



香港中文大學（深圳）
The Chinese University of Hong Kong, Shenzhen

香港中文大学（深圳）-慕思智慧家居物联网技术联合实验室
香港中文大学（深圳）-觅熵遥感技术联合实验室
香港中文大学-香港中文大学（深圳）空间智能实验室

智慧城市感知与通信实验室

潘文安

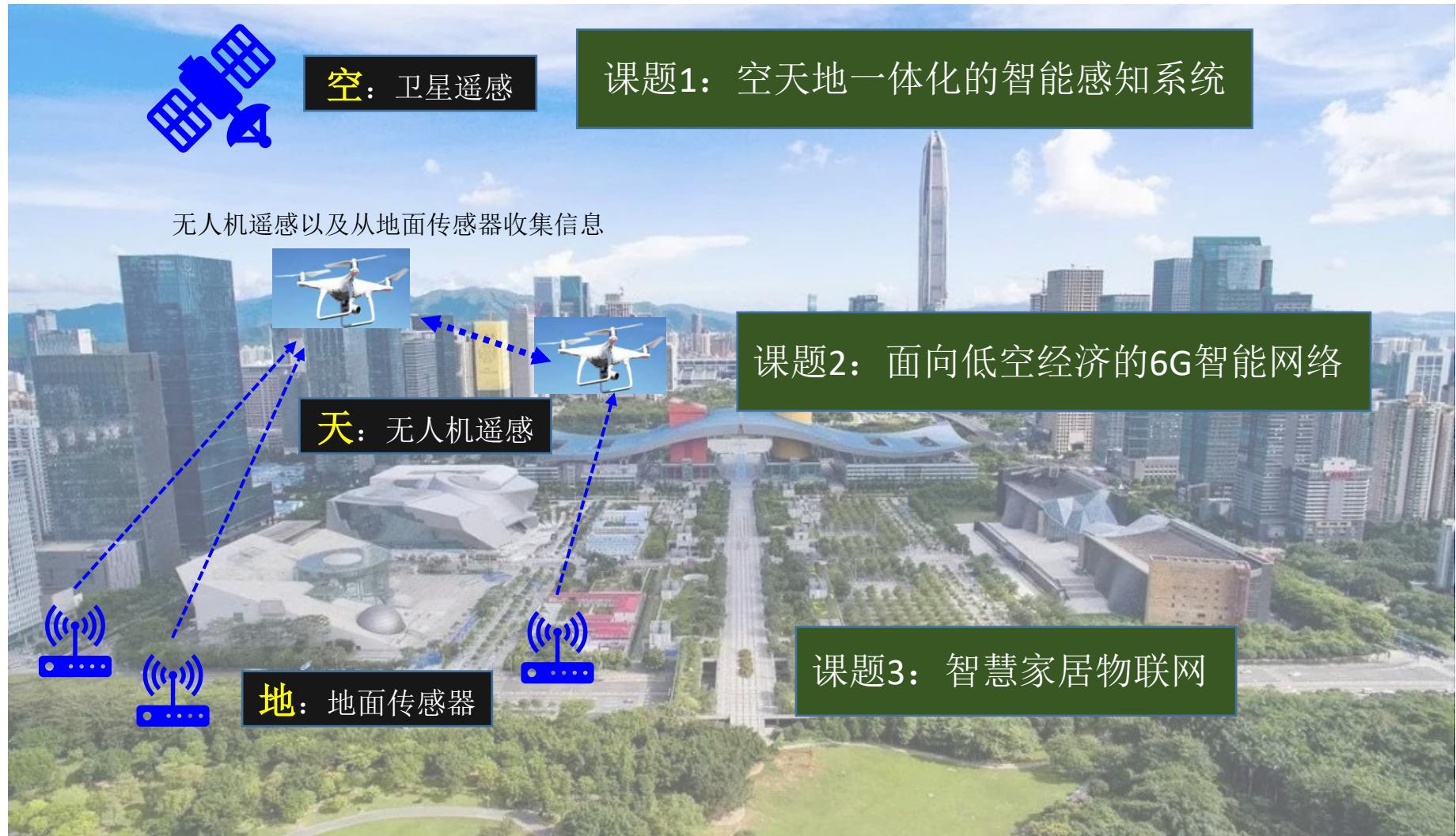
实验室主任

2025年9月20日



香港中文大學（深圳）
The Chinese University of Hong Kong, Shenzhen

面向智慧城市的智能感知与通信技术



博士生: 王茂南, 谢悦皎, 吴倩倩, 马丁, 庞奥宇, 王子尧, 罗艺欣, 王子溢, 章弘阳

硕士生: 汪思晋, 周芳冰, 苗纪凡

研究助理: 李斯玮

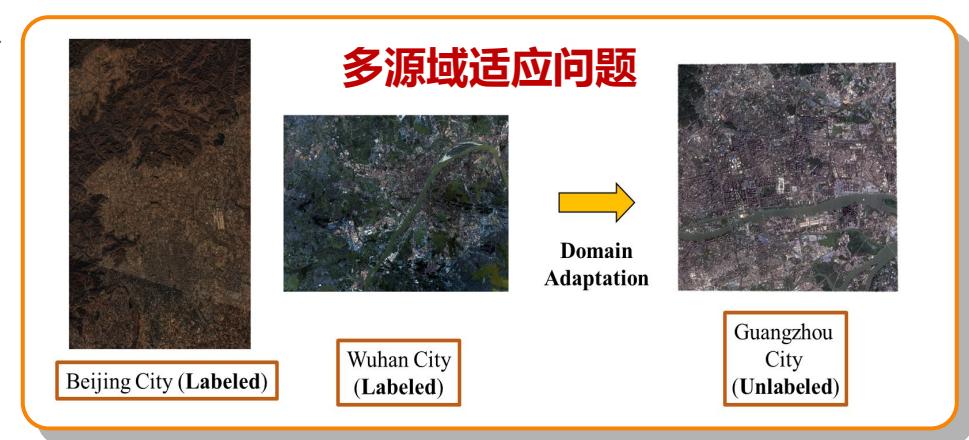


子课题1：基于多源域适应的遥感影像分类研究

研究问题：遥感影像的域适应研究在应对不同传感器、区域和采集时间下数据差异性方面具有重要意义，可提升模型跨域任务的泛化能力。然而，现有方法多集中于单一源域，忽略了多源域间的特征差异和协同效应。融合多源域能够整合不同源域数据，提升模型在目标域的适应性和分类精度。因此，本研究旨在通过该方法增强遥感影像分类在复杂环境下的鲁棒性与准确性。

主要挑战：

- ◆ 遥感数据来自不同传感器、区域和收集时间，导致源域间特征差异较大，需要提取共享特征；
- ◆ 源域数据的多样性可能带来噪声和类别不一致性，需有效处理以确保数据融合的有效性。



研究方法：

1. 提出一种多源无监督域适应方法，抑制源域特征差异与噪声影响，提升模型鲁棒性和跨域分类性能。
2. 在多种遥感任务上进行实验，包括农业监测、城市土地覆盖分类以及不同气候区的分类。通过覆盖不同应用场景，充分展示模型在复杂、多样化遥感数据上的适应能力和分类性能。

数据来源：

Jilin-1, Google Earth, Sentinel-2, Eurocrops

研究成果

Q. Wu, Y. Liu, Y. Zhong, K. Lin, X. Ma and M.O. Pun, "Source-free Multi-target Unsupervised Domain Adaptation for Cross-City Local Climate Zone Classification", major revision, *IEEE Transactions on Geoscience and Remote Sensing*, September 2025.

