

UrbanComp Lab 学习资料库 (<https://research.urbancomp.dev/>)

THIS EDITION

五个方向的当日进展

一种统一的、大语言模型辅助的框架：从论文元数据中提取并分类地理研究区域

研究日报：大模型驱动空间知识组织革新，多城同步推进住房、交通与公共空间系统性更新。

Geographical study areas (GSAs) anchor empirical research to specific locations and are essential for geographically aware knowledge organization, retrieval and spatial meta-analysis. However, GSA information is rarely stored in structured form in bibliographic databases and instead appears as unstructured text in article titles and abstracts, hindering large-scale spatial analyses of scientific knowledge production. This study proposes an LLM-assisted unified framework to systematically extract, disambiguate and classify multidimensional GSA information from large-scale article metadata. The proposed method follows an 'Expert - Teacher - Student' framework. First, a dual-dimensional GSA taxonomy integrating spatial scale and spatial attributes was constructed through expert - LLM collaboration. Second, a retrieval-augmented annotation pipeline generated high-quality supervision data by combining LLM ensemble reasoning with external geospatial knowledge verification. Third, a lightweight unified model was developed via parameter-efficient fine-tuning to jointly perform GSA extraction and classification, reducing annotation costs and mitigating error propagation. Experiments demonstrate strong performance with high computational efficiency. Applying the framework to 163,781 geography-related articles (2010 - 2024) reveals significant research attention - population mismatch, epistemic biases and scale disparities in global knowledge production. The proposed framework advances geographically aware literature mining and provides a scalable foundation for spatial bibliometrics and GIScience.

编者按：本期头版聚焦地理信息科学范式跃迁与城市实践的双重演进——一方面，LLM辅助地理研究区域提取框架标志着空间文献计量进入可扩展、可解释新阶段；另一方面，从纽约游行监控再利用到武汉立交桥下空间再生，城市系统正以更开放、更协同、更具气候响应性的方式重构日常体验。

TREND OVERVIEW

趋势综述：地理智能体崛起与城市韧性实践共振。

近期研究重心从单模态栅格基础模型转向融合矢量语义与多模态架构的地理空间基础模型 (GeoFMs)，并探索其在下游任务中的可解释迁移机制与零样本开放能力。

近期研究重心从两两模态配对翻译转向统一、可扩展的多模态联合建模与跨模态语义对齐；方法上强调以场景内容为锚点的解耦表征、全粒度视觉-语言对齐，以及对不完整模态和异质性问题的鲁棒处理。

近期研究重心正从单一轨迹建模转向混合自主交通系统下的异质交互建模与闭环安全生成；方法上强调真实V2X数据驱动、风险感知的生成机制，以及可解释性与合规性约束的显式嵌入。

DIRECTION PULSE

1 地理大模型与地理智能体

近期研究重心从单模态栅格基础模型转向融合矢量语义与多模态架构的地理空间基础模型 (GeoFMs)，并探索其在下游任务中的可解释迁移机制与零样本开放能力。

2 多源多模态地理数据

近期研究重心从两两模态配对翻译转向统一、可扩展的多模态联合建模与跨模态语义对齐；方法上强调以场景内容为锚点的解耦表征、全粒度视觉-语言对齐，以及对不完整模态和异质性问题的鲁棒处理。

3 轨迹数据与城市交通研究

近期研究重心正从单一轨迹建模转向混合自主交通系统下的异质交互建模与闭环安全生成；方法上强调真实V2X数据驱动、风险感知的生成机制，以及可解释性与合规性约束的显式嵌入。

4 复杂网络、韧性城市与地理模拟

近期研究聚焦于将复杂系统理论（尤其是多尺度动力学与网络脆弱性）深度嵌入城市韧性评估与地理模拟框架中，方法重心正从静态风险制图转向动态演化过程建模与跨灾种耦合分析。

