

THIS EDITION

五个方向的当日进展

浮亭：地形迹踪

研究日报 · 地理大模型、多模态融合与空间人文实践交汇之日。

浮亭是一座小型公共构筑物，嵌入浙江省松阳县梯田式茶园景观之中。

该构筑物坐落于一处仍在耕作的农业地形内，既作为观景装置，亦作为既有采茶路径的基础设施延伸。

场地特征为层叠茶园所构成的开阔视野，山谷底部随季节升腾起薄雾。

编者按：本期头版聚焦地理智能体范式跃迁——当ActCam重构视频中的空间运动控制，egenioussBench锚定城市级视觉定位基准，而‘浮亭’与‘隐秘花园’则以物质性回应同一套空间认知逻辑。五大学科趋势不再平行演进，而是于‘可感知—可建模—可干预’三层结构中形成闭环。

TREND OVERVIEW

趋势综述：地理智能体升维：从地形迹踪到城市感知的协同演进。

近期研究聚焦于地理空间基础模型（GFM）在真实地球观测任务中的适配性与泛化性验证，方法重心正从单纯架构创新转向对预训练数据构成、域偏移鲁棒性及多尺度嵌入融合机制的系统性解耦分析。

近期研究重心从单纯融合多源遥感图像模态，转向引入文本、时序、任务语义等异构模态，并强调面向真实部署的鲁棒性与可解释性。方法上更注重跨模态语义对齐机制的设计，而非简单特征拼接或对比学习。

近期研究重心正从单纯轨迹建模转向融合语义理解、行为推断与系统级动态建模的多粒度分析；方法上强调弱监督/无监督范式、统计可解释性、跨源标准化，以及将大模型或物理启发机制嵌入交通决策闭环。

DIRECTION PULSE

1 地理大模型与地理智能体

近期研究聚焦于地理空间基础模型（GFM）在真实地球观测任务中的适配性与泛化性验证，方法重心正从单纯架构创新转向对预训练数据构成、域偏移鲁棒性及多尺度嵌入融合机制的系统性解耦分析。

2 多源多模态地理数据

近期研究重心从单纯融合多源遥感图像模态，转向引入文本、时序、任务语义等异构模态，并强调面向真实部署的鲁棒性与可解释性。方法上更注重跨模态语义对齐机制的设计，而非简单特征拼接或对比学习。

3 轨迹数据与城市交通研究

近期研究重心正从单纯轨迹建模转向融合语义理解、行为推断与系统级动态建模的多粒度分析；方法上强调弱监督/无监督范式、统计可解释性、跨源标准化，以及将大模型或物理启发机制嵌入交通决策闭环。

4 复杂网络、韧性城市与地理模拟

近期研究聚焦于将城市基础设施建模为多层异质图（如道路-桥梁-建筑），并利用图神经网络（如R-GCN-VGAE）量化其在灾害中的功能角色；方法重心正从静态拓扑分析转向动态、场景驱动的元路径感知分类与反事实推演。

5 城市感知、街景感知与空间优化

近期研究正从单纯图像到主观判断的端到端相关性建模，转向对人类感知机制（如注视行为）和因果干预能力（如反事实编辑）的深入探索；方法重心由静态特征提取转向多模态耦合与可解释性驱动的空间操作。

HIGHLIGHTS

- 浮亭嵌入梯田茶园，将农业地形转化为可读的空间叙事界面。
- 隐秘花园以仿生贝壳外壳响应地中海气候，实现建筑微气候自主调节。
- ActCam首次支持视频生成中角色运动与相机参数的零样本联合控制。
- egenioussBench构建城市尺度机载三维网格与CityGML LoD2对齐基准。

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

近期研究聚焦于地理空间基础模型 (GFM) 在真实地球观测任务中的适配性与泛化性验证, 方法重心正从单纯架构创新转向对预训练数据构成、域偏移鲁棒性及多尺度嵌入融合机制的系统性解耦分析。

