

Профессиональное обучение больших моделей: интеграция знаний о проектировании городской пространственной формы в модель генерации изображений

НИУ Синьи, ЛЮ Сыхань, САН Тянь, ГУ Жуйсин, У Сюэфэй, ВАН Цзян

Аннотация: Основная трудность построения профессиональной большой модели на базе универсальной большой модели состоит в том, как встроить знания предметной области в универсальную модель. В проектировании городской пространственной формы существует большой объем неявных знаний, а публично-политическая природа городского планирования требует, чтобы такие знания эффективно передавались широкой общественности. Опираясь на особую способность больших моделей обучаться неявным знаниям и передавать их, статья предлагает техническую рамку, позволяющую посредством профессионального обучения больших моделей интегрировать знания о проектировании городской пространственной формы в универсальную модель генерации изображений. Взяв проектную ориентацию в качестве ядра, исследование систематизирует неявные и явные знания в проектировании пространственной формы, чтобы направлять профессиональное обучение большой модели, и посредством тонкой настройки реализует включение этих знаний в универсальную модель генерации изображений. На примере обновления общественных пространств сообщества построена профессиональная большая модель, включающая знания о проектировании городской пространственной формы. Практическое применение показывает, что такая профессиональная модель играет ключевую роль в эффективной передаче неявных знаний, повышает эффективность коммуникации при обновлении сообществ и действительно поддерживает партисипаторное проектирование.

Ключевые слова: профессиональная большая модель для городского планирования; знания о проектировании пространственной формы; неявные знания; модель генерации изображений; тонкая настройка большой модели; обновление общественных пространств сообщества

Классификационный номер Китайской библиотеки: TU984; код документа: A

DOI: 10.16361/j.urpf.202501008; номер статьи: 1000-3363(2025)01-0055-09

Сведения об авторах: НИУ Синьи - профессор и научный руководитель докторантов Колледжа архитектуры и городского планирования Университета Тунцзи, а также Ключевой лаборатории технологий интеллектуального пространственного планирования территорий Министерства природных ресурсов. E-mail: niuxinyi@tongji.edu.cn; ЛЮ Сыхань - докторант Колледжа архитектуры и городского планирования Университета Тунцзи; САН Тянь - магистрант Колледжа архитектуры и городского планирования Университета Тунцзи; ГУ Жуйсин - магистрант Колледжа архитектуры и городского планирования Университета Тунцзи; У Сюэфэй - магистрант Колледжа архитектуры и городского планирования Университета Тунцзи; ВАН Цзян - магистрант Колледжа архитектуры и городского планирования Университета Тунцзи.

Финансирование: проект Национальной ключевой программы исследований и разработок «Ключевые технологии оценки и регулирования комплексной эффективности территориального пространства в различных сценариях» (№ 2022YFC3800804).

С развитием генеративного искусственного интеллекта (GenAI) большие модели (Large Models) способны обучаться на массивных данных языку, изображениям, аудио и другим типам информации и генерировать новый контент. Они уже продемонстрировали мощные возможности в обработке естественного языка, компьютерном зрении и других областях и

Профессиональное обучение больших моделей: интеграция знаний о проектировании городской пространственной формы в модель генерации изображений быстро входят в разные дисциплины^[1]. Постоянное появление универсальных больших моделей вызвало широкий интерес и в планировочной науке^[2]. Модели генерации изображений являются сегодня важным типом больших моделей. Универсальные модели генерации изображений уже могут направляться инструкциями-подсказками (prompt) и напрямую применяются в прикладном искусстве, графическом дизайне, комиксах, иллюстрированных книгах и других дизайнерских сферах.

Создание профессиональных больших моделей стало темой, которую исследуют многие дисциплины. Профессиональные приложения имеют особые задачи, распределения данных и критерии оценки, поэтому прямое использование универсальной большой модели не может удовлетворить эти требования. С одной стороны, обучение больших моделей требует огромных массивов данных и вычислительных ресурсов, что ведет к крайне высоким затратам. С другой стороны, универсальная большая модель уже обладает свойствами базовой модели: в процессе предварительного самообучения она осваивает общие знания, что создает основу для профессиональных приложений. Поэтому в настоящее время профессиональные большие модели обычно строятся на базе универсальных. Универсальные модели генерации изображений уже в ходе масштабного предварительного обучения освоили многочисленные знания о генерации изображений; нет необходимости создавать профессиональную модель изображения с нуля. Построение профессиональной большой модели на основе универсальной модели генерации изображений уже имеет практические примеры: в архитектурном проектировании, например, изучается применение специально обученных моделей для генерации проектных предложений и визуализаций^[3-4].

Создание профессиональной большой модели для планирования также является актуальной потребностью дисциплины. Искусственный интеллект призван позволить машине овладеть человеческими знаниями и тем самым имитировать человеческое поведение. Поскольку универсальные большие модели не обучались специально на знаниях городского планирования, профессиональная большая модель для планирования также должна строиться на их основе. Если продолжить пример моделей генерации изображений, они способны понимать и создавать изображения широкого спектра тем и стилей - природные ландшафты, портреты, искусство и культуру, даже городскую архитектуру, - однако им по-прежнему недостает профессиональных знаний в области городского планирования. При прямом применении универсальной модели генерации изображений к проектированию городской пространственной формы она либо не понимает профессиональные термины планирования и не может сформировать соответствующее изображение, либо генерирует явно аномальные городские пространства. Главная причина состоит в том, что универсальная большая модель еще не обладает профессиональными знаниями городского планирования. Ключевая трудность создания профессиональной большой модели для планирования заключается в том, как встроить эти знания в универсальную модель.

Знания о проектировании городской пространственной формы являются важной частью профессиональных знаний городского планирования и специфическим знанием планировочной дисциплины. Цель этой статьи - исследовать технический путь построения профессиональной модели для планирования. Через систематизацию системы знаний о проектировании городской пространственной формы рассматривается способ их интеграции в универсальную модель генерации изображений посредством профессионального обучения. В качестве приложения выбран пример обновления общественных пространств сообщества, на основе которого

Профессиональное обучение больших моделей: интеграция знаний о проектировании городской пространственной формы в модель генерации изображений

создается профессиональная большая модель обновления общественных пространств сообщества на базе универсальной модели генерации изображений.

1 Знания о проектировании городской пространственной формы и большие модели

1.1 Явные знания, неявные знания и большие модели

Чтобы интегрировать профессиональные знания о проектировании городской пространственной формы в большую модель, прежде всего необходимо ясно понимать саму природу этих знаний. В управлении знаниями в менеджменте, в зависимости от того, можно ли знание четко выразить и эффективно передать, различают явные знания (explicit knowledge) и неявные знания (tacit knowledge)^[5]. Эти два типа представляют разные части человеческой системы знаний.

