

大模型融入城市规划专业工作的技术范式与模型构建*

Technical Paradigm and Model Construction for Embedding Large Models into Professional Urban Planning Practice

钮心毅 刘思涵 林诗佳 王晨亦

NIU Xinyi, LIU Sihan, LIN Shijia, WANG Chenyi

关键词 专业大模型; 城市规划; 领域知识; 诊断—推理—对策; 诊断训练; 对策训练

Keywords: professional large models; urban planning; domain knowledge; diagnosis-inference-treatment; diagnostic training; treatment training

提 要 从认识城市规划专业工作内涵和人工智能类型出发, 聚焦大模型融入城市规划专业工作的技术范式, 提出构建规划专业大模型的途径。城市规划专业工作采用“诊断—推理—对策”的工作模式, 运用领域知识提供解决方案。从感知智能与认知智能的差异出发, 论证大模型能支持“诊断—推理—对策”的工作模式。提出诊断训练、对策训练两个阶段的领域知识学习构建规划专业大模型的技术框架。以实证案例证明规划专业大模型能够具备“发现空间问题—推导问题原因—提出规划策略”的规划专业能力。研究表明, 大模型融入城市规划专业工作的技术范式是以具备“通用知识+领域知识”的专业大模型, 以“诊断—推理—对策”工作模式支持问题导向和规律导向的规划范式; 规划专业大模型是这一技术范式的载体, 其构建方式是以领域知识为核心, 在通用大模型基础上进行诊断训练、对策训练实现。具备领域知识的专业大模型能够有效融入规划专业工作, 起到增强规划师专业工作能力的作

Abstract: The paper starts by examining the nature of professional work in urban planning and the typology of artificial intelligence. It focuses on the technical paradigm through which large models can be embedded into professional urban planning practice and proposes an approach for constructing professional planning large models. The paper characterizes urban planning professional work as a "diagnosis—inference—treatment" professional working mode, in which domain knowledge is used to provide solutions. By distinguishing between perceptual intelligence and cognitive intelligence, the study demonstrates that large models are capable of supporting this professional working mode. Building on this understanding, a technical framework for constructing professional planning large models is proposed, emphasizing domain knowledge acquisition through two stages: diagnostic training and treatment training. An empirical case is presented to demonstrate that professional planning large models can acquire professional planning capabilities, including the ability to identify spatial problems, infer underlying causes, and propose planning strategies. The findings indicate that the technical paradigm for embedding large models into urban planning practice consists of professional large models equipped with both general knowledge and domain-specific knowledge, operating within the "diagnosis—inference—treatment" working mode to support problem-oriented and law-oriented planning paradigms. Professional planning large models serve as the carrier of this paradigm and are developed through domain-knowledge-centered learning based on general large models, implemented via diagnostic and treatment training. Such domain-specific professional large models can be effectively embedded into planning practice and function to enhance the professional capacity of planners.

中图分类号 TU984 文献标志码 A
DOI 10.16361/j.upf.202601003
文章编号 1000-3363(2026)01-0021-08

作者简介

钮心毅, 同济大学建筑与城市规划学院, 自然资源部国土空间智能规划技术重点实验室, 教授、博士生导师, niuxinyi@tongji.edu.cn

刘思涵, 同济大学建筑与城市规划学院, 自然资源部国土空间智能规划技术重点实验室, 博士研究生

林诗佳, 同济大学建筑与城市规划学院, 自然资源部国土空间智能规划技术重点实验室, 博士研究生

王晨亦, 同济大学建筑与城市规划学院, 自然资源部国土空间智能规划技术重点实验室, 硕士研究生

* 国家自然科学基金面上项目:“虚拟空间”和“流空间”双重作用对实体商业中心体系影响及机制研究(项目编号:52278072)资助

人工智能作为引领全球新一轮科技革命和产业变革的战略性技术，正深刻改变人类生产生活方式。大模型开启了人工智能发展的新阶段，不仅在迅速切入经济产业、社会治理、民生服务等行业领域，也在快速渗入自然科学、社会科学的科研领域。2025年以来，大模型进入了城市规划研究。大模型赋能城市规划研究与实践迅速成为规划学界和业界共同聚焦的议题^[1]。

综合近期出现的大模型支持城市规划研究，以大模型技术用途类型可划分为感知类与推理类两个类型。感知类用途是将大模型用于城市空间的形态特征、功能属性及其背后社会经济现象的感知，包括感知街道和公园空间特征主题^[2]、街道空间品质^[3]、城市空间安全性^[4]，进而感知贫困等社会经济现象^[5]。感知类用途研究均依托多模态大模型的视觉—语言理解能力，实现从图像到语义的转换。大模型能够实现与普通公众一致的空间感知^[6]，能用于城市空间评价等用途。推理类用途则是利用大模型对多模态信息的语义解析，实现对城市空间相关特征、需求或反馈的推断与预测。包括从街景图片推断城市功能^[7]、城市设计及街道更新等意象生成^[8-9]、规划政策文件解读及公众反馈意见分析^[10]、预测公共事件下人员流动特征^[11]等。推理类用途研究依托大模型的推理能力实现，有效支持了分析、预测等用途。

从大模型使用方式看，可划分为两种方式。第一种是将通用大模型作为规划研究中普通公众，起到模拟公众感知与参与等作用。包括公众对城市空间与景观偏好分析^[6,12]、模拟规划过程中普通公众参与^[13]等。第二种是在通用大模型基础上构建规划专业大模型，让大模型发挥规划专业人员的作用。专业大模型能够应用于规划设计理念的意象场景生成^[8-9]；完成土地利用分类、社会经济状况的空间分布分析等推理^[14]；根据多源数据预测交通出行时空需求^[11,15]。

一方面，既有研究在技术用途上探索了大模型支持规划研究的感知类用途、推理类用途。如何认识两类技术用途之间的关联和差异，明确人工智能技术用途与规划专业工作的适配关系，是推动大模型赋能城市规划专业工作的一个亟

