

李旭宸，本科生

2001.10.30 lixuchen2024@ia.ac.cn [li-xuchen](https://www.linkedin.com/in/li-xuchen)
<https://xuchen-li.github.io/> [Google Scholar](#) [Xuchen-Li](#)
单目标跟踪，自然语言跟踪，多模态目标跟踪



教育经历

2020.9 - 至今 **本科生，北京邮电大学，计算机学院（SCS, BUPT）**
专业：计算机科学与技术
毕业设计导师：黄凯奇研究员（IAPR Fellow）
论文题目：基于文本引导的多模态单目标跟踪系统的设计与实现

个人经历

2023.05 - 至今 **科研实习生**：在中国科学院自动化研究所（CASIA）开展多模态目标跟踪算法的研究（导师：黄凯奇研究员）
参与者：参与基于计算机视觉的跨领域研究（来自国内外10余所高校的22位参与者，一周一次，发起人和组织者：胡世宇博士）
2023.04 - 至今 **人工智能菁英班成员**：在中国科学院自动化研究所（CASIA）参与人工智能菁英班考核
2023.03 - 至今 **本科生党支部副书记**：担任北京邮电大学计算机学院（SCS, BUPT）本科生党支部副书记
2023.01 - 2023.05 **科研实习生**：在清华大学（THU）开展室内场景三维重建算法的研究（导师：王好谦教授）
2022.08 - 2023.01 **科研实习生**：在北京邮电大学（BUPT）开展个性化联邦学习算法的研究（导师：刘炳言副研究员）
2022.06 - 2023.06 **志愿者协会主席**：担任北京邮电大学计算机学院（SCS, BUPT）志愿者协会主席
2021.09 - 2023.07 **大班班长/副班长**：担任北京邮电大学计算机学院（SCS, BUPT）大班班长/副班长
2020.09 - 2021.07 **本科生团支部书记**：担任北京邮电大学计算机学院（SCS, BUPT）本科生团支部书记

论文成果

会议论文

1 S. Hu, D. Zhang, M. Wu, X. Feng, **X. Li**, X. Zhao, and K. Huang, "A multi-modal global instance tracking benchmark (mgit): Better locating target in complex spatio-temporal and causal relationship," in *the 37th Conference on Neural Information Processing Systems (NeurIPS, CCF-A Conference, Poster)*, 2023.

共同开发者：平台系统维护 & 升级

2023.12 - 至今 ■ **VideoCube: 大规模多维度全局实例跟踪任务智能评估平台**

链接：<http://videocube.aitestunion.com>.

描述：VideoCube 提出了基于影视学理论的视频叙事内容解耦框架，并基于此构建了一个包含 746 万视频帧的大规模多维度的全局实例跟踪任务智能评估平台。在此基础上，VideoCube 提出了一个借鉴人类层级化认知结构的语义标注框架，从多模态的视角将人类的长视频理解能力耦合进语义标签，从而为多模态研究提供一个良好的基准测试环境。

备注：VideoCube 是两篇研究工作的支撑平台，两篇工作分别被 CCF-A 期刊 TPAMI 和 CCF-A 会议 NeurIPS 接收。截止 2023 年 10 月，本平台获得全球 130+ 国家和地区 23.6 万次访问，1000 余次下载，360 余次算法测试。

■ **SOTVerse: 用户自定义的单目标跟踪任务空间**

链接：<http://metaverse.aitestunion.com>.

描述：SOTVerse 提出 3E 范式作为任务分析框架，其将单目标跟踪任务的代表性实验环境从静态封闭状态转为动态开放状态，实验环境总计包含 1256 万视频帧。此外，SOTVerse 包含了一个子空间构建策略，可以依据用户的研究目标快速构建子空间开展自定义的研究。最后，SOTVerse 也配备了多个评测指标和评价系统，旨在帮助研究者在多元的实验环境和多样化评测标准的帮助下快速发现算法的性能瓶颈，并开展针对性的研究。

备注：SOTVerse 是一篇 CCF-A 期刊 IJCV 工作的支撑平台。截止 2023 年 10 月，本平台获得全球 100+ 国家和地区 8.8 万次访问。

■ **GOT-10k: 大规模广覆盖单目标视觉跟踪评估平台**

链接：<http://got-10k.aitestunion.com/>.

描述：GOT-10k 是用来评估算法在未知目标类别和运动模式上泛化性的评估平台。该平台提供了 1 万段视频序列，包含 563 类目标类别和 87 种运动模式，总计 150 万个矩形框标注。和已有的单目标跟踪基准相比，GOT-10k 在目标类别的丰富程度上要远超过现有工作，其提出的训练-测试类别不重合的评测思路也成为单目标跟踪评估评测领域的研究规范。

备注：GOT-10k 是一篇 CCF-A 期刊 TPAMI 的支撑平台。截止 2023 年 8 月，本平台获得全球 150+ 国家和地区 244 万次访问，6 千余次下载，1.4 万余次算法测试。

独立研究者：算法设计 & 分析

2023.09 - 至今 ■ **基于文本引导的多模态单目标跟踪系统的设计与实现**

描述：本项目提出了一种简单、灵活、有效的视觉语言（VL）跟踪流程，它将 VL 跟踪视为 token 生成任务。提出的跟踪框架将语言描述和边界框序列化为一系列离散的 tokens。在这种新的设计范式中，所有的 token 查询都需要感知期望的目标，并以自回归的方式直接预测目标的空间坐标，允许跟踪器使用简单的交叉熵损失作为 VL 跟踪任务的统一优化目标。

备注：该项目是我在北京邮电大学的学士毕业论文的一部分。

科研项目 (continued)

共同研究者：数据整理 & 算法学习

2023.05 - 2023.09 ■ **MGIT：一种面向复杂时空因果关系的多模态视频跟踪评估基准**

链接：<http://videocube.aitestunion.com>

描述：VideoCube 提出了基于影视学理论的视频叙事内容解耦框架，并基于此构建了一个包含 746 万视频帧的大规模多维度的全局实例跟踪任务智能评估平台。在此基础上，MGIT 提出了一个借鉴人类层级化认知结构的语义标注框架，从多模态的视角将人类的长视频理解能力耦合进语义标签，从而为多模态研究提供一个良好的基准测试环境。

备注：MGIT 基准被 CCF-A 会议 NeurIPS 接收。

技能

语言 ■ 汉语普通话（母语）和英语。

编程 ■ Python, C++, C, \LaTeX 。

Linux ■ Shell。

其他 ■ 学术研究，团队协作能力，领导力。

奖励及荣誉

2023 ■ **国家奖学金**，中华人民共和国教育部（获奖率：1%，排名：1/455）。

■ **北京市三好学生**，北京市教育委员会。

■ **中国科学院大学大学生奖学金**，中国科学院大学。

■ **国家级三等奖**，华为ICT大赛（队长）。

■ **国家级三等奖**，中国机器人及人工智能大赛。

■ **国家级项目**，大学生创新创业训练计划项目（队长）。

■ **国家级项目**，大学生创新创业训练计划项目。

■ **省部级一等奖**，中国国际互联网+大学生创新创业竞赛（北京赛区）。

2022 ■ **国家奖学金**，中华人民共和国教育部（获奖率：1%，排名：2/430）。

■ **华为智能基座奖学金**，教育部-华为智能基座联合工作组。

■ **国际级一等奖**，美国大学生数学建模竞赛。

■ **国家级一等奖**，中国高校计算机大赛-人工智能创意赛（队长）。

2021 ■ **国家励志奖学金**，中华人民共和国教育部。

■ **浩瀚奖助学金**，北京浩瀚深度信息技术股份有限公司。