

面向多模态晶体结构预测的 LLM 代理框架^①

曹 芊^{1,3}, 徐 殷^{2,3}, 肖明军^{2,3}

¹(中国科学技术大学 人工智能与数据科学学院, 合肥 230026)

²(中国科学技术大学 计算机科学与技术学院, 合肥 230027)

³(中国科学技术大学 苏州高等研究院, 苏州 215004)

通信作者: 肖明军, E-mail: xiaomj@ustc.edu.cn



摘要: 高通量 X 射线衍射 (X-ray diffraction, XRD) 分析在加速材料发现方面至关重要, 但传统方法通常依赖大量人工解释, 且在处理复杂的 XRD 数据时容易忽视低强度峰值信息, 从而限制准确性的提升. 为解决这一问题, 本文提出了一个面向多模态晶体结构预测的大语言模型 (large language model, LLM) 代理框架, 该框架集成了 GPT-4 驱动的智能代理以及基于 XRD 和对分布函数的多模态投票模型, 能够自主执行晶体结构和空间群预测任务. 此外, 本文通过引入知识图谱来增强 LLM 的推理能力, 帮助其理解晶体特征之间的关系, 进一步提升预测的准确性. 实验结果表明, 该框架在晶体结构预测和空间群预测任务上的准确率分别达到 97.5% 和 98.7%. 这一设计显著提升了高通量分析的准确性和效率, 有望推动材料科学研究的进展, 为解决其他具有高度关联性的多任务问题提供宝贵的启示.

关键词: 大语言模型; 深度学习; 知识图谱; 多模态融合; 晶体结构预测

引用格式: 曹芊, 徐殷, 肖明军. 面向多模态晶体结构预测的 LLM 代理框架. 计算机系统应用, 2025, 34(8): 33-42. <http://www.c-s-a.org.cn/1003-3254/9893.html>

LLM-agent Framework for Multimodal Crystal Structure Prediction

CAO Qian^{1,3}, XU Yin^{2,3}, XIAO Ming-Jun^{2,3}

¹(School of Artificial Intelligence and Data Science, University of Science and Technology of China, Hefei 230026, China)

²(School of Computer Science and Technology, University of Science and Technology of China, Hefei 230027, China)

³(Suzhou Institute for Advanced Research, University of Science and Technology of China, Suzhou 215004, China)

Abstract: High-throughput X-ray diffraction (XRD) analysis plays a crucial role in accelerating material discovery. However, traditional methods often rely heavily on manual interpretation and tend to overlook low-intensity peak information when processing complex XRD data, thus limiting the potential for accuracy improvement. To address this issue, a large language model (LLM) agent framework for multi-modal crystal structure prediction is proposed. The framework integrates a GPT-4-driven intelligent agent with multi-modal voting models based on XRD and pair distribution functions, enabling autonomous crystal structure and space group prediction tasks. In addition, the reasoning capability of the LLM is enhanced through the introduction of knowledge graphs, which aid in understanding the relationships between crystal features, thus improving both prediction accuracy and reasoning performance. Experimental results demonstrate that the accuracy of this framework in crystal structure prediction and space group prediction tasks reaches 97.5% and 98.7%, respectively. This design significantly enhances the accuracy and efficiency of high-throughput analysis, with the potential to advance materials science research and provide valuable insights for addressing other highly interrelated multi-task problems.

Key words: large language model (LLM); deep learning; knowledge graph; multimodal fusion; crystal structure prediction

① 基金项目: 国家自然科学基金面上项目 (62172386); 江苏省自然科学基金面上项目 (BK20231212)

收稿时间: 2024-12-16; 修改时间: 2025-01-07; 采用时间: 2025-01-24; csa 在线出版时间: 2025-06-24

CNKI 网络首发时间: 2025-06-25

1 引言

高通量材料合成与快速表征是逆向设计和加速材料发现的关键环节^[1]. 在众多表征手段中, X射线衍射(X-ray diffraction, XRD)作为一种强大的表征材料的技术脱颖而出. XRD分析不仅能够解析晶体对称性和晶格参数, 还能深入揭示材料的原子排列与堆积方式, 对揭示材料的电磁特性及热力学性能具有重要作用. 近年来, XRD技术在识别新材料结构和指导材料设计与合成方面发挥了关键作用. 例如, 通过高通量XRD分析, 可以高效筛选出具有潜在优异性能的材料, 从而加速功能材料的发现进程^[2]. 此外, XRD分析还能够支持在极端条件下对材料的动力学性能进行研究^[3]. 因此, XRD高效分析对于加速新材料的发现与应用具有重要意义, 尤其是在新能源材料、催化剂和功能陶瓷等前沿领域, 具有广泛应用前景. 然而, 人工手动分析大量XRD数据极具挑战性. 开发基于XRD的快速、准确的晶体材料识别方法, 将有助于降低研究成本并提升效率^[4].

随着人工智能技术在材料科学领域的应用不断深入, 越来越多的研究利用深度学习算法解决XRD解析问题, 涵盖晶系预测^[5], 晶格常数预测^[6], 晶体结构^[7]和晶相预测^[8]等多个方面. 其中, 卷积神经网络(convolutional neural network, CNN)因其强大的特征提取能力, 能够从复杂的XRD数据中捕捉多尺度信息, 成为XRD晶体结构识别的主要工具. 然而, CNN算法仍然存在一些固有问题, 例如过于依赖衍射图案中的最大峰值, 忽略不显著特征, 从而在处理复杂或相似晶体结构时容易出错.

然而, 这些问题可以通过引入其他数据表示形式来缓解. 例如, 对分布函数(pair distribution function, PDF)数据作为衍射数据的实空间表示, 能够有效弥补传统XRD数据的不足. PDF通过傅里叶变换将衍射数据从倒空间转换到实空间, 不仅揭示了材料局部结构的短程有序性, 还能够捕捉衍射图案中细微但重要的特征. 这些特征在仅关注最大峰值的情况下往往被忽略. 已有研究表明, 利用PDF作为材料描述符的深度学习模型在晶体空间群和晶相的解析任务中表现优异^[8,9]. 因此, 将PDF与XRD数据结合, 可以为深度学习模型提供更加全面的信息支持, 从而提升晶体结构识别的准确性与鲁棒性. 尽管如此, 这些方法通常依赖于大量人工干预, 限制了其高通量和自动化应用中

的效率及可扩展性.

