

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

智慧城市感知与通信实验室

潘文安

实验室主任

2024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: 马献平, 王茂南, 谢悦皎, 吴倩倩, 马丁, 庞奥宇, 王子尧 硕士生: 谢炜枫

研究助理: 林可心, 胡尹, 罗艺欣, 汪思晋



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

子课题1：基于分割一切模型SAM辅助的地物识别研究

研究问题：遥感语义分割是许多地学下游任务的基础，它能够提供地面场景的精细的地物分类。近期出现了一种通用分割模型“任意分割模型”（SAM）再次改变了这一领域，它为准确、高效的分割提供了新的范式，将其以一种简洁的方式引入到遥感领域是十分迫切且重要的任务。

主要挑战：SAM 仅限于生成没有类别信息的分割结果。同时，当前方法预测的分割图通常存在过度碎片化和边界不准确等问题。

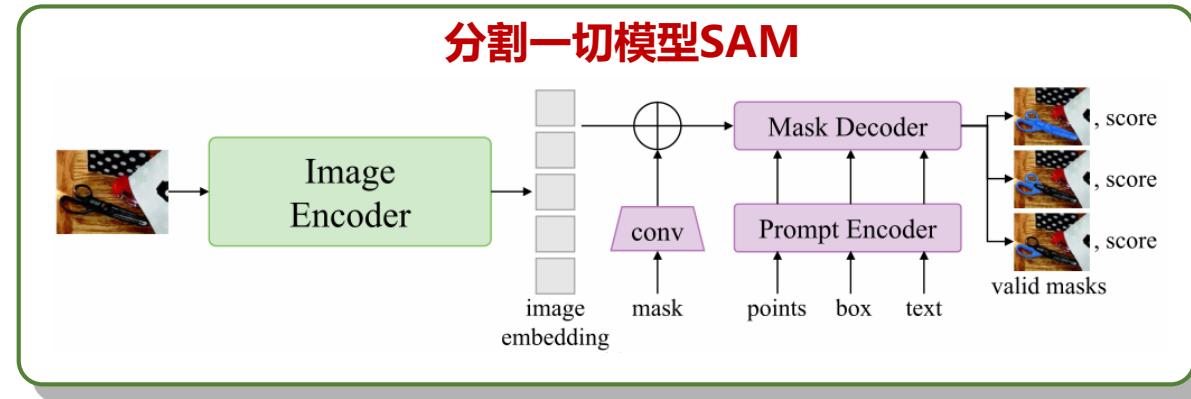
研究数据来源：ISPRS Vaihingen and LoveDA Urban

研究方法：本文提出了两个新概念，即 SAM 生成的对象 (SGO) 和 SAM 生成的边界 (SGB)，来充分利用 SAM 的原始输出。它们可以被广泛地、简洁地引入到通用的语义分割网络上，以辅助模型的优化。

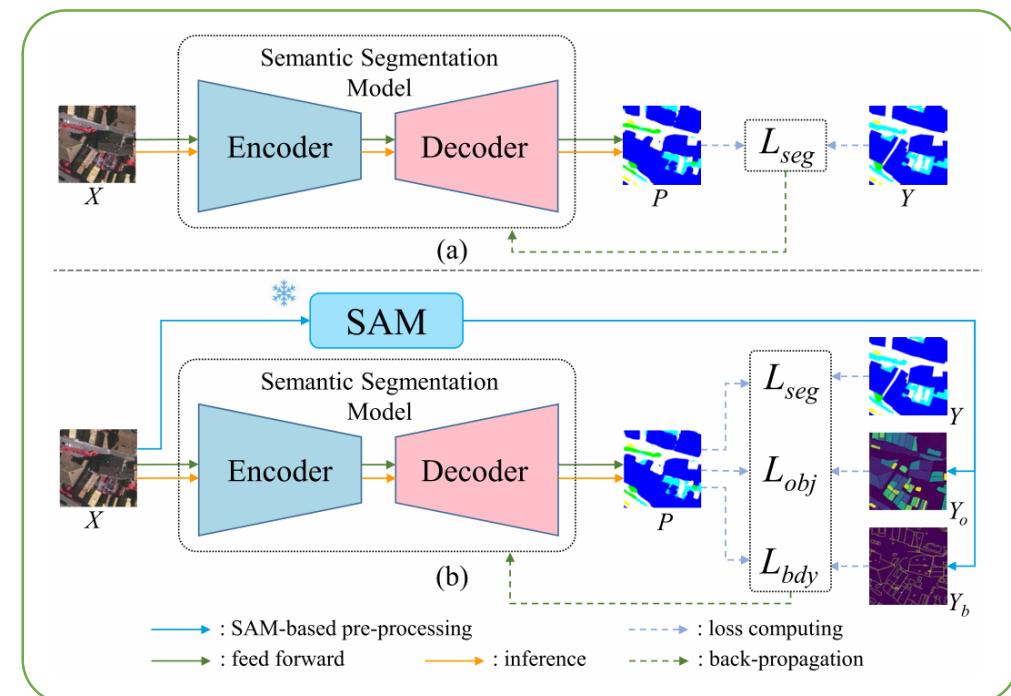
研究成果：

X. Ma, Q. Wu, X. Zhao, X. Zhang, M.O. Pun and B. Huang, "SAM-Assisted Remote Sensing Imagery Semantic Segmentation With Object and Boundary Constraints," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 62, pp. 1-16, 2024

