Yinuo Zhao

github | Homepage | linda.chao.007@gmail.com

Research Overview

I have focused on the study of reinforcement learning (RL) since obtaining my bachelor's degree. My research is driven by the observation that traditional RL algorithms often struggle with real-world challenges, particularly when faced with sparse feedback and complex action distributions. During my doctoral research, I addressed these challenges by developing novel algorithms aimed at improving both the learning efficiency and stability in environments with limited feedback. My work focuses on refining exploration strategies and utilizing auxiliary tasks to mitigate the negative effects of sparse rewards. Additionally, I am exploring solutions to complex multi-task manipulation problems by leveraging large-scale offline demonstrations. Ultimately, my goal is to develop general manipulation policies that can autonomously recover and improve themselves in dynamic and unpredictable environments.

Education

2021 - Now	PhD candidate at **Department of Computer Science, BIT**	(Grade: 3.9/4.0)
2019 - 2021	Master at **Department of Computer Science, BIT**	(Grade: 3.9/4.0)
Core Courses:	Stochastic Process, Matrix Analysis, Language Information Processing	
2015 - 2019	Bachelor's Degree at **Department of Computer Science, BIT**	(Grade: 3.9/4.0)
Core Courses:	Linear Algebra, Data Structures and Algorithms, Computer Network	

Employment

Midea Home Service Robotic Labs, research intern Sep 2021 - Now

- Worked on generalizable end-to-end visuomotor manipulation policies, especially for home service robot.
- Built a sim-to-real large-scale cabinet manipulation training and testing platform based on **Isaac**.
- Won the **first prize** (Rank 2nd) in No Interaction Track in ManiSkill2021 (Team Fattonny). [website]

DiDi AI Labs, research intern Mar 2019 - Jan 2021

- Worked on end-to-end autonomous urban driving policy based on single monocular camera.
- Built an urban driving policy's training platform that enables distributed DRL training based on CARLA. [github]

Professional Service

ICLR (2023-25), AAAI (2024-25), ICCV (2025), ICML (2024), IROS (2024), NeurIPS (2024-25), TMC (2025)

Selected Research

Empowering LLMs with DRL for Failure Recovery in Robot Manipulation Arxiv

We identify two main **sources of failure** in robot manipulation: external failures, caused by environmental changes that make current execution conditions unsuitable, and internal failures, due to the robot's inability to execute the current plan. Current off-the-shelf models lack in-context awareness of the environment, making them inefficient in recovering from failures. To address this, we introduce a self-recovery framework that helps RL agents learn and evaluate policies for specific tasks. Specifically, we design a zero-shot LLM-based agent that provides **subgoal-based rewards** for downstream training. Within this framework, we propose a self-adversarial module that autonomously generates and recovers from failures, enabling the composition of complex skills from a primitive skill library.

A Cascade DRL Framework for Vision-based Autonomous Urban Driving AAAI'21 Oral
We are **one of the first** to present a successful online DRL agent for urban driving, particularly for handling dense traffic. We decompose the learning of complex vision-based sensorimotor policies into two stages: 1) First, we train a Co-attention Perception Module (CoPM) that uses the **co-attention mechanism** to capture the relationships between visual and control information. 2) Next, based on frozen CoPM, we introduce an efficient **distributed proximal policy optimization** framework to learn the driving policy guided by specially designed reward functions. We evaluate our CADRE framework using the CARLA NoCrash benchmark **under dynamic road conditions and varying weather**.

Curiosity-Driven Energy-Efficient Worker Scheduling in Vehicular Crowdsourcing ICDE'20
In this paper, we address the **cooperation trajectory planning** problem for a swarm of aerial robots tackling unevenly distributed data-collection tasks. We introduce a novel DRL approach, "DRL-CEWS," which leverages curiosity-driven, energy-efficient worker scheduling to balance data collection, coverage fairness, and energy consumption. Our empirical results show that incorporating a spatial-curiosity model and **sparse** feedback leads to **better planning outcomes** than using only **dense rewards**. This research has deepened my interest in integrating human cognitive behaviors into intelligent systems.

Publications

Zhan, Y., Liu, C. H., **Zhao, Y.**, Zhang, J., Tang, J., (2019). "Free market of multi-leader multi-follower mobile crowdsensing: An incentive mechanism design by deep reinforcement learning". In: *IEEE Transactions on Mobile Computing (TMC),* **CCF-A** 19.10, pp. 2316–2329.

Liu, C. H., **Zhao, Y.**, Dai, Z., (2020). "Curiosity-driven energy-efficient worker scheduling in vehicular crowdsourcing: A deep reinforcement learning approach". In: *2020 IEEE 36th International conference on data engineering (ICDE),* **CCF-A**. IEEE, pp. 25–36.

Zhao, Y., Liu, C. H., (2020). "Social-aware incentive mechanism for vehicular crowdsensing by deep reinforcement learning". In: *IEEE Transactions on Intelligent Transportation Systems,* **SCI-Q1** 22.4, pp. 2314–2325.

Liu, C. H., Dai, Z., **Zhao, Y.**, Crowcroft, J., Wu, D., Leung, K. K., (2021). "Distributed and Energy-Efficient Mobile Crowdsensing with Charging Stations by Deep Reinforcement Learning". In: *IEEE TMC,* **CCF-A** 20.1, pp. 130–146.

Wu, K., **Zhao, Y.**, Xu, Z., **Zhao, Z.**, Ren, P., Che, Z., Liu, C. H., Feng, F., Tang, J., (2022). "A minimalist ensemble method for generalizable offline deep reinforcement learning". In: *ICLR 2022 Workshop on Generalizable Policy Learning in Physical World.*

Zhao, Y., Wu, K., Xu, Z., Che, Z., Lu, Q., Tang, J., Liu, C. H., (2022). "Cadre: A cascade deep reinforcement learning framework for vision-based autonomous urban driving". In: *Proceedings of the AAAI Conference on Artificial Intelligence,* **CCF-A**. Vol. 36. 3, pp. 3481–3489.

Wu, K., **Zhao, Y.**, Xu, Z., Che, Z., Liu, C. H., Feng, F., Tang, J., (2024). "An Adaptive Q-Learning Framework for Offline Reinforcement Learning". In: *TNNLS(In Press),* **SCI-Q1**.

Zhao, Y., Liu, C. H., Yi, T., Li, G., Wu, D., (2024). "Energy-Efficient Ground-Air-Space Vehicular Crowdsensing by Hierarchical Multi-Agent Deep Reinforcement Learning with Diffusion Models". In: *IEEE Journal on Selected Areas in Communications,* **CCF-A**.

Zhao, Y., Wu, K., Yi, T., Xu, Z., Ju, X., Che, Z., Qiu, Q., Liu, C. H., Tang, J., (2024). "Efficient Training of Generalizable Visuomotor Policies via Control-Aware Augmentation". In: *Under review.*