特征、矿物特征和光谱特征等知识所具有的多模态特性, 通过数据获取、预处理与信息抽取, 提炼并构建出实体、实体间的关系及实体属性等有效知识单元, 形成结构化的知识三元组. 此外, 采用分层架构构建“岩石—矿物—光谱”知识图谱, 并利用 Neo4j 图数据库实现高效的知识存储与语义查询, 有效克服传统关系型数据库在处理复杂语义关联方面的局限性.

首先, 依据从观测到解释的最短证据链与主流地学本体的一致性, 本文将岩石、矿物、光谱、成矿环境确定为四大核心类. 该框架既能让光谱特征指向矿物, 再以矿物组合刻画岩石并指示成矿环境, 又便于在岩浆岩、沉积岩与变质岩等子类以及硅酸盐与氧化物等门类上扩展. 同时与实验室和野外光谱、文献与遥感影像等数据源天然对接, 可将吸收峰与半峰高宽等属性以及“HAS\_MINERAL”、“HAS\_LITHOLOGY”等关系纳入计算与推理. 其中岩石类进一步细分为岩浆岩、沉积岩、变质岩等子类, 矿物类按化学组成划分为硅酸盐矿物、氧化物矿物等类别, 光谱类涵盖不同波段光谱数据, 成矿环境类则包括斑岩型、热液型等多种类型, 如下图 5 所示.

然后, 为各概念定义专属属性, 展示部分如下图 6 所示. 其中岩石类包含名称、颜色、纹理特征、形态等属性; 矿物类涵盖名称、颜色等属性; 光谱类包含波长、峰位、峰深、半高宽等属性; 成矿环境类具有类型、蚀变矿物组合等属性. 最后明确实体间的语义关系, 主要包括“岩石—矿物”间的“HAS\_MINERAL”关系, 表示岩石含有某种矿物; “矿物—光谱波长”间的“HAS\_ABS\_PEAK”、“HAS\_REF\_PEAK”关系, 表示矿物具有特定光谱特征; “矿物—成矿环境”间的“MINERAL\_CHAR”、“ALTER\_CHAR”、“TYPE\_MINERAL”关系, 表示矿物形成于特定成矿环境等.

