



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

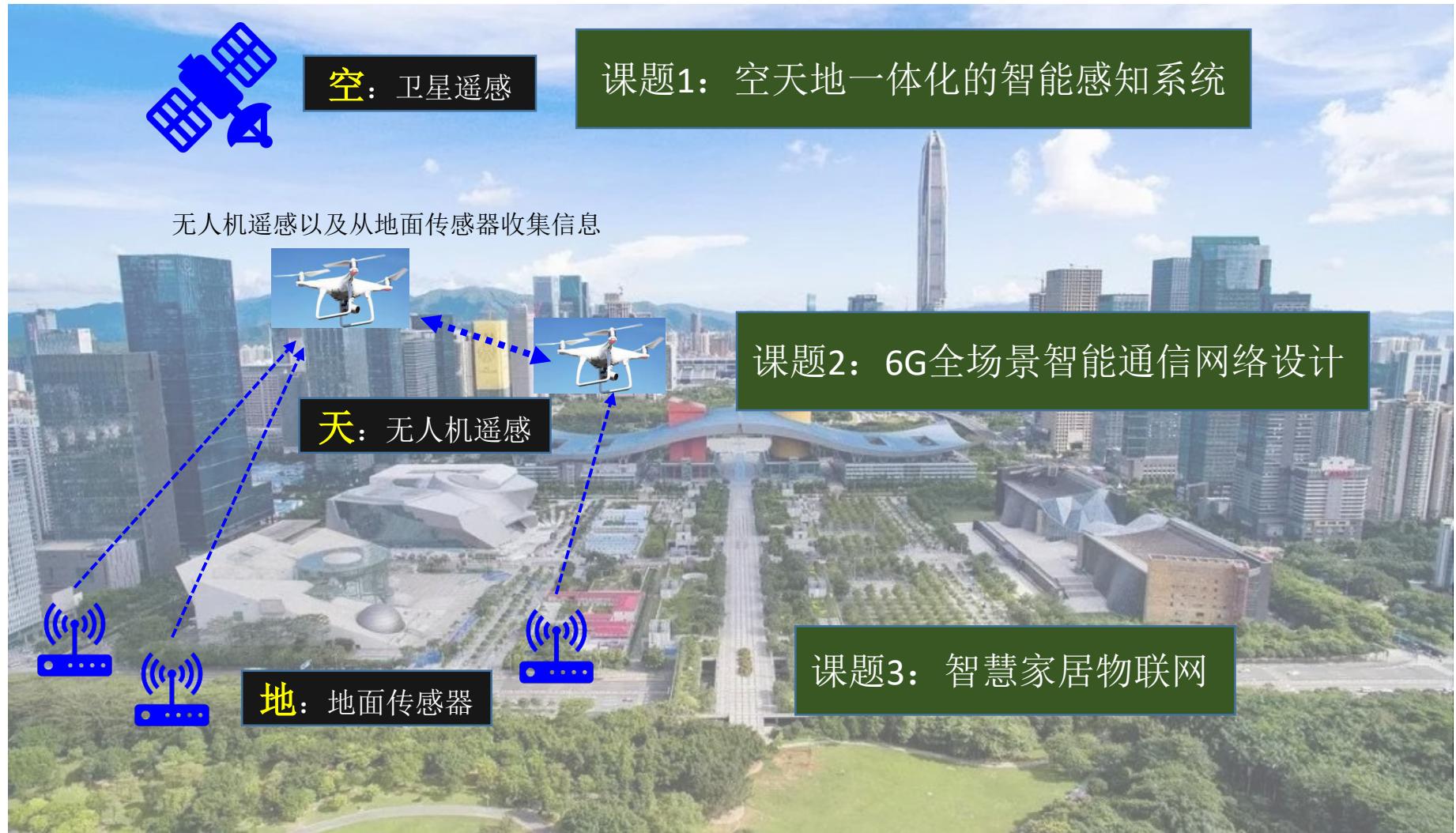
实验室主任

2023年9月12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



博士后：张永显博士(将于2023年10月加入)

