文章编号：1001-4632（2025）06-0044-11
引用格式：史小萌，徐功博，蔡国庆，等. 铁路路基压实施工知识图谱构建与检索增强方法[J]. 中国铁道科学，2025，46(6)：44-54.
Citation：SHI Xiaomeng，XU Gongbo，CAI Guoqing，et al. Knowledge Graph Establishment and Retrieval-Augmented Method for Railway Subgrade Compaction Construction [J]. China Railway Science，2025，46 (6)：44-54.

铁路路基压实施工知识图谱构建与检索增强方法

史小萌[1,2]，徐功博[1,2]，蔡国庆[1,2]，王子南[1,2]，邓志云[1,2]，刘保国[1,2]

（1. 北京交通大学 城市地下工程教育部重点实验室，北京 100044；
2. 北京交通大学 土木建筑工程学院，北京 100044）

摘 要：针对通用模型在铁路路基压实施工应用中存在的知识时效性滞后、专业知识理解不足及输出稳定性欠缺等问题，提出1种面向铁路路基压实施工领域的知识图谱构建与检索增强方法。首先，构建涵盖规范标准、专业书籍、学术论文及工程资料的专业语料库；其次，设计Bert-BiLSTM-CRF模型进行实体抽取，并融合层次聚类与DeepSeek-Prompt技术完成关系抽取与推理，构建包含1 315个实体与1 003组关系的铁路路基压实领域知识图谱；最后，通过检索增强生成（RAG）技术将该知识图谱集成至DeepSeek-R1模型，实现模型在铁路路基压实领域的检索增强。结果表明：融合知识图谱的模型在关键性能指标上显著超越了传统通用模型，其ROUGE-1值、ROUGE-L值和准确率分别达到0.321 2，0.383 4和0.822 9。该方法不仅可实现知识的系统化抽取与结构化表征，更能够有效提升模型对复杂专业知识的理解精准度与生成可靠性，为工程实践智能化水平的提升开辟了新的技术途径。

关键词：铁路路基；知识图谱；检索增强；路基施工；DeepSeek
中图分类号：U216.3 **文献标识码**：A **doi**：10.3969/j.issn.1001-4632.2025.06.05

近年来，ChatGPT，Claude和DeepSeek等大模型掀起技术变革浪潮，其强大的自然语言处理和数据解析能力，已广泛应用于医疗[1]、电力[2]等领域。然而，传统的通用大模型在复杂工程领域应用时仍存在局限性，通用模型在专业内容生成过程中准确性不足，常出现“幻觉”现象[3]。因此，探索适用于复杂工程领域的大模型优化方法具有重要意义。

知识图谱是1种结构化知识表达范式[4]，它以三元组作为逻辑载体，能够有效化解复杂工程领域知识复杂与碎片化问题。目前已有研究将知识图谱用于处理复杂工程知识，包括铁路隧道钻爆法施工安全质量进度智能管理[5]、隧道洞门[6]和支护设计[7]、地基加固方案智能决策[8]等方面。检索增强生成技术（RAG）是1种将结构化知识库与大语言模型深度融合的技术方法[9]。一些学者通过检索增强生成技术优化通用大模型，并在复杂工程中展开应用。王祥昊等[10]聚焦铁路应急处置领域，融合语义理解与生成模型构建专业知识库，有效提升了应急处置策略的生成效率和适用性；李子林等[11]提出了基于知识过滤的检索增强生成架构（KFRAG），通过减少噪声干扰使铁路领域问答任务的性能提升近10%，在多项评估指标上优于传统基线模型。检索增强生成技术能够为知识图谱以及大模型检索搭建技术桥梁，实现对专业领域知识的结构化组织与精准检索，增强大模型的专业性能。

路基压实作为铁路工程的核心环节，其智能化、数字化转型对提升工程质量[12]、施工效率以及保障安全运营意义重大[13]。本研究以铁路路基压实施工为研究对象，提出1种融合人工智能与大语言模型的知识图谱构建与检索增强方法，通过建

收稿日期：2025-04-01；修订日期：2025-11-25
基金项目：国家重点研发计划项目（2023YFB2603902）；国家自然科学基金青年项目（42307246）
第一作者：史小萌（1988—），男，江苏睢宁人，副教授，博士。E-mail：shixm@bjtu.edu.cn

立面向铁路路基压实施工领域的专业化语料库，并利用 Bert-BiLSTM-CRF 模型、层次聚类与 DeepSeek-Prompt 技术进行知识抽取，构建铁路路基压实领域知识图谱。在此基础上，以该知识图谱作为 DeepSeek-R1 大语言模型的外来知识库，通过 RAG 技术，实现模型对铁路路基压实领域知识的检索增强。

1　铁路路基压实知识维度

铁路路基压实领域的知识体系兼具专业性与复杂性，其知识广泛分布于规范标准、专业书籍等多种载体，单一知识源难以全面涵盖。将该领域的知识维度划分为 3 个层级："规范与标准""理论与观点"以及"方案与反馈"。其中，"规范与标准"是路基施工中的核心框架，任何工程都必须严格遵循这一限制框架。"理论与观点"是对众多前沿学术观点以及路基工程典型案例的提炼和总结，涵盖了多维度的知识。"方案与反馈"主要针对单个工程项目，将知识体系映射到具体的实际项目中，是对知识实践维度的有效补充。该知识维度体系见表 1。