待解决议题。另一方面，既有研究在使用方式上表明了大模型既能够承担公众参与等普通公众承担的任务，也能承担诸如交通预测等专业性较强的任务。如何认识这两类使用方式的差异，理清各自适用范畴，是推进大模型赋能城市规划专业工作的另一个亟待解决议题。本文围绕上述两大议题，从规划专业知识结构与空间规划范式、大模型所代表的人工智能类型阶段两个维度展开解析，提出大模型融入城市规划专业工作的技术范式，提出适用于城市规划专业工作的专业大模型构建方法，以期为人工智能赋能城市规划提供技术路径。

1 规划专业工作的内涵解析

1.1 知识结构：“通用知识+领域知识”

在认知科学、教育心理学等领域，知识通常被区分为通用知识（general knowledge, GK）与领域知识（domain knowledge, DK）^[16]。通用知识指跨学科、广泛适用的事实性与常识性知识，包括语言、历史、数学、物理、地理等。领域知识则指特定学科或行业内部高度结构化、深度组织的专业知识，例如城市规划中的城市规划原理、城市空间结构等专业理论。领域知识是在通用知识的基础上，通过在特定领域的持续学习与实践逐步形成。

从知识角度看待专业工作，其本质在于利用领域知识解决专业问题并创造专业价值。认知科学文献表明，“专家”与“新人”之间的差异主要体现在领域知识的掌握程度，而不是通用知识的差异^[17]。这一规律同样适用于人工智能，通用大模型在预训练阶段掌握了通用知识，但未能系统性涉及领域知识。有研究^[18]系统测试了通用大模型在556个城市规划专业任务中的表现，其在信息检索、规划文本辅助等由通用知识主导的任务上性能良好；但在规划决策、空间设计等高度依赖领域知识的任务上表现明显欠佳。

知识结构的差异可以解释大模型应用于规划专业工作的两种使用方式。第一种将通用大模型作为规划研究中普通公众使用，是在利用通用大模型的通用知识。这包含了模拟公众对城市空间感

知与偏好^[4-6]，公众意见的表达与参与等^[13,19-20]。第二种让大模型发挥规划专业人员的作用，是在利用大模型的“通用知识+领域知识”。这包含了规划意向生成^[8-9,21]、土地利用与社会经济特征推断^[7,14]、交通出行需求预测等^[11,15]。从知识结构来看，两类使用方式的根本差异在于调用通用知识或融入领域知识。

1.2 空间规划范式：“诊断—推理—对策”工作模式

问题导向的规划范式是现代规划思想的重要体现。随着工业化快速推进和城市化规模急剧扩张，城市病等空间问题大量涌现，城市规划需要从具体问题出发构建空间解决方案。Geddes^[22]提出了“survey before plan”的主张，并扩展类比为“diagnosis before treatment”，强调通过调查诊断城市运行状况、支撑规划决策。Mumford^[23]进一步强调基于第一手调查诊断“大都市病”，并提出对应规划策略。后人将其范式总结为“调查—评估—编制规划方案—实施”四个阶段。上述规划思想均强调以问题为导向，根据具体空间问题提出规划策略。在此基础上，吴志强等^[24]提出了规律导向的规划范式，强调依据城市发展规律开展空间诊断、推演与场景营造。在问题导向和规律导向的规划范式下，规划专业工作均遵循“发现空间问题—推导问题原因—提出规划策略”的逻辑，均可归纳为“诊断—推理—对策”的工作模式。

“诊断—推理—对策”的工作模式由空间诊断和空间对策两个阶段组成。空间诊断吸收了医学领域临床诊断的思想，是利用现有的理论知识和规律总结对城市问题进行认知和剖析^[25-26]。空间诊断对应“发现空间问题—推导问题原因”的过程，不仅强调诊断的结果，更强调推理具体的成因。空间对策对应“推导问题原因—提出规划策略”的过程，同样强调给出规划策略之前的推理过程。同样的空间问题在不同情境下可能具有不同成因，其对策也需要因地制宜。因此“推理”在两阶段中均具有关键作用。

综上，城市规划专业工作内涵可以概括为，在“通用知识+领域知识”的知识结构支撑下，遵循“诊断—推理—对

策”的工作模式，为城市空间问题提供解决方案。这一城市规划专业工作内涵是大模型融入城市规划专业工作技术范式的基础。

2 感知智能、认知智能与城市规划专业工作

2.1 从感知智能到认知智能

人工智能的发展通常被划分为计算智能、感知智能、认知智能与自主智能等阶段^[27]，当前最大量人工智能技术均属于感知智能（perceptual intelligence）。感知智能指机器获取并识别语音、图像、视频等感知数据的能力，例如使用深度学习实现语义分割、目标检测等任务，其典型能力是分类、检测与预测^[28]。以医学影像分析为例，基于深度学习模型能够完成病灶检测等任务，其输出是“客观观察结果”^[29]。但临床诊断仍必须综合病史、体格检查等信息，由医生在医学知识体系内进行因果推理和风险判断。感知智能不能完成后者任务。

大模型开启了人工智能向认知智能（cognitive intelligence）发展的新阶段。认知智能强调机器在感知信息之上形成理解、推理、判断和解释的能力，从识别走向理解。大模型能够展现“思维链”等多步推理能力，在一定程度上展现出“类认知能力”^[30]。虽然大模型也能具备“客观观察结果”的感知能力，但更具备从“看见”走向“主观分析推理”的技术用途能力。同样以医学领域为例，经过专业训练的医学专业大模型能够实现综合分析患者症状与病史，分析推理得到诊断结果乃至治疗方案^[31]。与上述病灶检测等感知智能相比，大模型认知能力可以将感知到的病灶与专业知识结构建立联系，能够实现医学诊断和治疗等“主观分析推理”过程。

2.2 城市规划中的感知智能和认知智能

感知智能已经广泛用于城市规划中城市空间要素识别、空间形态量化等任务。过去十余年中，以语义分割、目标检测、图像分类为代表的深度学习方法识别街道绿视率、天空暴露度、建筑形态指数等要素，为量化理解城市形态和空间品质提供了重要支撑^[32-33]。然而，