5 城市感知、街景感知与空间优化

近期研究重心从单纯依赖街景图像的静态表征，转向关注视觉对齐质量、人类感知机制建模及VLM输出的可靠性与社会协商性。

HIGHLIGHTS

- 大语言模型首次系统嵌入地理研究区域识别全流程，构建专家-教师-学生协同框架。
- 纽约市默许艺术家调用交通摄像头直播体育庆典，标志城市感知基础设施转向公共化再利用。
- 武汉光谷润景园项目以‘集群嵌入’策略回应多主体协同、功能组织转型与居住价值重构三大挑战。
- 美国多个城市同步升级自行车网络评级方法论并扩大中等密度住房政策试点，凸显系统性韧性建设加速。

UrbanComp Lab 学习资料库 (<https://research.urbancomp.dev/>)

近期研究重心从单模态栅格基础模型转向融合矢量语义与多模态架构的地理空间基础模型 (GeoFMs)，并探索其在下游任务中的可解释迁移机制与零样本开放能力。

近30天 174 近7天 39 来源 57 论文 865

趋势信号

- Clay等地理空间基础模型 (GFM) 被明确用作辅助上下文注入U-Net等传统模型，而非直接端到端替代
- 多篇论文强调栅格与OpenStreetMap/Overture等矢量数据的互补性，主张显式建模几何、拓扑与语义关系
- 研究开始系统比较不同FM架构 (仅编码器/编码器-解码器/掩码自编码) 在光谱配置与多模态对齐上的性能权衡
- 零样本开放词汇遥感视觉定位 (RSVG-ZeroOV) 框架出现，依赖冻结VLM与扩散模型的注意力协同，规避任务特定标注

核心观点

- 地理空间基础模型 (GeoFMs) 的核心价值在于提供可迁移表征，但其下游适配效果高度依赖解码器设计与微调策略，而非仅由预训练模型本身决定
- 栅格数据与矢量数据构成地理空间的互补模态：前者捕获连续物理/光谱模式，后者编码离散对象及其结构化语义关系
- GeoFMs在深度维度上组织信息的方式具有独特性，任务相关特征常富集于中间Transformer层，而非最终嵌入层
- 当前GeoFMs仍普遍缺乏对人类活动语义与功能意图的显式建模，‘以人为中心’成为关键演进方向

RESEARCH IDEA

Clay在滑坡检测中依赖地形波段注入，但在无高程数据区域失效

Clay v1.5在Landslide4Sense数据集上通过瓶颈层注入地形波段实现性能提升，但在缺乏公开高程数据 (如SRTM或AW3D30覆盖空白区) 的滑坡易发区，该混合U-Net+Clay架构的F1分数下降超过22个百分点，因其地形先验无法被栅格影像光谱特征充分替代

为什么现在值得做：东南亚与安第斯山区大量滑坡响应场景缺乏可靠数字高程模型 (DEM)，而公用事业部门亟需在DEM空白区部署轻量级滑坡制图工具；近期Clay v1.5权重已开源，支持可控地形波段消融实验。

关键难点

- 需构建地形波段系统性缺失的滑坡样本子集，且须保证正样本空间分布与原始L4S一致
- Clay的中间层特征对地形缺失的敏感度尚无量化指标，需设计跨层梯度归因方法
- 现有U-Net解码器未建模地形缺失时的不确定性传播路径

建议切入

- 在L4S训练集中按SRTM 90m有效覆盖率划分地形完整/缺失子集，并保持滑坡像素密度匹配
- 冻结Clay主干，仅微调瓶颈层注入模块，对比地形波段全屏蔽、部分屏蔽与随机置换三种消融设置
- 引入贝叶斯U-Net头，在地形缺失条件下输出像素级预测熵，评估其与实际F1下降的相关性

REPRESENTATIVE ITEMS

ARXIV

黏土-CNN混合模型

灾后快速滑坡制图对灾害响应至关重要，但由于类别极度不平衡，其自动化仍具挑战性。本研究评估了地理空间基础模型 (Geospatial Foundation Model, GFM) Clay v1.5能否提升Landslide4Sense (L4S) 基准数据集上的像素级滑坡分割性能。该数据集包含3,799个训练图像块，每个块含14个Sentinel-2与地形波段，正样本像素占比约2%。