表 1　铁路路基压实知识维度体系

维度	类别	形式	内容
规则与标准	规范	文本、图片、表格	国家和行业标准中规定的铁路路基压实技术的技术规范和质量标准
	技术标准	文本、图片、表格	具体的施工工艺要求、质量验收标准以及安全技术规程
理论与观点	专业书籍	文本	铁路路基压实的基础理论、关键技术以及实际工程经验
	专利	文本	铁路路基压实方面的技术创新和方法改进
	学术论文	文本	最新研究成果、技术创新以及工程案例研究
方案与反馈	调查报告	文本、表格	工程地质条件、岩土参数以及水文地质数据
	设计文件	文本、图片	路基设计方案、压实参数以及施工顺序安排
	施工日志	文本、表格	每日施工记录、质量控制数据以及异常情况处理
	组织文件	文本	施工组织设计、进度计划以及资源配置方案

2　融合人工智能与大语言模型的知识图谱构建方法

针对铁路路基压实知识体系及知识维度特点，提出 1 种融合人工智能与大语言模型的知识抽取方法，该方法分为语料基础、知识抽取、知识处理及知识储存 4 个环节，其技术路线如图 1 所示。在语料基础环节，选取相关规范标准、专业书籍、学术论文及工程资料形成专业语料库。在知识抽取环节，对实体与关系进行分类定义与人工标注，并使用 Bert-BiLSTM-CRF 模型进行预训练，在语料库中进行分类实体抽取，将抽取得到的实体进行层次聚类分析及关系推理，使用提示词模板 DeepSeek-Prompt 对关系抽取进行补充。在知识处理环节，对抽取到的实体和关系进行深度知识融合，整合碎片化知识，形成具有逻辑性和系统性的知识体系。在知识储存环节，使用 Cypher 语言将融合处理后的知识导入 Neo4j 数据库，实现知识存储与索引应用。

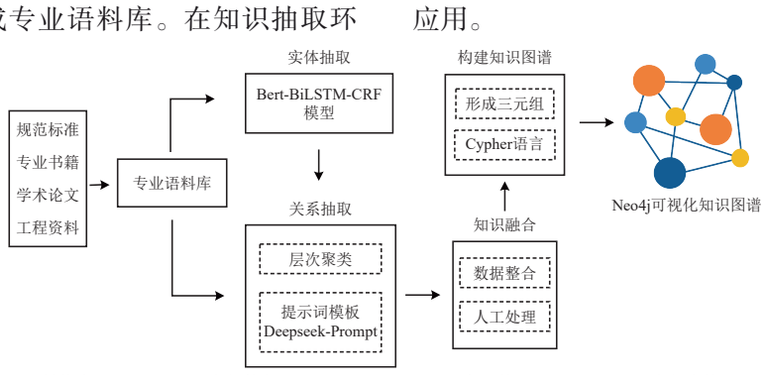


图 1　融合人工智能与大语言模型的知识抽取技术路线

2.1　基于 Bert-BiLSTM-CRF 模型的实体抽取

采用融合深度学习与概率图模型的 Bert-BiLSTM-CRF 实体抽取模型，通过多层次语义建模有效提升了实体识别的准确度与人工标注的一致

性。该模型包含5个层级：输入层、Bert编码层、BiLSTM特征提取层、CRF序列标注层和输出层。

模型训练依赖于适量的人工标注，因此需要先对铁路路基压实实体进行分类及定义，再通过实体定义与实体示例对类别边界进行设定，详情见表2。

表2 铁路路基压实实体分类与定义

实体类型	定义	示例
设计参数	设计阶段中铁路路基压实涉及的参数	高程、碾压面积
传感器参数	传感器可以感知到的相关参数	CMV、平整度
路基填料参数	填料相关性能指标	含水率、种类、级配
设备参数	可控制的施工设备相关参数	压实质量、碾压变数、行驶速度
压实质量	铁路路基压实施工的评价指标或评价描述	压实系数、承载力、地基系数
特殊现象与不良地质	施工中遇到的特殊现象及不良地质情况	暴雨天气、塌方
地质条件	工程及施工相关的地质条件与水文条件	降水量、土层性质
路基信息	描述路基性状及施工情况的表述	过渡段、填筑层
设备信息	描述设备性状的表述	最大质量、振动轮宽度

以输入“压实系数”为例，Bert-BiLSTM-CRF模型框架如图2所示。在模型计算过程中，输入文本首先被转换为插入特殊标记“CLS”和“SEP”的字符序列。Bert编码层通过多头自注意力机制对输入序列进行编码，生成包含全局上下文语义的词向量T_1，T_2，…，T_n[14]。BiLSTM特征提取层将Bert层输出的特征向量作为输入，通过协同的前向和后向LSTM网络进一步提取序列上下文依赖关系[15]。CRF序列标注层则基于BiLSTM层提取的特征序列生成标签预测[16]，采用BIO标注方案，为每个字符分配适当的实体类型标签，输出层中“O”表示该字符不属于表2中定义的实体类型，“B-压实质量”表示“压实系数”中字符“压”属于表2中定义的“压实质量”这一实体类型，且处在起始位置（B，Begin），“I-压实质量”表示字符“实”“系”“数”属于“压实质量”这一实体类型，但处在该实体的内部（I，Inside）。