近30天 161 近7天 28 来源 55 论文 648

趋势信号

- Prithvi 等 GFM 已进入在轨平台实际部署验证阶段, 标志从实验室向 operational AI 迈进
- 多篇论文明确将 GFM 与传统人工构建协变量 (如夜间灯光、聚居区范围) 进行控制性对比, 凸显其作为可复用表征源的定位
- 研究开始系统化 GFM 改进的不均衡性——如人口估计中提升幅度随地理背景强弱而显著变化
- 预训练数据的光谱多样性被识别为比地理覆盖或生物群系多样性更强的性能预测因子

核心观点

- 地理空间基础模型的核心价值在于提供跨任务、跨区域的可复用地表表征, 而非替代领域专用建模流程
- GFM 在下游任务中的性能瓶颈常源于地理与时间域偏移, 而非模型容量不足
- 全局 GFM 嵌入与高分辨率视觉特征存在语义-空间鸿沟, 需结构化解耦机制 (如 SSDM) 避免特征干扰
- 当前缺乏针对 GFM 的标准化适配范式, 低秩自适应 (LoRA) 等轻量微调策略正成为应对域偏移的主流探索路径

RESEARCH IDEA

GeoAgentBench 参数执行准确率在跨城市 GIS 工具链中失效

GeoAgentBench 定义参数执行准确率 (PEA) 在跨城市尺度迁移时无法反映真实 GIS 工作流失败原因, 因其依赖“最终尝试对齐”机制而忽略工具调用序列中早期参数漂移对后续拓扑操作的累积性破坏

为什么现在值得做: 城市计算与数字孪生项目亟需可迁移的 GIS 智能体部署方案, 当前缺乏能诊断跨城市工具链失败根源的评估指标; Prithvi 等在轨 GFM 已支持多源遥感输入, 但下游 GeoAgent 仍受限于本地化工具适配瓶颈。

关键难点

- 需构建跨城市 GIS 工具链故障注入框架, 覆盖 EPSG 代码混用、栅格像元大小不匹配、矢量几何精度截断三类典型偏移
- PEA 指标需解耦为序列级参数漂移度 (SPD) 与拓扑传导失效率 (TCFR) 两个新维度, 二者无现成计算范式
- 真实 GIS 工作流中参数错误的因果标注依赖专家回溯, 尚无公开的跨城市 GIS 操作日志数据集

建议切入

- 在 GeoAgentBench 沙箱中注入 3 类城市级 GIS 偏移 (投影不一致、分辨率跳跃、坐标系混用), 记录每步工具调用的参数偏差轨迹
- 基于 QGIS/GRASS 操作日志构建专家标注子集, 定义 SPD (参数偏离真值的标准化距离) 与 TCFR (前序参数错误导致后续拓扑操作返回空集的比例)
- 将 PEA 替换为 SPD-TCFR 联合约束优化目标, 在 AlphaEarth 与 Prithvi-v2 驱动的 GeoAgent 上重训练参数生成头

REPRESENTATIVE ITEMS

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。

NASA NEWS

NASA 的 Prithvi 成为首个在轨运行

一支研究团队在两个在轨平台上成功验证了 NASA 与 IBM 联合开发的开源地理空间人工智能基础模型 Prithvi。

ARXIV

地理空间基础模型嵌入在空间与尺度上不均衡地提升人口估计精度可靠的次国家级人口估计对诸多应用至关重要, 但在人口普查数据稀疏、过时或空间分辨率粗糙的地区仍难以实现。现有制图工作流依赖人工构建的地理空间协变量 (如聚居区范围、夜间灯光和环境条件), 这些变量需跨尺度与跨区域进行组装与标准化。地理空间基础模型则提供了一种替代方案, 其通过从更丰富、异构的数据源中学习可复用的地点表征来实现建模。

