

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

THIS EDITION

五个方向的当日进展

在家与学校之间：一道边界

研究日报头版聚焦地理智能体演进、城市空间再诠释与AI训练伦理张力。

对大多数儿童而言，上学之路是日常街景与交通的重复体验，频繁到几乎不可见，悄然融入童年与成长的背景之中。

然而，对于居住在深圳、在香港就学的跨境学生而言，上学日开始得更早，也远在教室之外——它始于边界。

编者按：本期头版以‘边界’为隐喻线索——既指跨境学童每日穿越的物理-制度边界，也指AI模型在语义与动作、视觉与几何、训练数据与版权之间的认知边界。五项主题趋势共同指向一个核心命题：智能系统正从‘识别空间’转向‘理解并介入空间’，而这一跃迁亟需跨学科的方法论校准。

TREND OVERVIEW

趋势综述：空间智能的双重边界：从城市肌理到机器感知。

近期研究重心正从单一模态地理空间基础模型（GeoFM）预训练，转向多模态融合、智能体式推理与下游任务适配方法的探索；问题焦点日益集中于模型可信性、跨域泛化能力及社区评估标准缺失等系统性挑战。

近期研究聚焦于将通用多模态基础模型（如SAM、MLLM）适配至遥感与地球观测场景，重心从模态间简单融合转向对齐机制诊断、场景中心表征建模与跨模态干扰抑制。

近期研究重心从单一轨迹建模转向以轨迹为驱动的城市交通系统级智能决策与科学发现；方法上强调边缘智能、代理式AI、端到端生成建模与真实V2X数据支撑的闭环验证。

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地理空间基础模型（GeoFM）预训练，转向多模态融合、智能体式推理与下游任务适配方法的探索；问题焦点日益集中于模型可信性、跨域泛化能力及社区评估标准缺失等系统性挑战。

2 多源多模态地理数据

近期研究聚焦于将通用多模态基础模型（如SAM、MLLM）适配至遥感与地球观测场景，重心从模态间简单融合转向对齐机制诊断、场景中心表征建模与跨模态干扰抑制。

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以轨迹为驱动的城市交通系统级智能决策与科学发现；方法上强调边缘智能、代理式AI、端到端生成建模与真实V2X数据支撑的闭环验证。

4 复杂网络、韧性城市与地理模拟

近期研究聚焦于将城市视为多层异质网络（如道路-桥梁-建筑），通过图神经网络等方法建模灾害扰动下的功能连通性；重心正从静态结构韧性转向动态功能韧性与政策干预的因果评估。

5 城市感知、街景感知与空间优化

近期研究重心从静态街景图像的表征学习，转向对感知过程机制（如注视行为）、空间推理可靠性（如VLMs的空间语义幻觉）及图像对齐质量等底层问题的深入探究。方法上强调多模态融合、结构化空间知识建模与工具链开源化。

HIGHLIGHTS

- 华集酒店改造实践体现历史空间审慎更新的范式转移。
- 跨境学童通勤揭示城市治理边界的日常化与身体化。
- Suno数据抓取事件引发AI训练数据来源合法性再审视。
- 多篇arXiv论文共同推进无GNSS无人机定位与自动驾驶语义-动作解耦。

近期研究重心正从单一模态地理空间基础模型（GeoFM）预训练，转向多模态融合、智能体式推理与下游任务适配方法的探索；问题焦点日益集中于模型可信性、跨域泛化能力及社区评估标准缺失等系统性挑战。

近30天 174 近7天 45 来源 59 论文 1013

趋势信号

- 多篇论文明确区分掩码自编码类可微调视觉模型与对比学习驱动的视觉-语言 GeoFM，并强调后者支持零样本开放词汇分析
- Clay等具体GeoFM被实证用于滑坡检测、野火过火区制图等任务，验证其作为辅助上下文或主干编码器的有效性
- 至少两篇论文指出矢量数据（如 OpenStreetMap、Overture）与栅格影像的互补性，呼吁构建融合栅格-矢量表征的GeoFM
- 多篇arXiv论文反复指出当前GeoFM领域缺乏统一评估协议、预训练配置复用率低、模型权重公开率不足（如39%未公开）

核心观点

- 地理空间基础模型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、高效的微调或提示工程
- GeoFM必须适配遥感特有的物理约束（如成像几何、大气校正、时序重访）与数据稀缺性，通用FM不能直接迁移
- 栅格与矢量数据构成地理空间的互补表征维度——前者捕获连续物理模式，后者编码离散对象及其语义-拓扑关系
- 当前尚无单一GeoFM在所有场景下表现最优，且评估标准不一致严重阻碍模型公平比较与可靠部署

