

Ruiyang Zhang

☎ (+853) 63680758 | ✉ yc47931@um.edu.mo | 🌐 ruiyang-061x.github.io
🔗 Ruiyang-061X | 🎓 Ruiyang Zhang

Introduction

Ruiyang Zhang is a Ph.D. candidate at the University of Macau. He receives his B.S. in Computer Science and Technology from Fudan University.

His research focuses on enhancing the **reasoning** and **tool use capabilities** of **multimodal agents** through reinforcement learning. He previously worked on exploring the uncertainty of large multimodal models and on 3D object detection in unsupervised settings.

Education

University of Macau

Ph.D. in Faculty of Science and Technology
Advisor: **Prof. Zhedong Zheng**

Aug. 2024 -
Macau SAR

Fudan University

B.S. in Computer Science and Technology

Sept. 2017 - Jun. 2021
Shanghai, China

Research Experience

Research Intern @ IDEA

Mentor: **Dr. Yiyang Qi**

Octo. 2025 -
Shenzhen, China

- Topics: Developing long-horizon multimodal search agent via reinforcement learning (RL)
- Responsibilities: (1) Build the RL infrastructure required to develop multi-turn tool-using multimodal search agent, including tool implementation, caching mechanism, and RL framework adaptation. (2) Build long-horizon, multi-turn multimodal search agent in real-world web environment via reinforcement learning. Propose a systematic post-training framework for multimodal search agents, including Iterative Injection-based Data Synthesis, Rejection Sampling Fine-tuning, and RL.

Research Intern @ Shanghai AI Lab

Mentor: **Dr. Dongzhan Zhou**

May 2025 - Sept. 2025
Shanghai, China

- Topics: Leveraging RL to incentivize sketch-style reasoning in large multimodal models
- Project: **SketchThinker-R1 (ICLR'26)**
- Responsibilities: (i) Devise Sketch-Mode Cold Start to instill initial sketch-style reasoning into base multimodal model; train SketchJudge reward model to pose supervision on thinking style during the RL stage; conduct Sketch-Thinking Reinforcement Learning to further generalize sketch-style reasoning. (ii) Evaluation on benchmarks across various domains, showing that SketchThinker-R1 achieves over 64% reduction in thinking cost without compromising answer accuracy.

Selected Award

- Second Prize, China Undergraduate Mathematical Contest in Modeling (CUMCM) 2019
- Third Class Scholarship for Outstanding Students, Fudan University 2018 & 2019
- First Prize, Chinese Mathematical Olympiad in Jiangsu Province (CMO) 2016
- First Prize, National Olympiad in Informatics in Jiangsu Province (NOIP) 2014 & 2015

Publication

1. [ICLR 2026] **SketchThinker-R1: Towards Efficient Sketch-Style Reasoning in Large Multimodal Models**
Ruiyang Zhang*, Dongzhan Zhou*, Zhedong Zheng
[\[Paper\]](#) | [\[Code\]](#)
- Incentivize sketch-style reasoning in large multimodal models to improve reasoning efficiency.
2. [ICCV 2025] **Harnessing Uncertainty-aware Bounding Boxes for Unsupervised 3D Object Detection**
Ruiyang Zhang, Hu Zhang, Hang Yu, Zhedong Zheng
[\[Paper\]](#) | [\[Code\]](#)
- Aim to mitigate the negative impact of noisy pseudo boxes.
3. [ECCV 2024] **Approaching Outside: Scaling Unsupervised 3D Object Detection from 2D Scene** [\[40+ Github Stars\]](#)
Ruiyang Zhang, Hu Zhang, Hang Yu, Zhedong Zheng
[\[Paper\]](#) | [\[Code\]](#)
- Enhance detection of far and small objects via LiDAR and 2D data fusion.

Preprint

1. [Preprint] **Uncertainty-o: One Model-agnostic Framework for Unveiling Epistemic Uncertainty in Large Multimodal Models**
Ruiyang Zhang, Hu Zhang, Hao Fei, Zhedong Zheng
[\[Paper\]](#) | [\[Code\]](#) | [\[Website\]](#)
- Applicable to both large comprehension model and large generation model.
2. [Preprint] **VL-Uncertainty: Detecting Hallucination in Large Vision-Language Model via Uncertainty Estimation** [\[30+ Citations\]](#)
Ruiyang Zhang, Hu Zhang, Zhedong Zheng
[\[Paper\]](#) | [\[Code\]](#) | [\[Website\]](#)
- Preliminary attempt for understanding uncertainty in LVLM.

Skill

Programming: Python (Pytorch), C/C++, Java, MATLAB

Language: English (IELTS 7.0/9.0), Chinese (Native)

Work Experience

Meituan

Backend Development Engineer

Jul. 2021 - Jun. 2023

Shanghai, China

Open Source

1. [Awesome List] **Awesome-MLLM-Uncertainty** [\[50+ Github Stars\]](#) [\[Code\]](#)
- Curated list of papers on the uncertainty in multi-modal large language model (MLLM).
2. [Awesome List] **Awesome-MLLM-Reasoning** [\[Code\]](#)
- Curated list about reasoning ability of MLLM, including OpenAI o1, OpenAI o3-mini, and Slow-Thinking.