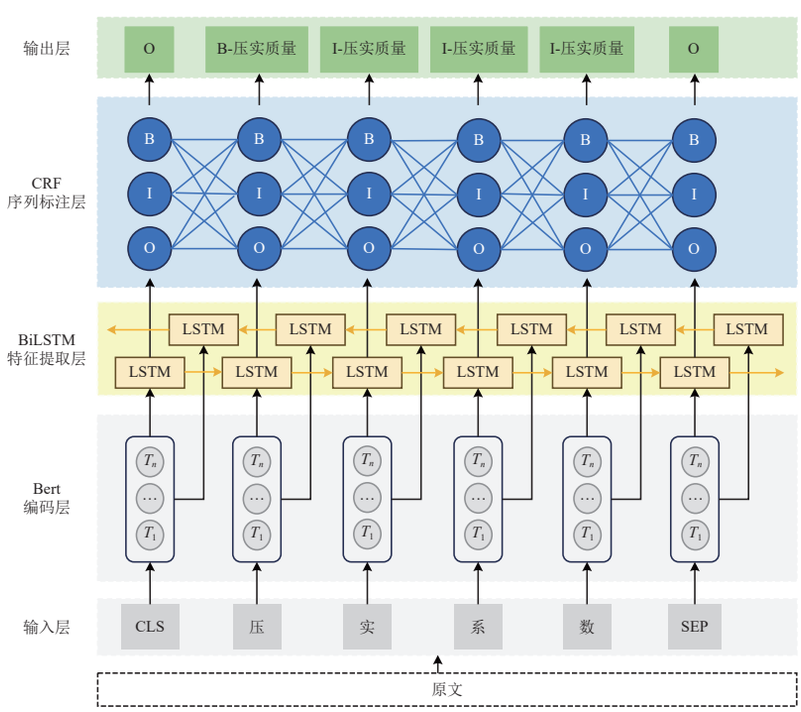


图2 Bert-BiLSTM-CRF模型框架

2.2 关系抽取

关系抽取首先要对关系进行定义，将铁路路基压实施工中的关系分为分类关系与非分类关系，对每种关系的定义见表3。

表 3　铁路路基压实关系定义

关系类型	关系信息	定义
分类关系	层级关系	描述实体间的从属或类别关系，即一个实体是另一个实体的类别或子集
	等同关系	描述两个实体在实际意义上是相同的，只是表达不同
非分类关系	时间关系	描述实体间的时间序列或阶段性顺序
	因果关系	描述一个实体或事件引发另一个实体或事件的关系
	限制关系	描述一个实体对另一个实体的约束或限制条件
	依据关系	描述一个实体或行为是基于另一个实体或规则执行的
	用途关系	描述一个实体的用途或功能
	属性关系	描述一个实体具有某种属性或特性
	空间关系	描述实体在空间上的相对位置关系
	影响关系	描述一个实体对另一个实体的作用或影响

2.2.1　基于层次聚类的关系推理

在铁路路基压实施工领域，实体间的关系往往表现为多样化的语义关联模式。采用层次聚类方法，通过分析实体在文本语料中的共现特征和语义分布，迭代合并最相似的实体，聚类发现实体间潜在的关系类型[17]，其示意图如图 3 所示。图中，A—F 表示铁路路基压实实体。初始阶段每个实体都是 1 个类，通过计算类与类之间的语义距离，将距离最近的类合并为 1 个类，不停重复合并，最终合成 1 个大类，形成层次结构。例如，“摊铺”“洒水”“压实”等实体可能因其在压实工序中的顺序关联而被聚为 1 类；“压实度”“承载力”“压实均匀性”等实体由于其描述压实效果的相关性而归为同组。这种基于语义关联的聚类结果，能够很好地反映铁路路基压实施工过程中实体间的内在联系。

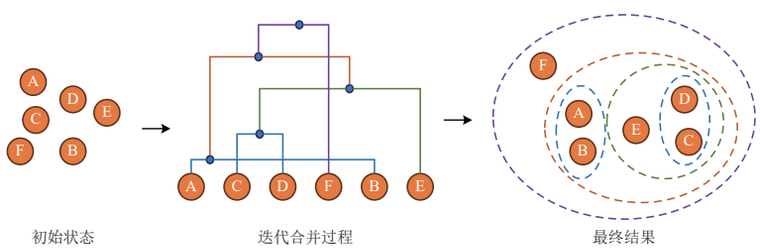


图 3　层次聚类方法示意图

2.2.2　基于DeepSeek-Prompt的关系抽取

层次聚类方法在处理分类关系推理时表现出色，但在非分类关系推理上效果不佳。为提升非分类关系的抽取准确性，并弥补层次聚类方法的应用局限，采用基于DeepSeek-Prompt的关系抽取方法作为补充，该方法通过结构化的“角色构建”方式，促进跨文本实体关系的拓展[18]。在关系抽取过程中，构建用于铁路路基压实关系抽取的Prompt模板，模板结构如图 4 所示。该模板主要由 3 个核心要素构成：角色定义、技能描述与约束条件。在

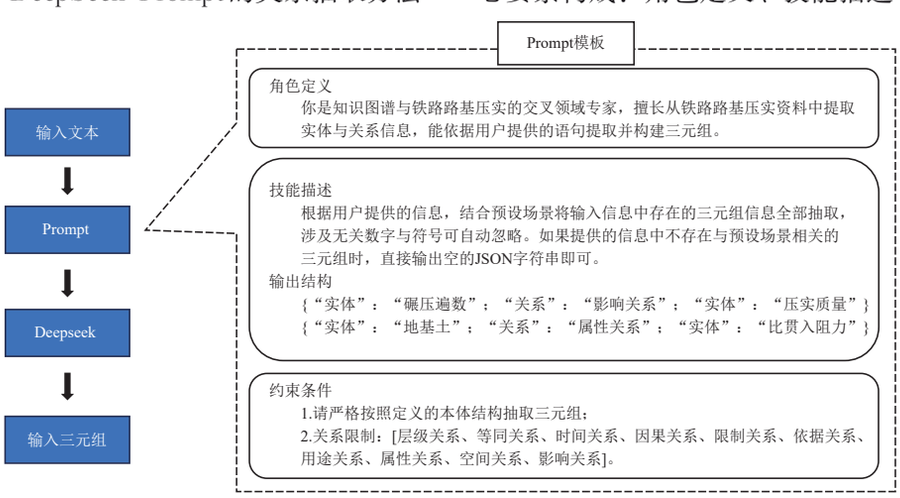


图 4　铁路路基压实关系抽取 Prompt 模板

角色定义中，将模型定位为知识图谱与铁路路基压实领域的交叉专家，擅长从专业语料中提取关键实体和关系。在技能描述中，模型根据用户提供的信息提取三元组信息，自动过滤无关数字和符号，聚焦语义关系抽取，结构化输出便于后续数据处理。约束条件明确限定抽取关系类型，仅涵盖层级、等同、时间等10种特定关系，使抽取工作更规范精准，契合知识图谱构建需求[19−20]。