Явное знание - это рациональное абстрактное знание, подчеркивающее определенность и системность. Оно может быть полностью выражено с помощью определенной системы знаков, например языка, математических формул или диаграмм^[6]. Типичным примером являются математические теоремы: теорема Пифагора в форме $a^2 + b^2 = c^2$ сформулирована ясно и легко передается и изучается через учебники. Достаточно понять эту формулу, чтобы применять ее к задачам, связанным с прямоугольными треугольниками.

Неявное знание, напротив, является опытным знанием, основанным на личном участии и главным образом зависящим от индивидуального опыта. Его невозможно полностью выразить с помощью четкой системы знаков^[6]. Типичным примером служат творческие навыки художника. Хотя художник может делиться базовыми приемами рисунка, подлинная художественная креативность и вдохновение обычно происходят из многолетней практики и накопленного опыта. Такие ощущения и опыт трудно передать языком, но именно они определяют уникальность произведения.

Традиционные технологии искусственного интеллекта долгое время с трудом справлялись с неявными знаниями^[7], пока появление больших моделей не изменило ситуацию. Обучение больших моделей основано на огромных объемах размеченных интернет-данных; в процессе обучения моделируется среда человеческого обучения и коммуникации в сообществе, поэтому модель способна усваивать не только явные, но и неявные знания. Например, естественный язык содержит большое количество неявных знаний. Большие языковые модели благодаря предварительному обучению осваивают обширные языковые знания и содержащиеся в них неявные знания^[8]. То, что универсальные модели генерации изображений уже используются в живописи и художественном творчестве, показывает, что большие модели также могут посредством масштабного предварительного обучения овладевать неявными знаниями, присущими изобразительному искусству.

1.2 Неявные знания в проектировании городской пространственной формы

Система знаний о проектировании городской пространственной формы включает как явные, так и неявные знания. Некоторые знания о городской пространственной форме можно выразить точными формулами, например отношением ширины улицы к высоте зданий по ее сторонам - отношением D/H. Однако большая часть знаний о проектировании пространственной формы происходит из опыта и плохо поддается явному описанию или записи. Так, в отношении понятия «уличная жизненность» опытный планировщик может спроектировать городское уличное пространство, соответствующее этой ориентации, но ему трудно точно описать это знание словами или формулами. В проектировании городской пространственной формы много

Профессиональное обучение больших моделей: интеграция знаний о проектировании городской пространственной формы в модель генерации изображений подобных неявных знаний, зависящих от многолетнего профессионального опыта планировщиков.

В профессиональных знаниях городского планирования существует большой объем неявных знаний; особенно в проектировании городской пространственной формы они занимают доминирующее положение. Именно поэтому такие знания трудно систематически извлекать и выражать через текстовые материалы, фиксированные правила или формулы^[9]. Осваивая проектные знания, планировщики должны много наблюдать, подражать и через практику постигать содержащиеся в них неявные знания. На протяжении длительного времени различные исследования изучали количественные методы анализа городской пространственной формы, включая методы анализа формальных характеристик^[10], пространственный синтаксис^[11] и другие. По сути, все они стремились перевести неявные знания проектирования городской пространственной формы в явные знания. Однако до сих пор невозможно преобразовать все знания проектирования пространственной формы в явные.

1.3 Необходимость передачи неявных знаний о проектировании городской пространственной формы

Выражение и передача неявных знаний всегда были трудной задачей в обмене знаниями. Поскольку неявное знание имеет невербальную форму выражения, на его коммуникацию влияют база знаний участников, их коммуникативные навыки и способность к пониманию^[6]. Хотя неявные знания трудно четко выражать и передавать, в городском планировании они должны широко распространяться и передаваться. Это определяется особенностями самой дисциплины.

Во-первых, публично-политическая природа планирования означает, что передача профессиональных знаний и информации не может происходить только между специалистами: неявные знания планировочной дисциплины должны передаваться неспециалистам. Неявные знания в проектировании пространственной формы также необходимо передавать и обсуждать с широкой общественностью. Особенно в процессе планирования полного цикла требуется эффективная коммуникация между жителями, планировщиками и лицами, принимающими решения; такое партисипаторное проектирование остро нуждается в передаче неявных знаний. Если сложные выражения вроде «дизайн, дружественный пожилым» или «стимулирование уличной жизненности» можно перевести в понятные изображения и язык, обсуждение планировочных стратегий станет гораздо эффективнее. Во-вторых, обучение неявным знаниям занимает значительную часть образования в области планирования и дизайна. «Проектная студия» в планировочном образовании фактически и есть преподавание и передача неявных знаний. Описания городского пространства вроде «приятный масштаб» или «ощущение пространственного давления» являются типичными примерами передачи неявных знаний. Если их можно быстро преобразовать в доступные графические и визуальные формы, эффективность передачи знаний существенно повысится.

Большие модели уже продемонстрировали особые способности к обучению и освоению неявных знаний. Распространение и передача неявных знаний в планировочной дисциплине настоятельно нуждаются в больших моделях; это специфическая потребность планирования. В проектировании городской пространственной формы неявных знаний значительно больше, чем явных знаний, передаваемых формулами. Уже существуют исследования применения моделей генерации изображений в обучении проектированию городской пространственной формы^[12], и именно такие ситуации являются теми областями, где большие модели могут сыграть важную роль.

2 Пути профессионального обучения моделей генерации изображений

2.1 Обучение универсальным знаниям и знаниям предметной области в моделях генерации изображений

Знания, которыми овладевают большие модели, обычно делят на две категории: универсальные знания (general knowledge, GK) и знания предметной области (domain knowledge, DK)^[13-14]. Универсальные знания вводятся в модель генерации изображений посредством предварительного обучения. В настоящее время диффузионные модели (diffusion model) являются основной технологией генерации изображений^[15]. Благодаря межмодальной связи текста и изображения диффузионные модели понимают отношения между ними и формируют модели, способные генерировать изображения по тексту^[16]. Генерация «текст - изображение» универсальной моделью по сути является процессом выражения универсальных знаний.

Знания предметной области необходимо вводить в универсальную модель генерации изображений посредством профессионального обучения. Тонкая настройка (fine-tuning) является широко используемым способом профессионального обучения моделей генерации изображений; она позволяет адаптировать модель к новым задачам и данным, изменяя небольшое число параметров без изменения структуры модели^[17]. Различные виды тонкой настройки подходят для разных типов профессиональных знаний в зависимости от степени их интеграции и специализации.

В существующих исследованиях профессиональное обучение больших моделей через тонкую настройку имеет два режима. Первый заключается в том, чтобы новым способом активировать уже имеющиеся у модели универсальные знания, позволяя ей выполнять «общие» задачи в профессиональной области; обычно это соответствует таким техникам, как инженерия подсказок (prompt engineering). Например, большая модель может играть разные роли в общественном участии, чтобы повышать эффективность планировочных решений^[18]; или через инженерную настройку подсказок на основе городской классификации абстрактные городские функции сопоставляются с конкретными визуальными элементами, что позволяет извлекать функции города из уличных изображений^[19]. Ролевое моделирование общественности и текстово-визуальное соответствие городских элементов относятся к универсальным знаниям, уже освоенным моделью. Второй режим состоит в том, чтобы дополнительно обучить модель более системным профессиональным знаниям, позволяя ей выполнять «экспертные» задачи; обычно он соответствует таким техникам, как низкоранговая адаптация (low-rank adaptation, LoRA) и генерация с дополнением через поиск (retrieval augmented generation, RAG). Например, тонко настроенные большие модели используются для прогнозирования финансовых временных рядов и финансовых рассуждений^[20]. Профессиональное обучение является эффективным способом включения знаний конкретной области в универсальную большую модель.