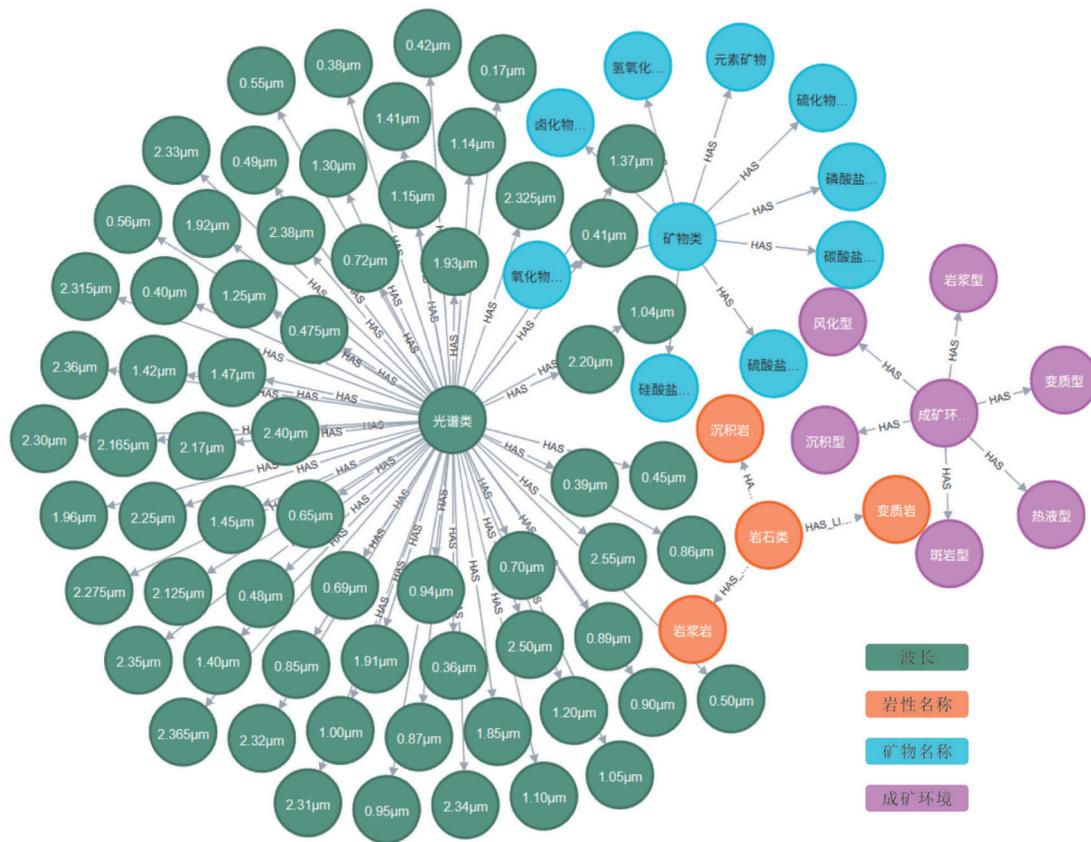


图 5 四大核心类概念定义

Fig. 5 Definitions of the four core class concepts

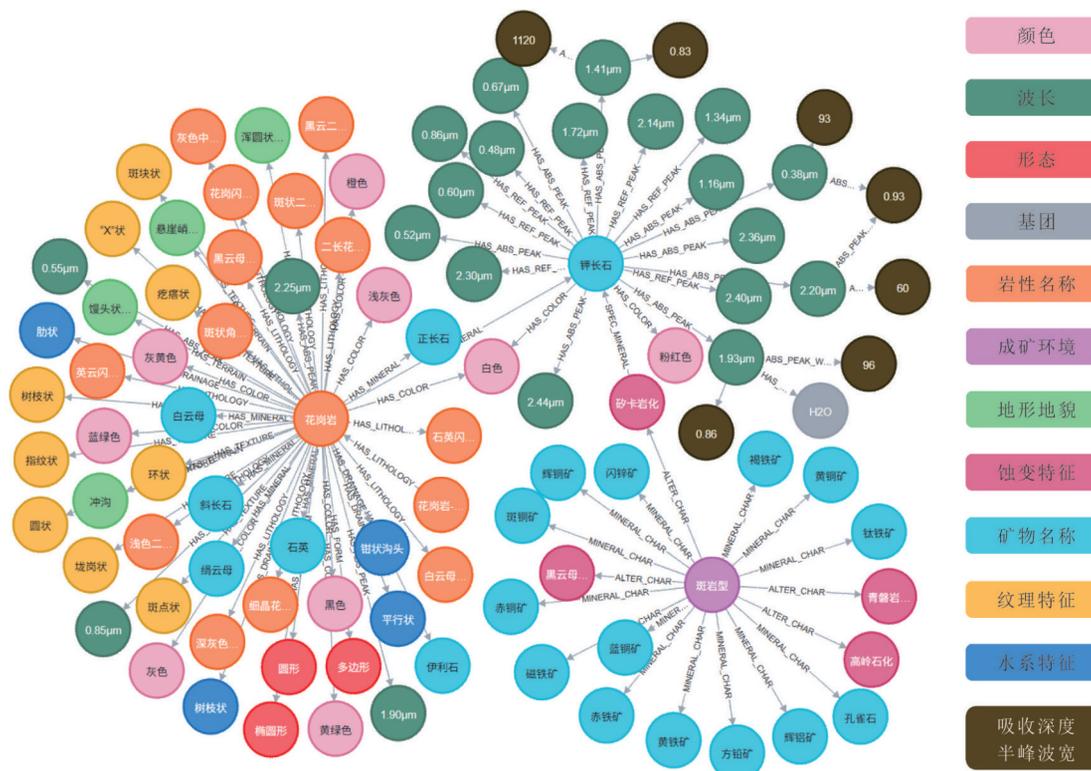


图 6 四大核心类属性

Fig. 6 Four core class attributes

### 3 “岩石—矿物—光谱”知识图谱应用

“岩石—矿物—光谱”知识图谱在地质领域中发挥着多源数据集成与语义推理的关键作用,可将岩石学、光谱学与矿物学领域内的碎片化知识转化为可计算、可推理的结构化知识网络.围绕知识图谱的地质应用价值,本章以光谱特征导向的矿物识别与岩石分类,以及基于岩石矿物组合的成矿类型推理两个典型应用场景为核心,通过对其理论基础、应用逻辑与前沿发展的系统阐述,全方位展示知识图谱如何实现从数据整合、知识表达到智能决策的全过程,体现其对地质认知与矿产勘探决策的综合赋能作用.

