

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

## THIS EDITION

五个方向的当日进展

# 亚马逊员工呼吁西雅图暂停新建数据中心

当数据中心扩张遭遇社区否决，当洪灾数据库延伸至2024年，地理大模型正从实验室走向市政厅与街角。

本周二，西雅图市议会将就是否实施为期一年的新建数据中心禁令进行投票——此举距多家公司提出在该市建设五座大型数据中心仅过去两个月。

该禁令最坚定的支持者之一，是来自本市最大科技企业亚马逊的现任员工，他们与其他人士共同出席听证会并作证。

编者按：本期头版聚焦地理智能体（Geo-Agent）范式在真实城市场景中的落地张力：它既驱动AI气候规划工具上线，也直面西雅图数据中心禁令这类社会技术争议；既依托FFEM-DB 2.0等高时效性灾害数据集，也需回应PREVI Lima所揭示的居民主体性这一深层空间政治命题。

## TREND OVERVIEW

趋势综述：地理智能体崛起：城市空间的多模态重构。

近期研究重心正从单模态栅格基础模型转向融合矢量语义的多模态地理空间基础模型（GeoMultimodal GFMs），同时显著加强对分布外鲁棒性、适配方法论及社区评估规范的系统性反思。

近期研究聚焦于在模态缺失、配对稀缺、语义鸿沟等现实约束下构建统一、可扩展、语义一致的多模态联合建模框架；方法重心正从两两模态翻译转向场景中心表征解耦、文本-视觉协同监督与模态感知动态融合。

近期研究重心正从单纯提升轨迹生成保真度，转向兼顾计算效率、路网拓扑约束、安全关键性与多智能体交互真实性的闭环建模；方法上扩散模型、流匹配与语言模型范式被系统性引入，但均需适配交通物理与城市空间逻辑。

## DIRECTION PULSE

### 1 地理大模型与地理智能体

近期研究重心正从单模态栅格基础模型转向融合矢量语义的多模态地理空间基础模型（GeoMultimodal GFMs），同时显著加强对分布外鲁棒性、适配方法论及社区评估规范的系统性反思。

### 2 多源多模态地理数据

近期研究聚焦于在模态缺失、配对稀缺、语义鸿沟等现实约束下构建统一、可扩展、语义一致的多模态联合建模框架；方法重心正从两两模态翻译转向场景中心表征解耦、文本-视觉协同监督与模态感知动态融合。

### 3 轨迹数据与城市交通研究

近期研究重心正从单纯提升轨迹生成保真度，转向兼顾计算效率、路网拓扑约束、安全关键性与多智能体交互真实性的闭环建模；方法上扩散模型、流匹配与语言模型范式被系统性引入，但均需适配交通物理与城市空间逻辑。

### 4 复杂网络、韧性城市与地理模拟

近期研究聚焦于将气候极端事件（如热浪、多灾种耦合）作为压力测试场景，驱动城市基础设施与空间系统的韧性建模；方法重心正从静态网络分析转向融合流式感知、时空推理与外源性空间映射的动态地理模拟。

### 5 城市感知、街景感知与空间优化

近期研究重心从单纯图像识别转向对齐性、主观性与测量可靠性的系统性建模；方法上强调多模态融合（尤其是注视、VLM、专家规则与行为信号）、视觉对齐工具开发，以及将人类判断的分歧与拒答纳入评估框架。

## HIGHLIGHTS

- 亚马逊员工公开支持西雅图数据中心建设禁令，技术部署遭遇在地化伦理审议。
- FFEM-DB 2.0更新至2024年，覆盖欧洲—地中海16个行政区域洪灾致死事件。
- HU-EnviroGrids发布匈牙利全国尺度栅格化环境数据集，支撑空间建模。
- SpatialWorld基准提出交互式空间推理新标准，突破静态视觉问答局限。

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

近期研究重心正从单模态栅格基础模型转向融合矢量语义的多模态地理空间基础模型 (GeoMultimodal GFM)，同时显著加强对分布外鲁棒性、适配方法论及社区评估规范的系统性反思。

