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

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

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.

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

Chinese Library Classification No.: TU984; Document Code: A

DOI: 10.16361/j.upf.202501008; Article No.: 1000-3363(2025)01-0055-09

Author Biography: NIU Xinyi, Professor and Doctoral Supervisor, College of Architecture and Urban Planning, Tongji University, and Key Laboratory of Spatial Intelligent Planning Technology, Ministry of Natural Resources. Email: niuxinyi@tongji.edu.cn; LIU Sihan, PhD Candidate, College of Architecture and Urban Planning, Tongji University; SANG Tian, Master's Student, College of Architecture and Urban Planning, Tongji University; GU Ruixing, Master's Student, College of Architecture and Urban Planning, Tongji University; WU Xuefei, Master's Student, College of Architecture and Urban Planning, Tongji University; WANG Jiang, Master's Student, College of Architecture and Urban Planning, Tongji University.

Funding: National Key Research and Development Program of China, "Key Technologies for Evaluating and Regulating the Comprehensive Performance of Territorial Space under Multiple Scenarios" (Project No. 2022YFC3800804).

With the development of generative artificial intelligence (GenAI), large models can learn language, images, audio, and other forms of information from massive datasets and generate new content. They have demonstrated powerful capabilities in natural language processing, computer vision, and related fields and are rapidly entering a wide range of disciplines[1]. The continued emergence of general-purpose large models has also attracted extensive attention in urban planning[2]. Image generation models are an important class of large models. Current general-purpose image generation models can be guided by prompts and other instructions and have already been applied directly in arts and crafts, graphic design, comics, picture books, and other design industries.

Developing domain-specific large models has become a major concern across disciplines. Professional applications involve particular task requirements, data distributions, and evaluation criteria that cannot be met simply by using a general-purpose model. Training a large model requires massive data and computing resources and therefore entails very high costs. At the same time, general-purpose large models function as foundation models: self-supervised pretraining has already endowed them with broad knowledge, providing a basis for domain applications. The prevailing approach is therefore to construct domain-specific models on top of general-purpose models. Because general image generation models have already learned extensive image-generation knowledge through large-scale pretraining, there is no need to retrain a professional model from scratch. Explorations of this route have already begun; in architecture, domain-specific training has been used to apply image generation models to design schemes and visualization production[3-4].

The planning discipline likewise requires domain-specific large models. Artificial intelligence enables machines to acquire human knowledge and thereby imitate human behavior. Because general-purpose large models have not been trained specifically on urban planning knowledge, planning-oriented models must also be constructed on top of general models. General image generation models can understand and generate images covering broad themes and styles, including natural landscapes, portraits, art and culture, and even urban architecture, yet they still lack professional planning knowledge. When directly applied to urban spatial form design, such models may fail to understand planning terminology or generate visibly abnormal urban spaces. The underlying reason is that they have not acquired urban planning knowledge. The central challenge in developing planning-specific large models is therefore how to integrate professional planning knowledge into a general-purpose model.

Urban spatial form design knowledge is a distinctive and important component of professional planning knowledge. This study seeks to develop a technical pathway for planning-specific models. It reviews the knowledge system of urban spatial form design, explores how domain-specific learning can integrate this knowledge into a general image generation model, and uses community public space renewal as an application case to construct a specialized model on top of a general-purpose image generation model.

1 Urban Spatial Form Design Knowledge and Large Models

1.1 Explicit Knowledge, Tacit Knowledge, and Large Models

Integrating professional urban spatial form design knowledge into a large model first requires a clear understanding of that knowledge. In knowledge management, knowledge is commonly divided according to whether it can be clearly articulated and effectively transferred into explicit knowledge and tacit knowledge[5]. The two represent different components of the human knowledge system.

Explicit knowledge is abstract and rational, emphasizing clarity and systematic organization. It can be fully expressed through symbolic systems such as language, mathematical formulas, and diagrams[6]. Mathematical theorems are typical examples. The Pythagorean theorem, expressed as $a^2 + b^2 = c^2$, is unambiguous and can be readily disseminated and learned through textbooks. Once understood, it can be used to solve problems involving right triangles.

Tacit knowledge is experiential knowledge grounded in personal participation. It depends primarily on individual experience and cannot be completely represented through an explicit symbolic system[6]. An artist's creative skill is a typical example. Although basic painting techniques can be shared, genuine creativity and inspiration often arise from years of practice and accumulated experience. Such experience and intuition are difficult to convey verbally but are essential to the distinctiveness of the resulting work.

Traditional artificial intelligence long struggled to address tacit knowledge[7], but large models have changed this situation. Large models are trained on vast quantities of unlabelled internet data in a process that simulates learning and communication within human communities, enabling them to acquire not only explicit but also tacit knowledge. Natural language, for example, contains extensive tacit knowledge. Through pretraining, large language models acquire large bodies of linguistic knowledge and the tacit knowledge embedded within them[8]. The successful use of general image generation models in artistic creation likewise demonstrates that large models can acquire tacit knowledge of painting through large-scale pretraining.