参与人员：马献平（博四）



◆ 如何利用SAM模型隐含的先验知识 ◆ 如何使本方法更具扩展性

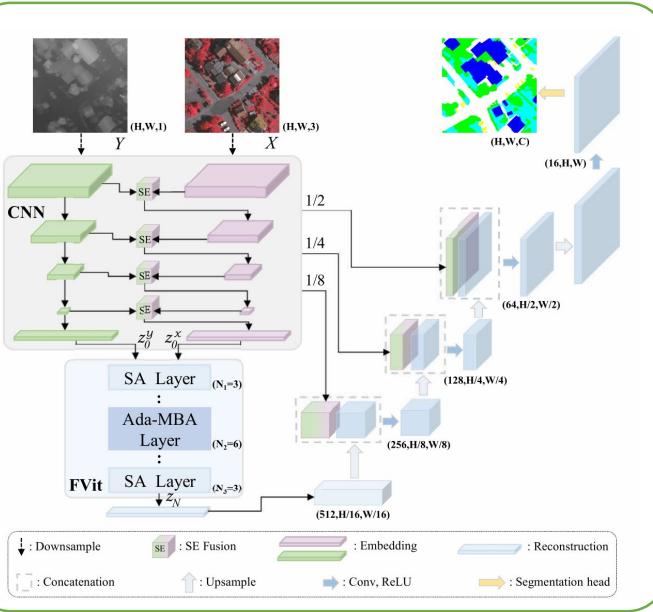


子课题2：基于多源遥感影像和多模态数据融合的地物识别研究

研究问题：在过去的几十年里，产生遥感数据的平台数量和产生其他地物信息的传感设备迅速增长，为该领域提供了丰富的多源多模态数据。基于不同传感器采集的影像是同一地物的事实，处理并融合多模态数据，使其相较于单一模态能提供更好的研究与实际应用价值已经成为一个重要的研究问题。

主要挑战：高分辨率遥感影像与自然图像相比，包含了丰富的地物信息，使用传统方法进行特征处理与融合容易损失较多的细节特征。

◆ 如何选择与处理多源遥感影像 ◆ 如何构造适用于遥感影像的特征提取网络 ◆ 如何设计多模态信息融合结构



研究方法：CNN作为一种自动化和层次化地提取特征的有效方法，已经成为计算机视觉相关任务的通用框架。然而卷积网络的感受野较小，无法高效处理高分辨率、远距离的遥感影像。本研究主要针对多模态融合这一方向，在卷积网络提取图像特征并进行初步融合的前提下，利用具有全局视野注意力机制的Transformer结构深度融合两种模态，以实现更好的融合效果。

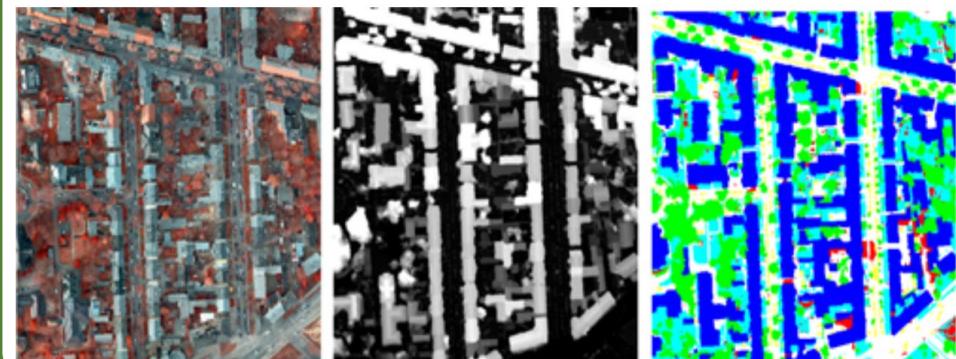
研究数据来源：ISPRS Potsdam and Vaihingen Dataset

研究成果：

X. Ma, X. Zhang, M. -O. Pun and M. Liu, "A Multilevel Multimodal Fusion Transformer for Remote Sensing Semantic Segmentation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 62, pp. 1-15, 2024

参与人员：马献平（博四）、胡尹（科研助理）

多源数据与地物识别



香港中文大學（深圳）

The Chinese University of Hong Kong, Shenzhen

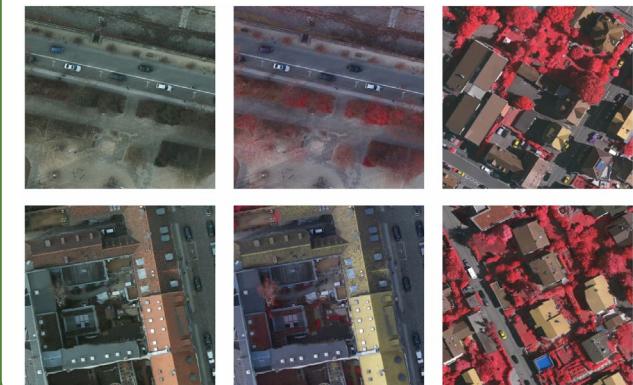
智慧城市感知与通信实验室

子课题3：基于分解的无监督域适应方法的跨域地物识别研究

研究问题：在过去的几十年里，产生遥感数据的平台数量和产生其他地物信息的传感设备迅速增长，为该领域提供了丰富的影像数据。基于不同域遥感影像之间存在的共性，在有标签的源域上训练的模型可以部分地适应于无标签的目标域场景。然后，由于跨域遥感数据的异构性，不同域之间存在极大的域差异。因此，解决域差异问题的无监督域适应方法已经成为一个重要的研究问题。

主要挑战：高分辨率遥感影像与自然图像相比，包含了丰富的语义信息，如何利用与解析丰富的遥感语义特征，是解决该问题的重点。

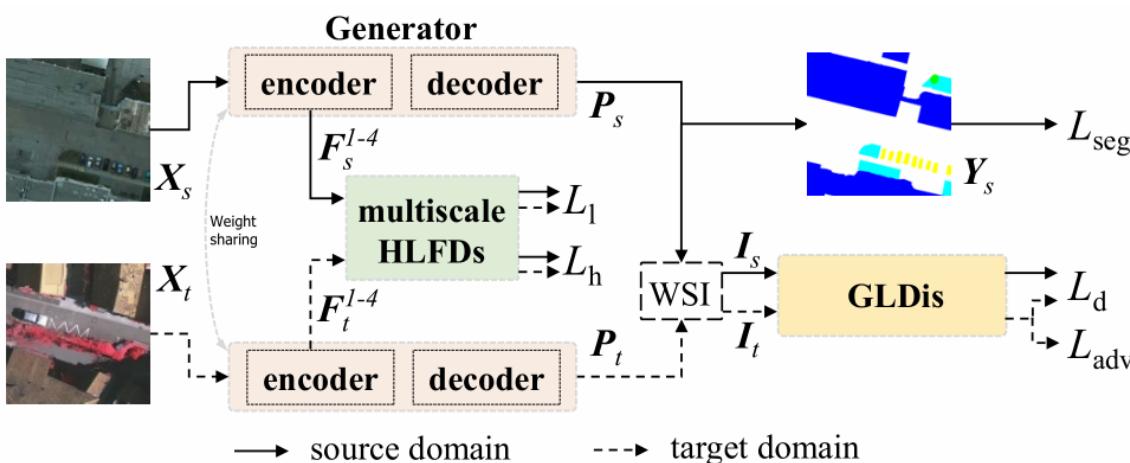
跨域遥感数据之间的域差异



◆ 如何构造适用于遥感影像的域适应网络

◆ 如何学习更具有迁移性的域通用特征

◆ 如何分解并对齐跨域遥感特征



研究成果：

Ma, X., Zhang, X., Ding, X., Pun, M. O., & Ma, S. (2024). "Frequency Decomposition-Driven Unsupervised Domain Adaptation for Remote Sensing Image Semantic Segmentation". submitted to *IEEE Transactions on Geoscience and Remote Sensing*, Sep. 2024.