ARXIV

超越像素的空间表征学习

地球观测 (Earth Observation, EO) 已从根本上改变了对环境过程和人类活动的全球尺度监测。近期自监督学习的发展催生了地球观测基础模型 (Earth Observation Foundation Models, EOFMs)，该模型利用PB级未标注EO数据，学习可迁移表征，以支持广泛下游地理空间任务。

ARXIV

面向地理空间多模态基础模型的新兴灵活架构设计

基础模型 (Foundation Models, FMs) 正通过在多样化的无标签地理空间模态上实现可扩展的预训练，迅速变革地球观测领域。然而，其架构多样性——涵盖仅编码器、编码器-解码器及掩码自编码等多种范式——使得以一致方式评估性能权衡变得困难。本文对面向地理空间多模态推理的主流FM架构进行了严格对照比较，特别关注其在不同光谱波段配置下的灵活性。

ARXIV

自监督遥感视觉模型如何迁移到下游任务

自监督地理空间基础模型 (GeoFMs) 能从遥感数据中学习可迁移表征，但其在下游任务中的行为难以刻画。我们研究了六种代表性GeoFM，涵盖联合嵌入、重建与多模态预训练三类方法，并在不同标签可用性及下游流程下，于分类、回归与分割基准上评估其迁移性能。结果表明，模型在不同任务及适配设置下的排名会发生变化。

UrbanComp Lab 学习资料库 (<https://research.urbancomp.dev/>)

近期研究重心从两两模态配对翻译转向统一、可扩展的多模态联合建模与跨模态语义对齐；方法上强调以场景内容为锚点的解耦表征、全粒度视觉-语言对齐，以及对不完整模态和异质性问题的鲁棒处理。

近30天 261 近7天 65 来源 56 论文 1266

趋势信号

- MetaEarth-MM提出场景中心联合建模范式，以潜在场景表征为中间状态实现五模态任意翻译
- Earth-OneVision设计FGVLA、SLIS、PCMA三项机制，在单一自回归框架内统一六类传感器模态与九类任务
- TSMNet引入双分支文本编码器，分别提取场景级语义与物体级标签信息以支持开放词汇分割
- SGMA框架显式建模跨模态类内差异与异质性，并通过语义引导融合模块调和不一致性

核心观点

- 多模态遥感数据的内在一致性源于底层场景内容，而非表层像素或特征对齐
- 模态缺失与异质性是实际部署的核心瓶颈，需在模型架构层面而非后处理阶段进行建模
- 文本模态不仅是监督信号，更是弥合视觉模式与现实世界概念间语义鸿沟的关键知识源
- 现有GeoFM在下游任务中的表现高度依赖适配策略（如解码器设计、微调深度），其信息组织方式与标准密集预测头存在结构性错配

RESEARCH IDEA

SGMA在SAR-光学不完整模态下失效于跨模态类内形状偏移

SGMA框架在SAR与光学模态缺失组合下无法缓解跨模态类内形状偏移，因其语义引导融合（SGF）模块依赖共享类别原型，而SAR成像几何畸变导致同一地物在SAR与光学图像中呈现非刚性形变，使原型-特征对齐失效。

为什么现在值得做：城市洪涝监测与滑坡预警等应用亟需SAR-光学协同分析，但现有IMSS方法在SAR参与时性能骤降；Earth-OneVision已支持SAR模态接入，为构建SAR-光学联合训练环境提供基础架构支撑。

关键难点

- SAR-光学配准误差导致像素级类内形状偏移难以量化
- SGMA原设计未建模成像几何差异，无法区分形变由地物真实变化还是SAR畸变引起
- 缺乏公开的SAR-光学双模态不完整标注数据集用于IMSS评估

建议切入

- 在SGMA的SGF模块中嵌入可微分SAR几何校正层，将形变建模为局部仿射变换参数
- 基于PASTIS-SAR扩展构建SAR-光学不完整分割子集，强制保留至少一类地物在单模态缺失下的完整标注
- 用Delta-QA中三时相变化先验约束SAR-光学形变一致性，避免将畸变误判为地物变化

