

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

Domain-specific Learning for Large Models: Integrating Urban Spatial Form Design Knowledge into Image Generation Model

钮心毅 刘思涵 桑田 顾睿星 吴雪菲 王 浆

NIU Xinyi, LIU Sihan, SANG Tian, GU Ruixing, WU Xuefei, WANG Jiang

关键词 规划专业大模型；空间形态设计知识；隐性知识；图像生成模型；大模型微调；社区公共空间更新

Keywords: domain-specific large model for urban planning; spatial form design knowledge; tacit knowledge; image generation model; large model fine-tuning; community public space renewal

中图分类号 TU984 文献标志码 A
DOI 10.16361/j.upf.202501008
文章编号 1000-3363(2025)01-0055-09

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博士研究生

桑田，同济大学建筑与城市规划学院硕士研究生

顾睿星，同济大学建筑与城市规划学院硕士研究生

吴雪菲，同济大学建筑与城市规划学院硕士研究生

王 浆，同济大学建筑与城市规划学院硕士研究生

提 要 在通用大模型基础上构建专业大模型，其难点在于将专业领域知识融入通用大模型。城市空间形态设计中存在大量隐性知识，而城市规划的公共政策属性决定了这类隐性知识必须向公众高效传递。利用大模型独特的学习和传递隐性知识的能力，提出以大模型专业学习方式将城市空间形态设计知识融入通用图像生成模型的技术框架。通过以设计导向为核心，梳理空间形态设计中的隐性知识和显性知识，指导大模型专业学习，以大模型微调实现城市空间形态设计知识融入通用图像生成模型。以社区公共空间更新为例，构建融入城市空间形态设计知识的专业大模型。案例应用表明专业大模型能够起到高效传递隐性知识的关键作用，提升社区更新中沟通交流效率，有效支持参与式设计。

Abstract: The primary challenge in developing a domain-specific large model based on a general large model lies in effectively integrating domain-specific knowledge into the general large model. There is a significant amount of tacit knowledge in urban spatial form design, and the public policy nature of urban planning necessitates the efficient communication of tacit knowledge to the public. Leveraging the unique ability of large model to learn and transfer tacit knowledge, this paper proposes a technical framework for integrating urban spatial form design knowledge into general image generation models through domain-specific learning. Tacit and explicit knowledge in spatial form design with design orientation constitutes the core components to guide the domain-specific learning of large models. The integration of urban spatial form design knowledge into a general image generation model is accomplished through the application of large model fine-tuning techniques. Using community public space renewal as a case study, a domain-specific large model incorporating the knowledge of urban spatial form design is constructed. The case shows that the domain-specific large model plays a crucial role in effectively transferring tacit knowledge, enhancing communication efficiency in community renewal, and supporting participatory design processes.

* 国家重点研发计划课题“国土空间多场景综合效能评价与调控关键技术”（项目编号2022YFC3800804）

随着生成式人工智能 (GenAI) 的发展, 大模型 (Large Models) 能够从海量数据中学习语言、图像、音频等, 生成新的内容, 已经在自然语言处理、计算机视觉等领域展示出了强大能力, 正迅速进入各个学科^[1]。多种通用大模型不断涌现也引发了规划学科广泛关注^[2]。图像生成模型是当前一类重要的大模型。当前通用图像生成模型可以通过提示词 (prompt) 指令等方式引导生成图像, 已在工艺美术、平面设计、漫画绘本等设计行业得到了直接应用。

构建专业大模型是众多学科正在探索的议题。由于专业领域应用具有特定的任务需求、特定的数据分布和评价标准, 直接使用通用大模型无法满足这些要求。一方面, 大模型训练需要海量的数据和算力资源, 会形成非常高昂的成本。另一方面, 通用大模型本身就具有基础模型的概念, 基于自监督学习的模型在预训练过程中已经具备了通用知识, 这为专业领域应用提供了基础。因此, 当前一般方式是在通用大模型的基础上构建专业大模型。通用图像生成模型已经能通过大规模预训练学习到众多图像生成知识, 没有必要重新从预训练开始构建专业领域图像生成模型。在通用图像生成模型基础上构建专业大模型, 这已经有了探索, 如建筑设计中已经在探索通过专业训练, 将图像生成模型用于设计方案和效果图生成等^[3-4]。

构建规划专业大模型也是当前规划学科的需求。人工智能是让机器掌握人类知识, 从而实现模仿人类行为。由于通用大模型没有针对城市规划专业知识的训练, 城市规划专业大模型同样需要在通用大模型基础上构建。继续以图像生成模型为例, 其能够理解和生成涵盖广泛主题和风格的图像, 包括了自然景观、人物肖像、艺术文化, 甚至城市建筑等, 但仍缺乏城市规划专业知识。通用图像生成模型直接用于城市空间形态设计, 或是模型不理解规划专业术语无法生成对应图像, 或是生成明显异常的城市空间图像, 根本原因是通用大模型尚不具备城市规划专业知识。构建规划专业大模型的关键难点在于如何将规划专业知识融入通用大模型。

城市空间形态设计知识是城市规划

专业知识中的重要组成部分, 是规划学科特有的专业知识。本文以构建规划专业模型技术途径为研究目的, 通过梳理城市空间形态设计的知识体系, 探索以专业学习方式将城市空间形态设计知识融入通用图像生成模型的途径, 以社区公共空间更新应用为例, 实现在通用图像生成模型基础上构建社区公共空间更新专业大模型。

1 城市空间形态设计知识与大模型

1.1 显性知识、隐性知识与大模型

让城市空间形态设计中的专业知识融入大模型, 首先需要对城市空间形态设计知识有清晰的认识。在管理学科的知识管理中, 依据知识能否被清晰地表述和有效地转移, 将其分为显性知识 (explicit knowledge) 和隐性知识 (tacit knowledge)^[5]。两者分别代表了人类知识体系中的不同部分。