近30天 176 近7天 36 来源 57 论文 811

趋势信号

- 多篇论文明确指出当前EOFMs/GFM过度依赖栅格EO数据，而OpenStreetMap、Overture等矢量数据源的结构化语义尚未被有效纳入主干建模
- DarkVesselNet等应用案例体现GFM正作为可插拔主干网络嵌入端到端遥感分析栈，与AIS航迹推理、TGARD间隙检测等传统GIS分析模块协同工作
- 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平均性能下降15-20%，凸显鲁棒性已成为核心评估维度
- 对152篇GFM论文的审计发现：46处跨论文结果不可比、94篇预训练配置唯一、39%未公开模型权重，暴露社区标准严重缺失

核心观点

- 栅格与矢量数据构成地理空间的互补表征：栅格捕获连续物理模式，矢量编码离散对象及其拓扑-语义关系，二者融合是构建以人为中心地理智能体的关键前提
- 当前GFM研究面临‘技术前沿不可知’困境——缺乏统一评估协议、预训练控制标准、模型权重公开义务及可复现的基准测试框架
- GFM在分布内任务上表现强劲，但在真实部署场景（新区域、新时段、新传感器）下存在系统性性能衰减，分布鲁棒性不能仅靠扩大数据或模型规模解决
- 多模态融合不是简单特征拼接，需针对性应对成像物理差异、模态异质性、语义鸿沟及地理域偏移等固有挑战

## RESEARCH IDEA

### 矢量语义注入如何影响GFM在跨城市交通噪声建模中的泛化性

将OpenStreetMap矢量语义显式注入地理空间基础模型主干后，其在AlphaEarth支持的跨城市交通噪声建模任务中，在非欧洲城市出现平均3.2分以上性能下降，因为矢量拓扑关系与本地路网密度、混合用地强度及声传播物理约束之间缺乏可微分对齐机制

为什么现在值得做：城市规划部门亟需可迁移的噪声不平等分析工具以支撑公平性政策制定，而当前开源GFM权重（如AlphaEarth）与OSM矢量数据均具备即用条件；EarthShift已提供跨城市分布偏移测试框架，使该问题可在标准化协议下量化验证。

关键难点

- 需构造矢量-物理耦合损失项，将OSM道路等级、建筑高度缓冲区、植被覆盖率等离散矢量属性映射至声衰减方程参数
- AlphaEarth原始权重未公开微调接口，须逆向解析其Hugging Face Space部署的tokenization与head结构
- 交通噪声真值数据在非欧美城市严重稀缺，无法直接监督矢量注入后的特征空间校准

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

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。

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。

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

近期研究聚焦于在模态缺失、配对稀缺、语义鸿沟等现实约束下构建统一、可扩展、语义一致的多模态联合建模框架；方法重心正从两两模态翻译转向场景中心表征解耦、文本-视觉协同监督与模态感知动态融合。

近30天 253 近7天 53 来源 56 论文 1172

趋势信号

- MetaEarth-MM 提出以底层场景内容为中介的解耦式生成范式，支持五模态任意翻译，替代传统两两映射
- TSMNet 显式引入双分支文本编码器，分别建模物体级标签与场景级语义，并通过文本引导视觉语义融合模块实现跨模态交互
- SGMA 针对不完整多模态语义分割 (IMSS) 提出语义引导融合 (SGF) 模块，利用类别相关语义原型对齐跨模态特征
- SMART-HC-VQA 将遥感活动检测重构为时空延展型 VQA 任务，将地理元数据、时间阶段与观测关系编码为自然语言问答三元组

核心观点

- 多模态遥感建模的核心挑战已从模态互补性挖掘转向场景一致性维持与模态特异性保留之间的平衡
- 文本模态不再仅作为辅助监督信号，而是被系统性用于弥合视觉表征与现实地理概念间的语义鸿沟
- 模态缺失是实际部署中的常态而非异常，因此鲁棒性设计（如模态感知、语义引导、非对称融合）已成为多模态架构的必要组件
- 统一多模态基础模型需具备解耦式结构：先推断共享场景表征，再条件化生成/推理目标模态，而非端到端外观映射

## 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：城市精细化治理亟需亚像元级地物边界支持，而Sentinel-1/2协同观测在建成区普遍存在单模态缺失；EarthMM与GeoHeight-Bench等新数据集已提供多分辨率、多几何约束的配对样本，支撑对空间关系与垂直梯度的联合建模。

关键难点

- 需在SGMA中嵌入可微分的多尺度空间关系建模模块，而非简单替换注意力机制
- 需构建SAR-光学边缘梯度响应差异的物理驱动标注协议，不能仅依赖人工勾绘边界
- 需在不增加模态专用参数的前提下，实现梯度响应差异的即插即用校准