近年来, 各种由大语言模型(large language model, LLM)驱动的人工智能代理(agent)被广泛开发并应用于化学、生物等领域^[10-14]. 这些代理通过思维链推理, 工具调用等方式减少人类参与, 提高了执行复杂实验的效率. 例如, Coscientist是一个由GPT-4驱动的人工智能系统, 该系统整合了互联网搜索、代码执行和实验自动化等工具, 能够自主设计、规划和执行复杂实验, 加速科学发现并降低实验门槛^[12]. 此外, ChemCrow集成了18种专用工具, 通过将这些工具与LLM相结合, 能够自主设计并执行复杂的化学任务^[13]. 针对基于XRD的材料分析, 传统方法往往依赖大量人工干预, 过程繁琐且效率低下. 而借助LLM可以减少人工参与, 提升自动化水平, 从而加快分析速度并提高准确性. 受这些进展的启发, 将LLM驱动的代理和有效工具相结合有望解决复杂的XRD分析任务. 然而, 如何使用LLM来快速且准确地识别晶体结构仍然面临着两个新的挑战.

虽然LLM已在化学和生物学等领域取得了显著进展, 但在晶体学领域, 其潜力还未得到充分开发. 现有方法通常独立预测空间群、晶系或晶体结构等特征, 未能将这些特征之间的物理约束和复杂依赖性融入预测模型, 这可能会限制预测结果的准确性与全面性. 进一步地, 如何使LLM理解这些特性之间的关系, 从而提升推理能力, 是一个极具前景的方向. 然而, 晶体学涉及复杂的物理规则和大量特定领域知识, 由于现有的通用LLM缺乏对这些知识的深度理解, 因此在处理这类知识密集型任务时不仅内存开销巨大, 而且更新知识库的微调过程资源消耗高. 此外, 现有基于搜索引擎的增强方法对于晶体学专业知识的补充也较为有限. 因此, 如何使得LLM理解晶体结构和空间群之间的复杂关系, 并基于这些关系合理规划路径, 从而执行有效推理与预测, 成为晶体结构预测面临的关键挑战.

其次, LLM代理在调用外部工具时, 工具模型的设计对于提升晶体结构预测的准确性至关重要. XRD中布拉格峰反映了材料的显著特征, 但对于复杂结构或相似相的区分, 这些显著特征可能不足以提供充分信息, 而PDF对衍射图谱中的细微特征更为敏感. 因此, 设计一个能够有效处理并融合XRD与PDF多模态数据的模型, 是实现高效晶体结构预测的核心挑战之一.

为了应对这些挑战, 我们提出了一个面向多模态

晶体结构预测的 LLM 代理框架 (LLM-agent framework for multimodal crystal structure prediction, LMM-CSP). 该框架集成了一个由 GPT-4 驱动的智能代理, 一个描述了晶体结构与空间群关系的知识图谱, 和两个精心设计的基于深度学习模型的预测工具, 以高效完成复杂的晶体结构分析任务. 在该框架中, 我们设计了一个多模态投票模型, 并基于该模型开发了两个预测工具: 晶体结构预测器用于预测晶体结构, 空间群预测器用于预测空间群, 这两个工具的分类准确率分别达到 95.4% 和 89.1%, 性能显著超越现有最先进方法. 随后, 我们将这两个预测工具、搜索引擎和其他通用工具集成到一个工具集中. 当科学家通过自然语言执行 XRD 分析任务时, 代理会调用工具集中的预测工具, 生成多个候选的晶体结构或空间群结果. 随后, 代理结

合知识图谱提取晶体结构与空间群之间的关联, 并通过综合分析这些信息, 动态调整推理路径, 从而优化预测过程. 最终, 经过这一综合推理流程, 晶体结构预测和空间群预测的准确率分别达到了 97.5% 和 98.7%. 这一设计不仅大大提升了 XRD 分析的准确性和效率, 有助于推动材料科学和晶体学领域的研究进展, 并为其他具有高度相关性的多任务问题提供了有价值的启示.

2 框架

本文提出了一个面向多模态晶体结构预测的 LLM 代理框架 LMM-CSP, 其结构见图 1. 旨在以文本的形式回答科学家的问題, 这些问題主要涉及晶体结构和空间群预测任务.

