

MINGHUAN LIU

Email: minghuanliu AT sjtu.edu.cn

Website: <http://minghuanliu.com> [Google Scholar]

EDUCATION

Shanghai Jiao Tong University (SJTU)

Sep. 2019 - June 2024 Expected

Ph.D. in Computer Science and Technology

- Apex Data & Knowledge Management Lab
- Prior Leader of the **ApexRL** research group
- Advisor: **Weinan Zhang**
- Member of Wu Wen Jun Honorary Doctoral Plan (Advisor: **Cewu Lu**)

Southwest Jiaotong University (SWJTU)

Sep. 2015 - July. 2019

B.S. in Computer Science and Technology

- Overall GPA: 3.84/4.0 Ranking: 1/98
- Key Lab of Cloud Computing and Intelligent Technology
- Advisor: **Tianrui Li**

RESEARCH INTERESTS

- My general research interest lies in developing principled and efficient **reinforcement learning (RL)** algorithms to tackle kinds of decision-making problems and build essential AI.
- I am sincerely devoted to **data-driven** RL methods, specifically, **imitation learning**, **offline RL**, that can make RL applicable for real-world challenges, like autonomous driving, sports analysis, healthcare, games, robotics, etc.

PUBLICATIONS

- 11 Visual Imitation Learning with Patch Rewards.** [[Project Page](#)]
Minghuan Liu, Tairan He, Weinan Zhang, Shuicheng Yan, Weinan Zhang.
The 11th International Conference on Learning Representations. **ICLR 2023**
- We propose to measure the expertise of various local regions of image samples, or called *patches*, and recover multi-dimensional *patch rewards* accordingly. Patch reward is a more precise rewarding characterization that serves as a fine-grained expertise measurement and visual explainability tool.
- 10 PerfectDou: Dominating DouDizhu with Perfect Information Distillation.** [[Project Page](#)]
Guan Yang*, **Minghuan Liu***, Weijun Hong, Weinan Zhang, Fei Fang, Guangjun Zeng, Yue Lin.
(*Equal Contribution)
The 36th Conference on Neural Information Processing Systems. **NeurIPS 2022**.
- We propose PerfectDou, a state-of-the-art DouDizhu AI system that beats all previous algorithms with perfect information distillation and a perfect-training-imperfect-execution.
- 9 Reinforcement Learning with Automated Auxiliary Loss Search.**
Tairan He, Yuge Zhang, Kan Ren, **Minghuan Liu**, Che Wang, Weinan Zhang, Dongsheng Li, Yuqing Yang.
The 36th Conference on Neural Information Processing Systems. **NeurIPS 2022**.
- We propose a principled and universal method for learning better representations with auxiliary loss functions, named Automated Auxiliary Loss Search (A2LS), which automatically searches for top-performing auxiliary loss functions for RL.
- 8 Plan Your Target and Learn Your Skills: Transferable State-Only Imitation Learning via Decoupled Policy Optimization.** [[Project Page](#)]
Minghuan Liu, Zhengbang Zhu, Yuzheng Zhuang, Weinan Zhang, Jun Wang, Yong Yu, Jianye Hao.

The 39th International Conference on Machine Learning. **ICML 2022**.

- We propose Decoupled Policy Optimization (DePO), a novel framework that explicitly decouples the state-to-action mapping policy as a high-level state planner and an inverse dynamics model. DePO allows for transferring to decision problems with different dynamics settings intuitively and generalizing the planner on out-of-demo state region.

7 Goal-Conditioned Reinforcement Learning: Problems and Solutions.

Minghuan Liu, Menghui Zhu, Weinan Zhang.

The 31st International Joint Conference on Artificial Intelligence, Survey Track. **IJCAI 2022**.

- We bring a brief survey for goal-conditioned reinforcement learning, containing the basic challenge, corresponding solutions, and future prospects.

6 Curriculum Offline Imitation Learning.

Minghuan Liu, Hanyue Zhao, Zhengyu Yang, Jian Shen, Weinan Zhang, Li Zhao, Tie-Yan Liu.

The 35th Conference on Neural Information Processing Systems. **NeurIPS 2021**.

- We propose curriculum offline imitation learning (COIL), a simple and practical imitation learning based method for offline reinforcement learning. COIL utilizes an experience picking strategy for imitating from adaptive neighboring policies with a higher return, and improves the current policy along curriculum stages.

5 MapGo: Model-Assisted Policy Optimization for Goal-Oriented Tasks.

Menghui Zhu*, **Minghuan Liu***, Jian Shen, Zhicheng Zhang, Sheng Chen, Weinan Zhang, Deheng Ye, Yong Yu, Qiang Fu, Wei Yang. (*Equal Contribution)

The 30th International Joint Conference on Artificial Intelligence. **IJCAI 2021**.

- We propose MapGo, a model-based framework for goal-oriented RL which involves a novel relabeling strategy FGI and a model-based training module UMPO.

4 Energy-Based Imitation Learning.

Minghuan Liu, Tairan He, Minkai Xu, Weinan Zhang.

The 20th International Conference on Autonomous Agents and Multiagent Systems. **AAMAS 2021**.

- We propose EBIL, a two-step solution for imitation learning: first estimate the energy of expert's occupancy measure, and then take the energy to construct a surrogate reward function as a guidance for the agent to learn the desired policy.

3 Multi-Agent Interactions Modeling with Correlated Policies.

Minghuan Liu, Ming Zhou, Weinan Zhang, Yuzheng Zhuang, Jun Wang, Wulong Liu, Yong Yu.

