

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

THIS EDITION

五个方向的当日进展

S-Agent：空间工具使用激发空间智能推理

本期《研究日报》头版聚焦地理智能体崛起、多模态城市感知深化与韧性治理范式迁移。

现实世界中的空间智能需对连续演化的三维环境进行推理，而现有视觉语言模型（VLM）及工具增强型智能体仍主要依赖于对孤立静态视觉观测的无状态推理。

我们提出 S-Agent ——一种面向连续多视角图像与视频理解与推理的空间工具使用型智能体范式。

通过将空间推理建模为时空证据累积过程，而非孤立帧级预测， S-Agent 将空间感知从以帧为中心的识别转向以场景为中心的理解。

编者按：本期头版紧扣五大研究方向趋势，突出‘空间智能’作为交叉枢纽的统领性——从S-Agent对三维环境的具身推理，到Text-to-SpatialSQL对地理语义的理解跃迁；从乌托邦式参与实践对空间权利的再主张，到NASA商业数据采购对全球地理信息基础设施的重构。所有条目均体现‘技术能力提升’与‘人文价值重置’的双重演进。

TREND OVERVIEW

趋势综述：空间智能觉醒：地理大模型驱动城市。

近期研究聚焦于地理空间基础模型（GeoFMs）在具体下游任务中的适配机制与多模态融合路径，方法重心从单模态表征学习转向架构可控性、跨模态对齐及低资源微调策略。

近期研究重心从两两模态配对融合转向统一、可扩展的多模态联合建模，强调以场景内容为锚点的跨模态协同生成与理解。

近期研究重心从单一轨迹建模转向混合自主交通系统下的异质交互建模与闭环安全生成；方法上正从黑箱深度学习向神经符号化、风险约束扩散、概率流建模等可解释、可验证、可引导的生成范式演进。

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（GeoFMs）在具体下游任务中的适配机制与多模态融合路径，方法重心从单模态表征学习转向架构可控性、跨模态对齐及低资源微调策略。

2 多源多模态地理数据

近期研究重心从两两模态配对融合转向统一、可扩展的多模态联合建模，强调以场景内容为锚点的跨模态协同生成与理解。

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混合自主交通系统下的异质交互建模与闭环安全生成；方法上正从黑箱深度学习向神经符号化、风险约束扩散、概率流建模等可解释、可验证、可引导的生成范式演进。

4 复杂网络、韧性城市与地理模拟

近期研究聚焦于将复杂系统理论（如多尺度动力学、人工生命演化机制）嵌入城市基础设施韧性评估与地理模拟框架中，方法重心正从静态网络拓扑分析转向动态过程建模与跨灾种耦合响应推演。

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静态表征，转向关注视觉对齐质量、人类感知机制建模及VLM输出的可靠性与社会协商性。

HIGHLIGHTS

- 空间智能体突破静态视觉推理局限，迈向连续三维环境中的具身化工具使用与因果推理。
- 地理大模型正推动空间查询语言生成、多模态出行预测与遥感耦合分析进入新阶段。
- 城市自然被重新构想为‘技术自然’，数字中介系统持续重塑生态感知与治理逻辑。
- 市民参与式城市营造重回理论前沿，列斐伏尔空间权利思想为协同设计提供根本依据。

近期研究聚焦于地理空间基础模型（GeoFMs）在具体下游任务中的适配机制与多模态融合路径，方法重心从单模态表征学习转向架构可控性、跨模态对齐及低资源微调策略。

近30天 173 近7天 42 来源 57 论文 873

趋势信号

- Clay等地理空间基础模型正被显式用作辅助上下文模块（如注入U-Net瓶颈层），而非直接替代主干网络
- 栅格遥感数据与OpenStreetMap/Overture等矢量数据的互补性被反复强调，成为构建以人为中心GeoFM的关键动因
- 研究开始系统比较不同FM架构（仅编码器/编码器-解码器/掩码自编码）在相同预训练设置下的灵活性与任务性能权衡
- 零样本开放词汇遥感视觉定位（RSVG-ZeroOV）等新范式出现，依赖冻结通用VLM+DM协同推理，规避标注依赖