该方法的应用示例如图5所示，将较长的文本内容输入模型，并给出诸如“根据已知文本内容，请推理并列举实体之间存在的关系”这样的开放式指令，以进行初始的关系识别。模型会以三元组形式进行输出，便于后续的归一化处理。为确保提取的可靠性，实施人工辅助审核机制，以便专家对模型生成的关系进行评估和筛选。若某些三元组未能依照铁路路基压实领域特定语义进行准确识别，则需通过人工审核进行手动筛选和修改。



图5 基于DeepSeek−Prompt的关系抽取应用示例

2.3 铁路路基压实知识图谱

采用上述方法，构建涵盖12部规范和标准、5本专著、296篇论文及多份工程资料的专业语料库。通过Bert−BiLSTM−CRF模型进行实体抽取，并融合层次聚类与DeepSeek−Prompt技术完成关系抽取与推理，经过知识处理与融合后共得到1 315个实体与1 003组关系，选取Neo4j作为知识图谱可视化的平台，使用Cypher语言将已经生成的三元组导入Neo4j的终端中。由于Neo4j的运行需求限制，随机选取300个实体及239组对应关系进行展示，形成的Neo4j可视化知识图谱如图6所示。

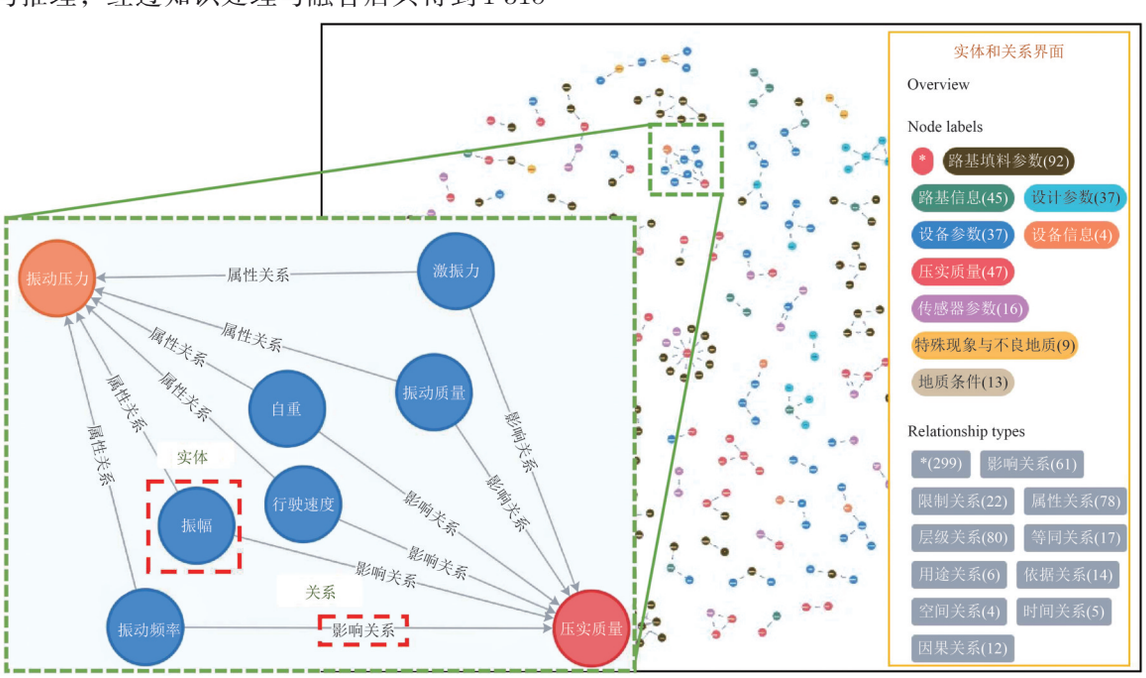


图6 Neo4j可视化知识图谱

3　基于知识图谱的检索增强生成技术框架

为充分发挥垂直领域知识图谱在智能化决策方面的效能，采用检索增强生成（RAG）技术将知识图谱嵌入 DeepSeek-R1 模型。该技术的核心在于通过动态联结外部知识库和生成模型，使模型输出结果更稳定可靠[21]。检索增强生成技术框架如图 7 所示，将构建好的知识图谱和相关文本资料通过向量嵌入模型转化为向量，储存至向量数据库。用户问题同样被向量嵌入模型转化为嵌入式向量，系统会在向量数据库中搜索与该问题向量语义上相似的实体以及与相似实体对应的知识文本，通过关系的链接将实体对应的知识文本整合并构建出 1 个提示模版，以大语言模型为载体输出回答。

