

# Hongrui Jia

✉ [jiahongrui@stu.pku.edu.cn](mailto:jiahongrui@stu.pku.edu.cn)

🐙 [GitHub](#)

🎓 [Google Scholar](#)

🌐 [hongruijia.github.io](http://hongruijia.github.io)

## Research Vision

My research focuses on **improving the reliability of multimodal reasoning in MLLMs and agents**. I study how to diagnose failure modes in multimodal reasoning, enforce evidence-grounded reasoning processes, and improve model capabilities through diagnostic-driven training. My goal is to **build multimodal AI systems that can reason reliably in complex open-world environments**.

## Education

- Jun, 2027**    **Peking University** Beijing, China  
(Expected)    *Master of Engineering in Software Engineering*  
**Advisors:** Prof. Wei Ye; Prof. Shikun Zhang  
**GPA:** 3.73/4.00    **Citations:** 186    **H-index:** 6 (May 2026)
- Jun, 2024**    **South China University of Technology** Guangzhou, China  
*Bachelor of Engineering in Software Engineering*  
**GPA:** 3.89/4.00

## Publications (Selected)

- [1] [From Blind Spots to Gains: Diagnostic-Driven Iterative Training for Large Multimodal Models](#)  
**Hongrui Jia**, Chaoya Jiang, Shikun Zhang, Wei Ye  
Under Review (ICML 2026) [🔗 Code](#)
- [2] [OSWorld-MCP: Benchmarking MCP Tool Invocation In Computer-Use Agents](#)  
**Hongrui Jia**, Jitong Liao, Xi Zhang, Haiyang Xu, Tianbao Xie, Chaoya Jiang, Ming Yan, Si Liu, Wei Ye, Fei Huang  
*International Conference on Learning Representations (ICLR 2026)* [🔗 Code](#)
- [3] [Decoupling Reasoning and Perception: An LLM-LMM Framework for Faithful Visual Reasoning](#)  
**Hongrui Jia**, Chaoya Jiang, Shikun Zhang, Wei Ye  
Preprint, 2025
- [4] [SymDPO: Boosting In-Context Learning of Large Multimodal Models with Symbol Demonstration Direct Preference Optimization](#)  
**Hongrui Jia**, Chaoya Jiang, Haiyang Xu, Wei Ye, Mengfan Dong, Ming Yan, Ji Zhang, Fei Huang, Shikun Zhang  
*Computer Vision and Pattern Recognition Conference (CVPR 2025)* [🔗 Code](#)
- [5] [MaVen: An Effective Multi-Granularity Hybrid Visual Encoding Framework for Multimodal Large Language Model](#)  
Chaoya Jiang\*, **Hongrui Jia\*** (\*Equal Contribution), Haiyang Xu, Wei Ye, Mengfan Dong, Ming Yan, Ji Zhang, Fei Huang, Shikun Zhang  
*Advances in Neural Information Processing Systems (NeurIPS 2024)* [🔗 Code](#)
- [6] [Hal-eval: A Universal and Fine-grained Hallucination Evaluation Framework for Large Vision Language Models](#)  
Chaoya Jiang\*, **Hongrui Jia\*** (\*Equal Contribution), Mengfan Dong, Wei Ye, Haiyang Xu, Ming Yan, Ji Zhang, Shikun Zhang  
*ACM International Conference on Multimedia (ACM MM 2024)*, **Oral**

[7] [ISC4DGF: Enhancing Directed Grey-box Fuzzing with LLM-Driven Initial Seed Corpus Generation](#)

Yijiang Xu, **Hongrui Jia**, Liguo Chen, Xin Wang, Zhengran Zeng, Yidong Wang, Qing Gao, Jindong Wang, Wei Ye, Shikun Zhang, Zhonghai Wu

*Journal of computer science and technology* (JCST)

## Honors (Selected)

Oct, 2025	National Scholarship	Ministry of Education, China
Oct, 2021	National Scholarship	Ministry of Education, China

## Industrial Experience

### Tongyi Lab, Alibaba Group

**Beijing**

*Research Intern · Mentored by: Haiyang Xu, Ming Yan*

*June 2025 – Jan 2026*

- Led research on enhancing the tool-use capabilities of computer-use agents.
- Contributed to the Mobile-Agent project, focusing on mitigating the reward sparsity issue of Multimodal Large Language Models in computer-use scenarios.

### JD EXPLORE ACADEMY, Jingdong Group

**Beijing**

*Research Intern · Mentored by: Zenan Zhou*

*Feb 2026 – Present*

- Led research on improving spatial reasoning and tool-use abilities of Multimodal Large Language Models, particularly in Vision-and-Language Navigation (VLN) scenarios.

## Academic Service

**Conference Reviewer:** CVPR 2026, ICLR 2026, EMNLP 2025, ACL 2025, NAACL 2025

**Open-source Service:** Main contributor to OSWorld-MCP Framework, DPE Framework and open-source VLM Training Methods SymDPO.