

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

## THIS EDITION

五个方向的当日进展

# 冰岛馆在2027年威尼斯建筑双年展上探讨沐浴文化作为公民基础设施

本期《研究日报》头版聚焦城市基础设施的再定义与地理智能的演进。

冰岛馆将在第二十届国际建筑展览——威尼斯双年展上呈现《SOAK：集体归属感的仪式》展览，该展览从建筑、公共空间与社会互动视角审视冰岛的沐浴文化。

本项目由冰岛设计与建筑机构（Iceland Design and Architecture）委托，由玄武岩建筑事务所（Basalt Architects）合伙人马科斯·佐特斯（Marcos Zotes）策展，并由玄武岩建筑事务所、设计工作室Gagarin及艺术家拉恩·弗莱根林（Rán Flygenring）开展跨学科合作完成。

《SOAK》是继s.ap architects事务所作品《Lavaforming》代表冰岛参加2025年威尼斯建筑双年展之后，第二个通过公开征集程序遴选产生的冰岛参展项目。

编者按：本期头版呼应五大研究趋势，突出‘城市即生命体’的跨尺度认知转向：从冰岛沐浴仪式的公民性重构，到热浪冲击下的系统韧性建模；从V2G电网协同的能源智能，到COGENT对地理物理过程的连续模拟。所有条目均体现‘空间—社会—技术’三重耦合的深化。

## TREND OVERVIEW

趋势综述：城市作为生命系统：韧性、仪式与智能的交汇。

近期研究重心正从单一栅格模态的地理空间基础模型（GFM）转向融合矢量语义的多模态表征学习，并愈发关注模型在真实分布偏移下的鲁棒性与适配方法；方法论上强调栅格-矢量互补建模、跨域泛化能力评估及任务适配协议标准化。

近期研究聚焦于构建统一、可扩展的多模态遥感基础模型，方法重心从两两模态翻译转向场景中心联合建模与跨模态语义对齐；问题重心从模态融合性能提升转向应对不完整性、异质性及知识可解释性。

近期研究重心从单一轨迹建模转向支撑城市尺度交通智能体协同仿真与安全评估的高保真、基础设施感知型轨迹数据构建与生成；方法上强调闭环动力学一致性、地图与车辆运动学约束嵌入，以及生成过程的可控性与计算效率。

## DIRECTION PULSE

### 1 地理大模型与地理智能体

近期研究重心正从单一栅格模态的地理空间基础模型（GFM）转向融合矢量语义的多模态表征学习，并愈发关注模型在真实分布偏移下的鲁棒性与适配方法；方法论上强调栅格-矢量互补建模、跨域泛化能力评估及任务适配协议标准化。

### 2 多源多模态地理数据

近期研究聚焦于构建统一、可扩展的多模态遥感基础模型，方法重心从两两模态翻译转向场景中心联合建模与跨模态语义对齐；问题重心从模态融合性能提升转向应对不完整性、异质性及知识可解释性。

### 3 轨迹数据与城市交通研究

近期研究重心从单一轨迹建模转向支撑城市尺度交通智能体协同仿真与安全评估的高保真、基础设施感知型轨迹数据构建与生成；方法上强调闭环动力学一致性、地图与车辆运动学约束嵌入，以及生成过程的可控性与计算效率。

### 4 复杂网络、韧性城市与地理模拟

近期研究聚焦于极端气候事件（如热浪）对城市物理与基础设施系统的冲击，方法重心正从静态脆弱性评估转向耦合多灾种、动态空间过程与系统级传播效应的整合建模。

### 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端到端建模，转向强调视觉对齐可靠性、人类感知机制建模及VLM评估的测量严谨性；方法上日益融合多模态信号（眼动、语义掩膜、人类标注分歧）与可解释性设计。

## HIGHLIGHTS

- 冰岛馆以沐浴文化为切口，重思公共基础设施的社会仪式维度。
- 地拉那国会宫改造标志后意识形态空间的再功能化实践。
- 通用汽车提出电动汽车作为动态能源节点，应对AI算力扩张。
- COGENT模型实现不规则地理网格上的长期物理过程连续模拟。