1.2 Tacit Knowledge in Urban Spatial Form Design

The knowledge system of urban spatial form design contains both explicit and tacit knowledge. Some knowledge can be expressed through clear formulas, such as the street enclosure ratio between street width and the height of buildings on both sides (D/H). Much more design knowledge, however, derives from experience and is difficult to articulate or formalize. For the concept of "street vitality," for example, an experienced planner can design a street space that responds to this orientation but may find it difficult to describe the underlying knowledge precisely in words or formulas. Many other concepts in spatial form design have a similar tacit character and depend on years of professional experience.

Urban planning contains large amounts of tacit knowledge, and tacit knowledge is especially dominant in urban spatial form design. This is why spatial form design knowledge is difficult to extract and express systematically through written materials, fixed rules, or formulas[9]. Planners acquire it through extensive observation, imitation, and experiential understanding. Although many studies have sought to quantify urban spatial form through shape-feature analysis[10], space syntax[11], and other methods, these approaches essentially attempt to convert tacit design knowledge into explicit knowledge. To date, however, it has not been possible to transform all spatial form design knowledge into explicit form.

1.3 The Need to Transfer Tacit Knowledge of Urban Spatial Form Design

The expression and transfer of tacit knowledge have always been difficult. Because tacit knowledge is non-verbal, communication depends on participants' knowledge backgrounds, expressive abilities, and capacity for comprehension[6]. Nevertheless, tacit knowledge in urban planning must be widely communicated and transferred because of the discipline's particular characteristics.

First, the public-policy nature of planning means that professional knowledge and information cannot circulate only among planning professionals; tacit knowledge must also be conveyed to non-specialists. Tacit knowledge in spatial form design must be communicated to the general public. Throughout the planning cycle, efficient exchange among residents, planners, and decision-makers is essential, and participatory design urgently requires the transfer of tacit knowledge. If opaque professional expressions such as "age-friendly design" or "stimulating street vitality" can be translated into accessible images and language, discussions of planning strategies will become far more efficient. Second, the teaching of tacit knowledge constitutes a substantial part of planning and design education. The planning design studio is fundamentally a setting for teaching and transferring tacit knowledge. Descriptions such as "human-scale" and "a sense of spatial oppression" are typical examples. Rapidly translating them into accessible graphic and visual expressions would greatly improve learning and communication.

Large models have demonstrated unique capabilities in learning tacit knowledge. The dissemination and transfer of tacit knowledge in planning therefore creates a distinctive demand for such models. Urban spatial form design contains far more tacit knowledge than explicit knowledge that can be conveyed through formulas. Existing studies have explored the use of image generation models in teaching urban spatial form design[12], precisely the context in which large models can play an important role.

2 Domain-specific Learning Pathways for Image Generation Models

2.1 General and Domain Knowledge Learning in Image Generation Models

The knowledge acquired by large models is generally divided into general knowledge (GK) and domain knowledge (DK)[13-14]. General knowledge is incorporated into image generation models through pretraining. Diffusion models are currently the mainstream technology for image generation[15]. By establishing cross-modal associations between text and images, diffusion models learn their relationships and acquire text-to-image generation capability[16]. Text-to-image generation by a general image model is, in essence, an expression of general knowledge.

Domain knowledge must be incorporated into a general image generation model through domain-specific learning. Fine-tuning is a widely used approach that adapts a model to new tasks and data by adjusting a small number of parameters while leaving the overall model structure unchanged[17]. Different fine-tuning techniques are suited to different kinds of domain knowledge, depending on the required degree of integration and specialization.

Existing work has followed two broad modes of domain-specific learning. The first activates general knowledge already contained in the model in a new way, enabling it to perform “general” tasks within a professional domain; this typically corresponds to prompt engineering and related techniques. Examples include asking a large model to role-play different participants in public engagement to improve planning decision-making[18], and using prompt engineering based on an urban taxonomy to map abstract urban functions onto concrete visual elements and infer functions from street-view images[19]. Role-playing and text-image correspondence are forms of general knowledge already possessed by the model. The second mode supplements the model with more systematic professional knowledge so that it can perform “expert” tasks, typically through low-rank adaptation (LoRA), retrieval-augmented generation (RAG), and similar techniques. Examples include fine-tuned models used for financial time-series prediction and financial reasoning[20]. Domain-specific learning is therefore an effective route for integrating professional knowledge into a general-purpose model.

2.2 Urban Spatial Form Design Knowledge and Domain-specific Learning for Image Generation Models

Domain-specific learning should respond to the characteristics of urban spatial form design knowledge and, in particular, to the challenge posed by tacit knowledge[21]. The knowledge system can be divided into spatial form design knowledge and spatial representation knowledge. The internal composition of urban spatial form consists of spatial elements and their relationships[22]. In planning and design, spatial form design knowledge is used to combine elements in different relationships and create a proposal, while spatial representation knowledge is used to express that proposal through drawings and images. Urban spatial form design knowledge is a typical form of domain knowledge that a general image model can acquire only through domain-specific learning.

Using the open-source image generation model Stable Diffusion① as an example, LoRA and ControlNet are two representative fine-tuning technologies. LoRA can connect general knowledge with professional knowledge through a relatively small amount of domain training and incorporate new knowledge into the model[23], making it suitable for learning spatial form design knowledge. ControlNet allows users to control particular structures and styles in generated images through images or image feature maps[24], making it suitable for learning spatial representation knowledge.

