

基于知识增强大模型的电催化剂设计

王露笛¹, 陈鸣^{1,2}, 崔文娟^{1,2}

1. 中国科学院计算机网络信息中心, 北京 100083;

2. 中国科学院大学, 北京 100049

摘要

作为实现可持续碳循环的重要手段, 开发高性能的电催化剂是未来可持续发展的关键, 而推荐具有创新性的有价值制备方案是提升电催化开发效率的有效手段之一。基于电催化领域的科技文献, 邀请领域专家进行知识体系的建设与知识抽取, 形成电催化剂领域知识库。此外, 利用相关文献数据对通用大语言模型进行微调与知识增强, 共同实现面向目标产物、材料类别和调控方法类别的制备方案推荐, 辅助电催化剂的设计。实验表明, 知识增强大模型推荐的制备方案的有效性和创新性都有一定的提升。

关键词

电催化剂; 知识增强; 大模型

中图分类号: TP391.1

文献标志码: A

doi:10.11959/j.issn.2096-0271.2025028

Design of electrocatalysts based on knowledge enhanced LLMs

WANG Ludi¹, CHEN Ming^{1,2}, CUI Wenjuan^{1,2}

1. Computer Network Information Center, Chinese Academy of Sciences, Beijing 100083, China

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

Abstract

As an important means of achieving sustainable carbon cycling, developing high-performance electrocatalysts is the key to the sustainable development in future, and recommending innovative and valuable preparation solutions is one of the effective ways to improve the efficiency of electrocatalytic development. This paper is based on scientific and technological literature in the field of electrocatalysis, inviting domain experts to construct a knowledge system and extract knowledge, forming a domain knowledge base. In addition, this paper utilizes literature data to fine tune and enhance knowledge of universal big language models, jointly completing preparation scheme recommendations for target products, material categories, and regulation method categories, and assisting in the design of electrocatalysts. Experimental results have shown that the preparation solutions provided by knowledge enhanced large language models have certain improvements in both effectiveness and innovation.

Key words

electrocatalysts, knowledge enhancement, large language model

0 引言

电催化剂和光催化剂是未来可持续发展的关键，它们可以产生清洁燃料，为全球变暖、环境污染提供解决方案。电催化剂作为推动未来可持续发展的重要因素，具备生成清洁燃料、缓解全球变暖效应以及解决环境污染问题的潜力。优化催化剂设计流程及深入理解光/电催化机制，对于提升催化剂效能具有决定性意义。尽管当前已有大量实验研究致力于开发高效的光/电催化剂，但实验试错不仅耗费大量人力，还存在可重复性方面的问题。理论计算与模拟虽然能加速催化剂的筛选过程，但实际催化剂表面结构远比理论模型假设的结构更复杂且非理想化，这无疑降低了计算与模拟结果在催化剂设计中的应用可靠性。此外，除了在原子/分子层面深入探究催化剂的细致性质外，亟须在更宏观的领域层面捕捉更多影响催化剂设计可靠性的因素，以指导催化剂的设计。实际上，海量的科技文献蕴藏着丰富的知识，如制备步骤、材料性质等。若能有效提取与催化剂相关的信息并整合成数据集，将极大提升新催化剂开发的效率与成功率。

1 研究背景

随着材料领域计算与模拟数据的不断累积，越来越多的材料分析软件被发布，如美国能源部开发的 Materials Project 网站、SUNCAT 机构构建的网络平台 Catalysis-Hub.org、IBM 设计的基于分子变压器的独特数字化学工具。但这些软件仍然以密度泛函理论（density functional

theory, DFT）计算为核心，以材料组成数据为基础，无法从语义层面对领域中已有的知识进行总结与分析，缺乏可解释性。

针对上述问题，研究人员尝试从海量领域科技文献中抽取相关领域知识，并借助自然语言处理技术推动“知识驱动”的创新发现。随着技术的不断发展，预训练模型在自然语言处理领域展现出强大的能力，可以被分为两大类：一类是 BERT 类模型^[1-3]，该类模型基于 Transformer 架构的编码器构建而成，主要用于语言理解任务；另一类是 GPT 类模型^[4-5]，该类模型基于 Transformer 架构的解码器构建而成，主要用于语言生成任务。随着 GPT3.5 的发布，后者在问答任务上展现出惊人的能力。尽管当前的大型语言模型（如 C4^[6]、Pile^[7]）在通用领域表现出色，但这些模型往往无法满足催化科学家的特定需求，这很大程度上归因于催化剂结构特征和性能分析等可靠知识的缺乏。