ARXIV

预训练在哪里? 探究预训练数据多样性对地理空间基础模型性能的影响

新兴的地理空间基础模型引入了新的模型架构与预训练数据集, 其数据采样常基于不同的数据多样性定义。当前性能差异主要归因于模型架构或输入模态, 而预训练数据集的作用却鲜有研究。为填补这一研究空白, 我们系统地探究了预训练数据的地理构成如何影响模型在下游任务中的性能。

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

近期研究重心从单纯融合多源遥感图像模态，转向引入文本、时序、任务语义等异构模态，并强调面向真实部署的鲁棒性与可解释性。方法上更注重跨模态语义对齐机制的设计，而非简单特征拼接或对比学习。

近30天 234 近7天 51 来源 53 论文 872

趋势信号

- 多篇论文明确将文本（如物体级标签、场景级语义）作为独立模态引入遥感分割与理解任务，以弥合视觉-概念语义鸿沟
- 变化检测研究开始构建专用多时相多模态基准（如Delta-QA）并设计具备时间感知能力的MLLM架构（如Delta-LLaVA）
- 针对实际部署场景，多个工作系统性建模真实扰动（云/雾/口语化指令），并提出语义等价簇训练范式（如RemoteShield）
- 智能体范式兴起，将多模态感知与领域知识驱动的元规划（如Meta Task Library）耦合，推动EO系统从感知向行动演进

核心观点

- 多模态融合必须尊重模态异质性，强行对齐易导致脆弱表征或主导模态压制；需通过语义引导（如SGMA）、动态交互（如TSMNet）等机制调和冲突
- 文本不仅是辅助监督信号，更是承载领域知识与人类认知结构的关键模态，其引入显著提升开放词汇泛化与决策可解释性
- 真实地球观测场景下的模态缺失（IMSS）、输入扰动、时相不一致等非理想条件，是当前多模态方法落地的核心瓶颈，而非性能上限问题
- 多时相理解不能依赖静态特征拼接，需建模差异先验与时空因果结构（如Change-Enhanced Attention），实现从‘检测变化’到‘理解变化’的跃迁

RESEARCH IDEA

SGMA框架在跨城市迁移时对SAR-光学模态缺失的泛化失效

SGMA框架在训练于中国东部城市SAR-光学配对数据后，迁移到中亚干旱区时因SAR散射机制差异导致模态感知权重崩塌，无法维持类内差异校正能力

为什么现在值得做：中亚国家正加速部署国产SAR星座（如KazEOSat-1），亟需可复用的IMSS模型；而当前开源遥感多模态数据集（如SEN12MS、RSDA）缺乏中亚典型地物（盐沼、干涸湖盆）与SAR成像参数标注，构成现实应用缺口。

关键难点

- 需构建中亚典型SAR-光学配对子集并标注跨模态类内差异（如盐壳在VV/VH极化下的纹理一致性）
- SGMA原架构中语义原型提取模块未建模雷达波长与地表介电常数的耦合响应
- 模态感知权重动态调整机制依赖本地统计先验，无法跨区域迁移

建议切入

- 在SGMA的语义引导融合（SGF）模块前插入物理驱动的SAR散射校准层，输入本地土壤湿度与粗糙度先验
- 基于RSDA数据集构建中亚子集（KazRSDA），强制保留3类以上盐渍化地物的SAR-光学像素级对齐
- 将SGMA的模态感知权重输出接入对抗域判别器，约束其在东部/中亚域间分布一致性

REPRESENTATIVE ITEMS

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

连接感知与行动：面向鲁棒地球观测智能体的轻量级多模态元规划框架

自主地球观测（Earth Observation, EO）智能体正从被动感知转向复杂、多步骤任务执行。然而，当前将规划与执行集成于单一模型的架构在动态EO场景中常面临组合爆炸与推理错误等挑战。为应对这些问题，我们提出轻量级多模态元规划框架（Lightweight Multimodal Meta-Planner, LMMP）。