Existing fine-tuning methods provide a technical foundation for integrating urban spatial form design knowledge into image generation models, but a corresponding technical framework must still be established. The following sections use community public space renewal as an application case to construct such a framework and realize the domain-specific learning of a general-purpose model.

3 Practice of Domain-specific Learning: A Community Public Space Renewal Model

3.1 Technical Framework

The renewal of existing community public space is an important component of contemporary urban regeneration[25-26]. Participatory design is increasingly encouraged in community renewal, with extensive resident involvement emphasized in communicating renewal intentions, refining renewal objects, and comparing alternative concepts[27]. The conventional approach relies on professional drawings such as plans, sections, and renderings. Because such drawings require specialist knowledge to read, communication barriers often arise between planners and residents. Residents may struggle to understand professional terminology such as design orientations or to recognize design elements, and may therefore be unable to establish a shared imagination of intended and actual outcomes during consultation[28]. Participatory community renewal is thus a typical context in which tacit spatial form design knowledge must be communicated.

This study constructs CommunityGenAI on top of a general-purpose large model to support participatory design in community renewal. The goal is to input several professional prompts, partially redraw photographs of community public space, and intelligently generate renewed scenes. The study first reviews urban spatial form design knowledge, maps professional knowledge to large-model fine-tuning technologies, and proposes a technical framework for integrating this knowledge into an image generation model. Within the framework, professional knowledge is treated as domain knowledge and divided into spatial form design knowledge and spatial representation knowledge. Spatial form design knowledge is dominated by tacit knowledge and is expressed as a design-orientation-centered knowledge graph that guides the construction of training datasets containing both explicit and tacit knowledge. LoRA fine-tuning then realizes domain-specific learning. Spatial representation knowledge is explicit and is learned through ControlNet image-condition control to guide generation (Fig.1). This framework integrates domain knowledge with general knowledge.

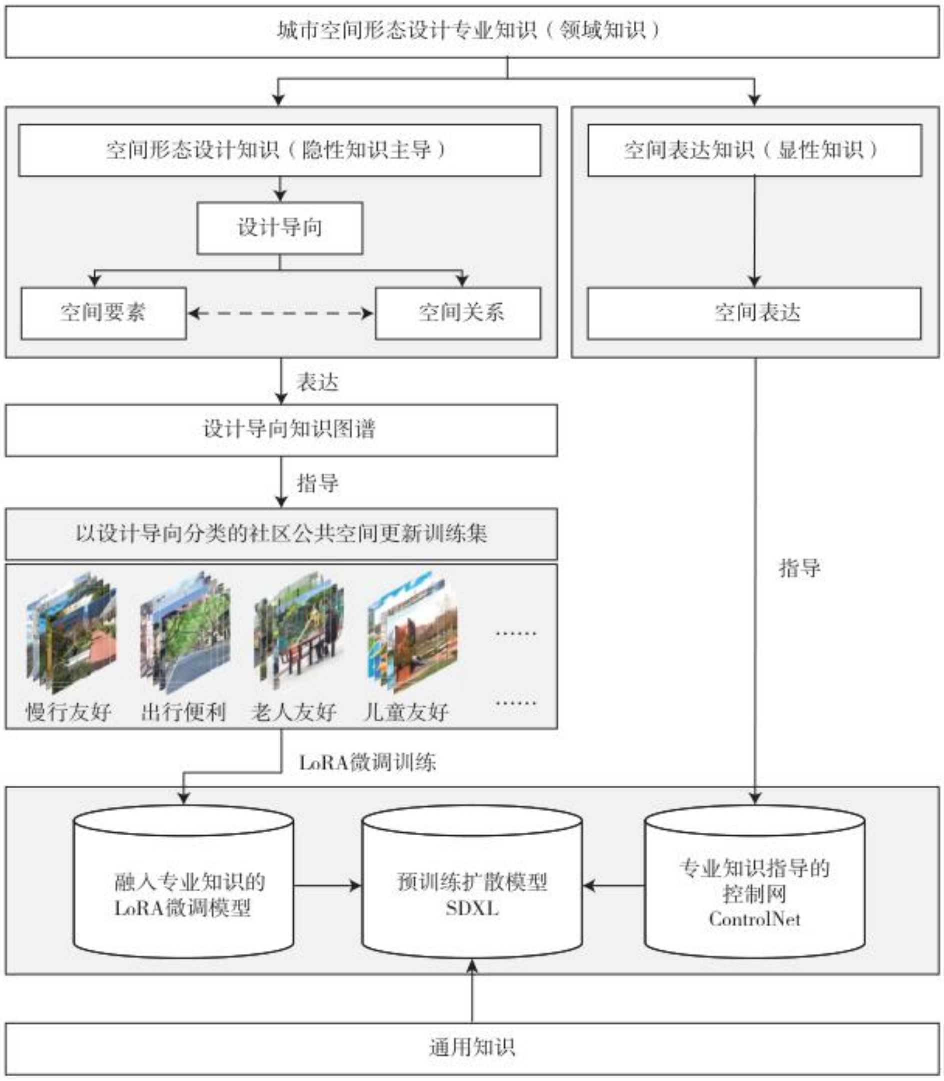

图1 技术框架
Fig.1 Technical framework

Fig.1 Technical framework

3.2 Knowledge Graph of Spatial Form Design

Community public space renewal contains extensive tacit knowledge. Graph databases such as semantic networks and knowledge maps can represent tacit knowledge more clearly[29]. Building on semantic networks, Google introduced the concept of the knowledge graph[30]. A knowledge graph structurally organizes concepts, objects, and their semantic associations as a directed graph. It is a semantic network that computers can understand and use, with nodes representing entities and edges representing semantic relations[31]. Knowledge graphs have already been used to analyze spatial elements and relationships in urban form[22,32]. This study therefore uses a knowledge graph to organize and express spatial form design knowledge.