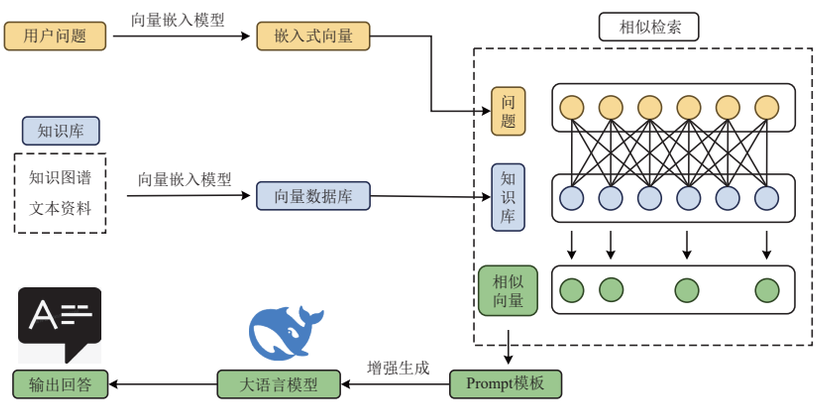


图 7　检索增强生成技术框架

4　指标与结果分析

4.1　Bert-BiLSTM-CRF 模型训练评价指标

Bert-BiLSTM-CRF 模型训练使用准确率 P、召回率 R 和综合评价指标 F_1 这 3 个指标评估模型性能。准确率为识别正确的实体占所有被识别出的实体的比例，召回率即为识别正确的实体数量占所有实体的比例，二者的计算式分别为

$$P=\frac{T_P}{T_P+F_P} \tag{1}$$

$$R=\frac{T_P}{T_P+F_N} \tag{2}$$

式中：T_P 为识别正确的实体数量；F_P 为识别出的非正确实体数量；F_N 为模型未正确识别的实体数量。

由于准确率与召回率往往存在一定的权衡关系，因此引入综合评价指标 F_1 对模型进行综合评价，其计算式为

$$F_1=2\frac{PR}{P+R} \tag{3}$$

使用人工标注的数据集对 Bert-BiLSTM-CRF 模型进行监督训练，得到 3 项评价指标 P、R 和 F_1 分数，其性能评估曲线如图 8 所示。由图 8 可知：各项评价指标整体随着迭代次数的增加而上升，但在迭代 20～30 轮次间，准确率出现异常下降，其原因主要由 2 方面引起：一方面由于模型可能过度拟合训练数据中的特征，导致对负例的错误判断增加；另一方面由于人工标注存在一定主观性和误差。若标注的样本特征存在错误或不规范情况，模型在学习过程中会受到干扰。在迭代到特定阶段时，这些标注误差带来的影响导致准确率暂时下降。最终模型在测试集上的准确率为 0.86，召回率为 0.81，F_1 分数为 0.83，整体性能表现良好，足以满足铁路路基领域自动标注的需求[22]。

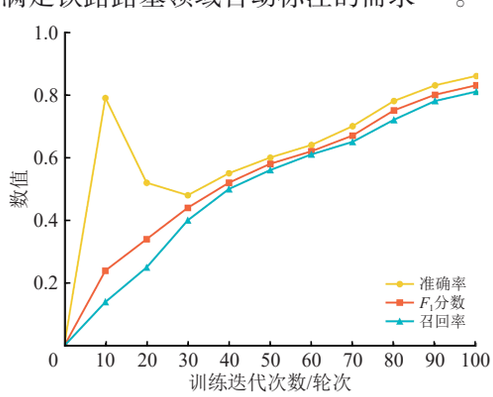


图 8　模型性能评估曲线

4.2　增强检索性能对比

为确保评估的科学性和可靠性，设计量化指标的对比试验对系统的问答能力进行评估。该试验聚

焦铁路路基压实施工构建编写了4类（基础概念类、施工工艺类、质量控制类、综合分析类）共48道典型问题。基础概念类聚焦路基压实的基础理论与核心定义，如“简述铁路路基压实度的含义及其对工程质量的影响”，旨在考察模型对专业术语与基础原理的掌握；施工工艺类围绕具体施工流程与技术要点，如“说明不同类型压路机在铁路路基压实施工中的适用场景与操作规范”，检验模型对施工实践知识的理解；质量控制类侧重质量问题的检测、评估与处理，如“当路基压实度检测结果不达标时，应如何排查原因并制定整改方案”，评估模型解决实际工程问题的能力；综合分析类则通过复杂工程情境，如“结合地质条件与工期要求，设计一套铁路路基压实施工的优化方案”，考查模型整合多维度知识、进行深度分析的水平。同时，为保证评估准确性，邀请专家依据行业标准与工程经验，针对每道问题撰写详细参考答案，构建测试评估体系，从而有效验证模型在铁路路基压实专业领域的问答能力。

选取传统DeepSeek-R1、基于传统检索增强的DeepSeek-R1、Claude3.5以及ChatGPT-4.0这4种模型与基于知识图谱检索增强的DeepSeek-R1模型进行对比。在评价指标选取方面，主要涵盖2个维度：一方面，将问答输出结果中涉及的领域实体数量作为衡量领域专业性的指标，统计每4 000个词元中所包含的领域实体数量，并取其平均值，记为$n_{4000\text{t}}$；另一方面，采用被广泛认可的基于召回率的摘要评估指标ROUGE以及BERT语义匹配指标BERTScore这2类指标体系[23]。其中，ROUGE-1基于单个词语的召回率进行计算，用于衡量生成文本与参考文本中单个词的重合程度，ROUGE-L基于最长公共子序列计算文本间的相似度，用于评估词汇表达准确性和文本连贯性；BERTScore利用预训练的BERT模型，从语义层面评估生成文本与参考文本的相似性，通过计算词嵌入的余弦相似度，得出准确率、召回率和F_1分数。不同模型量化评估对比见表4。