#### 3.1 光谱特征导向的矿物识别与岩石分类

在地质勘探与研究领域,光谱特征作为矿物和岩石的“指纹”信息,为精准识别矿物种类、划分岩石类型提供了重要依据.矿物光谱特征的形成,本质上源于分子键振动与电子跃迁等物理过程(吴立新等,2018).当矿物受到光辐射时,其内部特定官能团,如 Al-OH、Mg-OH、 $\text{CO}_3^{2-}$ 、 $\text{Fe}^{2+}$ 等,会吸收特定波长的光能量,从而在遥感光谱中呈现出特征吸收峰.这些吸收峰的位置、强度和形状,构成了光谱识别的物理基础,见下表4所示,可以看到有些矿物存在的不同光谱吸收峰,对应不同的官能团.

RMS-KG通过系统的本体构建,将矿物与光谱之间复杂的映射关系显式地表示为“矿物—存在吸收峰—波长”形式的三元组,明确矿物类型与其特征光谱的语义关联.具体而言,当遥感数据在1.93  $\mu\text{m}$ 处呈现明显吸收峰时,结合知识图谱中的光谱特征节点,可快速定位钾长石、斜长石、亚马逊石等矿物,如下图7所示,在图谱中定位矿物吸收峰的光谱特征.同时,通过光谱曲线的特征峰参数,如吸收峰深度、半峰波宽等,计算输入光谱与知识图谱内矿物光谱的相似度,从而实现矿物种类的智能识别.

进一步,在岩石分类方面,RMS-KG提供了一种“矿物识别—组合分析—岩石分类”的智能推理框架.首先,根据识别的矿物光谱特征,确定岩石的矿物组成;其次,通过矿物组合特征与知识图谱中岩石—矿物关联关系推断岩石类型.若某岩石样本中识别出石英、钾长石、斜长石等矿物,且其光谱特征与花岗岩的典型矿物光谱匹配,同时在知识图谱中,这些矿物的组合关系与花岗岩的矿物组成模式一致,则可判定该岩石为花岗岩.

表4 光谱特征可能对应的矿物示例

Table 4 Examples of minerals corresponding to spectral characteristics

基团	吸收峰( $\mu\text{m}$ )	可能存在矿物
$\text{H}_2\text{O}$	1.40、1.90	高岭石、蒙脱石、云母
	1.92	石英
	1.93	斜长石、钾长石、亚马逊石
Al-OH	2.20	高岭石、蒙脱石、云母
Mg-OH	2.30、2.40	角闪石
	1.40	角闪石
OH	1.41	石英
	1.42	亚马逊石
	1.90	刚玉
	2.30	斜长石
	2.32	黑云母
	2.40	钾长石
	1.85	白云石
$\text{CO}_3^{2-}$	1.96、2.17、2.55	方解石、白云石
	2.30	钾长石
	2.35	方解石
	2.36	黑云母
$\text{Fe}^{2+}$	0.45、1.10	刚玉
	0.95	斜长石
	1.00	橄榄石
	1.15	亚马逊石、刚玉
	1.85	石英、方解石
$\text{Fe}^{3+}$	0.86	钾长石
	0.50、0.90	褐铁矿、针铁矿、赤铁矿

#### 3.2 基于岩石—矿物组合的成矿类型推理

在地质勘探与矿产资源预测中,传统方法高度依赖专家经验,主观性强且效率受限,准确判定成矿类型是确定勘探方向与提升找矿效率的关键环节(王梅,2018).本研究依托构建的“岩石—矿物—光谱”知识图谱,提出以矿物组合为核心的成矿类型推理方案.知识图谱统一术语与关联规则,汇集实验室光谱、野外观测与文献证据,提供可追溯与可解释的推理链路,面向勘查决策输出候选成矿类型及其置信信息.为保障推理的准确性与可操作性,需梳理典型成矿环境及其稳定的矿物蚀变指示特征,不同成矿环境具有相对稳定的矿物组合与蚀变指示.