显性知识是一种形而上的理性知识, 注重知识的明确性和系统性。它能够以一定的符号系统, 如语言、数学公式、图表等加以完整的表述^[6]。数学中的定理是显性知识的典型例子, 例如勾股定理 ($a^2 + b^2 = c^2$) 的公式形式明确, 很容易地通过教科书等形式进行传播和学习。只要理解了公式, 就可以使用它来解决直角三角形相关的问题。

隐性知识则是一种亲身参与的经验知识, 主要依赖个人的经验。它不能使用明确的符号系统加以完整地表述^[6]。艺术家的创作技巧是隐性知识的典型。虽然艺术家会分享绘画等基本技能, 但真正的艺术创造力和灵感往往来自自身多年实践和经验积累。这些经验和感觉很难用语言表达传授, 但却是作品独特性的关键所在。

传统的人工智能技术一直难以应对隐性知识^[7], 直到大模型出现使得情况出现了转变。大模型的训练基于大量互联网无标注数据, 在训练过程模拟了人类在社群中学习和沟通的环境, 使其不仅能学习到显性知识, 也能学习到隐性知识。例如, 自然语言中包含了大量的隐性知识。大语言模型通过预训练学习到大量的语言知识, 掌握了其中的隐性知

识^[8]。通用图像生成模型已经能够在绘画等艺术创作中应用, 说明大模型也能通过大量的预训练掌握绘画中的隐性知识。

1.2 城市空间形态设计中的隐性知识

城市空间形态设计知识体系中同时包含了显性知识和隐性知识。虽然部分城市空间形态知识可以用明确的公式表示, 比如街道宽度与两侧建筑物高度的比值街廓比 (D/H), 但是空间形态设计中更多是源于经验、难以明确表达或书写的隐性知识。例如对于“街道活力”, 有经验的规划师可以根据这一提示设计出符合该导向的城市街道空间, 但很难用具体的语言、公式对其进行清晰的描述。城市空间形态设计中还有很多类似“街道活力”的隐性知识, 这些知识依赖于规划师多年的从业经验。

城市规划专业知识中存在大量隐性知识, 尤其城市空间形态设计中隐性知识占据了主导, 这也是空间形态设计知识难以采用文字资料、固定规则、公式等方式进行系统性的知识提取、知识表达的原因^[9]。规划师在学习掌握设计知识时, 需要通过大量观察、模仿等方式来领悟其中的隐性知识。长期以来, 虽然诸多研究在探索采用多种定量解析城市空间形态方法, 包括了形状特征解析法^[10]、空间句法^[11]等, 但实质上均在探索将城市空间形态设计中的隐性知识转为显性知识。然而, 至今尚无法将所有的空间形态设计知识均转为显性知识。

1.3 传递城市空间形态设计隐性知识的必要性

隐性知识的表述和传递一直是知识交流中的难点。由于隐性知识是一种非语言的表述形式, 参与交流的个体的知识背景、沟通表达能力、领悟能力等均影响了隐性知识的交流^[6]。虽然隐性知识难以清晰地表达和传递, 但城市规划学科中的隐性知识必须进行广泛的传播和传递。这是由规划学科自身特点决定的。

首先, 公共政策属性决定了规划专业知识、信息的传递不能只发生于规划专业人员之间, 规划学科的隐性知识必须向非专业人士传递。空间形态设计中隐性知识也需要向普通公众传递、交流。尤其全周期规划过程中, 需要居民、规

划师、决策者之间高效交流，这样的参与式设计迫切需要传递隐性知识。如果将“老年友好设计”“激发街道活力”等晦涩用词转译为通俗的图像和语言表达，规划策略讨论将会非常高效。其次，隐性知识教学占据了规划设计教学的相当大部分内容。规划教育中的“设计课”就是讲授、传递隐性知识。课堂上讲授“尺度宜人”“空间压抑感”等对城市空间的描述是典型的传递隐性知识，如果能迅速转译为通俗的图形图像表达，这样知识传递效果会大为提高。

大模型已经在学习掌握隐性知识方面显示出了独特的能力。规划学科隐性知识传播、传递迫切需要大模型，这是规划专业独特的需求。城市空间形态设计中包含隐性知识要远多于能用公式传达的显性知识。已有研究探索了图像生成模型应用于城市空间形态设计教学^[12]，这些隐性知识恰好是大模型能发挥重要作用的场合。

2 图像生成模型的专业学习途径

2.1 图像生成模型的通用知识学习和领域知识学习

一般将大模型学习掌握的知识区分为通用知识（general knowledge, GK）、领域知识（domain knowledge, DK）两大类^[13-14]。通用知识通过预训练方式输入图像生成模型。当前扩散模型（diffusion model）是图像生成模型中的主流技术^[15]。由于实现了文本与图像的跨模态关联，扩散模型能理解图像和文本之间的关系，形成了具备文字生成图像能力的图像生成模型^[16]。通用图像生成模型的文本到图像生成实质上是一种通用知识表达的过程。

领域知识需要通过专业学习方式输入通用图像生成模型。微调（fine-tuning）是图像生成模型广泛使用的专业学习方式，它可以在保持模型结构不变的情况下，通过调整少量参数来适应新的任务和数据^[17]。不同种类微调技术能够适应不同类型的专业领域知识，取决于领域知识的融入程度和专业化程度。

既有研究中，通过微调对大模型展开专业学习有两种模式：第一类是以一种新的方式激活大模型本身具有的通用

知识，从而使得大模型能够胜任专业领域中的“一般”任务，一般对应提示词工程（prompt engineering）等微调技术。例如：让大模型扮演公众参与中的不同角色，从而提升规划决策效率^[18]；通过构建城市分类法的提示词工程，把抽象的城市功能映射到具象的城市视觉要素上，从而实现从街景图片中提取城市功能^[19]。公众角色扮演、城市视觉要素的图文对应均是大模型已掌握的通用知识。第二类是让大模型补充学习专业领域更为系统的知识，从而使得大模型能够胜任专业领域中的“专家”任务，一般对应低秩适应（low-rank adaptation, LoRA）、检索增强生成（retrieval augmented generation, RAG）等微调技术。例如，通过微调的大模型进行金融时间序列的预测、金融推理等工作^[20]。专业学习就是特定专业领域知识融入通用大模型的一种有效途径。

