Jiwen Zhang

+86-18251180735 | jiwenzhang21@m.fudan.edu.cn

 IMNearth | G Jiwen Zhang

EDUCATION

- **School of Data Science, Fudan University** *Sep. 2021 - Now*
  PhD in Statistics, Supervisor: Prof. Zhongyu Wei Shanghai, China
  ◦ GPA: 3.83/4.00

- **School of Data Science, Fudan University** *Sep. 2016 - Jun. 2021*
  Bachelor of data science Shanghai, China
  ◦ GPA: 3.72/4.00   Rank: 03/62

- **National University of Singapore** *Jul. 2019 - Dec. 2019*
  Exchange student, supervised by Dr. Xiang Wang Singapore

RESEARCH INTERESTS

Large Multimodal Models, Embodied Agents, Reinforcement Learning

PREPRINTS

[1] Ding, Xuanwen*, Chengjun Pan*, Zejun Li*, **Jiwen Zhang***, Siyuan Wang, and Zhongyu Wei. *"AutoJudger: An Agent-Driven Framework for Efficient Benchmarking of MLLMs."* arXiv preprint arXiv:2505.21389 (2025).

[2] Li, Zejun*, **Jiwen Zhang***, Dianyi Wang*, Ye Wang*, Xuanjing Huang, and Zhongyu Wei. *"Continuous or Discrete, That Is the Question: A Survey on Large Multi-Modal Models from the Perspective of Input-Output Space Extension."* Preprints, manuscript/202411.0685 (2024).

[3] Yu, Ya-Qi, Minghui Liao, **Jiwen Zhang**, and Jihao Wu. *"Texthawk2: A large vision-language model excels in bilingual ocr and grounding with 16x fewer tokens."* arXiv preprint arXiv:2410.05261 (2024).

[4] **Zhang, Jiwen***, Yaqi Yu*, Minghui Liao, Wentao Li, Jihao Wu, and Zhongyu Wei. *"Ui-hawk: Unleashing the screen stream understanding for gui agents."* Preprints, manuscript/202408.2137 (2024).

[5] Luo, Ruipu, **Jiwen Zhang**, and Zhongyu Wei. *"Breaking Down the Task: A Unit-Grained Hybrid Training Framework for Vision and Language Decision Making."* arXiv preprint arXiv:2307.08016 (2023).

PUBLICATIONS

[1] Li, Zejun*, Ruipu Luo*, **Jiwen Zhang**, Minghui Qiu, and Zhongyu Wei. *"VoCoT: Unleashing Visually Grounded Multi-Step Reasoning in Large Multi-Modal Models."* In Proceedings of the 2025 Conference of the North American Chapter of the Association for Computational Linguistics, 2025.

[2] **Zhang, Jiwen**, Jihao Wu, Yihua Teng, Minghui Liao, Nuo Xu, Xiao Xiao, Zhongyu Wei, and Duyu Tang. *"Android in the zoo: Chain-of-action-thought for gui agents."* In Findings of the Association for Computational Linguistics: EMNLP 2024, pages 12016–12031, Miami, Florida, USA. Association for Computational Linguistics.

[3] Li, Zejun*, Ye Wang*, Mengfei Du*, Qingwen Liu*, Binhao Wu*, **Jiwen Zhang***, Chengxing Zhou et al. *"Reform-eval: Evaluating large vision language models via unified re-formulation of task-oriented benchmarks."* In Proceedings of the 32nd ACM International Conference on Multimedia, pp. 1971-1980. 2024.

[4] Du, Mengfei, Binhao Wu, **Jiwen Zhang**, Zhihao Fan, Zejun Li, Ruipu Luo, Xuan-Jing Huang, and Zhongyu Wei. *"DELAN: Dual-Level Alignment for Vision-and-Language Navigation by Cross-Modal Contrastive Learning."* In Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024), pp. 4605-4616. 2024.

[5] **Zhang, Jiwen**, Jianqing Fan, Jiajie Peng and Zhongyu Wei. *"Curriculum learning for vision-and-language navigation."* Advances in Neural Information Processing Systems 34 (2021): 13328-13339.

HONORS AND AWARDS

- Sep. 2022: National Graduate Scholarship for the 2021-2022 academic year

- Oct. 2019: Fudan University Excellent Undergraduate First-Class Scholarship for the 2018-2019 academic year

- Aug. 2019: TF LEaRN Scholarship, National University of Singapore

- May. 2017: Fudan University Longsheng-Yu Tongyin Scholarship (Named Scholarship)