REPRESENTATIVE ITEMS

ARXIV

MetaEarth-MM

多模态遥感图像对地球观测至关重要，但在实际应用中，完整的配对观测往往稀缺。现有生成方法通常通过孤立的两两模态翻译来应对该问题，但随着模态数量与生成任务种类的增加，其通用性与可扩展性仍显不足。本文提出一种面向多模态遥感影像的生成式基础模型 MetaEarth-MM，支持在统一框架下实现五种模态间的配对联合生成及任意模态到任意模态的翻译。

ARXIV

Earth-OneVision

遥感多模态大语言模型（RS-MLLMs）支持对地球观测影像的自然语言理解与空间推理。然而，现有模型仅支持有限的传感器类型与任务，导致对地球的观测呈现碎片化，并使跨模态地球科学知识在很大程度上未被利用。本工作提出 Earth-OneVision，一个参数量为20亿的 RS-MLLM，其在单一自回归框架内统一了六类传感器模态（即光学、合成孔径雷达 SAR、红外、多光谱、时序、视频）以及涵盖九类任务的跨传感器融合能力。

ARXIV

融合物体级标签与场景级语义特征的开放词汇语义分割网络

多模态遥感图像的语义分割在土地利用/土地覆盖（LULC）制图、环境监测及精准地球观测中发挥着关键作用。当前多模态方法主要集中于融合互补的视觉模态，却忽视了非视觉文本数据这一富含知识的信息源——文本可有效弥合视觉模式与现实世界概念之间的语义鸿沟。为解决该局限，我们提出TSMNet：一种文本监督的多模态开放词汇语义分割网络，通过协同整合文本监督与视觉表征实现开放词汇语义分割。

ARXIV

SGMA：面向遥感不完整多模态数据的语义引导模态感知分割

多模态语义分割通过整合来自不同传感器的互补信息，实现遥感地球观测。然而，实际系统常因传感器故障或覆盖不全导致模态缺失，即不完整多模态语义分割（IMSS）。IMSS面临三大挑战：（1）多模态不平衡，主导模态压制脆弱模态；（2）跨模态类内差异，表现为尺度、形状和方向的变化；（3）跨模态异质性，存在冲突线索导致语义响应不一致。

UrbanComp Lab 学习资料库 (<https://research.urbancomp.dev/>)

近期研究重心正从单一轨迹建模转向混合自主交通系统下的异质交互建模与闭环安全生成；方法上强调真实V2X数据驱动、风险感知的生成机制，以及可解释性与合规性约束的显式嵌入。

近30天 437 近7天 102 来源 69 论文 1828

趋势信号

- CAMASA等新型基础设施级真实轨迹数据集（如CAM/DENM消息）被持续构建，替代合成或传感器中心型基准
- 扩散模型、流匹配（flow-matching）等生成范式被系统引入城市尺度轨迹生成，并在CityTrajBench统一基准下横向比较
- 多篇工作（DRIFT、RiskFlow、TraCS）均将‘风险约束’‘合规性塑形’或‘安全关键场景闭环生成’作为核心设计目标
- 神经符号化方法（如TraCS）与动力学感知后处理（如RiskFlow的车辆动力学约束）成为提升生成轨迹物理合理性的共性技术路径

核心观点

- 真实世界V2X通信数据（如CAM/DENM）比传统GPS轨迹更能刻画车路协同动力学，是构建高保真交通仿真与预测模型的关键基础
- 混合自主交通（mixed-autonomy）下的异质性（HVs/AVs共存）导致行为分布动态偏移，必须在生成与预测中显式建模该异质性
- 仅优化轨迹拟合度（如ADE/FDE）不足以保障交通生成的安全性与可部署性，需在生成过程中嵌入法规合规性、风险感知与物理可行性约束
- 统一评估协议（如CityTrajBench）已成为领域共识，用以消解因数据预处理、表示方式与指标不一致导致的性能误判