2.2 城市空间形态设计知识与图像生成模型的专业学习方式

专业学习方式应当针对城市空间形态设计知识的特性，以应对隐性知识带来的挑战^[21]。城市空间形态设计知识体系可以分为空间形态设计知识和空间表达知识两类。城市空间形态的内在构成是空间要素及其空间关系^[22]。在规划设计中，通过运用空间形态设计知识将空间要素以不同的关系组合起来，形成规划设计方案；运用空间表达知识将规划设计方案表达为规划设计图纸、图像。城市空间形态设计知识是典型的领域知识，通用图像大模型必须进行专业学习才能实现学习掌握。

以通用开源图像生成模型Stable Diffusion^①为例，LoRA、控制网 ControlNet 是其两类型典型的微调技术。LoRA 微调技术能够以少量专业训练激发通用知识与专业知识的衔接，并完成新知识的融入^[23]，适用于空间形态设计知识的专业学习。控制网 ControlNet 允许用户使用图像或图像特征图作为条件来控制生成图像的特定结构和风格^[24]，适用于空间表达知识的专业学习。

图像生成模型既有的微调方式为城市空间形态设计知识融入大模型提供了技术基础，但还需建立两者之间的对应

技术框架。以下以社区公共空间更新为应用案例，构建城市空间形态设计知识融入通用图像生成模型的技术框架，实现通用大模型的专业学习。

3 大模型的专业学习实践案例：社区公共空间更新模型

3.1 技术框架

既有社区公共空间更新是当前城市更新中的重要内容^[25-26]。当前在社区更新中鼓励采用参与式设计的形式，在更新意愿沟通、细化更新对象、比选更新意向等环节均强调居民的广泛参与^[27]。通常的做法是规划师使用专业图纸，如平面图、剖面图、效果图等与居民交流。专业图纸存在阅读和理解的门槛，使得规划师与居民交流中存在沟通理解的障碍。居民一般不具备规划专业知识，会难以理解设计导向等规划专业术语，对设计要素的认知不足，在与规划师沟通的过程中无法在意向效果与实际效果之间建立统一的想象^[28]。社区更新中的参与式设计是典型的需要传递、交流空间形态设计隐性知识的场合。

本研究在通用大模型基础上构建社区公共空间更新模型 CommunityGenAI，用于支持社区更新中的参与式设计，目标是实现输入若干“专业”提示词，模型对社区公共空间照片进行局部重绘，智能生成更新后场景。首先，通过梳理城市空间形态设计知识，将专业知识与大模型微调技术对应，提出了融入城市空间形态设计的图像生成模型技术框架。在这个技术框架中，城市空间形态设计专业知识体系作为领域知识，分为空间形态设计知识和空间表达知识两大类。空间形态设计知识以隐性知识占主导，表达为以设计导向为核心的知识图谱，指导训练集的构建，使训练集中包含完整的显性、隐性知识，通过LoRA微调训练实现“专业学习”。空间表达知识是显性知识，通过ControlNet图像条件控制实现“专业学习”，引导图像生成（图1）。这一技术框架能够实现领域知识和通用知识的融合。

3.2 空间形态设计知识图谱

社区公共空间更新中包含了大量的

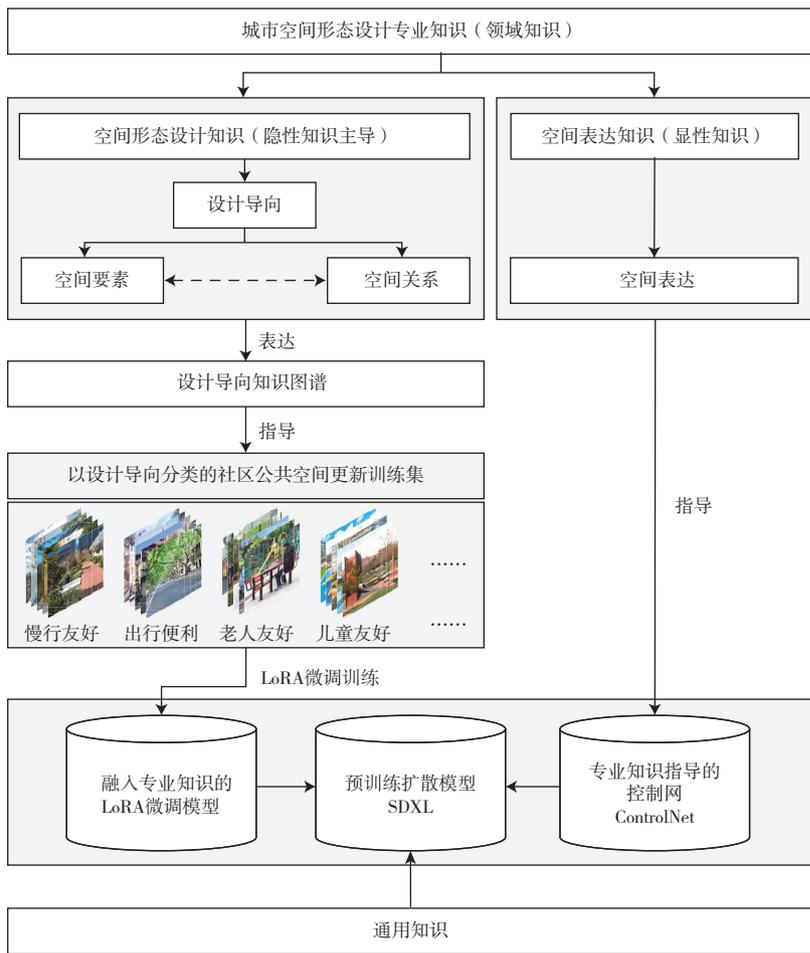


图1 技术框架
Fig.1 Technical framework

隐性知识。语义网络、知识地图等图数据库可以更清晰地表达隐性知识^[29]。在语义网络的基础上，谷歌公司提出了知识图谱(knowledge graph)的概念^[30]。知识图谱以有向图的方式对概念和事物及其相互之间的语义关联进行结构化组织，是一种可为计算机理解和利用的语义网络，其中，节点代表实体，边代表实体之间的语义关系^[31]。知识图谱已经能用于解析城市空间形态中的空间要素、空间关系^[22, 32]。本研究采用知识图谱形式梳理表达空间形态设计知识。

