Guohao Sun

School of Information, B. Thomas Golisano College of Computing and Information Sciences
Rochester Institute of Technology
20 Lomb Memorial Drive, Rochester, NY 14623

Email: gs4288@rit.edu   URL: www.guohaosun.com

Areas of specialisation

Research interest: I am drawn to simplicity, strive to understand things deeply, and enjoy building practical systems. Currently, I am fascinated by **multi-modal** foundation models like GPT-4o and CLIP, which can be adapted to a wide range of downstream tasks. My focus is on understanding how these models work, and improving their reasoning ability and robustness with **self-questioning**, **preference optimization** and **evidential knowledge**.
Topics: General-purpose Vision-Language Understanding, Evidential Learning
Skills: Python, PyTorch, C++

Work experience

2019-2021   **Software Engineer**, Robotrak Technologies Co., Ltd., China

Education

2023   **Ph.D. in Computing and Information Science**, Rochester Institute of Technology, **Advisor**: Dr. Zhiqiang Tao
2021   **M.S. in Computer Science**, Santa Clara University, **Advisor**: Dr. Zhiqiang Tao
2014   **B.S. in Computer Science and Engineering**, Michigan State University

Honours & awards

2022   4th place, Hackathon: Hack for Humanity, Santa Clara University
2015   Academic Excellence Award, Michigan State University

Publications

Peer-reviewed Conference articles

2024   ***Guohao Sun***, Can Qin, Huazhu Fu, Linwei Wang, Zhiqiang Tao. "STLLaVA-Med: Self-Training Large Language and Vision Assistant for Medical Question-Answering", *Empirical Methods in Natural Language Processing (EMNLP)*[code]

2024   **Guohao Sun**, Can Qin, Jiamian Wang, Zeyuan Chen, Ran Xu, Zhiqiang Tao. "SQ-LLaVA: Self-Questioning for Large Vision-Language Assistant", *European Conference on Computer Vision (ECCV)*[code]

2024   Cheng Han, Yawen Lu, **Guohao Sun**, James Chenhao Liang, Zhiwen Cao, Qifan Wang, Qiang Guan, Sohail Dianat, Raghuveer Rao, Tong Geng, Zhiqiang Tao, Dongfang Liu. "Prototypical Transformer As Unified Motion Learners", *International Conference on Machine Learning (ICML)*[code]

2024   Jiamian Wang, **Guohao Sun**, Pichao Wang, Dongfang Liu, Sohail Dianat, Majid Rabbani, Raghuveer Rao, Zhiqiang Tao. "Text Is MASS: Modeling as Stochastic Embedding for Text-Video Retrieval", *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*[code][poster][video]

2024   **Guohao Sun**, Yue Bai, Xueying Yang, Yi Fang, Yun Fu, Zhiqiang Tao. "Aligning Out-of-Distribution Web Images and Caption Semantics via Evidential Learning", *The ACM Web Conference (WWW)*[code]