

陈衍鑫 | 个人简历

🌐 <https://github.com/SSSSuperC> ✉ yanxinchen26@m.fudan.edu.cn

☎ (+86) 13363979511 (+852) 61587613

研究方向

多模态大模型, 视频语言模型 (VLM), 长视频理解, 检索增强生成 (RAG), 高效训练与推理, 模型训练 Infra

教育经历

复旦大学

人工智能博士

上海, 中国

2026.09 - 至今

香港城市大学

数据科学硕士

香港, 中国

2024.09 - 2025.10

绩点: 3.76/4.0

西北大学

软件工程本科

陕西西安, 中国

2020.09 - 2024.06

绩点: 3.35/4.0 (前 20%)

奖项 & 荣誉

- 杰出表现奖、最佳项目奖, 香港城市大学, 2025
- 校级一等奖学金 (2 次), 西北大学, 2022、2023
- 校级二等奖学金, 西北大学, 2021

科研与工程经历

复旦大学 NLP 组: MOSS-VL 训练与 Infra 优化

实习经历

- 千卡级训练经验: 参与 MOSS-VL 的大规模训练工作, 负责 pretrain 4 个 stage 以及 SFT 阶段的全流程消融实验与正式训练; 在千卡级集群环境中跟进训练稳定性、吞吐效率、数据异常与框架问题, 具备大规模任务排障和训练推进能力。
- VLM 架构探索: 使用基于 cross-attention 的图文交互方式降低原始 self-attention 交互的计算负载, 使模型更适用于流式视频理解与推理场景; 引入 cross-gate 机制以提升训练稳定性。
- 训练 Infra 优化: 适配 Megatron 中 cross-attention VLM 的并行策略与激活重计算, 修改 FA3 以支持 cross-attention mask, 从而降低显存占用并加速训练; 推动视频可训练帧数上限扩展和大规模训练高效运行。

香港城市大学 AML 实验室: 多模态 RAG 与长视频理解研究

科研经历

- 长视频 RAG: 面向超长视频理解中的上下文长度限制与浅层检索问题, 研究基于 probe 的检索与记忆增强方法, 通过探针生成、结构化注释和重排序提升模型对深层交互语义与跨模态上下文的定位能力。
- 评测与分析: 围绕长视频开放式理解任务设计 LLM-as-Judge 与人工评估对比实验, 并扩展到 LVBench、LongerBench 等 benchmark 设置, 验证方法在理解深度、信息密度和可信度等维度的提升趋势。

学术论文

- 共一作者, "ProEchoMem: Enhancing Long Video Understanding via Multi-Trace Probe-Echo Memory" SIGIR '26. CCF-A

核心内容: 针对长视频理解中上下文受限、浅层相似度检索难以覆盖深层交互语义的问题, 提出基于多轨迹

probe-echo memory 的检索与记忆增强框架，通过探针生成、结构化记忆组织与重排序机制提升超长视频场景下的语义定位与推理能力。

技术栈与能力

- **多模态大模型**：熟悉 VLM 训练流程、视频-文本跨模态交互、cross-attention 架构、长视频上下文建模与流式视频理解场景。
- **训练系统与 Infra**：熟悉 PyTorch、Megatron-LM、DeepSpeed，具备千卡级多机多卡训练、混合并行、激活重计算、显存优化、训练吞吐分析与分布式问题排查经验。
- **高效注意力与内核适配**：具备 cross-attention mask、FA3/FlashAttention 适配经验，能够围绕长视频帧数扩展和训练效率瓶颈进行工程优化。
- **长视频 RAG 与评测**：熟悉 probe-based retrieval、结构化记忆、重排序、LLM-as-Judge、多模型评测与人工评估对齐，能够完成从实验设计到论文撰写的科研闭环。
- **工程实现**：Python、C/C++、Linux，熟悉分布式训练脚本、实验编排、日志分析和大规模训练任务故障定位。