2.2 Знания о проектировании городской пространственной формы и профессиональное обучение моделей генерации изображений

Способ профессионального обучения должен соответствовать особенностям знаний о проектировании городской пространственной формы и отвечать на вызовы, связанные с неявными знаниями^[21]. Систему знаний о проектировании городской пространственной формы можно разделить на знания проектирования пространственной формы и знания пространственного представления. Внутреннюю структуру городской пространственной формы составляют пространственные элементы и их пространственные отношения^[22]. В планировочном

Профессиональное обучение больших моделей: интеграция знаний о проектировании городской пространственной формы в модель генерации изображений

проектировании знания проектирования пространственной формы используются для сочетания пространственных элементов в разных отношениях и формирования планировочного решения; знания пространственного представления используются для выражения решения в виде планов, чертежей и изображений. Знания о проектировании городской пространственной формы являются типичными знаниями предметной области, и универсальная модель генерации изображений должна пройти профессиональное обучение, чтобы их освоить.

На примере универсальной открытой модели генерации изображений Stable Diffusion^① можно выделить две типичные технологии тонкой настройки: LoRA и управляющую сеть ControlNet. Тонкая настройка LoRA способна при небольшом объеме специализированного обучения активировать связь между универсальными и профессиональными знаниями и включить новые знания^[23], поэтому она подходит для профессионального обучения знаниям о проектировании пространственной формы. ControlNet позволяет использовать изображения или карты признаков изображения как условия для управления конкретной структурой и стилем генерируемого изображения^[24], поэтому он подходит для профессионального обучения знаниям пространственного представления.

Существующие способы тонкой настройки моделей генерации изображений создают техническую основу для включения знаний о проектировании городской пространственной формы в большие модели, но между ними еще необходимо выстроить соответствующую техническую рамку. Ниже в качестве прикладного случая рассматривается обновление общественных пространств сообщества; на его основе создается техническая рамка интеграции знаний о проектировании городской пространственной формы в универсальную модель генерации изображений и реализуется профессиональное обучение универсальной большой модели.

3 Практический пример профессионального обучения большой модели: модель обновления общественных пространств сообщества

3.1 Техническая рамка

Обновление существующих общественных пространств сообщества является важной частью современного городского обновления^[25-26]. В настоящее время в обновлении сообществ поощряется применение партисипаторного проектирования: широкое участие жителей подчеркивается на этапах обсуждения намерений обновления, уточнения объектов вмешательства и сравнения вариантов обновления^[27]. Обычно планировщик общается с жителями с помощью профессиональных чертежей - планов, разрезов, визуализаций. Такие материалы имеют порог чтения и понимания, что создает препятствия в коммуникации между планировщиком и жителями. Жители обычно не обладают профессиональными знаниями планирования, поэтому им трудно понимать профессиональные термины вроде проектных ориентаций, недостаточно ясно распознавать проектные элементы и в ходе общения с планировщиком сформировать единое представление о желаемом и реальном результате^[28]. Партисипаторное проектирование в обновлении сообществ является типичной ситуацией, где необходимо передавать и обсуждать неявные знания проектирования пространственной формы.

На основе универсальной большой модели в данном исследовании строится модель обновления общественных пространств сообщества CommunityGenAI для поддержки партисипаторного проектирования. Цель состоит в том, чтобы при вводе нескольких «профессиональных» подсказок модель локально перерисовывала фотографии общественного пространства сообщества и интеллектуально генерировала сцену после обновления. Во-первых,

Профессиональное обучение больших моделей: интеграция знаний о проектировании городской пространственной формы в модель генерации изображений

путем систематизации знаний о проектировании городской пространственной формы профессиональные знания сопоставляются с технологиями тонкой настройки больших моделей, и предлагается техническая рамка модели генерации изображений, интегрирующей эти знания. В этой рамке система профессиональных знаний о проектировании городской пространственной формы рассматривается как знания предметной области и делится на две категории: знания проектирования пространственной формы и знания пространственного представления. Первые доминируются неявными знаниями и выражаются в виде графа знаний, ядром которого является проектная ориентация; граф направляет построение обучающей выборки, включающей полные явные и неявные знания, а «профессиональное обучение» реализуется через тонкую настройку LoRA. Знания пространственного представления являются явными; их «профессиональное обучение» осуществляется через условное управление изображением в ControlNet, направляющее генерацию изображения (рис. 1). Эта техническая рамка позволяет объединить знания предметной области и универсальные знания.