针对当前电催化材料设计过程中的幻觉和不可解释性问题，本文在电催化领域的科技文献上对通用大语言模型（large language model, LLM，简称大模型）进行微调，使模型能够更好地理解领域知识；同时，遵循催化剂制备过程中的知识体系，构建包括材料、产物、性质、方法在内的领域知识图谱，并通过提升工程将领域知识路径灌入微调后的大模型中，进一步减少大模型的幻觉、提升其可解释性。

本文的创新性如下：①利用电催化领域的科技文献对通用大模型进行微调，形成领域微调大模型；②基于领域知识体系形成面向电催化材料调控过程的领域知识图谱；③基于路径检索方法，利用领域知识图谱增强大模型的调控方法推荐能力，辅助电催化的设计。经过专家评估，知识嵌入的大模型具备更强的领域理解能力，

其推荐的设计方法符合科学原理，具备较强的可行性。

2 基于知识增强大模型的电催化设计

2.1 整体架构

基于知识增强大模型的电催化设计的整体框架如图1所示。首先，在领域文献数据上进行大模型的微调，得到电催化领域的垂类大模型；其次，邀请相关专家依据设定的制备过程标签体系对文献中的知识实体进行标注，并利用这些数据构建面向制备过程推荐的知识图谱；最后，基于该图谱进一步增强垂类大模型的领域知识的理解能力，共同完成制备方案的推荐。

2.2 领域数据准备

首先，使用关键词“CO₂”“reduction”和“electro*”作为主题索引，从Web of Science网站上导出22 000余篇文章的元数据（如标题、doi、摘要）。然后，根据专家定义的规则对导出的元数据进行逐步过滤；在此基础上，由领域专家进一步筛选，最终获得835篇与CO₂还原反应相关的电催化材料设计文献。本文使用PyMuPDF工具自动化批量提取文献的元数据，并将其处理成JSON格式用于大模型的预训练。

