

Guangyuan Hao

☎ (+1) 51 0775 8838 | ✉ guangyuan.hao@connect.ust.hk | 🏠 guangyuanhao.github.io

Education

Hong Kong University of Science and Technology

Hong Kong

Master of Philosophy (MPhil) in Individualized Interdisciplinary Program (Artificial Intelligence)

Sept. 2020 - Sept. 2022

- A two-year research-focused master's program.
- Supervisor: Prof. Dit-Yan Yeung (HKUST)
- Overall GPA: 3.95/4.30; Major GPA: 4.05/4.30

University of Electronic Science and Technology of China

Chengdu, China

Bachelor of Engineering (B.E.) in Information Display and Optoelectronic Technology

Sept. 2012 - Jul. 2016

- Equivalent to a degree in Electronic Engineering.
- GPA: 3.97/4.00
- Major Rank: 1st out of 184 students

Research Experience

MIT, Research Assistant, Prof. Paul Liang

Oct. 2024-Now.

- Researching theoretical foundations of multimodal learning under the mentorship of Prof. Paul Liang (MIT).

Berkeley, ML Alignment & Theory Scholars Program, Research Scholar

Jun. 2024-Aug. 2024

- Conducted AI safety research focused on mitigating backdoor attacks in LLMs in collaboration with Dr. Steven Basart (Center for AI Safety) and Mr. Andy Zou (CMU).
- Participated in workshops, talks, and networking events with the *Berkeley* alignment research community.
- Program acceptance rate: 87 out of 1,221 applicants.

CUHK&MBZUI, Research Assistant, Prof. Kun Zhang (CMU) and Prof. Jiji Zhang (CUHK)

Nov. 2022-Jun. 2024

- Conducted research in causal inference and causal discovery under the mentorship of Prof. Kun Zhang (CMU) and Prof. Jiji Zhang (CUHK), while collaborating with Prof. Biwei Huang (UCSD).

HKUST, MPhil Student, Prof. Hao Wang (Rutgers) and Prof. Dit-Yan Yeung (HKUST)

Jan. 2021-Jun. 2024

- Conducted research in trustworthy AI, with a focus on distribution-shift robustness, under the mentorship of Prof. Hao Wang (Rutgers) and Prof. Dit-Yan Yeung (HKUST).

Papers in Causal Inference & Causal Discovery

Natural Counterfactuals With Necessary Backtracking

Guang-Yuan Hao, Jiji Zhang, Biwei Huang, Hao Wang, Kun Zhang; **Thirty-Eighth Annual Conference on Neural Information Processing Systems (NeurIPS 2024)**; Previous version presented at ICML 2023 Workshop on Counterfactuals in Minds and Machines ([Oral Presentation](#))

A Conditional Independence Test in the Presence of Discretization

Boyang Sun, Yu Yao, **Guang-Yuan Hao**, Yumou Qiu, Kun Zhang; **The Thirteenth International Conference on Learning Representations (ICLR 2025)**; *This paper tackles a challenging causal discovery scenario, though **causal discovery** is not explicitly mentioned in the title.*

Permutation-based Rank Test in the Presence of Discretization and Application in Causal Discovery with Mixed Data

Xinshuai Dong, Ignavier Ng, Boyang Sun, Haoyue Dai, **Guang-Yuan Hao**, Shunxing Fan, Peter Spirtes, Yumou Qiu, Kun Zhang; Submitted to **ICML 2025**

Emerging Synergies in Causality and Deep Generative Models: A Survey

Guanglin Zhou, Shaoran Xie, **Guang-Yuan Hao**, Shiming Chen, Biwei Huang, Xiwei Xu, Chen Wang, Liming Zhu, Lina Yao, Kun Zhang; Submitted to **IEEE Transactions on Neural Networks and Learning Systems**

Papers in Robustness & Applications

Composite Active Learning: Towards Multi-Domain Active Learning with Theoretical Guarantees

Guang-Yuan Hao, Hengguan Huang, Haotian Wang, Jie Gao, Hao Wang; **Thirty-Eighth AAAI Conference on Artificial Intelligence (AAAI 2024)**

Domain-Indexing Variational Bayes: Interpretable Domain Index for Domain Adaptation

Guang-Yuan Hao*, Zihao Xu*, Hao He, Hao Wang (* indices equal contribution); **Eleventh International Conference on Learning Representations (ICLR 2023)** (Spotlight Presentation)

Taxonomy-Structured Domain Adaptation

Tianyi Liu*, Zihao Xu*, Hao He, **Guang-Yuan Hao**, Guang-He Lee, Hao Wang (* indices equal contribution); **Fortieth International Conference on Machine Learning (ICML 2023)**

MIXGAN: Learning Concepts from Different Domains for Mixture Generation

Guang-Yuan Hao, Hongxing Yu, Weishi Zheng; **27th International Joint Conference on Artificial Intelligence (IJCAI 2018)** (Oral Presentation)

DyLex: Incorporating Dynamic Lexicons into BERT for Sequence Labeling

Baojun Wang*, Zhao Zhang*, Kun Xu*, **Guang-Yuan Hao**, Yuyang Zhang, Lifeng Shang, Linlin Li, Xiao Chen, Xin Jiang, Qun Liu (* indicates equal contribution); **2021 Conference on Empirical Methods in Natural Language Processing (EMNLP 2021)**, Long Paper

Honors & Awards

- 2024 **MATS Scholarship**, Awarded to 87 out of 1,221 applicants by the ML Alignment & Theory Scholars (MATS) Program, Berkeley.
- 2020–2022 **Full Postgraduate Studentship**, Awarded by the Hong Kong University of Science and Technology for postgraduate studies.
- 2015 **China National Scholarship**, Awarded to the top 1% of students by the Ministry of Education, China.
- 2015 **Outstanding Graduate Award**, Top 1% of graduates at the University of Electronic Science and Technology of China.
- 2014 **Sekorm First-Class Scholarship**, Awarded to the top 2% of students by SEKORM LIMITED.
- 2013 **People's First-Class Scholarship**, Awarded to the top 10% of students at the University of Electronic Science and Technology of China.

Profession & Skills

Peer Reviewer: ICCV 2023, NeurIPS 2023, ICML 2024

Programming and Related: Linux, Python, PyTorch, TensorFlow, \LaTeX

Hobbies: Long-Distance Running, Boxing, Reading, Meditation