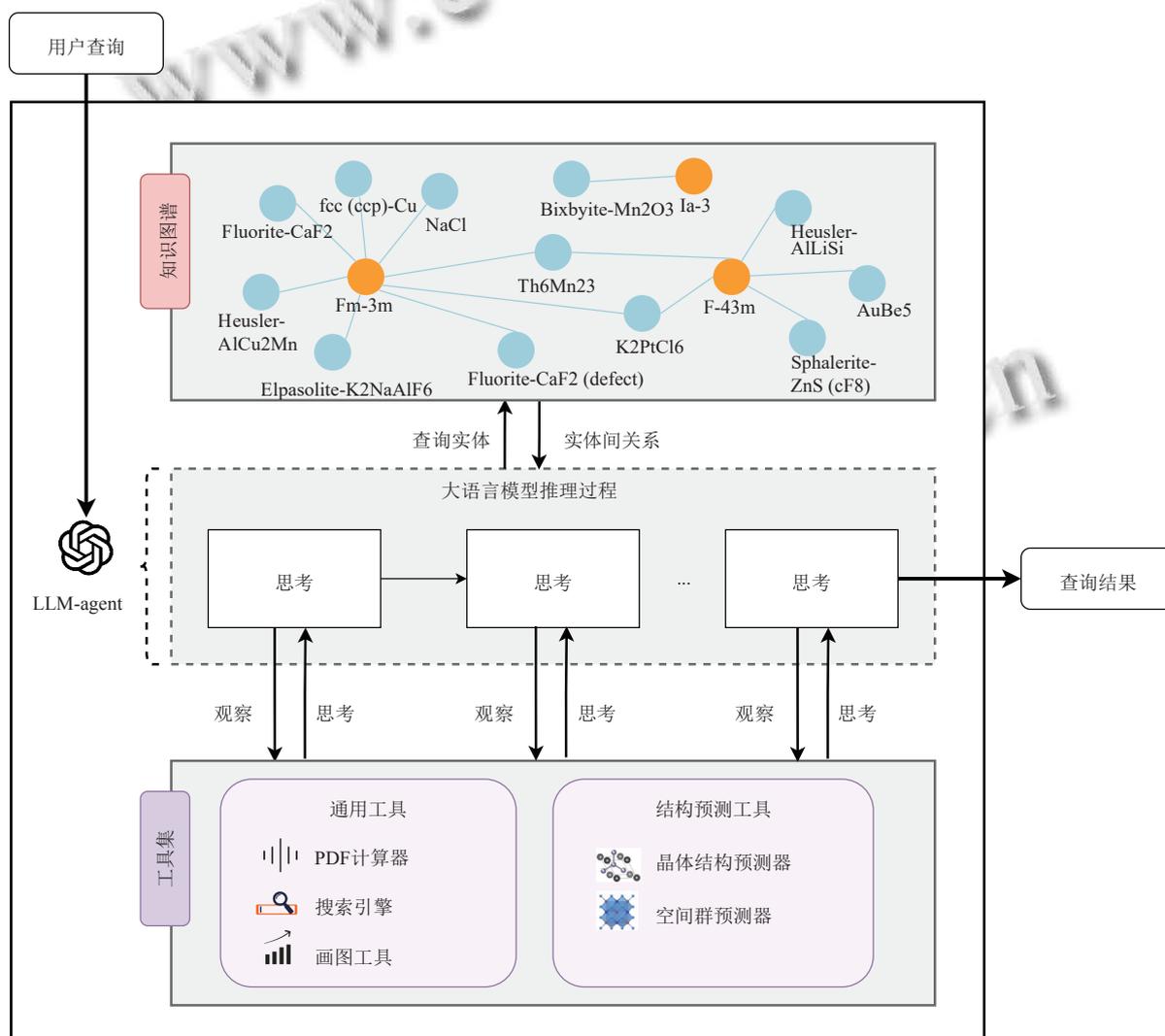


图 1 面向多模态晶体结构预测的 LLM 代理框架

如图 1 所示, LMM-CSP 框架包括 3 个主要模块: LLM-agent、工具集以及知识图谱. 由 GPT-4 驱动的代理是框架的核心, 它是基于 Python 环境下 OpenAI 提供的应用程序接口中的 GPT-4 模型实现的. 该代理可以理解科学家的请求, 规划推理过程, 并依据任务需求调用合适的工具来生成最终答案. 工具集提供了多个用于 XRD 分析的工具, 分为两类: 结构预测工具和通用工具. 结构预测工具由晶体结构预测器和空间群预测器组成. 这两个工具使用了本文设计的多模态加权投票融合模型来预测晶体结构和空间群 (见第 3 节). 通用工具包括 PDF 计算器、搜索引擎和画图工具, 分别用于从 XRD 数据中计算出 PDF 数据、检索晶体学领域知识以及可视化 XRD 和 PDF 数据. 这 3 个工具

均基于 Python 实现, 并封装为一个标准化接口, 供 LLM 代理调用. 其中, PDF 计算器实现了与傅里叶变换相关的计算; 搜索引擎工具通过 Google 提供的接口集成, 支持实时信息检索, 确保结果的时效性与准确性; 画图工具则用于绘制 XRD 和 PDF 数据曲线, 直观展示数据特征, 方便进一步分析. 知识图谱模块用于补充晶体结构与空间群之间的结构联系. 它为 LLM-agent 的推理过程提供了晶体结构和空间群的相关知识, 有助于增强预测结果的合理性和一致性.

本文提出的晶体结构预测的工作流旨在通过多阶段推理过程提升晶体结构与空间群预测的准确性. 如图 2 所示, 工作流包括 4 个步骤: 初始预测、知识检索、推理优化和最终结果预测.