建议切入

- 复现SGMA在LoveDA-City子集（含SAR+光学缺失组合）上的边界IoU退化曲线，定位失效频段（如0.5 - 2m空间频率）
- 基于Radarsat-2与Sentinel-2同步观测，构建边缘梯度矢量场差异图谱，量化模态间法向梯度偏移角与幅值比分布
- 设计梯度感知空间关系模块 (GSRM)，将梯度差异图谱作为软约束注入SGF的原型匹配过程，替代原始余弦相似度

## 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

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。

INFORMATION FUSION

一种受脑启发的低照度与红外多模态遥感检测网络用于夜间人员搜救

出版日期：2026年6月1日在线发表；来源：《Information Fusion》；作者：刘世宇、万雨婷、李克帆、马爱龙、钟燕飞。

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

近期研究重心正从单纯提升轨迹生成保真度，转向兼顾计算效率、路网拓扑约束、安全关键性与多智能体交互真实性的闭环建模；方法上扩散模型、流匹配与语言模型范式被系统性引入，但均需适配交通物理与城市空间逻辑。

近30天 435 近7天 116 来源 67 论文 1685

#### 趋势信号

- RiskFlow 提出用概率流传输替代迭代去噪，以单次前向传播生成加速度与横摆角速度指令，解决长时序运动失真问题
- CityTrajBench 建立统一基准协议，强制标准化数据接入、地图感知后处理与多层级评估，直指方法比较的碎片化困境
- TrajDLM 将轨迹建模为离散路段序列，融合路网编码器嵌入与拓扑约束采样，弥合连续空间生成效率与拓扑保真度之间的鸿沟
- SF-LIFE 发布含3.024万亿条记录、覆盖50万智能体70天多模态轨迹的高保真模拟数据集，明确依托GTFS与OpenStreetMap构建真实公共交通运动学

#### 核心观点

- 轨迹生成不再仅追求统计相似性，而必须满足车辆动力学约束、道路可行驶性及安全关键场景的可控引导
- 城市尺度轨迹建模亟需统一基准（如CityTrajBench）以剥离实验协议差异，使模型性能归因于方法本质而非工程选择
- 大规模真实轨迹数据稀缺且受隐私限制，高保真模拟数据集（如SF-LIFE）与基于被动GPS的群体行为推断（如EV充电活动分析）成为关键替代路径
- 奖励函数设计正从拥堵惩罚转向通行激励（如MBRF），强调交通信号控制中通行量与排放的权衡稳定性，反映对系统性目标的再定义

### 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：CityTrajBench已提供跨城市数据接入协议与地图感知后处理接口，支持快速部署至非网格城市；交通规划部门亟需在非标准路网中开展what-if仿真，而当前合成轨迹无法支撑高保真路权分配与冲突点识别。

#### 关键难点

- 需构建路段长度分布偏度量化指标以表征路网异质性，不能仅依赖OSM拓扑类型标签
- TrajDLM的拓扑约束采样机制未定义路段长度权重，无法直接适配变长路段序列
- 缺乏公开的非网格城市高精度轨迹-路网对齐标注数据用于验证连贯性退化程度

#### 建议切入

- 在SF-LIFE提供的旧金山路网基础上，人工构造三类控制变量路网：等长路段网格、指数分布路段长度的放射状路网、对数正态分布路段长度的历史中心路网
- 复现TrajDLM并替换原始分块逻辑为基于路段累积长度的动态窗口切分，验证其是否缓解连贯性下降
- 使用CityTrajBench的统一评估协议，在新构造路网上运行TrajDLM与基线模型，聚焦‘路段跳变率’与‘路径可行性比率’两个可解释指标

### REPRESENTATIVE ITEMS

#### 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。

#### ARXIV

利用大规模GPS数据揭示电动汽车驾驶员在充电会话之外的活动模式

准确理解电动汽车（EV）驾驶员行为对于长期基础设施规划、电网管理及评估下游经济影响至关重要，但目前仍缺乏个体层面的EV出行数据。本文构建了一个可扩展的分析框架，基于被动采集的高分辨率移动轨迹数据，覆盖美国四大主要都市区逾76万名驾驶员，推断EV拥有状况与充电行为。我们依据驾驶员对充电站与加油站的访问模式差异、访问频次及日常出行特征识别潜在EV驾驶员，并利用各邮政编码区域（zip code）的汇总EV注册统计数据校准推断群体规模。