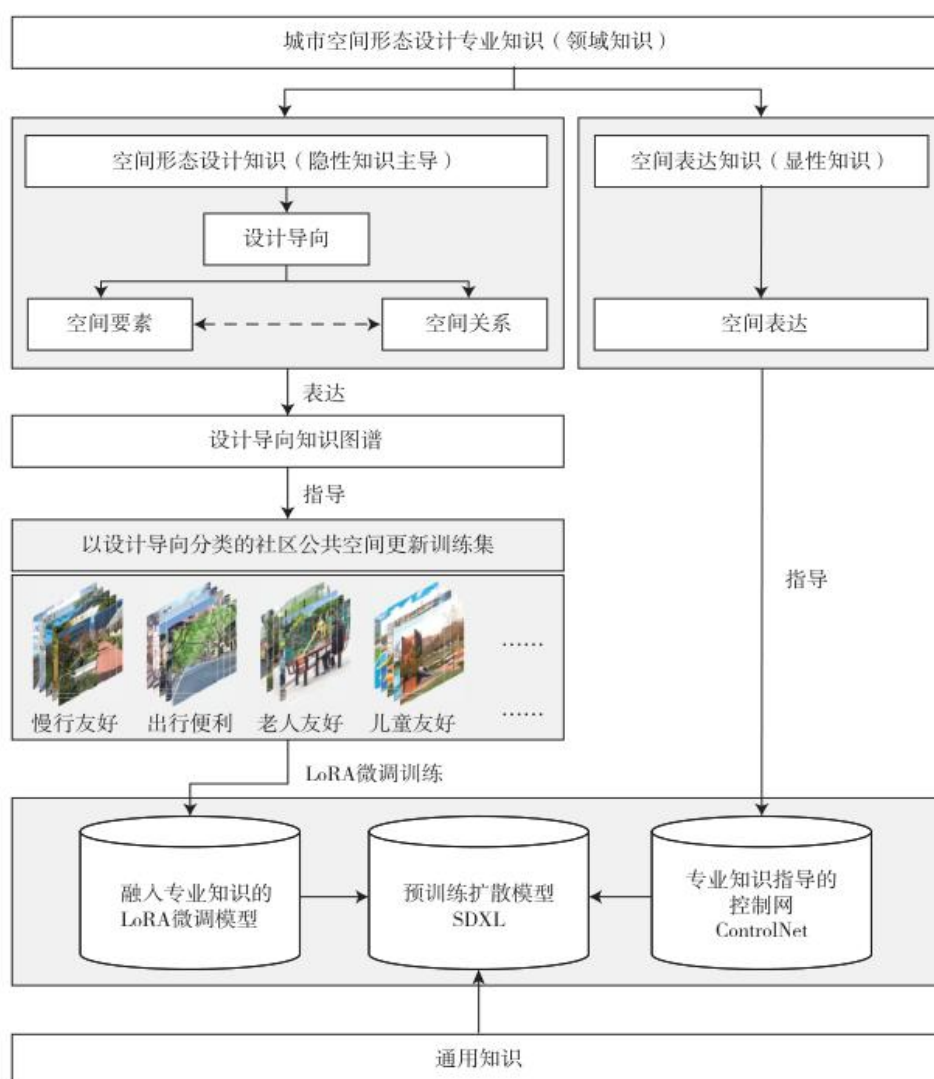


图1 技术框架

Fig.1 Technical framework

3.2 Граф знаний проектирования пространственной формы

В обновлении общественных пространств сообщества содержится большое количество неявных знаний. Графовые базы данных, такие как семантические сети и карты знаний, могут более ясно выражать неявные знания^[29]. На основе семантических сетей компания Google предложила понятие графа знаний (knowledge graph)^[30]. Граф знаний структурирует понятия, объекты и их семантические связи в виде ориентированного графа и представляет собой семантическую сеть, понятную и пригодную для использования компьютером: узлы обозначают сущности, а ребра - семантические отношения между ними^[31]. Графы знаний уже применяются для анализа пространственных элементов и пространственных отношений в городской пространственной форме^[22, 32]. В данном исследовании граф знаний используется для систематизации и выражения знаний о проектировании пространственной формы.

Знания о проектировании пространственной формы в обновлении общественных пространств сообщества разделяются на три категории: проектные ориентации, пространственные элементы и пространственные отношения. Проектная ориентация выражает цель обновления общественного пространства; пространственные элементы - это здания, дороги, фитнес-оборудование, озеленение и т. п. Пространственные отношения выражают связи между элементами, например соседство и включение, пространственный масштаб и др.: площадка может содержать несколько фитнес-устройств, они могут примыкать к дороге, а их пропорции соотносятся с человеческим телом. Под проектной ориентацией пространственные элементы и отношения организуются в общественное пространство сообщества. Например, ориентация «дружественно пожилым» организует проектирование общественного пространства вокруг потребностей пожилых групп и проявляется в предоставлении адресных сервисных объектов, покрытий площадок и элементов отдыха, направленных на удобство пожилых пользователей. Ориентация «спорт и отдых» продвигает массовую физическую активность и организует пространство вокруг спортивного здоровья, добавляя различные спортивные объекты и площадки для удобства занятий жителей.

Пространственные элементы и их атрибуты относятся к явным знаниям и могут быть четко описаны словами. Проектные ориентации и пространственные отношения являются типичными неявными знаниями и трудно поддаются точному словесному выражению. Граф знаний позволяет эффективно организовать и управлять указанными знаниями о городской пространственной форме^[33]. На рис. 2 показана часть графа знаний под ориентацией «дружественно детям». В графе узлы представляют пространственные элементы, а ребра - пространственные отношения. Выражение через граф знаний нужно для того, чтобы модель могла целенаправленно изучать содержащиеся в изображениях знания о проектировании пространственной формы.

3.3 Профессиональное обучение знаниям проектирования пространственной формы

Взяв проектную ориентацию в качестве ядра, исследование систематизирует знания проектирования пространственной формы в обновлении сообщества в виде графа знаний. Граф знаний направляет выбор изображений обучающей выборки, а тонкая настройка LoRA интегрирует явные и неявные знания о проектировании городской пространственной формы в модель генерации изображений. Для этого в модели CommunityGenAI установлены восемь типов проектных ориентаций, объединяющих различные пространственные элементы: удобство передвижения, дружелюбность медленному движению, искусство и культурное творчество,

Профессиональное обучение больших моделей: интеграция знаний о проектировании городской пространственной формы в модель генерации изображений спорт и отдых, дружелюбность пожилым, дружелюбность детям, коммерческая атмосфера и природная экология.

Под руководством графа знаний проектных ориентаций цель профессионального обучения состоит в том, чтобы большая модель понимала саму ориентацию, а также содержащиеся в ней знания о пространственных элементах и отношениях. На примере обновления площадки общественной активности в ориентации «дружелюбно детям» дети, детские игровые сооружения, резиновая дорожка, защитные устройства и другие элементы сами по себе относятся к явным знаниям, но их сочетание, указывающее на ориентацию «дружелюбно детям», является неявным знанием (рис. 2). Отбор изображений обучающей выборки под руководством графа знаний проектной ориентации направляет большую модель на установление связи между конкретными пространственными элементами и ориентацией «дружелюбно детям», тем самым интегрируя в модель неявное знание проектной ориентации.

Изображения, выбранные по графу знаний проектной ориентации, также содержат неявные знания пространственных отношений. Модель способна изучить отношения включения, соседства и масштаба, такие как «детские игровые сооружения находятся внутри пластикового покрытия» и «защитное устройство примыкает к пластиковому покрытию» (рис. 2). Эти неявные знания о пространственных отношениях передаются большой модели через обучающую выборку и реализуют включение в нее неявных знаний пространственных отношений под ориентацией «дружелюбно детям».