这些方法仅能测度城市空间要素特征，仍无法推理城市空间要素特征的形成原因，也无法依据原因生成优化策略，无法完整支持“诊断—推理—对策”的工作模式。

对照当前城市规划研究中的大模型两类技术用途，其核心差异在于对大模型感知智能与认知智能的能力侧重不同。感知类用途属于感知智能的应用范畴。感知类用途实质上是依托大模型的多模态感知能力，拓宽传统深度学习的识别边界与应用场景，这包括了对街景图像进行自然语言描述与空间品质评价^[2,34]，对城市环境的变化进行比对^[35]，以及对社会经济现象进行主观感知^[4-5]。此外，大模型的文本感知能力也被用于政策分析、公众意见解析和地理语义建模^[12,19,36]。感知类用途并未突破传统深度学习的核心范畴，仍归属于感知智能的技术体系。大模型的推理类用途属于认知智能的应用范畴，即利用大模型在感知要素及特征的基础上实现空间诊断、因果推理与策略生成等规划专业工作的任务。在融入规划领域知识的条件下，大模型可实现对城市、建筑功能的推断^[7,14]，生成规划设计意向图^[8,9,21]，对规划政策文件的专业解读^[10]，实现特殊公共事件下交通需求预测等^[11]。尽管上述工作仅完成了“诊断—推理—对策”全过程中的部分环节，但已经体现出大模型在规划专业工作中具有空间诊断与空间对策的潜力。

相比感知智能阶段的人工智能技术，大模型关键潜力在于其属于认知智能，能够从感知的现象推理形成原因、依据原因生成优化策略。这使得大模型能够具备完整支撑“诊断—推理—对策”工作模式的能力，为大模型融入城市规划专业工作的技术范式提供了技术可行性。

3 技术范式与模型框架

3.1 大模型融入城市规划专业工作的技术范式

大模型融入城市规划专业工作的技术范式是以具备“通用知识+领域知识”知识结构的专业大模型支撑“诊断—推理—对策”全过程工作模式，支持问题导向、规律导向的规划范式（图1）。面

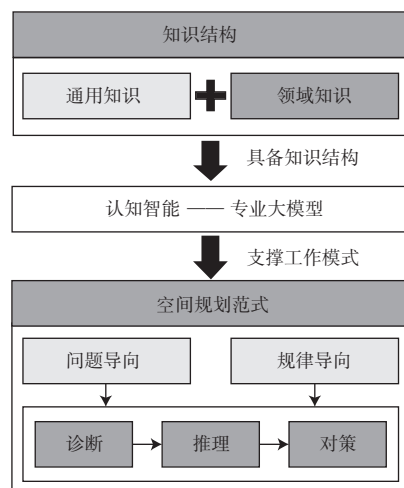


图1 大模型融入城市规划专业工作的技术范式

Fig.1 Technical paradigm for embedding large models into urban planning practice

向规划专业工作的规划专业大模型是上述技术范式的技术载体。

上述技术范式的可行性源于空间规划范式要求与大模型技术优势的契合。问题导向的规划范式核心是“通过数据分析诊断城市空间问题（诊断），运用规划领域知识推断原因（推理），提出空间规划对应策略（对策）”。大模型是认知智能的代表性进展，具有多步推理能力。在通用大模型基础上引入规划领域知识，构建规划专业大模型，即可在技术上支撑“诊断—推理—对策”的工作模式。

上述技术范式在存量规划时代更具必要性。面对城市存量空间更新提质的现实需求，传统规划技术面临着大规模和精细化的挑战。大规模体现在大规模存量城市空间更新需求，精细化体现在需要精准识别空间问题、提出差异化空间策略。运用规划专业大模型，依托领域知识实现大规模的存量空间诊断、推理形成差异化规划策略，其规模化处理与稳定性的优势就能弥补传统规划技术的局限。

3.2 规划专业大模型构建

落实上述技术范式的关键在于构建规划专业大模型。构建支撑“诊断—推理—对策”工作模式的规划专业大模型技术框架如图2所示。

规划专业大模型构建以“发现空间问题—推导问题原因—提出规划策略”

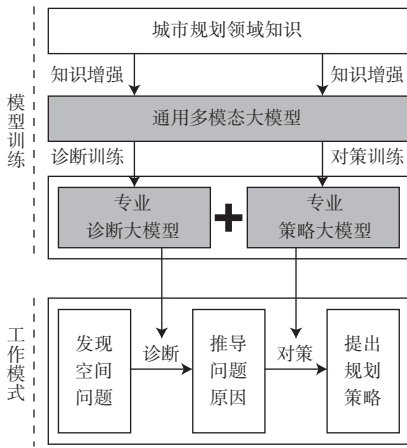


图2 构建规划专业大模型的技术框架
Fig.2 Technical framework for constructing professional planning large models

为核心，将规划领域知识划分为诊断领域知识、对策领域知识，分别融入通用大模型。诊断领域知识支撑模型完成“发现空间问题—推导问题原因”的专业判断，侧重于空间状态识别、问题诊断与原因分析；对策领域知识则支撑模型在诊断基础上完成“推导问题原因—提出规划策略”的专业推理，侧重于规划干预逻辑、策略选择与空间响应。

上述两类领域知识分别通过诊断训练与对策训练两个阶段实现，目标是让模型掌握空间诊断阶段、空间对策阶段所需的领域知识。诊断训练、对策训练均使用微调、提示词工程等知识增强方式实现。两类领域知识的学习解决了通用大模型在领域知识不足的短板，从而支撑“诊断—推理—对策”全过程工作模式。

3.3 规划专业大模型应用定位

规划专业大模型适宜的应用定位是规划专业工作助手，是将专业模型诊断、对策输出作为决策的参考。规划专业人员仍发挥主导作用，包括评估模型诊断结果、对方案对策进行质量评估，给出最后决定。规划专业工作中价值判断、公众利益权衡等核心人文属性未必均能包含在领域知识训练中。大模型具有高效率处理大规模海量数据的技术优势，能胜任大规模的初步诊断、提出供决策的参考方案，提供“循证”参考。规划专业工作性质和大模型技术特点共同决