将社区公共空间更新中的空间形态设计知识分为设计导向、空间要素、空间关系等三类。设计导向体现社区公共空间更新目标，空间要素指建筑、道路、健身设施、绿化等。空间关系表达空间要素之间的关系，比如空间相邻与包含关系、空间尺度等。诸如场地内包含了

若干健身设施，健身设施与道路相邻，健身设施与人的比例等。空间要素、空间关系在设计导向下组织为社区公共空间。例如：“老年友好”设计导向是在更新中以老年群体的需求为核心组织公共空间设计，体现为提供有针对性的服务设施、场地铺装与休闲装置等，以便利老年群体为宗旨组织空间。“运动休憩”设计导向倡导全民健身，以体育健康为目标组织公共空间设计，体现为增设各类运动设施和运动场地，为社区全民健身便利组织空间。

空间要素及其属性属于显性知识，可以清晰地用文字表述。设计导向、空间关系属于典型的隐性知识，很难用清晰文字表述。知识图谱可以有效组织和管理上述城市空间形态的知识^[33]。图2表示了“儿童友好”导向下部分知识图谱。知识图谱中，节点代表空间要素，边代

表空间关系。以知识图谱表达是为了模型能够有针对性地学习图像中的空间形态设计知识。

3.3 空间形态设计知识的专业学习

以设计导向为核心，采用知识图谱形式梳理社区更新中空间形态设计知识。以知识图谱指导选择训练集图片，通过LoRA微调训练，将城市空间形态设计中的显性知识、隐性知识融入图像生成模型。为此，在CommunityGenAI模型中设立了8类设计导向来组合各类空间要素，分别是出行便利、慢行友好、艺术文创、运动休憩、老年友好、儿童友好、商业氛围、自然生态。

以设计导向知识图谱为主导，专业学习目的是使大模型理解设计导向及其所包含的空间要素、空间关系知识。以“儿童友好”导向的社区活动场地更新为例，儿童、儿童游乐设施、橡胶跑道、防护装置等空间要素本身属于显性知识，但它们组合指向了“儿童友好”设计导向属于隐性知识(图2)。以设计导向知识图谱为指导进行训练集图片选择，引导大模型将具体的空间要素与“儿童友好”的设计导向建立联系，以实现设计导向隐性知识融入大模型。

依据设计导向知识图谱选择训练集图片，空间关系隐性知识也包含在训练集图片中，模型能够学习到“儿童游乐设施位于塑胶场地内”“防护装置与塑胶场地相邻”等空间包含、相邻的空间尺度关系(图2)。这些空间关系隐性知识通过训练集传递给大模型，实现“儿童友好”导向下的空间关系隐性知识融入大模型。

3.4 空间表达知识的专业学习

城市空间形态设计针对三维城市空间，但需要通过二维图纸进行表达。三维空间的二维表达也是设计中的重要知识。专业规划师必须学习画法几何等专业课程，就是在学习空间表达知识，实现三维空间转换为二维图像表达。空间表达知识是显性知识，可以用明确的公式和几何方法进行表述。

