Ao Zhang

Homepage: waxnkw.github.io
Github: github.com/waxnkw

Mobile: +65 80324197 / +86 18851831150
Email: aozhang@u.nus.edu

Experience

- NVIDIA Research Internship — Jul. 2024 – Now
  Research intern mentored by Dr. Jim Fan.
- National University of Singapore — Aug. 2021 – Now
  Ph.D. Candidate in Computer Science, supervised by Professor Chua Tat-Seng.
- Tsinghua University — Sep. 2020 – Jul. 2021
  Research Assistant supervised by Professor Zhiyuan Liu.
- Nanjing University — Sep. 2016 – Jun. 2020
  Bachelor in Software Engineering. GPA: 4.47/5.0 (top 10%). Top 1% Thesis.

Research Interest

My research interest mainly lies on multimodal understanding especially for MLLMs and VLA models.

Selected Publications & Preprints

Until April 15th, 2025, my Google Scholar citation has reached 2,008.

- <u>One of Core Contributors</u>. GR00T N1: An Open Foundation Model for Generalist Humanoid Robots. GTC 2025 highlighted.
- Yuan Yao*, Tianyu Yu*, <u>Ao Zhang*</u>, Chongyi Wang*, Junbo Cui*, Hongji Zhu*, Tianchi Cai*, Haoyu Li, Weilin Zhao, Ronghua Zhou, Zhihui He, Zhensheng Zou, Haoye Zhang, Shengding Hu, Zhi Zheng, Jie Zhou, Jie Cai, Xu Han, Guoyang Zeng, Dahai Li, Zhiyuan Liu, and Maosong Sun. MiniCPM-V: A GPT-4V Level MLLM on Your Phone GitHub 19k star, once 1st place in HuggingFace model trending.
- <u>Ao Zhang</u>, Yuan Yao, Wei Ji, Zhiyuan Liu, and Tat-Seng Chua. NExT-Chat: An LMM for Chat, Detection and Segmentation. ICML 2024.
- <u>Ao Zhang</u>, Hao Fei, Yuan Yao, Wei Ji, Li Li, Zhiyuan Liu, and Tat-Seng Chua. VPGTrans: Transfer Visual Prompt Generator across LLMs. NeurIPS 2023.
- <u>Ao Zhang</u>, Yuan Yao, Qianyu Chen, Wei Ji, Zhiyuan Liu, Maosong Sun, and Tat-Seng Chua. Fine-Grained Scene Graph Generation with Data Transfer. ECCV 2022 Oral.
- Yuan Yao*, <u>Ao Zhang*</u>, Qianyu Chen, Wei Ji, Zhiyuan Liu, Maosong Sun, and Tat-Seng Chua. CPT: Colorful prompt tuning for pre-trained vision-language models. Arxiv 2021.
- Yuan Yao*, <u>Ao Zhang*</u>, Xu Han, Mengdi Li, Cornelius Weber, Zhiyuan Liu, Stefan Wermter, and Maosong Sun. Visual Distant Supervision for Scene Graph Generation ICCV 2021.
- Yuan Yao*, Qianyu Chen*, <u>Ao Zhang</u>, Wei Ji, Zhiyuan Liu, Tat-Seng Chua, and Maosong Sun. PEVL: Position-enhanced Pre-training and Prompt Tuning for Vision-language Models. EMNLP 2022.

  (* denotes Equal Contribution or Core Contributor.)

Professional Service

- Conference Reviewer: CVPR, NeurIPS, ICLR, EMNLP, WWW
- Journal Reviewer: IJCV, TBD

Honors and Awards

- NUS Research Achievement Award — National University of Singapore, 2023
- NUS Research Scholarship — National University of Singapore, 2021
- First Prize of Peoples Scholarship — Nanjing University, 2019