Spatial form design knowledge in community public space renewal is divided into three categories: design orientations, spatial elements, and spatial relationships. Design orientations express renewal objectives. Spatial elements include buildings, roads, fitness facilities, and greenery. Spatial relationships describe relations among elements, including adjacency, containment, and spatial scale. A site may contain several fitness facilities; facilities may be adjacent to roads; and their dimensions may relate proportionally to the human body. Under a design orientation, spatial elements and relationships are organized into a community public space. An “age-friendly” orientation, for example, organizes public space around the needs of older residents through targeted services, paving, and leisure installations. A “sports and recreation” orientation promotes fitness and public health through sports facilities and activity grounds.

Spatial elements and their attributes are explicit knowledge and can be expressed clearly in words. Design orientations and spatial relationships are typical forms of tacit knowledge and are difficult to articulate precisely. Knowledge graphs can effectively organize and manage these kinds of urban spatial form knowledge[33]. Figure 2 presents part of the knowledge graph for a “child-friendly” orientation. Nodes represent spatial elements and edges represent spatial relations. The purpose is to enable the model to learn spatial form design knowledge in images in a targeted manner.

3.3 Domain-specific Learning of Spatial Form Design Knowledge

Using design orientations as the core, the study organizes spatial form design knowledge for community renewal in the form of a knowledge graph. The graph guides the selection of training images, and LoRA fine-tuning integrates explicit and tacit knowledge into the image generation model. CommunityGenAI establishes eight design orientations for combining spatial elements: convenient mobility, pedestrian and cycling friendliness, arts and creativity, sports and recreation, age-friendliness, child-friendliness, commercial atmosphere, and natural ecology.

With the design-orientation knowledge graph as the organizing framework, domain-specific learning aims to enable the model to understand each orientation and the spatial elements and relationships it contains. In a child-friendly community activity space, for example, children, play equipment, rubber tracks, and protective devices are explicit elements, but their combination as a child-friendly orientation constitutes tacit knowledge (Fig.2). Training images selected according to the knowledge graph guide the model in associating concrete elements with the orientation, thereby incorporating its tacit meaning.

The selected training images also contain tacit knowledge of spatial relationships. The model can learn containment, adjacency, and scale relationships such as “play equipment is located within a rubber-surfaced site” and “protective equipment is adjacent to the rubber-surfaced site” (Fig.2). These relationships are transferred to the model through the training set, thereby integrating tacit spatial-relation knowledge under the child-friendly orientation.