参与人员：马献平（博四）

智慧城市感知与通信实验室



课题1 空天地一体化的智能感知系统：

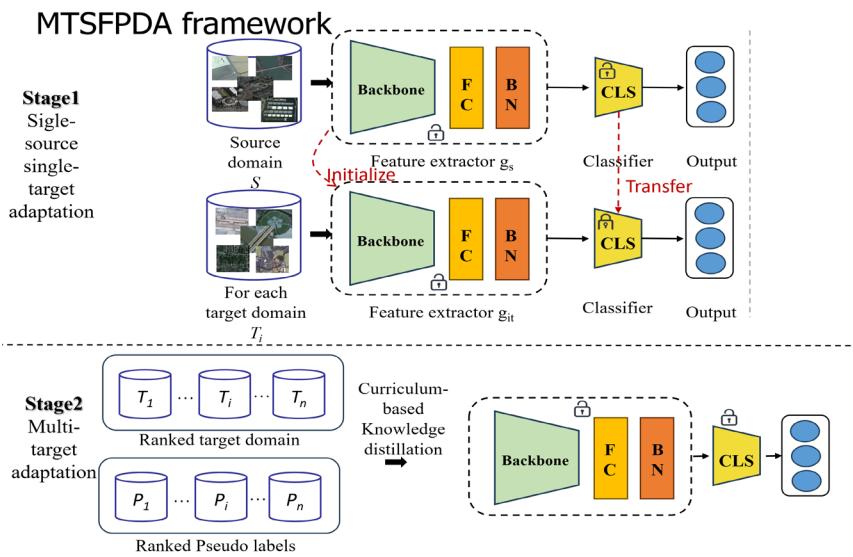
子课题4：基于无监督域适应的LCZ场景分类研究

研究问题：局部气候区 (Local climate zone, LCZ) 是具有统一地表覆盖、结构、材料和人类活动的区域。其分类为全球气候及城市热岛效应研究提供了基础框架。目前，针对基于遥感影像的LCZ场景分类方法，主要依赖于使用 Sentinel和Landsat数据的监督学习模型，但模型在不同城市之间的迁移性有限。本研究旨在减少目标城市标注的需求，提高模型在多个目标城市中的LCZ分类表现。

主要挑战：

- ◆ 传统的单一目标域适应方法无法同时适应多个城市的分布差异；
- ◆ 目前，适用于LCZ域适应的分类数据集相对较少，高分辨率影像数据集更为稀缺。

两阶段无监督域适应方法



数据集



研究方法：

1. 构建适用于 LCZ 分类的域适应数据集，获取来自不同城市的高分辨率遥感影像及标签；
2. 提出一种两阶段无监督域适应方法，针对多目标域适应问题，提升LCZ跨域分类效果。

数据来源： Jilin-1 image, Google Earth image, LCZ Generator

研究成果：

Q. Wu, X. Ma, J. Sui and M.O. Pun, "DF4LCZ: A SAM-Empowered Data Fusion Framework for Scene-Level Local Climate Zone Classification", *IEEE Trans. on Geoscience and Remote Sensing*, vol. 62, pp. 1-16, June 2024

参与人员： 吴倩倩 (博三)

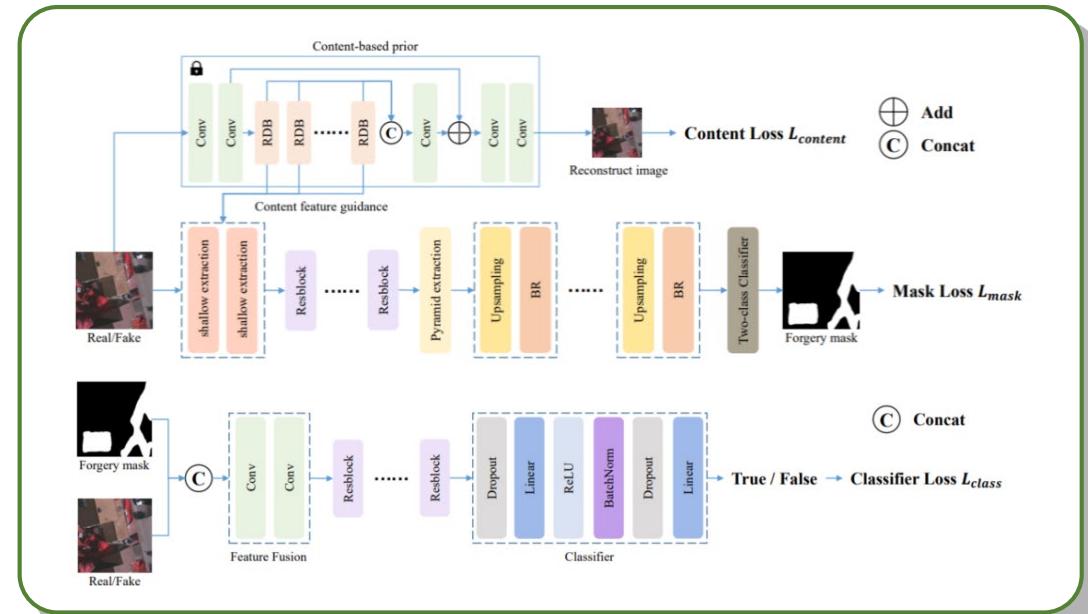
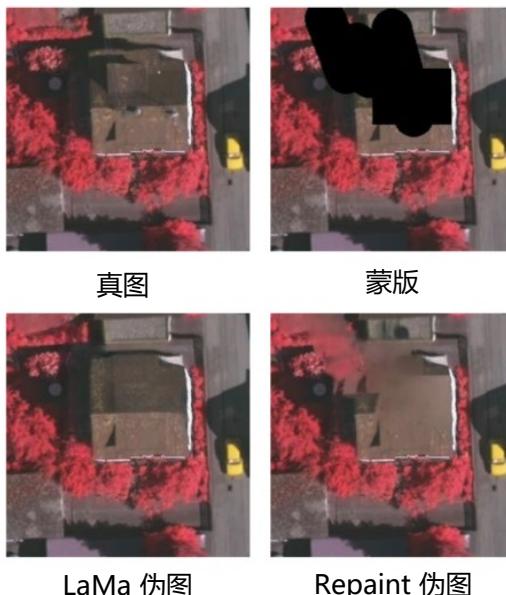


