Xiyao Wang

xywang@umd.edu · 1-240-467-1062

Education

University of Maryland, College Park, Maryland, United States 2022.8 – Present
Ph.D. in Department of Computer Science, expected in May 2026

University of Electronic Science and Technology of China, Chengdu, China 2015.9 – 2019.6
B.S. in Network Engineering

Research Interests

- Vision language and large language model for decision making
- Sample efficient reinforcement learning (especially model-based RL)
- Knowledge transfer in reinforcement learning (in-domain and cross-domain transfer learning)
- Representation learning for control problems

Publications

* is Equal contribution.

Conference Paper

- **Xiyao Wang**, Ruijie Zheng, Yanchao Sun, Ruonan Jia, Wichayaporn Wongkamjan, Huazhe Xu and Furong Huang. "COPlanner: Plan to Roll Out Conservatively but to Explore Optimistically for Model-Based RL."
  International Conference on Learning Representation (ICLR), 2024.

- Guowei Xu*, Ruijie Zheng*, Yongyuan Liang*, **Xiyao Wang**, Zhecheng Yuan, Tianying Ji, Yu Luo, Xiaoyu Liu, Jiaxin Yuan, Pu Hua, Shuzhen Li, Yanjie Ze, Hal Daumé III, Furong Huang, Huazhe Xu. "DrM: Mastering Visual Reinforcement Learning through Dormant Ratio Minimization."
  *International Conference on Learning Representation (ICLR), 2024 (**Spotlight**).*

- Ruijie Zheng, **Xiyao Wang**, Yanchao Sun, Shuang Ma, Jieyu Zhao, Huazhe Xu, Hal Daumé III, Furong Huang. "TACO: Temporal Latent Action-Driven Contrastive Loss for Visual Reinforcement Learning."
  Neural Information Processing Systems (NeurIPS), 2023.

- **Xiyao Wang**, Wichayaporn Wongkamjan, Ruonan Jia and Furong Huang. "Live in the Moment: Learning Dynamics Model Adapted to Evolving Policy."
  International Conference on Machine Learning (ICML), 2023.
  *Abridged in ICML Decision Awareness in Reinforcement Learning Worshop 2022 (**Spotlight**).*

- Ruijie Zheng*, **Xiyao Wang***, Huazhe Xu, and Furong Huang. "Is Model Ensemble Necessary? Model-based RL via a Single Model with Lipschitz Regularized Value Function."
  International Conference on Learning Representation (ICLR), 2023.
  *Abridged in NeurIPS 2022 DRL Workshop (**Spotlight**).*

- Yanchao Sun, Ruijie Zheng, **Xiyao Wang**, Andrew Cohen, and Furong Huang. "Transfer RL across Observation Representations via Model-Based Regularization."
  International Conference on Learning Representation (ICLR), 2022.

- Yankun Yu, Huan Liu, Minghan Fu, Jun Chen, **Xiyao Wang**, Keyan Wang "A Two-branch Neural Network for Non-homogeneous Dehazing via Ensemble Learning."
  New Trends in Image Restoration and Enhancement, CVPR workshop 2021.

Under review

- **Xiyao Wang**, Yuhang Zhou, Xiaoyu Liu, Hongjin Lu, Feihong He, Yuancheng Xu, Taixi Lu, Gedas Bertasius, Mohit Bansal, Furong Huang, Huaxiu Yao. "Mementos: A Comprehensive Benchmark for Multimodal Large Language Model Reasoning over Image Sequence"
- Yuancheng Xu, Chenghao Deng, Yanchao Sun, Ruijie Zheng, **Xiyao Wang**, Jieyu Zhao, Furong Huang. "Equal Long-term Benefit Rate: Adapting Static Fairness Notions to Sequential Decision Making" https://arxiv.org/abs/2309.03426

- Ruijie Zheng, Yongyuan Liang, **Xiyao Wang**, Shuang Ma, Hal Daumé III, Huazhe Xu, John Langford, Praveen Palanisamy, Kalyan Shankar Basu, Furong Huang. "PREMIER-TACO is a Few-Shot Policy Learner: Pretraining Multitask Representation via Temporal Action-Driven Contrastive Loss."

Ongoing projects

- Finetuning Large Language Model for low-level control
  Our goal is to enable Large Language Models (LLMs) to possess low-level control capabilities in robotics, allowing them to directly generate actions executable by robots rather than merely high-level objectives. By finetuning the LLM with data collected across various robotics tasks, we aim to endow it with the ability to generalize and solve unseen tasks.
- Vision Language Model for implicit reward design in reinforcement learning
  We finetune a vision-language model (CLIP) using in-domain data. Subsequently, we utilize the distance between the text embedding of the task description and the image embedding of the current observation as the reward function. Additionally, we pretrain a latent dynamics model in the image space of CLIP. This model, in conjunction with the defined reward function, is then employed for zero-shot planning on unseen tasks.

Work Experience

Lalamove Beijing, China 2021.3 – 2022.6

Machine Learning Engineer

• Responsible for the development and iteration of offline coupon distribution algorithm. Proposed using user elasticity model for 3D coupon issuance, and the user conversion rate is increased by 20%.

Professional Services

- Reviewer of International Conference on Learning Representations (ICLR), 2024
- Reviewer of Advances in Neural Information Processing Systems (NeurIPS), 2022, 2023
- Reviewer of International Conference on Machine Learning (ICML), 2022, 2023

Invited Talks

- Invited talk at Tsinghua University, 2022.

Skills

- Good at Python and Golang, have good machine learning and deep learning foundation, proficiency in using PyTorch and Jax, familiar with SQL and Spark.
- IELTS 7 (R:8 L:7 W:6 S:6)