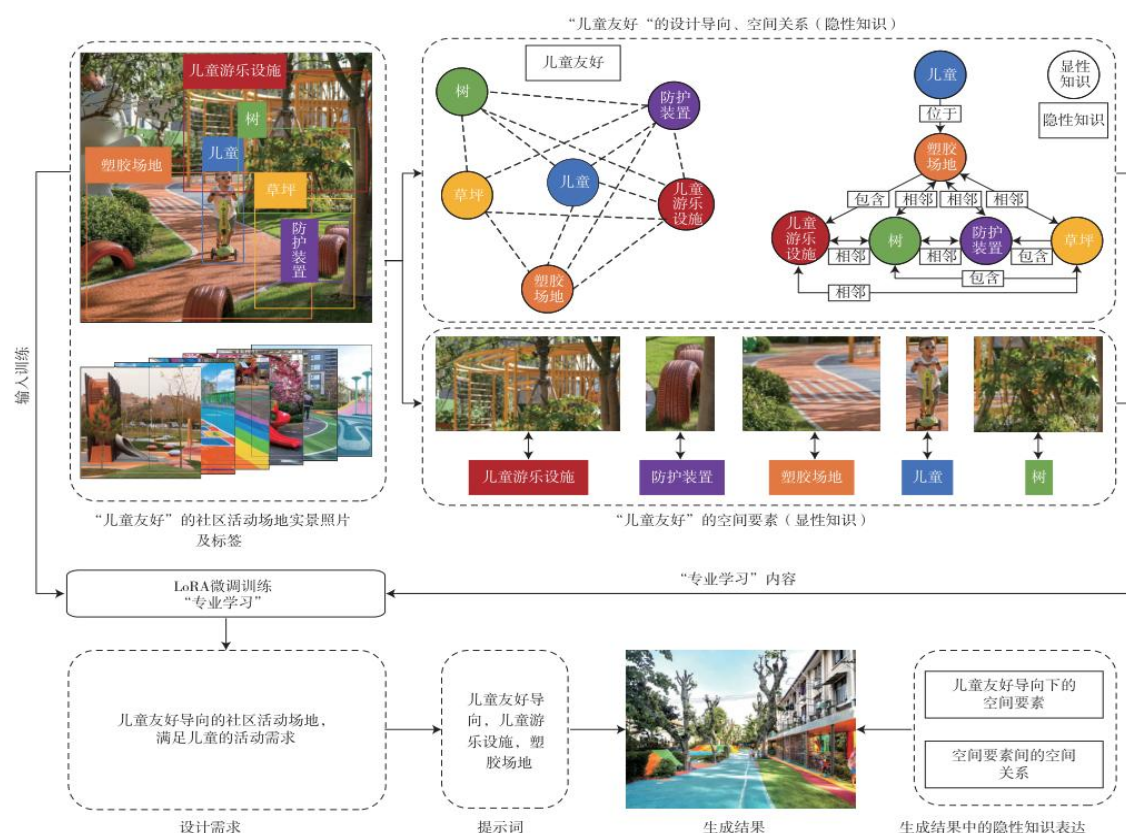


图2 空间形态设计知识“专业学习”和表达的路径
Fig.2 Pathway of "domain-specific learning" and expression of spatial form design knowledge

Рис. 2 Трактория «профессионального обучения» и выражения знаний проектирования пространственной формы

3.4 Профессиональное обучение знаниям пространственного представления

Проектирование городской пространственной формы относится к трехмерному городскому пространству, но должно выражаться через двумерные чертежи. Двумерное представление трехмерного пространства также является важным проектным знанием. Профессиональные планировщики обязательно изучают начертательную геометрию и другие специальные курсы именно для освоения знаний пространственного представления, позволяющих переводить трехмерное пространство в двумерные изображения. Эти знания являются явными и могут выражаться точными формулами и геометрическими методами.

Использование модели генерации изображений для выражения эффекта обновления общественного пространства сообщества означает интеллектуальную локальную перерисовку фотографии. Такая перерисовка должна сохранять общую согласованность пространственного представления изображения. Это требует «профессионального обучения» знаниям пространственного представления, чтобы сцена после обновления была выражена корректно.

Знания пространственного представления здесь подразделяются на отношения перспективы, перекрытия и относительного размера. Во-первых, на примере улицы сообщества реальная фотография на рис. 3 выражает одноточечную перспективу уличного пространства. Во-вторых, после проецирования трехмерных пространственных элементов на двумерную плоскость передние объекты перекрывают задние. Наконец, относительные размеры объектов в реальном трехмерном пространстве отражаются в двумерном изображении через проекцию: на рис. 3 правдиво выражено относительное соотношение высот человека и электровелосипеда. Все эти знания пространственного представления могут быть представлены теоремами начертательной геометрии и затем включены в большую модель через различные виды условного управления изображением в ControlNet.

В Stable Diffusion расширения ControlNet могут преобразовывать семантические карты, линейные схемы и другие материалы в условия управления изображением, что подходит для выражения и передачи знаний пространственного представления. Исследование предлагает использовать тонкую настройку условного управления изображением ControlNet как путь профессионального обучения знаниям пространственного представления и выстраивает рамку соответствия между существующими типами ControlNet и такими знаниями. В частности, отношения перспективы задаются через входные условия ControlNet типа обнаружения и управления границами; отношения перекрытия и относительного размера - через входные условия Segmentation (семантическая сегментация) и Tile (разбиение на блоки) (рис. 3).

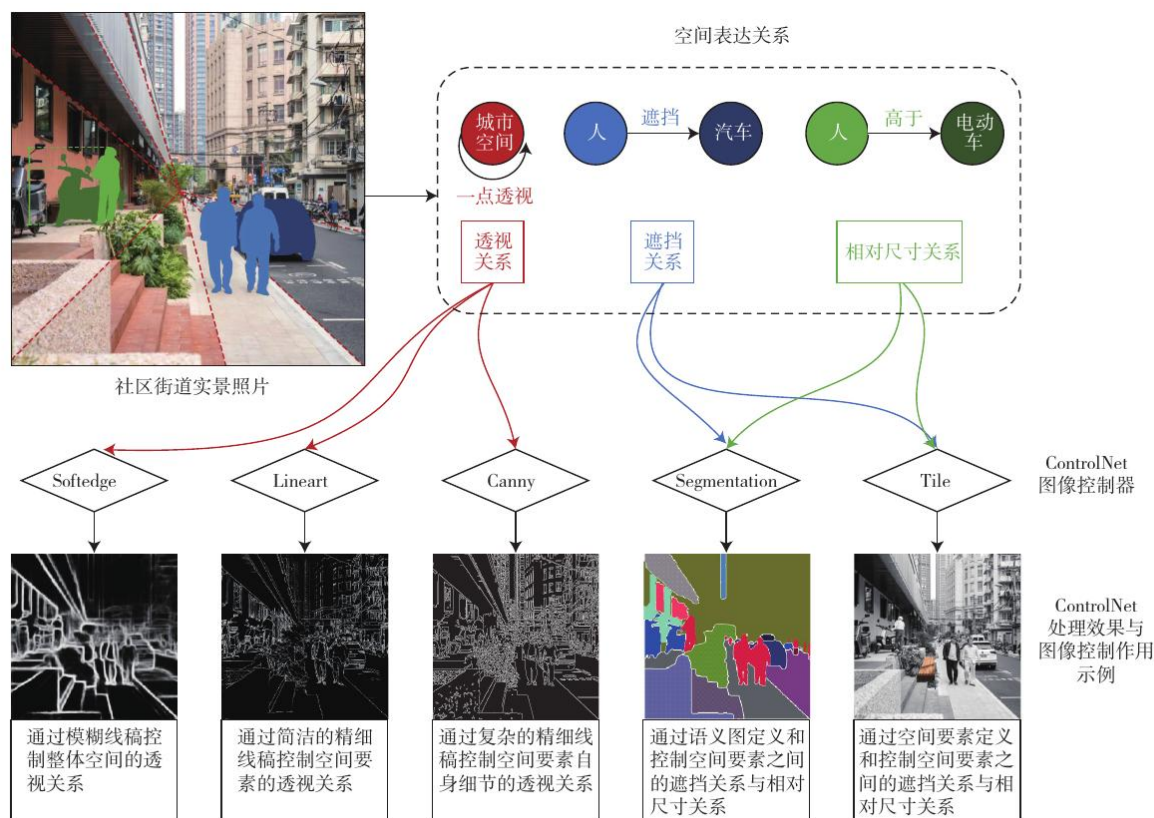


Рис. 3 Траектория «профессионального обучения» и выражения знаний пространственного представления