定了规划专业大模型的专业工作助手定位。

4 实证案例

4.1 研究案例

在“先体检、后更新，无体检、不更新”的导向下，超大、特大城市普遍面临大规模、动态化的城市体检与更新迫切需求。在上海内环线以内120 km²范围有1326条街道，以道路交叉口为界有5300个街段，均面临着街道空间更新提质的需求。将大模型融入存量空间更新中专业工作，能为街道空间更新大规模诊断空间问题、差异化提出规划策略提供支持。为此，以街道空间更新中提升街道步行安全性为案例，构建专业大模型支持街道空间更新中“发现空间问题—推导问题原因—提出规划策略”过程，以验证本文提出的技术范式和专业大模型构建方式。

4.2 模型训练

4.2.1 街道步行安全性的领域知识结构

提升街道步行安全性^①属于典型的城市规划专业性任务。为构建街道步行安全性专业大模型，首先梳理和区分诊断领域知识、对策领域知识（图3），以支撑诊断训练、对策训练两阶段的模型训练。

诊断领域知识是基于Alfonso^[37]的步行需求层级理论，该理论中包含了安全性层级，可将导致街道安全性差的问题原因归纳为自然监视不足（insufficient

natural surveillance）与环境失序（environmental disorder）两类。前者表现为街道步行环境缺乏被观察与干预的可能性，放大了行人对隐蔽性犯罪的风险预期，进而抑制步行意愿；后者表现为街道步行环境出现破损、脏乱、废弃等物质衰败现象，使得行人因这些环境特征高估安全风险，从而回避使用相关空间。这两类问题原因用以支撑“发现空间问题—推导问题原因”的诊断推理，对接空间诊断阶段的应用。

对策领域知识是基于通过环境设计预防犯罪（crime prevention through environmental design, CPTED）理论^[38]，针对上述两类问题原因提出了多种街道步行环境优化策略。以自然监视不足为例，根据问题原因不同，对策可概括为增强街道可见性与提升沿街功能支持，分别从“看得见”和“有人看”两个方面支撑“街道眼”的空间实现。上述对策领域知识为模型在诊断结果基础上开展“推导问题原因—提出规划策略”的对策推理提供专业依据，对接空间对策阶段的应用。

4.2.2 街道步行安全性的领域知识训练

以通用多模态大模型 Qwen3-VL-32B 为基础，通过领域知识训练构建街道步行安全性专业大模型。依据街道步行安全性的诊断领域知识、对策领域知识本身差异，在诊断训练与对策训练中分别采取了两种不同的知识增强路径实现领域知识的学习。

诊断训练是使专业大模型掌握步行需求层级理论中的街道步行安全性知识，

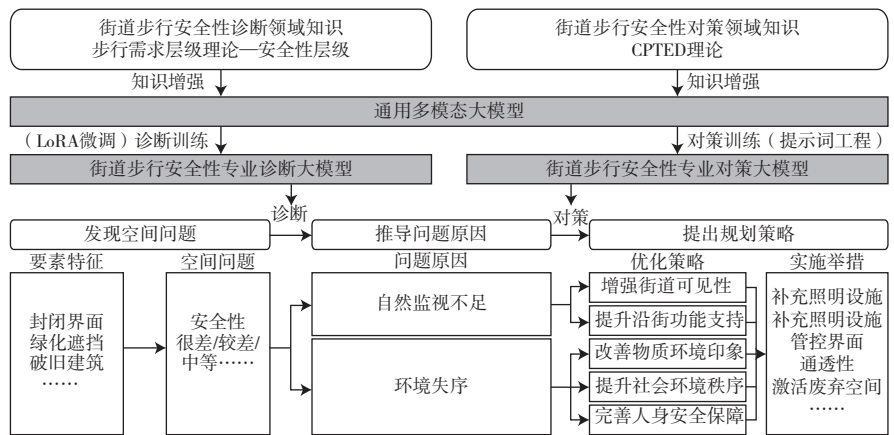


图3 街道步行安全性诊断—对策的训练框架
Fig.3 Training framework for diagnosis-treatment reasoning in street walking safety

通过低秩适应 (LoRA) 微调实现^[39]。步行需求层级理论虽然有清晰的概念体系和归因逻辑, 但现实街道空间感知涉及诸多空间要素的交织作用, 属于隐性知识, 难以通过明确规则表达^[8]。为此, 采用LoRA技术对通用大模型进行微调, 适用于隐性知识的训练学习。训练样本由“街景图像—诊断文本”构成, 其中诊断文本在对街道步行安全性进行评级的基础上, 依据步行需求层级理论明确指出主要问题原因, 并用自然语言描述其对应的具体空间表现特征。

对策训练是让专业大模型掌握CPTED理论中街道步行安全性的对策领域知识, 通过提示词工程 (prompt engineering) 实现。CPTED理论中街道步行安全性对策知识具有更为清晰的“问题原因—规划策略”知识体系, 表现为同一问题原因可直接对应若干优化策略, 具体举措也可以依据街道空间实际情况进行选择。因此, 适合在推理阶段显式引入结构化对策知识, 实现按需调用与动态适配。因此, 对策训练提示词由“任务指令+诊断文本+对策领域知识”三个部分构成 (图4): 其一, 通过“任务指令”明确模型的专业推理目标; 其二, 输入上一步的街道步行安全性诊断文本, 其中“问题原因”为大模型提供全面、具体的空间背景信息; 其三, 输入结构化的对策领域知识, 对不同问题原因、优化策略及实施举措之间的逻辑关系进行系统阐述。

通过诊断训练与对策训练两个阶段, 街道步行安全性专业大模型完成了步行需求层级理论、CPTED理论的学习。街道步行安全性专业大模型实现了对输入的每一幅街景图像, 以自然语言方式给出街道安全性诊断结果, 推断问题原因, 并进一步提出提升街道步行安全性的优化策略和实施举措。

4.3 模型应用

4.3.1 典型街道步行安全性诊断与对策

为验证规划专业大模型构建的框架可行性与有效性, 选取典型街道断面案例, 对比分析经两阶段训练的专业大模型与未经训练的通用大模型在街道步行安全性诊断生成中的表现差异 (表1)。结果表明, 经过训练的专业大模型能够