核心观点

- 地理空间基础模型（GeoFMs）的核心价值在于提供可迁移表征，但其下游有效性高度依赖适配方式（如LoRA微调、解码器设计、信息提取深度），而非仅模型选择本身
- 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，后者显式编码离散地理实体的几何、拓扑与语义关系
- 当前GeoFMs仍普遍缺乏对人类尺度语义（如功能、用途、社会属性）的建模能力，推动‘以人为中心’的多模态表征成为共识方向
- 自监督预训练虽能利用PB级未标注EO数据，但其表征在深度维度上的信息分布具有独特性，标准密集预测头常与其不匹配

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：公用事业部门亟需在山地城市快速识别滑坡风险点（GWF 2026），而当前 GeoFM 缺乏地形形变感知能力导致误报率高；Clay v1.5 与 OSM 数据均已开源，支持可控实验复现。

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

- 基于 L4S 元数据提取坡度与曲率统计量，按双阈值划分高形变子集并人工校验矢量-影像配准误差
- 在混合模型瓶颈层插入地形形变感知门控模块，以曲率梯度图作软掩码约束矢量注意力权重分布
- 采用 CKA 分析不同地形子集中栅格与矢量表征空间的线性可分性衰减程度，定位对齐失效深度层

REPRESENTATIVE ITEMS

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, 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

自监督遥感视觉模型如何迁移到下游任务

自监督地理空间基础模型（GeoFMs）能从遥感数据中学习可迁移表征，但其在下游任务中的行为难以刻画。我们研究了六种代表性GeoFM，涵盖联合嵌入、重建与多模态预训练三类方法，并在不同标签可用性 & 下游流程下，于分类、回归与分割基准上评估其迁移性能。结果表明，模型在不同任务及适配设置下的排名会发生变化。

GEOSPATIAL WORLD

面向公用事业与网络管理的GeoAI与数字孪生优化

全体会议3 | 2026年地理空间世界论坛（Geospatial World Forum 2026）| 阿姆斯特丹，4月28日至5月1日。本次全体小组讨论由2026年地理空间世界论坛主办，汇聚了来自电信、地理空间等领域的行业领袖。

近期研究重心从两两模态配对融合转向统一、可扩展的多模态联合建模，强调以场景内容为锚点的跨模态协同生成与理解。

近30天 260

近7天 68

来源 56

论文 1279

趋势信号	核心观点
- MetaEarth-MM 提出场景中心联合建模范式，解耦场景表征推断与模态生成，支持五种遥感模态间任意翻译 - Earth-OneVision 构建20亿参数RS-MLLM，在单一自回归框架内统一六类传感器模态与九类任务，引入FGVLA、SLIS、PCMA三项专用机制 - FusionRS 首次构建大规模RGB-红外-文本对齐数据集，显式区分scene captions与IR-aware captions两类文本监督 - TSMNet 和 SGMA 均强调文本模态或语义原型作为跨模态对齐的引导信号，而非仅依赖视觉特征级融合	- 多模态遥感建模正从‘外观映射’转向‘场景内容解耦’，底层场景表征被视为跨模态一致性的核心锚点 - 模态缺失（IMSS）与模态异质性（如光学/SAR/红外物理成像差异）是实际部署的关键瓶颈，需在模型设计中显式建模而非简单对齐 - 文本不仅是辅助监督信号，更是弥合视觉模态与现实世界概念语义鸿沟的知识载体，催生开放词汇与IR-aware等细粒度文本建模需求 - 统一多模态基础模型需兼顾通用性与模态特异性：既支持跨模态泛化能力，又保留各传感器物理特性所决定的判别性线索

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：城市洪涝应急响应需在SAR穿云观测受限时依赖光学补全，而当前LULC制图系统已部署SGMA类模型；FusionRS与Earth-OneVision提供的RGB-IR及六模态对齐数据集，使SAR-光学联合物理建模具备实证基础。