参与人员：吴倩倩 (博四)



课题1 空天地一体化的实时智能感知系统： 子课题2：面向下游任务的遥感图像去云

Funded by 联合实验室 &
HK RGC GRF Grant

研究问题：

现有的除云方法普遍依赖于图像相似性指标进行评价，但这些指标往往难以有效衡量，甚至可能产生误导。为改进除云效果的评估，本研究提出一种面向下游任务的评价思路，旨在通过下游任务的表现来检验除云算法是否真正恢复了对任务有用的信息。

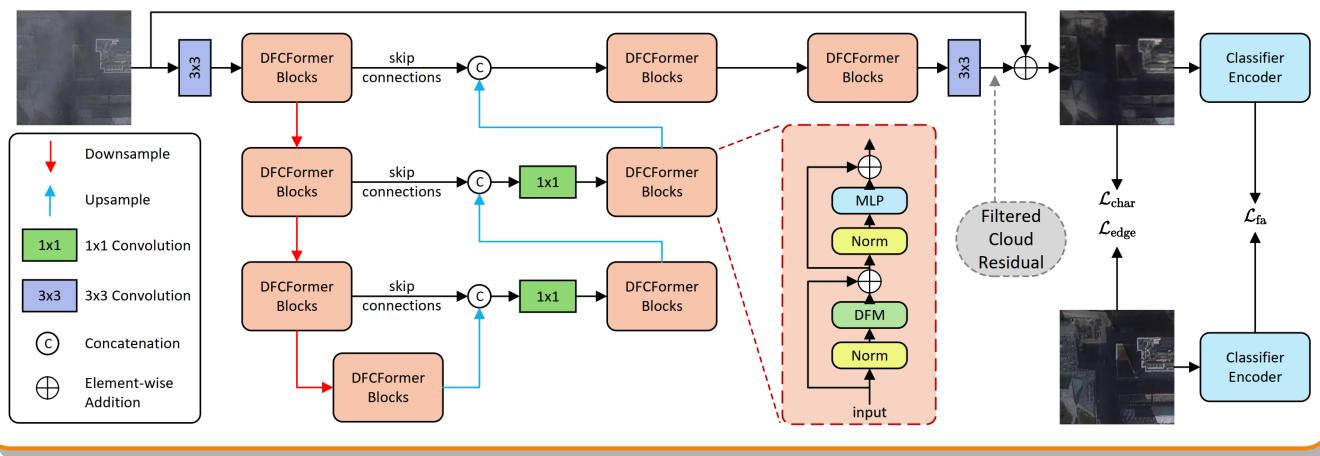
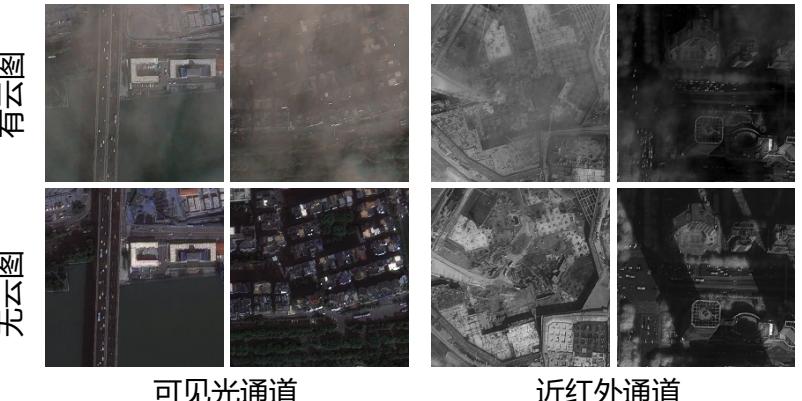
主要挑战：

- ◆ 云层形态复杂具有多样性
- ◆ 地物信息退化机理不明确
- ◆ 缺乏有效端到端评估数据

研究方法：本研究构建了 CUHKCR-EXT 数据集，并提出了一种基于动态滤波的 Transformer 模型，从而支持除云图像在下游任务中的质量评估，并兼顾薄云与厚云场景的普适性。

数据源：

- 自主研发高清数据集（30厘米空间分辨率）：
有云图、无云图、分类标签



研究成果

Z. Wang, X. Ma and M.O. Pun, "Downstream Task-Aided Cloud Removal for Ultra-High-Resolution Remote Sensing Images: An Information Loss Perspective", accepted for publication in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, Sep. 2025.

参与人员：王子尧（博二）



香港中文大學（深圳）
The Chinese University of Hong Kong, Shenzhen

智慧城市感知与通信实验室

课题1 空天地一体化的实时智能感知系统: Funded by 空间智能实验室

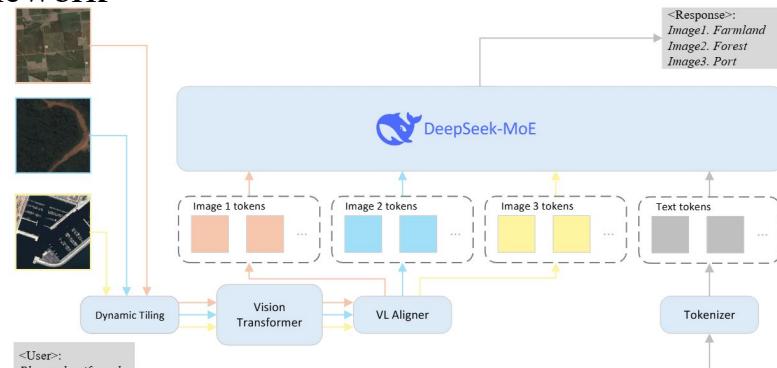
子课题3: 基于 DeepSeek-VL2 的遥感图像分类零样本域适应

研究问题:

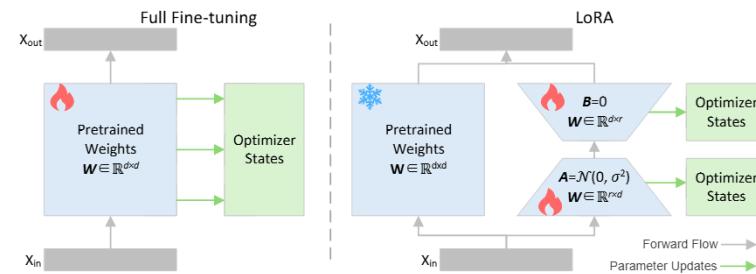
遥感图像分类在实际应用中常常面临 域间差异 (不同分辨率、成像条件) 和 标签空间不一致 的问题。传统方法在目标域数据不可用时失效，因此需要探索 零样本域适应 (ZSDA) 来提升模型的跨域泛化能力

研究方法

(a) framework



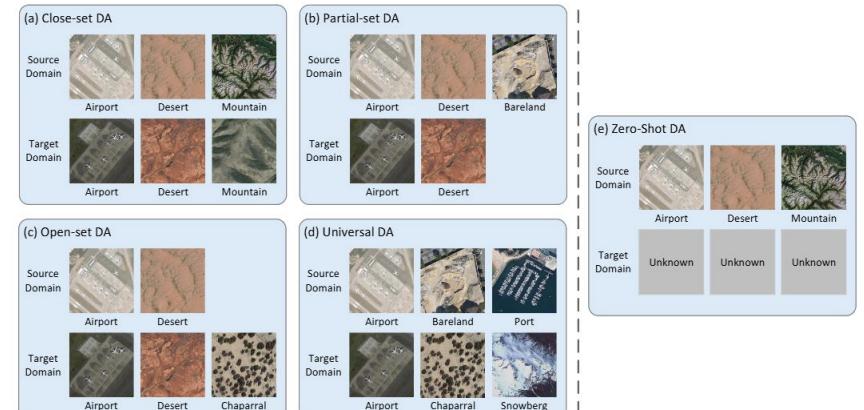
(b) finetuning



主要挑战:

- ◆ **遥感与自然图像差异:** 尺度、场景结构与背景纹理差异导致 VLM 直接迁移性能骤降。
- ◆ **标签空间动态变化:** 不同数据集定义的类别不一致，难以统一。
- ◆ **资源限制:** 主流 VLM 商业闭源、参数庞大，难以在有限算力条件下使用。

零样本域适应问题:



参与人员：王子尧（博二）

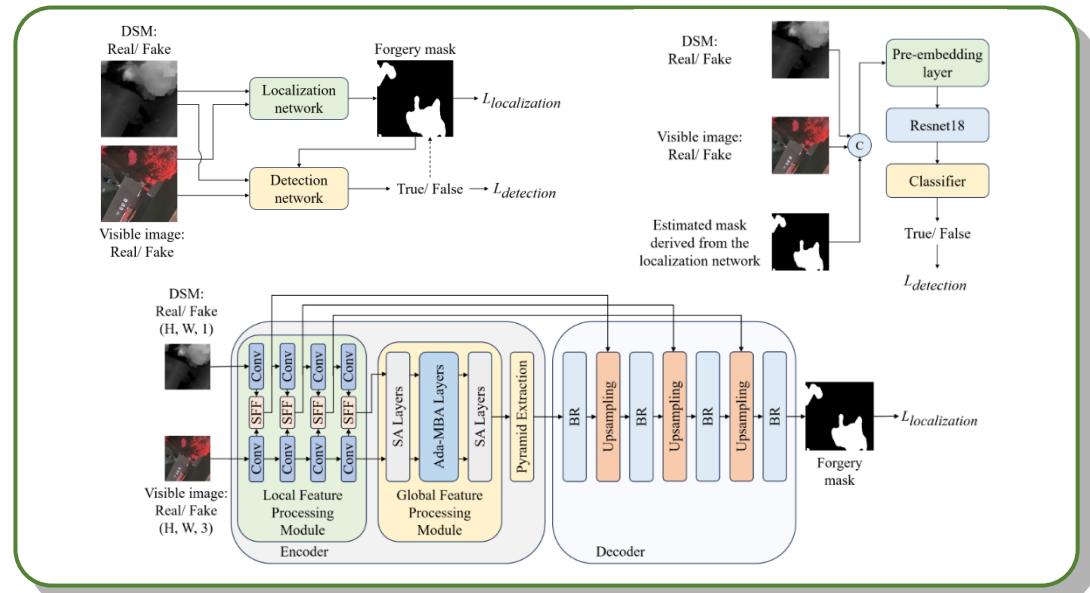
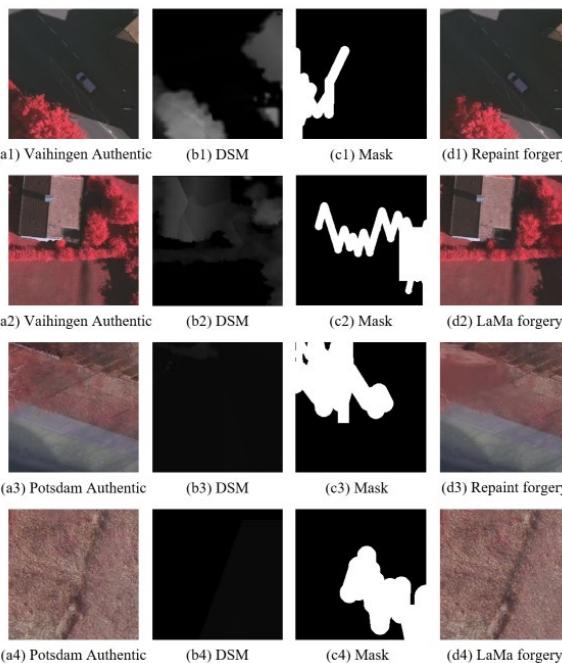


子课题4：卫星图像伪造检测与定位多模态融合技术

研究问题：尽管深度伪造技术已相当成熟，检测伪造的卫星图像的研究仍处于起步阶段。如何利用多模态数据实现针对深度伪造卫星图像的识别以及精确定位愈发成为一个重要的研究问题。

主要挑战：

- ◆ 如何充分利用多模态遥感数据
- ◆ 如何通过特征融合提高模型准确率
- ◆ 如何构建通用性识别模型



研究方法：本研究采用伪造定位与检测协同框架，基于可见光图像及数字表面模型数据，利用多模态结构分别实现浅层与深层特征的提取及融合。基于卷积神经网络与Transformer的多模态结构能够有效识别出图像中的伪造区域，整体的协同框架则进一步提升了模型整体的准确率。

研究数据来源：

ISPRS Vaihingen Dataset, Potsdam Dataset, Fake-Vaihingen-MM & Fake-Potsdam-MM Datasets

参与人员：马丁（博三）

研究成果：

J. Sui, D. Ma, C.-C. Jay Kuo and M.O. Pun, "FLDCF: A Collaborative Framework for Forgery Localization and Detection in Satellite Imagery", IEEE Trans. on Geoscience and Remote Sensing, vol.62, pp. 1-13, 2024.



香港中文大學（深圳）

The Chinese University of Hong Kong, Shenzhen

智慧城市感知与通信实验室

课题1 空天地一体化的实时智能感知系统: Funded by CUHK-CUHKSZ 联培

子课题5: 基于扩散模型的遥感影像生成

研究问题: 对于变化检测任务,当前的开源数据集主要集中在建筑物提取和城市发展监测。在面对自然灾害监测时(以滑坡为例),往往缺乏充分的数据,这种情况阻碍了针对自然灾害的变化检测技术进一步发展。本研究旨在利用扩散模型生成逼真的滑坡影像,解决滑坡影像稀缺的挑战。

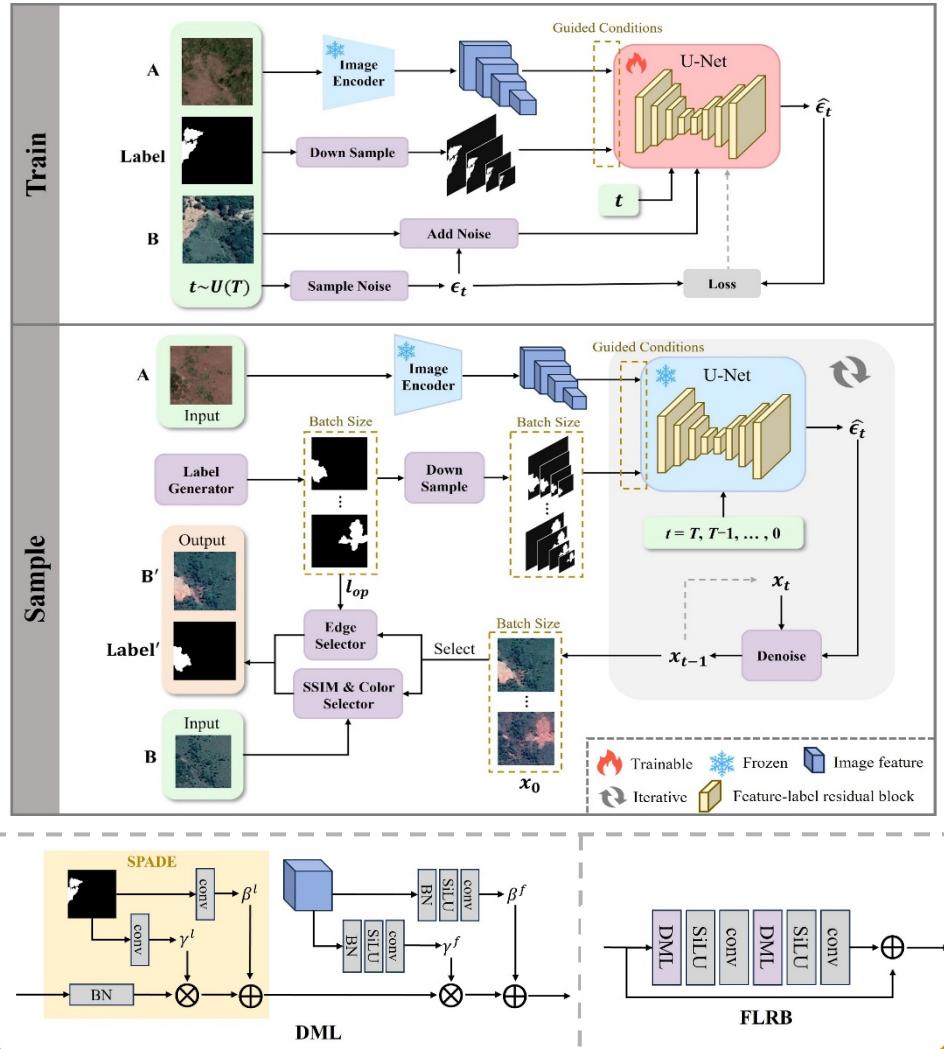
主要挑战

- 缺乏充足的滑坡正例数据
- 需要生成与滑坡前影像和掩膜对应的滑坡后影像
- 需要考虑生成数据的使用方法,即生成的数据用于变化检测任务时需要实现精度提升

研究方法:

1. 以DDPM为基础,将掩膜和滑坡前影像作为条件处理后输入模型,引导滑坡后影像的产生;
2. 在推理阶段,采用生成器自动生成一批次的掩膜,进而生成若干滑坡影像,最后采用选择机制选出最佳的伪影像;
3. 将生成的伪滑坡影像按照不同比例与真实影像一起训练,提出新的训练方法来训练变化检测网络,分析不同比例的伪影像对网络精确性的提升程度。

关键技术



参与人员: 罗艺欣 (博一)

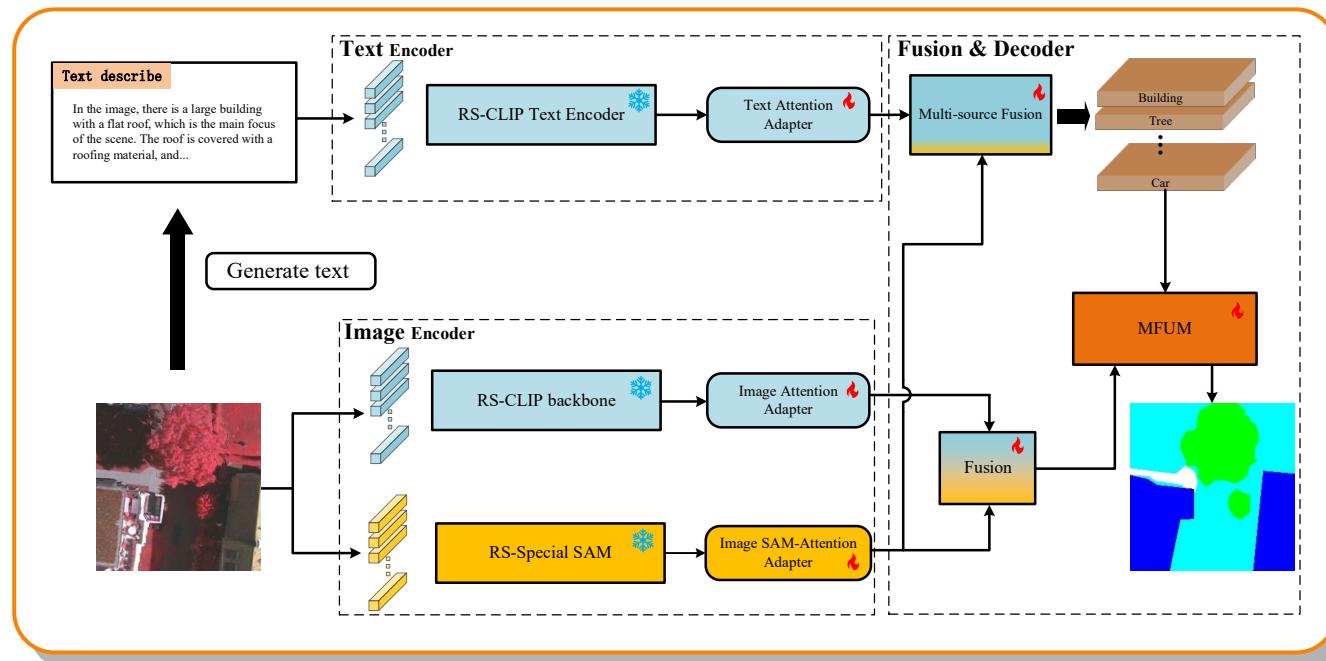


子课题6：基于CLIP的多模态视觉语言大模型遥感语义分割

研究问题：传统遥感语义分割研究主要集中在单模态视觉特征空间的表征学习，但往往忽略了图像之外的文本级语义信息。而随着大模型的快速发展，如何充分利用文本数据的潜力来融合构建多模态视觉语义遥感大模型成为了迫切需要解决的问题。

主要挑战

- 遥感任务中图像和文本数据源差异性较大
- 需要改构建适合的SEG-adapter来进行分割任务
- 如何生成更为有效果的复制遥感文本数据源

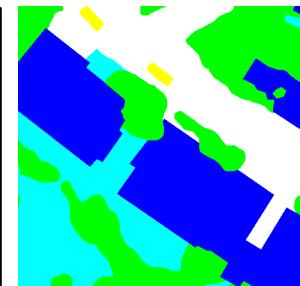


图像及文本数据



Text

In the image, there is a residential neighborhood with several houses, a road with a car traveling on it, and surrounding vegetation distributed around the buildings and along the road.



研究方法：

1. 以RS-CLIP作为主要的文本信息提取器，SAM做完主要的图像特征提取；
2. 构建多模态数据源的融合模块，以解决不同数据源在直接融合导致的负面影响。
3. 探索更为合适的遥感文本

参与人员：王子溢（博一）



课题1 空天地一体化的智能感知系统：

子课题7：3D知识引导的无人机-卫星跨视角定位技术

研究问题：基于无人机的跨视角地理定位任务是地学与遥感领域的重要基础任务，能够为下游应用提供精确的空间位置信息。近期随着3D视觉表征学习发展，通过多视角图像重建丰富精确的3D点云场景，提供了不同视角统一的表征信息。将3D先验知识引入到地理定位任务能够有效提升模型定位的鲁棒性。

主要挑战：现有跨视角定位方法往往依赖图像特征的直接相似性度量，难以抵御极端视角差异带来的语义退化。同时，模型在跨区域泛化时易受到域偏移影响，导致匹配结果不稳定且判别性不足。

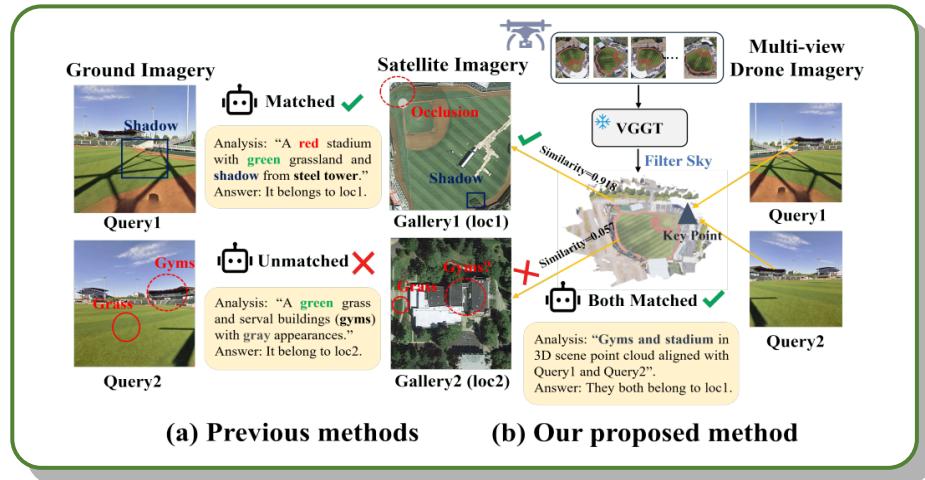
研究数据来源：University-1652, SUES-200, DenseUAV

研究方法：本研究主要围绕三个模块：首先，利用重建生成3D表征对各2D视角的特征；其次，为促进各视角表征的优化，进一步对各视角冗余信息进行剔除；最终，利用3D空间知识对各2D视角关系蒸馏。

