Zhenhao Zhang 章震豪

Email:zhangzhh2024@shanghaitech.edu.cn   WeChat: ACHEzzh   Homepage:zhangzhh.cn   Google Scholar   Openreview

EDUCATION

ShanghaiTech University,Master Student in Computer Science   2024.9 - now

MoE Key Laboratory, VDI Center, YesAI Lab. Advisor: Prof. Ye Shi and Prof. Jingya Wang
GPA: 3.85/4 (Specialized Course) Embodied AI: 4.0 (A+) Convex Optimization: 4.0 (A+) Deep learning: 3.7 (A-) Robotics: 3.7 (A-)

China University of Petroleum,B.E. Degree in Computer Science   2020.9 - 2024.6

Intern in the Department of Data Science and Statistics. Advisor by Prof. Yunquan Song
GPA: 3.3/5 C/C++ Program: 100 Matlab: 99 DIP: 95 Python Program: 95 Math Model: 95 Computer Vision: 90 Linear Algebra: 90

SELECTED PUBLICATIONS

Research Interest: Large Language Model, Spatial Intelligence, Embodied AI

[†]Equal Contribution [*]Corresponding author [#]Student First author [‡]Project Leader

OpenHOI: Open-World Hand-Object Interaction Synthesis with Multimodal Large Language Model[paper][page]

- **Zhenhao Zhang**, Ye Shi[*], Lingxiao Yang, Suting Ni, Qi Ye, Jingya Wang[*]
- Introduce the first Open-World Hand-Object Interaction (HOI) Synthesis framework that can generate Long-horizon HOI sequences of Unseen Objects from Open-vocabulary instructions with a 3D Multimodal Large Language Model.
- I am the **First Author**

DAG: Unleash the potential of Diffusion Model for Open-Vocabulary 3D Affordance Grounding[WAIC][page]

- Hanqing Wang[†*], **Zhenhao Zhang**[†#], Kaiyang Ji[†],et. al
- A novel framework designed to unlock affordance knowledge within diffusion models for 3D affordance grounding.
- We extract the implicit knowledge of Affordance from pre-trained stable diffusion, obtain implicit Condition using Self-Prompt, and transfer the knowledge of Affordance extracted from HOI Image to the point cloud
- DAG has been **Accepted** by **WAIC2025**! I am the **Co-first Author** and **Student First Author**

Spatial-ORMLLM: Improve Spatial Relation Understanding in the Operating Room with Multimodal Large Language Model

- Peiqi He[†], Zhenhao Zhang[†‡], Yixiang Zhang, Xiongjun Zhao[*], Shaoliang Pen[*]
- The first large vision-language model for 3D spatial reasoning in operating rooms using only RGB modality.
- Using 3D estimation algorithms and multimodal guidance enhancement, we aim to enhance the understanding of large multimodal models towards spatial relationships in clinical images.
- I am the **Co-first Author** and **Project Leader**

Diffusion-based Reinforcement Learning via Q-weighted Variational Policy Optimization[NeurIPS2024][page]

- Shutong Ding, Ke Hu, **Zhenhao Zhang**, Kan Ren, Weinan Zhang, Jingyi Yu, Jingya Wang, Ye Shi[†]
- We propose a novel diffusion-based online RL algorithm, conducting policy optimization with Q-weighted variational loss and diffusion entropy regularization to exploit the expressiveness and exploration capability of the diffusion policy.
- QVPO has been **Accepted** by **NeurIPS2024**!

SDEval: Safety Dynamic Evaluation for Multimodal Large Language Models

- Hanqing Wang, Yuan Tian, Mingyu Liu, **Zhenhao Zhang**, Xiangyang Zhu[*]
- We design a dynamic evaluation protocol for MLLM safety evaluation.

SELECTED HONORS AND AWARDS

ShanghaiTech University, Academic Scholarship	2024.9
China University of Petroleum, Academic Scholarship	2023.9
China University of Petroleum, Outstanding Scholarship(Competition)	2023.9
China Robotics and Artificial Intelligence Competition, National Second Prize	2023.7
LanQiao Cup, Shandong First Prize in Group A	2023.4
Kaggle Featured Code Competition, Bronze(SOLO)	2022.11
China University of Petroleum, Outstanding Scholarship(Competition)	2022.9
America Mathematics Contest in Modeling, Finalist, Top-1%	2022.6
LanQiao Cup, Shandong First Prize in Group A	2021.4

SELECTED INTERNSHIPS

SignalPlus, AI Agent System Development Intern   2025.6-2025.9

- Project: LLMs for Quantitative Trading. Mentor: Kaifeng Chen, Jingxiang Shan(**COO** of SignalPlus)
- Responsible for designing and developing a macroeconomic AI agent that monitors economic market events in real time and provides prompt data to downstream agents
- Establish a Macro Knowledge Base to enhance the performance of AI agents through RAG

ShanghaiTech University, Visiting Student   2024.3-2024.9
- VDI Center, YesAI Lab. Advisor: Prof. Ye Shi
- Research Area: Embodied AI, Reinforcement Learning
- Paper:NeurIPS2024[paper]

China University of Petroleum, Research Intern   2021.7-2023.9
- Department of Data Science and Statistics. Advisor: Prof. Yunquan Song
- Research Area: Machine Learning, Biomedical Information
- Paper:ARXIV2023[paper]

SERVICES
- PC Member: NeurIPS 2024, ICLR 2025, AISTATS 2025, ICML 2025, NeurIPS 2025, IJCNN 2025
- Association Member: CAA Member, CCF Student Member

SKILLS
- Coding: Python,C/C++,Matlab,Pytorch,OpenCV
- Math: Convex Optimzation(4.0,A+), Mathematical Experiment(99), Mathematics Modeling(95), Linear Algebra(90)
- Software: Latex, Blender, PPT, Word

OTHER EXPERIMENTS
- Predictive Algorithm Intern in HAOMO AI. Project: Hybrid Auto-driving based on rules and networks
- Research Intern in MIG@UPC, Project: Segmentation in the Dark.
- Research Intern in UPC-Artrobot Cooperation Item. Project: Intelligent Vehicle
- Member of ACM-ICPC Club@UPC and MCM Association@UPC