RESEARCH IDEA

轨迹方法跨城市迁移的首要失稳环节

轨迹方法迁移到另一座城市或极端天气场景后，最先失稳的通常不是模型结构，而是采样方式、路网约束和行为机制的变化。

为什么现在值得做：城市智能网联基础设施正加速部署，多地已开展RSU分阶段覆盖规划，亟需评估生成轨迹在不同覆盖密度下的空间鲁棒性；该结果可直接服务于交通仿真平台（如SUMO+CARLA联合仿真）中V2X感知模块的可信度校准。

关键难点

- 需从CAMASA原始消息中反演RSU地理坐标与瞬时有效覆盖半径，但论文未提供RSU部署元数据
- CityTrajBench的轨迹归一化流程默认全局坐标系一致性，未定义局部观测缺失条件下的插补/掩码策略
- CAM消息的时间戳精度与GNSS定位误差耦合，难以分离基础设施缺陷与个体运动建模误差

建议切入

- 基于CAMASA发布的路网拓扑与消息源ID映射，构建RSU空间可达性图，并以跳数距离量化各路段的平均观测强度
- 在CityTrajBench预处理环节引入‘观测置信度掩码’，将CAM消息密度低于阈值的时空片段标记为低置信区域，用于训练时加权损失与评估时分区统计
- 采用成分轨迹聚类（来自‘面向移动数据建模的成分轨迹聚类’）中的单纯形编码思想，将观测不确定性建模为道路类型兼容性分布，替代硬截断式缺失填充

REPRESENTATIVE ITEMS

ARXIV

CAMASA: 源自MASA Living Lab的基于CAM的数据集

轨迹预测是自动驾驶与协同驾驶系统的关键使能技术。然而，现有主流基准数据集大多以传感器为中心、地理范围受限，或基于合成移动轨迹，无法真实反映现实世界中车路协同（V2X）通信的动力学特性。

ARXIV

DRIFT: 面向混合自主交通生成的风险约束扩散模型与模仿先验

未来智能交通系统预计将长期演进为混合自主（mixed-autonomy）范式，即人类驾驶车辆（HVs）与自动驾驶车辆（AVs）在高度耦合的交通生态系统中共存。此类共存引发显著的异质性、加剧的不确定性以及日益复杂的车际交互动力学。在此背景下，如何同时刻画随AV渗透率动态变化而产生的异质行为分布偏移、在强车际耦合约束下生成多样且可执行的轨迹、并对罕见但高影响事件开展可靠的闭环安全与稳定性诊断，仍面临根本性挑战。

ARXIV

驾驶，快或慢？面向多模态地面移动的神经符号化运动预测引导

在包含行人、自行车、汽车和卡车等异构交通参与者的环境中，实现准确且可解释的运动预测，对安全自主导航至关重要。然而，当前最先进的方法仍主要为黑箱模型，缺乏对现实世界交通中法规性与行为性约束的显式建模。我们提出轨迹合规性塑形（Trajectory Compliance-Shaping, TraCS），一种神经符号化框架，通过可解释且概率化的谓词逻辑增强现有黑箱运动预测主干网络。

ARXIV

CityTrajBench: 面向城市尺度车辆轨迹生成的统一基准

城市轨迹生成是交通仿真、城市规划与移动性分析的一项基础任务。然而，由于现有研究常采用不同的数据集、预处理流程、轨迹表示方法及评估指标，轨迹生成方法间的系统性比较仍十分困难。这种碎片化使得难以判断所报告的性能差异究竟源于生成机制本身，还是源于实验协议的不一致。

UrbanComp Lab 学习资料库 (<https://research.urbancomp.dev/>)

近期研究聚焦于将复杂系统理论（尤其是多尺度动力学与网络脆弱性）深度嵌入城市韧性评估与地理模拟框架中，方法重心正从静态风险制图转向动态演化过程建模与跨灾种耦合分析。