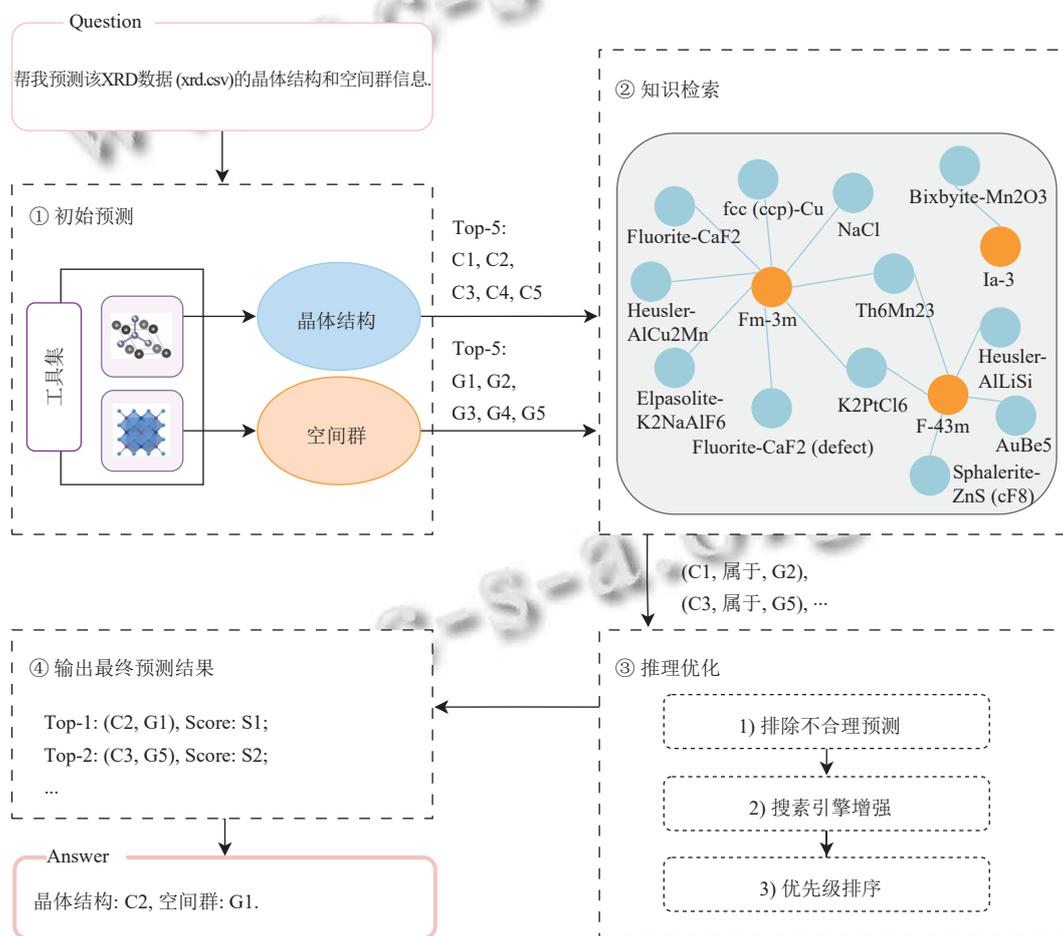


图 2 LMM-CSP 框架的工作流

初始预测阶段, LLM-agent 调用工具集中的 PDF 计算器从 XRD 数据中提取 PDF 信息, 并利用晶体结构预测器和空间群预测器从 XRD 和 PDF 两个模态数据中进行预测, 最终获得最可能的 5 个晶体结构及空

间群候选结果. 知识检索阶段, LLM-agent 通过查询预构建的知识图谱 (见第 4 节) 来获取晶体结构与空间群之间的关系. 推理优化阶段通过知识图谱的信息对初步预测结果进行进一步约束与精化. 首先, 从知识图谱

获取到的晶体结构与空间群之间的关系(如某些空间群仅适用于特定的晶体结构)会被应用来筛选不合理的预测。其次,对于知识图谱中未涵盖的晶体结构和空间群实体,LLM-agent通过搜索引擎增强来获取相关结构信息,填补知识空白。它能够从文献、材料数据库等非结构化数据中提取特性描述、对称参数等关键信息,并结合语义分析确保结果的准确性,从而对知识图谱内容进行补充,为后续推理提供支持。最后,综合上述结果,调整候选空间群和晶体结构的优先级排序。最终结果预测阶段输出优先级最高的晶体结构-空间群组合。这一过程不仅利用多模态数据丰富了模型的输入信息,还通过知识图谱强化了预测的物理一致性,同时借助LLM的规划和推理能力来优化决策,大幅提升了预测的准确性与效率。

3 结构预测工具设计

本文LLM代理调用的结构预测工具包括晶体结构预测器(CSPredictor)和空间群预测器(SGPredictor),

分别用于晶体结构预测和空间群预测任务。这两个工具都是基于本文设计的多模态投票融合模型所构建。第3.1、3.2节将分别介绍这两个工具的模型架构和训练数据。

3.1 模型架构和训练细节

多模态投票融合模型架构由多个模块组成:首先,XRD模块以长度为4096的XRD数据作为输入,并通过1个CNN网络提取潜在特征;其次,PDF模块处理由XRD数据经过傅里叶变换得到的序列长度为1000的PDF数据,并使用与XRD模块一致的CNN模型来执行分类任务;随后,两模块的输出特征在融合层进行投票融合,融合过程引入动态加权系数 α ,以灵活调整XRD与PDF数据的贡献,从而有效利用两种数据源的互补信息;最后,融合后的特征通过加权计算,前 K 类概率最高(即Top-K)的预测结果被输出。模型架构如图3所示,此模型通过对XRD和PDF数据进行加权融合,具备根据不同任务灵活调整对不同数据源关注程度的能力,且有望提升分类任务的准确性和鲁棒性。

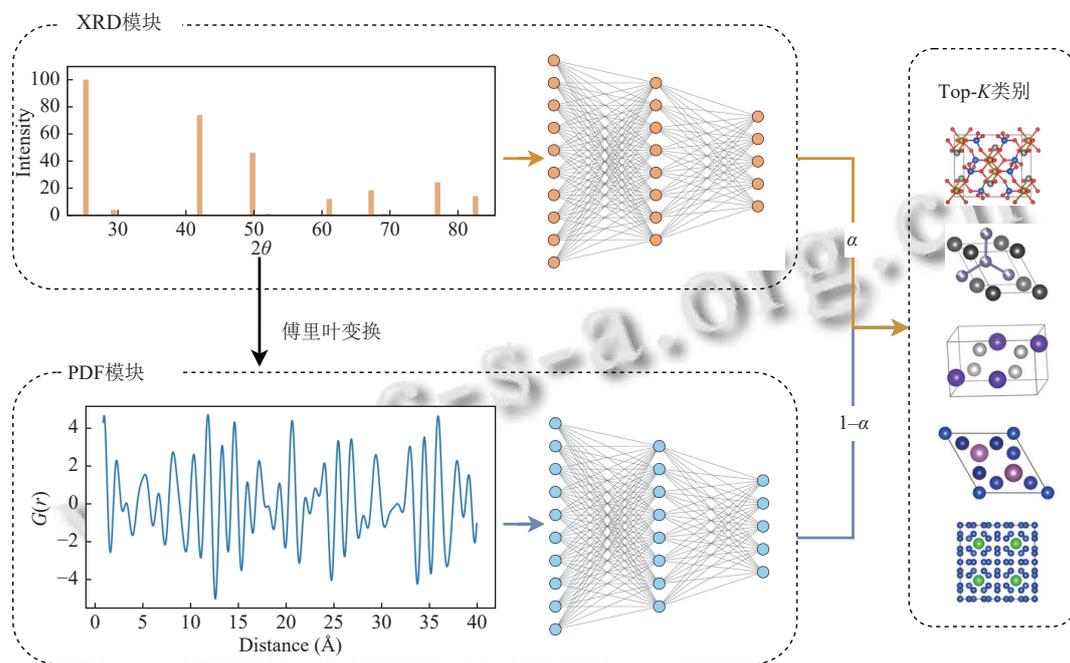


图3 基于XRD和PDF的多模态投票模型

XRD和PDF模块所使用的CNN模型由多个关键层组成。首先,模型包含1个一维卷积层,该层配备64个滤波器,卷积核大小为60,激活函数为ReLU。随后通过池化层(窗口大小为3,步幅为2)对特征进行下采样,减少维度的同时保留关键信息。接下来,模型通