关键难点	建议切入
- 需从SAR元数据中提取入射角、方位角、极化方式三元组并嵌入SGF模块的原型生成过程 - 缺乏公开标注的SAR-光学共视区方向性偏差真值（如建筑物边缘方位角误差分布 - SGMA原架构中SGF模块输出为固定维度原型向量，无法承载方向敏感的旋转等变表示	- 在SGMA的SGF模块前端插入SAR物理编码器，将入射角与方位角映射为二维正弦位置嵌入，并与视觉特征拼接 - 基于GeoHeight-Bench中已标注的建筑物三维轮廓，反演其在SAR图像中的方位角投影偏差，构建方向性类内差异监督信号 - 将SGF原型空间由欧氏空间替换为圆群SO(2)上的流形嵌入，使原型匹配具备旋转等变性

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。

近期研究重心从单一轨迹建模转向混合自主交通系统下的异质交互建模与闭环安全生成；方法上正从黑箱深度学习向神经符号化、风险约束扩散、概率流建模等可解释、可验证、可引导的生成范式演进。

近30天 440	近7天 102	来源 69	论文 1851
----------	---------	-------	---------

趋势信号

- 新数据集强调真实V2X通信动力学（如CAMASA），而非传感器中心或合成轨迹
- 多个工作（DRIFT、RiskFlow、TraCS）明确将‘混合自主’（mixed-autonomy）作为核心建模前提
- 统一基准（CityTrajBench）被提出以解决轨迹生成方法间因实验协议碎片化导致的不可比问题
- 奖励函数设计（如MBRF）开始显式耦合交通效率与环境目标（如CO2排放），超越传统延误/排队指标

核心观点

- 真实世界轨迹建模必须反映车路协同（V2X）通信机制与基础设施感知特性，而非仅依赖车载传感器数据
- 混合自主交通中的异质性（HVs/AVs共存）引发行为分布偏移与长尾风险，需在生成与评估中显式建模
- 轨迹生成不能止步于统计拟合，须嵌入交通法规、车辆动力学与安全约束，实现合规性（compliance）与可执行性（executability）双重保障
- 扩散模型、流匹配（flow-matching）、概率流传输等生成范式正替代GAN/VAE，因其更适配闭环仿真与风险引导需求

RESEARCH IDEA

CAMASA数据集的V2X轨迹在非基础设施密集城区失效

CAMASA所依赖的协同感知消息（CAM）在缺乏路侧单元（RSU）覆盖的低渗透率城区无法表征真实车际交互动力学，因其隐含假设所有车辆均持续广播高精度状态且通信延迟可忽略

为什么现在值得做：CAMASA：源自MASA Living Lab的基于CAM的数据集 与CityTrajBench：面向城市尺度车辆轨迹生成的统一基准 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

- 需定量界定RSU覆盖率阈值与CAM消息丢失率、时间戳偏移量、空间定位漂移量之间的函数关系
- 须在真实低RSU城区部署轻量级数据探针以校准CAM缺失模式，而非仅依赖SUMO仿真
- CAM消息中车辆ID匿名化处理导致无法追踪同一车辆在RSU覆盖空洞区间的轨迹连续性

建议切入

- 基于OpenStreetMap路网与公开RSU部署图（如工信部《车联网直连通信频率使用情况公示》）构建RSU覆盖率空间栅格层
- 在选定低RSU城区（如成都高新区非主干道片区）布设5台车载OBU设备，同步采集原始CAM与GNSS-RTK真值轨迹，用于量化消息缺失模式
- 设计CAM可观测性指标（COI）：定义为单位长度轨迹内有效CAM更新频次与理论最大频次之比，并与局部RSU密度做空间回归

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

DRIFT：面向混合自主交通生成的风险约束扩散模型与模仿先验
未来智能交通系统预计将长期演进为混合自主（mixed-autonomy）范式，即人类驾驶车辆（HVs）与自动驾驶车辆（AVs）在高度耦合的交通生态系统中共存。此类共存引发显著的异质性、加剧的不确定性以及日益复杂的车际交互动力学。在此背景下，如何同时刻画随AV渗透率动态变化而产生的异质行为分布偏移、在强车际耦合约束下生成多样且可执行的轨迹、并对罕见但高影响事件开展可靠的闭环安全与稳定性诊断，仍面临根本性挑战。