博士生：马献平，王茂南，谢悦皎，吴倩倩，马丁，庞奥宇

研究助理：林可心

硕士生：崔向宇，隋佳璐，梁文迪，孙宇超，谢炜枫



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

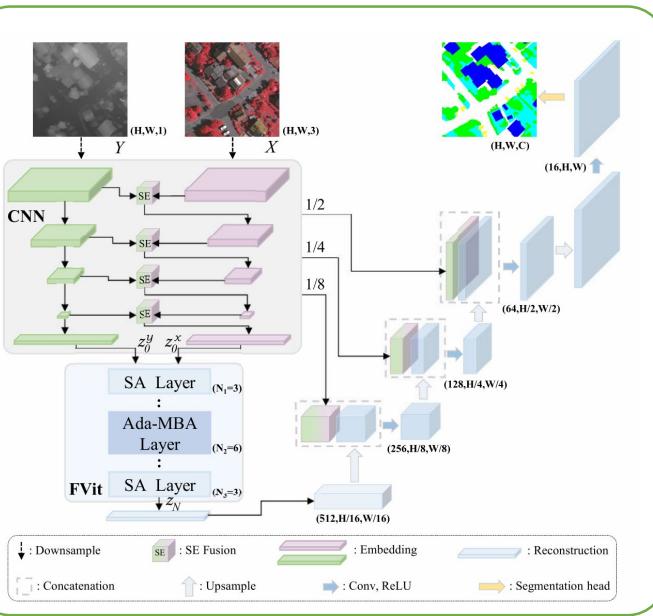
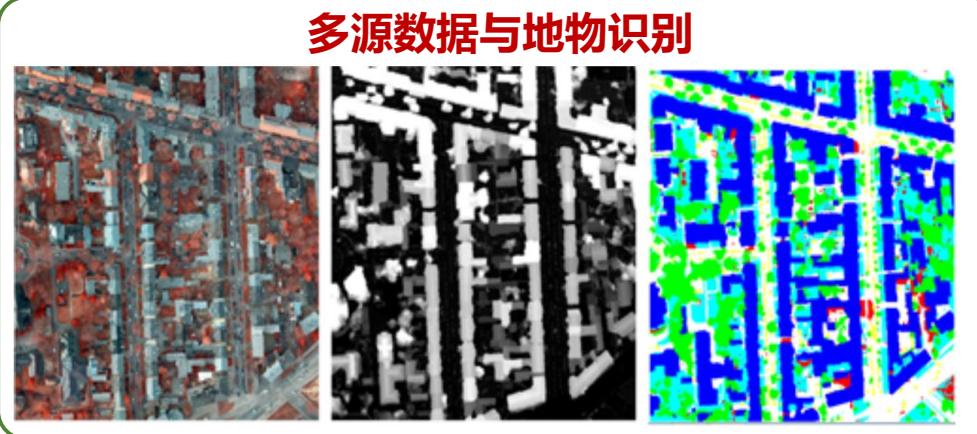
智慧城市感知与通信实验室

## 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 and M.O. Pun, “A Crossmodal Multiscale Fusion Network for Semantic Segmentation of Remote Sensing Data”, the IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, Vol. 15, pp. 3463-3474, May 2022.

**参与人员：**马献平（博士生）



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2：基于无监督域适应方法的跨域遥感影像识别研究

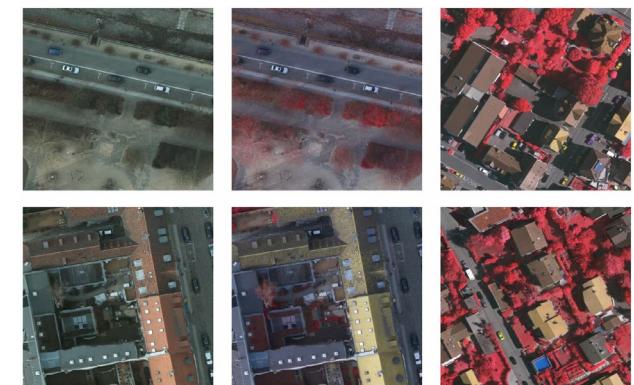
**研究问题：**在过去的几十年里，产生遥感数据的平台数量和产生其他地物信息的传感设备迅速增长，为该领域提供了丰富的影像数据。基于不同域遥感影像之间存在的共性，在有标签的源域上训练的模型可以部分地适应于无标签的目标域场景。然后，由于跨域遥感数据的异构性，不同域之间存在极大的域差异。因此，解决域差异问题的无监督域适应方法已经成为一个重要的研究问题。

**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，如何利用与解析高维的语义特征，是解决该问题的重点。

◆ 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

◆ 如何学习更具有判别性的域特殊特征



**研究方法：**本研究利用抗生成网络，研究针对高维特征的协同增强。具体而言，在卷积网络提取图像特征的前提下，利用协同增强注意力机制计算源域和目标域之间的共性，以对齐高维域通用特征。同时，设计全局的损失函数，从整体数据分布的角度让源域和目标域的特征图趋于一致。此外利用自注意力机制，提高判别器的能力，提高域特殊特征的提取能力。

**研究成果：**

X. Ma, X. Zhang, Z. Wang and M.O. Pun, "Unsupervised domain adaptation augmented by mutually boosted attention for semantic segmentation of VHR Remote Sensing Images", the IEEE Trans. on Geoscience and Remote Sensing, vol. 61, pp. 1-15, Feb. 2023.

**参与人员：**马献平 (博三)

### 研究数据来源：

ISPRS Potsdam and Vaihingen Dataset



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3：遥感影像超分辨率

Funded by 联合实验室

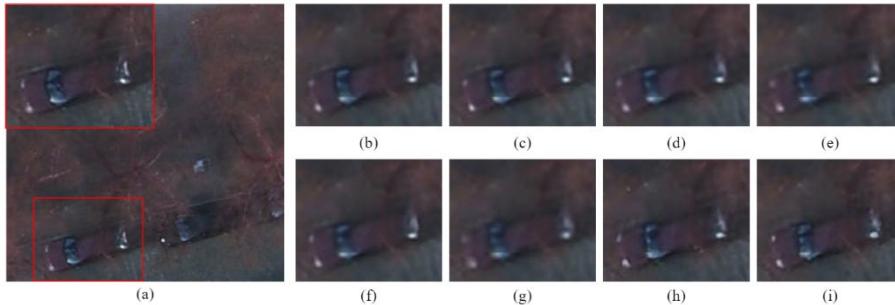
**研究问题：**单图像超分辨率观测是近几十年来图像处理领域的一个活跃的研究课题。本研究旨在从低分辨率图像重建高分辨率图像，利用深度学习提高合成数据的分辨率。