表4 不同模型量化评估对比

模型名称	ROUGE-1值	ROUGE-L值	BERTScore			$n_{4000\text{t}}$值
			准确率	召回率	F_1分数	
Claude 3.5模型	0.234 6	0.245 2	0.724 8	0.713 3	0.720 0	16
ChatGPT-4.0模型	0.246 8	0.305 5	0.742 1	0.712 9	0.720 6	14
传统DeepSeek-R1模型	0.258 6	0.309 4	0.744 3	0.731 6	0.737 8	15
基于传统检索增强的DeepSeek-R1模型	0.299 6	0.344 2	0.781 6	0.773 0	0.779 5	19
基于知识图谱检索增强的DeepSeek-R1模型	0.321 2	0.383 4	0.822 9	0.814 9	0.819 5	24

由表4可知：基于知识图谱检索增强的DeepSeek-R1模型在所有评估指标上均优于其他对比模型，在文本表达层面，该模型的ROUGE-1和ROUGE-L值分别达0.321 2和0.383 4，显著高于Claude 3.5模型的0.234 6和0.245 2、ChatGPT-4.0模型的0.246 8和0.305 5以及传统DeepSeek-R1模型的0.258 6和0.309 4，相较于传统检索增强的DeepSeek-R1模型（0.299 6和0.344 2）也有明显的提升，表明基于知识图谱的检索增强机制有效提高了模型在词汇表达准确性和文本连贯性方面的能力；在语义理解层面，该模型的BERTScore准确率达0.822 9，较传统通用模型提升约8个百分点，较传统检索增强的DeepSeek-R1模型提升约4个百分点，表明知识图谱的结构化信息有助于模型更准确地把握专业领域的语义内涵；在领域专业性评估方面，基于知识图谱检索增强的DeepSeek-R1模型的$n_{4000\text{t}}$值达到24，较其他模型平均提升约30%，体现了其在专业术语识别和应用方面的显著优势。

4.3 增强检索应用结果

模型性能指标表明，经过检索增强生成优化的模型性能明显优于传统通用模型。为研究知识图谱的检索增强机理，并直观展示模型的应用效果，对基于传统检索增强的DeepSeek-R1模型与基于知识图谱检索增强的DeepSeek-R1模型实际应用效果进行对比。

面向实际问题的模型应用效果如图9所示，向2种模型提出同一问题：“从路基压实施工角度，完成1个路堤的注意事项有哪些?”。结果显示，虽然二者均涵盖了“设备”“操作”“质量”和“环境”等关键事项，但在回答的准确性、完整性和效率上存在明显差异：①基于传统检索增强的模型提及了“地基处理”等内容，虽属于路基工程范畴，但偏离了“路基压实”这一核心主题，表现出一定

程度的AI幻觉，而基于知识图谱检索增强的模型则始终聚焦于路基压实施工相关内容，输出结果准确性更高；②基于知识图谱检索增强的模型不仅覆盖了基本事项，还额外纳入了“材料”“作业参数控制”及“水文地质”等关联因素，形成了更全面的知识体系；③基于传统检索增强模型响应时间为9～13 s，而基于知识图谱检索增强的模型响应时间为4～6 s，其生成效率提升约50%。

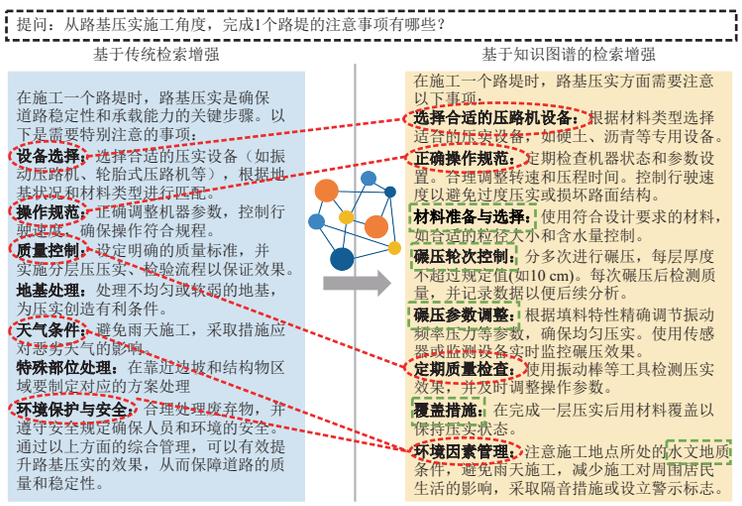


图9　面向实际问题的模型输出结果比较

由图9可知，传统的检索增强存在冗余信息过长，忽略知识库全局的上下文信息等问题，而基于知识图谱的检索增强能利用结构化知识图谱减少无关的冗余信息，降低AI幻觉，并提高大语言模型对外部知识的全局上下文理解能力。

5　结　语

本研究提出1种基于知识图谱的工程领域检索增强方法，以铁路路基压实为应用案例，对铁路路基压实领域知识进行“规范与标准”“理论与观点”以及“方案与反馈”三大维度的系统划分。基于此构建了包含12部规范、5本专业书籍、296篇学术论文与若干工程资料的专业语料库，最终形成包含1 315个实体与1 003组关系的铁路路基压实知识图谱。

在技术实现方面开发了融合人工智能与大语言模型的知识图谱构建技术路线，基于Bert-BiLSTM-CRF的实体抽取模型的精确率、召回率和F_1值分别达到0.86，0.81和0.78，可满足铁路路基压实领域自动标注的精度需求；通过融合层次聚类与DeepSeek-Prompt，实现了对文本中复杂关系的高效抽取以及跨文本关系的精准链接，有效解决了专业领域知识碎片化的问题。

将构建的知识图谱通过检索增强生成技术嵌入DeepSeek-R1模型，实现了模型在铁路路基压实领域的检索增强。模型性能评估表明，基于知识图谱检索增强模型的ROUGE-1，ROUGE-L，准确率和$n_{4\,000\,t}$值分别达到0.321 1，0.382 3，0.822 5和23，相较于传统模型与基于传统检索增强的DeepSeek-R1模型均获得显著提升，同时响应速度提高了近50%。

