

Zhan Ling

Research Scientist @ ByteDance

✉ zhan.ling@bytedance.com 🌐 <https://lz1oceani.github.io>

RESEARCH INTEREST

I am an explorer at heart, dedicated to pushing the frontiers of artificial intelligence. My long-term vision is to develop open-world, self-improving intelligent foundation models that drive us toward super-intelligence and expand the boundaries of human knowledge.

My career is defined by a relentless pursuit of diverse and challenging domains. During my PhD, I ventured broadly across imitation learning, reinforcement learning, robotics, 3D representation learning, and LLM reasoning. After that, I continued this exploration by driving LLM agent post-training at the cutting edge. I am continually exploring new technical frontiers, including training long-context LLM via reinforcement learning, and developing generative reward models for music generation.

EDUCATION

University of California San Diego, La Jolla, California, USA 2019.9–2024.12
Doctor of Philosophy, Computer Science, Advisor: Prof. Hao Su
Master of Science, Computer Science

Tsinghua University, Beijing, China 2015.8–2019.7
Bachelor of Engineering, Institute for Interdisciplinary Information Sciences (IIIS)
Yao Class (founded by Turing Award Laureate, Prof. Andrew Chi-Chih Yao)

PUBLICATIONS AND PREPRINTS

(*, **, ***, indicates equal contribution)

1. [Natural Language Actor-Critic: Scalable Off-Policy Learning in Language Space](#) Joey Hong, Kang Liu, **Zhan Ling**, Jiecao Chen, Sergey Levine. Preprint.
2. [Scaling Long-Horizon LLM Agent via Context-Folding](#) Weiwei Sun, Miao Lu, **Zhan Ling**, Kang Liu, Xuesong Yao, Yiming Yang, Jiecao Chen. *International Conference on Machine Learning (ICML) 2026*.
3. [Scaling LLM Multi-turn RL with End-to-end Summarization-based Context Management](#) Miao Lu, Weiwei Sun, Weihua Du, **Zhan Ling**, Xuesong Yao, Kang Liu, Jiecao Chen. *Annual Meeting of the Association for Computational Linguistics (ACL) 2026, Main*.
4. [Generalizable End-to-End Tool-Use RL with Synthetic CodeGym](#) Weihua Du, Hailei Gong, **Zhan Ling**, Kang Liu, Lingfeng Shen, Xuesong Yao, Yufei Xu, Dingyuan Shi, Yiming Yang, Jiecao Chen. *International Conference on Learning Representations (ICLR) 2026*.
5. [ORIC: Benchmarking Object Recognition in Incongruous Context for Large Vision-Language Models](#) Zhaoyang Li*, **Zhan Ling***, Yuchen Zhou, Litian Gong, Erdem Bıyık, Hao Su. *IEEE / CVF Computer Vision and Pattern Recognition Conference (CVPR) 2026*.

