

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

THIS EDITION

五个方向的当日进展

DARAD：面向持续遥感图文检索的双适配器与排序感知蒸馏方法

遥感图文持续检索、全局空间感知评估与电动微出行路径建模成为三大前沿支点。

随着地球观测技术的快速发展，遥感影像档案规模迅速扩大，使得遥感图文检索（RS-ITR）日益重要。

然而，持续RS-ITR仍面临挑战：遥感影像中的尺度变化与分布偏移加剧了跨模态对齐空间的失真，导致现有持续学习（CL）方法难以支持可靠的持续检索。

为应对该问题，我们提出DARAD——一种双适配器与排序感知蒸馏框架，在学习演化档案中新增视觉与文本概念的同时，保持历史跨模态排序结构。

编者按：本期头版聚焦地理智能体范式演进——从遥感多模态对齐到城市交通行为建模，体现‘感知-推理-行动’闭环在地理AI中的初步成型。五项主题趋势中，‘地理大模型与地理智能体’‘多源多模态地理数据’‘轨迹数据与城市交通研究’形成强共振，其余两条趋势亦在方法论层面提供支撑。

TREND OVERVIEW

趋势综述：地理智能体崛起：多模态空间推理。

近期研究重心正从单一模态栅格基础模型转向多模态、地理感知的生成式与推理型模型，并开始探索其作为地理智能体（GeoAgent）的推理能力；方法上强调领域原生预训练、地理空间约束建模与下游任务适配的协同优化。

近期研究聚焦于构建统一、可扩展的多模态遥感基础模型，方法重心正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。

近期研究正从传统车道化建模转向对无序、混合、异质性交通流的实证刻画，并强调轨迹数据与多源信息（如V2X网络参数、社会人口、POI）的跨域耦合与联合建模。

DIRECTION PULSE

1 地理大模型与地理智能体

近期研究重心正从单一模态栅格基础模型转向多模态、地理感知的生成式与推理型模型，并开始探索其作为地理智能体（GeoAgent）的推理能力；方法上强调领域原生预训练、地理空间约束建模与下游任务适配的协同优化。

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，方法重心正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。

3 轨迹数据与城市交通研究

近期研究正从传统车道化建模转向对无序、混合、异质性交通流的实证刻画，并强调轨迹数据与多源信息（如V2X网络参数、社会人口、POI）的跨域耦合与联合建模。

4 复杂网络、韧性城市与地理模拟

近期研究聚焦于将复杂网络分析与地理模拟方法（如ABM、Cellular Automata）耦合，用于量化评估城市在气候压力下的动态韧性；重心正从静态结构表征转向基于过程的、多尺度的时空演化建模。

5 城市感知、街景感知与空间优化

近期研究正从单纯依赖街景图像的静态表征，转向关注感知形成机制（如生理唤醒、注视行为）与测量鲁棒性（如天气干扰、跨平台对齐）。方法重心从端到端预测迁移至可解释建模与多源信号耦合。

HIGHLIGHTS

- DARAD方法提出双适配器与排序感知蒸馏，应对持续遥感图文检索中的尺度变化与分布偏移。
- GST-Bench首次构建面向视频流的全局空间感知能力评估基准，突破单视角局部感知局限。
- 两项华盛顿特区研究分别建模电动滑板车与电动自行车路径选择，推动无序混合交通流实证刻画。
- SpatioLM探索视觉-语言模型原生空间智能，避免引入外部3D先验导致通用能力削弱。

近期研究重心正从单一模态栅格基础模型转向多模态、地理感知的生成式与推理型模型，并开始探索其作为地理智能体（GeoAgent）的推理能力；方法上强调领域原生预训练、地理空间约束建模与下游任务适配的协同优化。

近30天 162 近7天 33 来源 63 论文 1110

趋势信号

- GEOID-Flood等新基准数据集强调多模态（SAR+光学+DEM）与灾害事件级覆盖，而非仅像素级标注
- GeoCore-9B首次实现完全基于EO数据从零训练的9B级生成式基础模型，并显式编码经纬度、GSD等地理元数据
- Clay等地理空间基础模型被用作辅助上下文注入U-Net等传统架构，体现‘基础模型+专业模型’混合范式兴起
- 多篇论文明确将GeoFM与Agent范式关联，提出‘从预训练到智能体推理’的演进路径