当前研究在铁路路基施工知识图谱构建及模型应用中，对施工工艺及流程图、设计图纸以及现场监理记录等多模态非结构化信息的处理仍不够细化，模型在专业术语的词汇表达准确性和文本连贯性方面也有待提高。未来研究拟引入跨模态模型深化多模态知识抽取，通过领域自适应训练与提示工程优化模型生成能力，并开发基于时间序列的关系推理模块以实现施工过程异常的动态知识建模。

参　考　文　献

[1] 吴天星，曹旭东，毕胜，等.基于大语言模型的重大慢病健康管理信息系统构建[J].计算机研究与发展，2025，62(7)：1653-1667.
(WU Tianxing，CAO Xudong，BI Sheng，et al. Constructing Health Management Information System for Major Chronic

Diseases Based on Large Language Model [J]. Journal of Computer Research and Development, 2025, 62 (7): 1653-1667. in Chinese)

[2] 马超,李武峰,陈羽飞,等.ChatGPT类大语言模型赋能电力标准数字化转型的核心技术、技术特征及应用展望[J].高电压技术,2025,51(4):1727-1746.
(MA Chao, LI Wufeng, CHEN Yufei, et al. Core Technologies, Technical Characteristics and Application Prospects of ChatGPT Class Big Language Models on the Digital Transformation of Standards in Electric Power Field [J]. High Voltage Engineering, 2025, 51 (4): 1727-1746. in Chinese)

[3] JI Z W, LEE N, FRIESKE R, et al. Survey of Hallucination in Natural Language Generation [J]. ACM Computing Surveys, 2023, 55 (12): 1-38.

[4] SINGHAL A. Introducing the Knowledge Graph: Things, not Strings [EB/OL]. USA: Official Google Blog, 2012 (2012-5-16) [2025-10-22]. https://blog.google/products/search/introducing-knowledge-graph-things-not/.

[5] 朱庆,王所智,丁雨淋,等.铁路隧道钻爆法施工智能管理的安全质量进度知识图谱构建方法[J].武汉大学学报(信息科学版),2022,47(8):1155-1164.
(ZHU Qing, WANG Suozhi, DING Yulin, et al. A Method of Safety-Quality-Schedule Knowledge Graph for Intelligent Management of Drilling and Blasting Construction of Railway Tunnels [J]. Geomatics and Information Science of Wuhan University, 2022, 47 (8): 1155-1164. in Chinese)

[6] WU J M, XIAO M Q, DAI L, et al. Intelligent Design Method of Mine Tunnel Portal Driven by Knowledge Graph [J]. Computers and Geotechnics, 2024, 173: 106431.

[7] LING J X, LI X J, LI H J, et al. Hybrid NLP-Based Extraction Method to Develop a Knowledge Graph for Rock Tunnel Support Design [J]. Advanced Engineering Informatics, 2024, 62 (Part B): 102725.

[8] 谢浩,姚洪锡,向子南,等.联合知识图谱与图神经网络的路基地基加固方案智能决策[J].铁道标准设计,2026,70(4):1-9.
(XIE Hao, YAO Hongxi, XIANG Zinan, et al. Intelligent Decision-Making for Subgrade Foundation Reinforcement Schemes by Combining Knowledge Graphs and Graph Neural Networks [J]. Railway Standard Design, 2026, 70 (4): 1-9. in Chinese)

[9] PROCKO T T, OCHOA O. Graph Retrieval-Augmented Generation for Large Language Models: a Survey [C]// 2024 Conference on AI, Science, Engineering, and Technology (AIxSET). New York: IEEE Press, 2024: 166-169.

[10] 王祥昊,杨怀志,王莉,等.基于语义理解与生成模型的铁路应急处置决策支持系统[J].铁道技术标准:中英文,2024,6(6):44-51.
(WANG Xianghao, YANG Huaizhi, WANG Li, et al. Railway Emergency Response Decision Support System Based on Semantic Understanding and Generation Models [J]. Railway Technical Standards: Chinese & English, 2024, 6 (6): 44-51. in Chinese)

[11] 李子林,王跃,刘庆猛,等.面向铁路知识问答的知识过滤检索增强生成方法[J].中国铁路,2024(12):121-129.
(LI Zilin, WANG Yue, LIU Qingmeng, et al. A Knowledge Filtering Retrieval-Augmented Generation Method for Railway Knowledge-Based Question Answering [J]. China Railway, 2024 (12): 121-129. in Chinese)

[12] 王同军,叶阳升,朱宏伟,等.铁路路基工程智能建造体系及关键技术[J].中国铁路,2023(1):1-9.
(WANG Tongjun, YE Yangsheng, ZHU Hongwei, et al. Intelligent Construction System and Key Technology of Railway Earthworks Project [J]. China Railway, 2023 (1): 1-9. in Chinese)

[13] 叶阳升,朱宏伟,尧俊凯,等.高速铁路路基振动压实理论与智能压实技术综述[J].中国铁道科学,2021,42(5):1-11.
(YE Yangsheng, ZHU Hongwei, YAO Junkai, et al. Review on Vibration Compaction Theory and Intelligent Compaction Technology of High-Speed Railway Subgrade [J]. China Railway Science, 2021, 42 (5): 1-11. in Chinese)

[14] 张念,周彩凤,万飞,等.基于BERT-BiLSTM-CRF的隧道施工安全领域命名实体识别[J].中国安全科学学报,2024,34(12):56-63.
(ZHANG Nian, ZHOU Caifeng, WAN Fei, et al. Tunnel Construction Safety Domain Named Entity Recognition Based on BERT-BiLSTM-CRF [J]. China Safety Science Journal, 2024, 34 (12): 56-63. in Chinese)