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：灾害应急响应需在多山地区（如喜马拉雅、安第斯）部署轻量化滑坡制图模块，而Sentinel-2全球重访数据已支持跨地形-光照组合的实证分析；Clay模型权重公开且支持特征级干预，使物理约束嵌入具备工程可行性。

关键难点

- 需从DEM与卫星轨道参数中联合推导每个像素的局部入射角与地形遮蔽掩膜
- Clay中间层特征与物理量之间缺乏可微分映射函数，需设计梯度感知的注意力门控机制
- Landslide4Sense未标注观测几何元数据，须复用ESA Sentinel-2 QA波段重建有效观测条件

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

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

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, EOFM_s），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。

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

近期研究聚焦于将通用多模态基础模型（如SAM、MLLM）适配至遥感与地球观测场景，重心从模态间简单融合转向对齐机制诊断、场景中心表征建模与跨模态干扰抑制。

近30天 236

近7天 104

来源 57

论文 1528

趋势信号	核心观点
- SAM 3 被系统评估其在遥感零样本/单样本任务中的泛化能力，并显式解耦其二元存在性预测头以重构为零样本分类器 - 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，替代传统两两映射 - TSMNet 引入双分支文本编码器，分别提取物体级标签与场景级语义，通过文本引导视觉语义融合模块实现开放词汇分割 - Earth-OneVision 设计全粒度视觉-语言对齐（FGVLA）、空间-语言同构序列化（SLIS）和渐进式跨模态适配（PCMA）三大机制，统一六类传感器模态与九类任务	- 多模态遥感建模正从外观层面的跨模态映射转向以底层场景内容为锚点的联合表征学习 - 文本模态不再仅作辅助监督，而是作为弥合视觉模式与现实概念语义鸿沟的关键知识源，支撑开放词汇理解 - 跨模态干扰（如视觉提示对文本解码的干扰）是影响多模态模型泛化的核心问题，需通过解耦或结构化适配显式诊断与缓解 - 多模态基础模型的可扩展性依赖于统一框架设计（如Earth-OneVision）与专用机制（如FGVLA、SLIS），而非堆叠两两翻译模块

RESEARCH IDEA

SAM 3 在零样本遥感分割中因视觉提示与文本提示冲突导致类别漏检

SAM 3 的二元存在性预测头在零样本遥感分割任务中，当视觉提示与文本提示指向同一地物但空间分布不一致时，会系统性忽略文本定义类别，因其解码器未显式建模跨模态提示冲突下的优先级仲裁机制。

为什么现在值得做：灾害应急制图等场景亟需零样本条件下依据自然语言指令快速定位未见地物（如‘被滑坡掩埋的光伏板’），而当前 SAM 3 类方法在提示冲突时可靠性不足；Earth-OneVision 和 MetaEarth-MM 提供了多模态对齐与联合表征的工程基础，使冲突建模具备可实现接口。

关键难点	建议切入
- 需构造可控的视觉-文本提示冲突样本集，要求同一地物在光学影像中不可见但在 SAR 或文本描述中明确存在 - SAM 3 解码器内部缺乏可微分的模态优先级门控模块，需在不破坏原始权重冻结前提下插入轻量仲裁层 - 评估需定义‘提示冲突强度’量化指标，不能仅依赖 IoU，须引入基于语义原型距离的漏检归因度量	- 基于 Moonstone 基准中月球多模态通道构建地球类比的提示冲突合成策略：利用 SAR 与光学影像的空间错位特性生成视觉-文本不一致样本 - 在 SAM 3 解码器前插入模态仲裁模块（MAM），以文本嵌入与视觉提示特征的余弦相似度差值为输入，动态调节二元存在性头的 logits 缩放系数 - 在 TSMNet 构建的开放词汇数据集上复现冲突场景，用其物体级标签作为 ground truth 评估漏检率，对比原始 SAM 3 与 MAM 增强版本

REPRESENTATIVE ITEMS

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
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。

ISPRS JOURNAL OF PHOTOGRAMMETRY AND REMOTE SENSING
先验引导的多领域混合专家模型用于多模态地球观测数据缺失填补
出版日期：2026年9月 来源：《ISPRS摄影测量与遥感杂志》（ISPRS Journal of Photogrammetry and Remote Sensing），第239卷 作者：刘闯、郭建华、皮英东、吴潇、张志奇、陈儒、王欣怡、王密。