过Flatten层将多维特征展平,并加入Dropout层(丢弃率为0.5)防止过拟合。紧接着是2个全连接层,分别包含128和64个神经元,每层后都附加了归一化层。最后是1个全连接层,该层使用Softmax激活函数来生成用于多分类的概率分布。训练过程中,模型的优化器

为 Adam (学习率为 0.001), 损失函数是交叉熵损失函数, 批次大小为 32, 训练轮数设定为 50。

晶体结构预测器和空间群预测器的模型架构共享相同的 XRD 模块和 PDF 模块, 不同的是最后的全连接层的输出维度, 晶体结构预测器中全连接层输出维度为 100, 空间群预测器中全连接层输出维度为 87。此外, 对于晶体结构预测器, 融合模块的 α 值为 0.4, 而空间群预测器的 α 值为 0.5。

3.2 训练数据

本研究使用的数据集包含 63 963 个无机化合物, 这些化合物来自无机晶体结构数据库 (inorganic crystal structure database, ICSD)^[15]中最常见的 100 种晶体结构。我们按照 8:1:1 的比例将其划分为训练集、验证集和测试集。我们使用 Pymatgen 包的 XRDCalculator^[16]模块来从这些化合物的晶体信息文件 (crystallographic information file, CIF) 中提取结构信息, 并生成相应的 XRD 数据 (衍射角 2θ 范围为 $0-90^\circ$)。最后将 XRD 谱图的衍射峰强度映射到一个长度为 4 096 的序列并将其作为模型的输入。

PDF 数据是从 XRD 谱图中计算得到的。计算流程: 首先, 将 XRD 数据中的衍射角 2θ 转换为 Q 空间, 公式如下:

$$Q = \frac{4\pi \sin \theta}{\lambda} \quad (1)$$

其中, λ 为入射 X 射线的波长。然后应用高斯平滑去除背景噪声, 以去除不相干散射的影响。将背景噪声平滑后的相干散射强度 $I(Q)$ 归一化得到散射函数 $S(Q)$:

$$S(Q) = \frac{I(Q)}{\max(I(Q))} - 1 \quad (2)$$

随后, 通过正弦傅里叶变换来生成 PDF 数据, 具体公式如下:

$$G(r) = \frac{2}{\pi} \int_{Q_{\min}}^{Q_{\max}} Q \cdot S(Q) \cdot \sin(Qr) dQ \quad (3)$$

其中, r 表示真实空间中的距离 ($1-40 \text{ \AA}$, $\text{\AA}=1\text{E}-10 \text{ m}$)。最后我们得到序列长度为 1000 的 PDF 数据。

4 知识图谱构建

为了帮助 LLM 理解晶体结构与空间群之间的关系, 并为后续的推理和分析任务提供支持, 我们基于 Neo4j 图数据库^[17]构建了一个知识图谱^[18]。构建知识图

谱以及 LLM 代理与知识图谱交互的过程主要包括以下 4 个步骤。

(1) 数据预处理: 从包含 63 963 个无机化合物的 ICSD 数据库^[15]中, 按照 8:1:1 的比例对数据进行训练集、验证集和测试集划分, 并从中提取出训练集数据作为本文知识图谱的数据来源。

(2) 实体和关系识别: 从训练集中获得相应的 CIF 文件, 从中抽取出晶体结构和空间群信息, 并将这两者作为实体, 然后将实体之间的关系 (例如“属于”) 转化为边, 在节点之间建立连接。

(3) 构建三元组并导入 Neo4j: 将识别出的实体和关系转化为三元组 (实体 1, 关系, 实体 2), 并将这些三元组导入到 Neo4j 图数据库中。最终, 该过程构建了一个包含 100 个晶体结构实体、73 个空间群实体和 156 条边的知识图谱。这一过程使得知识图谱得以存储, 并为 LLM 后续的图数据库查询和推理提供了基础。

(4) LLM 代理与知识图谱的交互过程: 首先, LLM 代理分析自然语言问题, 识别出需要查询的实体和关系, 并将其转化为 Neo4j 图数据库的查询语言; 然后, LLM 代理通过调用 Neo4j 图数据库的接口执行查询, 获取相关的实体和关系数据。最后, Neo4j 图数据库执行查询语句, 并将查询结果返回给 LLM 代理进行进一步的推理和分析。

5 实验分析

5.1 晶体结构预测器和空间群预测器的性能分析

5.1.1 对比实验

在晶体结构预测实验中, 我们分别比较了多种基线模型的分类准确率, 评估了基于 XRD 和 PDF 数据的随机森林 (XRD-RF, PDF-RF)、KNN (XRD-KNN, PDF-KNN)、决策树 (XRD-DT, PDF-DT)、CrySTINet^[7]以及本文提出的 CSPredictor 的分类性能。通过使用 6 个不同的随机种子划分训练集、验证集和测试集, 我们进行了多次实验, 并报告了多次实验结果的平均值, 实验结果如图 4 所示。

结果表明, 基于 PDF 数据的模型 (PDF-RF, PDF-KNN, PDF-DT) 性能均显著优于基于 XRD 数据的模型 (XRD-RF, XRD-KNN, XRD-DT), 这反映出 PDF 数据能够提供与晶体结构更相关的特征信息。此外, 我们提出的 CSPredictor 在实验中表现优异, 在晶体结构预测任务中达到了 95.4% 的 Top-1 准确率, 显著优于其

他模型. 这一优势源于 CSPredictor 通过融合 XRD 和 PDF 数据, 能够更全面地捕捉材料的特征信息, 从而有效提升模型在晶体结构预测任务中的表现.

