

钱姿

兴趣方向: 多模态学习 | 计算机视觉 | 机器学习

✉ qian-z20@mails.tsinghua.edu.cn · ☎ (+86) 130-0127-9855

🎓 教育背景

清华大学	硕士, 计算机科学与技术, 媒体所, 导师: 朱文武教授	2020 – 2023
清华大学	学士, 计算机科学与技术, GPA: 3.61/4.0	2016 – 2020
清华大学	辅修, I-Center, 智能交通创新创业	2017 – 2018

🔬 主要研究经历

用于视频问答与推理的动态时空模块化神经网络 2021.08 – 2022.06

清华大学媒体所

- 提出并设计用于视频问答与推理任务的动态时空模块网络, 第一次将模块化神经网络引入该任务。
- 设计并实现空间推理、时间推理和逻辑推理模块, 动态组装模块为树状逻辑推理树进行跨模态融合与推理, 有效提高了模型的可解释能力和文本泛化能力, 并且在公开数据集上取得当前最高准确率。
- 该工作相关论文以第一作者身份发表 ACM MM'22, 该工作相关专利正在申请中。

视频-文本任务在持续小样本设定下的研究 2022.06 – 至今

清华 - 阿里安全部合作项目, 学术合作实习生

- 研究视频-文本多模态风险数据的风险点分类问题, 随着风险类别增加, 持续更新模型。
- 基于大规模预训练多模态模型和提示学习, 设计可学习的提示向量池, 对于持续到来的新风险类别训练新的提示向量, 拟实现在新风险类别上的持续小样本学习。

机器人操作技能学习算法研究及仿真环境设计 2019.11 – 2020.05

本科毕业设计

- 使用 V-Rep 平台搭建包含机器人与机械臂的抓取任务仿真场景, 在搭建好的不同场景任务中分别设计反馈函数, 应用克隆学习、在宽约束下基于不完美的示教的强化学习进行训练, 有效实现抓取任务。

基于 RFID 传感器的普适计算应用场景研究 2019.02 – 2019.10

密西根大学安娜堡分校, 暑期访问学生

- 研究 RFID 传感器用于无穿戴实时定位场景, 设计三种绘制层析图像的矩阵算法用以显示多人的二维位置, 设计评估算法从层析图像获取用户二维位置, 搭建检测实时位置系统, 平均误差为 144cm。

清华大学媒体所, 科研助理

- 研究 RFID 传感器用于日常物品 DIY 场景, 设计基于 RFID 传感器的无线按钮和滑动条, 实现检测传感器状态和操作的低延迟双层算法, 搭建用户实验平台前端部分, 相关论文以第二作者身份发表 CCF TPCI'22。

🏆 学术成果

论文发表

- [Zi Qian](#), Xin Wang, Xuguang Duan, Hong Chen, Wenwu Zhu. “Dynamic Spatio-Temporal Modular Network for Video Question Answering.” *ACM International Conference on Multimedia*. 2022. (Full paper, CCF-A).
- Tengxiang Zhang, [Zi Qian](#), Hsuan Wei Fan, Jie Ren, Yuntao Wang, Yuanchun Shi. “Easily-add battery-free wireless sensors to everyday objects: system implementation and usability study.” *CCF Transactions on Pervasive Computing and Interaction* 4. 2022.

专利申请

- 用于视频问答的动态时空模块化网络方法。发明人: 朱文武, 王鑫, [钱姿](#)。(正在申请)

🏆 个人荣誉

清华大学第三十六届挑战杯三等奖	2018
学业优秀奖 (前 20%)	2016
全国中学生物理联赛省一等奖	2014

⚙️ 个人技能

- 编程语言: C/C++, Python, Java, JavaScript, Matlab, Shell
- 工具和框架: PyTorch, Scikit-Learn, OpenCV, Git, LaTeX, Google VR SDK, Vue, CAD, Sketch up
- 语言: TOEFL (107), 听说读写流利, 可独立阅读英文文献, 进行学术写作