

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: <https://heliossun.github.io/>

Areas of specialisation

Broadly: I am drawn to simplicity, strive to understand things deeply, and enjoy building practical systems. Currently, I am fascinated by foundation models like GPT-4, CLIP, and DALL-E. These models, trained on extensive data using self-supervision at an immense scale, can be adapted to a wide range of downstream tasks. Foundation models represent a paradigm shift in AI development and human-AI interaction. My focus is on understanding how these models work, and improving their intelligence and robustness with multi-modal information.

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
2021 **M.S. in Computer Science**, Santa Clara University
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)*
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)*, <https://github.com/heliossun/SQ-LLaVA>
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)*

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)*

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)*