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

近期研究重心正从单一栅格模态的地理空间基础模型 (GFM) 转向融合矢量语义的多模态表征学习, 并愈发关注模型在真实分布偏移下的鲁棒性与适配方法; 方法论上强调栅格-矢量互补建模、跨域泛化能力评估及任务适配协议标准化。

近30天 188 近7天 42 来源 57 论文 822

#### 趋势信号

- 多篇论文明确指出当前GFM主流仍局限于栅格模态, 而OpenStreetMap/Overture等矢量数据被反复强调为未被充分利用的关键结构化信息源
- EarthShift基准首次系统化GFM在地理、时间、传感器等真实分布偏移下的性能下降 (15 - 20%), 成为鲁棒性评估新焦点
- 至少4篇论文提及GFM领域缺乏统一评估标准、训练协议与模型权重公开规范, 导致跨研究结果不可比、复现困难
- DarkVesselNet等应用案例显示GFM主干正被嵌入端到端多模态智能体栈 (如 SAR+光学+AIS+轨迹推理), 体现从‘模型’向‘可部署智能体’演进迹象

#### 核心观点

- 栅格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, 后者编码离散对象的几何、拓扑与语义关系, 且更贴近人类认知与GIS实践
- 地理空间基础模型 (GFM) 的核心价值在于支持高风险EO任务 (如灾害响应、野火制图、暗船检测), 但其实际效用受限于下游适配方法 (如LoRA) 与域偏移鲁棒性
- 当前GFM研究存在严重社区标准缺失: 评估基准不统一、预训练配置不可复现、模型权重公开率低 (39%论文未公开), 阻碍技术可信演进
- 多模态融合不仅是数据拼接, 更需应对成像物理差异、模态异质性、语义鸿沟及分布偏移等根本性挑战, 对齐与知识迁移机制是关键技术瓶颈

## RESEARCH IDEA

### 地理基础模型跨场景迁移的关键瓶颈

地理基础模型迁移到新城市、新尺度或新数据源时, 关键瓶颈通常来自空间先验不足、标注差异和工具调用能力不稳定。

为什么现在值得做: 超越像素的空间表征学习: 融合栅格数据与矢量语义以构建以人为中心的地理空间基础模型 与 EarthShift: 面向地球观测中真实世界分布偏移鲁棒性的基准测试 已经提供了可复用的变量、数据或模型入口, 这使得问题不再停留在概念层面, 可以直接构造成小规模验证。

#### 关键难点

- 基础模型表征很强, 但如何落到可解释的地理任务指标上并不直接。
- 模型迁移成功时, 很难判断收益来自通用语义能力还是地理先验。

#### 建议切入

- 先选一个边界清楚的地理任务做轻量适配, 避免一开始铺得过大。
- 再把评价指标改成更符合 GIS 任务的空间约束和误差口径, 确认模型真正提升了什么。
- 最后把模型输出拆成检索、推理和空间约束三个环节看误差来源。

## REPRESENTATIVE ITEMS

### 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

无人知晓地理空间基础模型 (GFM) 的当前技术前沿

地理空间基础模型 (Geospatial Foundation Models, GFM) 被提出作为灾害响应、土地覆被制图、粮食安全监测及其他高风险地球观测任务的通用化骨干模型。然而, 现有已发表的相关研究未能向评审者或用户提供足够信息, 以判断何种模型适用于特定任务。我们认为, 目前尚无人确切知晓GFM的技术前沿究竟为何。

### ARXIV

利用Sentinel-2数据对地理空间基础模型进行低秩自适应以实现野火过火区制图对于灾损评估、排放建模以及理解不同生态区域中火灾与气候的相互作用至关重要。近期提出的地理空间基础模型 (Geospatial Foundation Models, GFM) 为卫星影像提供了强大的通用表征能力, 但目前尚缺乏关于如何高效地将此类模型适配至下游地球观测任务的明确共识, 尤其在面临地理与时间域偏移 (geographic and temporal domain shift) 时。