近期研究重心从单一轨迹建模转向以轨迹为驱动的城市交通系统级智能决策与科学发现；方法上强调边缘智能、代理式AI、端到端生成建模与真实V2X数据支撑的闭环验证。

近30天 474

近7天 202

来源 71

论文 2333

趋势信号	核心观点
- 强化学习（如PPO）被用于本地化、去中心化的交叉口信号控制，仅依赖实时轨迹观测，无需未来需求预测或中心协调 - AI代理（如TrafficSci）被构建为可审计的科学发现工作流，自主归纳并验证跨城市一致的交通规律，包括新发现的驾驶行为时间记忆尺度 - 面向隐私与可用性的离散扩散模型（如MobiDiff）直接建模多通道语义事件（空间/活动/时间），避免连续轨迹插值或潜在空间映射 - 新型真实世界V2X数据集（如CAMASA）强调基础设施侧协同感知消息（CAM/DENM）的大规模、长期、地理开放采集，弥补传感器中心与合成数据的局限	- 轨迹数据正从描述性分析工具升级为城市交通系统智能调控与科学发现的核心观测基础 - 边缘智能与分布式决策（如本地RL控制器）被视为提升交通系统鲁棒性与可扩展性的关键路径 - 交通规律具有跨城市统计一致性，但其发现需融合多尺度轨迹证据与可解释、可验证的AI代理工作流 - 合成移动性数据必须兼顾语义保真度（离散活动-时空结构）与隐私保护，不能仅追求几何轨迹逼真

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：城市规划部门亟需可泛化的合成轨迹支撑多城对比模拟，而当前真实轨迹数据因隐私与采集成本难以跨城共享；MobiDiff开源框架与CAMASA等新型基础设施轨迹数据集提供了可复用的训练-验证接口。

关键难点	建议切入
- 需构建统一尺度的城市功能语义图谱（如POI加权混合用地指数），且须与MobiDiff的三通道输入对齐 - 活动-时间联合分布偏移缺乏量化指标，需定义跨城市轨迹结构一致性损失函数 - 区域功能先验如何注入离散扩散过程尚无现成架构，需设计可微的土地利用条件掩码模块	- 基于OpenStreetMap与夜间灯光数据，在科威特与济南构建0.5km网格级功能混合度指数，并映射至MobiDiff空间通道的token embedding层 - 在MobiDiff事件级去噪目标中引入Jensen-Shannon散度约束，强制生成轨迹的活动×时间二维直方图与目标城市实测分布对齐 - 替换原通道级掩码为功能感知掩码：依据网格功能指数动态调整活动通道与时间通道的联合掩蔽概率

REPRESENTATIVE ITEMS

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
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。

ARXIV
从移动数据到商业洞察
源自移动应用的实时位置数据是应对多种城市挑战的有力工具，涵盖旅游规划、停车管理、公交线路优化及资源分配等领域；同时，该数据亦为基于位置的服务、市场份额分析和用户行为画像等商业战略决策提供宝贵洞见。本研究旨在通过系统考察智能手机用户在城市环境（特别是旅游、交通与零售领域）中的行为与模式，全面应对上述挑战。我们的方法涵盖从零构建并落地实施一个复杂的数据平台，包括用例定义、架构设计及模块实现。

ARXIV
优化公共交通网络的可达性与最大出行时间：枢纽覆盖选址问题
本研究提出一种优化模型，以支持一体化公共交通系统设计中的决策制定，旨在提升城市交通流动性。我们假设用户日常出行包含两类路段：短途路段（用户自行完成住所或工作地至公共交通接驳点之间的出行）和长途路段（在公共交通网络上的出行）。通过为短途路段设定最大出行时间上限，保障用户的可达性；同时，仅考虑位于给定覆盖半径内的用户使用该系统。

近期研究聚焦于将城市视为多层异质网络（如道路-桥梁-建筑），通过图神经网络等方法建模灾害扰动下的功能连通性；重心正从静态结构韧性转向动态功能韧性与政策干预的因果评估。

