

JIA, Mengzhao (she/her)

355 Fitzpatrick Hall, Notre Dame, IN 46556

jiamengzhao98@gmail.com

Education

Ph.D. in Computer Science and Engineering 08/2023 –
University of Notre Dame, USA. Advisor: Prof. Meng Jiang

M. S. in Computer Science and Engineering 09/2020 – 06/2023
Shandong University, China. Advisor: Prof. Liqiang Nie

B. Eng in Electronic Science and Technology 09/2016 – 06/2020
Shandong University, China.

Work Experience

Research Intern at Orby AI, Mountain View, US. *Worked on Pub.*[1] 03/2025 – 09/2025

Research Intern at Tencent AI Lab, Seattle, US. *Worked on Pub.*[3] 05/2024 – 09/2024

Research Intern at Alibaba DAMO Academy, Hangzhou, China. *Worked on Pub.*[6] 06/2022 – 01/2023

Research Intern at Kuaishou Technology, Beijing, China. *Worked on Pub.*[5] 04/2021 – 01/2022

Research Interest: Multimodal Reasoning. Multimodal Large Language Models.

Google Scholar: <https://scholar.google.com/citations?user=E332upAAAAAJ>

Citations: 508; h-index: 12; i10-index: 13 (07/2026)

Recent Work

1. **Mengzhao Jia**, Zhihan Zhang, Ignacio Cases, Zheyuan Liu, Meng Jiang, Peng Qi. *AutoRubric: Rubric-Based Generative Rewards for Faithful Multimodal Reasoning*. **ACL 2026 Findings** [ACL] [GitHub] *TL;DR:* AutoRubric uses automatically generated rubrics in multimodal reinforcement learning to improve both accuracy and reasoning faithfulness.
2. **Mengzhao Jia**, Zhihan Zhang, Meng Jiang. *Prioritizing the Best: Incentivizing Reliable Multimodal Reasoning by Rewarding Beyond Answer Correctness*. arXiv 2026. [Arxiv] [Google Scholar] *TL;DR:* Groupwise Ranking Reward ranks verifier-passed trajectories to improve the reliability of multimodal reasoning beyond final-answer correctness.
3. **Mengzhao Jia**, Wenhao Yu, Kaixin Ma, Tianqing Fang, Zhihan Zhang, Siru Ouyang, Hongming Zhang, Dong Yu, Meng Jiang. *Leopard: A Vision Language Model for Text-Rich Multi-Image Tasks*. **TMLR 2025**. [Arxiv] [GitHub] [HuggingFace] *TL;DR:* Leopard outperforms the best multimodal large language models by an average of +9.61 points on 5 text-rich, multi-image benchmarks. The constructed multi-image reasoning dataset got 200k+ downloads.
4. Zifeng Zhu*, **Mengzhao Jia***, Zhihan Zhang, Lang Li, Meng Jiang. *MultiChartQA: Benchmarking Vision-Language Models on Multi-Chart Problems*. **NAACL 2025**. [ACL Anthology] [HuggingFace] *TL;DR:* MultiChartQA evaluates how MLLMs perform multi-hop, comparative, and sequential reasoning across multiple charts.

Peer-Reviewed Publications

5. **Jia**, Wei, Song, Sun, Zhang, Nie. *Query-Oriented Micro-Video Summarization*. IEEE **TPAMI**, 46 (6), 4174–4187, 2024.
6. **Jia**, Xie, Jing. *Debiasing Multimodal Sarcasm Detection with Contrastive Learning*. **AAAI** 2024.
7. **Jia**, Chen, Jing, Fu, Li. *Knowledge-Enhanced Memory Model for Emotional Support Conversation*. Machine Learning for Cognitive and Mental Health Workshop, **AAAI** 2024.
8. Yang, Guo, **Jia**, Su, Liu, Zhang, Jiang. *MMTutorBench: The First Multimodal Benchmark for AI Math Tutoring*. **ACL** 2026.
9. Liu, Dou, **Jia**, Tan, Zeng, Yuan, Jiang. *Protecting Privacy in Multimodal Large Language Models with MLLMU-Bench*. **NAACL** 2025.
10. Ouyang, Yu, Ma, Xiao, Zhang, **Jia**, et al. *RepoGraph: Enhancing AI Software Engineering with Repository-Level Code Graph*. **ICLR** 2025.
11. Zhang, Ge, Liang, Yu, Yu, **Jia**, Yu, Jiang. *Learn Beyond The Answer: Training Language Models with Reflection for Mathematical Reasoning*. **EMNLP** 2024.
12. Zhang, Lee, Fang, Yu, **Jia**, Jiang, Barbieri. *PLUG: Leveraging Pivot Language in Cross-Lingual Instruction Tuning*. **ACL** 2024.
13. Jing, Song, Ouyang, **Jia**, Nie. *Multi-source Semantic Graph-based Multimodal Sarcasm Explanation Generation*. **ACL** 2023.
14. Nie, **Jia**, Song, Wu, Cheng, Gu. *Multimodal Activation: Awakening Dialog Robots without Wake Words*. **SIGIR** 2021.

Service

Reviewer for ACL, ICLR, CVPR, AAAI