ARXIV

解码变化量：利用多模态大语言模型统一遥感变化检测与理解

尽管多模态大语言模型（MLLMs）在通用视觉-语言任务中表现优异，但其在遥感变化理解中的应用受限于一种根本性的“时间盲性”。现有架构缺乏内在的多时相对比推理机制，且难以实现精确的空间定位。为此，我们首先提出Delta-QA——一个包含18万条视觉问答样本的综合性基准。

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纯轨迹建模转向融合语义理解、行为推断与系统级动态建模的多粒度分析；方法上强调弱监督/无监督范式、统计可解释性、跨源标准化，以及将大模型或物理启发机制嵌入交通决策闭环。

近30天 398 近7天 89 来源 64 论文 1195

趋势信号

- 多篇论文聚焦GPS轨迹中停留点 (staypoint) 的出行目的弱监督推断，依赖POI语义区域与宏观统计数据（如家庭出行调查）进行分布对齐而非个体标签。
- 强化学习与大语言模型 (LLM) 开始被引入交通信号控制 (TSC)，但并非端到端生成动作，而是通过DQN批评器提供稠密状态级监督信号以引导GRPO微调。
- 统计推断框架被显式构建用于日际路径选择动态建模，强调参数可识别性、不确定性量化及对匿名化/异质性/需求变化等现实约束的理论包容性。
- 统一平台（如Ozone）和新型度量（如PASS）同步涌现，分别针对数据互操作瓶颈与瞬时-行程级评估割裂问题，体现基础设施层与评估层的协同演进。

核心观点

- GPS轨迹的语义解读面临三重根本挑战：个体标签缺失、空间不确定性（GPS噪声+POI覆盖不全）、以及必要性/非必要性活动的行为模式本质差异。
- 交通系统建模需兼顾微观个体行为可解释性（如学习机制、调整策略）与宏观涌现现象的物理类比（如热力学相变、自由能景观），二者不可相互替代。
- 轨迹数据驱动的研究正经历‘标准化危机’：异构数据源（NGSIM/highD/CitySim/UTE）在坐标系、目标表征、元数据上的不兼容，已实质性阻碍可复现性与跨区域迁移。
- 瞬时驾驶效率评估不能脱离行程级目标孤立设计；有效度量必须在运动学约束（如可达速度空间）与交通情境响应之间建立耦合映射。

RESEARCH IDEA

帕累托校准的出行目的推断在低POI覆盖城市失效

基于POI语义区域与帕累托校准的GPS轨迹出行目的推断方法在印度国家首都辖区等低POI覆盖城市中会失效，因为其依赖邻域级POI语义区域的完整性以构建距离加权空间似然，而该区域POI密度不足导致空间似然分布严重偏置。

为什么现在值得做：印度国家首都辖区无人机轨迹数据集 (arXiv:2512.11898v1) 提供了带人工验证的微观车辆轨迹与真实道路功能标注，可支撑POI覆盖度量化与目的标签反事实构造；城市规划部门亟需在POI基础设施薄弱地区开展通勤链识别，以支撑公交线路优化与非机动车道配置。

关键难点

- 需建立POI覆盖度的空间显式度量（如半径500m内POI类别熵与数量双指标），不能仅用全局密度替代
- 需从无人机轨迹中提取可信停留点并映射至非POI锚点（如路口拓扑、路侧设施图像识别结果）以构造弱监督真值
- 帕累托优化目标函数中多任务权重在低POI条件下不可泛化，需重新设计约束项以解耦空间似然偏差与行为先验偏差

建议切入

- 首先在印度国家首都辖区六路段无人机数据上运行标准停留点检测 (DBSCAN+速度-加速度联合阈值)，并人工校验200个停留点的空间锚定合理性（是否邻近路口/公交站/建筑出入口）
- 其次基于OpenStreetMap与本地POI爬取数据计算每个停留点半径500m内的POI类别熵与数量，划分高/中/低覆盖子集（依据三分位数）
- 然后在低覆盖子集中冻结POI语义区域模块，改用路网拓扑中心性 (betweenness centrality) 与影像分割所得路侧设施掩膜联合构建空间似然基底