斑岩型环境通常与黄铁矿、黄铜矿、辉钼矿及绢云母等矿物共生,并伴随硅化与钾化等强烈蚀变(Mao *et al.*, 2011).热液型环境以热液流体为载体,常见黄铁矿、方铅矿、闪锌矿、黄铜矿、毒砂以及石英与碳酸盐矿物等组合(Song *et al.*, 2021).沉积型

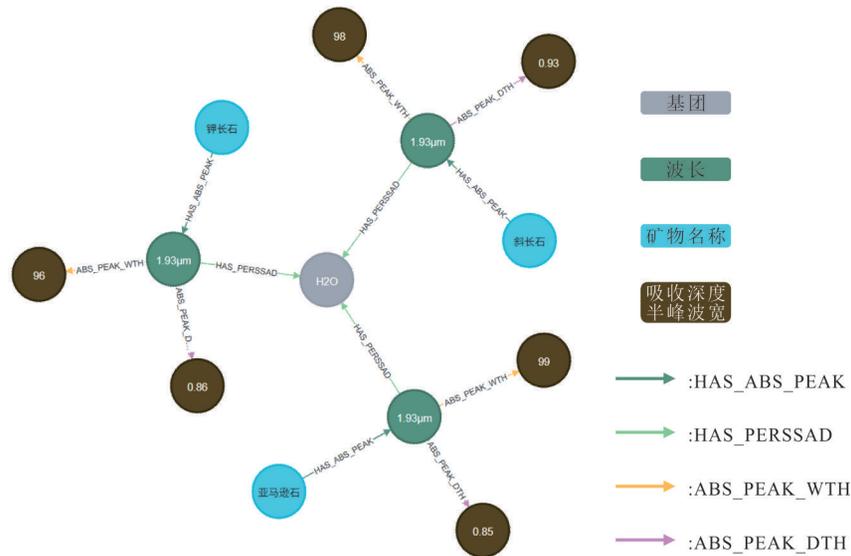


图 7 定位矿物吸收峰

Fig. 7 Location of mineral absorption peaks

表 5 成矿环境规则

Table 5 Metallogenic environment rules

序号	规则内容
规则 1	若矿物组合中含有碳酸盐矿物,如方解石、白云石等,则排除斑岩型铜矿成矿环境
规则 2	当矿物组合出现“黄铁矿+黄铜矿+石英”时,应优先考虑斑岩型铜矿与热液型铜矿成矿环境
规则 3	若矿物组合呈现“赤铁矿+针铁矿+方解石”时,则应优先考虑沉积型铁矿成矿环境
规则 4	若矿物组合中出现橄榄石、铬铁矿或铂族矿物等特征性矿物,应优先考虑岩浆型成矿环境
规则 5	若矿物组合中出现红柱石、蓝晶石、矽线石等特征变质矿物,或石墨、石棉等非金属矿物,则考虑变质型成矿环境
规则 6	若矿物组合中出现三水铝石、硬水铝石等铝土矿矿物,或稀土元素矿物与黏土矿物共生,且位于地表风化带,则优先考虑风化型成矿环境

环境在近地表形成,砂矿常见锆石、金红石与独居石,化学沉积可形成铁锰磷铝土等矿床. 风化型环境受表生淋滤与氧化作用控制,常见三水铝石、一水硬铝石、孔雀石与蓝铜矿等矿物组合(Song *et al.*, 2019). 岩浆型环境由分异与结晶作用主导,可形成铬铁矿、钽铁磁铁矿与铜镍硫化物等,典型矿物包含磁铁矿、钛铁矿、铬铁矿与含硫化物组合(许立权等, 2016). 变质型环境在温压变化下重结晶与再分配,代表性矿床包括变质铁矿、石墨与滑石等类型(Tian *et al.*, 2024). 上述相对稳定的组合关系为专家经验规则的提炼与知识图谱中的可解释推理提供了依据.