[15] 赵贵中,黄淼华.基于BERT-BiLSTM-CRF的电力事故信息抽取方法[J].综合智慧能源,2024,46(11):19-28.
(ZHAO Guizhong, HUANG Miaohua. An Information Extraction Method for Electric Power Accidents Based on BERT-

BiLSTM-CRF Model [J]. Integrated Intelligent Energy, 2024, 46 (11): 19-28. in Chinese)

[16] 姚松林,梁敦毫,罗振营.基于BERT-BiLSTM-CRF模型的荔枝命名实体识别[J].中国科技信息,2025(3):110-113.
(YAO Songlin, LIANG Dunhao, LUO Zhenying. Litchi Named Entity Recognition Based on BERT-BiLSTM-CRF Model [J]. China Science and Technology Information, 2025 (3): 110-113. in Chinese)

[17] EZUGWU A E, IKOTUN A M, OYELADE O O, et al. A Comprehensive Survey of Clustering Algorithms: State-of-the-Art Machine Learning Applications, Taxonomy, Challenges, and Future Research Prospects [J]. Engineering Applications of Artificial Intelligence, 2022, 110: 104743.

[18] 董磊,吴福居,史健勇,等.基于大语言模型的施工安全多模态知识图谱的构建与应用[J].计算机工程与应用,2025,61(9):325-333.
(DONG Lei, WU Fuju, SHI Jianyong, et al. Construction and Application of Multimodal Knowledge Graph in Construction Safety Field Based on Large Language Model [J]. Computer Engineering and Applications, 2025, 61 (9): 325-333. in Chinese)

[19] 刘启刚,王尹凡,牟立峰,等.大语言模型赋能"问题与方法"学术图谱构建与应用研究[J/OL].数据分析与知识发现,(2025-02-07)[2025-04-01].https://link.cnki.net/urlid/10.1478.G2.20250207.1425.010.
(LIU Qigang, WANG Yinfan, MU Lifeng, et al. Research on LLM Empowering Construction and Application of Problem and Method-Oriented Academic Graph [J/OL]. Data Analysis and Knowledge Discovery, (2025-02-07) [2025-04-01]. https://link.cnki.net/urlid/10.1478.G2.20250207.1425.010. in Chinese)

[20] 夏江镧,李艳玲,葛凤培.基于大语言模型的实体关系抽取综述[J].计算机科学与探索,2025,19(7):1681-1698.
(XIA Jianglan, LI Yanling, GE Fengpei. Survey of Entity Relation Extraction Based on Large Language Models [J]. Journal of Frontiers of Computer Science and Technology, 2025, 19 (7): 1681-1698. in Chinese)

[21] BAHR L, WEHNER C, WEWERKA J, et al. Knowledge Graph Enhanced Retrieval-Augmented Generation for Failure Mode and Effects Analysis [J], Journal of Industrial Information Integration, 2025, 45: 100807.

[22] VIKTORIN A, HRABEC D, NEVRLÝ V, et al. Hierarchical Clustering-Based Algorithms for Optimal Waste Collection Point Locations in Large-Scale Problems: a Framework Development and Case Study [J]. Computers & Industrial Engineering, 2023, 178: 109142.

[23] 刘耀文,夏一雪,张鹏,等.国家安全情报战略知识图谱构建与检索增强问答框架研究[J].情报杂志,2025,44(7):165-173.
(LIU Yaowen, XIA Yixue, ZHANG Peng, et al. Research on Knowledge Graph Construction and Retrieval-Enhanced Question Answering Framework for National Security Intelligence Strategy [J]. Journal of Intelligence, 2025, 44 (7): 165-173. in Chinese)

Knowledge Graph Establishment and Retrieval-Augmented Method for Railway Subgrade Compaction Construction

SHI Xiaomeng[1,2], XU Gongbo[1,2], CAI Guoqing[1,2], WANG Zinan[1,2], DENG Zhiyun[1,2], LIU Baoguo[1,2]

(1. Key Laboratory for Urban Underground Engineering of the Ministry of Education, Beijing Jiaotong University, Beijing 100044, China;
2. School of Civil Engineering, Beijing Jiaotong University, Beijing 100044, China)

Abstract: To address the issues such as delays of knowledge timeliness, insufficient understanding of professional knowledge, and lack of output stability in the application of general models to railway subgrade compaction construction, a knowledge graph establishment and retrieval-augmented method specifically for the railway subgrade compaction construction domain was proposed. First, a specialized corpus encompassing standards and specifications, professional books, academic papers, and engineering documents was built. Then, a

Bert-BiLSTM-CRF model was designed for entity extraction from this corpus. By integrating hierarchical clustering and DeepSeek-Prompt techniques, relation extraction and reasoning were accomplished, resulting in a knowledge graph comprising 1 315 entities and 1 003 relations within the railway subgrade compaction domain. Finally, this knowledge graph was integrated into the DeepSeek-R1 model through Retrieval-Augmented Generation (RAG) technology to enable retrieval-augmented in the field of railway subgrade compaction. The results show that the model integrated with the knowledge graph significantly outperforms traditional general models in key performance metrics, achieving ROUGE-1, ROUGE-L, and accuracy rate of 0.321 2, 0.383 4, and 0.822 9, respectively. This approach not only successfully achieves systematic extraction and structured representation of knowledge, but also effectively improves the model's precision in understanding complex professional knowledge and the reliability of its generation. It opens up a new technical approach for enhancing the level of intelligence in engineering practice.

Key words: Railway subgrade; Knowledge graph; Retrieval-augmented; Subgrade construction; DeepSeek

（责任编辑　杨婧婕）