2.3 垂直领域大模型预训练

将获取的文献数据作为无监督语料，对通用开源大模型进行微调，完成领域知识的注入。Vicuna 13b模型^[8]采用Transformer结构，以其高质量的回答和

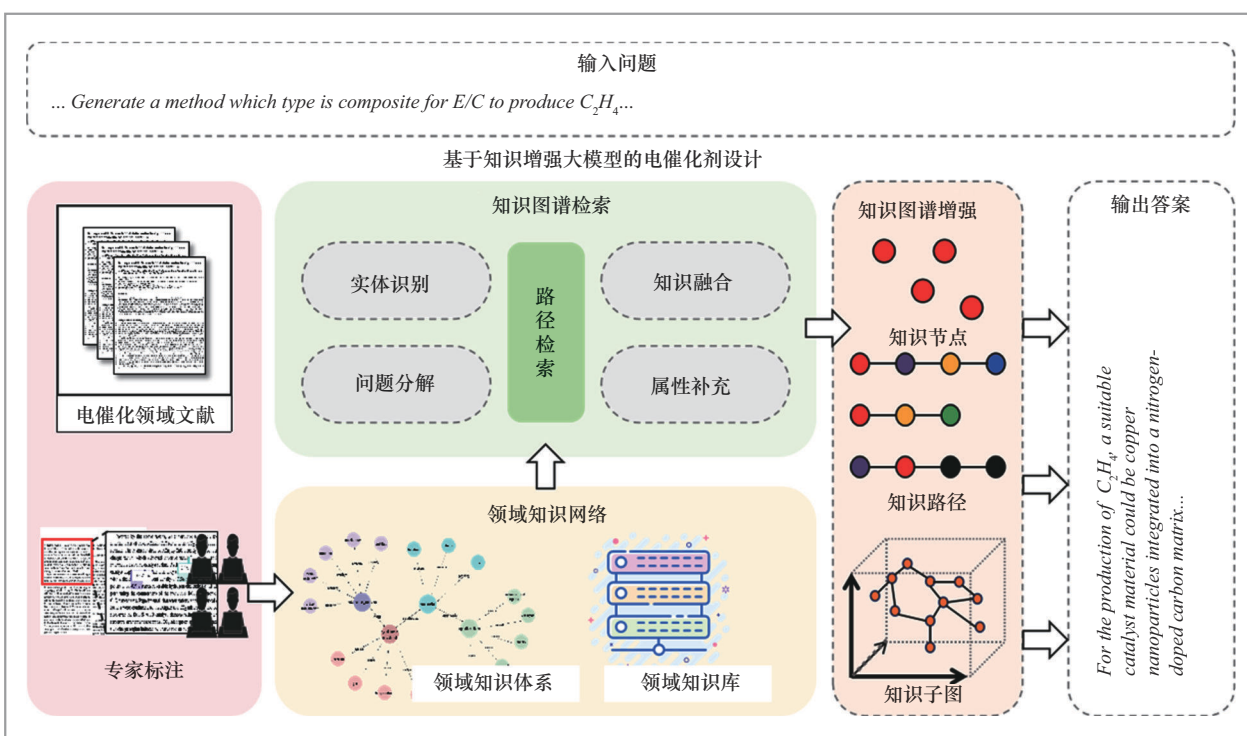


图1 基于知识增强大模型的电催化设计的整体框架

接近 OpenAI ChatGPT 的性能而闻名。因此，选取该模型作为底座大模型，并使用 LoRA^[9] 方法对该模型进行微调。LoRA 是由微软研究人员开发的一种技术，用于大模型的微调。该方法冻结预训练模型的权重参数，然后在每个 Transformer 块中注入可训练的层。该方法不需要重新计算模型权重参数的梯度，大大减少了训练的计算量。研究表明，LoRA 的微调质量与全模型微调相当，因此，本文采用 LoRA 方法对模型进行微调。

2.4 电催化领域知识图谱构建

电催化剂的制备过程比较复杂，涉及材料类别、产物、制备方法分类、催化效率等一系列要素，单纯依靠无监督语料的微调无法使大模型充分理解这些要素之

间的关系。因此，邀请领域专家依据构建的领域知识体系（如图 2 所示）对上述要素进行标注，形成电催化领域知识图谱。电催化领域知识体系包含三大类，分别为电催化反应的产物种类、电催化剂的材料种类，以及催化调控因素。笔者通过专家标注和训练相关知识抽取模型的方式抽取领域知识，最终形成了拥有 6 086 个知识实体的领域知识图谱。领域知识图谱的示例如图 3 所示。

2.5 知识增强大模型的制备方案推荐

针对制备方案推荐，利用领域大模型和领域知识图谱进行协同推荐。该流程的输入为预计产物、基底材料类别和调控方法类别，基于设计好的提示将数据填入指定位置，形成输入文本数据；同时，在领

