

付临

☎ 19857109661 · ✉ 748544794@qq.com · 📍 杭州
当前状态：在校生

教育背景

浙江大学 硕士，电子信息工程，集成电路学院，研究方向：多模态大模型、强化学习 2024.09 – 2027.06
浙江大学 本科，信息与计算科学，数学科学学院，GPA: 3.52 / 4.00 (保研) 2020.09 – 2024.07

实习经历

阿里巴巴 淘天，多模态大模型算法工程师 2026.06 – 至今

- **AI 视频的异常检测**: 面向淘宝商家 AIGC 商品展示视频的上线质检，聚焦肢体畸变、时序突变及物体异常出现或消失等视觉缺陷；针对现有多模态大模型对细粒度生成异常感知能力不足的问题，开展异常识别能力增强与业务落地，提升问题视频的自动筛查与过滤能力。

阿里巴巴 达摩院，研究型实习 2026.02 – 2026.06

- **GUI Agent World Model Benchmark**: 参与 GUI-CC 的基准设计、数据构建与模型评测，搭建离线轨迹回放和在线 Agent-loop 流程。详细内容见研究经历。

华为技术有限公司 终端 BG，算法工程师 2025.04 – 2025.08

- **AIGC 辅创工具**: 集成扩图、重绘、擦除、抠图、色彩增强与超分能力，完成前后端推理闭环及多模型接口、mask 和分辨率适配，沉淀可复用的 Demo 链路。
- **自动配图优化**: 针对关键帧不佳、主体截断等封面 badcase，引入大模型判断抠图需求与裁剪幅度，结合规则/检测模块升级主体保留策略，减少线上异常。

研究经历

VideoKR: 知识与推理密集型视频理解数据与后训练 2025.07 – 2026.02

- **研究动机**: 针对现有视频理解数据缺少结构化领域知识与高质量推理监督、难以支撑科学场景后训练的问题，提出 VideoKR，构建知识驱动的视频理解数据体系。
- **数据与实验**: 构建覆盖 82 个学科、6K 个知识点的课程级知识库，知识点驱动检索收集 145K 个 YouTube(CC) 视频；扩展生成 VR / KV / KVR 三类样本，得到 201K SFT 数据与 114K RL 数据，并基于 Qwen2.5-VL-7B、Qwen3-VL-8B 开展 SFT+GRPO 训练。
- **评估与分析**: 在 Video-MME、MVBench、VideoMMMU、SciVideoBench 等基准上，SFT+RL 相比仅 SFT / 仅 RL 带来一致增益，模型能够更好融合视觉证据与领域知识完成推理。
- **成果**: **Lin Fu***, Zheyuan Yang*, Yang Wang, Tingyu Song, Arman Cohan, Yilun Zhao. *VideoKR: Towards Knowledge- and Reasoning-Intensive Video Understanding*. **ICML 2026 Spotlight**.

Sci-VBench: 面向科学推理的视频生成评测基准 2025.12 – 2026.04

- **研究动机**: 针对现有视频生成评测偏重视觉质量与文本对齐、缺少对科学知识约束、多步因果推理与时序一致性衡量的问题，提出 Sci-VBench。
- **数据与实验**: 构建覆盖自然科学、医疗健康、人文社科、工程四大领域、78 个学科的 1500 条专家标注样本，并为每个样本配套 high-level reference guide 与 1-5 分 anchored rubric，系统评测 11 个前沿闭源/开源文本到视频模型。
- **评估与分析**: 提出结合自动感知指标与 rubric-conditioned MLLM-as-Judge 的评测框架，验证引入 evaluation specification 后非专家与 MLLM 与专家评分的一致性提升；结果表明当前模型在科学因果正确性与时空一致性上仍存在明显不足。
- **成果**: Diandian Zhang*, Tingyu Song*, **Lin Fu***, Zheyuan Yang, Yilun Zhao. *Sci-VBench: Evaluating Knowledge- and Reasoning- Intensive Video Generation in Science Domains*. **COLM 2026**.

GUI-CC: 面向智能体环境的 GUI 世界模型上下文一致性评测 2026.02 – 2026.06

- **研究动机**: 针对单步预测难以衡量 GUI 世界模型长期可靠性的问题，提出 GUI-CC，评估多步交互中的上下文一致性、状态保持与任务可执行性。
- **基准构建**: 设计离线参考动作与在线 Agent-loop 两类轨道；基于 GUIOdyssey 构建 500 个离线轨迹任务（130 个应用、4,905 次转移），并构建 200 个经 Android 模拟器验证的在线任务（30 个应用、6 类场景、18 种任务模板），标注有序任务里程碑。
- **评估与分析**: 从转移保真度、合理性、上下文一致性与任务进展四个维度评测 18 种模型设置，揭示单步视觉可用性与长期环境可靠性之间的差距，并归纳世界知识缺失、误差累积与上下文不一致三类主要失败。
- **成果**: **Lin Fu***, Zheyuan Yang*, Tianhui Zhang*, Jinbiao Wei, Guo Gan, Boxu Liu, Yilun Zhao, Yu Rong. *GUI-CC: Benchmarking Contextual Consistency of GUI World Models as Agent Environments*. **EMNLP 26 Findings**.

专业技能

编程与框架: Python, PyTorch, LLaMA-Factory, Verl
训练与微调: 熟悉多模态大模型的 SFT、LoRA、QLoRA 等微调方法
后训练与对齐: 熟悉 DPO、RLHF、RLVR 等常见后训练与对齐范式
工程与数据: 熟悉机器学习、深度学习、强化学习与常用训练流程，具有大规模数据集集成经验