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

近期研究聚焦于将气候极端事件（如热浪、多灾种耦合）作为压力测试场景，驱动城市基础设施与空间系统的韧性建模；方法重心正从静态网络分析转向融合流式感知、时空推理与外源性空间映射的动态地理模拟。

近30天 36 近7天 7 来源 37 论文 167

#### 趋势信号

- 2026年欧洲多地破纪录热浪事件被多篇博客与政策评估文献用作气候适应能力的现实检验场
- OVO-S-Bench基准明确将‘空间模拟与推理’和‘外源性空间映射（allocentric mapping）’列为流式空间智能的高阶能力层级
- 美国高压输电网络研究构建了统一框架，对九类单一灾害及一类复合灾害（冻雨—强阵风）进行跨灾种风险比较评估
- HME-CA模型将人类移动性增强机制嵌入元胞自动机，用于城市群尺度的经济与人口协同模拟

#### 核心观点

- 城市韧性不再仅指物理抗毁性，而是表现为在多重、时序叠加的气候与基础设施扰动下维持关键功能的空间—系统响应能力
- 地理模拟正从离散状态转移（如传统CA）向支持流式输入、时间戳约束推理与跨尺度空间表征（egocentric→allocentric）的动态架构演进
- 复杂网络分析需与灾害物理过程（如热传导、电力故障传播、洪水淹没动力学）耦合，而非仅依赖拓扑指标
- 多灾种风险评估必须打破‘单灾种孤立建模’范式，建立统一脆弱性表征与影响传播链路

#### RESEARCH IDEA

### HME-CA模型在热浪胁迫下无法表征应急服务节点失效的级联传播

基于人类移动性增强的元胞自动机（HME-CA）模型在模拟热浪期间城市应急服务网络时，因未嵌入设施物理脆弱性阈值与服务响应延迟反馈机制，导致其无法复现高温导致的救护车调度失败向消防站覆盖盲区的级联扩散过程。

为什么现在值得做：美国高压输电网络的多灾种风险比较评估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

- 需构建热敏感型设施脆弱性函数，但现有遥感反演地表温度与设施运行状态间缺乏实测标定数据
- HME-CA元胞状态空间需扩展至包含服务可用率、响应延迟、资源重分配三类新维度
- 级联传播验证依赖高精度救护车GPS轨迹与医院接诊时间戳配对数据，该类数据在中国尚未开放共享

#### 建议切入

- 先对两篇代表论文做变量和评价口径对照，确认差异到底来自设定还是方法。
- 再做一个小规模复现或案例迁移，先找出最先失稳的部分。
- 最后根据失稳环节反推真正需要补的数据或约束，而不是直接堆方法。

#### 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。

##### 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。

##### 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。

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

近期研究重心从单纯图像识别转向对齐性、主观性与测量可靠性的系统性建模；方法上强调多模态融合（尤其是注视、VLM、专家规则与行为信号）、视觉对齐工具开发，以及将人类判断的分歧与拒答纳入评估框架。

近30天 19 近7天 7 来源 38 论文 163

趋势信号

- 出现专用于街景图像视觉对齐的开源工具（PairWise Image Finder），强调特征匹配与语义掩膜联合量化对齐质量
- 多个研究引入眼动追踪（Place Pulse-Gaze 数据集）和人类标注信度（inter-annotator reliability）作为建模与评估的必要输入
- VLM被广泛用于生成结构化描述与无障碍性评分，但均强调需嵌入领域知识（如 ADA 准则）或行为代理信号（如轮椅驻留时间）进行引导
- 虚拟角色提示（persona prompting）被实证用于探测多模态 LLM 在解释生成中的社会属性偏差，而非仅追求描述准确性

核心观点

- 街景图像的元数据对齐不可靠，视觉级对齐是开展纵向城市变化分析的前提
- 主观城市感知不能脱离人类感知过程建模，注视行为本身即携带可预测的感知信号
- VLM在城市感知任务中的有效性高度依赖其与人类判断分布（含分歧与拒答）的一致性，而非单一准确率
- 评价性城市感知标签具有情境敏感性与价值负载性，其标注空间与评分策略应被视为可协商的技术产物