严格遵循步行需求层级理论中街道安全性知识进行诊断推理。专业大模型明确指出导致该街道步行安全性较差的核心原因在于自然监视不足, 并结合空间界面形态与沿街活动特征进行了清晰、可解释的专业归因。在此基础上, 专业大模型依据CPTED理论, 在空间优化策略层面提出提升沿街功能支持与增强街道可见性的干预方向、具有实施指向性的建议举措。与之相对, 未经训练的通用

大模型虽能准确感知街道步行环境空间要素, 但其诊断结果停留在共享单车乱停、界面老旧、照明不足等表象性问题描述, 未能基于步行需求层级理论中街道安全性知识对问题进行专业推理。其对策输出亦属于城市运行管理的建议, 非空间优化策略。这一实证结果从应用层面验证了大模型融入城市规划专业工作的技术范式有效性。

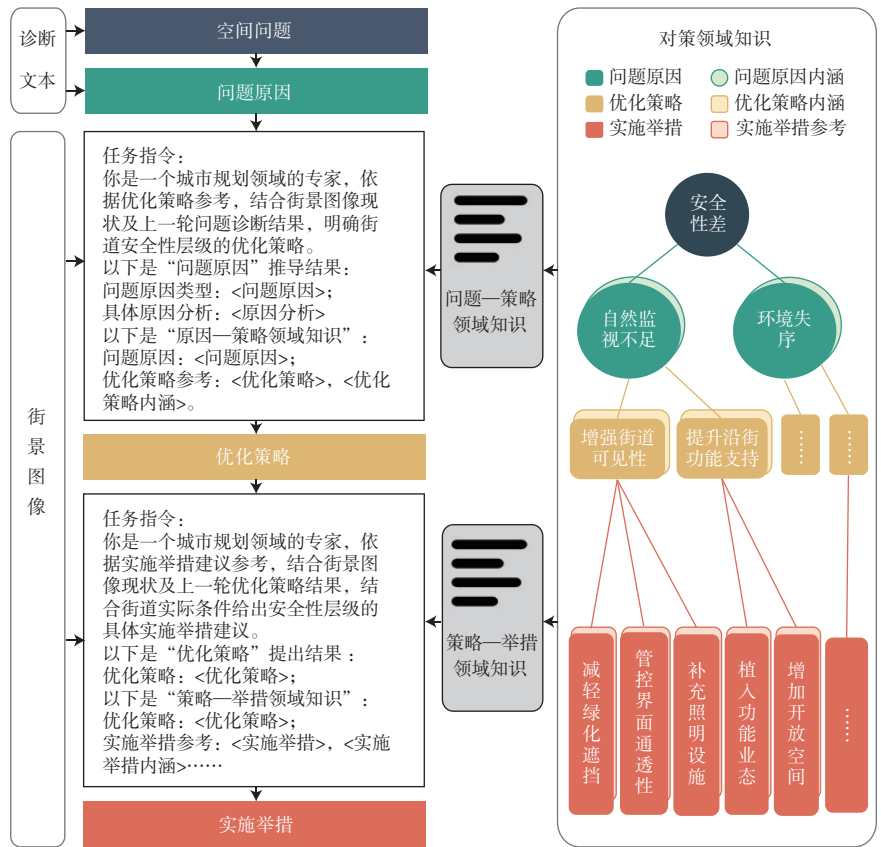


图4 基于“街景图像+诊断文本+对策领域知识”的对策训练提示词设计

Fig.4 Prompt design for treatment training based on "SVI + diagnostic text + treatment domain knowledge"

表1 专业大模型的典型街道诊断输出结果

Tab.1 Outputs of diagnosis reasoning for a typical street segment generated by the professional large models

诊断	专业大模型的输出结果	通用大模型的输出结果
发现空间问题	该街道步行安全性较差	
推导问题原因	问题原因为自然监视不足。具体表现为街道右侧存在封闭界面, 造成视线通透性差, 缺乏有效的自然监视; 同时, 未见路灯、监控、商铺等安全支持要素, 也无明显日常活动, 导致环境缺乏活力与心理安全感支撑。上述因素共同削弱了步行者对街道安全性的感知, 构成明显的自然监视不足问题。环境失序方面, 未见破败建筑、杂乱街道、垃圾或施工行为等要素, 故不构成环境失序问题	该街道人行道被大量共享单车侵占, 迫使行人进入车道, 增加碰撞风险; 步行空间连续性中断, 影响通行安全与应急疏散; 右侧建筑立面老旧破损, 环境杂乱, 引发心理不安与犯罪感知; 两侧树木过密, 夜间或阴雨天易造成视线遮蔽, 形成安全隐患。这些问题相互叠加, 削弱了步行者的物理安全与心理安全感, 故该街道步行安全性评价为“较差”

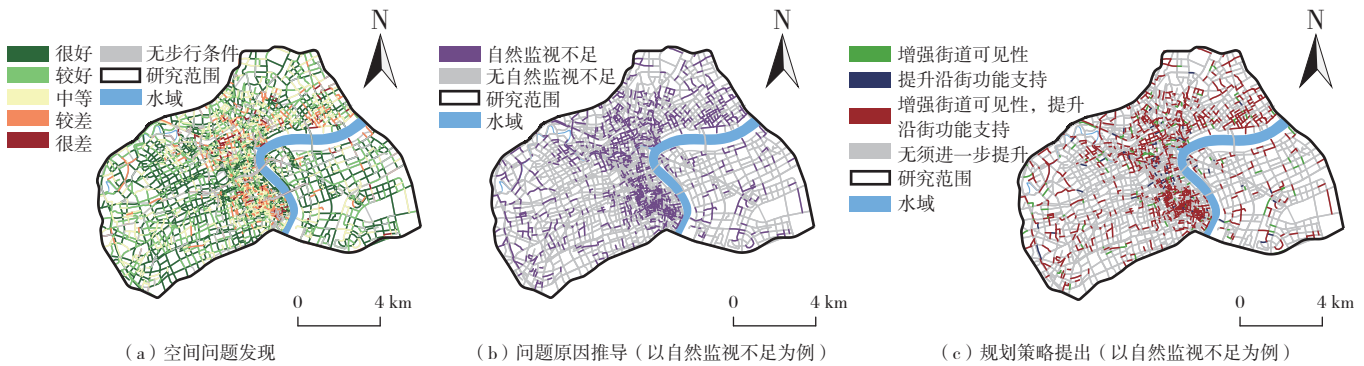


图5 上海内环内城区的街道步行安全性诊断和策略

Fig.5 Diagnosis and treatment for walking safety across streets within Shanghai's Inner Ring

