Yiwen Qiu

Room524A, Zijing 5#, Tsinghua University, Beijing, 100084, P. R. China | (+86)13521166552 |
qyw19@mails.tsinghua.edu.cn

EDUCATION

Department of Automation, Tsinghua University, Beijing, China *Aug 2019 - Jul 2023(Expected)*

- Major GPA: 3.83/4.0 Ranking: Top 15% in 160+ students (Top 3 in Female Students)
- **Core Courses:**
  Operations Research / Pattern Recognition and Machine Learning / Intelligent Optimization Algorithms and Its Applications / *Big Data and Machine Intelligence / Foundation of Artificial Intelligence / Automatic Control Theory / Computer Principles and Applications / Signals and System Analysis*
- **Honors and Awards:**
  Person of the Year in Department of Automation (10 in 500+) *Dec.2021*
  Scholarship for overall excellence (**highest scholarship for comprehensive performance**) *Oct.2021*
  Excellent Youth League member, Tsinghua University (15/105, awarded for **leadership** in teamwork) *Oct.2020*
  Weichai Scholarship, for academic excellence and excellent social work *Oct.2020*

RESEARCH INTERESTS

- Applying transfer learning to various applications, e.g. reinforcement learning (RL) & imitation learning (IL) in sim2real settings.
- Combination of control theory & deep learning.

PUBLICATIONS

- **Yiwen Qiu**, Jialong Wu, Zhangjie Cao, Mingsheng Long, "Out-of-Dynamics Imitation Learning from Multimodal Demonstrations", *accepted by Conference on Robot Learning (CoRL), 2022*
- Haoyi Niu, Shubham Sharma, **Yiwen Qiu**, Ming Li, Guyue Zhou, Jianming Hu, Xianyuan Zhan, "When to Trust Your Simulator: Dynamics-Aware Hybrid Offline-and-Online Reinforcement Learning", *submitted to Neural Information Processing Systems (NeurIPS), 2022* **arxiv**

PROJECT EXPERIENCES

Out-of-Dynamics(OOD) Imitation Learning from Multimodal Demonstrations | RA *March 2022 - Present*
*Advisor: **Mingsheng Long**, Associate Professor, Machine Learning Group, School of Software, Tsinghua University*

- Studied **out-of-dynamics** imitation learning (OOD-IL): the assumption in Imitation Learning(IL) is that the demonstrator who collects demonstrations **share the same dynamics** as the imitator **limits the usage of IL**. Aimed at enabling a wider usage of a mixture of mutimodal demonstrations in IL.
- Developed a novel **sequence-based contrastive clustering** algorithm to tackle the **multimodal distribution** problem in demonstrations collected under multiple sources and mitigated their negative mutual influence.
- Developed an adversarial-based transferability measurement to down-weight non-transferable demonstrations for OOD-IL which enables agents to learn from a mixture of source data under **different dynamics.**
- Conducted experiments on **3 MuJoCo** environments, a **driving** and **a simulated robot** environment, showing that the proposed approach outperforms prior works on final IL performance by 100 ~ 300%.

Modularized Out-of-Dynamics(OOD) Imitation Learning | RA *July 2022 - Present*
*Advisor: **Kun Zhang**, Associate Professor, Carnegie Mellon University*

- Considered a more general case where trajectories are **composed of multiple sub-tasks** based on the previous work on OOD-IL. Trying to generalize policies to novel scenarios with **arbitrary recombination** of sub-tasks.
- Designed an algorithm to shed light on the **hidden structures of sub-tasks** from their resulting state-action trajectory sequences under multiple dynamics from a causal view by leveraging mutual information theory.

- Developing a **hierarchical conditional policy** to generate trajectories in accord with the target new environment by appropriately disambiguating between different sub-tasks and constraining the transition to be smooth.

Dynamics-Aware Hybrid Offline-and-Online Reinforcement Learning | RA *Feb 2022 - April 2022*

*Advisor: **Xianyuan Zhan,*** Institute for AI Industry Research (AIR), Tsinghua University, China

- Combined learning from **limited** real data in offline RL and **unrestricted exploration** of imperfect simulators in online RL, which is a novel scenario.
- Proposed the Dynamics-Aware **H**ybrid **O**ffline-and-**O**nline Reinforcement Learning(H2O) framework, theoretically proved it can allow learning with high-fidelity from both offline-dataset and online-exploration.
- Designed a practical implementation with PyTorch through **an adversarial training** process, adaptively penalizing the learning on simulated state-action pairs with large dynamics gaps.
- Conducted experiments in **4** datasets of MuJoCo **each with 3 unreal dynamics** (Gravity / Friction / Joint-Noise) and a **real wheel-legged robot**, and achieved results beat all existing baselines.

Universal Domain Adaptation with Meta-learning | RA *Aug 2021-Dec 2022*

*Advisor: **Mingsheng Long**, Associate Professor, Machine Learning Group, School of Software, Tsinghua University*

- Aimed to eliminate the **label category gap** on sources and target domains in Domain Adaptation (DA) tasks, called Universal DA by identifying outlier samples without the need for prior knowledge.
- Conducted experiments with PyTorch and achieved improving performance on Office31, OfficeHome settings. (1~2% in accuracy, 8% in h-score) by utilizing **a meta-learning method.**
- Demonstrated that identifying outlier samples through distributional distance measurement is beneficial. Detecting outlier is not enough, intended to consider harder circumstances like **long-tail distribution** in real-world settings.

Modular Networks for Domain Generalization | RA *Nov 2021-Jan 2022*

A*dvisor: **Mingsheng Long**, Associate Professor, Machine Learning Group, School of Software, Tsinghua University*

- Considered enabling the model to have the ability to solve problems for **any target domain** (while DA algorithm aims to solve domain gap for a **specific single target**) with the access to an abundance of source domains, called Domain Generalization (DG)
- Designed a novel mixture-of-experts **modular structure with attention mechanism** for models to merge domain-generic and domain-specific information **selectively produce knowledge** in a more flexible way.
- Conducted experiments on **OfficeHome** and **WILDS** datasets for image classification task **in unseen domain(DG tasks)** showing that the modularized design significantly boosts the performance by 1%, while there are currently no DG approaches proved to be effective on the WILDS dataset.