

Mingyang Song

Ph.D. candidate (Year 3), Fudan University
mysong23@m.fudan.edu.cn — (+86)15246834828 — Homepage — Google Scholar

BIOGRAPHY

Mingyang Song is currently pursuing his Ph.D. at Fudan University, advised by Professor Yu Cheng. His research interests include **Visual Agents**, aiming to improve models' perception and reasoning capabilities through the invocation of external agents, and **Multi-step Reasoning**, focusing on evaluating and enhancing step-level reasoning accuracy in LLMs and VLMs.

EDUCATION

Fudan University, Shanghai, China Aug. 2023 — Present
Ph.D. Candidate
Research field: Artificial Intelligence

Harbin Institute of Technology, Harbin, China Aug. 2019 — Jun. 2023
B.Eng. in Computer Science and Technology Percentage Grade: 96.2/100 (ranked 3/92)

ACADEMIC EXPERIENCE

Shanghai Artificial Intelligence Laboratory Shanghai, China
Research Intern Feb. 2023 — Aug. 2025

- **LVLMM Hallucination**: Focused on investigating and mitigating hallucination within Large Vision and Language Models (LVLMMs) to improve model reliability and performance through data balancing and targeted mitigation techniques.
- **Long Sequence Reasoning**: Towards building a comprehensive multi-step reasoning evaluation framework to thoroughly assess the reasoning capabilities of LLMs, and subsequently enhance the reasoning capabilities of LLMs.

The Chinese University of Hong Kong Hong Kong SAR
Research Assistant Sep. 2025 — Present

- **Visual Agents**: Focus on enhancing the perception and reasoning capabilities of LVLMMs through supervised fine-tuning and reinforcement learning on visual agent trajectories.

PUBLICATIONS

AdaReasoner: Dynamic Tool Orchestration for Iterative Visual Reasoning ICLR 2026

- Proposed a unified framework for data curation, reinforcement learning, and evaluation of tool-planning models.
- Our model develops emergent, self-adaptive behaviors, learning to autonomously adopt beneficial tools, discard irrelevant ones, and modulate its usage frequency, achieving substantial improvements (+38.7% on VSP tasks) over baselines. [Paper][Project Page][Code]
- Authors: **Mingyang Song**^{*}, Haoyu Sun^{*}, Jiawei Gu^{*}, Linjie Li^{*}, Ranjay Krishna, Yu Cheng

PRMBench: A Fine-grained and Challenging Benchmark for Process-Level Reward Models ACL 2025 Main

- Introducing the first comprehensive process-level reward model benchmark, comprising 6,216 carefully curated samples and 83,456 step-level labels for a series of evaluations on PRMs. Through comprehensive evaluation and analysis, we reveal the partial availability and the tendency to false positives in current PRMs. [Paper][Project Page][Code]
- Authors: **Mingyang Song**, Zhaochen Su, Xiaoye Qu, Jiawei Zhou, Yu Cheng

From Head to Tail: Towards Balanced Representation in Large Vision-Language Models through Adaptive Data Calibration CVPR 2025

- Conducted the first in-depth analysis of the long-tail problem within the training data of current LVLMMs from four key aspects, including tokens, objects, co-occurrences, and interrogations. Subsequently, the ADR framework was designed and introduced to mitigate the long-tail problem through a fine-grained strategy. [Paper][Project Page][Code]
- Authors: **Mingyang Song**, Xiaoye Qu, Jiawei Zhou, Yu Cheng

Mitigating Multilingual Hallucination in Large Vision-Language Models TOMM

- Developed the first approach to mitigate multilingual hallucinations in LVLMMs. Our work reveals the severe multilingual hallucination problem when querying the existing LVLMMs using non-English languages. Meanwhile, we analyze the factors of causing multilingual hallucinations in LVLMMs. [Paper][Code]

- Authors: Xiaoye Qu*, Mingyang Song*, Wei Wei, Jianfeng Dong, Yu Cheng

GenReasoner: Generative Visual Agent Planning

Nature Portfolio Submission

- We investigate the generalization ability of visual agents, examining whether tool-planning capabilities learned under specific task–tool combinations can generalize to unseen tasks.
- We present a generalizable transfer learning method, and our latest experiments reveal that tool-calling capabilities acquired from one task can robustly generalize to other tasks.

AWARDS

Finalist of The Interdisciplinary Contest in Modeling, COMAP

2022

First-Class People's Scholarship

2021

National Encouragement Scholarship

2020,2021,2022

BD scholarship

2025