REPRESENTATIVE ITEMS

ARXIV

基于POI语义区域与帕累托校准的GPS轨迹不确定性感知出行目的推断

大规模GPS轨迹数据为人类移动性提供了丰富的观测信息，但因缺乏个体层面的真实标签、GPS噪声导致的空间不确定性以及兴趣点 (POI) 覆盖不全，且不同出行目的在行为模式上存在根本差异，故对检测出的停留点分配出行目的仍具挑战性。

ARXIV

DGLight: 基于DQN引导的GRPO微调大型语言模型用于交通信号交通信号控制 (TSC) 在缓解拥堵和维持城市交通流动性方面发挥着核心作用。本论文提出DGLight，一种基于批评器引导的强化学习框架，用于将预训练大型语言模型适配至TSC任务。DGLight首先训练一个基于CoLight的深度Q网络 (DQN) 批评器，以从结构化的路口状态中估计交通感知的动作值；随后，该冻结的批评器被用于对候选语言模型动作进行打分，并利用组相对策略优化 (GRPO) 更新策略。

ARXIV

日际交通动态的统计推断

日际交通动态被广泛用于建模出行者学习与调整行为所导致的交通流演化，但此类模型的经验分析通常依赖于描述性校准，推断内容有限。本文基于随机个体层面调整模型，构建了一个面向日际路径选择动态的统计推断框架。该框架支持从轨迹数据中对行为参数进行不确定性量化与形式化推断。

ARXIV

投影可达速度空间：一种连接瞬时评估与行程时间的驾驶效率度量过于保守的让行等低效驾驶行为仍是自动驾驶汽车 (AV) 部署的关键障碍。瞬时驾驶效率度量对自动驾驶决策至关重要，因其直接影响实时性能评估与控制优化。然而，常用指标（如车速、相对车速和车间距）在刻画交通情境方面存在局限，且难以保证瞬时输出结果与行程级结果之间的一致性。

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

近期研究聚焦于将城市基础设施建模为多层异质图（如道路-桥梁-建筑），并利用图神经网络（如R-GCN-VGAE）量化其在灾害中的功能角色；方法重心正从静态拓扑分析转向动态、场景驱动的路径感知分类与反事实推演。

近30天 29 近7天 5 来源 33 论文 133

趋势信号

- 多篇论文显式构建跨物理层级的异质图结构（如道路层-桥梁层-建筑层），以刻画基础设施的功能耦合关系
- ‘元路径’成为关键建模范式，用于定义桥梁在灾害中对医院/商铺/住宅等终端节点的可达性传导路径
- 出现面向灾害情景的轻量级系统级仿真器（如TN-NTN混合网络仿真器），强调故障传播与服务迁移的动态建模
- ‘If-then’反事实框架被引入城市韧性评估，用于形式化表达灾害扰动下的功能恢复路径

核心观点

- 城市韧性不能仅由单一指标（如连通性或冗余度）表征，而需在具体灾害场景下解耦基础设施的多维功能角色
- 基础设施的功能韧性高度依赖其在网络中的位置与上下文语义（如通往医院的桥梁 ≠ 通往商铺的桥梁），需结合地理语义与拓扑结构联合建模
- 异质图与元路径是连接GIS空间数据、复杂网络理论与GeoAI模型的关键抽象接口
- 韧性评估正从‘是否失效’转向‘失效后能否维持关键功能流’，强调服务连续性而非结构完整性

RESEARCH IDEA

R-GCN-VGAE元路径分类在中小城市失效因桥梁-建筑拓扑稀疏性

R-GCN-VGAE模型在桥梁-建筑异质图中依赖高密度元路径连接以支撑关系型图卷积，当应用于桥梁数量低于200座且医院/商铺POI密度低于每平方公里3个的中小城市时，其三类灾害准备等级分类准确率下降超18个百分点，因局部子图结构退化导致关系嵌入不可靠