用图像生成模型表达社区公共空间更新效果，体现对图片智能局部重绘。局部重绘必须保证图像整体上空间表达

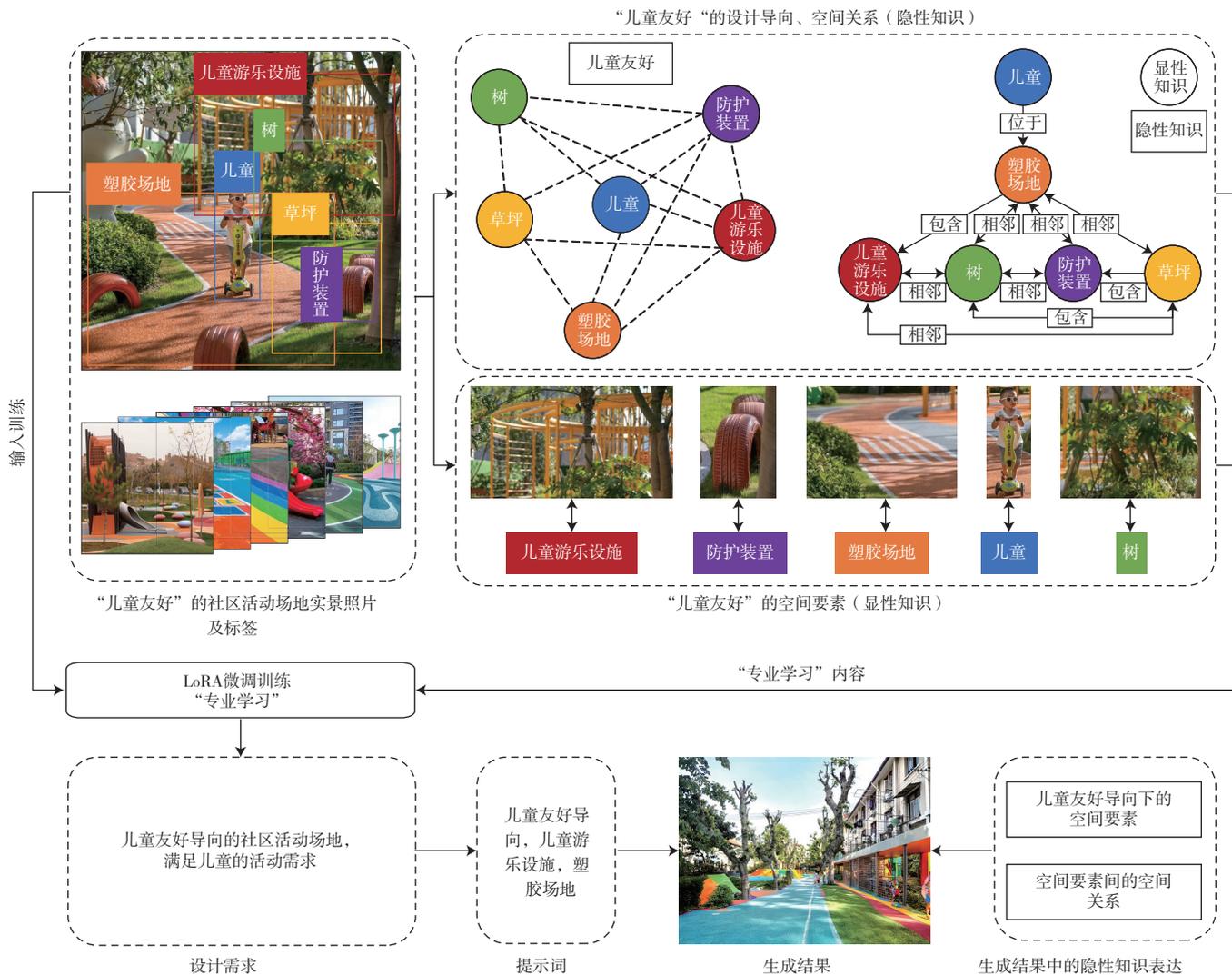


图2 空间形态设计知识“专业学习”和表达的路径

Fig.2 Pathway of "domain-specific learning" and expression of spatial form design knowledge

一致性。这就需要空间表达知识“专业学习”，确保更新后场景的正确空间表达。

将空间表达知识归为透视关系、遮挡关系与相对尺寸关系。首先，以社区街道空间为例，图3实景图首先表达了街道空间的一点透视关系。其次，三维空间要素投影到二维图面后，前方物体会遮挡后方物体。最后，物体在真实三维空间中的相对尺寸会通过投影体现在二维图面上，图3中人与电动车之间的相对高度关系真实地表达在了图面上。上述空间表达知识均能使用画法几何中的定理表示，进而通过不同的ControlNet

图像条件控制微调融入大模型。

在Stable Diffusion中，ControlNet扩展能够通过语义图、线框图等方式转为图像条件控制，适宜用于表达和传递空间表达知识。本研究提出以ControlNet图像条件控制微调作为空间表达知识专业学习的途径，在既有ControlNet图像条件控制与空间表达知识之间建立对应框架。其中，透视关系使用边缘检测和控制类ControlNet输入图像条件控制，遮挡和相对尺寸关系使用Segmentation（语义分割）与Tile（分块）输入图像条件控制（图3）。

4 社区公共空间更新模型 CommunityGenAI 构建与实践案例

4.1 标签集构建与数据准备

CommunityGenAI模型专业学习的训练集由标签和图像两部分组成。训练集和标签的构建依托设计导向知识图谱。其中标签用于表达城市空间形态设计中设计导向、空间要素的知识。在模型的应用流程中，标签将直接作为提示词对图像生成过程进行文本控制。标签集的构造对照设计导向知识图谱及其对应的空间要素。

图像是模型专业学习的对象，是城

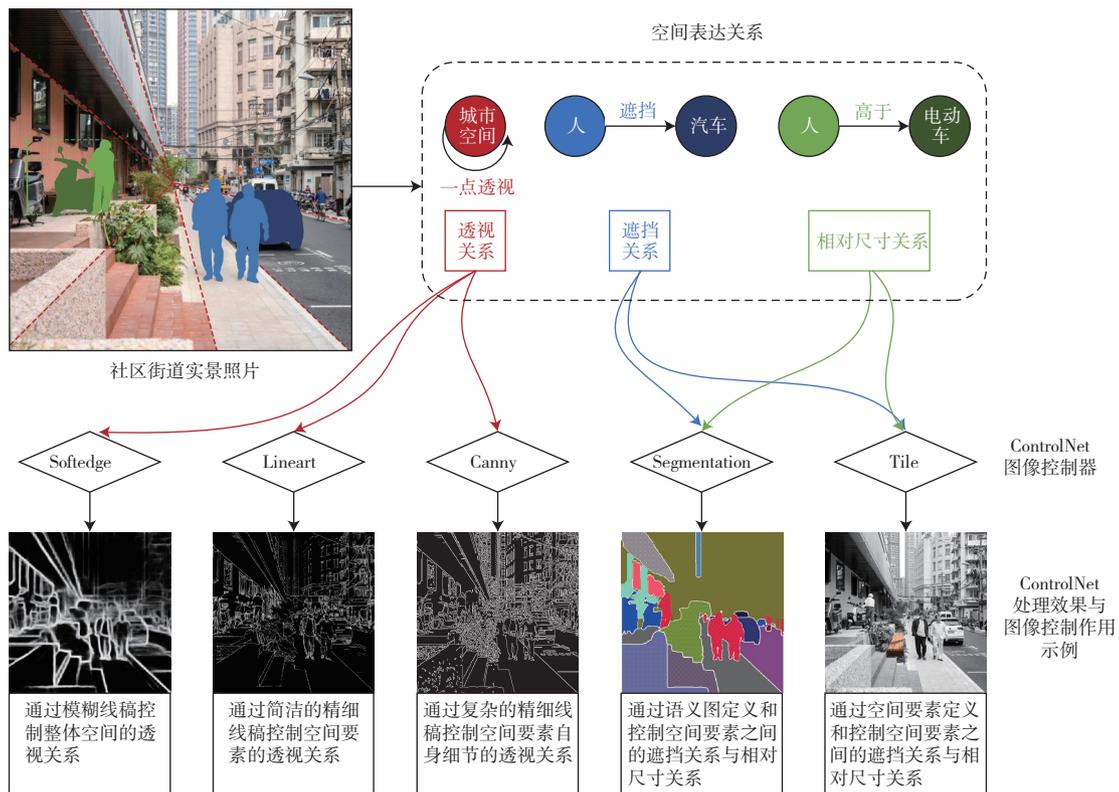


图3 空间表达知识“专业学习”和表达的路径

Fig.3 Pathway of "domain-specific learning" and expressing spatial expression knowledge