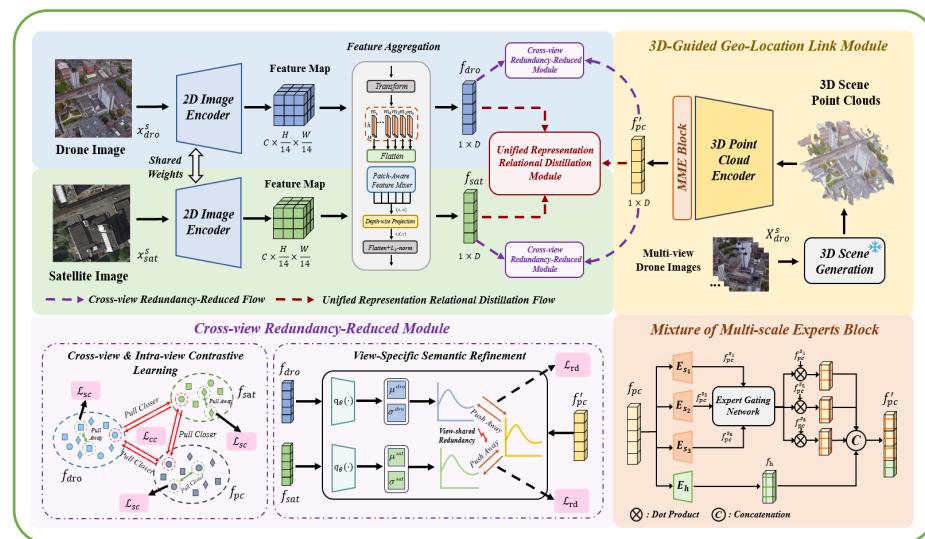
研究成果：

H. Zhang, Y. Liu, Z. Kuang, "SkyLink: Unifying Street-Satellite Geo-Localization via UAV-Mediated 3D Scene Alignment," *Proceedings of the 3rd Workshop on UAVs in Multimedia: Capturing the World from a New Perspective*. 2025.

参与人员：章弘阳 (博一)



- ◆ 如何克服一对匹配中存在的视角间隔问题
- ◆ 如何使本方法能够有效利用3D先验知识



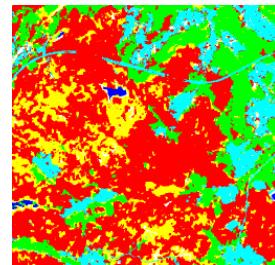
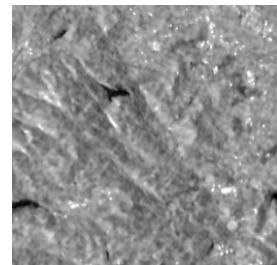
子课题8：基于SAR和带云光学图像融合的土地覆盖识别

研究问题：光学遥感因其成像机制存在易受云层遮挡导致地面部分信息缺失的问题，严重制约了大区域连续监测的可靠性。相比之下，SAR具备全天时、全天候的工作能力，能有效穿透云层获取地表信息。因此，融合SAR影像进行光学遥感影像缺失区域语义重建的方法受到广泛关注。本研究旨在云遮蔽区域多模态遥感协同解译，实现对受云层覆盖导致的数据空洞区域进行语义重建。

主要挑战

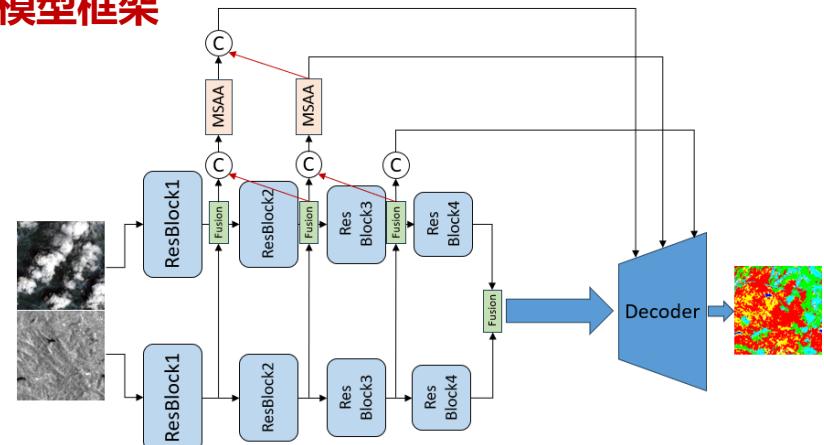
- ◆ 云层遮蔽区域地表信息缺失，需充分挖掘 SAR 结构特征融合光学模态，实现云下信息的精准重建。
- ◆ SAR与光学影像存在显著模态差异，跨模态特征对齐与一致性保持困难
- ◆ 云层作为强噪声扰乱融合，云下信息重建困难

数据集：



SEN12MS-CR/SEN12MS

模型框架



研究方法：

1. 沿通道维度对双模态特征图进行拼接，随后通过压缩-激励机制削弱光学影像云遮蔽区域的影响，并强化 SAR 数据权重，以更精准地还原云层下方结构细节。
2. 将深层特征图上采样至浅层特征图的空间分辨率，聚合多尺度信息；借助高层语义引导，有效抑制浅层特征中的云层噪声，减少其经跳跃连接传入解码器所带来的干扰，从而提升语义分割的准确性与鲁棒性。

参与人员：李斯玮（科研助理）



香港中文大學（深圳）

The Chinese University of Hong Kong, Shenzhen

智慧城市感知与通信实验室

10

课题2 面向低空经济的6G智能网络：

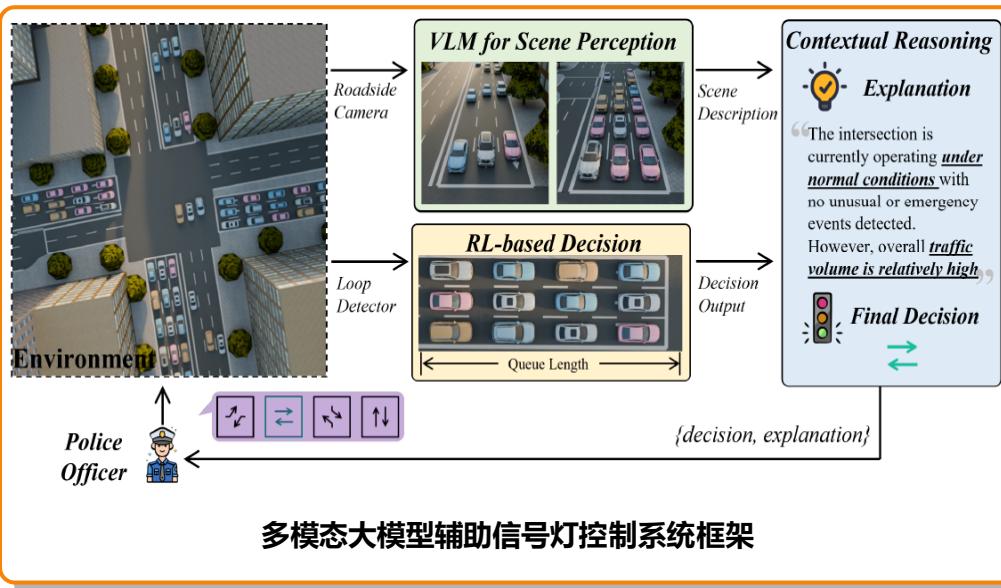
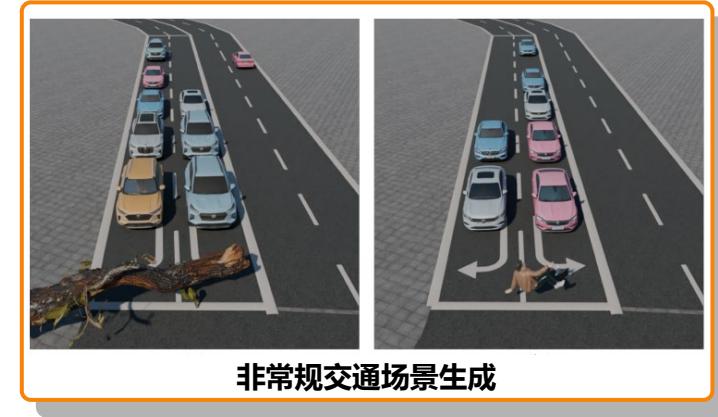
子课题1：基于多模态大模型辅助的交通信号控制方法