6. [MIR-Bench: Benchmarking LLM’s Long-Context Intelligence via Many-Shot In-Context Inductive Reasoning](#). Kai Yan, **Zhan Ling**, Kang Liu, Yifan Yang, Ting-Han Fan, Lingfeng Shen, Zhengyin Du, Jiecao Chen. *Neural Information Processing Systems (NeurIPS) Datasets and Benchmarks Track 2025*.
7. [LongReason: A Synthetic Long-Context Reasoning Benchmark via Context Expansion](#). **Zhan Ling**, Kang Liu, Kai Yan, Yifan Yang, Weijian Lin, Ting-Han Fan, Lingfeng Shen, Zhengyin Du, Jiecao Chen. Preprint.
8. [ActPlan-1K: Benchmarking the Procedural Planning Ability of Visual Language Models in Household Activities](#). Ying Su, **Zhan Ling**, Haochen Shi, Cheng Jiayang, Yauwai Yim, Yangqiu Song. *Empirical Methods in Natural Language Processing (EMNLP) 2024 Main*.
9. [Unleashing the Creative Mind: Language Model As Hierarchical Policy For Improved Exploration on Challenging Problem Solving](#). **Zhan Ling**, Yunhao Fang, Xuanlin Li, Tongzhou Mu, Mingu Lee, Reza Pourreza, Roland Memisevic, Hao Su. Preprint.
10. [DreamFuser: Value-guided Diffusion Policy for Offline Reinforcement Learning](#) Kairong Luo, Caiwei Xiao, Zhiao Huang, **Zhan Ling**, Yunhao Fang, Hao Su. Preprint.
11. [Deductive Verification of Chain-of-Thought Reasoning](#). **Zhan Ling***, Yunhao Fang*, Xuanlin Li, Zhiao Huang, Mingu Lee, Roland Memisevic, Hao Su. *Neural Information Processing Systems (NeurIPS) 2023*.
12. [On the Efficacy of 3D Point Cloud Reinforcement Learning](#). **Zhan Ling***, Yunchao Yao*, Xuanling Li, Hao Su. Preprint.
13. [Distilling Large Vision-Language Model with Out-of-Distribution Generalizability](#). Xuanlin Li*, Yunhao Fang*, Minghua Liu, **Zhan Ling**, Zhuowen Tu, Hao Su. *IEEE / CVF International Conference on Computer Vision (ICCV) 2023*.
14. [Reparameterized Policy Learning for Multimodal Trajectory Optimization](#). Zhiao Huang, Litian Liang, **Zhan Ling**, Xuanlin Li, Chuang Gan, Hao Su. Deep Reinforcement Learning Workshop, Neural Information Processing Systems(NeurIPS) 2022; *International Conference on Machine Learning (ICML) 2023*, **Oral**.
15. [PartSLIP: Low-Shot Part Segmentation for 3D Point Clouds via Pretrained Image-Language Models](#). Minghua Liu, Yin hao Zhu, Hong Cai, Shizhong Han, **Zhan Ling**, Fatih Porikli, Hao Su. *IEEE / CVF Computer Vision and Pattern Recognition Conference (CVPR) 2023*.
16. [ManiSkill2: A Unified Benchmark for Generalizable Manipulation Skills](#). Jiayuan Gu*, Fanbo Xiang*, Xuanlin Li**, **Zhan Ling****, Xiqiang Liu**, Tongzhou Mu**, Yihe Tang**, Stone Tao**, Xinyue Wei**, Yunchao Yao**, Xiaodi Yuan, Pengwei Xie, Zhiao Huang, Rui Chen, Hao Su. *International Conference on Learning Representations (ICLR) 2023*.
17. [Frame Mining: a Free Lunch for Learning Robotic Manipulation from 3D Point Clouds](#). Minghua Liu*, Xuanlin Li*, **Zhan Ling***, Yangyan Li, Hao Su. *Conference on Robot Learning (CoRL) 2022*.
18. [Improving policy optimization with generalist-specialist learning](#). Zhiwei Jia, Xuanlin Li, **Zhan Ling**, Shuang Liu, Yiran Wu, Hao Su. *International Conference on Machine Learning (ICML) 2022*.
19. [Close the Visual Domain Gap by Physics-Grounded Active Stereovision Depth Sensor Simulation](#). Xiaoshuai Zhang*, Rui Chen*, Ang Li**, Fanbo Xiang**, Yuzhe Qin**, Jiayuan Gu**, **Zhan**

- Ling****, Minghua Liu**, Peiyu Zeng**, Songfang Han***, Zhiao Huang***, Tongzhou Mu***, Jing Xu, Hao Su. *IEEE Transactions on Robotics (T-RO)* 2023.
20. [Approximate Convex Decomposition for 3D Meshes with Collision-Aware Concavity and Tree Search](#). Xinyue Wei*, Minghua Liu*, **Zhan Ling**, Hao Su. *ACM Transactions on Graphics (Proceedings of SIGGRAPH)* 2022.
 21. [ManiSkill: Generalizable Manipulation Skill Benchmark with Large-Scale Demonstrations](#). Tongzhou Mu*, **Zhan Ling***, Fanbo Xiang*, Derek Yang*, Xuanlin Li*, Stone Tao, Zhiao Huang, Zhiwei Jia, Hao Su. *Neural Information Processing Systems (NeurIPS) Datasets and Benchmarks Track* 2021.
 22. [State Alignment-based Imitation Learning](#). Fangchen Liu, **Zhan Ling**, Tongzhou Mu, Hao Su. *International Conference on Learning Representations (ICLR)* 2020.

WORK EXPERIENCES

ByteDance Inc., San Jose

2024.12-Now

Research Scientist

End-to-end Multi-turn Function Call Training with Reinforcement Learning

- Developed a robust synthetic data generation pipeline to significantly enhance end-to-end function call capabilities. ([4])
- Engineered RL pipelines with dynamic context management, successfully empowering LLM agents to execute long-horizon tasks. ([3], [2])

Generative Reward Modeling for Music Generation (In Progress)

- Built an RL pipeline to train a speech-to-text reasoning model as a generative reward model for high-fidelity music generation.