市空间形态设计知识的直观表达。以设计导向知识图谱为指导，由规划专业人员选取设计导向体现完整、空间要素表现清晰、空间品质较高的社区场景实景照片作为训练集。8类设计导向知识图谱指导下，共收集436张社区场景实景照片，涵盖多种类型居住社区，以保证模型的专业学习能够适应多种不同年代居住社区的更新需求。

依据设计导向知识图谱，将实景照片按照设计导向进行分类，并对照设计导向下属的空间要素对每张图片进行标注，保证每个标签及其所代表的空间要素在训练集中至少出现5次，形成标签文件。

4.2 LoRA 微调训练融入城市空间形态设计知识

通过 LoRA 微调训练将城市空间形态设计中的设计导向、空间要素、空间关系知识融入大模型。其中：空间要素作为显性知识，能够用标签集进行直接标注；设计导向、空间关系则属于隐性知识，通过训练集实景照片表达。由于

表1 模型训练的软、硬件环境

Tab.1 Software and hardware environment for model training

硬件	配置
CPU	Intel Xeon Platinum 8352V, 36核, 2.10 GHz, 54 MB
GPU	NVIDIA RTX 4090, 24GB
内存	120G
操作系统	Linux
基础通用大模型	Stable Diffusion XL base 1.0

表2 训练参数设置

Tab.2 Training parameter settings

模型参数	含义	取值
epoch	循环次数	100
batch_size	批大小	1
unet_lr	U-Net 模型学习率	1
text_encoder_lr	文本编码器学习率	0.5
optimizer_type	优化器类型	AdaFactor
network_Dim	网络规模	128

LoRA 微调具有以少量训练集完成新知识融入的特征，由此实现空间形态设计知识的融入。

LoRA 训练使用 Stable Diffusion XL 版本下的“base 1.0”大模型进行软硬件环境和参数设置如表1和表2所示。在

进行了100轮的训练迭代后，损失函数(Loss)值稳定在0.08左右，模型展现出最佳的训练效果。通过X/Y/Z图表对不同权重下LoRA微调模型进行性能测试。基于测试结果，确定LoRA的最优权重为0.4，并选取最后一轮训练的模型作为最终LoRA微调结果。

4.3 空间形态设计专业知识在CommunityGenAI模型中的体现

CommunityGenAI模型通过专业学习融入了设计导向、空间要素、空间关系的专业知识，也融入了空间表达知识。在社区更新中，选取既有社区公共空间的图片，输入设计导向提示词、关键空间要素的提示词，就能智能化生成局部更新后公共空间场景意向图。

以“儿童友好”的社区活动场地更新为例，通过输入“儿童友好、儿童友好色彩、儿童游乐设施、沙坑”作为提示词，得到生成符合以上设计导向的意向图(图4)。模型生成的意向图中不仅包含了儿童游乐设施、沙坑，色彩导向符合儿童友好色彩，场景中还包含了儿童及其身边的家长、塑胶场地等要素。提示词并没有包含“家长”“塑胶场地”，这就是“儿童友好导向”一词所包含的隐性知识，游乐设施、树木、沙坑等要素之间相对空间位置、相对尺寸也是专业学习中掌握的隐性知识。CommunityGenAI模型融入了社区更新中的空间形态设计知识，尤其是融入了其中的隐性知识。

由于按照设计导向进行了专业学习，CommunityGenAI模型能依据特定设计导向提示词，在同一个社区更新案例中智能化生成不同设计导向的意向图，以供居民比较选择。图5显示了以某社区内部宅间道路作为更新对象，针对“商业氛围”“老人友好”“出行便利”等3类设计导向生成的3个更新意向图。“商业氛围”导向所生成的桌椅、树下的花坛显著区别于“老人友好”导向的桌椅与“出行便利”导向的树下花坛(图5)。各个设计导向下的意向图所包含的空间要素有显著差异，均能体现该导向下社区公共空间的典型特征，尺度均符合社区公共空间意向图的空间表达要求。这表明CommunityGenAI模型通过专业学习掌