近30天 24 近7天 1 来源 38 论文 170

趋势信号

- 多尺度路径发散 (MSPD) 等受重整化群启发的标量被提出，用于量化CA类地理模拟中局部转移律的异质性组织结构
- 高压输电网络研究构建统一框架，耦合灾害表征、脆弱性建模与宏观经济影响传播，实现九类单一灾害及复合灾害的可比性风险评估
- 城市洪涝与应急服务研究强调基础设施属性与地理空间特征的集成学习，以支撑精细化、可操作的风险表征
- 热浪事件频发推动城市气候适应能力讨论从环境议题转向对建筑与公共基础设施空间运行表现的实证关切

核心观点

- 城市韧性不能仅通过静态指标或单灾种分析评估，必须纳入多灾种交互、基础设施网络级联失效与宏观经济传播机制
- 地理模拟（如Cellular Automata）正从经验规则驱动转向具备可解释复杂度度量（如MSPD）的开放式演化建模
- 关键基础设施的空间分布与道路/电力/应急服务网络的拓扑特性共同构成城市韧性的结构性约束
- 气候极端事件（如热浪、洪涝）的集中发生已实质性重塑人类聚居空间的功能逻辑与规划前提

RESEARCH IDEA

MSPD度量在城市交通网络韧性模拟中失效于节点功能异质性建模

MSPD作为多尺度路径发散度量，在城市路网级灾害传播模拟中无法区分医院、消防站与普通交叉口的功能权重，因其仅依赖局部转移律异质性而忽略设施语义约束下的拓扑嵌入差异

为什么现在值得做：城市应急管理机构亟需可解释的、支持优先级排序的动态韧性指标，而非仅统计意义上的失效概率；OSMnx与开放建筑数据已支撑大规模语义化路网构建，使功能标注与动力学度量联合建模具备数据基础。

关键难点

- 需重新定义MSPD在异质图上的泛函形式，使其输入包含节点类型张量而非纯邻接矩阵
- OSM提取的设施语义存在粒度不一致问题（如‘医院’标签未区分急诊科与门诊部）
- 缺乏真实灾害下城市路网多时间步功能状态观测数据用于验证语义增强型MSPD

建议切入

- 基于R-GCN-VGAE在茨城县桥梁分类任务中构建的元路径模板，迁移定义‘道路-关键设施’二跳语义路径作为MSPD滑动窗口的约束边界
- 在MSPD滑动窗口估计器中嵌入节点类型掩码，使局部转移律异质性计算仅在同功能组内进行归一化
- 使用石家庄应急服务供需格局论文中的空间匹配结果，校准不同设施类型在路径中断时的服务替代弹性阈值

REPRESENTATIVE ITEMS

SUSTAINABLE CITIES AND SOCIETY

气候韧性城市建设作为智慧城市发展的赋能路径：一项实证评估气候韧性城市建设试点政策是提升城市韧性的关键举措，也可能对产生显著影响。

ARXIV

通过多尺度路径发散引导人工生命中的开放式进化

人工生命中的开放式进化 (OEE) 通常依赖于不可解释的黑箱神经网络复杂度度量，致使类生命系统与物理复杂性理论脱节。我们提出MSPD（多尺度路径发散，记为DP），一种受重整化群启发的标量，用于量化局部转移律异质性在时间维度上的多尺度组织结构。MSPD在种群层面定义为实际演化轨迹的泛函，并以滑动窗口有限分辨率估计器形式计算；其两种实现方式之间的一致性被表述为一条命题。

ARXIV

美国高压输电网络的多灾种风险比较评估

现代经济高度依赖高压输电网络，但该基础设施频繁遭受地震、洪水、龙卷风和地磁暴等自然灾害的破坏。传统风险评估通常孤立地分析各类灾害，因而缺乏统一基准以比较全灾种组合下的经济损失。本研究通过构建一个整合框架弥补这一空白，该框架耦合灾害表征、脆弱性建模与宏观经济影响传播模型。