子课题5：基于多模态数据融合的遥感图像真伪识别研究

研究问题：由于缺乏合适的伪卫星图像数据集，对真伪卫星图像检测的研究目前处于起步阶段。如何实现针对深度伪造卫星图像的识别，以及如何对造假区域进行精确定位已成为一个重要的研究问题。

主要挑战：

- ◆ 如何构建应用范围更广泛的识别模型
- ◆ 如何通过特征融合提高模型准确率
- ◆ 如何融合多模态结构与现有模型



研究方法：本研究分为定位及检测两部分。具体而言，定位模块利用特征提取及边缘细化模块对伪造部分进行定位，识别模块融合定位特征及二分类算法输出最终的识别结果，二者结合以提升判别器的判断性能，提高伪造特征的提取能力。

研究数据来源：

ISPRS Vaihingen Dataset, LoveDA Dataset,
Fake Vaihingen & Fake LoveDA Dataset

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", submitted to *IEEE Trans. on Geoscience and Remote Sensing*, Sep. 2024.

参与人员：马丁（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： 子课题6：针对端到端任务的遥感图像去云

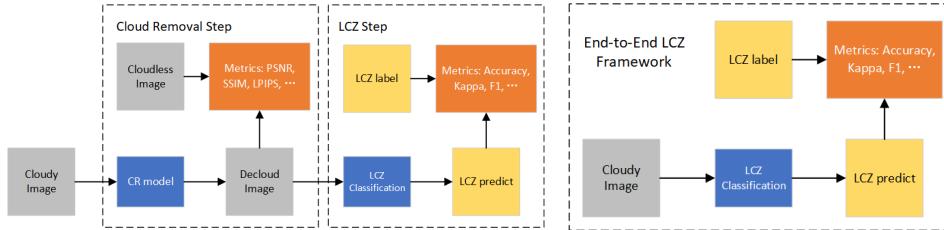
Funded by 联合实验室

研究问题：

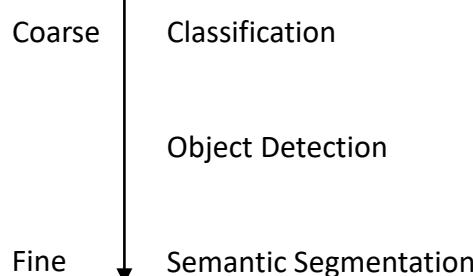
云层遮挡严重影响了遥感图像的质量并限制了对遥感数据的进一步分析和应用。本研究旨在通过利用深度学习技术，从有云图像中恢复地表信息。同时，从端到端任务角度，分析除去算法在下游任务中的可用性。

研究内容

(1) 端到端任务除云框架



(2) 下游任务性能评估



研究成果：

J. Sui, Y. Ma, W. Yang, X. Zhang, M.O. Pun and J. Liu, ``Diffusion Enhancement for Cloud Removal in Ultra-Resolution Remote Sensing Imagery'', *IEEE Trans. on Geoscience and Remote Sensing*, vol. 62, pp. 1-14, 2024.



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

- 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

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



可见光通道

近红外通道

研究方法：采用端到端学习框架，结合扩散模型技术模拟图像退化过程，利用反向扩散去除噪声并恢复地面信息。对于除云结果，使用传统图像除云指标和下游任务表现性能对模型进行综合评估。

参与人员：王子尧（博一）

课题1 空天地一体化的智能感知系统：

子课题7：基于深度学习的遥感视觉问答

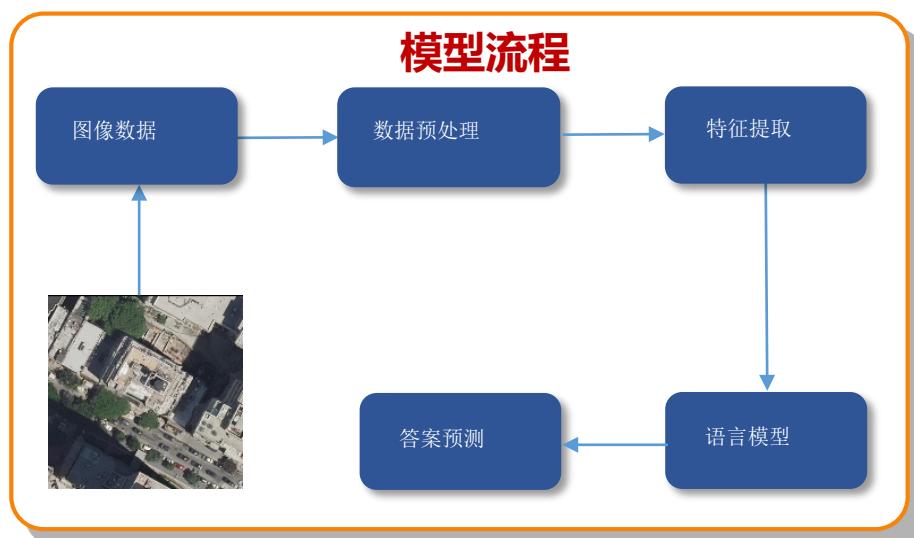
Funded by 联合实验室

研究问题

遥感视觉问答作为一种新兴的研究方向，旨在探索如何将遥感图像和文本信息进行有效的融合，对用户提出的自然语言问题作出准确的回答，提高地理信息的获取效率，并拓宽遥感技术的应用范围。