4.3.2 全部街道步行安全性诊断与对策

专业大模型能够在1 h内完成上海内环以内全部 5300 个街段的空间诊断与策略推演^②。结果显示,上海内环以内街道步行安全性总体较好,65.47%街道步行安全性为很好或较好[图5(a)]。在安全性判断为中等及以下的34.53%街段中,专业大模型进一步诊断出存在自然监视不足[图5(b)]、环境失序两类主要原因。针对自然监视不足带来的负面影响,专业大模型提出了差异化的空间优化策略[图5(c)],并对每个街段进一步给出了推荐的实施举措。以上案例表明,专业大模型具备了大规模快速诊断、差异化提出规划策略的能力,具备了承担存量城市空间中大规模、常态化的更新体检专业工作的能力。

4.4 模型测评

为验证所构建的街道可步行性专业大模型在认知推理能力上的可靠性,邀请了13位具有完整城市规划专业教育背景的专业人员参与模型测评。基于52张随机抽取的街景图像构成了测试集,使用准确率(accuracy)和 F_1 分数(F_1 -Score)两个指标^③,对专业大模型、通用大模型与专业人员在空间诊断与空间对策两个环节的表现进行对比分析(图6)。

测评结果显示,在推导问题原因方面,专业大模型准确率得分为0.88,仅与专业人员平均水平(0.94)存在 ± 2 标准差区间内的差异,表明其能力处于专业人员群体的正常波动范围之内。相比之下,通用大模型准确率仅为0.61,显著低于专业人员均值,难以胜任专业推

理任务。进一步分析专业大模型与专业人员共识不一致的案例发现:专业大模型未出现诊断问题“漏报”现象;差异仅源于存在极少量的“误报”。这意味着专业大模型尤其适用于城市更新工作的前期预警环节——通过大规模、自动化地诊断潜在问题空间,辅助专业人员快速锁定需深入研判的对象,实现“先预警、后确诊”的高效工作链条。

在提出规划策略方面,专业大模型与专业人员平均水平基本相当,在准确率和 F_1 分数两个指标上,二者均只有0.1的轻微差异。值得注意的是,相较于诊断推理阶段,专业人员在策略制定阶段的个体差异显著增大,反映出在空间优化方案选择上存在更强的主观偏好与经

验依赖。相比之下,专业大模型展现出更高的稳定性,能够持续输出接近专业人员平均水平的策略建议,为专业规划师提供可参考策略建议。

专业大模型专项测评结果证实,本文提出的技术范式具备了可靠承担规划专业性工作的能力。本案例构建的专业大模型适用于存量更新背景下大规模且稳定实现“诊断—推理—对策”的规划专业工作流程,其能力满足了规划专业助手的应用定位。

5 结论和讨论

5.1 结论

从规划专业工作内涵、人工智能类

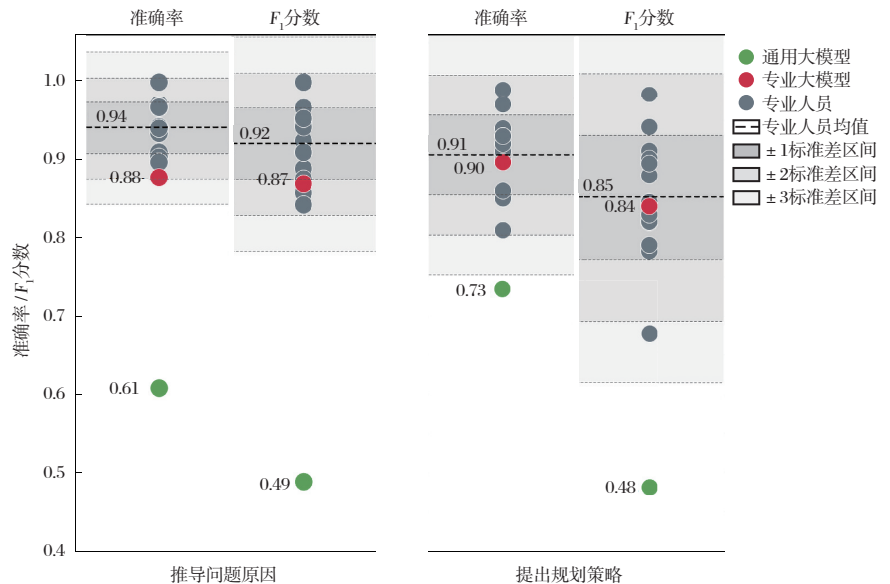


图6 “诊断—推理—对策”专业大模型与专业人员的比较

Fig.6 Comparison between the "diagnosis—inference—treatment" professional large model and human professionals

型两个维度展开,解析了大模型在城市规划研究和实践中的技术用途、使用方式,提出了大模型融入城市规划专业工作的技术范式、模型构建方法。以街道步行安全性专业大模型为案例,验证了技术范式和模型构建有效性,得出以下结论。

大模型融入城市规划专业工作的技术范式是以具备“通用知识+领域知识”知识结构的专业大模型支撑“诊断—推理—对策”全过程工作模式,支持问题导向、规律导向的规划范式。问题导向、规律导向的规划范式均遵循“诊断—推理—对策”的工作模式。大模型属于认知智能,具备依据知识推理进行“诊断—推理—对策”的能力。具备规划领域知识的专业大模型能支持城市规划专业工作中“诊断—推理—对策”的工作模式。