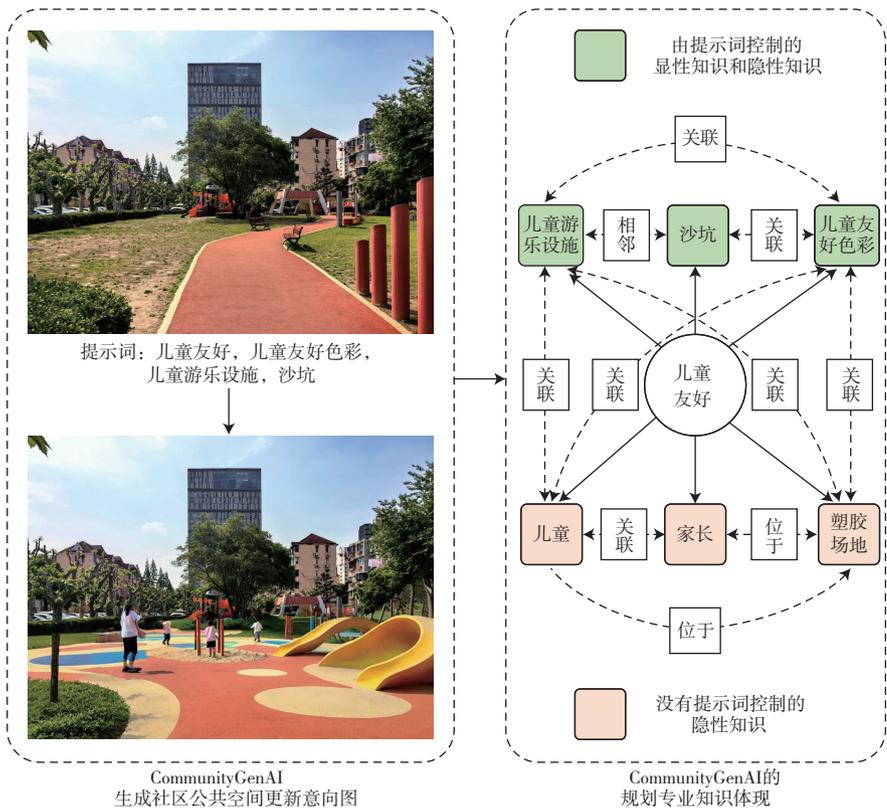


图4 CommunityGenAI模型的专业知识体现
Fig.4 Representation of domain knowledges in CommunityGenAI Model



图5 CommunityGenAI模型智能生成3类设计导向的意向图
Fig.5 CommunityGenAI Model intelligently generates concept maps for three design orientations

握了特定设计导向所包含的隐性知识。

CommunityGenAI模型对于社区公共空间更新中的设计导向、空间要素、空间关系、空间表达知识进行了有效的专业学习,掌握了社区公共空间更新中的显性知识和隐性知识,具备了社区公共空间更新专业大模型的能力。

4.4 CommunityGenAI模型中的专业知识传递

融入城市空间形态设计知识的图像生成模型能够实现隐性知识的表达和传递,在社区更新的前期沟通阶段、方案交流反馈阶段均能发挥有效作用,支持社区更新中居民、规划师之间的交流。以某老旧小区中硬质活动场地更新为例:在规划师与居民更新意愿沟通的环节,模型能够帮助规划师和居民确认更新对象和范围;在细化设计导向环节、比选更新意向环节,规划师均能通过CommunityGenAI将居民需求迅速转为“儿童友好导向、儿童游乐设施”等规划专业术语以及相应的控制网,模型能够以平均40 s/张图的速度智能生成一系列符合设计导向的更新意向图。这一过程完成了从居民需求到规划师专业术语的快速对接,并且将专业知识中难以表述的隐性知识以易于理解的图像反馈呈现给居民。CommunityGenAI实现了社区公共空间更新专业知识在规划师、居民之间高效传递,尤其实现了隐性知识的有效表达和传递,支持居民和规划师之间达成一致的更新意向,进而支持社区更新中的参与式设计。

5 结论和讨论

5.1 结论

本文以“领域知识”认识城市空间形态设计专业知识,进一步引入了管理学科的隐性知识、显性知识的概念,分类认识城市空间形态设计专业知识,分别对应大模型微调技术,提出了两类专业知识学习途径,构建了城市空间形态设计专业知识融入的通用图像生成模型的技术框架。以社区公共空间更新为应用案例,构建了专业大模型CommunityGenAI。研究得到以下三点结论。

第一,规划学科知识体系中存在大

量隐性知识,尤其在城市空间形态设计中隐性知识占据了主导。城市空间形态设计中隐性知识的交流和传递需要大模型发挥作用。传递城市空间形态设计中的隐性知识是规划学科对大模型特有的需求。构建空间形态设计大模型的关键难点在于如何将城市空间形态设计知识融入通用大模型。

第二,提出了以专业学习方式将城市空间形态设计知识融入通用图像生成模型的技术框架,将专业知识分为空间形态设计知识和空间表达知识。前者以设计导向为核心形成隐性知识和显性知识的知识图谱,后者以显性知识的方式指导图像表达,使用两类微调技术实现城市空间形态设计知识融入通用图像生成模型。

第三,以通用图像生成模型Stable Diffusion为基础,通过大模型微调,将社区公共空间更新设计专业知识融入通用模型,构建了支持社区更新中参与式设计的专业大模型CommunityGenAI。训练得到CommunityGenAI模型实现了专业知识的学习,具备了社区公共空间设计导向知识、空间要素知识、空间关系知识、空间表达知识。在参与式设计场景中,CommunityGenAI实现了专业知识的传递,能起到支持高效传递交流隐性知识的关键作用,提升了社区更新中沟通、交流的效率,能有效支持参与式设计。

5.2 讨论

在规划领域,通用大语言模型已可以在规划设计过程模拟利益相关方,起到有效模拟公众参与的作用^[34]。这一方式是将大模型视为仅具备“通用知识”的公众。本文在通用大模型基础上构建规划专业模型,是通过大模型专业学习方式,让专业模型成为同时具备“通用知识+领域知识”的专家,从而能在专业领域发挥独特作用。以上两种途径均将是是是是大模型技术应用于规划学科的有效途径,均值得深入探索。

大模型专业学习也是构建其他类型规划专业模型的技术途径,如何将规划专业知识融入通用大模型会成为构建规划专业大模型的关键技术。从知识管理的维度出发,规划学科不同领域的专业知识本身就有差异。不只是显性知识、

隐性知识的专业学习途径会有不同;不同种类隐性知识专业学习途径也会不同。各类规划专业知识融入通用大模型方法必然是不一样的。为此,分领域、分门别类地构建规划专业大模型可能是比较适宜的路径。这方面值得进行系统性的研究。