近30天 25 近7天 5 来源 41 论文 205

趋势信号

- 多篇论文在《Sustainable Cities and Society》集中发表，主题均关联气候/灾害韧性政策的实证影响评估
- arXiv上出现基于R-GCN-VGAE的异质图建模工作，明确以元路径刻画基础设施的功能角色
- Nature子刊论文强调整合地理空间属性与基础设施属性以提升风险表征的操作性
- Smart Cities Dive博客反映政策趋势：关键基础设施韧性责任正向州/地方层级下放

核心观点

- 城市韧性本质是功能连通性而非单纯结构冗余，需通过多层网络（如道路-桥梁-建筑）建模保障基本服务可达性
- 单一指标（如度中心性）无法支撑灾害场景下的差异化维护决策，必须融合空间、拓扑与功能语义信息
- 政策干预（如AI政策、气候试点）被视作可识别的准自然实验，用于因果推断韧性提升效果
- 开放数据（OSMnx、开放基础设施数据）已成为构建城市异质图的基础支撑条件

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：长三角城市群多尺度路网开放数据已覆盖全部地级市（OSMnx+高德API），且地方政府正试点基于桥梁功能分类的差异化养护预算分配，亟需可迁移的分类框架。

关键难点

- 需定义跨城市路网子图的聚类系数阈值切片标准，而非直接复用全局指标
- 医疗可达型标签在长三角缺乏统一灾时通行能力实测验证，仅能依赖POI密度与应急响应时间仿真代理
- R-GCN-VGAE的元路径长度固定为3（道路-桥-建筑），无法适配长三角更密集中间节点层级

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

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。

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

人工智能政策与城市韧性：来自中国一项准自然实验的证据
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：凌志鹏（Ling-Zhi Peng）、金阳蔡（Jinyang Cai）。

SMART CITIES DIVE

新型关键基础设施网络安全框架包含政府与地区委员会选项
特朗普政府表示，希望将更多网络安全韧性责任下放至州级和地方层级。

近期研究重心从静态街景图像的表征学习，转向对感知过程机制（如注视行为）、空间推理可靠性（如VLMs的空间语义幻觉）及图像对齐质量等底层问题的深入探究。方法上强调多模态融合、结构化空间知识建模与工具链开源化。

近30天 10 近7天 3 来源 41 论文 176

趋势信号

- 出现专门面向街景图像视觉对齐的开源工具PairWise Image Finder，聚焦纵向变化分析中的图像配准质量量化
- 新数据集Place Pulse-Gaze同步引入眼动追踪与主观感知标签，推动人类感知过程建模成为城市感知研究新支点
- DM-KG提出方向-度量知识图谱（Direction-Metric Knowledge Graph），显式建模街景中实体间方位与距离关系以缓解VLMs的空间语义幻觉
- 多篇论文将街景感知指标（如视觉围合度、安全感知）与居民心理健康（elderly mental health, urban residents' mental health）进行实证关联

核心观点

- 街景图像的主观城市感知不能仅依赖端到端视觉建模，需纳入人类感知机制（如注视行为）作为建模范式的关键变量
- 现有VLMs在街景理解中存在系统性空间语义幻觉，其根源在于2D图像到3D空间关系的隐式推断不可靠，需结构化空间知识（如方位、距离）显式注入
- 高质量纵向对比的前提是视觉对齐而非元数据对齐，视角一致性与匹配特征覆盖度是街景变化检测中被低估但决定性的技术前提
- 城市感知指标（如安全、围合感）已实质性进入公共卫生与城市健康研究链条，成为连接建成环境与居民心理健康的可计算中介变量

RESEARCH IDEA

DM-KG在非均匀高密度建成区失效

DM-KG方法在武汉等非均匀高密度建成区街景中因深度估计失准导致实体三维坐标偏移，进而使方向-度量知识图谱中钟面方位角与欧氏距离关系系统性偏离真实空间拓扑

为什么现在值得做：PairWise Image Finder：一种面向城市感知研究的开源街景图像对齐视觉对齐查找工具 与 利用人类注视建模主观城市感知 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

- 需在无真值深度图条件下，构建适用于高密度建成区的深度误差敏感性评估指标
- 钟面方位角计算依赖稳定实体中心点定位，而骑楼立柱与广告牌支架在分割掩膜中易被误切为碎片
- DM-KG图谱构建依赖实体对间欧氏距离排序，但遮挡导致的深度塌缩会扭曲距离相对序