4 Построение модели CommunityGenAI для обновления общественных пространств сообщества и практический пример

4.1 Построение набора меток и подготовка данных

Обучающая выборка профессионального обучения модели CommunityGenAI состоит из двух частей: меток и изображений. Построение обучающей выборки и меток опирается на граф знаний проектных ориентаций. Метки выражают знания о проектных ориентациях и пространственных элементах в проектировании городской пространственной формы. В процессе применения модели они непосредственно служат подсказками, осуществляющими текстовый контроль генерации изображения. Набор меток формируется в соответствии с графом знаний проектных ориентаций и соответствующими пространственными элементами.

Изображения являются объектом профессионального обучения модели и наглядным выражением знаний о проектировании городской пространственной формы. Под руководством графа знаний проектных ориентаций специалисты-планировщики отбирают реальные фотографии сцен сообщества с полно выраженной проектной ориентацией, четко представленными пространственными элементами и высоким пространственным качеством в качестве обучающей выборки. Под руководством восьми типов графов знаний проектных ориентаций было собрано 436 реальных фотографий сцен сообщества, охватывающих разные типы жилых сообществ, чтобы профессиональное обучение модели могло адаптироваться к потребностям обновления жилых сообществ разных периодов.

В соответствии с графом знаний проектных ориентаций реальные фотографии классифицируются по ориентациям, и каждое изображение маркируется с учетом пространственных элементов, подчиненных соответствующей ориентации. Это обеспечивает

Профессиональное обучение больших моделей: интеграция знаний о проектировании городской пространственной формы в модель генерации изображений

появление каждой метки и обозначаемого ею пространственного элемента в обучающей выборке не менее пяти раз и формирует файл меток.

4.2 Интеграция знаний о проектировании городской пространственной формы через тонкую настройку LoRA

Через тонкую настройку LoRA в большую модель интегрируются знания о проектных ориентациях, пространственных элементах и пространственных отношениях в проектировании городской пространственной формы. Пространственные элементы как явные знания могут напрямую размечаться набором меток; проектные ориентации и пространственные отношения относятся к неявным знаниям и выражаются реальными фотографиями обучающей выборки. Поскольку тонкая настройка LoRA способна включать новые знания при небольшом объеме обучения, с ее помощью реализуется интеграция знаний о проектировании пространственной формы.

Обучение LoRA использует большую модель Stable Diffusion XL «base 1.0». Программно-аппаратная среда и параметры приведены в табл. 1 и табл. 2.

Табл. 1 Программно-аппаратная среда обучения модели

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

Табл. 2 Параметры обучения

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

После 100 циклов обучающих итераций значение функции потерь (Loss) стабилизируется примерно на уровне 0,08, и модель демонстрирует наилучший результат обучения. С помощью таблицы X/Y/Z была протестирована производительность тонко настроенной модели LoRA при разных весах. На основе результатов тестирования оптимальный вес LoRA определен как 0,4, а модель последней итерации выбрана в качестве итогового результата тонкой настройки LoRA.

4.3 Проявление профессиональных знаний проектирования пространственной формы в модели CommunityGenAI

Профессиональное обучение больших моделей: интеграция знаний о проектировании городской пространственной формы в модель генерации изображений

Модель CommunityGenAI посредством профессионального обучения интегрирует профессиональные знания о проектных ориентациях, пространственных элементах и пространственных отношениях, а также знания пространственного представления. При обновлении сообщества достаточно выбрать изображение существующего общественного пространства, ввести подсказки проектной ориентации и ключевых пространственных элементов, чтобы интеллектуально сгенерировать концептуальное изображение общественного пространства после локального обновления.

Возьмем пример обновления площадки общественной активности с ориентацией «дружественно детям». При вводе подсказок «дружественно детям», «цвета, дружественные детям», «детские игровые сооружения», «песочница» создается концептуальное изображение, соответствующее этим проектным ориентациям (рис. 4). В сгенерированном изображении модель включает не только детские игровые сооружения и песочницу; цветовое решение соответствует ориентации «дружественно детям», а в сцене присутствуют дети и сопровождающие их родители, пластиковое покрытие и другие элементы. Подсказки не содержали слов «родители» и «пластиковое покрытие»; это и есть неявные знания, заключенные в слове «ориентация, дружественная детям». Относительное положение и размеры игровых сооружений, деревьев, песочницы и других элементов также являются неявными знаниями, освоенными в ходе профессионального обучения. Модель CommunityGenAI интегрировала знания о проектировании пространственной формы в обновлении сообщества, особенно неявные знания.