首先,依据矿物共生组合理论,构建专家经验驱动的成矿环境初筛规则库,快速缩小候选成矿类型的范围. 如下表 5 所示,当矿物组合包含“黄铁矿+黄铜矿+石英”时,初筛规则将斑岩型铜矿与热液型铜矿环境纳入候选;若组合中含有橄榄石、铬铁矿或铂族矿物时,则优先考虑岩浆型成矿环境.

其次,针对经规则初筛后的候选成矿环境,采用图注意力网络(graph attention network, GAT)进行深度智能推理(Velickovic *et al.*, 2017). 具体做法是将岩石矿物组合编码为计算机可识别的 One-Hot 向量形式,随后以岩石—矿物组合为起始节点,以成矿环境为目标节点构建局部子图结构. 利用 GAT 中的多头注意力机制,挖掘矿物组合节点与邻居成矿环境节点之间的潜在关联关系. 通过计算节点之间的注意力权重,突出了关键矿物组合对成矿环境推断的重要性. 最终, GAT 将节点特征进行聚合处理,并通过全连接层与 Softmax 分类函数,输出各候选成矿环境的概率得分,实现精准的成矿类型预测.

#### 4 总结与展望

本研究以“岩石—矿物—光谱”知识图谱的构建与应用为核心,旨在解决传统地质研究领域数据

碎片化、知识关联薄弱和推理效率低下等问题。研究通过整合遥感影像、地质文献等多源异构数据,融合知识图谱与深度学习技术,以期推动地质领域的理论与技术创新。在知识图谱的构建层面,本研究采用模式层与数据层的分层架构。其中,模式层负责标准化定义地质领域知识的概念体系与语义关系;数据层则基于 ResNet-50、BERT 等深度学习模型,对多源数据进行特征提取,并以三元组的结构化形式存储于 Neo4j 图数据库中,实现多源地质知识结构化整合与语义关联。在知识图谱的应用层面,本研究探讨了光谱导向矿物识别与岩石-矿物组合成矿推理两个关键场景。在光谱导向矿物识别中,通过构建光谱-矿物映射关系,结合图结构计算,实现从光谱数据到矿物类别的精准反演;在岩石-矿物组合成矿推理方面,融合矿物共生理论与图神经网络技术,提出“规则初筛+深度学习”的混合推理方法,有效提升矿床潜在成矿预测的准确性与效率。当然,本研究仍存在若干局限性,包括复杂地质数据处理精度有待提升、知识图谱更新过于依赖人工干预,以及跨领域应用场景扩展能力较弱等问题。未来的研究工作可考虑引入隐变量图模型等先进方法以优化复杂数据处理流程,开发自动化知识更新算法并结合区块链技术实现知识图谱的动态管理,以进一步增强知识图谱在地质领域中的实用价值。