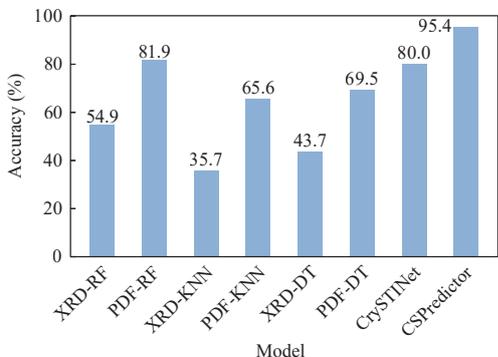


图4 不同方法的晶体结构预测准确率对比

对于空间群预测任务, 我们比较了相同的基线模型、已有研究中使用的 CNN^[9]方法, 以及本文提出的 SGPredictor 模型. 实验数据和设置与晶体结构预测实验一致, 实验结果如图5所示. 首先, 在晶体的空间群预测任务中, 对于单模态数据, 在随机森林、KNN 和决策树3种模型下, PDF 数据的表现均优于 XRD 数据. 这反映了 PDF 数据通过傅里叶变换能够体现更丰富的晶体结构信息. 其次, SGPredictor 模型相较于其他模型表现更为突出, 达到了 89.1% 的 Top-1 准确率, 显著优于其他基线模型. 这一结果验证了 SGPredictor 模型在空间群预测任务中的优势, 尤其是其对多模态数据的有效融合, 提升了预测的准确性和鲁棒性.

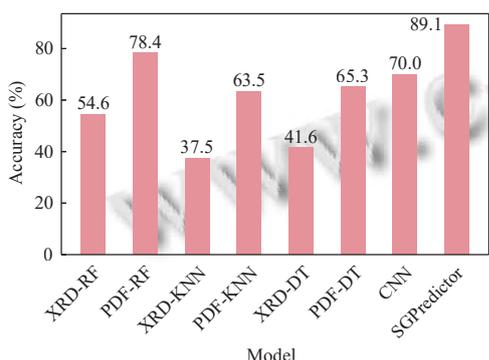


图5 不同方法的空间群预测准确率对比

5.1.2 消融实验

(1) 单模态模型性能对比: 为评估单模态模型在晶体结构和空间群预测任务中的性能, 我们对比了3种模型的性能: 基于 XRD 数据的 CNN 模型 (XRD-CNN)、基于 PDF 数据的 CNN 模型 (PDF-CNN) 及结合两种

模态的多模态融合模型 (CSPredictor 和 SGPredictor). XRD-CNN 模型以长度为 4096 的 XRD 数据作为输入, 通过 CNN 提取特征并输出分类结果; PDF-CNN 模型以长度为 1000 的 PDF 数据作为输入, 通过 CNN 模型完成分类任务; CSPredictor 和 SGPredictor 融合了两种模态的特征. 为确保评估结果的可靠性, 实验使用 6 个随机种子对数据集进行划分, 并取 6 组实验结果的平均值进行分析. 图6和图7分别展示了3种模型在晶体结构预测和空间群预测任务上的实验结果.

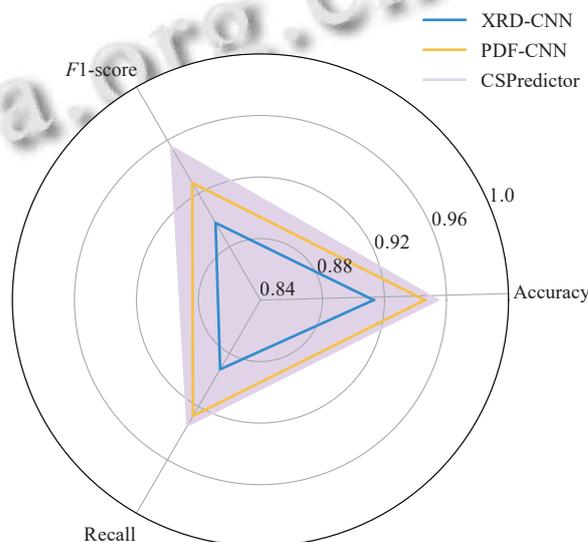


图6 使用不同模态场景下晶体结构预测性能对比

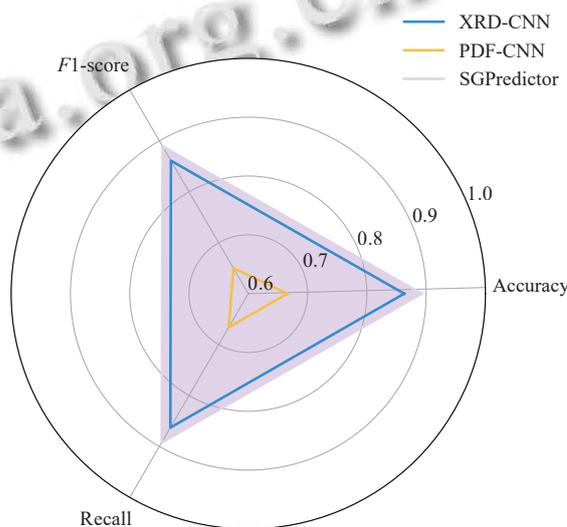


图7 使用不同模态场景下空间群预测性能对比

对于晶体结构预测任务, CSPredictor 在分类任务的准确率 (95.4%)、F1 分数 (0.955) 和召回率 (93.4%) 均优于 XRD-CNN 和 PDF-CNN. 具体来说, PDF-CNN

的性能优于 XRD-CNN, 这表明 PDF 数据在此任务中的潜在特征信息更为丰富. 同时, CSPredictor 利用多模态数据的互补性及动态加权机制进一步提升了模型的准确率与稳定性.

