

# 刘熠

+86-13006363023

yiliu61richard@gmail.com

www.yiliu.uno



## 基本信息

出生年月: 1995年08月

籍贯: 湖北黄冈

政治面貌: 中共党员

Base: 上海

目前在字节Commercial AI-AIGC团队负责广告原料Agent的建设, 包括数据构造、模型训练及业务落地。

研究方向包括: 多模态理解大模型、多模态数据合成、长视频理解、视频时序检测等。

## 教育经历

中国科学院大学	模式识别与智能系统	博士	2019.09 – 2024.07
深圳先进技术研究院	博士导师: 乔宇教授/王亚立研究员		北京/深圳
华中科技大学	自动化理工交叉创新实验班	本科	2015.09 – 2019.07
人工智能与自动化学院	本科导师: 肖阳副教授		湖北武汉

## 工作经历

字节跳动 中国交易与广告-Commercial AI-AIGC 原料Agent 负责人 2025.11 – 至今

AIGC算法工程师, 负责广告原料Agent的建设, 包括数据构造、模型训练及业务落地。

- 模型训练: 主导原料 Agent 模型建设, 基于大盘数据构建高质量商品原料片段, 增强模型在创意画面设计任务中的泛化能力; 面向 Seedance-2.0 构建原料生成强化学习训练数据, 持续提升创意模型生成质量。
- 业务落地: 负责干川原料生成链路的搭建与优化, 沉淀 Agent 通用 Skill, 提升多业务场景下的原料生成质量; 将微剧情生成能力迁移至干川业务场景, 持续打磨干川微剧情原料 Skill, 推动生成效果稳定迭代。

荣耀终端 大模型能力平台部 多模态理解与生成组组长 2024.04 – 2025.10

多模态理解与生成组组长, 实线管理 10 余人, 负责荣耀端侧 VLM 的技术预研以及业务落地。

- 多模态大模型: 主导端侧多模态大模型 MagicVL-2B 研发。相比同尺寸 InternVL2.5-2B 模型, 功耗降低 40%, 同时具备更强的通用多模态理解能力; 模型已落地于荣耀旗舰折叠屏手机 Magic V5。
- 多模态数据合成: 主导端侧多模态大模型的训练数据的清洗和构建。构建短视频结构化描述数据合成 Pipeline、屏幕知识 GUI 数据合成 Pipeline, 支撑视频描述、GUI Agent 业务。
- 视频理解大模型: 主导短视频细粒度描述大模型 VideoCap-R1 与长视频高效理解大模型 E-VRAG 研发, 覆盖短视频精细化描述与长视频高效理解场景。

上海人工智能实验室 通用技术研究部 见习研究员 (实习) 2022.03 – 2023.12

- 参与视频理解通用表征模型 InternVideo 的模型研发, 并参与时序定位等下游任务的对齐训练。
- 参与视频理解领域大模型评测基准构建, 包括短视频理解基准 MVBench 与长视频理解基准 LVBench。

## 论文成果

[1] MagicVL-2B: Empowering Vision-Language Models on Mobile Devices with Lightweight Visual Encoders via Curriculum Learning, arXiv, 2025 (Tech report, 第1作者)

[2] E-VRAG: Enhancing Long Video Understanding with Resource-Efficient Retrieval Augmented Generation, arXiv, 2025 (Tech report, 第1通讯)

[3] VideoCap-R1: Enhancing MLLMs for Video Captioning via Structured Thinking, arXiv, 2025 (under review, 共同通讯)

- [4] LvBench: A Benchmark for Long-form Video Understanding with Versatile Multi-modal Question Answering, International Journal of Computer Vision, 2025 (IJCV, 中科院1区, IF=9.3, 共一第3)
- [5] MLLM-TA: Leveraging Multimodal Large Language Models for Precise Temporal Video Grounding, IEEE Signal Processing Letters, 2024 (SPL, 中科院2区, IF=3.9, 第1作者)
- [6] MVBench: A Comprehensive Multi-modal Video Understanding Benchmark, Conference on Computer Vision and Pattern, 2024 (CVPR, CCF-A会议, 第6作者)
- [7] F2S-Net: Learning Frame-To-Segment Prediction for Online Action Detection, Journal of Real-Time Image Processing, 2024 (JRTIP, 中科院3区, IF=3.0, 第1作者)
- [8] Dual masked modeling for weakly-supervised temporal boundary discovery, IEEE Transactions on Multimedia, 2023 (TMM, 中科院1区, IF=9.7, 共一第2)
- [9] InternVideo: General Video Foundation Models via Generative and Discriminative Learning, arXiv, 2022 (SCIS, 第9作者)
- [10] Learning Discriminative Feature Representation for Open Set Action Recognition, ACM International Conference on Multimedia, 2023 (ACM MM, CCF-A会议, 共一第2)
- [11] FineAction: A Fine-Grained Video Dataset for Temporal Action Localization, IEEE Transactions on Image Processing, 2022 (TIP, 中科院1区, IF=13.7, 第1作者)
- [12] VideoPipe 2022 Challenge: Real-World Video Understanding for Urban Pipe Inspection, International Conference on Pattern Recognition, 2022 (ICPR, CCF-C会议, 第1作者)
- [13] 短视频场景在线起始检测任务及方法研究, 集成技术, 2021 (共一第2)

## 相关荣誉

---

### 【工作阶段】

- 2025年 荣耀终端-AI平台部-闪耀之星
- 2024年 荣耀终端-研发管理部总裁团队-AI多模态竞争力突破团队

### 【博士阶段】

- 2023年 中国科学院大学-三好学生标兵
- 2022年 ECCV Ego4D Episodic Memory Challenge (Looking at Me Track) 第1名
- 2022年 ECCV Ego4D Episodic Memory Challenge (Moment Queries Track) 第1名
- 2022年 中科院深圳先进技术研究院-院长优秀奖
- 2021年 中科院深圳先进技术研究院-院长优秀奖
- 2020年 中科院深圳先进技术研究院-院长优秀奖
- 2020年 中国科学院大学-三好学生
- 2019年 慧科杯人工智能应用创新挑战赛一等奖

### 【本科阶段】

- 2019年 华中科技大学-优秀毕业生
- 2018年 亚太大学生机器人大赛(ABU Robocon)-全国赛二等奖
- 2018年 亚太大学生机器人大赛(ABU Robocon)-南方赛区亚军
- 2018年 机械创新设计大赛-湖北省一等奖
- 2017年 挑战杯全国大学生课外学术科技作品竞赛-湖北省二等奖
- 2017年 全国大学生数学建模竞赛-湖北省二等奖
- 2017年 美国大学生数学建模竞赛-H奖