ARXIV

驾驶，快或慢？面向多模态地面移动的神经符号化运动预测引导
在包含行人、自行车、汽车和卡车等异构交通参与者的环境中，实现准确且可解释的运动预测，对安全自主导航至关重要。然而，当前最先进的方法仍主要为黑箱模型，缺乏对现实世界交通中法规性与行为性约束的显式建模。我们提出轨迹合规性塑形（Trajectory Compliance-Shaping, TraCS），一种神经符号化框架，通过可解释且概率化的谓词逻辑增强现有黑箱运动预测主干网络。

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。

近期研究聚焦于将复杂系统理论（如多尺度动力学、人工生命演化机制）嵌入城市基础设施韧性评估与地理模拟框架中，方法重心正从静态网络拓扑分析转向动态过程建模与跨灾种耦合响应推演。

近30天 24 近7天 1 来源 38 论文 170

趋势信号

- 多尺度路径发散（MSPD）等受重整化群启发的标量被引入GeoSimulation，用于量化CA类模型中局部转移律的异质性组织结构
- 高压输电网络研究构建统一框架，耦合灾害表征、脆弱性建模与宏观经济影响传播模型，实现九类单一灾害及复合灾害的可比性风险评估
- 城市洪涝韧性研究强调集成学习融合多源地理空间与基础设施属性，以提升关键设施风险表征的空间粒度与操作性
- 气候韧性试点政策评估与热浪事件响应分析均将城市作为‘物理-社会-技术’耦合系统，关注极端事件对空间功能网络（如应急服务供需、道路连通性）的级联扰动

核心观点

- 城市韧性不能仅依赖静态网络连通性指标，而需建模其在多时间尺度、多灾种扰动下的动态响应与恢复路径
- 地理模拟（尤其是CellularAutomata和基于代理的模型）正成为连接复杂网络理论与城市韧性实证评估的关键方法接口
- 关键基础设施（能源、交通、应急服务）的空间分布与网络拓扑共同构成韧性瓶颈，其风险需在‘设施属性+空间关系+外部冲击’三维框架下联合评估
- 气候适应已从环境议题升维为城市空间生产逻辑的根本约束，热浪等极端事件暴露出既有建成环境与公共服务网络的结构性脆弱

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：城市应急服务供需空间格局与网络韧性：中国石家庄多系统案例研究 与 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决策 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

- 需构造非标签依赖的桥梁功能ground truth：无法直接使用OSM标签，但石家庄无现成桥梁-医院实时通行时长或急救车GPS轨迹数据
- 元路径语义漂移量化困难：同一‘hospital’标签在茨城对应三级综合医院，在石家庄可能对应社区卫生站，二者在应急响应链中层级不同
- 异质图对齐不可行：茨城用道路层+桥梁层+建筑层三元组，石家庄应急设施含消防站、避难所、方舱医院等新增节点类型，图结构不兼容

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

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

通过多尺度路径发散引导人工生命中的开放式进化

人工生命中的开放式进化（OEE）通常依赖于不可解释的黑箱神经网络复杂度度量，致使类生命系统与物理复杂性理论脱节。我们提出MSPD（多尺度路径发散，记为DP），一种受重整化群启发的标量，用于量化局部转移律异质性在时间维度上的多尺度组织结构。MSPD在种群层面定义为实际演化轨迹的泛函，并以滑动窗口有限分辨率估计器形式计算；其两种实现方式之间的一致性被表述为一条命题。

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。

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。

近期研究重心从单纯依赖街景图像的静态表征，转向关注视觉对齐质量、人类感知机制建模及VLM输出的可靠性与社会协商性。方法上强调多模态融合（如眼动+图像、VLM+人类标注信度）、工具链开源化，以及将主观性、测量偏差和虚拟角色效应显式纳入建模框架。

近30天 18 近7天 4 来源 38 论文 169

趋势信号

- 出现面向纵向变化分析的视觉对齐工具（PairWise Image Finder），强调特征匹配与语义掩膜联合量化对齐质量
- 引入人类眼动追踪数据构建Place Pulse-Gaze数据集，将注视行为作为建模主观城市感知的显式输入
- VLM基准测试研究明确提出需报告inter-annotator reliability，并将标签空间与评分策略视为可协商的技术产物
- 多模态LLM代理研究系统考察虚拟角色（persona）对生成解释内容的社会经济与政治主题倾向的影响