ByteDance Inc., San Jose

2024.6-2024.12

Student Researcher, Host: Jiecao Chen

Building a Synthetic Benchmark/Dataset for Long-Context Reasoning

- Developed a novel algorithm to synthesize high-quality, long-context reasoning question-and-answer pairs.
- Introduced a comprehensive synthetic benchmark to rigorously evaluate and compare the performance of open-source and closed-source models. An internal version of the benchmark was used to evaluate LLMs' long-context reasoning ability.
- Details can be found in the paper [7].

Qualcomm AI Research, San Diego

2023.6-2023.9

System 2 Team

Interim Engineering Intern, Host: Mingu Lee, Reza Pourreza, Roland Memisevic

Unleashing the Creative Mind: Language Model As Hierarchical Policy For Improved Exploration on Challenging Problem Solving

- Addressed the limited exploration capabilities of LLMs in reasoning tasks by identifying their bottleneck in high-level tactic spaces.
- Developed a novel hierarchical policy framework to enhance exploration: a high-level policy dictates strategic direction while a low-level policy executes detailed reasoning steps.
- Details can be found in the paper [9].

Qualcomm AI Research, Remote

2023.3-2023.6

System 2 Team

Support Engineering, Host: Mingu Lee, Roland Memisevic

Deductive Verification of Chain-of-Thought Reasoning

- Proposed "Natural Program," a verifiable reasoning format to address LLM hallucinations and logical errors.
- Leveraged LLMs for step-by-step self-verification, successfully filtering invalid chains and boosting reasoning reliability.
- Details can be found in the paper [11].

Qualcomm AI Research, Remote

2022.6-2023.3

System 2 Team

Interim Engineering Intern, Host: Mingu Lee, Roland Memisevic

Learning Code Execution with Language Model

- Investigated the capability of LLMs to execute code as a Turing machine by training models (Like 6.7B models) on python code execution traces.
- Developed a random code generator supporting loops, conditionals, and function calls to synthesize robust training data.
- Engineered a modified interpreter to automatically extract execution traces from generated code for model training.
- Trained and evaluated LLMs on code execution tasks, achieving high accuracy and strong generalization to unseen programs.
- Resulted in a filed US patent [US20250355632A1].

X, the moonshot factory, Remote

2020.6-2020.9

Mineral Project

AI Resident, Host: Lianghao Li, Kangkang Wang

Lane detection for agriculture images

- We develop a lane detection algorithm using the Hough transform in Python and accelerate its speed with Cython.

Improved weed detection with lane detection

- We create a weed detection model based on EfficientDet.
- We integrate the lane detection algorithm with the weed detection model to enhance overall weed detection performance.

University of California San Diego, San Diego

2019.9-Now

Su Lab

Graduate Research Assistant, Host: Hao Su

Embodied AI

- We develop generalizable manipulation skill benchmarks, ManiSkill, to evaluate the generalization ability of manipulation skills.
- We develop reinforcement learning algorithms and imitation learning algorithms for robotic tasks.

Computer Vision

- We develop 3D perception algorithms, such as 3D mesh convex decomposition, and 3D part segmentation.
- We integrate 3D representation learning with reinforcement learning algorithms for robotic tasks.

Reasoning

- We develop algorithms to enhance reasoning abilities using LLMs.

AWARDS AND SERVICES

Conference Reviewer: CVPR 2022-2026; ICCV 2021, 2023, 2025; ECCV 2022, 2024, 2026; NeurIPS 2023-2025; ICLR 2024-2026; ICML 2024-2026; ACCV 2024; ICRA 2025; AISTATS 2025; ACL RR 2026.

Journal Reviewer: T-RO; RA-L.

Program Committee: AAAI 2025-2026.

Awards:

Outstanding Freshman Scholarship, Tsinghua University, 2015

First Prize (Gold Medal) in National Olympiad in Informatics (NOI), China, 2014

Workshop Reviewer:

Generalizable Policy Learning in the Physical World, ICLR 2022.

Interdisciplinary Exploration of Generalizable Manipulation Policy Learning: Paradigms and Debates, RSS 2023.

Challenge Organizer or Contributor:

[SAPIEN ManiSkill Challenge 2021](#)

[SAPIEN ManiSkill Challenge 2022](#)

TECHNIQUE SKILLS

Programming Language: Python, C, C++.

Deep Learning: PyTorch, Tensorflow, Jax, Pytorch CUDA extension.

Large Models Related Skills: Quantization, Distributed training, Fine-tuning, Pre-training dataset processing.

Parallel Computing: MPI, OpenMP, CUDA, AVX.

Simulation: SAPIEN, MuJoCo.

Additional Skills: Pybind, ROS.

last updated: May 5, 2026