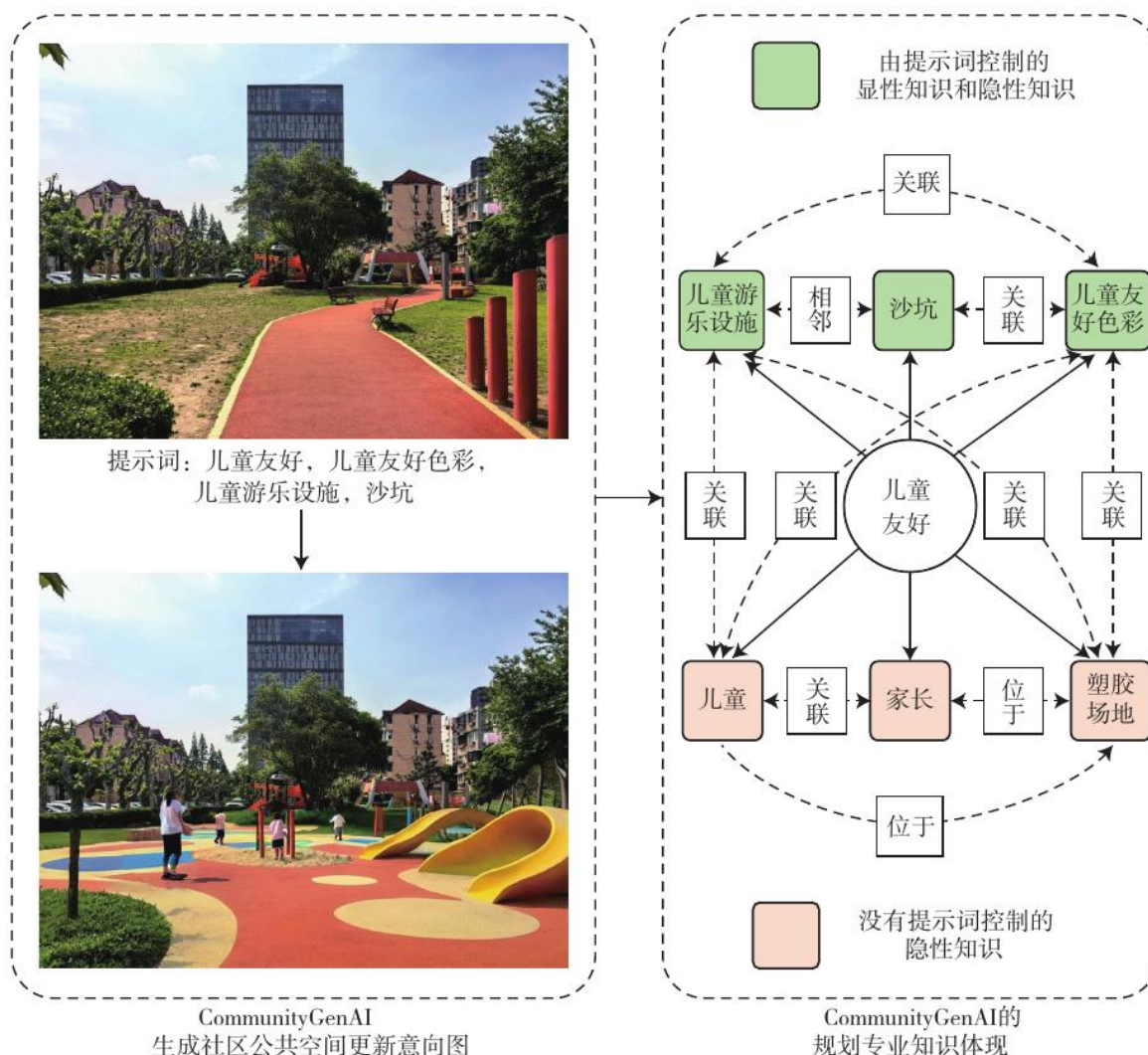


Рис. 4 Проявление профессиональных знаний в модели CommunityGenAI

Поскольку модель прошла профессиональное обучение по проектным ориентациям, CommunityGenAI может в одном и том же случае обновления сообщества интеллектуально генерировать концептуальные изображения разных проектных ориентаций, чтобы жители могли сравнивать и выбирать варианты. На рис. 5 показаны три концептуальных изображения обновления для внутренней дороги между жилыми домами в одном сообществе по ориентациям «коммерческая атмосфера», «дружественно пожилым» и «удобство передвижения». Столы и стулья, а также цветники под деревьями, сгенерированные по ориентации «коммерческая атмосфера», заметно отличаются от столов и стульев в ориентации «дружественно пожилым» и от цветников под деревьями в ориентации «удобство передвижения» (рис. 5). Концептуальные изображения разных ориентаций содержат существенно различные пространственные элементы, но каждое отражает типичные характеристики общественного пространства сообщества в соответствующей ориентации; масштаб при этом соответствует требованиям пространственного представления концептуального изображения. Это показывает, что CommunityGenAI посредством профессионального обучения освоила неявные знания, содержащиеся в конкретных проектных ориентациях.

Профессиональное обучение больших моделей: интеграция знаний о проектировании городской пространственной формы в модель генерации изображений

Модель CommunityGenAI эффективно прошла профессиональное обучение по проектным ориентациям, пространственным элементам, пространственным отношениям и знаниям пространственного представления в обновлении общественных пространств сообщества. Она овладела явными и неявными знаниями такого обновления и приобрела способности профессиональной большой модели для обновления общественных пространств сообщества.



Рис. 5 Интеллектуальная генерация CommunityGenAI концептуальных изображений по трем проектным ориентациям

4.4 Передача профессиональных знаний в модели CommunityGenAI

Модель генерации изображений, интегрирующая знания о проектировании городской пространственной формы, способна выражать и передавать неявные знания. Она может эффективно работать как на ранней стадии коммуникации по обновлению сообщества, так и на стадии обмена и обратной связи по проектным решениям, поддерживая общение между жителями и планировщиками. На примере обновления жестко мощеной площадки в старом жилом квартале: на этапе обсуждения намерений обновления модель помогает планировщику и жителям уточнить объект и границы обновления; на этапах детализации проектной ориентации и сравнения вариантов планировщик с помощью CommunityGenAI быстро переводит потребности жителей в профессиональные термины, такие как «ориентация, дружественная детям» и «детские игровые сооружения», а также в соответствующую

Профессиональное обучение больших моделей: интеграция знаний о проектировании городской пространственной формы в модель генерации изображений управляющую сеть. Модель способна в среднем за 40 секунд на изображение интеллектуально генерировать серию концептуальных карт обновления, соответствующих проектной ориентации. Этот процесс обеспечивает быстрое сопряжение потребностей жителей с профессиональным языком планировщика и представляет трудно выразимые неявные знания в понятной для жителей визуальной форме обратной связи. CommunityGenAI реализует эффективную передачу профессиональных знаний об обновлении общественных пространств сообщества между планировщиками и жителями, особенно эффективное выражение и передачу неявных знаний, поддерживает достижение согласованного намерения обновления и тем самым способствует партисипаторному проектированию в обновлении сообществ.

5 Заключение и обсуждение

5.1 Заключение

В статье профессиональные знания о проектировании городской пространственной формы рассматриваются как «знания предметной области». Далее вводятся понятия неявных и явных знаний из менеджмента знаний, что позволяет классифицировать профессиональные знания проектирования пространственной формы и соотнести их с технологиями тонкой настройки больших моделей. На этой основе предложены два пути профессионального обучения и построена техническая рамка интеграции профессиональных знаний о проектировании городской пространственной формы в универсальную модель генерации изображений. В качестве прикладного примера выбрано обновление общественных пространств сообщества, и создана профессиональная большая модель CommunityGenAI. Исследование приводит к следующим трем выводам.

Во-первых, в системе знаний планировочной дисциплины существует большое количество неявных знаний, особенно доминирующих в проектировании городской пространственной формы. Коммуникация и передача этих неявных знаний требуют участия больших моделей. Передача неявных знаний проектирования городской пространственной формы является специфической потребностью планировочной дисциплины в больших моделях. Ключевая трудность построения большой модели для проектирования пространственной формы заключается в том, как интегрировать знания о проектировании городской пространственной формы в универсальную большую модель.

Во-вторых, предложена техническая рамка, позволяющая посредством профессионального обучения интегрировать знания о проектировании городской пространственной формы в универсальную модель генерации изображений. Профессиональные знания разделяются на знания проектирования пространственной формы и знания пространственного представления. Первые формируют граф знаний явных и неявных знаний с проектной ориентацией в качестве ядра; вторые как явные знания направляют визуальное выражение. Две технологии тонкой настройки используются для включения знаний о проектировании городской пространственной формы в универсальную модель генерации изображений.