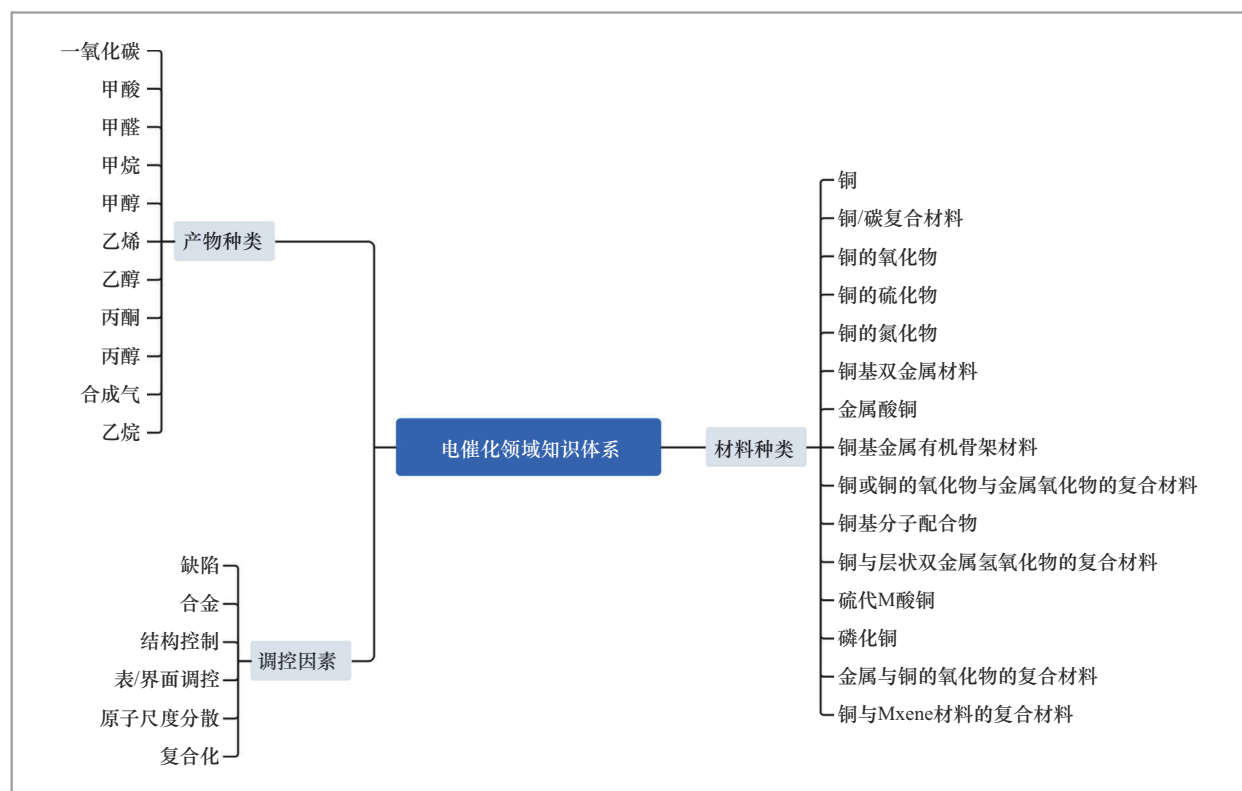


图2 电催化领域知识体系

域知识图谱中查询出符合输入数据的知识路径，将其填充到大模型的上下文中作为参考信息。知识增强大模型的制备方法推荐流程如图4所示。

3 实验结果

3.1 实验参数设置

本文实验在 NVIDIA A100 GPU 上训练，实验参数见表1。其中，Batch size是在一次训练迭代中使用的样本数量；Learning rate用于控制模型在训练过程中更新参数的步长；Lora r 表示低秩矩阵的秩；Lora alpha是一个缩放因子，用于调整LoRA模块输出的缩放比例；Lora dropout表示在训练过程中随机丢弃的参数比例；Temperature用于调整概率分布平滑程度或随机性的参数，较高的Temperature使生成结果更具随机性和多样性；Top p 是核采用概率阈值；Top k 为备选词数量；Beams表示搜索的候选数量；Max tokens是允许的最大标记数量。

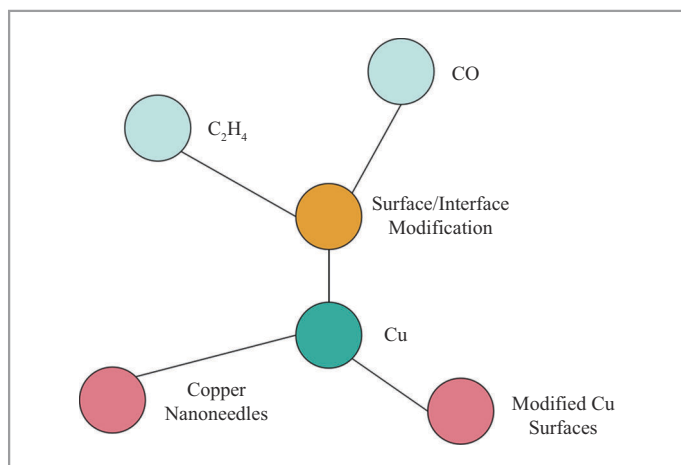


图3 领域知识图谱示例

表1 实验参数

通用大模型微调阶段		制备方法推荐阶段	
Batch size	10	Temperature	0.5
Learning rate	3×10^{-4}	Top p	0.75
Lora r	8	Top k	40
Lora alpha	32	Beams	4
Lora dropout	0.1	Max tokens	1 248