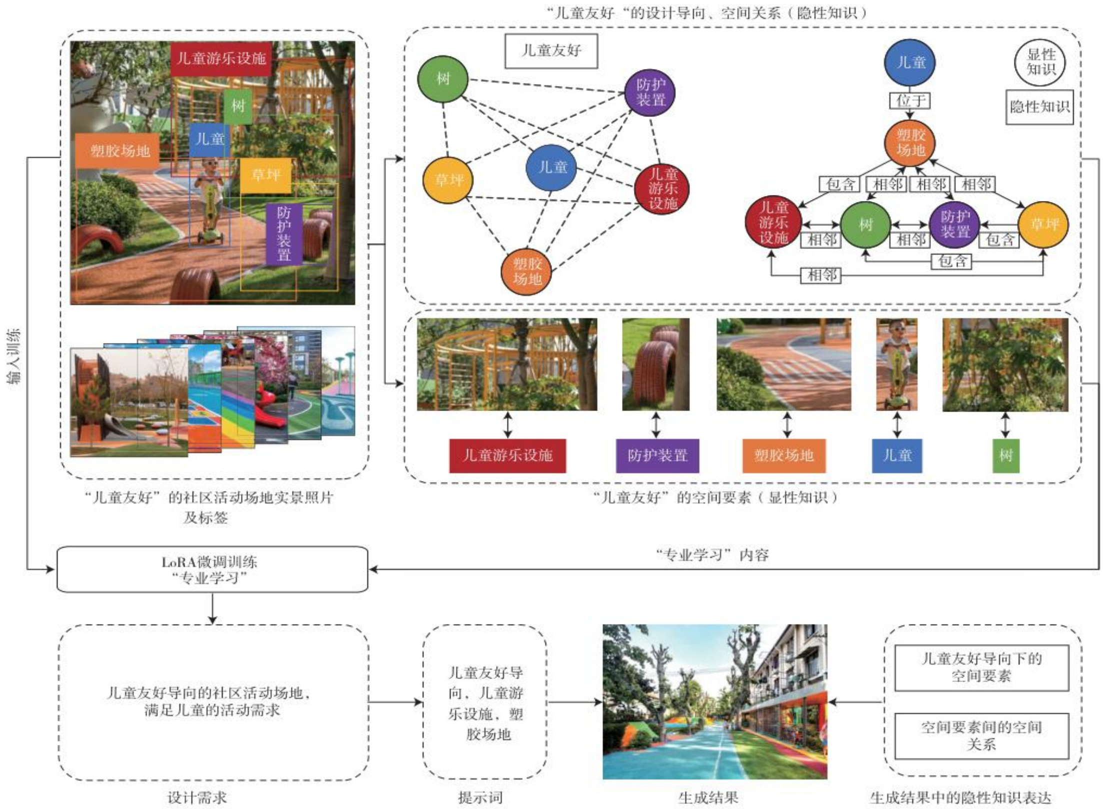

图2 空间形态设计知识“专业学习”和表达的路径

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

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

3.4 Domain-specific Learning of Spatial Representation Knowledge

Urban spatial form design addresses three-dimensional urban space but must be expressed through two-dimensional drawings. This conversion is itself an important body of design knowledge. Professional planners study descriptive geometry and related courses in order to learn spatial representation and translate three-dimensional space into two-dimensional images. Spatial representation knowledge is explicit knowledge and can be expressed through clear formulas and geometric methods.

Using an image generation model to visualize community public space renewal involves intelligent local redrawing. The resulting scene must maintain consistent spatial representation, which requires domain-specific learning of spatial representation knowledge.

This knowledge is divided into perspective, occlusion, and relative-size relationships. In the community street example, the photograph in Fig.3 first expresses one-point perspective. When three-dimensional elements are projected onto a two-dimensional plane, foreground objects occlude objects behind them. Relative dimensions in real three-dimensional space are likewise represented through projection, as illustrated by the relative heights of pedestrians and electric bicycles in Fig.3. These relationships can all be expressed through theorems of descriptive geometry and incorporated into a large model through different forms of ControlNet image-condition fine-tuning.

In Stable Diffusion, ControlNet extensions can transform semantic maps, line drawings, and other representations into image conditions and are therefore suitable for expressing and transferring spatial representation knowledge. This study uses ControlNet image-condition fine-tuning as the domain-specific learning pathway and establishes a correspondence between existing ControlNet controls and spatial representation knowledge. Perspective uses edge-detection and control-oriented ControlNet inputs, while occlusion and relative-size relationships use Segmentation and Tile controls (Fig.3).

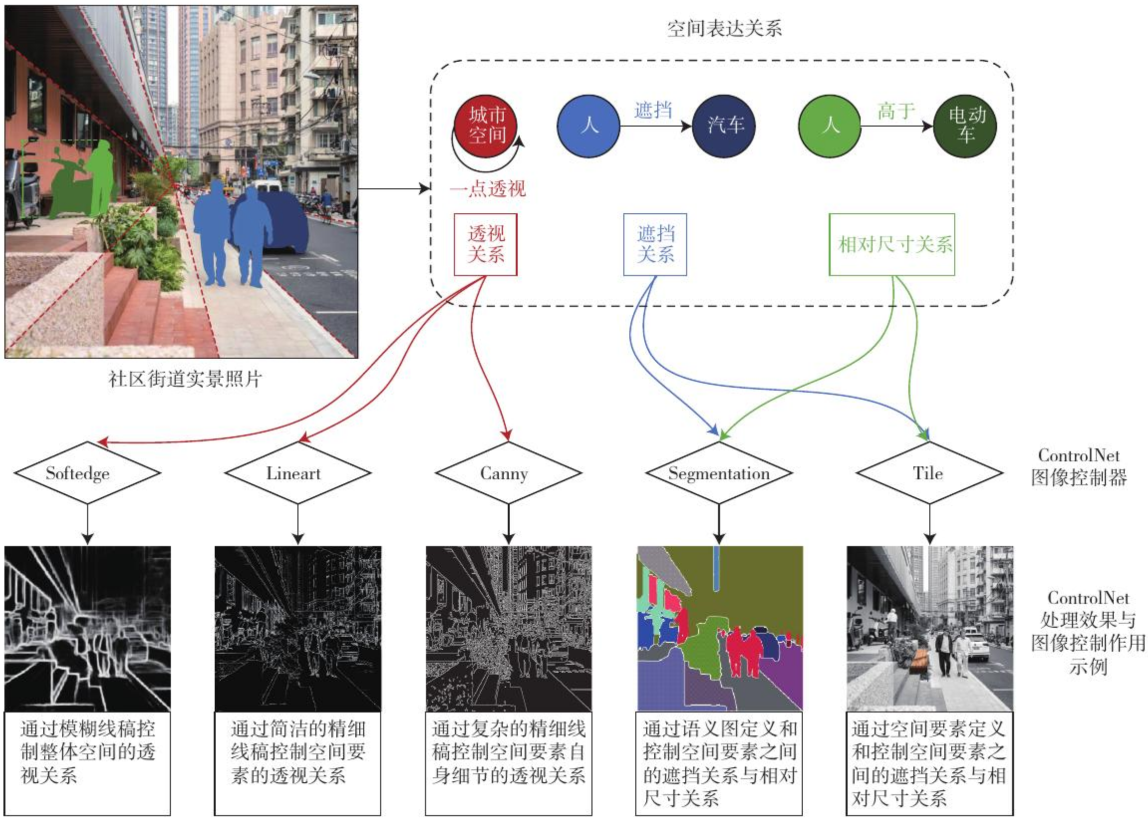

Fig.3 Pathway of "domain-specific learning" and expressing spatial expression knowledge

4 Construction and Application of the CommunityGenAI Community Public Space Renewal Model

4.1 Label Set Construction and Data Preparation

The training set for CommunityGenAI consists of labels and images, both constructed with reference to the design-orientation knowledge graph. Labels express knowledge of design orientations and spatial elements. During model application, they serve directly as prompts that provide textual control over image generation. The label set is therefore constructed by matching the orientations and their corresponding elements in the knowledge graph.