ARCHDAILY

2026年世界环境日恰逢创纪录热浪，再度聚焦城市气候适应能力

随着欧洲遭遇近年来最早且最强烈的热浪之一，2026年世界环境日的到来，再度引发关于气候适应、城市韧性以及城市应对日益极端高温能力的讨论。葡萄牙、法国、意大利、西班牙、德国、瑞士、爱尔兰和英国多地气温远超季节性均值，促使各地发布高温预警、关闭学校、启动应急规划，并加剧了对建筑及公共基础设施在持续高温压力下运行表现的担忧。此类事件的集中发生凸显了一种日趋全球化的现实：气候变化已不再仅是环境议题，更从根本上重塑着人类居住、工作与聚集的空间。

UrbanComp Lab 学习资料库 (<https://research.urbancomp.dev/>)

近期研究重心从单纯依赖街景图像的静态表征，转向关注视觉对齐质量、人类感知机制建模及VLM输出的可靠性与社会协商性。方法上强调多模态融合（如眼动+图像、VLM+人类标注信度）、工具链开源化，以及将主观性、测量偏差和虚拟角色效应显式纳入建模框架。

近30天 18 近7天 4 来源 38 论文 169

趋势信号

- 出现面向纵向变化分析的视觉对齐工具（PairWise Image Finder），强调特征匹配与语义掩膜联合量化对齐质量
- 引入人类眼动追踪数据构建Place Pulse-Gaze数据集，将注视行为作为建模主观城市感知的显式输入
- VLM基准测试研究明确提出需报告inter-annotator reliability，并将标签空间与评分策略视为可协商的技术产物
- 多模态LLM代理研究系统考察虚拟角色（persona）对生成解释内容的影响，发现合理性说明存在社会经济与政治属性相关的系统性差异

核心观点

- 街景图像元数据不足以支撑可靠的的城市变化检测，视觉层面的对齐是开展精确纵向分析的前提
- 主观城市感知不能仅由图像端到端建模，必须嵌入人类感知过程（如注视行为）或其代理信号
- VLM在城市感知任务中的评估不能脱离人类判断的固有分歧与拒答行为，标注信度是解释模型一致性的必要参照
- 模型输出（尤其是解释性文本）具有可塑性，虚拟角色提示会系统性影响其评价性语言的焦点与倾向，但对基础感知标签影响微弱

RESEARCH IDEA

注视引导模型在非西方城市失效的主因

基于Place Pulse-Gaze数据集训练的注视引导城市感知框架，在首尔或重庆等非西方城市街景上预测安全性与宜人性时，因眼动模式与语义先验的地域错配而系统性低估低收入街区感知得分

为什么现在值得做：首尔、重庆等城市正推进基于街景AI的街道微更新决策支持系统，亟需可解释、低偏差的感知预测模块；多国城市计算团队已开源本地化街景采集协议与初步标注，为跨域验证提供基础数据接口。

关键难点

- 需复现并冻结Place Pulse-Gaze原始模型权重，排除训练数据量与架构差异干扰
- 须在目标城市完成严格控制曝光时长、视角与任务指令的眼动实验，而非直接迁移公开眼动数据
- 安全性与宜人性标签需采用与Place Pulse一致的二元成对比较范式，不可替换为Likert量表

建议切入

- 先复刻已有论文中的视觉指标，确认哪些变量在原始设定中真正起作用。
- 再补入人口、设施和可达性控制项，避免把社会经济差异误判成视觉效应。
- 最后在另一座城市做小规模外部验证，判断结论是否具有迁移性。

REPRESENTATIVE ITEMS

ARXIV

PairWise Image Finder

变化检测与场景识别技术已被广泛应用于街景影像（SVI），以理解跨年度场景的变化。然而，仅依赖元数据往往不足以可靠地找到视觉上对齐的图像对。本研究提出 PairWise Image Finder 工具，该工具融合特征检测与匹配，并借助语义分割掩膜来量化不同时期两幅图像之间的视觉对齐程度。

