

Haoxiang Wang

+1(217)904-6521
✉ haoxiang096@gmail.com
📄 haoxiang-wang.github.io
in haoxiang-wang-071414ab

Education

- 2019–2024 **University of Illinois, Urbana-Champaign, Urbana, IL, USA.**
Ph.D. *Electrical & Computer Engineering* (GPA: 4.0/4.0)
Research Area: *Machine Learning*. Thesis Advisors: Dr. Han Zhao and Dr. Bo Li
- 2015–2019 **University of Illinois, Urbana-Champaign, Urbana, IL, USA.**
B.S. with Highest Distinctions. Double Majors:
(1) *Statistics & Computer Science* (2) *Physics* (GPA: 3.89/4.0)

Professional Experience

- 2023 Summer **Apple AI/ML.**
Research Intern on Machine Learning
Paper: *Merging Vision Foundation Models towards Semantic and Spatial Understanding*
Managers: Dr. Hadi Pouransari & Dr. Oncel Tuzel
- 2022 Summer **Amazon Web Services.**
Research Intern on Quantum Machine Learning
Paper: *Predicting Properties of Quantum Systems with Conditional Generative Models*
Managers: Dr. Cedric Yen-Yu Lin & Dr. Peter Komar
- 2021 Summer **Waymo LLC**, (formerly the Google self-driving car project).
Research Intern on Machine Learning
Project: *Improving Robustness of Perception Systems with Mixture-of-Experts (MoE)*
Manager: Dr. Zhao Chen

Research Interests

Multi-Modal AI, Large Language Models (LLM), Robust and Trustworthy AI, Reinforcement Learning from Human Feedback (RLHF)

I am currently interested in:

- Improving the multi-task capabilities of foundation models (LLM and multi-modal models)
- Enhancing the controllability of foundation models via more fine-grained RLHF

Publications

(* indicates equal contribution.)

Foundation Models (Multi-Modal Models and Large Language Models)

- 2024 **Arithmetic Control of LLMs for Diverse User Preferences: Directional Preference Alignment with Multi-Objective Rewards.**
Haoxiang Wang*, Yong Lin*, Wei Xiong*, Rui Yang, Shizhe Diao, Shuang Qiu, Han Zhao, Tong Zhang
Under review of **ACL**, 2024.
- 2024 **Enhancing Compositional Generalization via Compositional Feature Alignment.**
Haoxiang Wang*, Haozhe Si*, Han Zhao
Under review of Transactions on Machine Learning Research (**TMLR**)
- 2023 **SAM-CLIP: Merging Vision Foundation Models towards Semantic and Spatial Understanding.**
Haoxiang Wang, Pavan Kumar Anasosalu Vasu, Fartash Faghri, Raviteja Vemulapalli, Mehrdad Farajtabar, Sachin Mehta, Mohammad Rastegari, Oncel Tuzel, Hadi Pouransari
Under review of **CVPR**, 2024.

2023 **Reward vs. Tax: Towards Mitigating the Alignment Tax of RLHF.**
Yong Lin, Hangyu Lin, Wei Xiong, Shizhe Diao, Jianmeng Liu, Jipeng Zhang, Rui Pan, Haoxiang Wang, Wenbin HU, Hanning Zhang, Hanze Dong, Han Zhao, Nan Jiang, Heng Ji, Yuan Yao, Tong Zhang
Under review of **ICML**, 2024.

2023 **Speciality vs generality: An empirical study on catastrophic forgetting in fine-tuning foundation models.**

Yong Lin, Lu Tan, Hangyu Lin, Zeming Zheng, Renjie Pi, Jipeng Zhang, Shizhe Diao, Haoxiang Wang, Han Zhao, Yuan Yao, Tong Zhang
Preprint

[Multi-Task Learning and Meta-Learning](#)

2022 **Predicting Properties of Quantum Systems with Conditional Generative Models.**

Haoxiang Wang*, Maurice Weber*, Josh Izaac, Cedric Yen-Yu Lin
Submitted to **Nature Communications**

CVPR 2022 **Global Convergence of MAML and Theory-Inspired Neural Architecture Search for Few-Shot Learning.**

Haoxiang Wang*, Yite Wang*, Ruoyu Sun, Bo Li
Conference on Computer Vision and Pattern Recognition (**CVPR**), 2022.

ICML 2021 **Bridging Multi-Task Learning and Meta-Learning: Towards Efficient Training and Effective Adaptation.**

Haoxiang Wang, Han Zhao, Bo Li
International Conference on Machine Learning (**ICML**), 2021.

[Robust and Trustworthy AI under Data Distribution Shifts](#)

2023 **Gradual Domain Adaptation: Theory and Algorithms.**

Yifei He*, Haoxiang Wang*, Bo Li, Han Zhao
Submitted to Journal of Machine Learning Research (**JMLR**)

2023 **Invariant Feature Subspace Recovery: A New Class of Provable Domain Generalization Algorithms.**

Haoxiang Wang*, Gargi Balasubramaniam*, Haozhe Si, Bo Li, Han Zhao
Submitted to Journal of Machine Learning Research (**JMLR**)

ICML 2022 **Provable Domain Generalization via Invariant-Feature Subspace Recovery.**

Haoxiang Wang, Haozhe Si, Bo Li, Han Zhao
International Conference on Machine Learning (**ICML**), 2022.

ICML 2022 **Understanding Gradual Domain Adaptation: Improved Analysis, Optimal Path and Beyond.**

Haoxiang Wang, Bo Li, Han Zhao
International Conference on Machine Learning (**ICML**), 2022.

UAI 2022 **Future Gradient Descent for Adapting the Temporal Shifting Data Distribution in Online Recommendation System.**

Mao Ye, Ruichen Jiang, Haoxiang Wang, Dhruv Choudhary, Xiaocong Du, Bhargav Bhushanam, Aryan Mokhtari, Arun Kejariwal, Qiang Liu
Conference on Uncertainty in Artificial Intelligence (**UAI**), 2022.

[Optimization](#)

NeurIPS 2019 **Learning Positive Functions with Pseudo Mirror Descent.**

Yingxiang Yang, Haoxiang Wang, Negar Kiyavash, Niao He
Neural Information Processing Systems (**NeurIPS**), 2019. (**Spotlight Presentation**)

Service

2020-Now **Reviewer**, *ICML, NeurIPS, ICLR, ICCV, CVPR, TMLR, AISTATS, AAAI.*

Skills

Programming, *Python, PyTorch, Jax.*