核心观点

- 地理空间基础模型（GeoFM）的核心价值在于职责分离：模型提供商承担预训练，领域专家专注安全、可控的下游适配
- 遥感物理规律（如成像几何、辐射传输）和地理空间约束（如坐标系、尺度依赖性）必须内化于模型设计，不能仅依赖自然图像先验
- 栅格与矢量数据是互补的地理空间表征范式——前者捕获连续场模式，后者编码离散实体及其语义拓扑关系
- 当前尚无通用最优GeoFM，评估标准不一致与基准缺失严重制约模型公平比较与可靠部署

RESEARCH IDEA

GeoCore-9B在洪水制图中无法生成符合水文连通性的淹没矢量边界

GeoCore-9B以文本和地理元数据为条件生成的洪水淹没面，在GEOID-Flood基准上无法保证生成矢量边界的水文连通性约束，因其扩散过程未显式建模地表流向与汇流路径。

为什么现在值得做：灾害应急部门亟需可直接输入GIS工作流的拓扑合法矢量产品，而非仅栅格掩膜；GEOID-Flood首次提供带水文约束的真值矢量边界，使该问题可量化验证。

关键难点

- 需将水文连通性（如D8流向、汇流累积量）编码为可微分几何约束，嵌入Flow Matching损失
- GeoCore-9B输出为栅格，而水文连通性是矢量拓扑属性，需设计栅格到矢量的可导映射层
- GEOID-Flood中仅部分样本提供矢量级真值，需构建水文一致性判别器替代全监督

建议切入

- 在GeoCore-9B解码器后接入轻量级水文物理模块（基于DEM梯度计算D8流向），将生成栅格投影至流向空间并计算汇流一致性得分
- 用GEOID-Flood中已有的矢量真值训练一个图卷积判别器，区分连通/非连通淹没区域，将其梯度反传至生成主干
- 在地理空间语义对齐损失中增加一项水文先验蒸馏项，从冻结的HydroSHEDS水文模型提取流向场作为教师信号

REPRESENTATIVE ITEMS

ARXIV

GEOID-Flood: 面向洪水分割的大规模多模态基准数据集

地理空间基础模型旨在学习可跨区域与传感器迁移的表征，但对其在具体任务上的评估需依赖大规模、高质量、多模态的基准数据集，以衡量此类模型从数据中提取价值的能力。在洪水制图领域，现有数据集极少在大规模上同时整合双时相合成孔径雷达（SAR）影像与配准的光学影像，致使基础模型在此下游任务中的价值 largely 未被验证。

ARXIV

地理空间基础模型的新兴范式：从预训练到智能体推理

卫星与航空影像分析已随着基础模型的出现进入新纪元。本文阐述了地理空间基础模型（GeoFM）的概念，即通过多种方法在海量地理空间数据集上进行预训练的人工智能/机器学习（AI/ML）模型。我们首先阐明GeoFM所推动的核心范式转变：职责分离——大规模模型提供商承担计算密集型的预训练任务，使领域专家能够快速对这些模型进行微调或提示工程，以完成特定的关键任务。

ARXIV

可扩展且可信的地球观测基础模型

基础模型（FM_s）已将机器学习范式从孤立的任务专用模型开发，转向基于广泛数据预训练、并适配多种下游任务的通用模型。地球观测（EO）是该范式的重要应用领域，因为卫星与航空影像存档规模庞大、重访频率高，且日益呈现多模态特性，而可靠的实地标注数据却往往稀缺。遥感基础模型（RSFM_s）若未经领域特定适配，则无法被可靠或最优地迁移应用。

ARXIV

GeoCore-9B: 面向地球观测的地理感知生成式基础模型

现有面向地球观测（Earth Observation, EO）的生成式模型主要依赖对自然图像先验进行微调，这限制了其可扩展性，并引入与地理空间约束相冲突的视角偏差。为解决此问题，我们提出GeoCore-9B——一个参数量达90亿的生成式基础模型，也是首个完全基于EO数据从零训练、且达到该规模的模型。

近期研究聚焦于构建统一、可扩展的多模态遥感基础模型，方法重心正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。

近30天 239

近7天 52

来源 58

论文 1672

趋势信号	核心观点
- MetaEarth-MM 提出以潜在场景表征为中介的解耦式生成架构，支持五种模态间任意配对生成 - TSMNet 和 Earth-OneVision 均引入双路径或多粒度文本-视觉融合机制，显式利用文本模态弥合语义鸿沟 - SAM 3 在遥感领域的评估揭示其多模态解码器存在跨模态干扰，促使研究者诊断模态对齐机制 - SGMA 和 FBA 方法均针对不完整模态或能力缺口设计分阶段适配策略，强调模态特异性与场景专用性	- 多模态遥感建模的核心挑战在于跨模态异质性（如成像物理差异、视角差异）与类内差异，而非单纯特征融合 - 场景一致性是多模态联合建模的关键先验，应作为中间表征或对齐锚点，而非仅在像素/特征层面进行映射 - 文本模态正成为不可或缺的语义监督源，用于支撑开放词汇任务与人类可解释决策 - 现有方法普遍面临模态不平衡问题：鲁棒模态压制脆弱模态，需通过语义引导或渐进式适配实现均衡学习

