



教育背景

电子科技大学 英才实验学院（荣誉学院） 2024.9 - 2028.6
专业: 数理基础科学（“成电英才计划”拔尖创新人才实验班） 方向: 人工智能
GPA: 3.96/4.0 加权均分: 91.55 专业排名: 2/24 CET-4: 581 CET-6: 561
核心课程: 高等微积分、高等代数与几何、大学物理、概率论与数理统计、机器学习等。

科研经历

► Computational Pathology & Multiple Instance Learning 计算病理与多示例学习 2025.3 - 2025.8

CKMIL: Cascaded Key-Instance Attention Multiple Instance Learning for Histopathology Whole Slide Image Analysis [pdf] Submitted to AAAI 2026 第一作者

指导教师: 电子科技大学 数学科学学院 邓良剑教授 香港科技大学 陈浩助理教授

- 针对病理全切片图像（WSI）关键诊断信号稀疏、全局交互易稀释关键信息的问题，提出级联关键实例注意力框架 CKMIL，以关键实例引导高效依赖建模；在 BRACS 与 TCGA 多队列癌症分型和生存预测任务上取得领先性能。
- 个人贡献: 项目负责人，统筹研究进度，参与数据处理、方法构思、实验与论文撰写，主导项目全流程推进。

► LVLML Safety & Jailbreak Attacks 多模态大模型安全与越狱攻击 2025.9 - 2025.12

Persuasion in Scene: A Multi-Agent Typographic Jailbreak Crew against Large Vision-Language Models [pdf] Submitted to CVPR 2026 共同第一作者

Models as Lego Builders: Assembling Malice from Benign Blocks via Semantic Blueprints [pdf] Accepted by CVPR 2026 第三作者

- 围绕 LVLML 黑盒越狱安全，提出多智能体跨模态协同框架 PIS 与基于语义槽填充漏洞的单次、免优化攻击框架 StructAttack；在主流开源 LVLML 上取得领先攻击成功率，揭示场景排印与结构化视觉提示的安全风险。
- 个人贡献: 作为 PIS 工作共同第一作者，核心参与研究构思、方法设计、实验推进与论文撰写；作为 StructAttack 工作第三作者，负责方法在 SafeBench 数据集上的适配与评测实验。

► Long-Context Test-Time Training 长上下文测试时训练 2026.1 - 2026.8

SPLIT: Write to Learn, Query to Retrieve in Long-Context Test-Time Training Submitted to ICLR 2027 第一作者

- 揭示长上下文测试时训练中内容归纳与稀疏检索的更新冲突，从 Where to Update 出发，提出双路径框架 SPLIT，以中层 MLP “写”路径学习可复用结构、深层 Query “读”路径增强长程检索；在 RULER、HELMET-ICL 与 LongBench 上取得显著提升。
- 个人贡献: 项目负责人，设计 Where、When、How to TTT 预实验并提出核心 Idea，推动方法从后训练方案向 Train-free 验证框架调整，负责全流程实验与论文撰写。

► Pathology Foundation Models & Efficient WSI Analysis 病理基础模型与高效全切片分析 2025.12 - 至今

From Full-Bag Diagnosis to Budgeted Patch Selection: Role-Decoupled Benchmarking and Adaptation of Pathology Foundation Models in Whole-Slide Imaging Expected Submission to IEEE TMI 2027 第一作者

指导教师: 电子科技大学 数学科学学院 邓良剑教授 香港科技大学 陈浩助理教授

- 从病理图像压缩出发，发现部分形态学任务无需依赖图像重建，并将其归因于病理图像的高度冗余性与基础模型（FM）的鲁棒性；基于这两项性质，研究 Selector×Consumer 场景下 FM 的能力解耦，并提出适配修复方案。

► Multimodal Post-Training & Chain-of-Thought Reasoning 多模态模型后训练与思维链推理 2026.4 - 至今

Seeing Again, Seeing Differently: ReMAP for Reasoning-Time Visual Memory Beyond One-Shot Encoding Submitted to ICLR 2027 第一作者

- 通过预实验量化一次性视觉编码瓶颈：随长链推理推进，初始视觉证据的因果影响持续衰减；固定特征回放虽可恢复已保留信息，却无法补回首次编码遗漏的关键证据，形成 replay ceiling。受感知循环理论启发，提出 ReMAP 推理时自适应视觉工作空间，由当前推理状态在 Continue、Global Reorientation 与 Local Replay 间调节感知范围，以全局重定向缓解局部固着、恢复场景一致性，以局部回放返回原始像素并重编码细节证据，最终由模型按需 invoke 相应的 latent memory 支持后续推理。

专业技能

- 编程与工具: Python; 熟练使用 PyTorch、Linux、Git、LaTeX、Codex、Claude Code; 具备 Idea 验证、实验统筹与论文写作能力。
- 深度学习: 熟悉 Transformer、Mamba 架构; 具备多模态大模型测试时训练、后训练及系统评测经验; 数理、学术英语基础扎实。

荣誉奖项

- 电子科技大学优秀学生奖学金 一等 2025 年
- 美国大学生数学建模竞赛（MCM/ICM） Meritorious Winner（一等奖） 2026 年