对于空间群预测任务, XRD-CNN 的性能显著优于 PDF-CNN, 这表明 XRD 数据保留了更多与晶体对称性和空间群相关的细节信息, 而 PDF 数据在生成过程中可能损失了部分关键信息. 此外, 融合了 XRD 和 PDF 两种模态的 SGPredictor 模型表现最佳, 这表明通过结合两种模态的优势, 模型能够更全面地捕捉晶体的全局和局部结构特征. XRD 数据提供了有关晶体对称性和空间群的丰富信息, 而 PDF 数据则补充了晶体的局部结构信息, 二者的融合使得 SGPredictor 能够在空间群预测任务中实现更高的准确性和鲁棒性.

(2) 多模态投票融合中权值的影响: 为了研究多模态投票融合中权值 α 的影响, 我们在不同 α 值 (范围为 0–1, 步长为 0.1) 设置下进行多次实验, 并在 6 个不同随机种子划分的数据集上评估了模型的性能.

图 8 和图 9 的箱线图分别展示了 CSPredictor 模型在不同 α 值设置下的晶体结构预测的准确率和 F1 分数. 实验结果表明, 当 α 从 0 增大到 1 时, 模型性能先升后降, 并在 $\alpha=0.4$ (即 XRD 数据的权重为 0.4, PDF 数据的权重为 0.6) 时, 模型的性能达到最佳, 此时多次实验的准确率平均值为 95.4%, F1 分数平均值为 0.951. 当过度依赖单一模态 (如 $\alpha=0.0$ 或 $\alpha=1.0$) 时, 模型性能有所下降. 这表明适当的 α 值可充分发挥多模态信息的互补性, 体现了本文多模态投票融合机制在平衡 XRD 和 PDF 数据特征方面的优势.

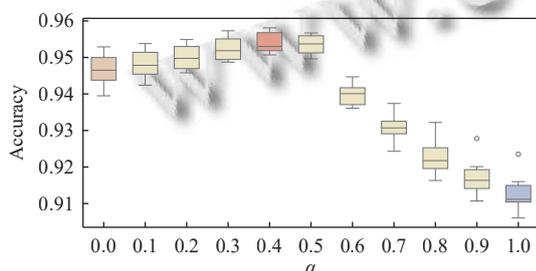


图 8 CSPredictor 在不同 α 值下的晶体结构预测准确率

对于空间群预测任务, 图 10 和图 11 分别展示了 SGPredictor 在不同 α 值设置下的准确率和 F1 分数. 根据实验结果, 当 $\alpha=0.5$, 即 XRD 和 PDF 数据的权重相等时, 模型的性能达到了最佳水平. 具体而言, 在此

配置下, 模型的准确率和 F1 分数均表现出显著的提升, 分别达到了 89.1% 和 88.9%. 在单模态情境下, 分别使用 XRD 和 PDF 模态进行空间群预测时, 模型的性能相对较低. 这些结果表明, 尽管单一数据源 (XRD 或 PDF) 能够提供一定的预测能力, 但缺乏对另一模态的补充时, 模型的性能受到限制. 综上所述, 通过合理的多模态融合和权值调整, 空间群预测任务的准确性得到了显著提升, 这验证了多模态方法在处理复杂晶体结构问题中的潜力.

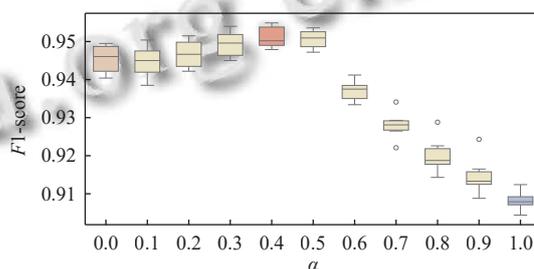


图 9 CSPredictor 在不同 α 值下的晶体结构预测 F1 分数

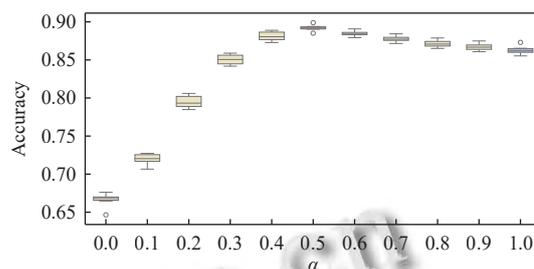


图 10 SGPredictor 在不同 α 值下的空间群预测准确率

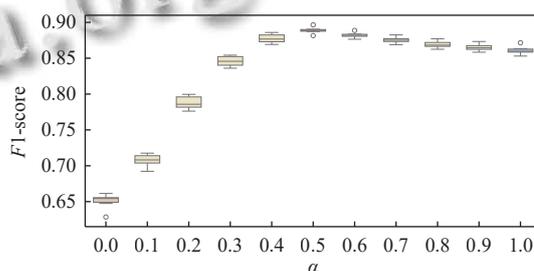


图 11 SGPredictor 在不同 α 值下的空间群预测 F1 分数

5.2 基于 LLM 的晶体结构和空间群预测结果

为了验证 LLM 推理机制通过利用工具 Top-5 预测结果来提升晶体结构和空间群预测准确率的能力, 我们设计了一组对比实验来评估 LLM 推理优化的有效性. 本实验基于晶体结构预测器和空间群预测器的预测结果设置了如下 3 个对比.

(1) 工具 Top-1 预测: 直接使用工具的 Top-1 结果

作为最终预测;

(2) 工具 Top-5 预测: 使用工具的 Top-5 结果作为最终预测结果, 以此评估潜在准确率;

(3) LMM-CSP 预测: 将工具的 Top-5 预测结果与知识图谱信息相结合, 通过 LLM 推理优化, 生成最终的 Top-1 预测结果.

实验在相同的测试数据集上进行, 以确保结果的可比性, 实验结果如表 1 所示. 工具 Top-5 预测结果具有较高的潜在准确率, 显著高于工具 Top-1 预测的准确率. 通过 LLM 结合知识图谱对 Top-5 结果进行优化 (LLM-CSP), 晶体结构和空间群的最终预测准确率分别提升至 97.5% 和 98.7%, 接近 Top-5 的理想情况. 分析表明, 在 LLM 基于潜在候选结果进行推理时, 知识图谱提供的物理约束信息有效减少了预测误差, 提升了模型的预测性能. 综上所述, 利用 LLM 结合多模态数据和知识图谱进行推理的方法, 显著提升了预测准确率, 为材料科学中的复杂推理任务提供了高效、可靠的解决方案.