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：Earth-OneVision：将遥感多模态大语言模型扩展至更多传感器模态与任务 与 SGMA：面向遥感不完整多模态数据的语义引导模态感知分割 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	建议切入
- 需从SAR成像方程（如RCS模型）导出形状敏感度指标，而非仅依赖统计特征对齐 - SGF模块中语义原型的类别粒度（如‘集装箱船’）与SAR散射主导结构（如甲板/舷墙）不匹配 - HarborEval中SAR图像存在方位向压缩与距离向模糊，光学图像对应区域存在视角畸变，二者无法直接像素级对齐	- 基于SAR成像物理模型（如Extended Bragg散射近似）构建船舶部件级散射响应图谱，作为SGF模块的先验约束 - 在SGMA原有语义原型空间中引入物理感知投影层，将光学特征映射至散射主导的几何子空间 - 利用HarborEval中提供的传感器元数据（入射角、极化、分辨率）构造模态差异加权损失项，替代原始SGF中的L2一致性约束

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

自上而下的可提示概念分割

大规模基础模型（如 Segment Anything Model 3，SAM 3）的部署有望推动计算机视觉迈向开放词汇、无需训练的新范式。然而，其在分布外遥感影像——尤其是具有复杂俯视几何结构的地观测图像——上的泛化能力迄今尚未得到量化评估。受 SAM 3 在高度专业化领域中性能差异的驱动，我们在严格零样本与单样本约束下，对遥感场景分类、目标检测与实例分割三项任务开展了全面的多任务实证评估。

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。

近期研究正从传统车道化建模转向对无序、混合、异质性交通流的实证刻画，并强调轨迹数据与多源信息（如V2X网络参数、社会人口、POI）的跨域耦合与联合建模。

近30天 497 近7天 71 来源 70 论文 2631

趋势信号

- 高分辨率无人机轨迹被用于验证无序交通中横向交互对宏观动力学的持续影响，挑战Edie一维基本图假设
- 强化学习控制器在真实流量驱动的仿真中仅依赖本地轨迹观测实现信号自适应优化，脱离中心化预测依赖
- MINT-V2X等新数据集显式耦合SUMO交通仿真与OMNeT++/Simu5G网络仿真，填补移动性-通信参数联合建模的数据缺口
- AI代理（TrafficSci）基于多尺度轨迹证据自主归纳并验证交通规律，推动从专家驱动向可审计AI科学发现范式迁移

核心观点

- 无序交通流中车辆异质性与持续横向交互不可忽略，一维交通流模型存在根本性表征局限
- 轨迹数据的价值不仅在于描述移动性，更在于作为连接微观交互、中观控制与宏观规划的统一观测基底
- 真实世界部署的交通智能需兼顾边缘实时性（如PPO本地决策）与多源异构数据融合能力（如轨迹+POI+INSEE）
- 可解释性（XAI）与可复现性（模块化ETL/分析构件）已成为从轨迹数据生成政策或商业洞见的关键前提

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：城市精细化治理亟需中小尺度路口的交通状态实时表征，而MINT-V2X等新数据集提供了含RSU定位与几何拓扑的同步仿真轨迹，使局部几何-运动耦合建模成为可能。

关键难点

- 需从SUMO仿真中提取亚米级路口几何参数（如曲率半径、视距三角区顶点坐标）并与轨迹点精确对齐
- Edie框架要求定义二维速度-密度平面，但小尺度路口存在瞬时多向流叠加，导致速度场方向离散度超阈值
- 横向再分布被截断后，传统稳态跟驰-领航识别方法无法收敛，需重构交互事件检测逻辑

建议切入

- 基于MINT-V2X中15个RSU的空间坐标与SUMO路网XML，提取每个交叉口的几何约束要素并构建缓冲区掩膜
- 在掩膜内对轨迹点进行局部主成分分析（PCA），量化横向运动自由度衰减率作为截断强度指标
- 将截断强度作为协变量嵌入Edie二维基本图的密度维度，重新拟合速度-密度-截断强度三元关系