## 主要挑战

- ◆ 算法效果在真实场景中的LR图像会出现退化
- ◆ 需要能够适应不断变化的真实世界图像的模型
- ◆ 需要特定于任务的评估指标

## 数据源：

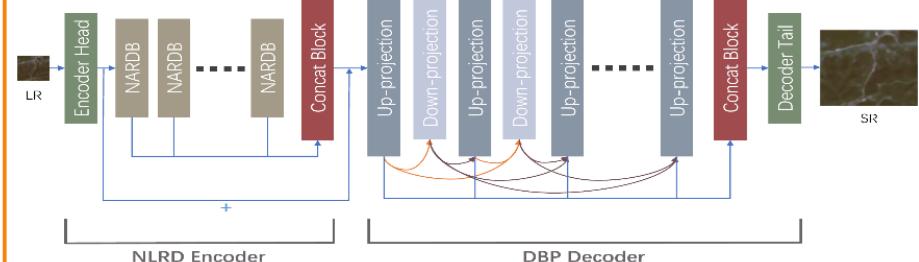
- 遥感卫星影像数据OLI2MSI, Alsat



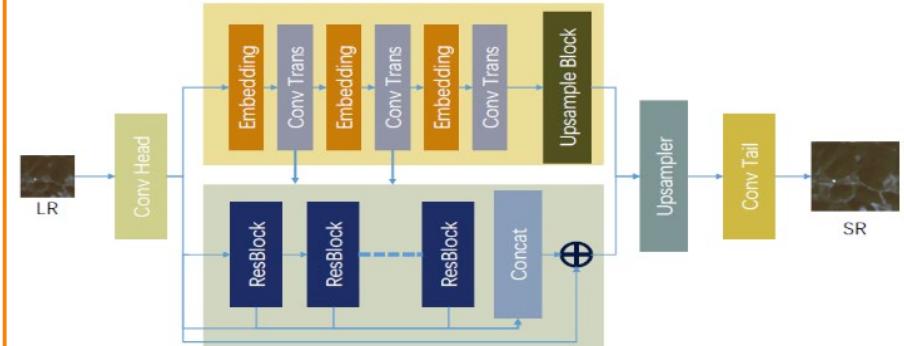
## 研究成果：

J. Sui, X. Ma, X. Zhang and M.O. Pun, "GCRDN: Global Context-driven Residual Dense Network for remote sensing image super-resolution", IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 16, pp. 4457-4468, 2023.

## 关键技术



(1) 基于遥感图像超分的全局上下文驱动残差密集网络



(2) 基于遥感超分的双分支残差网络

**研究方法：**使用传统的编码器解码器架构，向编码器的密集残差网络中增添注意力机制模块，从而将注意力机制的全局信息与卷积层的局部信息相融合，得到更好的超分辨率结果。

**参与人员：**隋佳璐（硕二）



## 子课题4：基于分割大模型与多源信息的LCZ场景分类研究

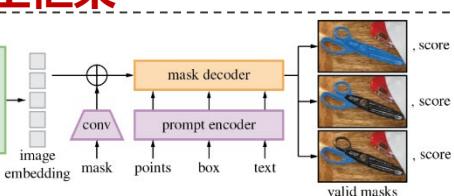
**研究问题：**局部气候区(Local climate zone, LCZ)是具有统一地表覆盖、结构、材料和人类活动的区域。其分类为全球气候及城市热岛效应研究提供了基础框架。目前，针对基于遥感的LCZ分类方法，其研究重点正从像素级分类方法转向对象级和场景级分类方法；同时，深度学习在LCZ分类中呈现快速增长趋势。本研究旨在结合图像分割大模型与多源信息提升LCZ场景级分类效果。

**主要挑战：**1.现有场景级LCZ分类方法未考虑与图像分割大模型结合；2.现有分类数据集较少，且缺乏空间坐标信息，难以获得空间范围对齐的其他多源数据。

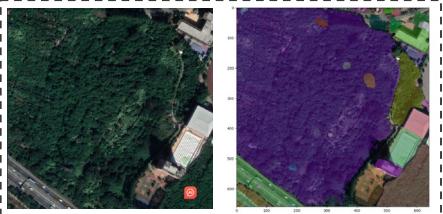
### 分割大模型框架



image



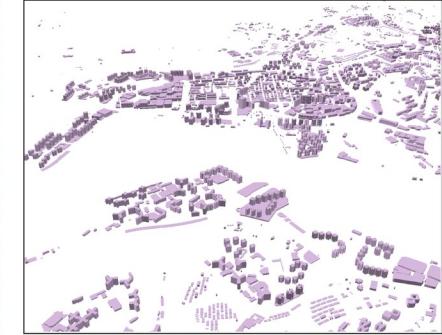
### 分割大模型下的遥感影像分割效果



### 数据集



Sentinel-2影像



OSM buildings

### 研究方法：

- 1.进行LCZ分类数据集构建，获取时间属性相近、空间坐标对齐的多源数据；
- 2.提出个性化的分类算法，将分割大模型和多源信息融合相结合，提升LCZ分类效果。