大模型融入城市规划专业工作的模型构建方式是以规划领域知识为核心,在通用大模型基础上进行专业性诊断训练、对策训练实现。本研究提出了构建规划专业大模型技术框架,包括梳理“空间问题—问题原因”的诊断领域知识、“问题原因—规划策略”的对策领域知识两类领域知识,完成诊断、对策两步训练。街道步行安全性专业大模型案例表明专业大模型能够达到“发现空间问题—推导问题原因—提出规划策略”的规划专业能力。

规划专业大模型的应用定位是规划专业工作助手,在存量规划时代更具必要性。人工智能专业大模型具有大规模、稳定性的优势,其应用定位是规划专业工作助手。街道步行安全性专业模型案例表明,专业大模型作为规划专业工作的助手,能高效处理海量信息,提供“循证”参考,有效提升规划师的诊断、施策效率和技术,实现了以人工智能增强规划师的专业工作能力。

5.2 讨论

随着人工智能技术发展,基础大模型的能力在迅速迭代,性能在迅速提升。近期测试表明部分通用大模型已能通过法律、财会等基础性资格考试,但是针对特定专业的领域知识仍将存在^[40-41]。在不断迭代通用大模型基础上再进行规划

领域知识学习,适用于城市规划专业工作的专业模型会有更为优异的“发现空间问题—推导问题原因—提出规划策略”能力。通用大模型能力迅速迭代并不意味着未来不需要专业大模型,而是能使专业大模型能力更加接近于人类专家。

规划专业人员需要关注规划专业工作中的人机交互复杂性。规划专业大模型应用定位是专业工作助手,这并不意味着规划专业工作难度降低。人机协作复杂性包含了两个方面:其一是专业人员需要具备评估大模型输出质量的能力,评估专业大模型所给出的空间问题诊断、规划策略的质量;其二是专业人员需要有梳理规划领域知识,才有可能通过诊断训练、对策训练构建“诊断—推理—对策”专业大模型。人机交互复杂性对专业人员提出了更高的要求。

本文讨论的大模型融入城市规划专业工作技术范式是将既有领域知识融入模型,让模型掌握城市规划专业知识,高效处理大规模专业工作任务。这类规划领域人工智能可归入工程智能(AI for Engineering)。人工智能驱动科学(AI for Science)已经是支撑新知识科学发现的一种新的科学研究范式^[42]。城市规划领域一样也需要有人工智能驱动城市科学的研究,让人工智能帮助挖掘和探知城市发展、城市运行的新知识。这将是另一个值得探索的研究领域。

注释

- ① 本研究所讨论的街道步行安全性聚焦行人对于犯罪威胁及相关人身伤害的感知,关注的是步行环境及其营造的社会空间安全性。交通安全不在讨论范围之内。
- ② 专业大模型的推理环境配备2张H20-NVLink显卡(单卡显存96 GB),搭配AMD EPYC 9K84处理器及300 GB系统内存。
- ③ 准确率(accuracy)衡量所有预测中正确预测的比例, F_1 分数(F_1 -Score)是精确率(precision)与召回率(recall)的调和平均。在街道步行安全性诊断与对策任务中,阳性样本(即存在特定问题原因或需空间干预的街道)通常占比较低,导致类别分布高度不均衡。在此情形下,accuracy虽能反映整体预测性能,但易受多数类(无问题街道)主导,难以捕捉对少数关键问题的漏判;相比之下, F_1 分数通

过兼顾精确率与召回率,能更有效地评估模型对问题街道的识别能力,对实际应用中的敏感性与可靠性更具指示意义。因此,本研究同时采用accuracy与 F_1 分数,以全面衡量模型在整体稳定性与关键问题检出能力之间的平衡表现。

参考文献

- [1] XIA J H, TONG Y, LONG Y. Advancements in the application of large language models in urban studies: a systematic review[J]. Cities, 2025, 165: 20.
- [2] ZHU H X, CHANG J, AN X Y, et al. Global and local feature extraction of urban historical spatial perception using large language models: a case study of Harbin Central Street District[J]. Cities, 2025, 165: 16.
- [3] KI D, LEE H, PARK K, et al. Measuring nuanced walkability: leveraging Chat GPT's vision reasoning with multisource spatial data[J]. Computers Environment and Urban Systems, 2025, 121: 13.
- [4] ZHANG J X, LI Y Q, FUKUDA T, et al. Urban safety perception assessments via integrating multimodal large language models with street view images[J]. Cities, 2025, 165: 13.
- [5] WU C, LIANG Y X, ZHAO M W, et al. Perceiving the fine-scale urban poverty using street view images through a vision-language model[J]. Sustainable Cities and Society, 2025, 123: 11.
- [6] LIU J X, ZHANG T C, MA Y C, et al. Generative artificial intelligence perspectives on typical landscape types: can ChatGPT compete with human insight? [J]. Landscape and Urban Planning, 2025, 264: 18.
- [7] HUANG W, WANG J, CONG G. Zero-shot urban function inference with street view images through prompting a pre-trained vision-language model[J]. International Journal of Geographical Information Science, 2024, 38 (7): 1414-1442.
- [8] 钮心毅,刘思涵,桑田,等.大模型的专业学习:构建融入城市空间形态设计知识的图像生成模型[J].城市规划学刊,2025(1): 55-63.
- [9] HUANG L, OKI T. Enhancing people's walking preferences in street design through generative artificial intelligence