The 8th International Conference on Learning Representations. **ICLR 2020**.

- We propose CoDAIL, which cast the multi-agent interactions modeling problem into a multi-agent imitation learning framework with explicit modeling of correlated policies by approximating opponents' policies.

2 Towards Applicable Reinforcement Learning: Improving the Generalization and Sample Efficiency with Policy Ensemble.

Zhengyu Yang, Kan Ren, Xufang Luo, **Minghuan Liu**, Weiqing Liu, Jiang Bian, Weinan Zhang, Dongsheng Li.

The 31st International Joint Conference on Artificial Intelligence **IJCAI 2022**.

- We propose Ensemble Proximal Policy Optimization (EPPO), a policy ensemble framework that improves the generalization and sample efficiency, especially in real-world tasks like financial trading.

1 Learning to Build High-fidelity and Robust Environment Models.

Weinan Zhang, Zhengyu Yang, Jian Shen, **Minghuan Liu**, Yimin Huang, Xing Zhang, Ruiming Tang, Zhenguo Li.

The 20th European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases. **ECML-PKDD 2021**.

- We propose robust learning to simulate (RL2S), a new problem of RL that focuses on learning a high-fidelity environment simulator for serving diverse downstream tasks; we further transform

RL2S as a novel robust imitation learning problem and propose efficient algorithms to solve it.

PREPRINTS

5 MADiff: Offline Multi-agent Learning with Diffusion Models.

Zhengbang Zhu*, Minghuan Liu*, Liyuan Mao, Bingyi Kang, Minkai Xu, Yong Yu, Stefano Ermon, Weinan Zhang (*Equal Contribution)

In submission. Arxiv

- We propose MADiff, a novel generative multi-agent learning framework to tackle this problem. MADiff is realized with an attention-based diffusion model to model the complex coordination among behaviors of multiple diffusion agents.

4 Planning Immediate Landmarks of Targets for Skill Transfer across Agents.

Minghuan Liu, Zhengbang Zhu, Menghui Zhu, Yuzheng Zhuang, Weinan Zhang

In submission. Arxiv

- We propose PLoT, i.e., Planning Immediate Landmarks of Targets. PLoT utilizes the universal decoupled policy optimization to learn a goal-conditioned state planner; then, distills a goal-planner to plan immediate landmarks in a model-free style that can be shared among different agents.

3 Is Risk-Sensitive Reinforcement Learning Properly Resolved?

Ruiwen Zhou, Minghuan Liu, Kan Ren, Xufang Luo, Weinan Zhang, Dongsheng Li

In submission. Arxiv

- We prove that the existing RSRL methods do not achieve unbiased optimization and can not guarantee optimality or even improvements regarding risk measures. To remedy this issue, we further propose Trajectory Q-Learning (TQL), for RSRL problems with provable convergence to the optimal policy. Based on our new learning architecture, we are free to introduce a general and practical implementation for different risk measures to learn disparate risk-sensitive policies.

2 Looking Ahead to Avoid Being Late: Solving Hard-Constrained Traveling Salesman Problem

Jingxiao Chen, Ziqin Gong, Minghuan Liu, Yong Yu, Weinan Zhang

In submission

- We aim to derive an effective solution against Traveling Salesman Problem (TSP) with hard constraints. Specifically, we propose a novel learning-based method that uses looking-ahead information as the feature to guarantee the legality of TSP with Time Windows (TSPTW) solutions.

1 RITA: Boost Autonomous Driving Simulators with Realistic Interactive Traffic Flow

Zhengbang Zhu, Yuzheng Zhuang, Yuecheng Liu, Minghuan Liu, Liyuan Mao, Ziqin Gong, Weinan Zhang, Qiang Gu, Bin Wang, Yong Yu

In submission. Arxiv

- We propose Realistic Interactive Traffic flow (RITA) as an integrated component of existing driving simulators to provide high-quality traffic flow for the evaluation and optimization of the tested driving strategies.

AWARDS & HONORS

ByteDance Scholarship (Only 10 in China)	2022
NeurIPS 2022 Top Reviewer	2022
China National Scholarship for Ph.D. (1%)	2022
TOP 1, Ubiquant Special Retro Snake Challenge (Bonus ~ \$10k)	2022
TOP 6, Finalist of Sports Analytics Challenge (sponsored by PSG)	2019
TOP 10, SCADA Data Missing Repair Competition	2019
TOP 3, AI Challenger 2018 in Weather Forecasting	2018
Sishiyanghua Medal (Only 10 in university)	2019

Outstanding Graduate	<i>2019</i>
National First Prize, China Undergraduate Mathematical Contest in Modeling	<i>2017</i>
Meritorious Winner, Mathematical Contest In Modeling	<i>2017</i>
China National Scholarship × 2 (1%)	<i>2016&2017</i>
Tang Lixin Scholarship (1‰)	<i>2017</i>
IBM Scholarship (1‰)	<i>2017</i>
Special Grade Comprehensive Scholarship × 4 (1%)	<i>2016 - 2018</i>

SERVICES

Conference Reviewer: NeurIPS'21'22'23, ICML'21'22'23, ICLR'23

SKILLS

Machine Learning: Pytorch, Jax, Tensorflow, Scikit-Learn, LightGBM

Programming Languages: Python, JavaScript, C / C++, Java, MATLAB

Standard Tests: CET-6(574), CET4(616)

Hobbies and Interests: Soccer, Tennis, Gym, Swimming