

张沐阳

年龄: 26岁

性别: 男

联系电话: 17303642317

联系邮箱: muyang.zhang@mais.ia.ac.cn

所在机构: 中科院自动化所 多模态人工智能系统实验室 (原模式识别国重)



教育背景

2017-09 ~ 2021-07	哈尔滨工程大学	机械工程 (本科)
2022-09 ~ 至今	中国科学院大学	电子信息 (硕士)

研究方向

图像生成; 视频生成; 多模态生成; 模型轻量化

科研成果

第一/共一作者:

- 1.PanoDiT: Panoramic Videos Generation with Diffusion Transformer (AAAI, 2025, CCF-A, 第一作者)
- 2.AG-SDM: Aquascape Generation based on Stable Diffusion Model with Low-Rank Adaptation (CASA, 2024, CCF-C, 第一作者, 获Wiley中国高贡献作者奖-2024第二季度)
- 3.PDFT: Parameter-Diminish Fine-Tuning for Transformer-based Models (The Visual Computer, SCI二区, 第一作者)
- 4.Controllable Panoramic Video Generation with 360-Degree Motion Consistency for Multiple Control Tasks using a Unified Framework (ICCV, 2025, CCF-A, 共一作者, 在投)
- 5.AccidentX: A Large-Scale Multimodal BEV Dataset for Traffic Accident Analysis and Prevention (IROS, 机器人顶会, 第一作者, 在投)

其他作者:

- 1.HumanDreamer: Generating Controllable Human-Motion Videos via Decoupled Generation (CVPR, 2025, CCF-A, 第七作者)
- 2.AASD: Accelerate Inference by Aligning Speculative Decoding in Multimodal Large Language Models (DAC, 2025, CCF-A, 第三作者)
- 3.FeaCo: Reaching Robust Feature-Level Consensus in Noisy Pose Conditions (ACMMM, 2023, CCF-A, 第三作者)
- 4.HTCViT: An effective network for image classification and segmentation based on natural disaster datasets (CGI, 2023, CCF-C, 第三作者)
- 5.Vision Also You Need: Navigating Out-of-Distribution Detection with Multimodal Large Language Model (ICCV, 2025, CCF-A, 第四作者, 在投)
- 6.TR-DQ: Time-Rotation Diffusion Quantization (ICCV, 2025, CCF-A, 第五作者, 在投)
- 7.EventVAD: Event aware Train-Free Video Anomaly Detection (ACM MM, 2025, CCF-A, 第七作者, 在投)
- 8.Memory Efficient Point Cloud Segmentation with Spatial Group Attention (ACM MM, 2025, CCF-A, 第六作者, 在投)
- 9.Robust Detection in Complex Construction Sites: HiPA-DETR with Weather-Aware and Cross-Domain Generalization (ACM MM, 2025, CCF-A, 第五作者, 在投)

项目经历

2022-09 ~ 2023-08	大尺度灾害识别定位装备 国家任务	合作单位: 交通运输部科学研究院
-------------------	------------------	------------------

- 项目描述: 在无人机航拍图像上实现对受灾地区的场景重建, 废墟、滑坡、洪水等受灾区域精确识别。
- 项目职责: 完成其中的废墟区域识别任务; 基于Yolo模型在大尺寸航拍图像上识别废墟建筑和车辆达到90%以上准确度; 该项目产出的论文被《Vision Computer》期刊和机器人顶级会议IROS接收。

🏆 学习成果

- 2023-2024 中国科学院大学三好学生
- 2022-2023 中国科学院大学人工智能论坛 三等奖
- 2022-2023 中国科学院大学三好学生
- 2020-2021 CET-6 485
- 2022-2023 JLPT N2 159

🎓 实习经历

2025-02 ~ 至今

北京智谱华章科技有限公司

算法实习生

主要负责2D数字人的算法优化和落地，
结合全景视频生成和数字人技术探索虚拟直播技术，[NeurIPS 2025](#)

⚙️ 学生工作

2023-2024学年 中科院自动化所研究生会 宣传部部员

2024-2025学年 中科院自动化所研究生会 实践部副部长