研究问题：

交通拥堵是全球性问题，导致环境恶化、经济损失和生活质量下降。优化交通信号控制（TSC）是缓解拥堵的有效途径。但是现有的方法异常事件处理能力弱，且决策缺乏可解释性，无法和人类协同工作。

主要挑战：

- ◆ **异常事件处理能力弱**：无法有效应对道路树木遮挡、行人摔倒等非常规交通场景；
- ◆ **决策缺乏可解释性**：在安全关键场景中，系统决策过程不透明，无法向交通管理人员或说明决策依据；
- ◆ **人机协同缺失**：决策未充分贴合人类逻辑，无法与交通管理人员形成有效协作；



研究方法（双核心模块设计）：

- ◆ **VLM-based 场景感知模块**：以路边摄像头实时图像为视觉输入，基于 VLM 生成场景描述；
- ◆ **上下文推理模块**：接收 VLM 生成的场景描述和强化学习的决策建议，结合交通规则模仿人类推理，输出决策，同步生成决策依据说明，确保透明度，支持人机协同验证。

研究成果：

Wang, Maonan, Yirong Chen, Aoyu Pang, Yuxin Cai, Chung Shue Chen, Yuheng Kan, and Man-On Pun. "VLMLight: Traffic Signal Control via Vision-Language Meta-Control and Dual-Branch Reasoning." accepted for publication at the Advances in Neural Information Processing Systems (NIPS), San Diego, U.S.A, Dec. 2-7, 2025.

参与人员：王茂南（博四）



子课题2：整合 RL 和 VLM 的交通信号控制算法

研究问题

随着城市车辆激增，交通拥堵加剧，传统信号控制方法，因依赖预定义参数难以适应动态流量。强化学习（RL）虽能利用实时数据实现自适应优化，但在通信延迟、数据缺失、噪声和应急车辆等复杂场景下鲁棒性不足。已有研究引入大语言模型（LLM），但受限于提示词依赖和缺乏视觉理解，难以应对新型交通场景。因此，构建兼具适应性与泛化性的智能信号控制框架仍是亟待解决的挑战。

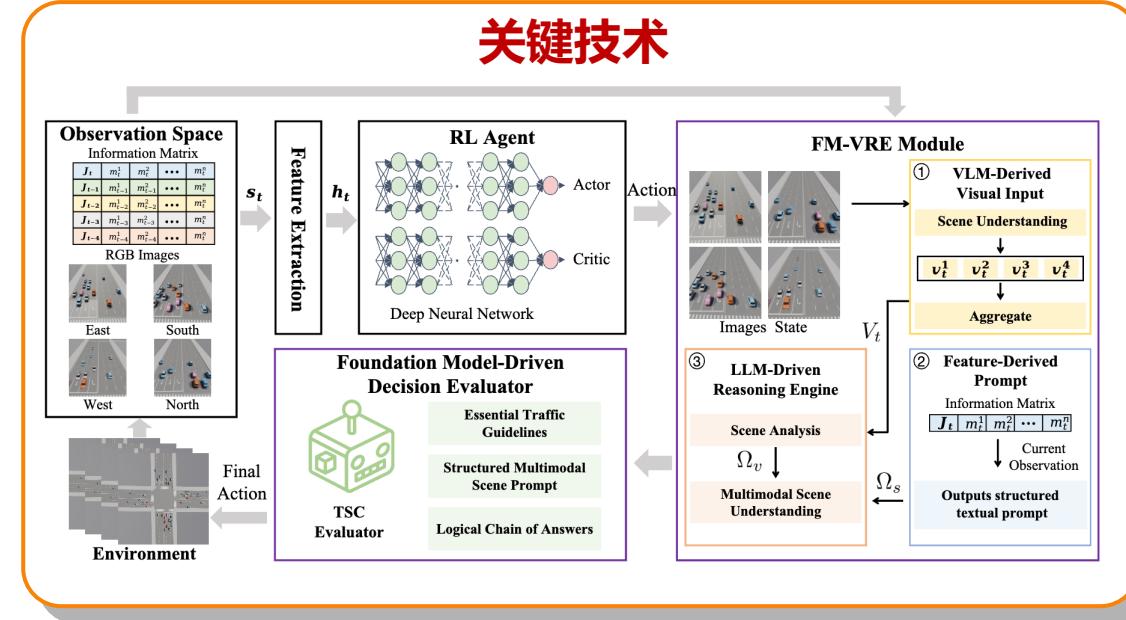
研究方法

- 融合视觉与语义：使用 VLM 直接处理路口摄像头图像，生成高层次语义表示。利用链式推理，提升因果推理能力，预测候选动作的效果。
- 强化学习结合：RL 负责高效的策略优化与环境自适应决策。VLM 提供语义增强的环境建模，辅助 RL 提升鲁棒性与泛化性。

研究成果

A. Pang, M. Wang, M.O. Pun, C.S. Chen and X. Xiong, "iLLM-TSC: Integration reinforcement learning and large language model for traffic signal control policy improvement", submitted to *IEEE Trans. on Vehicular Technology*, June 2025..

关键技术



主要挑战

- ◆ 动态与不确定性：现实交通环境高度动态且不可预测，RL 方法难以保持稳定性。
- ◆ 大模型的局限性：现有 LLM 缺乏视觉理解，无法直接处理路口摄像头视频；依赖手工 prompt，泛化性不足，难以部署在开放环境。

参与人员：庞奥宇（博三）、王茂南（博四）



课题2 面向低空经济的6G智能网络： 子课题3：基于DRL的按需空中出租车选择优化

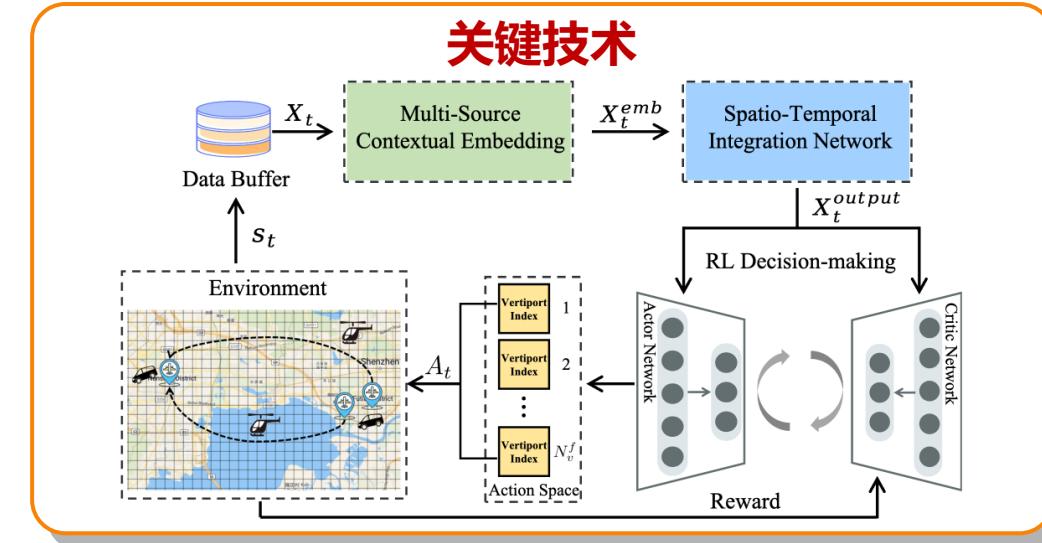
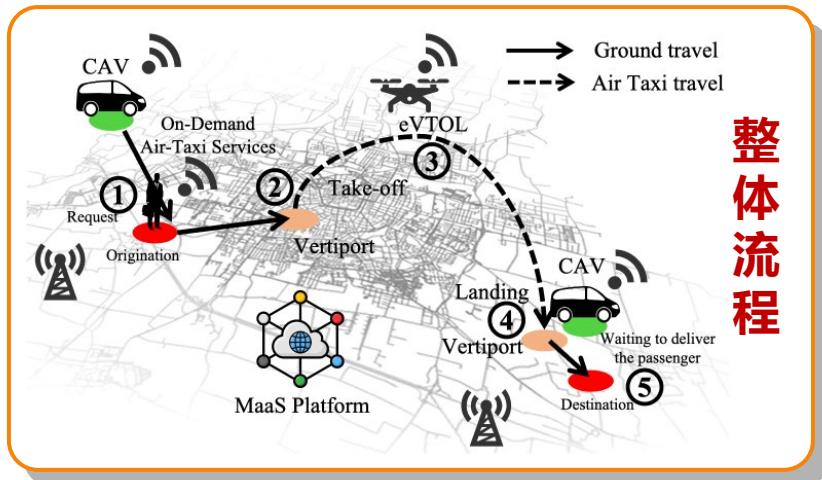
Funded by 科技部