REPRESENTATIVE ITEMS

ARXIV

超越车道：基于高分辨率轨迹数据的无序交通流动力学研究
无序交通流以弱化或缺失的车道规则为特征，在车辆异质性显著且持续发生横向交互的条件下形成，从而挑战了传统基于车道的建模假设。本研究利用无人机采集的城市主干道高分辨率轨迹数据，对无序交通的宏观与微观特性开展实证分析。采用Edie框架的二维扩展方法量化集总交通变量，并构建二维基本图，结果表明一维公式不足以充分表征交通状态，并凸显横向再分布的持续作用。

ARXIV

基于强化学习的物联网赋能交叉口交通信号控制
城市交通拥堵在以汽车为主导的城市中仍是长期存在的挑战，带来显著的经济与社会成本。交通信号系统正日益作为网络化信息物理系统组件部署于智慧城市基础设施中，分布式感知与边缘智能支撑了自适应交通管理。本文研究强化学习（RL）作为一种边缘智能方法，用于科威特某信号控制城市交叉口的自适应交通信号运行。

ARXIV

MINT-V2X

车联万物（V2X）通信系统依赖于既包含车辆轨迹数据、又涵盖具备真实保真度的无线网络参数的数据集，以支撑预测与优化模型的构建。当前存在一项极为关键的研究基础设施缺口：公开可用的数据集通常仅覆盖移动性或网络参数二者之一，极少提供将两者统一整合的单一视图。本文提出MINT-V2X，该数据集通过耦合SUMO交通动力学仿真与OMNeT++/Simu5G网络仿真生成。

ARXIV

利用AI交通科学家自主发现交通规律

普适性交通规律描述了拥堵、移动性及驾驶行为在不同城市中反复出现的模式，为交通规划、管理和控制提供了科学基础。然而，此类规律的发现目前仍依赖专家主导，需从异构观测证据中识别候选规律，或通过干预实验加以验证。尽管自主人工智能（AI）系统已在受控实验室环境中推动科学发现，但将其拓展至复杂的交通领域仍具挑战。

近期研究聚焦于将复杂网络分析与地理模拟方法（如ABM、Cellular Automata）耦合，用于量化评估城市在气候压力下的动态韧性；重心正从静态结构表征转向基于过程的、多尺度的时空演化建模。

近30天 27 近7天 4 来源 40 论文 228

趋势信号

- 多篇论文集中发表于《Sustainable Cities and Society》，主题均围绕洪水等气候压力下城市韧性的动态评估方法
- arXiv新论文提出无需可调参数的引力-拟拉普拉斯节点排序方法，强调计算高效性与可解释性以适配大规模城市网络
- Nature子刊《Scientific Reports》论文将集成学习与地理空间+基础设施属性融合，提升关键设施风险表征的操作性
- 基于智能体的建模（ABM）被明确用于模拟出行行为干预对交通方式转移的影响，体现微观行为与系统韧性间的因果推演需求

核心观点

- 城市韧性不能仅依赖静态网络拓扑指标，必须嵌入时空动态过程（如洪涝传播、出行响应）进行评估
- 复杂网络方法（如关键节点识别）需兼顾计算可扩展性与地理可解释性，避免过度依赖黑箱参数
- 地理模拟（ABM、CA）正成为连接微观个体行为与宏观系统韧性的重要桥梁
- 关键基础设施的空间嵌入性（如道路网络可达性）是影响城市整体韧性的结构性约束

RESEARCH IDEA

引力-拟拉普拉斯节点排序在地铁内涝韧性评估中失效于拓扑扰动敏感性

引力-拟拉普拉斯方法在地铁网络内涝扰动模拟中会因忽略时空耦合边权重动态衰减而失效，因其仅依赖静态k-壳与度数，未建模水位上升导致的边容量非线性退化过程。

为什么现在值得做：城市地铁运营方亟需可部署的实时韧性诊断工具，而现有遥感+IoT融合数据已支持分钟级水位-流量-客流联合观测，填补了动态边权建模的数据基础。

关键难点

- 现有论文结论大多成立在特定场景里，换尺度或换样本后未必还稳定。
- 很多关键变量只在论文里被隐含处理，真正复用时需要重新显式定义。

建议切入

- 基于徐飞等（2026）地铁站点与管廊GIS拓扑，叠加Sentinel-1 SAR反演的地表积水深度时序栅格，生成边权衰减曲线
- 将引力-拟拉普拉斯框架中的邻域聚合操作替换为带衰减因子的时变邻接矩阵幂次迭代
- 在珠江三角洲城市群案例中，用池慧等（2026）识别的网络结构转型阶段划分训练集，校准不同发展阶段下的R自适应规则

