

Zhuohang Jiang

(+86) 152-6704-1093 | zhuohang.jiang@connect.polyu.hk | zhuohangjiang.com | jzzzzh | Zhuohang Jiang

个人简介

现为香港理工大学计算机科学博士研究生，导师为李青教授（Prof. Qing Li）与范文琦教授（Prof. Wenqi Fan）。主要研究方向包括大语言模型（LLMs）、检索增强生成（RAG）、推荐系统与多模态智能体，重点研究大模型在层次结构推理、多模态知识密集型问答及工业推荐系统中的应用。博士阶段累计 GPA 为 3.60/4.00。

2024 年获四川大学计算机科学与技术工学学士学位，专业 GPA 为 3.79/4.00，综合 GPA 为 3.78/4.00。本科期间于四川大学 MachineLab / DICALab 担任科研助理，师从周吉喆教授（Prof. JiZhe Zhou），参与国家自然科学基金及国家重点研发计划支持的科研项目。

研究概述

近期研究主要关注如何连接大模型的语义推理能力与实际智能系统。代表性方向包括：以 HiBench 为代表的大语言模型层次结构推理评测；以 QA-Dragon、SUPERGLASSES / SUPERLENS 为代表的多模态 RAG 与智能眼镜智能体；以及以 AIR 为代表的大模型增强工业跨域推荐。

早期研究主要集中于计算机视觉，包括图像篡改检测与定位和隐私敏感目标识别。相关成果发表于 KDD、CVPR、NeurIPS、AAAI、ICONIP 等国际会议及 Track，目前论文累计引用超过 500 次，h-index 为 7。

教育经历

香港理工大学 2024.09 – 至今
中国香港
计算机科学博士 · 导师：李青教授、范文琦教授

- 累计 GPA：3.60/4.00。
- 主要课程：Advanced Machine Learning (A)、Advanced Topics in Optimization (A-)、Advanced Topics in Visual Computing (A-)。
- 已完成 Research Seminar I & II、Practicum I、Engineering Ethics、Presentation Skills 等培养环节。

四川大学 2020.09 – 2024.06
成都
计算机科学与技术 · 工学学士

- 专业 GPA：3.79/4.00 (89.39/100)；综合 GPA：3.78/4.00 (89.25/100)。
- 2022–2024 年于 MachineLab / DICALab 担任科研助理。

代表性论文

- A Survey on AI Smart Glasses for Wearable Intelligence: From Egocentric Sensing to Agentic Personalization*
X. Yuan, Y. Wang, Z. Jiang, H. Qu, Y. Ding, S. Lin, G. Xing, H. Yang, J. Cao, Q. Li, W. Fan. **Preprints.org preprint**, 202608.0648, 2026.
- A/B Agent: A Self-Evolving Agent for Strategy Iteration in Industrial A/B Testing*
Z. Jiang, Y. Chen, Y. Pan, Z. Hu, W. Fan, Q. Li, H. Wang, J. Wang, W. Ou. **arXiv preprint**, arXiv:2608.04625, 2026.
- Atomic Intent Reasoning: Bringing LLM Semantics to Industrial Cross-Domain Recommendations*
Z. Jiang, Y. Chen, S. Wang, H. Qu, J. Zhou, W. Fan, Q. Li, D. Liang, J. Wang. **KDD 2026**, Applied Data Science Track. CCF-A.
- SUPERGLASSES: Benchmarking Vision Language Models as Intelligent Agents for AI Smart Glasses*
Z. Jiang, X. Yuan, H. Qu, S. Lin, K. Liu, W. Fan, Q. Li. **CVPR 2026 Findings**. CCF-A.
- QA-Dragon: Query-Aware Dynamic RAG System for Knowledge-Intensive Visual Question Answering*
Z. Jiang, P. Wu, X. Yuan, W. Fan, Q. Li. **KDD 2025 Workshop on Multimodal Retrieval Augmented Generation**. CCF-A.
- A Survey of WebAgents: Towards Next-Generation AI Agents for Web Automation with Large Foundation Models*
L. Ning, Z. Liang, Z. Jiang, H. Qu, Y. Ding, W. Fan, X.-Y. Wei, S. Lin, H. Liu, P. S. Yu, Q. Li. **KDD 2025 Tutorial Track**. CCF-A.
- HiBench: Benchmarking LLMs Capability on Hierarchical Structure Reasoning*
Z. Jiang, P. Wu, Z. Liang, P. Q. Chen, X. Yuan, Y. Jia, J. Tu, C. Li, P. H. F. Ng, Q. Li. **KDD 2025 Benchmark & Dataset Track**. CCF-A.
- Mesoscopic Insights: Orchestrating Multi-Scale & Hybrid Architecture for Image Manipulation Localization*
X. Zhu, X. Ma, L. Su, Z. Jiang, B. Du, X. Wang, Z. Lei, W. Feng, C.-M. Pun, J. Zhou. **AAAI 2025**. CCF-A.