研究问题

- 随着城市化进程的不断深入，交通拥堵、环境污染、通勤时间过长等问题也日趋严重
- 城市空中出行 (UAM) 利用低空空域，提供按需空中出租车 (Air Taxi) 服务
- 如何在空中交通 (eVTOL) 与地面交通 (CAV/公共交通) 之间实现高效衔接？
- 关键问题：乘客如何选择起降点 (Vertiport)，并完成端到端出行策略优化？

主要挑战

- 构建地面交通和低空交通的联乘模型
- 涉及大量动态和异构数据，难以全面获取
- 强化学习面对高维异构观测训练困难



研究方法

- 本研究首次提出了一个动态的空中出租车路线推荐系统，该系统同时考虑了实时的地面和空中交通状况；
- 使用强化学习框架，实现全局的乘客的动态线路推荐；
- 设计了多源数据编码和时序信息编码模块，以实现高效提取和融合高维异构数据的关键特征。

参与人员：庞奥宇（博三）、王茂南（博四）

研究成果：

A. Pang, M. Wang, Z. Sha, W. Yue, C. Li, C.S. Chen and M.O. Pun, "Heterogeneous Vertiport Selection Optimization for On-Demand Air Taxi Services: A Deep Reinforcement Learning Approach", submitted to *IEEE Trans. on Intelligent Transportation Systems*, Apr. 2025.



课题2 面向低空经济的6G智能网络：

Funded by 科技部

子课题4：基于无线电地图和乘客动态需求的UAM轨迹规划

研究问题

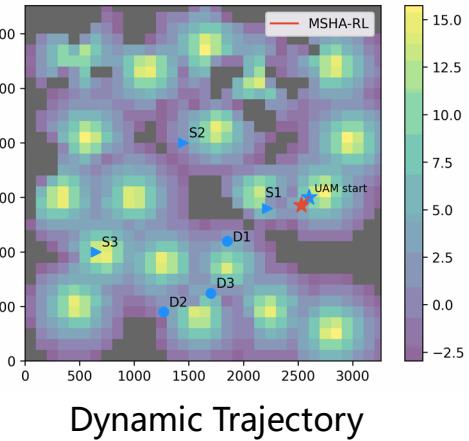
UAM(Urban Air Mobility)系统是近年来解决地面交通拥堵极具前景的方案。其中，UAM的轨迹规划是一个重要的研究领域。在实际生活中，确保系统安全并减少出行时间和距离是一个重大的挑战。我们构建了一个无线电地图(Radio Map)，并使用了SINR作为评价通信质量的指标。

因此，本研究以最小化旅行时间为目地同时考虑通信质量，优化UAM作为空中出租车的共乘轨迹。

实验结果



Radio Map



Dynamic Trajectory

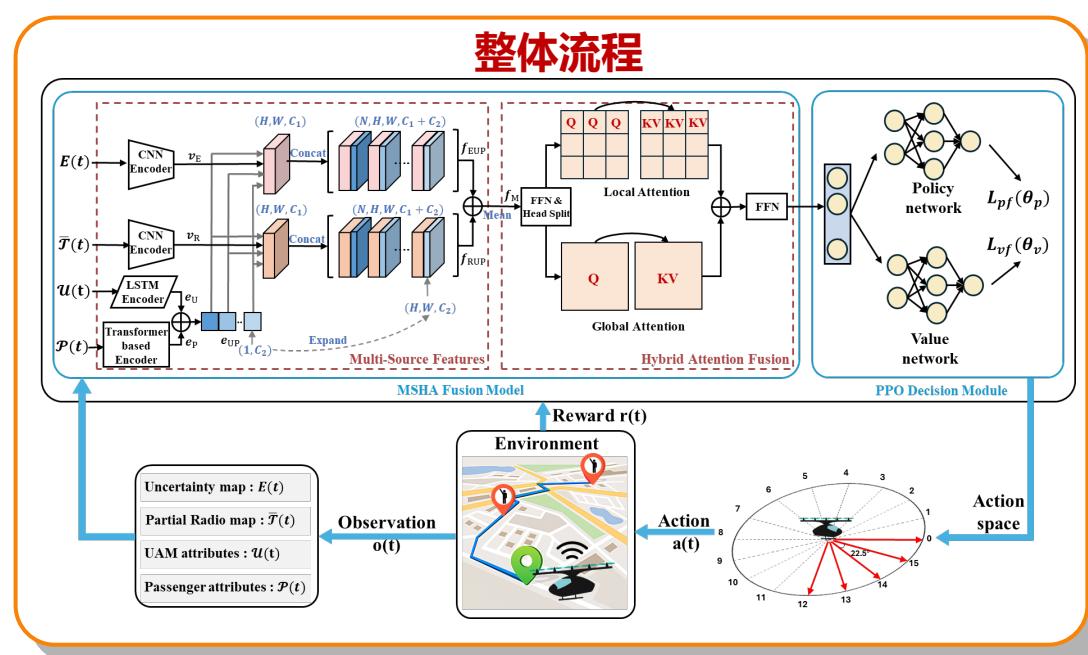
研究成果：

Y. Xie, M. Wang, D. Zhou, M.O. Pun and Z. Han, "Real-Time Communication-Aware Ride-Sharing Route Planning for Urban Air Mobility: A Multi-Source Hybrid Attention Reinforcement Learning Approach", submitted to *IEEE Trans. on Intelligent Vehicles*, July 2025.



香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen

整体流程



主要挑战

- ◆ 构建Radio Map，提取地图中每个点的SINR的值
- ◆ 大量地图数据，用户数据和UAM数据的多源融合
- ◆ 解决在线问题中用户时间和空间分布的不确定性和随机性

研究方法

传统的算法无法处理用户的动态需求，在实际中很难应用。我们设计了一种自适应强化学习(RL)框架，融合不同来源的环境信息，并结合全局和局部视角，实现实时的确保通信质量的路径规划。