### 数据来源：

- ESA sentinel-2 image
- Google satellite image
- OSM buildings

**参与人员：**吴倩倩 (博二)



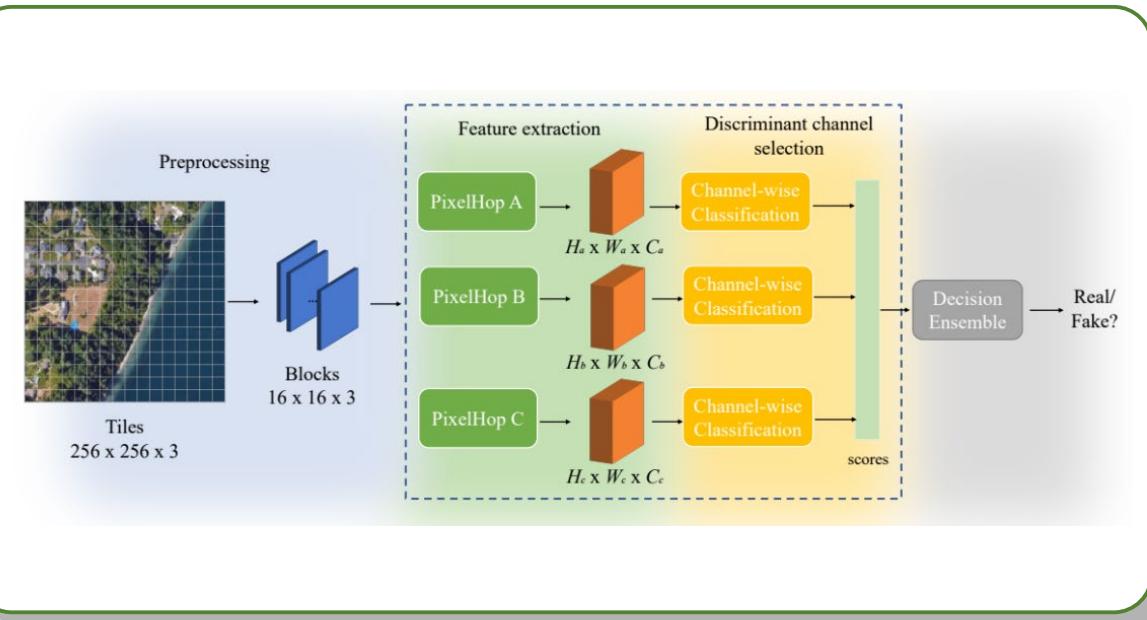
# 课题1 空天地一体化的智能感知系统：

## 子课题5：基于Geo-DefakeHop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。南加州大学的陈宏硕团队曾于2021年提出Geo-DefakeHop模型，该模型对于基于CycleGAN生成对抗网络生成的伪遥感图有较好的识别能力，但针对其他途径生成的伪图仍有较大的改良空间。

### 主要挑战：

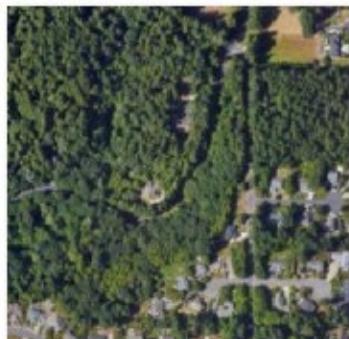
- ◆ 如何构建应用范围更广泛的识别模型
- ◆ 如何识别嵌套于真图中的伪造部分
- ◆ 如何提升高分辨率图像的检测正确率