## References

- Brodaric, B., Richard, S.M., 2021. The GeoScience Ontology Reference. Geological Survey of Canada, Open File, 8796, 34. Natural Resources Canada. <https://doi.org/10.4095/328296>
- Chen, Y., Tian, M., Wu, Q. R., et al., 2024. A Deep Learning-Based Method for Deep Information Extraction from Multimodal Data for Geological Reports to Support Geological Knowledge Graph Construction. *Earth Science Informatics*, 17(3): 1867–1887. <https://doi.org/10.1007/s12145-023-01207-0>
- Ding, Y., Teng, F., Zhang, P., et al., 2021. Research on Text Information Mining Technology of Substation Inspection Based on Improved Jieba. 2021 *International Conference on Wireless Communications and Smart Grid (ICWCSG)*. August 13–15, 2021, Hangzhou, China. IEEE: 561–564. <https://doi.org/10.1109/ICWCSG53609.2021.00119>
- Enkhsaikhan, M., Holden, E. J., Duuring, P., et al., 2021. Understanding Ore-Forming Conditions Using Machine Reading of Text. *Ore Geology Reviews*, 135: 104200. <https://doi.org/10.1016/j.oregeorev.2021.104200>
- Guan, S. P., Jin, X. L., Jia, Y. T., et al., 2018. Knowledge Reasoning over Knowledge Graph: a Survey. *Journal of Software*, 29(10): 2966–2994(in Chinese with English abstract).
- Garcia, L. F., Abel, M., Perrin, M., et al., 2020. The GeoCore Ontology: a Core Ontology for General Use in Geology. *Computers & Geosciences*, 135: 104387. <https://doi.org/10.1016/j.cageo.2019.104387>
- Hu, X. M., Xu, Y. W., Ma, X. G., et al., 2023. Knowledge System, Ontology, and Knowledge Graph of the Deep-Time Digital Earth (DDE): Progress and Perspective. *Journal of Earth Science*, 34(5): 1323–1327. <https://doi.org/10.1007/s12583-023-1930-1>
- Jin, X. B., Shen, L., Huang, X. J., et al., 2024. Empowering High-Quality Management of Natural Resources with New Quality Productive Forces: Logic and Path. *Journal of Natural Resources*, 39(9): 2011–202(in Chinese with English abstract).
- Kong, J. Y., Gao, Y. R., Zhang, Y. J., et al., 2021. Improved Attention Mechanism and Residual Network for Remote Sensing Image Scene Classification. *IEEE Access*, 9: 134800–134808. <https://doi.org/10.1109/ACCESS.2021.3116968>
- Li, D. R., Wang, S. L., Shi, W. Z., et al., 2001. On Spatial Data Mining and Knowledge Discovery (SDMKD). *Geomatics and Information Science of Wuhan University*, 26(6): 491–499(in Chinese with English abstract).
- Liu, G. Q., Gong, R. B., Shi, Y. J., et al., 2022. Construction of Well Logging Knowledge Graph and Intelligent Identification Method of Hydrocarbon-Bearing Formation. *Petroleum Exploration and Development*, 49(3): 502–512 (in Chinese with English abstract).
- Lei, X. Y., Song, W. J., Fan, R. Y., et al., 2023. Semi-Supervised Geological Disasters Named Entity Recognition Using few Labeled Data. *GeoInformatica*, 27(2): 263–288. <https://doi.org/10.1007/s10707-022-00474-1>
- Le, B. M. J., Streckeisen, A. L., 1991. The IUGS Systematics of Igneous Rocks. *Journal of the Geological Society*, 148(5): 825–833. <https://doi.org/10.1144/gsjgs.148.5.0825>
- Liu, Q., Li, Y., Duan, H., et al., 2016. Knowledge Graph Construction Techniques. *Journal of Computer Research and Development*, 53(3): 582–600(in Chinese with English abstract).
- Ma, X. G., 2022. Knowledge Graph Construction and Application in Geosciences: a Review. *Computers & Geosciences*, 161: 105082. <https://doi.org/10.1016/j.cageo.2022.105082>

