

## 教育背景

西安电子科技大学

本科 - 人工智能学院 - 智能科学与技术

中国西安

2022–2026

- 绩点 (GPA): 3.70/4.00 加权平均分: 85.00/100.00 排名: 专业课 6.95%, 必修课 12.88%
- 核心课程: 算法 (99), 数据结构 (98), 电路基础 (96), 模式识别 (94), 高等数学 (94), 工程概论 (93), 程序设计 (93), 机器学习 (92), 最优化方法 (90), 概率论与数理统计 (90), 数值计算 (90).

## 研究兴趣

我的研究兴趣及希望探索的深度学习方向主要集中在以下几个方面:

- 基于推理的学习 (Reasoning-based Learning):** 通过与真实世界的交互来推断知识, 例如从观察中推断物体的运动、接触、力学交互、功能可见性 (affordances) 等。利用对真实世界进行推理的能力, 机器人或智能体可以获得更强的泛化能力, 并可能在未来发展出主动学习的能力。
- 生成式建模 (Generative Modeling):** 旨在使用生成式方法对受物理定律约束的物体进行建模, 包括特征级、语义级和场景级的表示, 最终目标是构建世界模型。

## 研究经历

- 机器视觉与智能研究组** 上海交通大学  
研究助理 (导师: 卢策吾教授, 杨理欣教授) 2024 年 7 月至今  
从事机器人学习领域的研究, 重点关注对真实世界的感知与推理。
- 协同智能系统教育部重点实验室** 西安电子科技大学  
研究助理 (导师: 公茂果教授, 李豪教授) 2023 年 9 月至 2024 年 7 月  
参与针对计算机视觉系统的对抗攻击研究。

## 发表论文

- Dense Policy: Bidirectional Autoregressive Learning of Actions**  
苏悦\*, 詹欣宇\*, 方泓杰, 薛寒, 方浩树, 李永露, 卢策吾, 杨理欣<sup>†</sup>  
提出 Dense Policy, 一种双向自回归机器人策略, 通过从稀疏关键帧逐步扩展动作来推断轨迹, 展现了超越基于扩散模型的策略的能力。  
[\[arXiv\]](#) [\[代码\]](#) [\[项目主页\]](#)
- Motion Before Action: Diffusing Object Motion as Manipulation Condition**  
苏悦\*, 詹欣宇\*, 方泓杰, 李永露, 卢策吾, 杨理欣<sup>†</sup>  
提出 MBA, 一种新颖的模块, 采用两个级联扩散过程, 在物体运动引导下生成机器人动作。MBA 被设计为即插即用组件, 可灵活集成到现有的具有扩散动作头的机器人操纵策略中。  
[\[arXiv\]](#) [\[代码\]](#) [\[项目主页\]](#)
- Generative Adversarial Patches for Physical Attacks on Cross-Modal Pedestrian Re-Identification**  
苏悦, 李豪<sup>†</sup>, 公茂果<sup>†</sup>  
一种针对可见光-红外跨模态行人重识别 (VI-ReID) 模型的生成式对抗攻击方法, 通过扰动模态不变特征生成对抗补丁, 暴露了当前最优模型 (SOTA) 的脆弱性, 并强调了增强特征提取能力的必要性。  
[\[arXiv\]](#)
- AdvDisplay: Adversarial Display Assembled by Thermoelectric Cooler for Fooling Thermal Infrared Detectors**  
李豪<sup>†</sup>, 万方高, 苏悦, 武越, 张明阳, 公茂果<sup>†</sup>  
以往的红外对抗攻击是一次性的且不灵活。我们利用热电制冷器 (TEC) 实现了可适应不同场景的攻击, 导致行人检测模型产生误判。  
[\[AAAI-2025\]](#)

## 项目经历

---

- **MetaPalace: 带你进入故宫元宇宙**

我们实现了故宫官网未能做到的功能：通过单视图重建提供文物的 3D 视图，并利用检索增强生成（RAG）技术开发了交互式大语言模型导览员。

[\[项目主页\]](#) [\[前端代码\]](#) [\[后端代码\]](#)

- **U-pre: U-Net 作为时间序列预测的优秀学习器**

时间序列预测适用于 U-Net 架构，因其输入输出分布一致且具有很强的数学一致性。将 U-Net 与 Bert-Encoder 相结合，通过融合局部和全局注意力提升了性能。

[\[代码\]](#) [\[报告-中文\]](#)

- **AgentCrossTalk: 两个大语言模型智能体表演相声**

本项目使用 Google Gemini 创建了一个简单的聊天机器人应用，模拟两个相声演员根据用户提供的主题进行表演。

[\[代码\]](#) [\[项目主页\]](#)

- **FGSM3D: 点云梯度扰动攻击是否可行？**

我们尝试将 FGSM 扩展到 3D 领域，并在一定梯度范围内取得了显著成功，但 3D 模型的采样方式告诉我们事情似乎没那么简单……

[\[代码\]](#) [\[报告-中文\]](#)

- **AcoFlow: 最大流问题的启发式搜索**

寻找最大流问题的关键在于如何设计更好的启发式信息来找到增广路径。我们通过蚁群算法大胆地挑战了这个问题。

[\[代码\]](#) [\[报告-中文\]](#)

## 获奖情况

---

- 2023 年全国大学生数学建模竞赛省级一等奖
- 2024 年全国大学生数学建模竞赛省级一等奖
- 2024 年中国大学生计算机设计大赛西北赛区二等奖

## 社区经历

---

- 西安电子科技大学微软俱乐部 - 科研部部长