Xunzhe Zhou

(+86)180-9600-2356 • zhouxunzhe.github.io • xunzhe_zhou@outlook.com

EDUCATION

Fudan University <div align="right">Shanghai, China</div>
- B.S. in Computer Science and Technology, GPA 3.55/4.00 (Average grade 89/100) — 2021.09 - now
- Natural Science Experimental Class, GPA 3.58/4.00 — 2020.09 - 2021.06

Explanation: *First admitted to Natural Science Experimental Class, later transferred to Computer Science and Technology.*

University of California, Berkeley <div align="right">Berkeley, CA, USA</div>
- Exchange student**,** Department of EECS, GPA 4.00/4.00 — 2023.08 - 2023.12

Graduate courses: *CS282A Deep Learning, EECS227A Optimization;* ***Undergraduate course:*** *CS188 Artificial Intelligence.*

PUBLICATIONS (* denotes equal contribution)

- **<u>Xunzhe Zhou</u>***, J. Chen*, C. Yu*, T. Xu, Y. Mu, M. Hu, W. Shao, Y. Wang, G. Li, L. Shao, "EMOS: Embodiment-aware Heterogeneous Multi-robot Operating System with LLM Agents", *in submission*.
- Q. He, J. Zeng, W. Huang, L. Chen, J. Xiao, Q. He, **<u>Xunzhe Zhou</u>**, J. Liang, Y. Xiao "Can Large Language Models Understand Real-World Complex Instructions?" accepted by *AAAI 2024*.
- **<u>Xunzhe Zhou</u>***, R. Cao*, J. Hou, C. Guan, S. Leng, "Reservoir computing as digital twins for controlling nonlinear dynamical systems", *in submission*.

RESEARCH EXPERIENCE

School of Computing, National University of Singapore <div align="right">Singapore, Singapore</div>
Advisor: Prof. Lin Shao <div align="right">2024.05 - now</div>

- Constructed a heterogeneous multi-robot collaborating system with LLM agents in *Habitat-lab*. I was responsible for from low-level robot actions to high-level multi-agent task planning. I also constructed the *Habitat-MAS* benchmark dataset.
- Built a task planning framework with embodiment-aware visual prompting in *Habitat-lab*. Now working on aggregating learned skills to novel skills and environments with *Diffusion Policies* by distilling task-specific features.
- Co-first authored the paper *EMOS: Embodiment-aware Heterogeneous Multi-robot Operating System with LLM Agents*.

Shanghai Key Laboratory of Data Science, Fudan University <div align="right">Shanghai, China</div>
Advisor: Prof. Yanghua Xiao <div align="right">2023.06 - 2023.08</div>

- Conducted a benchmark to evaluate LLMs' capabilities of real-world complex instructions following. I was responsible for 1) constructing the dataset, 2) designing the evaluation criteria, and 3) evaluating the LLM models.
- Drafted the project proposal *A Practical Benchmark for Evaluating Large Language Models' Understanding of Complex Instructions under Hard Constraints* to apply for the National Natural Science Foundation of China (youth projects).
- Co-authored the paper *Can Large Language Models Understand Real-World Complex Instructions?*

Institute of AI and Robotics, Fudan University <div align="right">Shanghai, China</div>
Advisor: Prof. Siyang Leng <div align="right">2022.11 - 2023.05</div>

- Constructed an *Echo State Network* as the digital twin to predict and control the behavior of nonlinear dynamical (chaotic) systems. I was responsible for 1) constructing the model, 2) conducting experiments, and 3) revising the paper.
- Evaluated the model performance on 3 chaotic systems and 3 control strategies, collectively. Conducted experiments to validate the prediction accuracy, control efficiency, and noise robustness of the model.
- Co-first authored the paper *Reservoir Computing as Digital Twins for Controlling Nonlinear Dynamical Systems*.

SELECTED PROJECTS

Mobile Manipulation and Hierarchical Task Planning <div align="right">Fudan University, China</div>
Advisors: Prof. Yanwei Fu and Prof. Xiangyang Xue <div align="right">2024.03 - 2024.05</div>

- Constructed a mobile manipulation system with the robot assembled with *Franka Panda* arm and *Hermes* mobile base.
- I was responsible for constructing 1) semantic grasping pose estimation, 2) semantic mobile base navigation, and 3) hierarchical task planning. The follow-up work: *TaMMa* (Hou et al.) was accepted by *CoRL 2024*.

Resolving Knowledge Conflicts in Vision-Language Models <div align="right">Fudan University, China</div>
Advisors: Prof. Xiangyang Xue <div align="right">2024.03 - 2024.04</div>

- Constructed a VQA dataset with images involving knowledge conflicts from the Internet or generated with *DALL·E 3*.
- Evaluated 8 SOTA VLMs on the dataset, and resolved knowledge conflicts in *LLaVA-1.5* with contrastive decoding.

Neural Style Transfer Based on Fine Tuning Vision Transformer <div align="right">UC Berkeley, USA</div>
Advisors: Prof. Anant Sahai <div align="right">2023.11 - 2023.12</div>

- Replaced the content and style encoders of $StyTr^2$ with pre-trained ViT to improve the task of *Neural Style Transfer*.
- Leveraged a two-stage training strategy with *COCO* datasets: First freeze the pre-trained ViT, and just train the decoders. Then wrap LoRA to fine-tune ViT for joint training.

HONOR &AWARDS

- Second prize of scholarship in Outstanding Students 2021
- Third prize of scholarship in Outstanding Students 2023
- Second award in the National High School Mathematics League 2019
- Honor roll of distinction certificate in The Mathematics League (World Top 8%) 2016
- Champion of Soccer League, Fudan University 2023 & 2024

COMMUNITY SERVICE

- Fudan University Recruit Voluntary Group 2022
- Covid-19 Voluntary Service 2022
- Guizhou Province Voluntary Service 2019

SKILLS

- Relevant coursework: Deep Learning, Artificial Intelligence, Machine Learning, Deep Reinforcement Learning, Convex Optimization, Intro to Robotics, Data Mining.
- Programming Languages: Python, C/C++, ROS, Matlab, Verilog.
- Software: Pytorch, Git, L^AT_EX, COLMAP.
- Robots: Franka Emika Panda, Kinova Gen2, HERMES.
- AI Models: Vision-Language Models, Large Language Models, Diffusion Models, NeRF, 3DGS.
- Simulator: Habitat-lab, RoboCasa, PyBullet, MuJoCo, Gazebo, AI2-THOR, IssacSim, ThreeDWrold.

STANDARDIZED TESTS

- IELTS: Overall 7.0 (Listening 6.5+Reading 7.5+Writing 6.5+Speaking 6.5).
- Duolingo: Overall 120 (Literacy 120+ Comprehension 125+ Conversation 105+ Production 90).