传统上，遥感图像的处理主要依赖于计算机视觉技术。本研究结合自然语言处理技术，生成与遥感图像内容相关的自然语言描述，引入大语言模型进行数据融合及语义关系分析，以优化遥感图像分析和理解方式，提升问答准确率。

模型流程



参与人员：林可心（科研助理）

研究数据



主要挑战

- ◆ 从图像中提取出文本特征
- ◆ 文本语义关系分析及特征融合
- ◆ 针对大规模的图像数据，要求模型有较强的泛化能力

研究方法

- 基于ResNet、Transformer等深度学习方法，对遥感图像进行特征提取
- 通过Prompt、LoRA等，对大语言模型进行微调，提升回答的质量
- 将模型应用于大规模数据，提升模型泛化能力



子课题8：通过语义生成扩散模型制作滑坡影像

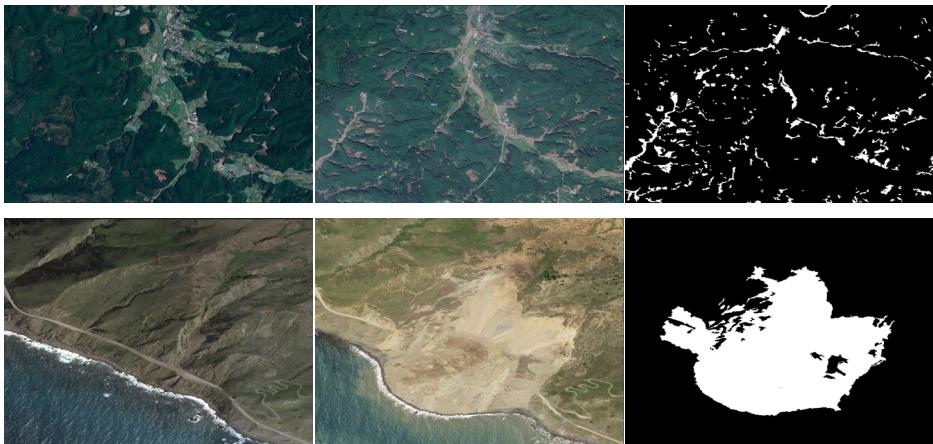
研究问题：对于变化检测任务，当前的开源数据集主要集中在建筑物提取和城市发展监测。在面对自然灾害监测时（以滑坡为例），往往缺乏充分的数据，这种情况阻碍了针对自然灾害的变化检测技术进一步发展。本研究旨在利用扩散模型生成逼真的滑坡影像，解决滑坡影像稀缺的挑战。

主要挑战

- ◆ 缺乏充足的发生滑坡的数据
- ◆ 需要生成与滑坡前影像和掩膜对应的滑坡后影像
- ◆ 生成的数据用于变化检测任务时需要实现精度提升

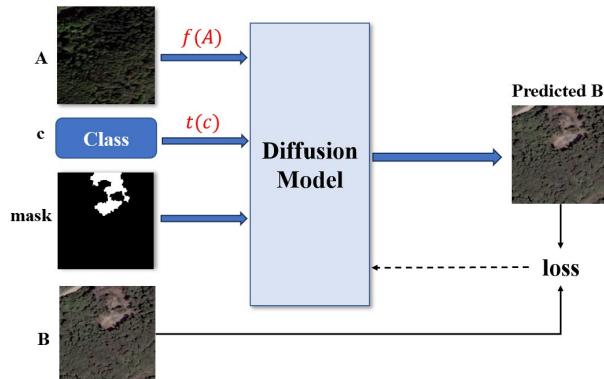
数据集：

GVLM：包括17组双时相超高分辨率影像对，空间分辨率为0.59m。影像通过 Google Earth获取。

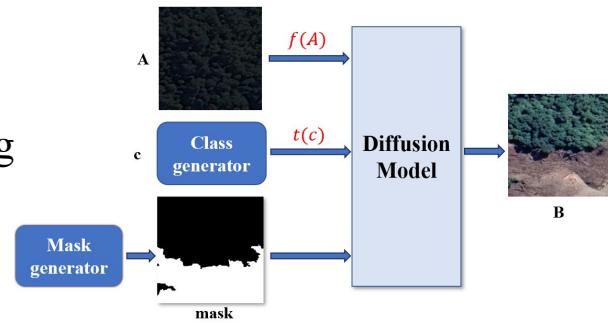


关键技术

(a) Training



(b) Sampling



研究方法：

1. 以扩散模型为基础，采用生成器得到若干掩膜，将其作为条件，和滑坡前图像一起输入条件扩散模型中，得到按照语义生成的滑坡影像；
2. 将生成的滑坡影像按照不同比例替换真实影像，构造数据集并用来训练变化检测网络，分析不同比例的生成图像对网络精确性的提升程度。