- Mao, J. W., Zhang, J. D., Pirajno, F., et al., 2011. Porphyry Cu-Au-Mo-Epithermal Ag-Pb-Zn-Distal Hydrothermal Au Deposits in the Dexing Area, Jiangxi Province, East China: A Linked Ore System. *Ore Geology Reviews*, 43(1): 203—216. <https://doi.org/10.1016/j.oregeorev.2011.08.005>
- Mikolov, T., Chen, K., Corrado, G., et al., 2013. Efficient Estimation of Word Representations in Vector Space. *arXiv preprint arXiv: 1301.3781*. <https://doi.org/10.48550/arXiv.1301.3781>
- Niu, F.G., Zhang, B., Chen, S., 2024. Review and Perspective of Earth Science Knowledge Graph in Big Data Era. *Acta Seismologica Sinica*, 46(3): 353—376 (in Chinese with English abstract).
- Qiu, Q. J., Wang, B., Ma, K., et al., 2023. A Practical Approach to Constructing a Geological Knowledge Graph: a Case Study of Mineral Exploration Data. *Journal of Earth Science*, 34(5): 1374—1389. <https://doi.org/10.1007/s12583-023-1809-3>
- Qiu, Q. J., Wu, L., Ma, K., et al., 2023. A Knowledge Graph Construction Method for Geohazard Chain for Disaster Emergency Response. *Earth Science*, 48(5): 1875—1891(in Chinese with English abstract).
- Shen, Z. H., Zhu, X. J., Wang, H. J., et al., 2024. Research Data Network: Concept, Systems and Applications. *Frontiers of Data & Computing*, 6(4): 3—21(in Chinese with English abstract).
- Song, M. C., Ding, Z. J., Zhang, J. J., et al., 2021. Geology and Mineralization of the Sanshandao Supergiant Gold Deposit (1 200 t) in the Jiaodong Peninsula, China: a Review. *China Geology*, 4(4): 686—719. <https://doi.org/10.31035/cg2021070>
- Song, Y. C., Liu, Y. C., Hou, Z. Q., et al., 2019. Sediment-Hosted Pb-Zn Deposits in the Tethyan Domain from China to Iran: Characteristics, Tectonic Setting, and Ore Controls. *Gondwana Research*, 75: 249—281. <https://doi.org/10.1016/j.gr.2019.05.005>
- Tian, J. P., Wang, J. H., Tian, T. L., et al., 2024. In-Situ Geochemical and Rb-Sr Dating Analysis of Sulfides from a Gold Deposit Offshore of Northern Sanshandao, Jiaodong Peninsula, North China: Implications for Gold Mineralization. *Minerals*, 14(5): 456. <https://doi.org/10.3390/min14050456>
- Velickovic, P., Cucurull, G., Casanova, A., et al., 2017. Graph Attention Networks. *Stat*, 1050(20):10—48550.
- Wang, C. S., Hazen, R. M., Cheng, Q. M., et al., 2021. The Deep-Time Digital Earth Program: Data-Driven Discovery in Geosciences. *National Science Review*, 8(9): nwab027. <https://doi.org/10.1093/nsr/nwab027>
- Wu, L. X., Mao, W. F., Liu, S. J., et al., 2018. Mechanism of Infrared and Microwave Radiation Variation of Rock Stress and Key Problems of In-Situ Stress Remote Sensing. *National Remote Sensing Bulletin*, 22(S1): 146—161(in Chinese with English abstract).
- Wang, M., 2018. Metallogenic Type and Prospecting Characteristics of Bauxite in a Certain Area. *World Nonferrous Metals*, 43(12): 102—103 (in Chinese with English abstract).
- Xu, Q., Cui, S. H., Huang, W., et al., 2023. Construction of a Landslide Knowledge Graph in the Field of Engineering Geology. *Geomatics and Information Science of Wuhan University*, 48(10): 1601—1615(in Chinese with English abstract).
- Xu, L. Q., Zhang, T., Zhang, M., et al., 2016. Summary of Ore Regularity of Important Mineral Resources in Inner Mongolia. *Mineral Deposits*, 35(5): 966—980(in Chinese with English abstract).
- Xie, T., Yang, J.A., Liu, H., 2020. Chinese Entity Recognition Based on BERT-BiLSTM-CRF Model. *Computer Systems & Applications*, 29(7): 48—55(in Chinese with English abstract).
- Ye, X., Shen, H., Ma, X., et al., 2016. From Word Embeddings to Document Similarities for Improved Information Retrieval in Software Engineering. *Proceedings of the 38th International Conference on Software Engineering*. Austin Texas, . ACM, 404—415. <https://doi.org/10.1145/2884781.2884862>
- Zhai, M. G., Yang, S. F., Chen, N. H., et al., 2018. Big Data Epoch: Challenges and Opportunities for Geology. *Bulletin of Chinese Academy of Sciences*, 33(8): 825—831(in Chinese with English abstract).
- Zhou, C.H., Wang, H., Wang, C.S., et al., 2021. Research on Geoscience Knowledge Graph in the Era of Big Data. *Chinese Science: Earth Sciences*, 51(7): 1070—1079 (in Chinese).
- Zhang, C. J., Liu, W. C., Zhang, X. Y., et al., 2023. Knowledge Graph Construction Method of Gold Mine Based on Ontology. *Journal of Geo-Information Science*, 25(7): 1269—1281(in Chinese with English abstract).
- Zhou, Y.Z., Zhang, Q.L., Huang, Y.J., et al., 2021. Constructing Knowledge Graph for the Porphyry Copper Deposit in the Qingzhou-Hangzhou Bay Area: Insight into Knowledge Graph based Mineral Resource Prediction and Evaluation. *Earth Science Frontiers*, 28(3): 67—75 (in Chinese with English abstract).
- Zhang, Q. L., Zhou, Y. Z., Yu, P. P., et al., 2024. Ontology Construction of Multi-Level Ore Deposit and Its Applica-