### 利用CycleGAN生成的真伪图对比



伪图



真图

**研究方法：**使用CycleGAN以外的多个生成对抗网络生成伪图并分别训练现有识别模型，研究输出结果并对模型参数进行调整。同时优化模型的通道选择策略及结果集成策略，以实现针对高分辨率图像的识别以及对于复杂合成图像中伪造部分的识别。

**研究数据来源：**美国华盛顿大学西雅图UW Fake Satellite Image dataset

**参与人员：**马丁（博一）



# 课题1 空天地一体化的智能感知系统：

## 子课题6：基于深度学习的遥感视觉问答

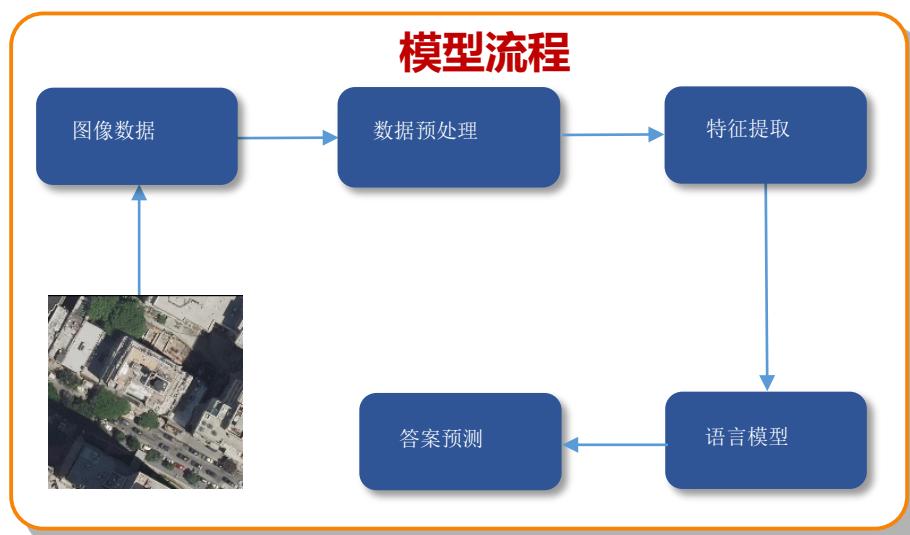
Funded by 联合实验室

### 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（访问学生）

### 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



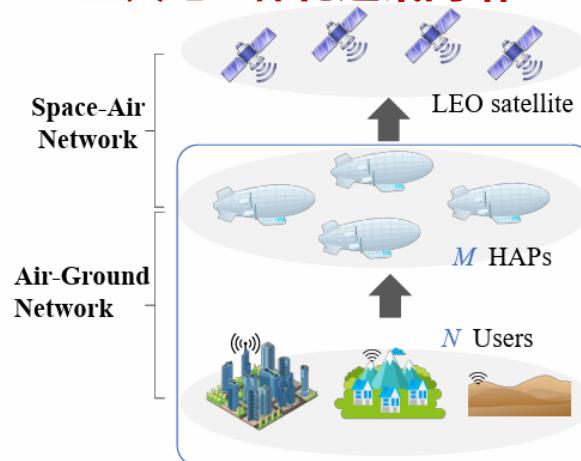
## 子课题1：基于强化学习的空天地一体化接入网中的资源优化

### 研究问题

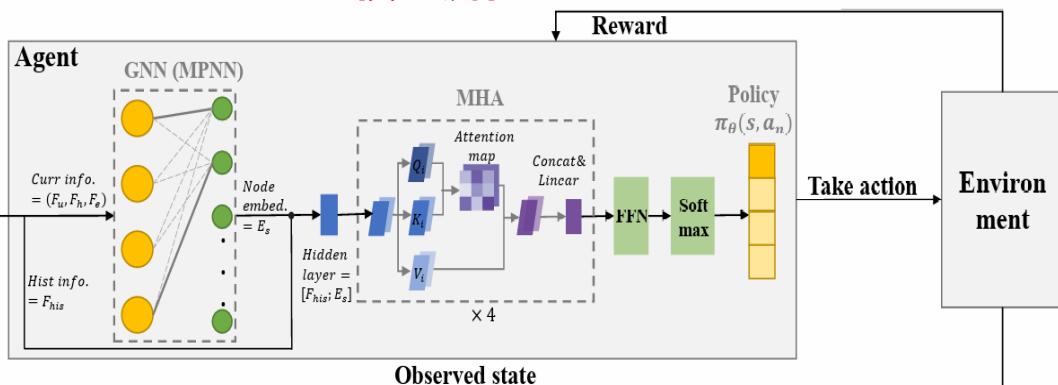
构建空天地一体化网络，可以提高网络的覆盖率，主要针对在没有被地面基站覆盖的偏远地区，海域，或者人员密集的区域，实现6G网络全覆盖的目标。网络的信号传递过程为，地面用户向HAP(High Altitude Platform)发送信号，HAP收到后需要将数据传回地面，此时将卫星作为接收站，HAP将数据发送LEO(Low Earth Orbit) satellites，再由LEO将数据传回其地面接收站。

我们主要研究的是HAP和用户，HAP和卫星之间的匹配问题。考虑到用户连接的连接收益和成本，我们将用户的接入或HAP的能量资源分配建模为在线二部匹配问题(Online bipartite matching)。

### 空天地一体化通讯网络



### 模型流程



### 主要挑战

- ◆ 大规模用户接入网匹配优化问题。
- ◆ 解决在线问题中的随机性和不确定性

### 研究方法

对用户-HAP匹配建模为在线二部匹配问题，利用GNN网络+强化学习框架进行优化。

参与人员：谢悦皎（博二）

### 研究成果：

Y. Xie, G. Niu, M.O. Pun and Z. Han, "Online Bipartite Matching for HAP Access in Space-Air-Ground Integrated Networks using Graph Neural Network-Enhanced Reinforcement Learning", Proc. the 2023 IEEE International Conference in Communications (ICC), Rome, Italy, May 28 – June 1, 2023.



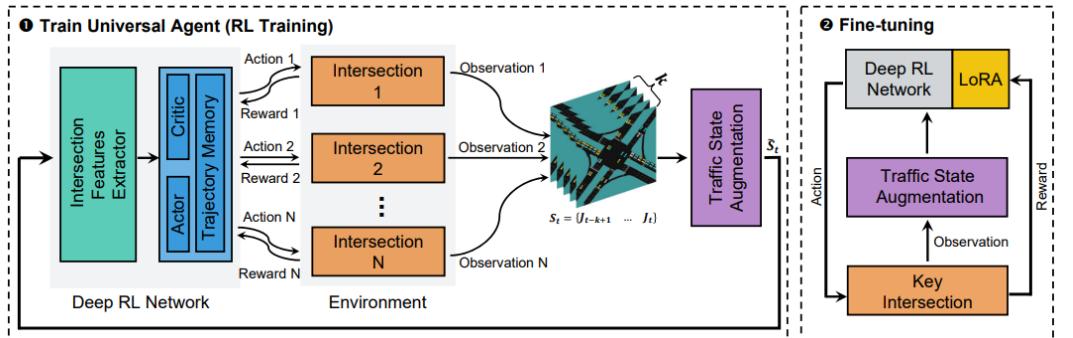
## 子课题2：基于强化学习的通用交通信号控制研究

### 研究问题：

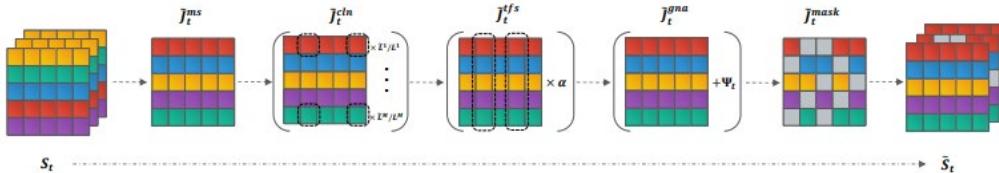
旨在解决城市交通拥堵问题，重点关注交通信号控制（TSC）系统的优化。具体而言，研究问题是如何开发一个通用的强化学习（RL）框架，用于处理不同结构的交叉路口，以改善交通信号控制的效能。

