邹元昊 Yuanhao Zou

16680807958 | yuanhaoz@ucf.edu | 微信：AlexWhyNot
个人主页：alexco1d.github.io | 学术主页 | GitHub：AlexCo1d

求职意向：多模态 / 视频理解 / 高效视觉语言模型 / 强化学习智能体

实习安排：2027 年 4 月起，每周 5 天，可连续实习 4-6 个月

教育背景

中佛罗里达大学（UCF） 计算机科学 · 博士 2025.08 - 2028.05（预计）
GPA：4.0/4.0 | 导师：Prof. Chen Chen | 计算机视觉研究中心（CRCV）

密歇根大学安娜堡分校 电子与计算机工程 · 硕士 2023.08 - 2025.05
GPA：4.0/4.0

中南大学 计算机科学与技术 · 学士 · 88/100 2019.09 - 2023.06

实习经历

OpusClip 人工智能研究实习生 | 硅谷 2026.10 - 2027.02（预计）
负责核心产品 **Opus Clip 视频剪辑智能体**的训练，聚焦强化学习后训练与推理能力优化，围绕视频理解、片段选择和剪辑规划中的多步决策开展研究。

联影智能（UII） 多模态算法研究实习生 | 波士顿 2026.05 - 2026.08
面向热成像、深度及多视角视频，训练基于**强化学习的多模态推理智能体**，研究跨模态信息融合与复杂视频问答，支持多模态联合推理。

科研与合作项目

高效视频理解与长视频推理 UCF · Prof. Chen Chen 课题组 2025.02 - 至今
自适应选帧（A.I.R.）：以轻量视觉模型引导大模型迭代选帧，平均 **32 帧替代 128 帧**；在 NExT-QA / LongVideoBench 上最高提升 7.0% / 4.5%，推理耗时降低约 **74%**（162 秒降至 42 秒）。
长视频多事件定位（CoMET）：构建 600 个长视频、2,789 条多条件查询的基准；图搜索智能体将 GPT-5 的 F1@0.5 从 **10.1% 提升至 16.2%**，与高通人工智能研究院合作。

端侧视频时序定位 Axon 科研合作 | 远程（奥兰多） 2025.10 - 2025.12
面向 Axon AB4 等端侧设备研究连续时间解码与非自回归定位；**15 亿参数**模型在 Charades-STA 达到 **57.1 mIoU**，超过 70 亿参数基线的 27.5，并降低推理延迟。

医学视觉语言模型 纽约州立大学石溪分校 · Zhaozheng Yin 教授课题组 2024.02 - 2024.11
提出硬负样本挖掘、表示对齐与选择性知识融合方法；VQA-RAD 准确率 **80.5%**（较 BiomedCLIP 提高 5.3%），SLAKE / PathVQA 达到 **86.7% / 65.3%。**

医学图像分割 宁波诺丁汉大学 · Xiangjian He 教授课题组 2022.09 - 2023.08
融合 U-Net、Transformer 与多尺度注意力；DDTI 上达到 **85.2% Dice / 74.6% mIoU**，并在三个公开数据集上验证泛化能力。

代表论文

NeurIPS 2026 E&D • 第一作者 CoMET：长视频条件多事件时序定位（已接收）
ICLR 2026 • 共同第一作者 A.I.R.：自适应、迭代式推理引导的视频选帧
CVPR 2025 • 第一作者 AMiF：医学视觉问答的表示对齐、负样本挖掘与知识融合
ECCV 2026 • 第二作者 CLARITY：面向疾病轨迹模拟与治疗决策的医学世界模型
CVPR 2026 研讨会 • 共同第一作者 视频大模型高效时序定位范式研究
CVPR 2025 研讨会 • 第一作者 MVCM：医学视觉问答与图文检索的多视角跨模态对齐
Knowledge-Based Systems • 共同第一作者 HFA-UNet：甲状腺结节分割的混合注意力网络

技术能力与学术服务

技术：Python、C++、PyTorch、Transformers、DeepSpeed、vLLM；多卡分布式训练、Linux、Docker、Slurm、OpenCV。
会议审稿：ICLR 2027（5 篇）、NeurIPS 2026（4 篇）。
CVPR 2026 ECV 研讨会（2 篇）、ECCV 2026 MedVidU 研讨会（1 篇）、WACV 2027 MAVU 研讨会（2 篇）。
期刊审稿：Medical Image Analysis、Pattern Recognition、IEEE TMM 等 8 本期刊；累计完成 **23 次审稿**（含修订轮次）。