### ARXIV

DarkVesselNet: 面向暗船检测的多模态遥感与航迹推理方法

暗船检测需融合船舶通过AIS上报的信息与卫星通过雷达和光学传感器观测到的信息。DarkVesselNet是一种多模态遥感技术栈, 整合了Sentinel-1 SAR数据、Sentinel-2光学影像、地理空间基础模型主干网络、AIS航迹推理、TGARD风格的间隙检测, 以及受Pi-DPM启发的异常检测头。该代码库以经过测试的Python软件包及公开的Hugging Face Space形式发布。

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

近期研究聚焦于构建统一、可扩展的多模态遥感基础模型，方法重心从两两模态翻译转向场景中心联合建模与跨模态语义对齐；问题重心从模态融合性能提升转向应对不完整、异质性及知识可解释性。

近30天 260 近7天 59 来源 56 论文 1185

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介实现五模态任意翻译
- Earth-OneVision 在单一自回归框架中统一六类传感器模态与九类任务，并设计全粒度视觉-语言对齐 (FGVLA) 等专用机制
- TSMNet 引入双分支文本编码器，显式融合物体级标签与场景级语义特征以支持开放词汇分割
- SGMA 框架针对不完整多模态数据 (IMSS) 提出语义引导融合 (SGF) 与模态感知解耦模块，直面跨模态类内差异与异质性

核心观点

- 多模态遥感建模正从‘外观映射’范式转向‘场景内容中心’范式，强调底层语义一致性而非像素级对齐
- 文本模态作为先验知识源被系统性引入，用于弥合视觉表征与现实概念间的语义鸿沟，支撑开放词汇与可解释决策
- 不完整 (IMSS)、跨模态异质性与类内差异是多模态遥感落地的核心瓶颈，现有方法易因过度对齐而牺牲模态特异性
- 统一架构 (如自回归大模型、解耦式生成框架) 成为主流技术路径，旨在提升多模态任务的通用性、可扩展性与跨传感器泛化能力

## RESEARCH IDEA

### 方法迁移到新场景后的失效边界

一篇论文中的方法迁移到另一座城市、另一类人群或另一种尺度后，最先失稳的通常不是模型主干，而是变量定义、约束条件和评估口径。

为什么现在值得做：城市应急响应与军事侦察场景亟需在部分传感器失效时维持LULC制图稳定性，而现有IMSS方法缺乏面向物理差异的失效归因分析；Earth-OneVision提出的PCMA机制与SGMA的SGF模块存在可比接口，支持模块级替换实验。

关键难点

- 需构造SAR-光学联合缺失的可控仿真数据集，不能仅依赖真实缺失样本
- 需量化同一地物在SAR与光学图像中的形变向量场，现有遥感配准工具不支持跨物理模态形变建模
- SGMA原代码未开放SGF模块的中间原型可视化接口，须逆向工程其类别原型生成逻辑

建议切入

- 复现SGMA并在EarthMM数据集上注入SAR-光学联合缺失掩码，固定其余模态输入
- 基于SAR成像方程与BRDF模型合成同场景下SAR与光学图像对，提取对应像素级形变残差作为监督信号
- 将Earth-OneVision的PCMA中物理差异分解模块接入SGMA的SGF前端，替换原始原型聚合方式

## 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/>)

近期研究重心从单一轨迹建模转向支撑城市尺度交通智能体协同仿真与安全评估的高保真、基础设施感知型轨迹数据构建与生成；方法上强调闭环动力学一致性、地图与车辆运动学约束嵌入，以及生成过程的可控性与计算效率。

近30天 445 近7天 128 来源 68 论文 1711

#### 趋势信号

- 新数据集（如CAMASA、SF-LIFE）强调真实V2X通信消息（CAM/DENM）或基于GTFS/OSM的多模态智能体级仿真，而非仅GPS轨迹采集
- 生成方法（如RiskFlow）显式规避扩散模型的迭代去噪缺陷，转向单步前向的动作空间概率流建模
- 基准建设（如CityTrajBench）聚焦统一协议以解决轨迹生成评估碎片化问题，覆盖VAE/GAN/diffusion/flow-matching等异构生成器
- 奖励设计（如MBRF）从传统拥堵惩罚转向动量引导，强调通行连续性与排放-通行量权衡