### 主要挑战：

- ◆ 1. 传统方法处理动态路况表现不佳，性能有限。
- ◆ 2. RL 方法在新环境需要重新训练，限制了扩展性。
- ◆ 3. 现有的通用模型遇到未知结构路口时性能下降。



UniTSA 整体框架，①通用模型；②关键路口微调



UniTSA 中数据流（路口矩阵与五种数据增强）

### 研究方法：

- ✓ 引入了交叉路口矩阵来描述交叉路口状态，使不同结构的交叉路口可以用相同大小的矩阵进行描述。
- ✓ 使用“Keep or Not”的动作，确保不同配置的交叉路口具有一致的模型结构。
- ✓ 开发了五种交通状态增强方法，丰富了代理的数据收集，提高了在训练和测试集中的性能。
- ✓ 使用Low-Rank Adaptation (LoRA) 进一步微调模型以提高关键交叉路口的性能。

### 研究成果：

M. Wang, Y. Xu, X. Xiong, Y. Kan, C. Xu and M.O. Pun, "A Universal Approach of Traffic Signal Control with Augmented Data Using Reinforcement Learning", Proc. the Transportation Research Board (TRB) Annual Meeting 2023, Washington, D.C., U.S.A., Jan. 8–12, 2023.

参与人员：王茂南（博二）



## 子课题3：V2X 延迟观测下基于强化学习的两阶段控制系统

### 研究问题

大多数现有的智能交通信号控制系统设计，都假设交通流信息是可实时获取的，但在实际部署中，因为传输时间和数据预处理的影响，模型获取的观测信息与实际情况存在时延。

为了解决上面的问题，本研究将观测延迟的因素考虑在信号灯控制系统中，提出了一个基于强化学习（RL）的两阶段智能交通信号控制系统，明确地考虑了智能交通信号控制系统获取观测信息的延迟。

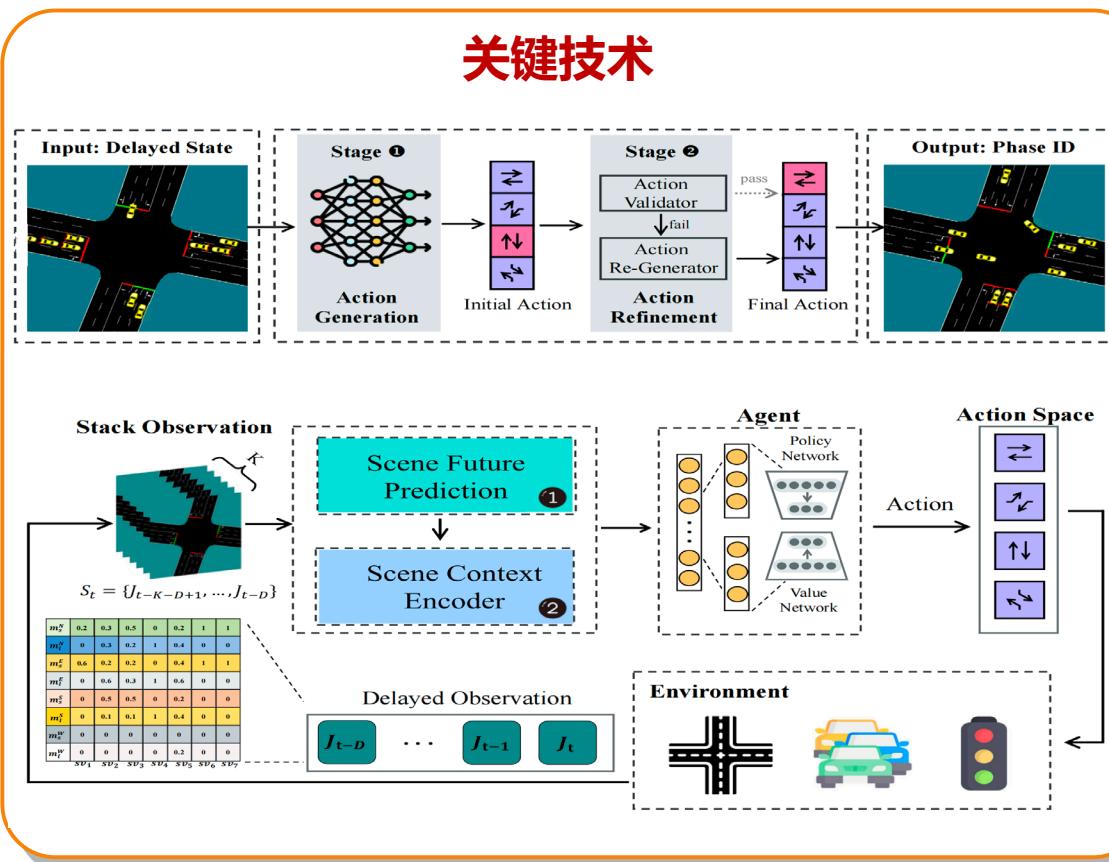
### 主要挑战

- 交通信息复杂，难以捕捉交通状态的时间依赖性和空间依赖性。
- 高维状态空间训练困难，智能体在延迟后的高维观测下可能会出现决策失误。

### 研究方法

- 建立一个两阶段的控制系统，分别是动作生成阶段和动作校准阶段。
- 基于CNN，Transformer等深度学习方法，对观测信息进行特征提取。
- 基于一般的交通规则对智能体给出的动作进行判断，如误差过大，则进行校准。

### 关键技术



**参与人员：**庞奥宇（博一）、王茂南（博二）

### 研究成果

A. Pang, Z. Xu, M. Wang, M.O. Pun and Y. Kan, "Reinforcement Learning-based Traffic Signal Control Using Delayed Observations for V2X", Proc. the 2023 IEEE International Conference in Communications (ICC), Rome, Italy, May 28 – June 1, 2023.