Images are the objects of domain-specific learning and provide an intuitive expression of urban spatial form design knowledge. Guided by the design-orientation knowledge graph, planning professionals selected photographs of community scenes that fully embodied the orientation, clearly represented spatial elements, and demonstrated high spatial quality. Under the eight orientation graphs, 436 photographs were collected from multiple types and periods of residential communities, enabling the model to adapt to diverse renewal needs.

The photographs were classified by design orientation and annotated against the subordinate spatial elements. Each label and the spatial element it represented appeared at least five times in the training set, producing the final label file.

4.2 Integrating Urban Spatial Form Design Knowledge through LoRA Fine-tuning

LoRA fine-tuning incorporates design orientations, spatial elements, and spatial relationships into the large model. Spatial elements are explicit knowledge and can be directly annotated in the label set. Design orientations and spatial relationships are tacit knowledge and are expressed through the real-scene training images. Because LoRA can incorporate new knowledge with a relatively small training set, it enables spatial form design knowledge to be integrated efficiently.

LoRA training used the Stable Diffusion XL “base 1.0” model. The software and hardware environment and parameter settings are shown in Tabs.1 and 2.

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

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

After 100 training epochs, the loss function stabilized at approximately 0.08 and the model achieved its best training performance. An X/Y/Z plot was used to test the LoRA fine-tuned model under different weights. The optimal LoRA weight was determined to be 0.4, and the model from the final training epoch was selected as the final LoRA fine-tuned result.

4.3 Representation of Spatial Form Design Knowledge in CommunityGenAI

Through domain-specific learning, CommunityGenAI incorporates professional knowledge of design orientations, spatial elements, spatial relationships, and spatial representation. In community renewal, users can select an image of an existing public space and input prompts for the desired orientation and key elements; the model then intelligently generates a locally renewed scene.

For a child-friendly community activity space, prompts such as “child-friendly, child-friendly colors, children’s play facilities, sandpit” produce an image consistent with the intended orientation (Fig.4). The generated concept contains play equipment and a sandpit, adopts child-friendly colors, and also includes children, accompanying parents, and a rubber-surfaced site. The prompts do not include “parents” or “rubber-surfaced site.” These are forms of tacit knowledge embedded in the child-friendly orientation. The relative positions and dimensions of play equipment, trees, sandpits, and other elements are also tacit knowledge acquired through domain-specific learning. CommunityGenAI has thus incorporated spatial form design knowledge for community renewal, especially its tacit components.