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。

SUSTAINABLE CITIES AND SOCIETY

揭示城市在洪水压力下的韧性：一种基于正态云评分的动态评估方法

出版日期：2026年10月1日 来源：《可持续城市与社会》（Sustainable Cities and Society），第149卷 作者：张翔、黄建华、揭琼楠。

SUSTAINABLE CITIES AND SOCIETY

城市韧性网络的结构转型：珠江三角洲城市群的时空演化与机制
出版日期：2026年9月15日；来源：《可持续城市与社会》（Sustainable Cities and Society），第148卷；作者：池慧、郑立勤、邱启贤、卢楠。

NATURE

集成学习提升关键基础设施应对城市洪涝的韧性

这表明，整合多种地理空间与基础设施属性可改善精细化风险表征，为城市规划者和应急管理人员提供更具操作性的信息。我们的结果表明，华盛顿特区超过40%的能源设施和应急服务机构位于高风险道路网络内，这对韧性建设具有关键意义。

近期研究正从单纯依赖街景图像的静态表征，转向关注感知形成机制（如生理唤醒、注视行为）与测量鲁棒性（如天气干扰、跨平台对齐）。方法重心从端到端预测迁移至可解释建模与多源信号耦合。

近30天 20 近7天 5 来源 41 论文 188

趋势信号

- 多篇论文聚焦街景图像中感知测量的外部效度问题，包括天气导致的偏差、跨平台图像视觉对齐不足等系统性误差来源
- 出现专门面向城市感知研究的开源工具（如 PairWise Image Finder），强调纵向变化分析中视角一致性和视觉对齐质量
- 新数据集（如 Place Pulse-Gaze）显式引入人类眼动追踪与主观标签，推动‘人类感知过程’成为建模先验而非黑箱输出
- 期刊来源覆盖 GIS、Urban Systems、Transport & Health 等跨学科阵地，反映城市感知正作为连接空间形态、行为响应与健康结果的中介变量被系统考察

核心观点

- 街景图像虽是城市感知建模主流输入，但其语义丰度与人类主观判断之间存在显著鸿沟，需引入生理或行为信号弥合
- 城市感知不是图像属性的直接映射，而是受观测者状态（如注视路径、生理唤醒）、环境条件（如天气、光照）和图像采集元信息（如视角、平台一致性）共同调制的复合过程
- 高质量纵向分析的前提是视觉对齐而非仅地理/时间元数据匹配，这对变化检测与政策效果评估构成基础性约束
- 城市感知指标正从描述性景观特征（如绿化率、整洁度）向功能性空间效应（如步行友好性、老年人心理健康关联）延伸，驱动其在规划决策中的实证嵌入

RESEARCH IDEA

街景感知指标对行为差异的解释边界

街景感知指标可以描述城市体验，但它解释的究竟是行为与健康差异，还是社会经济差异的替代变量，仍然需要进一步辨析。

为什么现在值得做：城市更新决策需可迁移的感知代理指标，而全球超百个新增街景覆盖城市（如科索沃）缺乏本地标注数据；PairWise Image Finder等工具使跨城市对齐图像获取可行，为系统性检验形态变量迁移偏差提供数据基础。

关键难点

- 需构建多城市统一标定的视觉封闭度真值——依赖LiDAR或高精度3D重建，非所有城市具备
- 现有感知模型未输出视觉封闭度梯度，须重新设计可微分几何约束模块
- 跨城市建筑规范差异导致高度-宽度比分布偏移，无法仅靠域自适应缓解

建议切入

- 使用PairWise Image Finder在5个对比城市（含北京、首尔、奥斯陆、利马、普里什蒂纳）提取视角一致的街景图像对
- 基于OpenStreetMap建筑轮廓与Street View深度估计，逐帧计算视觉封闭度（垂直视野角占比），构建城市级分布基准
- 冻结主干网络，在感知模型最后一层引入视觉封闭度回归分支，并以多城市分布差异为正则项进行联合微调

REPRESENTATIVE ITEMS

TRANSACTIONS IN GIS

利用街景图像估计客观城市感知：以生理唤醒为例

《国际地理信息系统汇刊》（Transactions in GIS），第30卷，第5期，2026年8月。

COMPUTERS, ENVIRONMENT AND URBAN SYSTEMS

深度学习揭示街景影像在城市感知制图中的跨平台一致性

出版日期：2026年12月；来源：《Computers, Environment and Urban Systems》，第130卷；作者：白婷婷、徐东、高峰、刘佳伟、宋永泽。