众多学科关注专业大模型,大致有两个方面的原因。一是在科学研究中,将大模型用于处理人工难以完成的任务,如实现海量数据处理分析等,从而提升科研效率;二是在工程应用中,将大模型用于提升设计效率,优化生产流程,从而降低工作成本,提升工作效率。对于城市规划学科,除了存在以上两个方面的需求,还存在规划学科特有的需求。这就是规划学科的隐性知识传播、传递需要大模型,这是规划学科的公共政策属性所决定的。如何让大模型学习和掌握规划学科隐性知识,在全周期空间规划中发挥作用,这将是规划专业大模型的独特需求。这方面值得进行深入探索和研究。

注释

① Stable Diffusion是Stability AI公司发布的通用图像大模型。Stable Diffusion XL 1.0是2023年7月发布的开源版本,其基本模型(base model)拥有35亿个参数。

参考文献

- [1] WANG H, FU T, DU Y, et al. Scientific discovery in the age of artificial intelligence [J]. Nature, 2023, 620: 47-60.
- [2] Otthein HERZOG, 潘海啸, 邓智团, 等. 新一代人工智能赋能城市规划:机遇与挑战[J]. 城市规划学刊, 2023(4): 1-11.
- [3] 袁潮, 郑豪. 生成式人工智能影响下的建筑设计新模式[J]. 建筑学报, 2023(10): 29-35.
- [4] 何宛余, 杨良崧. 生成式人工智能在建筑设计领域的探索:以小库AI云为例[J]. 建筑学报, 2023(10): 36-41.
- [5] POLANYI M. The tacit dimension[M]. Roulledge and Kegan Paul, 1966.
- [6] 姚伟. 知识管理[M]. 北京:清华大学出版社, 2020.
- [7] SANZOGNI L, GUZMAN G, BUSCH P. Artificial intelligence and knowledge management: questioning the tacit dimen-

- sion[J]. Prometheus, 2017, 35(1): 37–56.
- [8] JIN K, ZHUO H H. Integrating AI planning with natural language processing: a combination of explicit and tacit knowledge[J]. ArXiv, 2022, 2202.07138.
- [9] 钮心毅, 林诗佳, 桑田, 等. 数字化规划技术: 数据与知识 [J]. 城市规划学刊, 2024 (2): 18–24.
- [10] 林炳耀. 城市空间形态的计量方法及其评价[J]. 城市规划汇刊, 1998(3): 42–45.
- [11] HILLIER B. The hidden geometry of deformed grids: or, why space syntax works, when it looks as though it shouldn't[J]. Environment and Planning B: Planning and Design, 1999, 26(2): 169–191.
- [12] KAPSALIS T. UrbanGenAI: reconstructing urban landscapes using panoptic segmentation and diffusion models[J]. ArXiv, 2024, 2401.14379.
- [13] WANG H, FU T, DU Y, et al. Scientific discovery in the age of artificial intelligence [J]. Nature, 2023, 620 (7972): 47–60.
- [14] MECKLENBURG N, LIN Y, LI X, et al. Injecting new knowledge into large language models via supervised fine-tuning [J]. ArXiv, 2024, 2404.00213.
- [15] HO J, JAIN A, ABBEEL P. Denoising diffusion probabilistic models[J]. Advances in Neural Information Processing Systems, 2020, 33: 6840–6851.
- [16] RADFORD A, KIM J W, HALLACY C, et al. Learning transferable visual models from natural language supervision[C]//International Conference on Machine Learning. PMLR, 2021: 8748–8763.
- [17] RUIZ N, LI Y, JAMPANI V, et al. Dreambooth: fine tuning text-to-image diffusion models for subject-driven generation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 22500–22510.
- [18] ZHOU Z, LIN Y, JIN D, et al. Large language model for participatory urban planning[J]. ArXiv, 2024, abs/2402.17161.
- [19] 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.
- [20] NIE Y, KONG Y, DONG X, et al. A survey of large language models for financial applications: progress, prospects and challenges[J]. ArXiv, 2024, abs/2406.11903.
- [21] 甘惟, 吴志强, 王元楷, 等. AIGC辅助城市设计的理论模型建构[J]. 城市规划学刊, 2023(2): 12–18.
- [22] 杨俊宴, 邵典, 汪鹏, 等. 集成·拓扑·转译: 一种基于知识图谱的城市形态深度解析方法[J]. 城市规划, 2023, 47(6): 57–67.
- [23] HU E J, SHEN Y, WALLIS P, et al. Lora: low-rank adaptation of large language models[J]. ArXiv, 2021, 2106.09685.
- [24] ZHANG L, RAO A, AGRAWALA M. Adding conditional control to text-to-image diffusion models[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2023: 3836–3847.
- [25] 王嘉, 白韵溪, 宋聚生. 我国城市更新演进历程、挑战与建议[J]. 规划师, 2021, 37 (24): 21–27.
- [26] 华霞虹, 庄慎. 以设计促进公共日常生活空间的更新: 上海城市微更新实践综述[J]. 建筑学报, 2022(3): 1–11.
- [27] 李晴, 林妮. “人民城市”视角下社区微更新参与式规划设计的新模式探索: 以上海市YF里弄微更新为例[J]. 城市规划学刊, 2023(6): 87–96.
- [28] 杨俊宴, 张钟虎, 史宜. 面向城市更新的城市设计公众参与数字平台探索: 以南京阅江楼片区为例[J]. 城市规划学刊, 2024(3): 74–81.
- [29] 张凯. 信息资源管理[M]. 3版. 北京: 清华大学出版社, 2013.
- [30] SINGHAL A. Introducing the knowledge graph: things, not strings[J]. Official Google Blog, 2012, 5(16): 3.
- [31] HOGAN A, BLOMQVIST E, COCHEZ M, et al. Knowledge graphs[J]. ACM Computing Surveys (CSUR), 2021, 54(4): 1–37.
- [32] CONG Y, YANG M Y, ROSENHAHN B. Reltr: relation transformer for scene graph generation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(9): 11169–11183.
- [33] 刘宇, 李勇. 面向城市可持续发展的城市商圈/街区知识图谱构建方法与应用展望 [J]. 地球信息科学学报, 2023, 25(12): 2374–2386.
- [34] 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.