参与人员：罗艺欣（科研助理）



课题2 面向低空经济的6G智能网络：

Funded by 商汤科技

子课题1：面向不同交互频率的交通信号控制方法

研究问题：

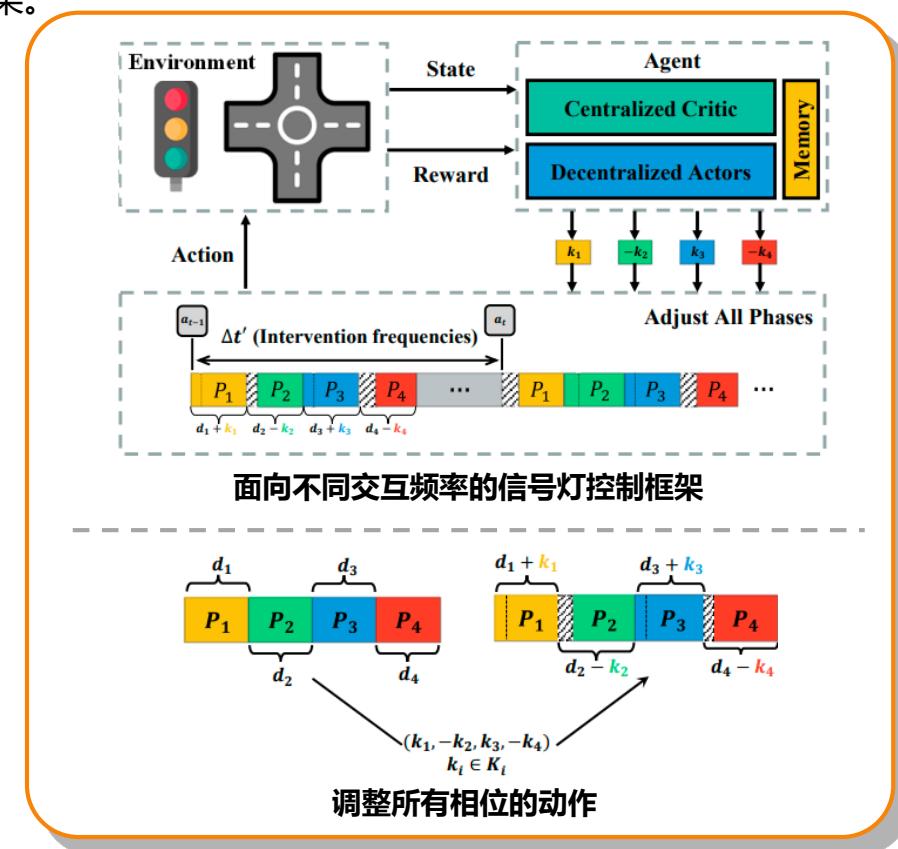
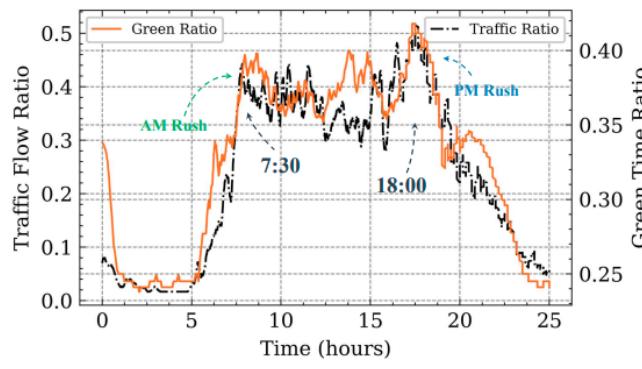
交通拥堵是全球性问题，导致环境恶化、经济损失和生活质量下降。优化交通信号控制（TSC）是缓解拥堵的有效途径。但是目前基于强化学习的方法需要高频的对信号灯进行控制才可以获得好的效果。

主要挑战：

- ◆ 传统方法依赖简单交通模型，无法适应复杂场景。
- ◆ 现有的 RL 方法在需要频繁对信号灯进行修改。
- ◆ 频繁修改信号灯会导致安全，稳定性等问题。

研究方法：

- ◆ **创新动作设计**：提出“调整所有相位”的新设计，允许同时优化所有信号相位。
- ◆ **CCDA框架**：采用中心化评论者与去中心化行动者的协同方法，有效减少动作空间并提高信号控制效率。
- ◆ **适应性控制**：方法适应不同的干预频率，减少队列长度并提高信号系统的稳定性。



研究成果：

Wang, M., Chen, Y., Kan, Y., Xu, C., Lepech, M., Pun, M. O., & Xiong, X. (2024). Traffic Signal Cycle Control with Centralized Critic and Decentralized Actors under Varying Intervention Frequencies. Accepted for publication in *IEEE Transactions on Intelligent Transportation Systems*, Sep. 2024.

参与人员：王茂南（博三）



子课题 2：复杂交通环境下基于大语言模型的交通信号控制

研究问题：

城市交通拥堵对经济、环境和社会产生重大影响，传统交通信号控制系统在应对复杂的交通状况时表现有限，亟需更加智能的解决方案。

主要挑战：

- ◆ **规则固化**：传统基于规则的信号控制系统无法灵活应对复杂和动态的交通情况。
- ◆ **强化学习局限**：强化学习系统可能会过拟合特定的交通模式，难以适应突发情况。
- ◆ **事件处理不足**：现有方法无法有效应对突发的紧急事件，如突发事故或紧急车辆的到来

研究方法：

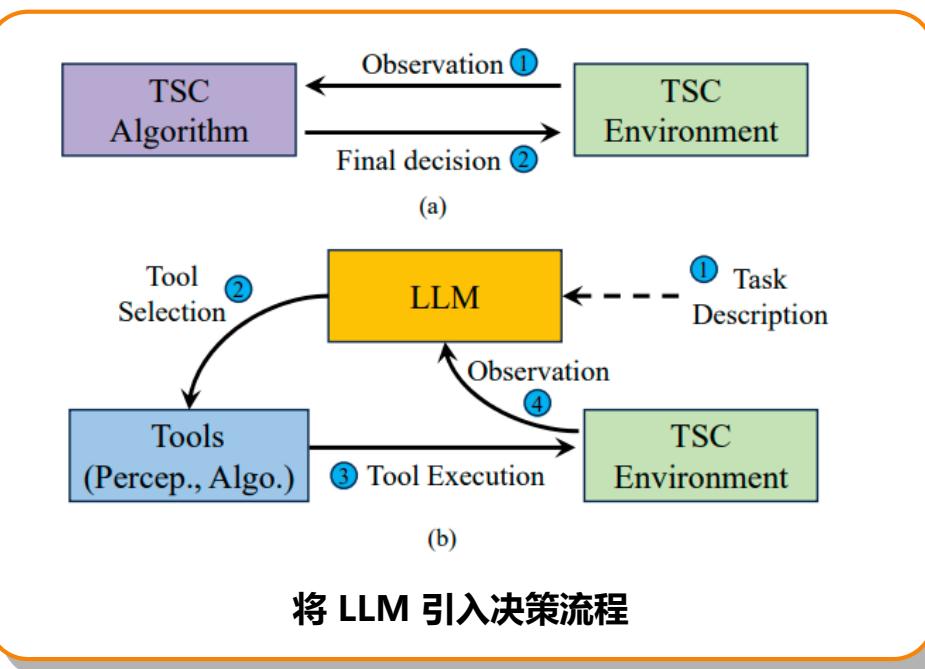
- ◆ **LLM集成**：通过引入大型语言模型（LLM）来增强交通信号控制系统的决策能力，应对复杂和罕见的交通场景。
- ◆ **工具集成**：开发了一套用于收集静态和动态交通信息的工具，作为LLM与系统的桥梁。
- ◆ **仿真验证**：通过仿真平台进行实验，验证LLM辅助系统在多种交通环境中的高效表现。

研究成果：

Wang, M., Pang, A., Kan, Y., Pun, M. O., Chen, C. S., & Huang, B. (2024). LLM-assisted light: Leveraging large language model capabilities for human-mimetic traffic signal control in complex urban environments. Submitted to the *IEEE Transactions on Intelligent Transportation Systems*, Sep. 2024.