## RESEARCH IDEA

### 注视引导模型在老年群体街景感知中失效于语义掩膜对齐偏差

注视引导的城市感知框架在面向中国老年人群的街景评估中，因眼动轨迹与语义分割掩膜的空间覆盖不一致导致感知标签预测偏差显著高于青年组

为什么现在值得做：中国城市老龄化加速推动适老化空间评估需求，而现有街景感知工具缺乏面向老年视觉生理特征的校准机制；PairWise Image Finder 提供的语义掩膜对齐度指标可被复用为偏差量化基准，使该问题具备可测量性。

关键难点

- 需构建老年与青年被试在相同街景样本下的眼动-标注双模态配对数据
- 通用语义分割模型（如Mask2Former）在老年注视热点区域的掩膜召回率未知，需实证评估
- 注视轨迹与掩膜像素级对齐的度量方式尚未标准化，需定义空间重叠衰减函数

建议切入

- 复用 Place Pulse-Gaze 的注视引导框架作为基线，在北京、成都招募60+与25-35岁各30名被试完成相同100张街景的感知打分与眼动采集
- 在相同街景上运行 PairWise Image Finder 输出的语义掩膜，计算各被试注视热图与掩膜交集面积占热图总面积的比例，作为对齐偏差代理变量
- 以该比例为协变量，构建分层回归模型检验其对感知标签预测误差（MAE）的解释力，控制图像亮度、对比度等低阶视觉特征

## 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

##### 危险品 (Hazmat) 十年事故汇总报告——数据挖掘工具

一系列事故数据与汇总统计报告，提供按事故类型、年份、地理区域等维度划分的统计信息。所用数据源自《危险品事故报告表》(Form 5800.1)。

#### USDOT OPEN DATA

##### 承运人安全评估系统 (CSMS, 或SMS) ——原始数据

美国联邦汽车运输安全管理局 (FMCSA) 的安全管理体系 (SMS) 是一个自动化数据系统，用于监测机动车承运人的路面安全表现。FMCSA 通过将承运人数据在 SMS 中归类至七个行为分析与安全改进类别 (BASICs)，分析其安全表现；这些类别进而用于识别单个承运人潜在的安全问题，并判断何时宜采取执法干预措施。

#### SCIENTIFIC DATA

##### HU-EnviroGrids: 面向匈牙利全国尺度空间分析与建模的栅格化环境数据集

日益复杂的环境挑战正推动对综合性环境信息需求的增长。应对这些挑战需要多学科方法，并依赖于源自多样化数据源的标准化数据集。我们开发了 HU-EnviroGrids，这是一个空间全覆盖、栅格化的环境数据集，覆盖匈牙利全境。

#### CITIES TODAY

##### 新推出的谷歌支持 AI 工具助力气候规划

CDP——一家全球环境信息披露组织——已推出一款由人工智能驱动的平台，旨在帮助城市。该文章《新推出的谷歌支持 AI 工具助力气候规划》首发于《Cities Today》。

#### USDOT OPEN DATA

##### 驾驶员安全评估系统 (DSMS)

非公开资料——驾驶员安全评估系统旨在识别具有安全违规记录的驾驶员。该信息用于在调查员对机动车承运人开展合规性审查或其他干预行动时，辅助确定执法重点。该信息不向公众提供。

#### SCIENTIFIC DATA

##### FFEM-DB 2.0

本数据描述论文介绍了《欧洲—地中海地区洪灾致死数据库》(FFEM-DB 2.0) 的第二版，该多国数据集记录了 1980 至 2024 年间发生在 16 个行政区域（其中 12 个为完整国家）的 3,737 起洪灾致死事件。

#### ARXIV

##### SpatialWorld: 面向真实世界任务的多模态智能体交互式空间推理基准

空间推理是多模态大语言模型 (MLLM) 感知并操作物理世界的一项基础能力。然而，现有基准主要依赖被动评估（例如静态视觉问答）或特定仿真器的流水线，无法有效衡量通用的交互式空间理解能力。我们提出 SpatialWorld，一个专为评估多模态智能体在复杂真实世界任务中交互式空间理解能力而设计的统一基准。

#### CITIES TODAY

##### 我们在出行战略中需要弥补的差距

作者：里克斯·沙阿 (Rikesh Shah)，交通创新联盟 (Transport Innovation Alliance) 主席。五月，作为主席 本文首发于《城市今日》(Cities Today)。