为什么现在值得做：国家‘城市更新行动’要求将韧性评估下沉至县级市与县城，而当前无适配小尺度、低密度建成环境的基础设施角色识别工具；OSMnx+OpenStreetMap可稳定获取中小城市桥梁与POI数据，使该问题具备实证基础。

关键难点

- 需定义并标定‘桥梁-建筑元路径稀疏性’的操作化指标（如平均路径长度、连通分量占比、POI覆盖半径内桥梁数
- R-GCN-VGAE原始代码未开源，重实现需复现其异质图构建逻辑与损失函数中的关系正则项
- 守谷市桥梁数量在原文摘要中截断为‘14’，缺失完整统计，需通过JPN-National Road Database交叉验证

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

LANDSCAPE AND URBAN PLANNING

ReTrEAT城市

出版日期：2026年9月；来源：《景观与城市规划》（Landscape and Urban Planning），第273卷；作者：Malay Pramanik、Kyaw Zabu Tun、Md Saify Ababil。

ARXIV

基于R-GCN-VGAE的桥樑中心化元路径分类方法用于灾害韧性维护

面向灾害的日常基础设施管理对城市韧性至关重要。当桥梁在灾害引发的外部作用力下保持韧性时，经由元路径通往医院、商铺与住宅的通行能力得以维持，从而保障城市基本功能。然而，在预算有限条件下优先开展桥梁维护，需量化桥梁在灾害场景中的多维角色——这一挑战是现有单一指标方法所无法应对的。

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

面向灾害感知的天地一体化TN-NTN系统级仿真器

非地面网络（NTN）已被第三代合作伙伴计划（3GPP）标准化，作为未来6G系统的关键组成部分，以增强覆盖范围与网络韧性。具体而言，低地球轨道（LEO）卫星、高空平台站（HAPS）及无人机（UAV）等NTN技术有望在极端事件与灾害期间为地面网络（TN）提供支撑。本文提出一种轻量级系统级仿真器，用于在部分失效型灾害模型下评估天地一体化TN-NTN无线网络的故障后回退行为。

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

近期研究正从单纯图像到主观判断的端到端相关性建模，转向对人类感知机制（如注视行为）和因果干预能力（如反事实编辑）的深入探索；方法重心由静态特征提取转向多模态耦合与可解释性驱动的空间操作。

近30天 20 近7天 5 来源 32 论文 148

趋势信号

- Place Pulse-Gaze 数据集引入眼动追踪与个体感知标签同步记录，显式锚定人类感知过程
- 出现基于提示词引导的街景局部编辑框架，以识别影响安全感知的语义杠杆（如交通基础设施、物理维护）
- 多篇论文将街景感知指标与居民心理健康（如老年人精神健康、城市居民心理状态）建立实证关联
- LLMs 开始被整合进街景理解流程，用于增强语义推理与上下文生成能力

核心观点

- 城市感知本质上是主观的，不能仅通过图像统计特征建模，必须纳入人类感知行为或认知反馈机制
- 当前主流街景感知模型仍是相关性建模，缺乏对因果干预路径的刻画与验证能力
- 街景图像承载的空间信息具有跨领域效用，尤其在公共健康、老龄化支持等社会议题中已形成稳定实证链
- 局部化、语义可控的图像编辑正成为检验感知模型鲁棒性与可解释性的新范式

RESEARCH IDEA

注视引导模型在老年群体中失效

注视引导的城市感知框架在老年被试群体上预测性能显著下降，因其未建模年龄相关的注视模式退化（如扫视幅度减小、中心凹固定延长）与感知标签分布偏移的耦合效应

为什么现在值得做：利用人类注视建模主观城市感知 与 *The relationship between visual enclosure for neighbourhood street walkability and elders' mental health in China: Using street view images* 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