В-третьих, на основе универсальной модели генерации изображений Stable Diffusion и посредством тонкой настройки большой модели в универсальную модель интегрированы профессиональные знания проектирования обновления общественных пространств сообщества. Построена профессиональная большая модель CommunityGenAI, поддерживающая партисипаторное проектирование при обновлении сообществ. Обученная модель CommunityGenAI освоила профессиональные знания и обладает знаниями проектных ориентаций, пространственных элементов, пространственных отношений и пространственного

Профессиональное обучение больших моделей: интеграция знаний о проектировании городской пространственной формы в модель генерации изображений представления для проектирования общественных пространств сообщества. В сценарии партисипаторного проектирования CommunityGenAI реализует передачу профессиональных знаний, играет ключевую роль в эффективной коммуникации и передаче неявных знаний, повышает эффективность общения при обновлении сообществ и действительно поддерживает партисипаторное проектирование.

5.2 Обсуждение

В области планирования универсальные большие языковые модели уже могут моделировать заинтересованные стороны в процессе планирования и проектирования, эффективно имитируя общественное участие^[34]. Такой подход рассматривает большую модель как «общественность», обладающую только универсальными знаниями. В этой статье профессиональная модель планирования строится на базе универсальной большой модели посредством профессионального обучения, благодаря чему она становится экспертом, обладающим одновременно «универсальными знаниями + знаниями предметной области» и способным выполнять особые функции в профессиональной сфере. Оба пути являются эффективными способами применения технологий больших моделей в планировочной дисциплине и заслуживают дальнейшего глубокого изучения.

Профессиональное обучение больших моделей также является техническим путем создания других типов профессиональных моделей для планирования. То, как интегрировать профессиональные знания планирования в универсальные большие модели, станет ключевой технологией построения таких моделей. С точки зрения управления знаниями профессиональные знания разных областей планирования сами по себе различаются. Отличаться будут не только пути профессионального обучения явным и неявным знаниям; различными будут и пути обучения разным видам неявных знаний. Методы включения различных планировочных знаний в универсальные большие модели неизбежно будут неодинаковыми. Поэтому более целесообразным путем может быть пообластное и категориальное построение профессиональных больших моделей для планирования. Это направление заслуживает систематического исследования.

Интерес многих дисциплин к профессиональным большим моделям в целом объясняется двумя причинами. Первая связана с научными исследованиями: большие модели применяются для выполнения задач, которые трудно выполнить вручную, например для обработки и анализа массивных данных, тем самым повышая эффективность исследований. Вторая связана с инженерной практикой: большие модели используются для повышения эффективности проектирования, оптимизации производственных процессов, снижения затрат и повышения производительности. Для городского планирования, помимо этих двух потребностей, существует и специфическая дисциплинарная потребность. Она состоит в том, что распространение и передача неявных знаний планировочной дисциплины нуждаются в больших моделях, что определяется публично-политической природой планирования. Как обеспечить обучение больших моделей неявным знаниям планировочной дисциплины и их действие в пространственном планировании полного цикла - это особая потребность профессиональных больших моделей для планирования. Данное направление требует дальнейшего углубленного изучения и исследования.

Примечания

Профессиональное обучение больших моделей: интеграция знаний о проектировании городской пространственной формы в модель генерации изображений

- ① Stable Diffusion - универсальная большая модель изображений, выпущенная компанией Stability AI. Stable Diffusion XL 1.0 - открытая версия, опубликованная в июле 2023 г.; ее базовая модель (base model) имеет 3,5 млрд параметров.

Список литературы

- [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 Otthein, PAN Haixiao, DENG Zhituan, et al. Искусственный интеллект нового поколения на службе городского планирования: возможности и вызовы [J]. Urban Planning Forum, 2023(4): 1-11. (на китайском)
- [3] YUAN Chao, ZHENG Hao. Новые модели архитектурного проектирования под влиянием генеративного искусственного интеллекта [J]. Architectural Journal, 2023(10): 29-35. (на китайском)
- [4] HE Wanyu, YANG Liangsong. Исследование применения генеративного искусственного интеллекта в архитектурном проектировании: на примере Xkool AI Cloud [J]. Architectural Journal, 2023(10): 36-41. (на китайском)
- [5] POLANYI M. The Tacit Dimension [M]. Routledge and Kegan Paul, 1966.
- [6] YAO Wei. Управление знаниями [M]. Beijing: Tsinghua University Press, 2020. (на китайском)
- [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. Цифровые технологии планирования: данные и знания [J]. Urban Planning Forum, 2024(2): 18-24. (на китайском)
- [10] LIN Bingyao. Количественные методы городской пространственной формы и их оценка [J]. Urban Planning Forum, 1998(3): 42-45. (на китайском)
- [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 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. Построение теоретической модели поддержки городского дизайна средствами AIGC [J]. Urban Planning Forum, 2023(2): 12-18. (на китайском)
- [22] YANG Junyan, SHAO Dian, WANG Peng, et al. Интеграция, топология и перевод: метод углубленного анализа городской формы на основе графов знаний [J]. City Planning Review, 2023, 47(6): 57-67. (на китайском)
- [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. Эволюция, вызовы и рекомендации по городскому обновлению в Китае [J]. Planners, 2021, 37(24): 21-27. (на китайском)
- [26] HUA Xiahong, ZHUANG Shen. Продвижение обновления общественных пространств повседневной жизни средствами дизайна: обзор практик микровозобновления в Шанхае [J]. Architectural Journal, 2022(3): 1-11. (на китайском)
- [27] LI Qing, LIN Ni. Исследование новой модели партисипаторного планирования и проектирования микровозобновления сообщества с точки зрения «народного города»: пример микровозобновления жилого переулка YF в Шанхае [J]. Urban Planning Forum, 2023(6): 87-96. (на китайском)
- [28] YANG Junyan, ZHANG Zhonghu, SHI Yi. Исследование цифровой платформы общественного участия в городском дизайне для городского обновления: пример района Юэцзянлоу в Нанкине [J]. Urban Planning Forum, 2024(3): 74-81. (на китайском)
- [29] ZHANG Kai. Управление информационными ресурсами [M]. 3-е изд. Beijing: Tsinghua University Press, 2013. (на китайском)
- [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. ReITR: 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. Методы построения и перспективы применения графов знаний городских торговых районов и кварталов для устойчивого городского развития [J]. Journal of Geo-information Science, 2023, 25(12): 2374-2386. (на китайском)
- [34] TIAN Li, YANG Xin, ZHANG Yudi, et al. Инновации в обучении городскому и сельскому планированию и проектированию на основе двойного двигателя «профессиональные знания + искусственный интеллект»: пример планирования жилых районов [J]. Urban Planning Forum, 2024(5): 71-78. (на китайском)

Доработано: январь 2025

Профессиональное обучение больших моделей: интеграция знаний о проектировании
городской пространственной формы в модель генерации изображений
(Статья переведена при содействии ИИ.)