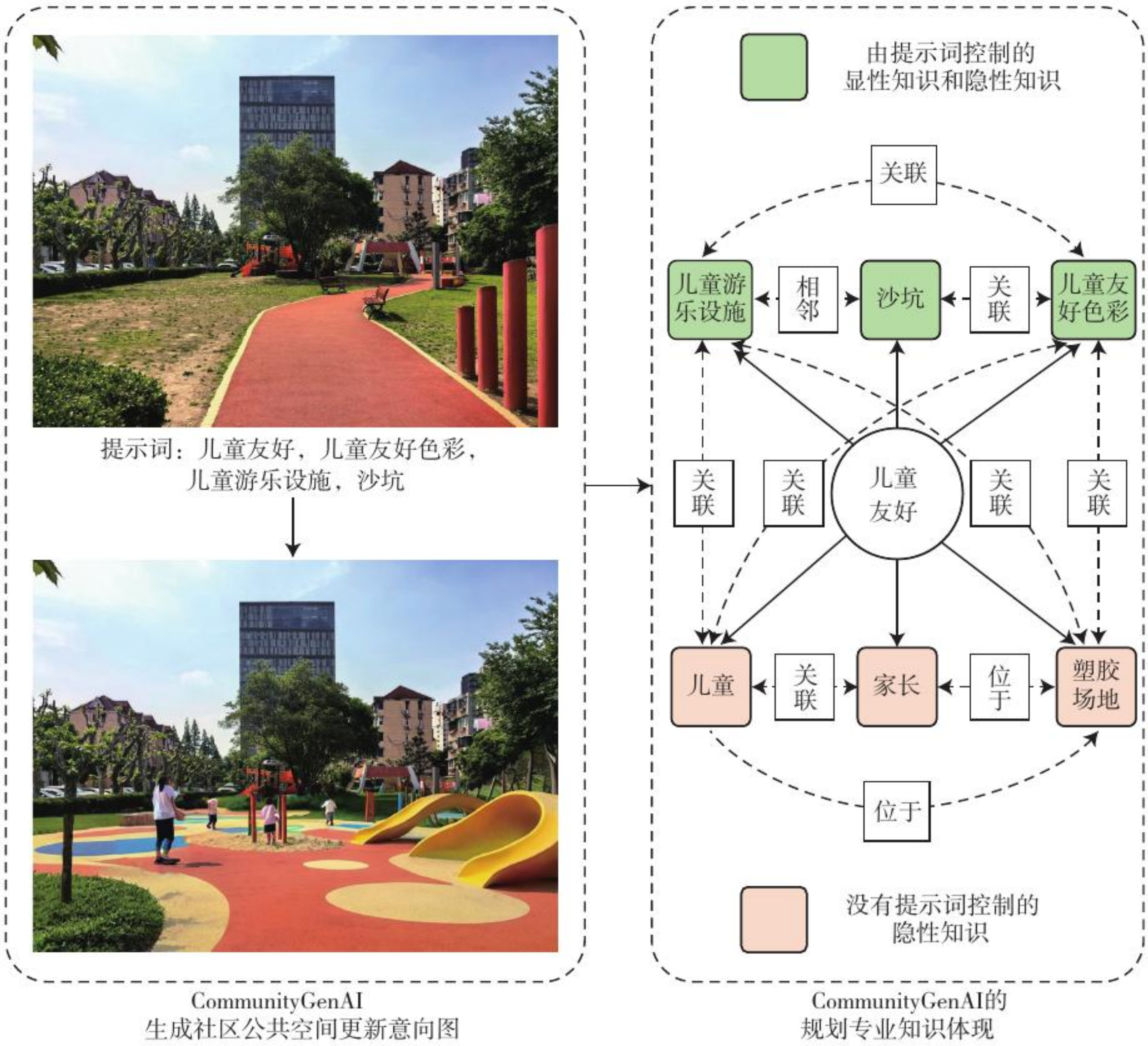

Fig.4 Representation of domain knowledges in CommunityGenAI Model

Because the model was trained by design orientation, it can generate contrasting renewal concepts for the same site and allow residents to compare alternatives. Figure 5 uses an internal road between residential buildings as the renewal object and presents three concepts generated for “commercial atmosphere,” “age-friendly,” and “convenient mobility.” Tables and chairs and flowerbeds under trees in the commercial-atmosphere concept differ visibly from those in the age-friendly concept, while the tree-bed treatment under convenient mobility differs again (Fig.5). The spatial elements in the three concepts are clearly differentiated and each reflects typical features of its orientation. Their scales also satisfy the spatial-representation requirements of a community public-space concept image. This confirms that CommunityGenAI acquired the tacit knowledge embedded in particular design orientations through domain-specific learning.

CommunityGenAI has effectively learned design orientations, spatial elements, spatial relationships, and spatial representation for community public space renewal. It has acquired both explicit and tacit knowledge and therefore possesses the capabilities of a domain-specific large model for community public space renewal.

Fig.5 CommunityGenAI Model intelligently generates concept maps for three design orientations

4.4 Professional Knowledge Transfer in CommunityGenAI

An image generation model incorporating urban spatial form design knowledge can express and transfer tacit knowledge. It can support communication between residents and planners both during early consultation and during the review and feedback of design alternatives. In the renewal of a hard-surfaced activity area in an older residential community, for example, the model helps planners and residents confirm the object and scope of renewal during initial consultation. When refining design orientations and comparing alternatives, planners can rapidly translate residents' needs into professional terms such as "child-friendly orientation" and "children's play facilities," together with corresponding control networks. The model can then generate a series of compliant renewal concepts at an average rate of approximately 40 seconds per image. This process rapidly links residents' needs to professional planning terminology and presents difficult-to-articulate tacit knowledge through accessible visual feedback. CommunityGenAI thus transfers professional community-renewal knowledge efficiently between planners and residents, especially by expressing and communicating tacit knowledge. It supports the formation of a shared renewal intention and, in turn, participatory design.

5 Conclusions and Discussion

5.1 Conclusions

This study treats professional urban spatial form design knowledge as domain knowledge and introduces the concepts of tacit and explicit knowledge from knowledge management. By classifying professional knowledge and matching it to large-model fine-tuning technologies, the study proposes two domain-specific learning pathways and constructs a technical framework for integrating urban spatial form design knowledge into a general image generation model. CommunityGenAI is developed as a specialized model for community public space renewal. Three main conclusions are drawn.

First, the planning discipline contains extensive tacit knowledge, particularly in urban spatial form design, where it is dominant. Large models are needed to communicate and transfer this tacit knowledge. Its transfer is a distinctive planning requirement for large models, and the central difficulty in developing a spatial form design model lies in integrating professional knowledge into a general-purpose model.

Second, the study proposes a technical framework for integrating urban spatial form design knowledge through domain-specific learning. Professional knowledge is divided into spatial form design knowledge and spatial representation knowledge. The former is organized as a knowledge graph centered on design orientations and contains both tacit and explicit knowledge; the latter guides image expression as explicit knowledge. Two forms of fine-tuning integrate these knowledge types into a general image generation model.

Third, using Stable Diffusion as the general image generation foundation and large-model fine-tuning, the study integrates professional community public-space renewal knowledge and constructs CommunityGenAI to support participatory design. The trained model learns design-orientation knowledge, spatial-element knowledge, spatial-relationship knowledge, and spatial representation knowledge. In participatory-design scenarios, it transfers professional knowledge, plays a key role in efficiently communicating tacit knowledge, improves communication in community renewal, and effectively supports participatory design.