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

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。

USDOT OPEN DATA

运输口袋指南：货物运输更新版

提供基于数据的国家货运物流基础设施概览。本版汇集了所有主要运输方式——公路、铁路、水运、航空和管道——在货运量、货值及吨英里方面的最新统计数据。该指南是政策制定者与行业专业人士的重要参考，用于直观呈现商业流通状况及供应链所面临的不断变化的需求。

ARXIV

基础设施与街道景观如何影响电动滑板车路径选择：来自华盛顿特区的证据

电动滑板车已成为城市短途出行的重要微型交通方式，但关于其路径选择行为的实证研究仍较为有限。本研究利用GPS轨迹数据和路径规模Logit（Path Size Logit）模型，分析华盛顿特区的电动滑板车路径选择行为。除道路及基础设施特征外，该模型还纳入了通过计算机视觉技术从Google街景图像中提取的视觉街道景观特征。

ARXIV

SpatioLM：面向视觉-语言模型的通用物理空间智能

视觉-语言模型（VLM）在常识推理任务上表现良好，但在视觉空间推理方面存在困难。现有大多数解决方案需引入额外的3D先验输入或外部空间编码器，这不仅增加了模型复杂度，还会在空间微调后削弱VLM原有的通用能力。为此，我们提出一种参数高效的 $\textit{\textbf{Spatio}}\text{-vision} \textit{\textbf{L}}\text{anguage} \textit{\textbf{M}}\text{odels}$ （SpatioLM），在不依赖额外3D先验输入或第三方空间编码器的前提下增强空间智能。

ARCHDAILY

BIAU | 第十四届伊比利亚美洲建筑与城市主义双年展公布174个入选项目 |

第十四届伊比利亚美洲建筑与城市主义双年展（BIAU）公布了入选项目。在“作品与出版物”类别中，共收到711份提案，最终选出174个项目；决赛入围名单将于9月公布。本届双年展以“关爱生命”为主题，旨在构建一个针对当代领土变迁进行批判性反思的平台，并围绕四大主题轴线展开：水、植被、社区与住房。第十四届BIAU将于11月3日至7日在巴西利亚举行。

ARXIV

GST-Bench：多模态大模型能否从视频中发展出全局空间感知能力

空间智能是具身智能体的基础能力，但现有基准测试主要集中于单视角或少数视角下的局部空间感知，忽视了对连续、长时序视觉流的全局空间感知能力评估。为弥补这一局限，我们提出全局-空间-时间基准（GST-Bench），这是一个面向视频理解中全局空间智能的视觉问答（VQA）基准，包含基于6,790分钟合成生成视频人工校验的问题。该基准要求模型能够从输入视频中未出现过的新视角进行精确的空间推理，并将第一人称视角观测映射至全局俯视图像。

ARXIV

华盛顿特区电动自行车路径选择建模：一种路径规模Logit方法

理解电动自行车路径选择对于构建有效的骑行基础设施至关重要，但实证证据仍较为有限。本研究利用Capital Bikeshare系统提供的全球定位系统（GPS）轨迹数据，考察华盛顿特区共享电动自行车的路径选择行为。采用混合选择集（包含观测路径及对应最短路径）估计路径规模Logit（Path Size Logit）模型，并整合地理信息系统（GIS）表征的基础设施变量与基于街景图像（Street View images, SVI）的计算机视觉提取的街道级视觉特征。

ARCHDAILY

Helix住宅 / SAOTA

该私人住宅位于洛杉矶日落大道上方的‘鸟街’（Bird Streets）区域，承袭了洛杉矶山地现代主义居住传统，强调建筑与景观的融合，以彰显光线、视野与开放性。设计直面陡峭地形，在保障私密性的同时，实现与下方城市全景的视觉联结。

ARXIV

克服注意力漂移：基于同质性-异质性引导的特征聚合方法用于低光照遥感图像增强

从极端低光照退化中恢复高保真度遥感影像，对可靠的地球观测及下游机器视觉任务至关重要。然而，在严重噪声与光照失真条件下，现有方法易出现注意力漂移现象，错误地跨不同物理边界聚合特征，导致严重的结构模糊与色彩失真。为此，我们提出HALO——一种双先验驱动的特征聚合框架，将增强建模为由基础模型先验引导的特征聚合问题。