- *IMDL-BenCo: A Comprehensive Benchmark and Codebase for Image Manipulation Detection & Localization*
X. Ma[†], X. Zhu[†], L. Su[†], B. Du[†], **Z. Jiang[†]**, B. Tong[†], Z. Lei[†], X. Yang[†], C.-M. Pun, J. Lv, J. Zhou. **NeurIPS 2024 Benchmark Track – Spotlight**. [†] Co-first authors. CCF-A.
- *Beyond Visual Appearances: Privacy-Sensitive Objects Identification via Hybrid Graph Reasoning*
Z. Jiang, B. Tong, X. Du, A. Alhammedi, J. Zhou. arXiv:2406.12736, 2024.
- *TPTGAN: Two-Path Transformer-Based Generative Adversarial Network Using Joint Magnitude Masking and Complex Spectral Mapping for Speech Enhancement*
Z. Liu, **Z. Jiang**, W. Luo, Z. Fan, H. Di, Y. Long, H. Wang. **ICONIP 2023**, Lecture Notes in Computer Science. CCF-C.
- *IML-ViT: Benchmarking Image Manipulation Localization by Vision Transformer*
X. Ma, B. Du, **Z. Jiang**, A. Y. Al Hammadi, J. Zhou. arXiv:2307.14863, 2023.
- *Perceptual MAE for Image Manipulation Localization: A High-level Vision Learner Focusing on Low-level Features*
X. Ma, **Z. Jiang**, X. Xu, C.-M. Pun, J. Zhou. arXiv:2310.06525, 2023.

科研与工业经历

快手科技 · 电商推荐中心
研究实习生

2025.10 – 至今
北京

- 研究面向短视频电商及内容-电商跨域场景的大语言模型增强推荐系统。
- 研发 AIR (Atomic Intent Reasoning)，将大模型语义推理迁移至离线阶段，通过在线检索与组合高效构建用户意图表示。
- 研究用户意图建模、用户画像、语义特征工程及转化导向推荐模型，并应用于工业级推荐排序系统。

四川大学 · MachineLab / DICALab
科研助理 · 导师：周吉喆教授

2022.09 – 2024.06
成都

- 研究图像篡改检测与定位、隐私敏感目标识别等计算机视觉问题。
- 参与国家自然科学基金及国家重点研发计划支持的科研项目。
- 相关科研成果发表于 NeurIPS、AAAI 等国际人工智能会议及 Benchmark Track。

新加坡国立大学 · School of Computing
暑期学校

2023.08
新加坡

- 参加计算机科学暑期研究项目，并完成基于 CNN 特征提取与相似度匹配的人脸识别系统。

代表性项目

AIR：面向工业跨域推荐的原子意图推理
KDD 2026 Applied Data Science Track · 第一作者

2025 – 2026

- 提出基于大语言模型的用户意图推理框架，用于连接短视频内容行为与电商推荐场景。
- 将大模型语义推理迁移至离线阶段，在线通过目标相关检索与组合构建用户意图表示，实现约 **400× 推理加速**。
- 应用于快手电商大规模推荐系统，并在真实线上 A/B 实验中实现 **GMV +3.446%**。

SUPERGLASSES / SUPERLENS：智能眼镜多模态智能体
CVPR 2026 Findings · 第一作者

2026

- 构建真实智能眼镜视觉问答 Benchmark，包含 **2,422 个**第一视角图像-问题对，覆盖 14 个领域与 8 类问题。
- 提出 SUPERLENS 多模态 RAG 智能体，集成目标检测、问题分解与多模态网络搜索能力。

HiBench：大语言模型层次结构推理 Benchmark
KDD 2025 Benchmark & Dataset Track · 第一作者 / 项目负责人

2025

- 构建大语言模型层次结构推理综合 Benchmark，覆盖 6 类场景、30 项任务和 **39,519 条**测试样例。
- 主导评测体系设计、核心代码实现、数据集与工具库开源以及多机构科研合作。

Meta CRAG-MM：多模态检索增强生成挑战赛

KDD Cup 2025 · 队长

2025

- 带领团队在数百支国际参赛队伍中获得**全球第 12 名**。
- 构建知识密集型视觉问答多模态 RAG 系统，设计领域路由、搜索路由与多跳推理模块。

IMDL-BenCo：图像篡改检测与定位 Benchmark

NeurIPS 2024 Benchmark Track – Spotlight · 共同第一作者

2024

- 构建图像篡改检测与定位综合 Benchmark 及模块化代码库。
- 实现 GPU 加速评测指标，并参与 Benchmark 设计、实验与论文撰写。

受邀报告

Understanding Hierarchical Data with Large Language Models

Reasoning Day @ KDD 2025 · Invited Talk

2025.08
Toronto, Canada

- 围绕大语言模型的层次数据理解、RAG、结构化推理及推理增强智能系统介绍相关研究进展与未来方向。

荣誉与奖励

KDD Cup 2025 Meta CRAG-MM 全球第 12 名	2025
多模态检索挑战赛	
NeurIPS 2024 Spotlight	2024
IMDL-BenCo · Benchmark Track · 共同第一作者	
四川省优秀毕业生、四川大学优秀毕业生	2024
中国大学生计算机设计大赛省级一等奖	2023
腾讯奖学金	2023
四川大学 Top 2%	
综合素质评价 A 级证书	2023
综合一等奖学金	2022
四川大学 Top 1%	
优秀学生	2022
四川大学 Top 5%	

学术服务与教学

国际会议与期刊审稿

Reviewer

2023 – 至今

- 担任 TIP、CVPR、ICCV、ECCV、NeurIPS、KDD、AAAI、IoTJ 等国际会议及期刊审稿人。

Frontiers in Artificial Intelligence / Frontiers in Big Data

Topic Coordinator

2025 – 2026

- 担任 Frontiers in Artificial Intelligence 与 Frontiers in Big Data 专题协调人。

- 协助 Artificial Intelligence (COMP4431)、NLP Practicum (COMP5423) 及 Database System (COMP2411) 等课程教学。

专业技能

- 研究方向** 大语言模型、检索增强生成、推荐系统、多模态智能体、跨域推荐、计算机视觉
- 框架与工具** PyTorch、Hugging Face、NumPy、OpenCV、Docker、Git、Anaconda、LaTeX
- 语言能力** 中文（母语）、英语（熟练，IELTS 6.5）