项目地址：[Traffic-Alpha \(github.com\)](https://Traffic-Alpha.github.com)

参与人员：王茂南（博三），庞奥宇（博二）



子课题3：整合RL和LLM的交通信号控制算法

研究问题

现有的智能交通信号控制系统通常假设交通信息能够实时准确地获取，但在实际部署中，由于数据传输延迟、丢失以及通信退化等因素，观测信息与实际交通状况存在差异，导致信号控制系统决策的有效性下降。本研究提出了一种基于强化学习（RL）与大语言模型（LLM）相结合的智能交通信号控制框架，明确地考虑通信退化和长尾场景，以提升交通信号控制在实际应用中的泛化性和鲁棒性。

主要挑战

- ◆ 通信退化问题：实际应用中，由于通信链路的延迟、丢包和噪声等问题，系统获取的状态信息可能不完整或失真，这会直接影响RL模型的决策性能
- ◆ 长尾场景挑战：当面对训练中未见过的复杂或罕见场景（即长尾场景）时，模型可能无法有效泛化，从而导致决策失效甚至系统崩溃

研究成果

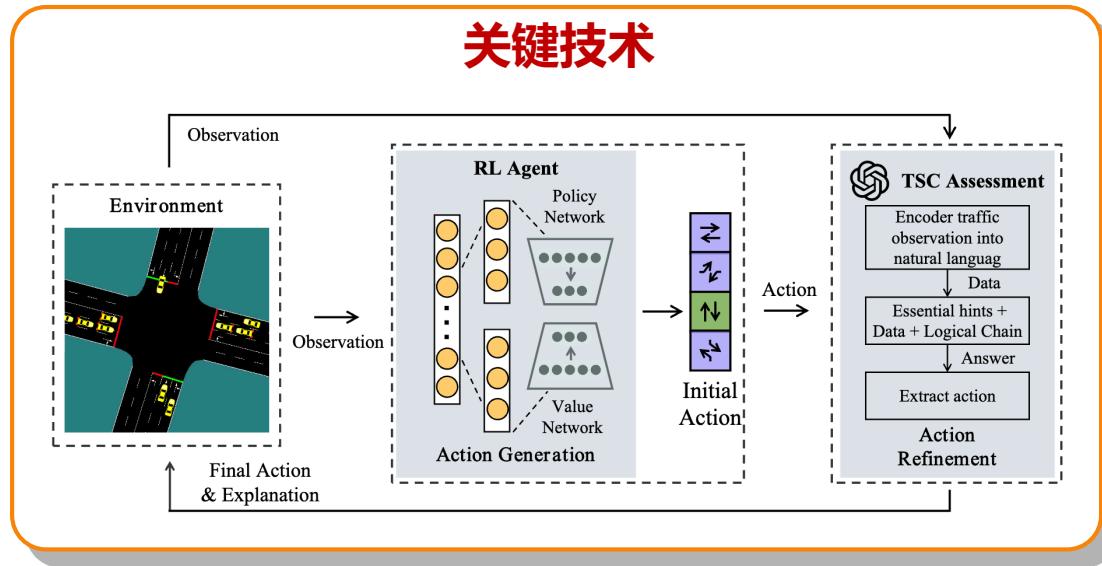
- ◆ A. Pang, M. Wang, Y. Chen, M.O. Pun and M. Lepechand, "Scalable Reinforcement Learning Framework for Traffic Signal Control under Communication Delays", *IEEE Open Journal of Vehicular Technology*, vol. 5, pp. 330 - 343, March 2024.
- ◆ Pang A, Wang M, Pun M O, et al. "iLLM-TSC: Integration reinforcement learning and large language model for traffic signal control policy improvement", submitted to *IEEE Transactions on Vehicular Technology*, Sep. 2024.



香港中文大學（深圳）

The Chinese University of Hong Kong, Shenzhen

关键技术



研究方法

- ◆ RL基于环境特定的交通数据进行学习，在实时应用中做出初步交通信号控制决策。
- ◆ LLM对RL的初步决策进行二次评估和优化，提升应对复杂场景的能力。
- ◆ iLLM-TSC框架可无缝集成现有RL系统，增强RL在不确定环境中的决策精度和鲁棒性