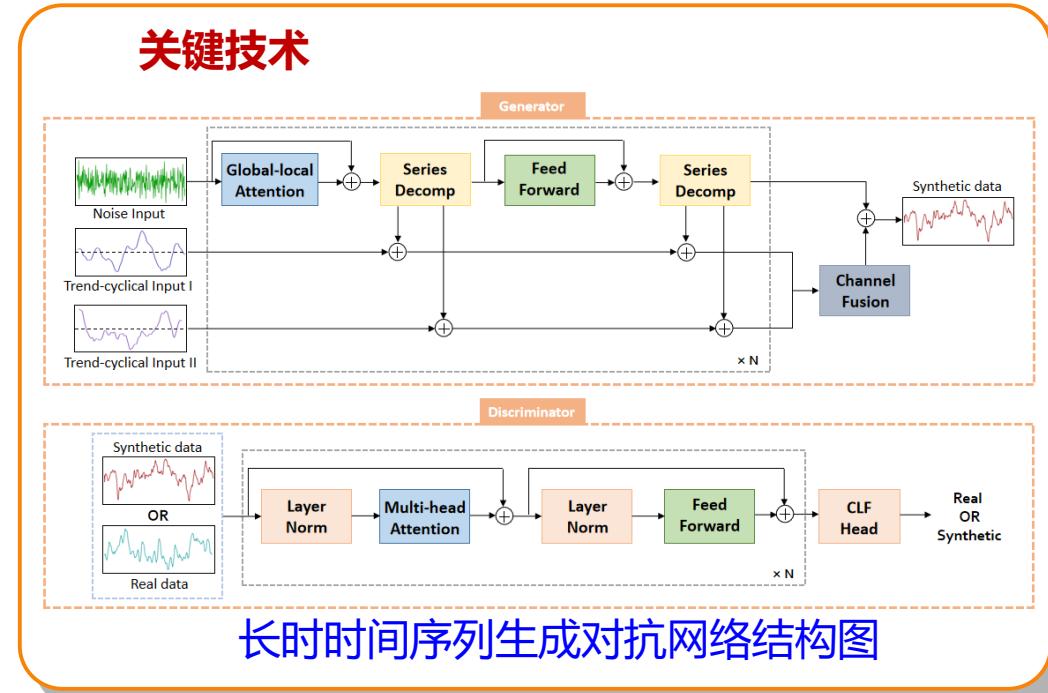
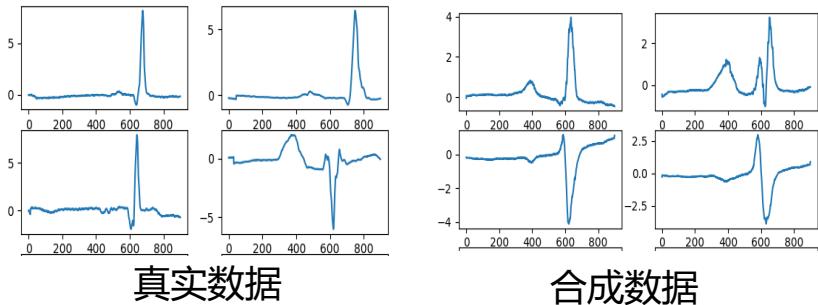
## 子课题1：基于GAN的长时时间序列生成

**研究问题：**时间序列数据是一种常见且重要的数据类型，在金融、医学等领域有重要应用。本研究旨在使用基于Transformer的生成对抗网络捕获长时时间序列的动态特性，以此来生成高质量的、具有实际应用价值的时间序列数据。生成的数据可以用来进行后续的分类、异常检测等工作。

### 主要挑战

- ◆ 以往的生成模型很难捕获较长序列的特征
- ◆ 使用Self-Attention机制会面临平方复杂度以及Loss不收敛的问题
- ◆ 需要同时兼顾全局信息和局部细节
- ◆ 缺少合理的模型性能评价指标

**数据源：**两个开源数据集(UniMib, Clinic ECG)与一个自采集数据集



### 研究方法：

- 改进生成器架构，提出了一种多通道渐进分解的生成器架构，充分利用真实数据的趋势信息
- 引入Auto-Correlation机制，该机制在无缝替代Self-Attention机制的基础上降低了复杂度
- 使用Global-local Attention来平衡全局信息与局部细节之间的关系
- 提出了一种在真实任务中验证生成数据质量的方法

