Yinuo Zhao

github | Google Scholar | linda.chao.007@gmail.com

Summary

I am a Ph.D. student advised by Prof. Chi Harold Liu at the School of Computer Science, Beijing Institute of Technology (BIT). I have been working on vision deep reinforcement learning (RL) since I got my bachelor degree at the School of CS, BIT in 2019. I'm interested in embodied Artifical Intelligence and doing research on robot learning with end-to-end RL. My IELTS score is 7.

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, Software requirements engineering and modeling, Data Structures and Algorithms, Computer Network	

Experience

Midea Home Service Robotic Labs, internship 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, internship 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]

Computer Programming on Algorithm Lab Mar 2015 - Jan 2017

- Learned basic optimization algorithms (e.g., dynamic programming, segment tree, graphs)
- Took part in programming competitions, like "Blue Bridge" programming design competition.

Researches

Empowering LLMs with DRL for Failure Recovery in Robot Manipulation CoRL Under Review

In this paper, we introduce DRL agents to bridge the gap between the **common knowledge** embedded in LLMs and **executable skills** in specific tasks. Specifically, we design an LLM-based zero-shot agent to provide **subgoal-based rewards** for downstream training. Under this framework, we further propose a **self-adversarial module** to compose complex skills from primitive skill library by autonomously generating and **recovering from failures**.

Energy-Efficient VCS by Hierarchical MADRL with Diffusion Models JSAC Under Review

In this paper, we consider a vehicular crowdsensing (VCS) campaign where **unmanned ground vehicles (UGVs) periodically dispatch and recall unmanned aerial vehicles (UAVs)** at various workzone stops to gather unevenly distributed data. We propose an energy-efficient, **goal-directed hierarchical** MADRL method that uses discrete diffusion models. This method optimizes high-level navigation policies for UGVs and long-term sensing strategies for UAVs. Experimental results on **real-world datasets** from Rome, Italy, and Hong Kong SAR, China, demonstrate the effectiveness of our method.

An Adaptive Q-Learning Framework for Offline Reinforcement Learning TNNLS Under review

Offline RL suffers from the overestimation of values induced by the distributional shift. Previous work necessitated prior knowledge of the dataset's quality, resulting in substantial trial-and-error efforts to fine-tune hyperparameters. In this paper, we provide more **adaptive and flexible control over the conservative level of the Q-function for offline RL**. Specifically, we present two adaptive weight functions to shape the Q-values for collected and out-of-distribution data. We then give theoretical analysis about the changing conditions of the conservative level and formally define the monotonicity with respect to data quality and similarity. Motivated by the theoretical analysis, we propose a novel algorithm AdaQL under the framework, which leverages neural networks as the adaptive weight functions. We evaluate AdaQL on the commonly-used **D4RL benchmark** and conduct extensive ablation studies to illustrate the effectiveness and state-of-the-art (SOTA) performance compared to existing offline DRL baselines.

An Efficient Generalizable Framework for Visuomotor Policies via ... arXiv

In this paper, we aim to solve the **zero-shot generalization problem** for **visuomotor policies** from two aspects: 1) We learn a control-aware mask by a self-supervised attention mechanism and then apply strong augmentation based on the mask to reduce the generalization gaps. 2) To eliminate the training instability problem resulting from the data augmentations, we directly use a privileged expert to guide the visual control policy. Extensive experiments on **DMC-GB**, **Robot Manipulation Benchmark** and our self-designed **cabinet-opening task** demonstrate SOTA generalization ability of our method.

A Cascade DRL Framework for Vision-based Autonomous Urban Driving AAAI'21 Oral

In this work, we present the first successful online DRL agent on urban driving, especially with dense traffic handling. We first train a Co-attention Perception Module (CoPM) that leverages the **co-attention mechanism** to learn the inter-relationships between the visual and control information. Then, we present an efficient **distributed proximal policy optimization** framework to learn the driving policy under the guidance of particularly designed reward functions. We evaluate our framework CADRE on **CARLA NoCrash benchmark** with dynamic road conditions and various weather. The experimental results well justify the effectiveness of CADRE.

Curiosity-Driven Energy-Efficient Worker Scheduling in Vehicular Crowdsourcing ICDE'20

In this paper, we explicitly consider the use of unmanned vehicular workers, e.g., drones and driverless cars, to carry on **long-term and hash tasks as a vehicular crowdsourcing (VC) campaign**. We propose a novel DRL approach for curiosity driven energy-efficient worker scheduling, called "DRL-CEWS", to achieve an optimal trade-off between maximizing the collected amount of data and coverage fairness, and minimizing the overall energy consumption of workers. Specifically, we first utilize a **chief-employee distributed computational architecture** to stabilize and facilitate the training process. Then, we propose a **spatial curiosity model** with a sparse reward mechanism to help derive the optimal policy in large crowdsensing space with unevenly distributed data.

References

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,* **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 Trans. Mob. Comput,* **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(Under review),* **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 (In Press),* **CCF-A**.

Zhao, Y., Wu, K., Yi, T., Xu, Z., Ju, X., Che, Z., Qiu, Q., Liu, C. H., Tang, J., (2024). "An Efficient Generalizable Framework for Visuomotor Policies via Control-aware Augmentation and Privilege-guided Distillation". In: *arXiv preprint arXiv:2401.09258*.

Zhao, Y., Yuan, J., Xu, Z., Tang, J., Liu, C. H., (n.d.). "Empowering LLMs with Deep Reinforcement Learning for Failure Recovery in Robot Manipulation". In: *(Under review)*.