Guanyu Zhou

Homepage
GitHub
guanyuzhou.ai@gmail.com
+86 18066288008

Education & Honors

B.Sc. in Data Science and Big Data Technology | GPA: 87.62/100 Sep. 2022 – Present
Wuhan University of Technology (WHUT) *Wuhan, China*

Core Courses (Total score: 100): Foundations of C Language Programming (99.76), Machine Learning and Data Mining Experiment (92.35), Comprehensive Experiments of Foundation of Computer and C Language Programming (100), Computer Organization and Architecture Experiment (95.78), CUDA High Performance Scientific Computing (92), High Level Language Programming (93.25), Experiments on Data Structure and Algorithm (95)

Teaching Assistant

- Business English Course Sep. 2023 – Dec. 2023

Selected Honors & Awards

- Outstanding Student Leader Award (Top 1 out of 75) Sep. 2023
- Bronze Medal, China Collegiate Programming Contest (CCPC) Hubei Provincial Competition Nov. 2023

Publications

- **Guanyu Zhou**, Yibo Yan, Xin Zou, Kun Wang, Aiwei Liu, and Xuming Hu. "Mitigating modality prior-induced hallucinations in multimodal large language models via deciphering attention causality." **International Conference on Learning Representations (ICLR)**, 2025.
- Junzhe Chen, Tianshu Zhang, Shiyu Huang, Yuwei Niu, Chao Sun, Rongzhou Zhang, **Guanyu Zhou**, Lijie Wen, and Xuming Hu. "OmniDPO: A preference optimization framework to address omni-modal hallucination." **Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)**, 2026.
- **Guanyu Zhou**, Wenxuan Liu, Wenxin Huang, Xuemei Jia, Xian Zhong, and Chia-Wen Lin. "OccludeNet: A causal journey into mixed-view actor-centric video action recognition under occlusions."arXiv:2411.15729, 2024.
- Mingzhe Zheng, Dingjie Song, **Guanyu Zhou**, Jun You, Jiahao Zhan, Xuran Ma, Xinyuan Song, Ser-Nam Lim, Qifeng Chen, and Harry Yang. "CML-Bench: A framework for evaluating and enhancing LLM-powered movie scripts generation.", **submitted to ICLR 2026**.
- **Guanyu Zhou**, Yonghua Hei, Yibo Yan, Xin Zou, Junzhe Chen, and Xuming Hu. "ManiLens: Visual perception in multimodal large language models via intrinsic dimension analysis.", **submitted to CVPR 2026**.

Research Experience

Data-centric Approach for VLM Weakness | Project Lead (Remote) | Princeton University Apr. 2025 – Present
Advisor: Zhuang Liu, Assistant Professor of Computer Science at Princeton University

- Address the common weakness of Vision-Language Models (VLMs, i.e., MLLMs) in vision-centric tasks through data-centric approaches.
- Validate the effectiveness of traditional deep learning within the existing VLM SFT framework.
- A comprehensive evaluation of VLM training data, models, and training methods.
- Explore the impact of CLIP-Blind pairs in the MMVP on text-to-image models that also utilize the CLIP encoder.
- Prepare first-author submission to the International Conference on Machine Learning (ICML) 2026.

Mitigating MLLM Hallucinations | Project Lead | AI Thrust, HKUST(GZ) Jul. 2024 – Mar. 2025
Advisor: Xuming Hu, Assistant Professor at The Hong Kong University of Science and Technology

- Developed CausalMM, a novel causal-inference framework (structural causal model + counterfactual reasoning) to treat visual and language modality priors as confounders, thereby isolating the true causal effect of attention mechanisms in multimodal large language models (MLLMs).

- Applied interventions at both visual-attention and language-attention levels (e.g. randomized / shuffled / reversed attention) to generate counterfactual outputs, enabling back-door adjustment and substantially reducing prior-induced bias in model decoding process.
- Demonstrated effectiveness across multiple benchmarks: up to 65.3% improvement on 6 VLind-Bench metrics, 164 points on MME Benchmark, and consistent gains on POPE tasks — as a plug-and-play, training-free method.

Action Recognition under Occlusion | Project Lead | XIAN Group, WHUT Jul. 2023 – Jun. 2024

Advisor: Xian Zhong, Professor at the School of CS and AI at Wuhan University of Technology

- Constructed OccludeNet, a large-scale mixed-view video dataset for occlusion-aware human action recognition, covering 424 action classes and a variety of realistic occlusion types (dynamic tracking, static scene, interactive, and multi-view) using both real and synthetic data.
- Proposed a novel Causal Action Recognition (CAR) framework based on a structural causal model of occluded scenes, treating occlusion as a confounder, and applied back-door adjustment plus counterfactual intervention to isolate the causal effect of visible actor features on predictions — thereby significantly improving robustness under occlusion.
- Demonstrated that occlusion impacts action classes unevenly, such as actions with low scene relevance or partial body visibility suffering larger accuracy drops, highlighting the need for class-dependent modeling.

Services

Peer Review & Academic Service

- Reviewer, *IEEE International Conference on Computer Vision (ICCV) 2025*
- Reviewer, *IEEE International Conference on Multimedia & Expo (ICME) 2025*

Administrative & Leadership Service

- Deputy Director, Competition Management Center, University's Innovation and Entrepreneurship Department