**参与人员：**崔向宇 (硕二)



## 子课题2：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。

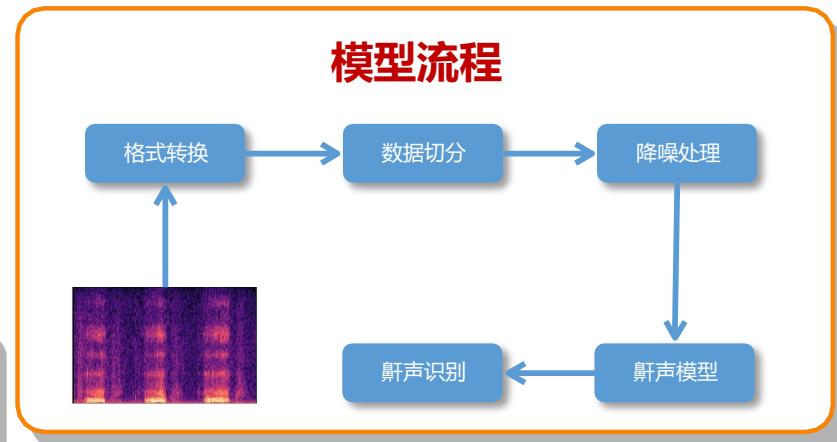
### 主要挑战

- ◆ 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：

崔向宇（硕二）、谢炜枫（硕一）



**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
- 使用numpy改写librosa库中的一些函数以及卷积神经网络，将语音处理与深度学习移植到单片机中
- 采用麦克风阵列实现声源定位，以确定鼾声的方位
- 实现python转c部署，优化内存和性能
- 录制不同大小不同类型的鼾声数据加入训练

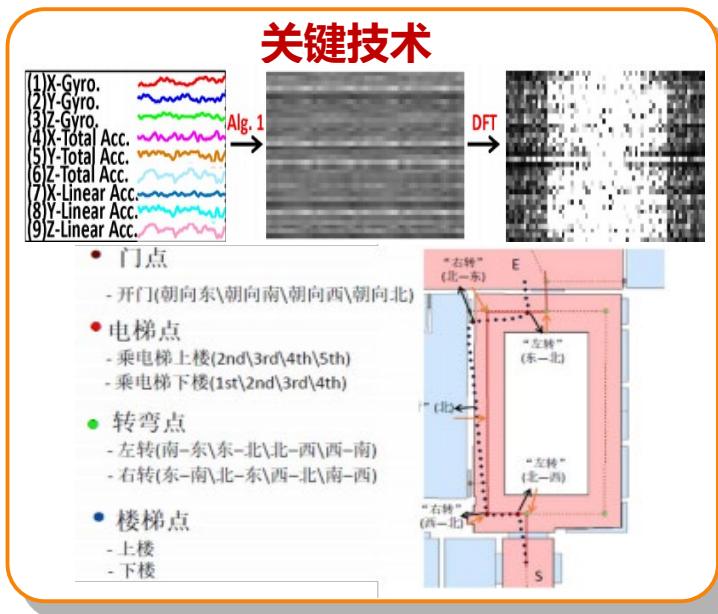


## 子课题3：基于深度神经网络的智能手机室内定位

**研究问题：**定位技术是实现万物智联的关键技术。特别是在智慧城市中，导购、导览、停车场寻车等应用都离不开可靠的室内定位技术。基于卫星信号的GPS定位只能在室外使用，这使得对室内定位的需求非常迫切。

**主要挑战：**传统基于惯性导航的室内定位技术需要依赖高端传感器以达到较高的定位精度。但是通过手机内置的廉价传感器采集到的信号的信噪比较低，难以满足亚米级的定位精度要求。另一方面，基于无线网络信号的指纹定位方法需要对定位场地做大范围的预测量，耗时费力。

- ◆ 室内场所面积大、空间布局结构复杂
- ◆ 数据来源众多，质量不一，难以偶合
- ◆ 实时性以及定位精度要求较高
- ◆ 手机有计算量和耗电的限制



**研究方法：**提出了一种基于深度神经网络的方法，关注行人的活动转换来执行HAR。使用傅里叶变换等算法提取原始时序数据的特征。最后将神经网络的输出与行人航位推算（PDR）技术相结合，以实现行人位置的实时估计和校正。

**研究数据来源：**利用自主开发的APP，在香港中文大学（深圳）的学生活动中心（6层，总面积2500m<sup>2</sup>）采集。采集的数据包括三轴加速度、陀螺仪和磁力计数据，采样频率设置为50Hz。

**参与人员：**梁文迪（硕二）

S. Guo, G. Niu, Z. Wang, M.O. Pun and K. Yang, "An Indoor Knowledge Graph Framework for Efficient Pedestrian Localization", the IEEE Sensors Journal, Vol. 21, No. 4, pp. 5151-5163, Feb. 2021.



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

14



## 子课题4：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

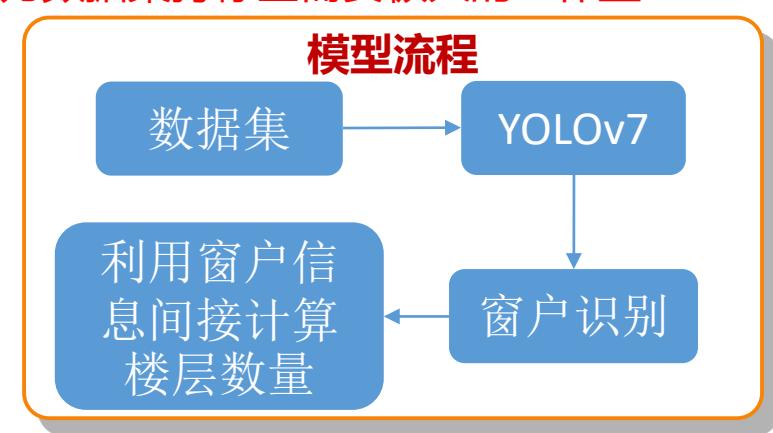


百度地图  
获取街景照片

### 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校正。
- 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。

参与人员：孙宇超（硕二） 卢梓涵（本科生） 吕咏荷（本科生）