3.2 实验结果

为了验证本文方法的有效性，对比原始

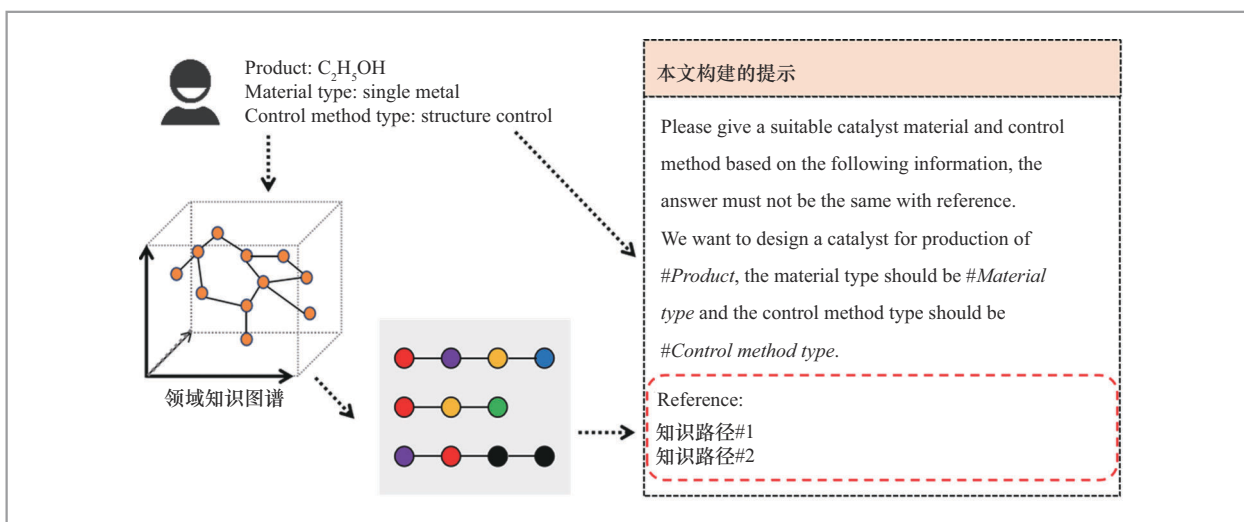


图4 知识增强大模型的制备方法推荐流程

大模型、检索增强大模型和知识增强大模型的回答，并邀请领域专家对推荐的制备方案进行评估。以 C_2H_4 为目标产物，组合不同的材料类别与调控类别对以上模型进行提问，大模型推荐的制备方案示例如图5所示。经过专家评判，知识增强大模型在制备方案的创新性和价值性上都有所提升。

为了进一步证明本文所提方法的有效性，邀请领域专家对30组回答进行评估。在评估过程中，隐藏了模型信息并打乱了答案顺序，专家根据问题中的参数设置（包括材料类别、产物、调控方法类别）评估答案的可行性，领域专家对制备方案的评估结果见表2。

由表2可知，相比于原始大模型，检索增强大模型和知识增强大模型推荐的制备方案的可行性更高。这可能是因为后两个大模

型根据问题中的产物、材料类别和调控方法类别，进行了相关科技文献或知识路径的搜索与提示填充，使原始大模型能够以相同或相似的推荐制备方案为参考进行提示生成。相比于检索增强大模型，知识增强大模型能够更加准确地搜索出相关知识路径，其推荐的制备方案的可行性更高。

4 结束语

作为实现可持续碳循环的重要手段，开发高性能的电催化剂是未来可持续发展的关键。然而，影响电催化剂的合成、组成、结构及性能的因素较多，难以从有限的实验中揭示复杂的关联关系。理论计算和模拟可加速新型催化剂的筛选和合成，但难以对催化剂的表面结构和合成机理进行有效描述。