#### 核心观点

- 真实世界轨迹建模必须耦合基础设施语义（如路网、信号相位、V2X协议）与车辆动力学约束，否则生成结果易失真
- 城市尺度轨迹生成的有效性评估高度依赖标准化基准，当前性能差异常源于实验协议不一致而非模型本质优劣
- 安全关键场景生成的核心挑战在于闭环可控性与物理合理性之间的平衡，而非单纯提升似然或多样性
- 轨迹数据增强的价值不仅取决于扰动方式，更取决于被增强轨迹的选择策略——离群性、多样性与不确定性等系统化准则显著优于随机选择

### 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：CAMASA是首个大规模真实城市V2X基础设施轨迹数据集，填补了车路协同场景下高时间分辨率轨迹数据空白；交通仿真与自动驾驶验证亟需能复现协同感知动力学的合成轨迹，以支撑安全关键场景测试，而当前所有生成模型均缺乏对消息时序一致性的建模能力。

#### 关键难点

- CAMASA中CAM消息包含位置、速度、加速度、航向角及时间戳等多维异步字段，需定义跨字段的联合时序一致性约束
- TrajDLM的离散路段token化方案与CAM消息的连续状态流存在语义鸿沟，无法直接映射消息状态变化率
- 缺乏V2X通信协议（ETSI TS 102 637）驱动的轨迹合理性评估指标，现有CityTrajBench指标不适用

#### 建议切入

- 将CAMASA原始消息流按车辆ID与时间戳对齐，构建每辆车的多变量状态时间序列（含消息发送间隔、状态更新延迟等协议特征）
- 修改TrajDLM的分块扩散主干，使其输入为连续状态序列而非路段ID序列，并嵌入ETSI协议定义的状态更新规则作为硬约束
- 设计协议一致性损失项：对生成序列强制满足CAM消息中相邻帧间最大加速度/角速度变化阈值（依据ETSI标准）

### 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

RiskFlow: 快速且保真的安全关键交通场景生成方法

安全关键交通场景生成对于在罕见但高风险交互下评估自动驾驶系统至关重要。现有基于扩散模型的方法在闭环生成中具备较强的可控性，但其迭代去噪过程计算开销大，且在长时序展开过程中易累积采样误差与引导误差，导致运动失真，例如抖动、异常加速度及驶离道路行为。为解决上述问题，我们提出 RiskFlow——一种面向安全关键多智能体交通场景的闭环生成框架，将未来轨迹生成建模为动作空间中的概率流传输。

#### 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。

#### ARXIV

基于动量的奖励设计用于低排放交通信号控制

城市交通拥堵是一个日益严重的全球性问题，显著加剧了通勤时间延长与环境污染。传统交通信号控制系统往往难以适应动态变化的交通状况。自适应交通信号控制可在不改变道路基础设施的前提下改善城市交通。

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

近期研究聚焦于极端气候事件（如热浪）对城市物理与基础设施系统的冲击，方法重心正从静态脆弱性评估转向耦合多灾种、动态空间过程与系统级传播效应的整合建模。

近30天 35 近7天 8 来源 38 论文 169

趋势信号

- 2026年欧洲多地破纪录热浪事件被作为典型场景，触发对建筑与公共基础设施高温耐受性的实证关切
- arXiv新发布的OVO-S-Bench基准明确将‘空间模拟与推理’列为多模态智能体能力的第四层级，强调流式视频输入下的动态GeoSimulation需求
- 美国高压输电网络研究构建了耦合灾害表征—脆弱性建模—宏观经济影响传播的统一框架，突破单一灾种孤立分析范式
- Nature Cities论文指出城市土壤修复的成本维度‘迄今尚未得到系统研究’，揭示韧性建设中生态基础设施经济可及性的知识缺口

核心观点

- 城市韧性已不再仅指环境适应能力，而是表现为复杂网络（能源、交通、土壤、建筑）在多重扰动下维持功能与恢复能力的系统属性
- 地理模拟（GeoSimulation）正从传统CellularAutomata等离散模型，向支持流式感知、时空上下文追踪与外源性空间映射（allocentric mapping）的动态智能体范式演进
- 气候适应政策（如韧性城市试点）需与微观基础设施性能（如土壤生态功能、电网抗灾能力）和宏观经济影响传导机制形成闭环验证
- 多灾种风险评估缺乏统一基准是当前关键方法论瓶颈，亟需跨灾害类型、跨尺度、跨部门的数据—模型—决策链整合