5.2 Discussion

In planning, general-purpose large language models can already simulate stakeholders in planning and design processes and thereby support simulated public participation[34]. This approach treats the model as a member of the public equipped only with general knowledge. The present study instead constructs a planning-specific model through domain-specific learning, turning it into an expert that possesses both general and domain knowledge. Both approaches are valid pathways for applying large-model technology in planning and merit further exploration.

Domain-specific learning can also be used to construct other kinds of planning-specific models. How professional planning knowledge is integrated into general-purpose models will become a key technical issue. From a knowledge-management perspective, the knowledge structures of different planning fields vary. Learning pathways differ not only between explicit and tacit knowledge, but also among different kinds of tacit knowledge. Methods for incorporating different planning knowledge domains into general-purpose models will therefore differ. Constructing planning-specific large models by field and category may be an appropriate strategy and deserves systematic investigation.

The widespread interest in domain-specific large models across disciplines is driven broadly by two considerations. In scientific research, large models can perform tasks that are difficult for humans, such as processing and analyzing massive datasets, thereby improving research efficiency. In engineering applications, they can increase design efficiency, optimize production processes, reduce costs, and improve productivity. Urban planning has these needs but also possesses a distinctive one:

because planning is a form of public policy, its tacit knowledge must be disseminated and communicated. How large models can learn and master the tacit knowledge of planning and apply it throughout the planning cycle is therefore a unique requirement for planning-specific large models and deserves further research.

Notes

① Stable Diffusion is a general-purpose image model released by Stability AI. Stable Diffusion XL 1.0 is an open-source version released in July 2023; its base model contains 3.5 billion parameters.

References

[1] WANG H, FU T, DU Y, et al. Scientific discovery in the age of artificial intelligence [J]. Nature, 2023, 620: 47-60.

[2] HERZOG O, PAN Haixiao, DENG Zhituan, et al. Next-generation artificial intelligence empowering urban planning: Opportunities and challenges [J]. Urban Planning Forum, 2023(4): 1-11. (in Chinese)

[3] YUAN Chao, ZHENG Hao. New modes of architectural design under the influence of generative artificial intelligence [J]. Architectural Journal, 2023(10): 29-35. (in Chinese)

[4] HE Wanyu, YANG Liangsong. Exploring generative artificial intelligence in architectural design: The case of Xkool AI Cloud [J]. Architectural Journal, 2023(10): 36-41. (in Chinese)

[5] POLANYI M. The tacit dimension [M]. Routledge and Kegan Paul, 1966.

[6] YAO Wei. Knowledge Management [M]. Beijing: Tsinghua University Press, 2020. (in Chinese)

[7] SANZOGNI L, GUZMAN G, BUSCH P. Artificial intelligence and knowledge management: questioning the tacit dimension [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] NIU Xinyi, LIN Shijia, SANG Tian, et al. Digital planning technologies: Data and knowledge [J]. Urban Planning Forum, 2024(2): 18-24. (in Chinese)

[10] LIN Bingyao. Quantitative methods for urban spatial form and their evaluation [J]. Urban Planning Forum, 1998(3): 42-45. (in Chinese)

[11] HILLIER B. The hidden geometry of deformed grids: or, why space syntax works, when it looks as though it should not [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 pretrained 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] GAN Wei, WU Zhiqiang, WANG Yuankai, et al. Constructing a theoretical model for AIGC-assisted urban design [J]. Urban Planning Forum, 2023(2): 12-18. (in Chinese)

[22] YANG Junyan, SHAO Dian, WANG Peng, et al. Integration, topology, and translation: A knowledge-graph-based method for deep analysis of urban form [J]. City Planning Review, 2023, 47(6): 57-67. (in Chinese)

[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] WANG Jia, BAI Yunxi, SONG Jusheng. Evolution, challenges, and recommendations for urban renewal in China [J]. Planners, 2021, 37(24): 21-27. (in Chinese)

[26] HUA Xiahong, ZHUANG Shen. Promoting the renewal of everyday public space through design: A review of micro-renewal practices in Shanghai [J]. Architectural Journal, 2022(3): 1-11. (in Chinese)

[27] LI Qing, LIN Ni. A new participatory planning and design model for community micro-renewal from the perspective of the people-centered city: The YF lilong micro-renewal in Shanghai [J]. Urban Planning Forum, 2023(6): 87-96. (in Chinese)

[28] YANG Junyan, ZHANG Zhonghu, SHI Yi. Exploring a digital platform for public participation in urban design for urban renewal: The Yuejianglou area in Nanjing [J]. Urban Planning Forum, 2024(3): 74-81. (in Chinese)

[29] ZHANG Kai. Information Resource Management [M]. 3rd ed. Beijing: Tsinghua University Press, 2013. (in Chinese)

[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, 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] LIU Yu, LI Yong. Methods and prospects for constructing knowledge graphs of urban commercial districts and blocks for sustainable urban development [J]. Journal of Geo-information Science, 2023, 25(12): 2374-2386. (in Chinese)

[34] TIAN Li, YANG Xin, ZHANG Yudi, et al. Innovation in urban-rural planning and design education driven by "professional knowledge + artificial intelligence": The case of residential-area planning [J]. Urban Planning Forum, 2024(5): 71-78. (in Chinese)

Revised: January 2025