COMPUTERS, ENVIRONMENT AND URBAN SYSTEMS

天气有影响吗？基于街景图像的城市感知评估中的测量偏差探究

发表日期：2026年7月 来源：《计算机、环境与城市系统》，第127卷 作者：金东焕，李承敏，韩彩妍，金友贞，高奉宇，黄义正。

ARXIV

利用人类注视建模主观城市感知

城市感知描述了人们如何主观评估城市环境，从而塑造城市被体验与理解的方式。现有计算方法主要直接从街景图像建模城市感知，却在很大程度上忽略了形成此类判断所依赖的人类感知过程。本文提出 Place Pulse-Gaze 数据集，该数据集在街景图像基础上同步增加了眼动追踪记录及个体感知标签。

ARXIV

面向城市感知的视觉-语言模型基准测试应具备可靠性意识并经协商确立

视觉-语言模型（VLMs）正日益被用于生成街景图像的结构化描述，以支持街道环境评估、制图及公众咨询等任务。此类应用将可观属性与评价性类别相结合，其目标人群常表现为存在分歧与明确拒答的判断分布。

UrbanComp Lab 学习资料库 (<https://research.urbancomp.dev/>)

USDOT OPEN DATA

长期路面性能 (LTTP)

涵盖美国和加拿大逾2500个路面路段的长期路面性能、施工、交通及环境数据。十余种试验设计方案针对专门修建及既有沥青与混凝土路面，以及养护与修复策略展开。数据采集自1990年起持续进行。约三分之一的路面路段目前仍在研究中。当前正招募并修建新型温拌沥青混凝土路面加铺路段。

USDOT OPEN DATA

美国国家公路交通安全管理局缺陷调查办公室 (NHTSA-ODI)

录入美国国家公路交通安全管理局缺陷调查办公室 (NHTSA-ODI) 车辆所有者投诉数据库的投诉信息，将与其他数据源结合使用，以识别需开展调查的安全问题，并判定是否存在与安全相关的缺陷趋势。投诉信息还用于监测现有召回措施的覆盖范围是否恰当、整改措施是否充分。

ARCHDAILY

华润武汉光谷润景园商业更新 / 大川设计 + WTD纬图设计

2025年武汉荷叶山立交桥下空间更新项目，旨在高密度开发背景下，通过“集群嵌入”策略应对三大相互关联的城市更新挑战：① 多主体协同——消融社区发展、市政基础设施与运营主体之间的边界；唯有基于场地条件与多元利益相关方需求共同演化的组织机制，方能实现真正意义上的整合；② 范式转型——从“形式构图”转向“功能组织”，确保每栋建筑的布局、转角处理及尺度均针对邻近绿地、林地、高架道路、街道及社区所施加的具体约束作出差异化响应；③ 居住价值重构——摒弃物质堆砌与过度奢华，转向务实的公共性，引入紧凑、轻量、可运营的公共空间，推动居住边界公共化，并向城市开放。

SMART CITIES DIVE

美国城市自行车友好度提升：2026年报告

根据PeopleForBikes年度评级，自行车基础设施评分为50分及以上（满分100分）的城市数量增加了一倍多。

USDOT OPEN DATA

机动车乘员安全调查2007 (MVOSS) ——合并问卷 (版本1与版本2) (SAS)

该调查数据由相关机构用于追踪公众在乘员保护方面的态度、知识与行为的变化。MVOSS还收集与紧急医疗服务 (Emergency Medical Services) 及碰撞经历相关的信息。调查包含两份问卷，一份聚焦于安全带使用，另一份聚焦于儿童乘员保护。

WIRED

如何通过纽约市交通监控摄像头观看尼克斯队游行

艺术家Morry Kolman将利用纽约市交通摄像头的实时画面直播NBA冠军球队的彩带游行；此次，纽约市交通局并未要求其停止该行为。

SMART CITIES DIVE

马里兰州一城市考虑放宽分区限制，房地产经纪人协会研究报告强调其益处

马里兰州洛克维尔市 (Rockville) 一项“缺失的中等密度” (missing middle) 住房提案将使该市住房存量增加4.5%，当地房地产经纪人协会表示。

NASA NEWS

厄尔尼诺现象已开始

卫星对海面高度的观测表明，2026年事件于6月初持续增强。