## RESEARCH IDEA

### HME-CA模型在热浪胁迫下无法表征人口迁移的临界响应阈值

基于增强人类移动性的元胞自动机（HME-CA）模型在模拟热浪期间人口空间重分布时，因未嵌入热应激生理阈值与避暑设施可达性约束，导致其迁移路径与驻留时长预测在35°C以上连续高温条件下系统性偏离真实GPS轨迹数据。

为什么现在值得做：OVO-S-Bench：面向多模态大语言模型的流式空间智能分层基准与基于增强人类移动性的元胞自动机（HME-CA）模型对城市群开展经济与人口模拟已经提供了可复用的变量、数据或模型入口，同时近期资讯说明现实需求已经出现，适合把问题往应用场景推进。

关键难点

- 需将UTCI网格与HME-CA格网进行亚像元级对齐，避免热暴露指标在离散元胞中平均化失真
- 需定义热应激触发迁移的个体异质性参数（如年龄、基础疾病），但当前缺乏城市级分人群热响应实证标定数据
- HME-CA原框架不支持事件驱动型状态切换（如‘热警报发布→启动避暑迁移’），需重构状态转移矩阵

建议切入

- 先把原论文任务拆成预测、识别或匹配等可比较子任务，明确误差发生在哪一层。
- 再选一类公开轨迹场景做跨城市或跨系统复现，判断模型最先失效的条件。
- 最后把误差与路网结构、采样方式和出行约束对齐，确认问题不是预处理造成的。

## REPRESENTATIVE ITEMS

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。

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。

NATURE CITIES

深入探究城市土壤修复的成本

城市化导致土壤压实、污染和封盖，削弱了土壤的生态功能及对城市韧性至关重要的生态系统服务，例如洪水缓解与气候调节。因此，城市土壤的生态修复对于建设可持续且适应气候变化的城市至关重要。然而，城市土壤修复的经济维度，尤其是其成本，迄今尚未得到系统研究。

ARXIV

OVO-S-Bench：面向多模态大语言模型的流式空间智能分层基准机器人、增强现实与自动驾驶中的多模态智能体需基于连续的第一人称视角视频流推理场景与空间布局，且常需依赖当前视野之外的证据。现有基准或针对完整视频进行离线评估，或聚焦于事件识别而非空间结构理解。我们提出OVO-S-Bench——一个完全由人工标注的流式空间智能基准，涵盖348个源视频上的1,680个问题。

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端到端建模，转向强调视觉对齐可靠性、人类感知机制建模及VLM评估的测量严谨性；方法上日益融合多模态信号（眼动、语义掩膜、人类标注分歧）与可解释性设计。

近30天 18 近7天 5 来源 38 论文 164

趋势信号

- 出现开源工具PairWise Image Finder，聚焦街景图像对的视觉对齐量化而非元数据匹配
- 新数据集Place Pulse-Gaze引入眼动追踪与感知标签同步采集，将人类注视行为显式建模为城市感知生成的关键中间变量
- VLM基准测试研究明确提出须将inter-annotator reliability与拒答率纳入评估指标，反对将人类判断简化为单一真值
- 虚拟角色提示（persona prompting）被系统用于探查multimodal LLM在城市感知解释中的社会属性偏差，而非仅优化准确性

核心观点

- 街景图像的空间可比性前提在于视觉对齐，元数据匹配不可替代像素级对齐验证
- 城市感知是主观建构过程，其计算建模需锚定人类感知机制（如注视行为），而非仅拟合静态图像统计模式
- VLM在城市感知任务中的有效性必须置于人类判断分布（含分歧与弃答）中评估，标签空间本身具有协商性
- 多模态大语言模型生成的城市感知解释存在隐性社会角色偏差，尤其体现在合理性说明而非感知标签层面

## RESEARCH IDEA

### PairWise对齐在非结构化旧城街景中失效

PairWise Image Finder 在缺乏规则路网与建筑立面统一性的中国历史街区街景中，因特征检测器对杂乱招牌、临时摊贩与非刚性遮挡物响应不稳定，导致匹配关键特征比例低于0.3且语义掩膜对齐度标准差超0.22，无法支撑纵向变化量化分析。