本文基于电催化领域的科技文献，邀请领域专家进行知识体系设计与标注，形成领域知识图谱；在领域科技文献上进行模型的预训练，为模型灌入领域知识；用

表2 领域专家对制备方案的评估结果

模型	可行方案数量/个	可行方案占比
原始大模型	14	47%
检索增强大模型	19	63%
知识增强大模型	21	70%

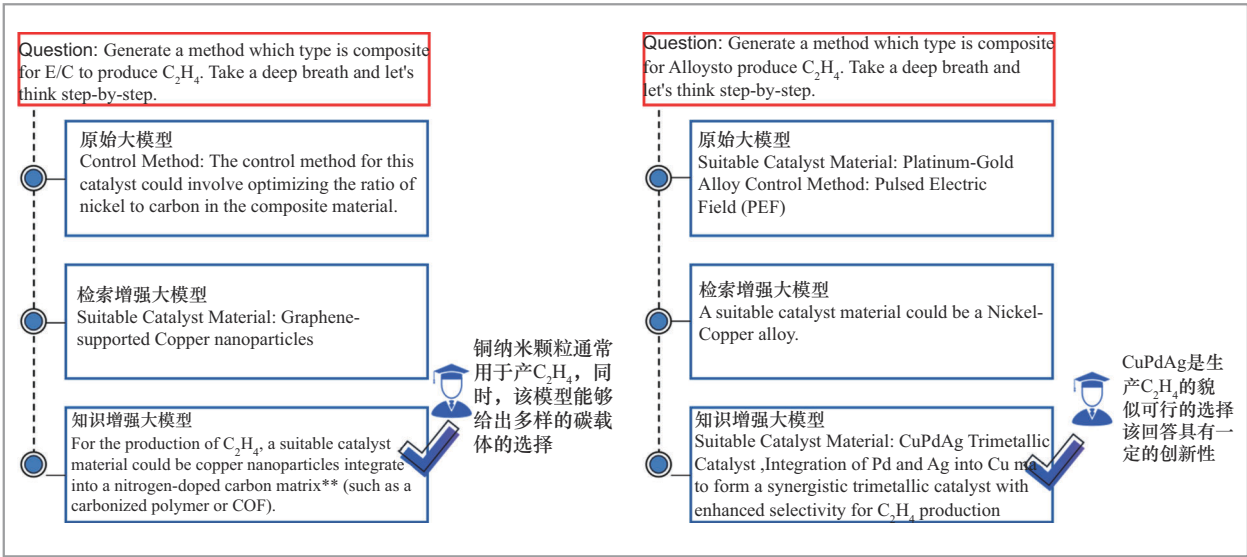


图5 不同大模型推荐的制备方案示例

知识图谱进一步增强模型在特定条件下的推荐能力,实现制备方案的推荐。实验表明,知识增强大模型推荐的制备方案更加丰富,在一定程度上提高了大模型对新型电催化材料制备方案的生成能力。未来将引入以谱学数据为代表的多模态信息,实现更加具有针对性的制备方案推荐,进一步激发大模型的“涌现”能力。

参考文献:

- [1] DEVLIN J, CHANG M W, LEE K, et al. BERT: pre-training of deep bidirectional Transformers for language understanding[EB]. arXiv preprint, 2018, arXiv: 1810.04805.
- [2] LIU Y H, OTT M, GOYAL N, et al. RoBERTa: a robustly optimized BERT pretraining approach[EB]. arxiv preprint, 2019, arxiv: 1907.11692.
- [3] CLARK K, LUONG M T, LE Q V, et al. ELECTRA: pre-training text encoders as discriminators rather than generators [EB]. arXiv preprint, 2020, arXiv: 2003.10555.
- [4] BABES-VROMAN M, MACGLASHAN J, GAO R Y, et al. Learning to interpret natural language instructions[C]//Proceedings of the Second Workshop on Semantic Interpretation in an Actionable Context. Stroudsburg: ACL, 2012: 1-6.
- [5] BROWN T B, MANN B, RYDER N, et al. Language models are few-shot learners[EB]. arXiv preprint, 2020, arXiv: 2005.14165.
- [6] RAFFEL C, SHAZEER N, ROBERTS A, et al. Exploring the limits of transfer learning with a unified text-to-text transformer[EB]. Journal of Machine Learning Research, 2020, 21(1): 5485-5551.
- [7] GAO L, BIDERMAN S, BLACK S, et al. The pile: an 800 GB dataset of diverse text for language modeling[EB]. arXiv preprint, 2020, arXiv: 2101.00027.
- [8] ZHENG L M, CHIANG W L, SHENG Y, et al. Judging LLM-as-a-judge with MT-bench and Chatbot Arena[EB]. arXiv preprint, 2023, arXiv: 2306.05685.
- [9] HU E J, SHEN Y L, WALLIS P, et al. Lora: low-rank adaptation of large language models[EB]. arXiv preprint, 2021, arXiv: 2106.09685, 2021.

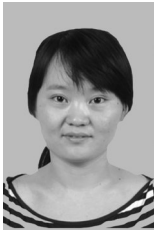
作者简介



王露笛 (1992-), 女, 博士, 中国科学院计算机网络信息中心高级工程师, 主要研究方向为交叉领域人工智能应用。



陈鸣 (2002-), 女, 中国科学院计算机网络信息中心硕士生, 主要研究方向为人工智能算法研究。



崔文娟（1986-），女，博士，中国科学院计算机网络信息中心副研究员，主要研究方向为知识图谱构建与应用。

收稿日期: 2024-09-25

通信作者: 崔文娟, wenjuancui@cnic.cn

基金项目: 中国科学院重点部署项目 (No.RCJJ-145-24-20)

Foundation Item: Key Research Program of the Chinese Academy of Sciences (No.RCJJ-145-24-20)