核心观点

- 街景图像的元数据不足以支撑可靠的纵向变化分析，视觉对齐必须通过特征级与语义级联合评估来保障
- 城市感知本质上是主观的，其计算建模需嵌入人类感知过程（如注视机制）而非仅拟合图像统计模式
- VLM在城市感知任务中的有效性不能脱离人类判断的分歧性与拒答行为来评估，标注信度是模型性能解释的必要上下文
- 模型输出（尤其是解释性文本）受提示中虚拟角色的社会属性影响，表明城市感知建模需警惕隐含的价值预设与立场偏移

RESEARCH IDEA

注视引导模型在非西方城市失效因眼动模式与语义权重错配

注视引导的城市感知框架在首尔或上海街景中会因眼动轨迹分布与语义场景表征权重不匹配而显著降低预测一致性，因为其训练数据仅覆盖蒙特利尔与波士顿社区且未解耦文化特异性注视偏好与通用视觉显著性

为什么现在值得做：东亚超大城市正加速部署基于街景的街道品质评估系统，亟需可解释、可审计的感知模型；开源 PairWise Image Finder 工具已支持多城市街景对齐，为跨城眼动-图像配对提供数据基础。

关键难点

- 需构建首尔/上海本地眼动采集协议，覆盖不同年龄与通勤方式人群
- 须解耦 gaze map 中文化特异性成分（如招牌阅读路径）与通用显著性成分（如运动物体）
- 缺乏跨城市统一标注的感知维度真值集，无法区分模型失效源于 gaze 建模偏差还是语义表征偏移

建议切入

- 使用 PairWise Image Finder 对齐首尔江南区与蒙特利尔Plateau区同位点街景，控制视角与天气变量
- 在对齐图像上复现 Place Pulse-Gaze 实验范式，采集 60 名本地居民眼动与 5 维感知评分
- 构建双分支 gaze encoder：一支冻结预训练权重提取通用显著性，另一支微调以拟合本地眼动残差

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/>)

TRANSACTIONS IN GIS

Text-to-SpatialSQL

《GIS事务》（Transactions in GIS），2026年6月，第30卷第4期。

NATURE CITIES

将城市自然重新构想为技术自然（technonature）

数字中介系统正将城市环境转变为持续被感知与调节的生态体系，从而重塑自然在城市中的生产、治理与争议方式。

WIRED

本届世界杯：规模更大未必更好

史上规模最大的世界杯正使球迷、球员和主办城市承受极限压力，专家指出这仅是开端。

TRANSACTIONS IN GIS

一种用于多模态起讫点出行需求协同预测的多视角时空学习模型

《GIS学报》（Transactions in GIS），2026年6月，第30卷第4期。

ARCHDAILY

通过参与实现城市营造：来自《乌托邦时光2026》的启示

谁拥有城市的权利？亨利·列斐伏尔（Henri Lefebvre）的著述质疑控制城市空间的结构力量，转而将市民置于决策核心。其思想深刻影响了建筑学与城市设计的实践方式，推动了社区参与和协同设计的发展。这些理念构成了城市营造主题节《乌托邦时光2026》（Utopian Hours 2026）的核心议题；该活动首阶段于荷兰鹿特丹市举办，以纪念其创办十周年。

OGC BLOG

了解你的数据：地理空间集成的首要挑战

在我之前的博客文章中，我介绍了一些在地理空间集成项目中常见的挑战。本文将聚焦于其中的第一个挑战：了解你的数据。在我们开展数据集集成、发布或分析之前，必须首先理解所处理的数据内容。这包括回答一系列问题。

GISCIENCE & REMOTE SENSING

全球城市化与生态环境长期耦合协调分析：遥感视角

Volume 63, Issue 1, December 2026 .。

NATURE CITIES

背景气溶胶减排对地表城市热岛效应的差异化影响

许多国家正面临加剧的城市热岛（UHI）效应与空气污染的复合风险。尽管严格的气溶胶排放管控已改善背景空气质量，但其对UHI的影响仍缺乏深入研究。本研究利用大样本全球气候模拟表明，背景气溶胶减排在湿润气候区会增强暖季地表UHI，而在干旱地区则削弱该效应。气溶胶减少导致背景向下太阳辐射增加，进而使城市与乡村地表均升温。