tion in Knowledge Graph. *Bulletin of Mineralogy, Petrology and Geochemistry*, 43(1): 211–217(in Chinese with English abstract).

Zhao, Z. F., Han, S. K., So, I. M., 2018. Architecture of Knowledge Graph Construction Techniques. *International Journal of Pure and Applied Mathematics*, 118(19): 1869–1883.

Zhang, F., Yang, L. Y., Li, J. W., et al., 2022. An Overview of Entity Alignment Methods. *Chinese Journal of Computers*, 45(6): 1195–1225(in Chinese with English abstract).

### 中文参考文献

官赛萍, 靳小龙, 贾岩涛, 等, 2018. 面向知识图谱的知识推理研究进展. *软件学报*, 29(10):2966–2994.

金晓斌, 沈镭, 黄贤金, 等, 2024. 新质生产力赋能自然资源高质量管理: 逻辑与路径. *自然资源学报*, 39(9):2011–2028.

李德仁, 王树良, 史文中, 等, 2001. 论空间数据挖掘和知识发现. *武汉大学学报(信息科学版)*, 26(6):491–499.

刘国强, 龚仁彬, 石玉江, 等, 2022. 油气层测井知识图谱构建及其智能识别方法. *石油勘探与开发*, 49(3):502–512.

刘峤, 李杨, 段宏, 等, 2016. 知识图谱构建技术综述. *计算机研究与发展*, 53(3):582–600.

牛凤桂, 张贝, 陈石, 2024. 大数据时代的地球科学知识图谱研究现状与展望. *地震学报*, 46(3):353–376.

邱芹军, 吴亮, 马凯, 等, 2023. 面向灾害应急响应的地质灾害链

知识图谱构建方法. *地球科学*, 48(5):1875–1891.

王梅, 2018. 某区铝土矿成矿类型以及找矿特点初探. *世界有色金属*, 43(12):102–103.

吴立新, 毛文飞, 刘善军, 等, 2018. 岩石受力红外与微波辐射变化机理及地应力遥感关键问题. *遥感学报*, 22(S1):161.

许强, 崔圣华, 黄维, 等, 2023. 面向工程地质领域的滑坡知识图谱构建方法研究. *武汉大学学报(信息科学版)*, 48(10):1601–1615.

许立权, 张彤, 张明, 等, 2016. 内蒙古自治区重要矿种成矿规律综述. *矿床地质*, 35(5):966–980.

谢腾, 杨俊安, 刘辉, 2020. 基于BERT-BiLSTM-CRF模型的中文实体识别. *计算机系统应用*, 29(7):48–55.

翟明国, 杨树锋, 陈宁华, 等, 2018. 大数据时代: 地质学的挑战与机遇. *中国科学院院刊*, 33(8):825–831.

周成虎, 王华, 王成善, 等, 2021. 大数据时代的地质知识图谱研究. *中国科学: 地球科学*, 51(7):1070–1079.

张春菊, 刘文聪, 张雪英, 等, 2023. 基于本体的金矿知识图谱构建方法. *地球信息科学学报*, 25(7):1269–1281.

周永章, 张前龙, 黄永健, 等, 2021. 钦杭成矿带斑岩铜矿知识图谱构建及应用展望. *地学前缘*, 28(3):67–75.

张前龙, 周永章, 虞鹏鹏, 等, 2024. 多层次矿床本体的构建及在知识图谱中的应用. *矿物岩石地球化学通报*, 43(1):211–217.

张富, 杨琳艳, 李健伟, 等, 2022. 实体对齐研究综述. *计算机学报*, 45(6):1195–1225.