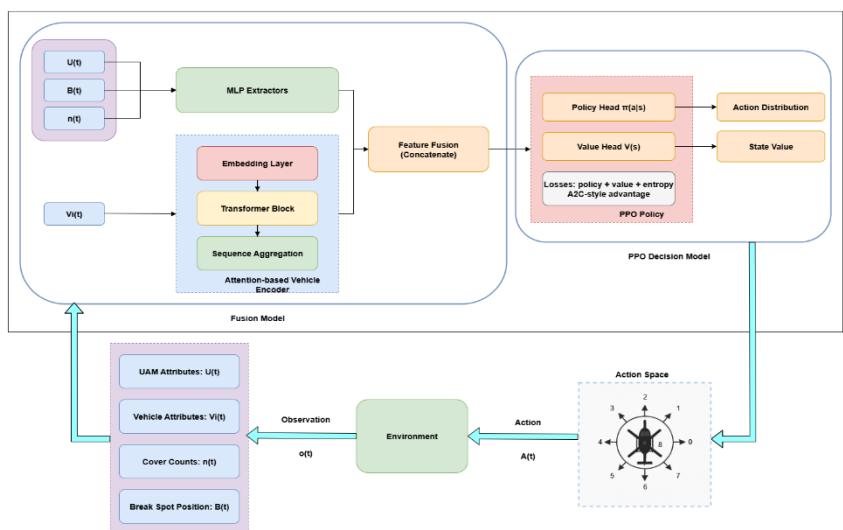
参与人员：谢悦皎（博四），胡泽云（大四）

子课题5：基于无线电地图和车辆动态位置的UAM轨迹规划

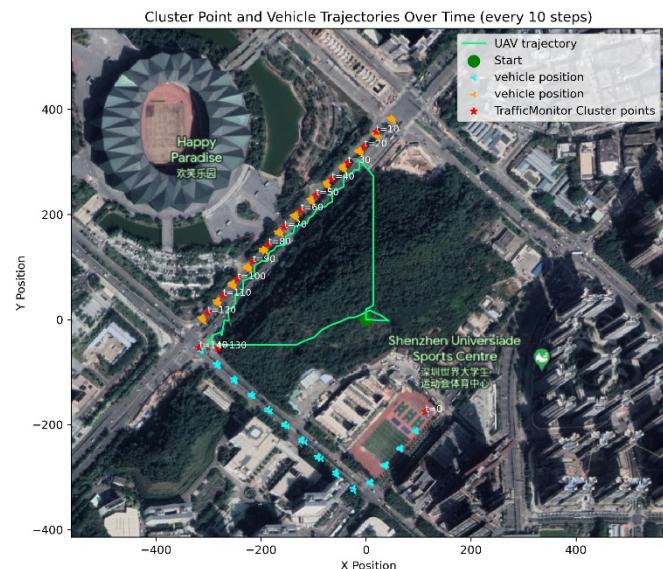
研究问题

在智慧交通与无人机应用场景下，如何利用无人机对道路车辆进行实时监控仍是一个重大挑战。当前，传统方法在动态环境中难以保证对车辆集群的稳定覆盖与跟踪。我们聚焦于在有限高度与覆盖半径条件下，最大化无人机监控范围内的车辆数目，并解决监控过程中因车辆分布动态变化带来的不确定性。

整体流程



实验结果



主要挑战

- ◆ 奖励函数的设计需在覆盖率、路径稳定性与能耗之间权衡。
- ◆ 非平稳环境导致训练收敛速度慢，策略易过拟合特定场景。

研究方法

我们设计了一个基于强化学习 (RL) 的框架，训练无人机智能体学习动态轨迹策略，以覆盖道路网络中最密集的车辆簇。通过奖励函数引导智能体在 3D 空间 (x, y, z) 中移动，并利用聚类和 KD-tree 等方法帮助识别最优覆盖簇，从而提升监控效率。

参与人员：谢悦皎（博四）



子课题1：基于深度学习的鼾声识别

研究问题：打鼾是睡眠中的一种非意识行为，它既能反映打鼾者本身的一些生命体征，也能监测人的生理健康状况。本研究以单片机作为硬件平台，使用深度学习与声源定位的方法，实现实时判断是否有鼾声并判断鼾声的方位。

研究数据

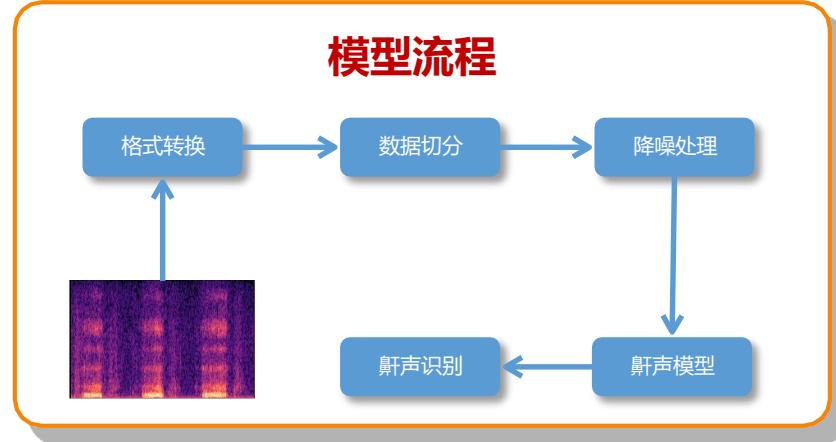
本研究的研究数据，部分使用公开数据集，共包括15000秒（4.16小时）左右的鼾声数据和15000秒左右的非鼾声数据。通过数据预处理将其统一为每帧16bit, 16000Hz采样频率，单声道的wav文件，方便后续处理。

主要挑战

- ◆ 研究数据的质量不一致，需要筛选
- ◆ 单片机性能有限，部署需要对内存和性能优化
- ◆ 实际应用场景中，多种类的环境噪音会对鼾声识别造成干扰
- ◆ 提高鼾声不同大小不同类型的识别率

参与人员：

汪思晋（硕二）



数据处理：主要使用短时傅里叶变换（STFT）将音频数据转换为图像数据，再利用CNN进行训练，窗函数为汉明窗。除此之外，使用了带通滤波、阈值分割等方法对数据进行降噪处理。

研究方法：

- 使用CNN在电脑进行训练，并保存训练模型
- 实现python转C部署，并进行模型量化，优化内存和性能
- 完成PCB集成，可实现鼾声识别及定位、气囊控制等功能
- 采用麦克风阵列实现声源定位，以确定鼾声的方位
- 在真实环境中进行实验测试

实验结果： 鼾声识别准确率可达90%以上



香港中文大学（深圳）

The Chinese University of Hong Kong, Shenzhen

智慧城市感知与通信实验室

16

子课题2：基于大语言模型(ChatGPT)的学习平台开发

研究问题：本研究专注于基于大语言模型（如ChatGPT）的学习平台开发，旨在利用先进的人工智能技术，通过个性化的学习内容和反馈机制，提高语言学习的效果与效率。相比传统的统一教学模式，该平台能够针对学生的学习情况进行精准分析，自动生成符合个体需求的练习和指导，为学习者提供更智能、更有效的学习体验。

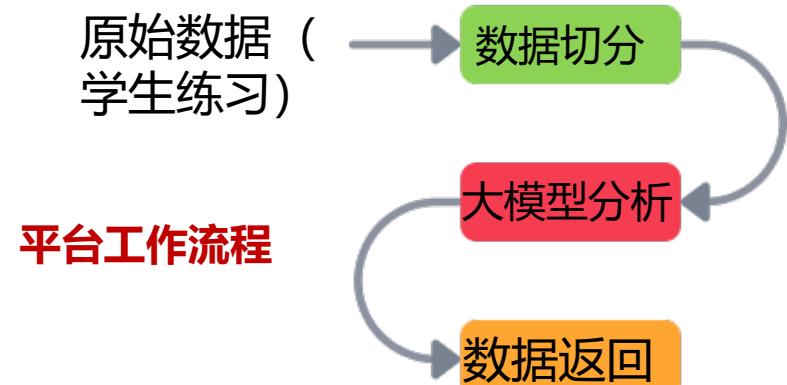
研究数据

本研究的数据主要由香港中文大学提供，包括参与者的日语练习题（N4难度）和相应的日语练习样卷。这些数据为平台开发和评估提供了重要的实证依据。

主要挑战

- ◆ ChatGPT在处理大量错误知识点时，分析深度不足，需要进一步细化和优化调试。
- ◆ 在为学生提供提升建议时，ChatGPT的回答针对性不够，需要增强个性化反馈。
- ◆ 新题目的难度和质量控制有待加强

数据处理：使用Langchain工具对学生的学习练习数据进行预处理，将其切分为知识点后输入模型。通过ChatGPT的API接口分析学生在练习中的错误，识别其错误的知识点并生成针对性的个性化练习题，帮助学生复习和巩固弱点知识。



研究方法：

- ◆ 个性化练习生成：基于模型的分析结果，为不同学生的错误和薄弱环节生成个性化练习题目。通过错题练习，帮助学生有效归纳错误、复习相关知识点，从而逐步提升语言学习水平。
- ◆ 学习效果反馈：平台定期分析学生的练习数据，生成可视化学习报告，帮助学生了解自身的学习进度。同时，利用ChatGPT的反馈功能，为学生提供针对性的讲解和改进建议。
- ◆ 实验结果：能够生成对应的知识点分析和新的练习题目，增强学生的学习体验。

参与人员：

刘宇（大四）、沈正一（大四）、尹泓睿（大三）