参与人员：庞奥宇（博二）、王茂南（博三）

课题2 面向低空经济的6G智能网络：

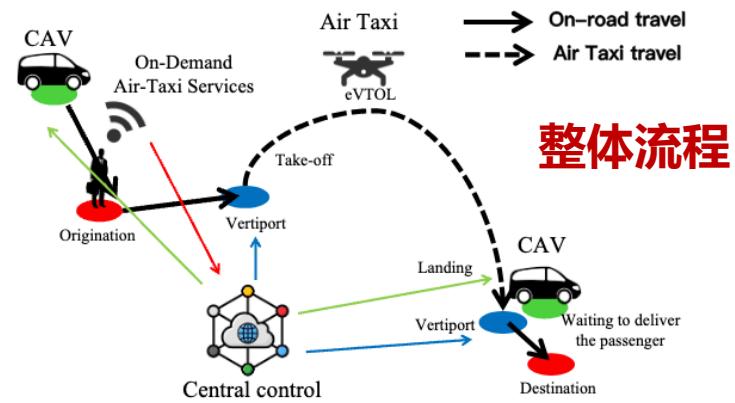
子课题4：基于RL的联网自动驾驶车辆与按需空中出租车服务策略优化

研究问题

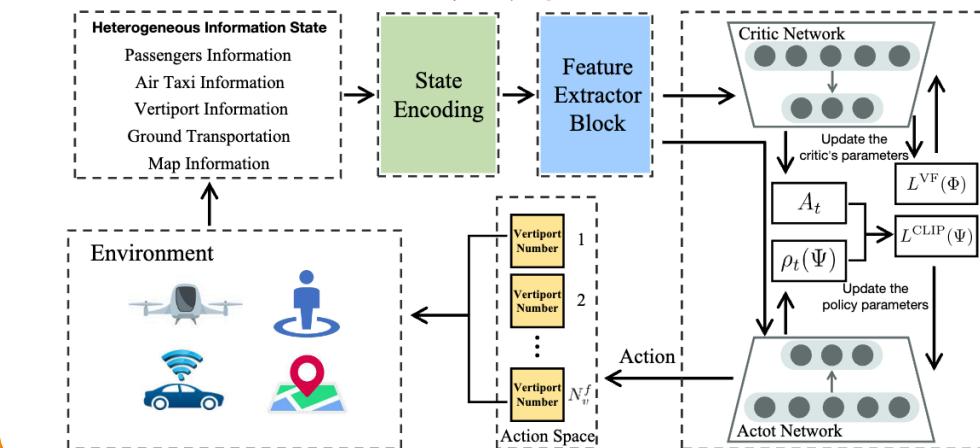
城市低空出行缓解城市拥堵提供了新的解决方案，但在实际应用中，如何将空中出租车服务与地面交通系统无缝集成，实现“门到门”的出行仍然是一个挑战。为了解决上述问题，本研究提出了一个基于RL的动态的空中出租车路线推荐系统，该框架通过实时的空中和地面交通状况，优化联网自动驾驶车辆(CAV)与eVTOL之间的协作，动态推荐空中出租车的最佳路线，显著减少了整体出行时间。

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



关键技术



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上空交通状况；
- ◆ 基于Actor-Critic的强化学习框架，实现全局的乘客的动态线路推荐；
- ◆ 设计了一个深度神经网络，以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 (博二)

研究成果：

Wang, M., Chen, Y., Kan, Y., Xu, C., Lepech, M., Pun, M. O., & Xiong, X. (2024). "Traffic Signal Cycle Control with Centralized Critic and Decentralized Actors under Varying Intervention Frequencies", submitted to the IEEE Transactions on Intelligent Transportation Systems, Sep. 2024.



课题3 智慧家居物联网

Funded by 联合实验室

子课题1：基于深度学习的鼾声识别

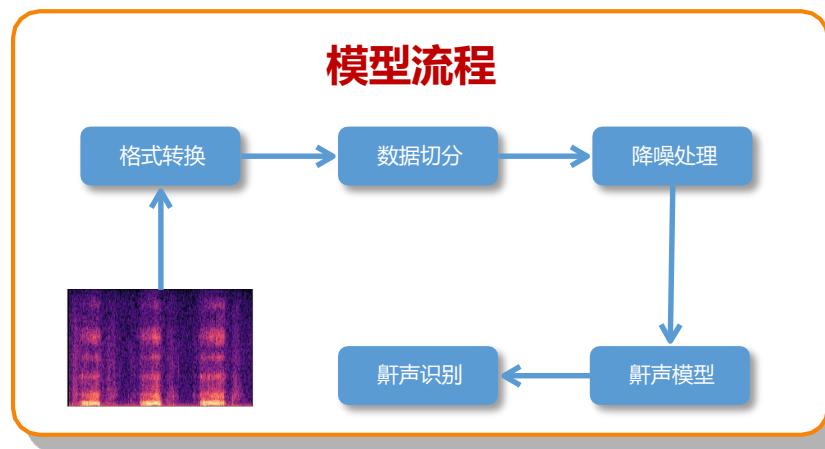
研究问题：打鼾是睡眠中的一种非意识行为，它既能反映打鼾者本身的一些生命体征，也能监测人的生理健康状况。本研究以单片机作为硬件平台，使用深度学习与声源定位的方法，实现实时判断是否有鼾声并判断鼾声的方位。

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



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



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。

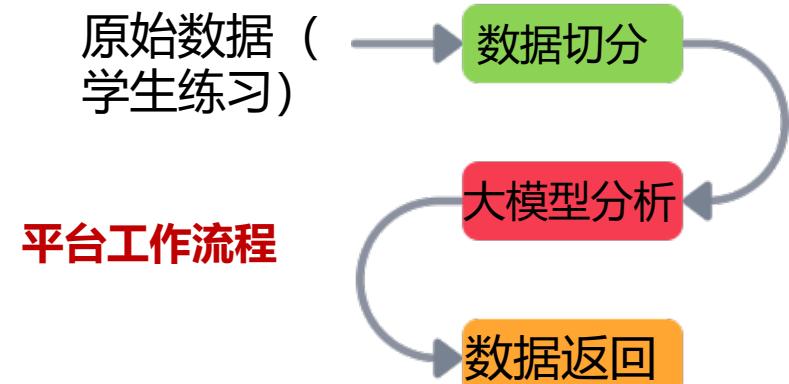
研究数据

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：

康子墨（大四）、刘宇（大三）



子课题3：基于街景的建筑物窗户数据集和楼层数量识别

研究问题

- ◆ 碳排放会导致全球变暖、极端恶劣天气等重要影响，许多研究表明人口因素是影响碳排放的重要因素之一。但由于人口密度等指标难以直接计算，因此通过窗户数量和楼层数量来间接计算成为了一种可取的方案。
- ◆ 但是这项工作并没有开源的数据集，并且该问题常用的OSM在深圳市具有极大的误差，因此该项目的数据集需要从零开始收集，并据此制定有针对性的收集方案。

数据来源



高德地图API
获取建筑物信息



百度地图
获取街景照片

主要挑战

- ◆ 高德地图的经纬度及建筑物信息与百度街景存在一定差异
- ◆ 大部分高楼层的窗户因为图形畸变导致不能使用
- ◆ 为数据集打标签需要极大的工作量

模型流程



研究成果

- ◆ Y. Sun, X. Ma, Y. Yan, M.O. Pun and B. Huang, "A Novel Automated Urban Building Analysis Framework Based on GPT and SAM", Proc. the 2024 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Athens, Greece, July 7-12, 2024.

研究方法

- ◆ 通过高德地图的API与OSM的API提供的建筑物信息与OSM的API，对获取的百度地图街景照片进行交叉矫正。
- ◆ 利用LabelImg工具为数据集的窗户打标签。
- ◆ 使用YOLOv7进行训练识别，并间接计算楼层的高度信息。