- and crowdsourcing surveys: the case of Tokyo[J]. *Environment and Planning B: Urban Analytics and City Science*, 2025: 18.
- [10] WANG T Y, CHEN J, DENG Z Y, et al. Natural language processing for planning policy identification: a benchmarking study using 113 Chinese cities between 2011 and 2019[J]. *Environment and Planning B: Urban Analytics and City Science*, 2025: 18.
- [11] LIANG Y B, LIU Y C, WANG X H, et al. Exploring large language models for human mobility prediction under public events[J]. *Computers Environment and Urban Systems*, 2024, 112: 16.
- [12] LUO H Y, ZHANG Z D, ZHU Q, et al. Using large language models to investigate cultural ecosystem services perceptions: a few-shot and prompt method[J]. *Landscape and Urban Planning*, 2025, 258: 16.
- [13] ZHANG Y, LIN Y, TIAN L, et al. Leveraging LLM-based multi-agent simulations to boost participatory design education: an experimental exploration in residential area design[J]. *Sustainable Cities and Society*, 2025, 131: 16.
- [14] WANG X L, CHENG T, LAW S, et al. Multi-modal contrastive learning of urban space representations from POI data[J]. *Computers Environment and Urban Systems*, 2025, 120: 16.
- [15] SHAO Z Q, XI H N, LU H H, et al. A spatial-temporal large language model with denoising diffusion implicit for predictions in centralized multimodal transport systems[J]. *Transportation Research Part C-Emerging Technologies*, 2025, 179: 32.
- [16] TRICOT A, SWELLER J. Domain-specific knowledge and why teaching generic skills does not work[J]. *Educational Psychology Review*, 2014, 26 (2): 265 - 283.
- [17] DE GROOT A D. *Thought and choice in chess*[M]. Amsterdam: Amsterdam University Press, 2008.
- [18] ZHAO X K, HUANG H, YANG T, et al. Urban planning in the age of large language models: assessing OpenAI o1's performance and capabilities across 556 tasks [J]. *Computers Environment and Urban Systems*, 2025, 121: 17.
- [19] FU X, SANCHEZ T, LI C, et al. Deciphering public voices in the digital era[J]. *Journal of the American Planning Association*, 2024, 90 (4): 728-741.
- [20] 田莉, 杨鑫, 张雨迪, 等. “专业知识+人工智能”双驱动的城乡规划设计教育创新探索: 以住区规划为例[J]. *城市规划学刊*, 2024(5): 71-78.
- [21] YE X, HUANG T, SONG Y, et al. Generating conceptual landscape design via text-to-image generative AI model[J]. *Environment and Planning B: Urban Analytics and City Science*, 2025, 52 (8): 1903-1919.
- [22] GEDDES P. *Cities in evolution: an introduction to the town planning movement and to the study of civics*[M]. London: Williams, 1915.
- [23] MUMFORD L. *The culture of cities*[M]. New York: Open Road Media, 2016.
- [24] 吴志强, 张修宁, 鲁斐栋, 等. 技术赋能空间规划: 走向规律导向的范式[J]. *规划师*, 2021, 37(19): 5-10.
- [25] 叶锺楠, 吴志强. 城市诊断的概念、思想基础和发展思考[J]. *城市规划*, 2022, 46(1): 53-59.
- [26] 刘晓畅, 吴志强. 人工智能辅助国土空间诊断的理论范式建构与思行方法变革[J/OL]. *城市规划学刊*, (2026-1-20)[2026-01-05]. <https://link.cnki.net/urlid/31.1938.TU.20251111.1703.018>.
- [27] ZHU S, YU T, XU T, et al. Intelligent computing: the latest advances, challenges, and future[J]. *Intelligent Computing*, 2023, 2: 0006.
- [28] PENTLAND A. Perceptual intelligence [C]. *International Symposium on Handheld and Ubiquitous Computing*, 1999: 74-88.
- [29] WANG J, YU Y, TAN Y, et al. Artificial intelligence enables precision diagnosis of cervical cytology grades and cervical cancer [J]. *Nature Communications*, 2024, 15(1): 4369.
- [30] LAMPINEN A K, DASGUPTA I, CHAN S C, et al. Language models, like humans, show content effects on reasoning tasks[J]. *PNAS Nexus*, 2024, 3 (7): 233-245.
- [31] WANG H, ZHAO S, QIANG Z, et al. Knowledge-tuning large language models with structured medical knowledge bases for trustworthy response generation in Chinese[J]. *ACM Transactions on Knowledge Discovery from Data*, 2025, 19 (2): 1-17.
- [32] LI X, ZHANG C, LI W, et al. Assessing street-level urban greenery using Google street view and a modified green view index[J]. *Urban Forestry & Urban Greening*, 2015, 14 (3): 675-685.
- [33] 郝新华, 龙瀛. 街道绿化: 一个新的可步行性评价指标[J]. *上海城市规划*, 2017(1): 32-36.
- [34] LIN S, NIU X, ZHOU S, et al. Street walkability inference with street view imagery through fine-tuning and prompting pre-trained large models[J/OL]. *Environment and Planning B: Urban Analytics and City Science*, 2025: 17. <https://doi.org/10.1177/23998083251395490>.
- [35] XIAO Y, TANG Y W. Can ChatGPT-4o assess the perceptions of streetscape change? evidence from Shanghai, China[J]. *Sustainable Cities and Society*, 2025, 130: 16.
- [36] SUN P J, ZHAO H X, ZHONG J Q, et al. Popularity influence mechanism of coastal spaces in urban areas: insights from multi-modal large language models[J]. *Cities*, 2025, 161: 13.
- [37] ALFONZO M A. To walk or not to walk? the hierarchy of walking needs[J]. *Environment and Behavior*, 2005, 37(6): 808-836.
- [38] COZENS P, LOVE T. A review and current status of crime prevention through environmental design (CPTED) [J]. *Journal of Planning Literature*, 2015, 30(4): 393-412.
- [39] HU E J, SHEN Y, WALLIS P, et al. Lora: Low-rank adaptation of large language models[C]. *International Conference on Learning Representations*, 2022.
- [40] KATZ D M, BOMMARITO M J, GAO S, et al. GPT-4 passes the bar exam[J]. *Philosophical Transactions of the Royal Society A*, 2024, 382 (2270): 20230254.
- [41] CALLANAN E, MBAKWE A, PAPADIMITRIOU A, et al. Can GPT models be financial analysts? an evaluation of CHAT-GPT and GPT-4 on mock CFA exams [C]. *Proceedings of the Eighth Financial Technology and Natural Language Processing and the 1st Agent AI for Scenario Planning*, 2024: 23-32.
- [42] WANG H, FU T, DU Y, et al. Scientific discovery in the age of artificial intelligence [J]. *Nature*, 2023, 620 (7972): 47 - 60.

修回: 2026-02