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

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。

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

DM-KG：一种提升视觉-语言模型在街景图像中空间认知能力的新方法

随着视觉-语言模型（VLMs）日益应用于地理空间问答与视觉场景理解任务，提升其在街景图像中进行复杂逻辑推理的空间认知能力已成为关键研究方向。然而，现有VLMs在感知真实街景中物体的位置、距离与方向时，常出现“空间语义幻觉”；此类错误往往难以追溯与校准，构成其在地理空间任务中实际部署的关键瓶颈。

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

USDOT OPEN DATA

车辆安全技术挽救的生命数量及相关联邦机动车安全标准

本报告（编号 DOT HS 812 069）所用 SAS 程序及辅助文件：《车辆安全技术挽救的生命数量及相关联邦机动车安全标准（FMVSS），1960 至 2012 年乘用车与轻型商用车》。

URBAN NEXT

展开、重织、连接：华集酒店室内空间的审慎更新

一项敏感的适应性再利用项目将南京小西湖片区内的四栋历史建筑改造为酒店及多功能复合目的地，同时保留了古城的风貌特征与空间品质。既有建筑状况各异：其中两栋为单层木构建筑，木构架与砖墙保存完好；另两栋则为窄进深的多层砖砌建筑，开窗有限，层高仅2.6米。

ARXIV

AeroMap3D：基于视觉-几何-语义地图先验锚定单目无人机6自由度定位

我们提出AeroMap3D，一种单目6自由度（6-DoF）无人机（UAV）定位系统，通过将机载图像锚定至视觉、几何与语义地图先验，实现无全球导航卫星系统（GNSS）环境下的导航。AeroMap3D应对地图参考式空中定位中的两大根本挑战：无人机影像与卫星地图之间的跨视角差异，以及裸地数字高程模型（DEM）与城市场景之间的结构不一致性。首先，我们引入一种轻量级适配器，使在互联网规模通用数据上预训练的稠密匹配器无需微调即可可靠执行无人机至地图的配准；该适配器通过估计无人机图像与地图瓦片之间的尺度比和偏航偏移，在稠密对应关系估计前消除由飞行高度、相机视场角及航向引起的主导几何错位。

ARXIV

从持续部署到可查询数据集：TB级AIS对齐的被动声学标注

长期被动声学监测部署产生大量录音存档，但这些录音未与船舶轨迹或相遇结构关联，导致作用距离与接触条件无法作为变量使用，且分析需依赖人工筛选。为克服这一局限，我们提出一种数据库原生工作流，将水听器录音与自动识别系统（AIS）位置报告对齐，生成距离解析数据。固定时长录音窗口与AIS报文被存储为持久化地理空间表，并通过索引化的时空连接进行关联，以单次、可扩展的基于集合的数据库操作取代内存中嵌套迭代，从而在不耗尽可用内存的前提下处理连续多年、含百万级窗口的归档部署。

USDOT OPEN DATA

美国国家公路交通安全管理局缺陷调查办公室（ODI）

外国召回行动数据包含美国国家公路交通安全管理局（NHTSA）当前及过往外国召回行动的信息，包括数据接收日期及具体制造商。用户可使用以下一项或多项检索条件在数据库中查询外国召回行动信息：— NHTSA 接收外国召回行动信息的日期；— 参与该召回行动的本土制造商名称。

THE VERGE AI

Suno 从 YouTube、Genius 和 Deezer 抓取了数百万首歌曲

据 404 Media 报道，一次黑客事件中获取的 Suno 数据显示，这款 AI 音乐生成器通过爬取 YouTube Music、Deezer 和 Genius 等在线音频平台上的数百万首歌曲及歌词进行训练。鉴于 Suno 一直未公开其训练数据集的具体内容及获取方式，此次事件提供了罕见的披露。

ARXIV

S²-VLA：面向自动驾驶的视觉-语言-动作模型中语义流与空间流的解耦

视觉-语言模型（VLM）在自动驾驶的高层推理任务中展现出显著潜力，但其根本性局限在于难以生成精确的低层控制动作。这一局限源于语义与物理世界之间的鸿沟，其根源在于离散的语言标记与连续的轨迹规划之间固有的不匹配。尽管视觉-语言-动作（VLA）架构试图通过将感知与控制统一为单一策略来弥合该鸿沟，但这种耦合反而引发新的瓶颈：标准VLA模型普遍遭遇严重的空间表征坍塌，不可逆地削弱了安全导航所必需的细粒度空间与几何先验知识。

CITIES TODAY

政府软件公司 Accela 收购 AI 平台 Civira

Accela 已收购 AI 平台 Civira，将人工智能能力整合至其配置与 本文首发于 Cities Today。