为什么现在值得做：住建部‘城市更新行动’要求对1200+历史街区开展微更新效果评估，亟需可靠纵向街景对比工具；开源工具可直接部署于国产GPU集群，无需标注即支持无监督对齐质量筛选。

关键难点

- 同一套街景指标在不同城市的拍摄时段、道路尺度和绿视率条件下可能不稳定。
- 视觉特征与真实行为或健康结果之间隔着社会经济背景和空间选择机制，不能直接等同。

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。此类应用将可观测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

##### 公路性能监测系统 (HPMS) ——2011年密苏里州数据

HPMS 汇集有关公路路网规模、使用情况、技术状况及运行性能的数据。该系统由一个具备地理空间功能的数据库构成，用于生成报告并提供数据分析工具。HPMS 数据被美国交通部 (U.S. DOT)、行政管理部、国会以及交通运输界众多利益相关方所使用。

#### USDOT OPEN DATA

##### 交通运输业就业——公共交通与地面旅客运输

就业人员指16岁及以上、属于民用非机构人口、且在调查参考周内满足以下任一条件的人员：作为受薪雇员从事任何工作；自营职业、专业服务或经营自家农场；或作为家庭成员所经营企业中的无薪工作者工作15小时及以上；以及虽未工作但拥有工作岗位或自营生意、仅暂时缺勤者。美国劳工统计局 (BLS) 通过《当前人口调查》 (CPS) 发布非农工资就业的行业估计数据。

#### SMART CITIES DIVE

##### 世界杯揭示城市日益加剧的高温风险

极端高温与野火烟雾正威胁运动员与观众的健康，科学家指出，这凸显了为应对气候变暖而改造基础设施与公共安全预案的必要性。

#### ARXIV

##### COGENT: 面向长期物理预测的基于神经常微分方程的连续图模拟器

本文提出 COGENT，一种面向不规则地理空间网格上长期物理预测的、基于神经常微分方程 (Neural ODE) 的连续图模拟器。COGENT 利用基于图的历史编码器，对系统状态的有限历史序列及相应的强迫场与外部强迫进行编码，生成节点级上下文向量，以同时刻画局部空间相互作用与时间演化特征。这些上下文向量用于初始化并调节一个潜在神经常微分方程，其动力学由插值后的未来强迫项及显式的相对滚动时间驱动。

#### USDOT OPEN DATA

##### 航空旅行——国内——经季节性调整

美国运输统计局根据美国商业航空公司提交的月度报告，发布经季节性调整的航空客运量数据。

#### ARCHDAILY

##### 赫尔佐格与德梅隆将改造地拉那共产主义时期国会宫

2026年6月3日，赫尔佐格与德梅隆 (Herzog & de Meuron) 被选定负责阿尔巴尼亚地拉那国会宫 (Pallati i Kongreseve) 的改造项目。该项目协作方包括朱利安·贝基里 (Julian Beqiri)、玛赛拉·德马伊 (Marsela Demaj)、米歇尔·德维涅景观事务所 (Michel Desvigne Paysagistes, MDP)、奥雅纳 (ARUP)、LDK、根蒂安·什库尔蒂 (Gentian Shkurti)、SUEB Industries sh.p.k.、The Space Factory Ltd、MBBM以及KLAR sh.p.k.。

#### THE VERGE AI

##### 通用汽车认为电动汽车可通过车网互动 (V2G) 技术抵消人工智能的能耗

今日在旧金山举行的一场活动中，通用汽车 (General Motors) 围绕电动汽车电池、能源存储及电网韧性发布了一系列公告，以应对人工智能数据中心日益增长的电力需求。该公司宣布将为其现有电动汽车及家庭能源客户启用新的车网互动 (vehicle-to-grid, V2G) 功能。

#### NATURE CITIES

##### 人迹罕至之路交汇于彭博城市实验室2026

城市必须创新，并日益在全球节奏日益紧迫的背景下发挥引领作用。《自然·城市》 (Nature Cities) 参加了城市实验室2026会议，与共同直面挑战的领导和开拓者一道学习、共同探索。