- 缺乏公开的老年群体街景注视-感知同步标注数据集
- 注视模式退化参数（如平均扫视幅度、注视持续时间方差）与感知维度（安全/宜人）间的非线性映射关系未知
- 青年与老年群体在相同街景下感知标签的分布偏移方向不一致，无法直接迁移校准

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

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。

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

驱动城市感知的视觉杠杆有多少？基于多重局部化编辑的干预性反事实分析

街景感知模型可大规模预测安全等主观属性，但其本质仍为相关性建模：无法识别针对特定场景、可能改变人类判断的局部视觉变化。我们提出一种基于杠杆的干预性反事实框架，将场景级可解释性重构为在结构化反事实编辑空间内的有界搜索。每个杠杆定义一个语义概念、空间支持范围、干预方向及受约束的编辑模板。

JOURNAL OF TRANSPORT & HEALTH

The relationship between visual

Journal of transport & health; 卷 13; 页码 90-102; 发表于 2019 年; 出版机构 Elsevier。

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

USDOT OPEN DATA

月度交通量趋势报告——2012年6月

《交通量趋势》月度报告是一份国家级数据报告，为各州所有道路提供经质量控制的车辆行驶里程（VMT）数据。

USDOT OPEN DATA

月度交通量趋势——2003年1月

《交通量趋势》月度报告是一份国家级数据报告，为各州所有道路提供经质量控制的车辆行驶里程（VMT）数据。

ARXIV

ActCam：面向视频生成的零样本联合相机与3D运动控制

在艺术应用中，视频生成需对表演与电影摄影（即角色运动与相机轨迹）进行细粒度控制。本文提出 ActCam，一种零样本视频生成方法，可将驱动视频中的角色运动联合迁移至新场景，并支持逐帧控制相机的内参与外参。ActCam 基于任意预训练的、以场景深度与角色姿态为条件输入的图像到视频扩散模型。

ARXIV

通过联合嵌入实现零样本卫星图像检索：面向危机响应的应用

地球观测档案的语义搜索仍具挑战性。CLAY 等视觉基础模型可生成丰富的卫星影像嵌入，但缺乏支撑直观自然语言查询所需的语义对齐能力；而构建遥感领域 CLIP 风格的全对比式联合编码模型，则需大规模配对数据与算力，这在全球尺度上尚不可行。本文提出 GeoQuery——一种零样本检索系统，通过提示词对齐的文本代理（prompt-aligned text proxies）绕过该限制。

USDOT OPEN DATA

州级交通公共财政统计汇总（TPFS）

州级交通公共财政统计汇总（TPFS）提供各级政府（包括联邦、州和地方政府）以及所有交通方式相关的收入与支出信息。

ARCHDAILY

隐秘花园生物气候外壳 / Vincent Callebaut

住宅以仿生贝壳状的生物气候外壳形式展开，其形态由地中海、阳光与风塑造，营造出可呼吸的微气候。——Vincent Callebaut。

ARXIV

egenioussBench：一种面向地理空间视觉定位的新数据集

我们提出 egenioussBench，一个基于地理空间参考数据构建的视觉定位基准：一座城市尺度的机载三维网格（3D mesh）与一个 CityGML LoD2 模型。该组合反映了实际可部署的测绘资产，并支持超越传统基于运动恢复结构（SfM）方法的真实可扩展性。查询数据由智能手机拍摄的图像构成，其地面真值具备厘米级精度且不依赖于地图，通过后处理动态定位（PPK）及地面控制点/连接点（GCP/CP）辅助平差获得。

ARCHDAILY

空间连续体 / Abin Design Studio

该办公空间以宁静为根基，将极简主义理念转化为温暖且以人为本的工作环境。项目面向一家具有前瞻视野的开发公司设计，体现了对清晰性、专注力与平衡感的追求——这些特质既契合组织文化，也呼应了员工日常工作的实际需求。