

Kaixuan Fan

📍 Hong Kong ✉ kxfan127@gmail.com ☎ 15653290228

Education

The Chinese University of Hong Kong

First year Ph.D. student of MMLab, Department of Information Engineering

Aug.2025 - June

2029(Expected)

- **Supervisor: Prof. Xiangyu Yue**
- Recipient of the **Hong Kong PhD Fellowship Scheme**.

Fudan University

Bachelor of Science | Computer Science and Technology

Sept. 2021 – June 2025

- **GPA: 3.86/4.0**
- **Rank: 1/101 (Department of Computer Science) 1/195(School of Computer Science)**
- Recipient of the **National Scholarship** (2022,2024)

Honors and Awards

Hong Kong PhD Fellowship Scheme (HKPFS)

2025

China National Scholarship

2024, 2022

Scholarship for Outstanding Students at Fudan University

2023

Meritorious Winner of Mathematical Contest In Modeling

2024

Publication

SophiaVL-R1: Reinforcing MLLMs Reasoning with Thinking Reward. Kaixuan Fan^{*}, Kaituo Feng^{*}, Haoming Lyu, Dongzhan Zhou, Xiangyu Yue. *Accepted By the Fourteenth International Conference on Learning Representations (ICLR) 2026.*

Exploring Reasoning Reward Model for Agents. Kaixuan Fan, Kaituo Feng, Manyuan Zhang, Tianshuo Peng, Zhixun Li, Yilei Jiang, Shuang Chen, Peng Pei, Xunliang Cai, Xiangyu Yue. *Accepted By the 64th Annual Meeting of the Association for Computational Linguistics (ACL findings) 2026.*

Onethinker: All-in-one reasoning model for image and video. Kaituo Feng, Manyuan Zhang, Hongyu Li, Kaixuan Fan, Shuang Chen, Yilei Jiang, Dian Zheng, Peiwen Sun, Yiyuan Zhang, Haoze Sun, Yan Feng, Peng Pei, Xunliang Cai, Xiangyu Yue. *Accepted By The IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) 2026.*

Diffusion Autoencoders for Few-shot Image Generation in Hyperbolic Space. Lingxiao Li, Kaixuan Fan, Boqing Gong, Xiangyu Yue. *Accepted By the International Conference on Computer Vision (ICCV) 2025.*

Research Experience

Research Intern

Remote

Meituan Inc. Longcat group

Sept. 2025 – April 2026

- We study the limitation of sparse outcome-based rewards in agentic RL and propose Agent-RRM, a reasoning reward model that provides structured feedback for agent trajectories, including explicit reasoning traces, targeted critiques, and holistic quality scores.
- We systematically investigate three ways of integrating reasoning rewards into agent training, namely Reagent-C (text-augmented refinement), Reagent-R (reward-augmented guidance), and Reagent-U (unified feedback integration).
- We curate four specialized datasets for training both reasoning agents and reward models, covering mathematical reasoning, multimodal understanding, web-based information seeking, and complex tool use.
- Experiments across 12 diverse benchmarks show that Reagent-U consistently outperforms strong baselines, achieving 43.7% on GAIA and 46.2% on WebWalkerQA, demonstrating the effectiveness of combining scalar rewards with textual critiques in agentic RL.
- Accepted as Findings of the 64th Annual Meeting of the Association for Computational Linguistics (**ACL findings**) 2026

Research Intern

Shanghai Artificial Intelligence Lab

Shanghai

May. 2025 – June 2025

- We focus on the “wrong thinking, correct answer” phenomenon in current reasoning models and propose SophiaVL-R1, an MLLM that enhances reasoning by integrating model-generated thinking rewards with rule-based outcome rewards in RL training.
- We propose the Trust-GRPO training algorithm to mitigate the impact of potentially unreliable thinking rewards during RL training.
- Experiments show that our SophiaVL-R1 surpasses a series of reasoning MLLMs on various benchmarks (e.g., MathVisita, MMMU), even outperforms LLaVA-OneVision-72B with only 7B parameters.
- Accepted by the Fourteenth International Conference on Learning Representations (**ICLR** 2026).

Research Intern

Shanghai Artificial Intelligence Lab

Shanghai

Nov. 2024 – Feb. 2025

- Constructed and processed the International Chemistry Olympiad (ICHO) dataset for benchmarking the complex chemistry problem-solving capabilities of multimodal large language models.
- Utilized MinerU and Mathpix to extract text, formulas, and images from collected PDF files, and developed scripts to standardize the dataset format.
- Designed a pipeline to split lengthy ICHO problems with multiple sub-parts into independent sub-questions and categorize them by type.
- Developed tailored evaluation methods for each question type to more accurately assess model performance.

Research Intern

MMLab, CUHK

Hong Kong

Sept. 2024 – Feb. 2025

- Worked as an intern on few-shot image generation based on the Diffusion Model, supervised by Prof. Xiangyu Yue.
- Proposed a novel network *HypDAE*, which leveraged hyperbolic space to capture hierarchical relationships among images and textual descriptions, improving the diversity and quality of generated images in low-data settings.
- The hyperbolic representation introduced an additional degree of control over semantic diversity through the adjustment of radii within the hyperbolic disk.
- Accepted by the International Conference on Computer Vision (**ICCV**) 2025.

Internship

Longcat group intern at Meituan Inc.

Sept. 2025 - April. 2026

- Mentor: Manyuan Zhang

AI for Science intern at Shanghai Artificial Intelligence Lab

Oct. 2024 - June 2025

- Mentor: Dongzhan Zhou

Skills

English: IELTS (Overall 7.5, Writing 7.0, Speaking 7.0, Listening 8.0, Reading 8.5), CET-6 (666)

Programming: Python (NumPy, Pytorch, Matplotlib), C, C++, System Verilog, LaTeX