表 1 使用 LLM 推理前后准确率对比 (%)

测试模型和指标	晶体结构预测	空间群预测
工具Top-1准确率	95.4	89.1
工具Top-5准确率	99.6	98.8
LMM-CSP准确率	97.5	98.7

6 结论与展望

本文提出了一个面向多模态晶体结构预测的 LLM 代理框架, 该框架利用知识图谱来增强 LLM 的推理能力, 使其能够深入理解材料的不同特征之间的内在关联, 并将这些关系融入推理过程, 从而保证分析的全面性. 同时, 该框架通过多模态投票模型融合 XRD 和 PDF 数据, 在晶体结构和空间群预测任务中表现出色, 展示出其在提高晶体学预测效率和准确性方面的巨大潜力. 这一进展为自动化晶体分析的研究和创新开辟了新的方向, 提供了一种可靠且可扩展的高通量材料分析解决方案. 随着未来更多工具的不断加入以及更大规模的数据集被整合, 该框架有潜力革新基于 XRD 的材料学研究, 并推动材料和晶体学领域的发展.

参考文献

- Long CJ, Bunker D, Li X, *et al.* Rapid identification of structural phases in combinatorial thin-film libraries using X-ray diffraction and non-negative matrix factorization. *Review of Scientific Instruments*, 2009, 80(10): 103902. [doi: 10.1063/1.3216809]
- Li MX, Sun YT, Wang C, *et al.* Data-driven discovery of a universal indicator for metallic glass forming ability. *Nature Materials*, 2022, 21(2): 165–172. [doi: 10.1038/s41563-021-01129-6]
- 王倩男, 胡建波. 基于大型激光装置的材料塑性变形行为微观研究进展. *强激光与粒子束*, 2020, 32(11): 112010. [doi: 10.11884/HPLPB202032.200116]
- Szymanski NJ, Fu SA, Persson E, *et al.* Integrated analysis of X-ray diffraction patterns and pair distribution functions for machine-learned phase identification. *npj Computational Materials*, 2024, 10(1): 45. [doi: 10.1038/s41524-024-01230-9]
- Ziletti A, Kumar D, Scheffler M, *et al.* Insightful classification of crystal structures using deep learning. *Nature Communications*, 2018, 9(1): 2775. [doi: 10.1038/s41467-018-05169-6]
- Chitturi SR, Ratner D, Walroth RC, *et al.* Automated prediction of lattice parameters from X-ray powder diffraction patterns. *Journal of Applied Crystallography*, 2021, 54(6): 1799–1810. [doi: 10.1107/S1600576721010840]
- Chen L, Wang BX, Zhang WT, *et al.* Crystal structure assignment for unknown compounds from X-ray diffraction patterns with deep learning. *Journal of the American Chemical Society*, 2024, 146(12): 8098–8109. [doi: 10.1021/jacs.3c11852]
- Szymanski NJ, Bartel CJ, Zeng Y, *et al.* Adaptively driven X-ray diffraction guided by machine learning for autonomous phase identification. *npj Computational Materials*, 2023, 9(1): 31. [doi: 10.1038/s41524-023-00984-y]
- Liu CH, Tao Y, Hsu D, *et al.* Using a machine learning approach to determine the space group of a structure from the atomic pair distribution function. *Acta Crystallographica Section A: Foundations and Advances*, 2019, 75(4): 633–643. [doi: 10.1107/S2053273319005606]
- Zhao A, Huang D, Xu Q, *et al.* ExpeL: LLM agents are experiential learners. *Proceedings of the 38th AAAI Conference on Artificial Intelligence*. Vancouver: AAAI, 2024. 19632–19642. [doi: 10.1609/aaai.v38i17.29936]
- Guan YC, Wang D, Chu ZX, *et al.* Intelligent agents with LLM-based process automation. *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. Barcelona: ACM, 2024. 5018–5027. [doi: 10.1145/3637528.3671646]
- Boiko DA, MacKnight R, Kline B, *et al.* Autonomous

- chemical research with large language models. *Nature*, 2023, 624(7992): 570–578. [doi: [10.1038/s41586-023-06792-0](https://doi.org/10.1038/s41586-023-06792-0)]
- 13 Bran AM, Cox S, Schilter O, *et al.* Augmenting large language models with chemistry tools. *Nature Machine Intelligence*, 2024, 6(5): 525–535. [doi: [10.1038/s42256-024-00832-8](https://doi.org/10.1038/s42256-024-00832-8)]
- 14 Szymanski NJ, Rendy B, Fei YX, *et al.* An autonomous laboratory for the accelerated synthesis of novel materials. *Nature*, 2023, 624(7990): 86–91. [doi: [10.1038/s41586-023-06734-w](https://doi.org/10.1038/s41586-023-06734-w)]
- 15 Bergerhoff G, Hundt R, Sievers R, *et al.* The inorganic crystal structure data base. *Journal of Chemical Information and Computer Sciences*, 1983, 23(2): 66–69. [doi: [10.1021/ci00038a003](https://doi.org/10.1021/ci00038a003)]
- 16 Ong SP, Richards WD, Jain A, *et al.* Python materials genomics (pymatgen): A robust, open-source python library for materials analysis. *Computational Materials Science*, 2013, 68: 314–319. [doi: [10.1016/j.commatsci.2012.10.028](https://doi.org/10.1016/j.commatsci.2012.10.028)]
- 17 Robinson I, Webber J, Eifrem E. *Graph Databases: New Opportunities for Connected Data*. 2nd ed. Sebastopol: O'Reilly, 2015. 10–238.
- 18 张天成, 田雪, 孙相会, 等. 知识图谱嵌入技术研究综述. *软件学报*, 2023, 34(1): 277–311. [doi: [10.13328/j.cnki.jos.006429](https://doi.org/10.13328/j.cnki.jos.006429)]

(校对责编: 王欣欣)