

SHENYUAN GAO

3rd-year Ph.D. student at HKUST

Homepage: github.com/Little-Podi ◇ Email: sygao@connect.ust.hk

RESEARCH INTEREST

Embodied AI Generalist Robots, Learning from Internet Videos

Generative AI Generalizable World Models, Foundation Models

PUBLICATION

AdaWorld: Learning Adaptable World Models with Latent Actions

Shenyuan Gao, Siyuan Zhou, Yilun Du, Jun Zhang, Chuang Gan

Accepted by ICML 2025.

[\[Paper\]](#) [\[Web\]](#) [\[Code\]](#)

Learning to Act Anywhere with Task-Centric Latent Actions

Qingwen Bu*, Yanting Yang*, Jisong Cai*, **Shenyuan Gao**, Guanghui Ren, Maoqing Yao, Ping Luo, Hongyang Li

Accepted by RSS 2025.

Vista: A Generalizable Driving World Model with High Fidelity and Versatile Controllability

Shenyuan Gao, Jiazhi Yang, Li Chen, Kashyap Chitta, Yihang Qiu, Andreas Geiger, Jun Zhang, Hongyang Li

Accepted by NeurIPS 2024.

[\[Paper\]](#) [\[Web\]](#) [\[Code\]](#) (300+ stars within 2 weeks)

Generalized Predictive Model for Autonomous Driving

Jiazhi Yang*, **Shenyuan Gao***, Yihang Qiu*, Li Chen, Tianyu Li, Bo Dai, Kashyap Chitta, Penghao Wu, Jia Zeng, Ping Luo, Jun Zhang, Andreas Geiger, Yu Qiao, Hongyang Li

Accepted by CVPR 2024 as *Highlight* (**Top 2.8%**).

[\[Paper\]](#) [\[Dataset\]](#)

AgiBot World Colosseo: A Large-Scale Manipulation Platform for Scalable and Intelligent Embodied Systems

Qingwen Bu, Jisong Cai, Li Chen, Xiuqi Cui, Yan Ding, Siyuan Feng, **Shenyuan Gao**, Xindong He, Xu Huang, Shu Jiang, Yuxin Jiang, Cheng Jing, Hongyang Li, Jialu Li, Chiming Liu, Yi Liu, Yuxiang Lu, Jianlan Luo, Ping Luo, Yao Mu, Yuehan Niu, Yixuan Pan, Jiangmiao Pang, Yu Qiao, Guanghui Ren, Cheng Ruan, Jiaqi Shan, Yongjian Shen, Chengshi Shi, Mingkang Shi, Modi Shi, Chonghao Sima, Jianheng Song, Huijie Wang, Wenhao Wang, Dafeng Wei, Chengen Xie, Guo Xu, Junchi Yan, Cunbiao Yang, Lei Yang, Shukai Yang, Maoqing Yao, Jia Zeng, Chi Zhang, Qinglin Zhang, Bin Zhao, Chengyue Zhao, Jiaqi Zhao, Jianchao Zhu

[\[Paper\]](#) [\[Web\]](#) [\[Dataset\]](#) (the most popular robotics dataset on Hugging Face)

Generalized Relation Modeling for Transformer Tracking

Shenyuan Gao, Chunluan Zhou, Jun Zhang

Accepted by CVPR 2023.

[\[Paper\]](#) [\[Code\]](#)

AiATrack: Attention in Attention for Transformer Visual Tracking

Shenyuan Gao, Chunluan Zhou, Chao Ma, Xinggang Wang, Junsong Yuan

Accepted by ECCV 2022.

[\[Paper\]](#) [\[Code\]](#)

Content-Aware Masked Image Modeling Transformer for Stereo Image Compression

Xinjie Zhang, **Shenyuan Gao**, Zhening Liu, Jiawei Shao, Xingtong Ge, Dailan He, Tongda Xu, Yan Wang, Jun Zhang

Accepted by AAAI 2025.

[\[Paper\]](#)

EXPERIENCE

OpenDriveLab

Research Intern

April 2023 - present

Shanghai, China

- Leader: Prof. Hongyang Li
- Worked on foundation models for autonomous driving.
- Developed a driving world model with strong generalization to diverse scenarios and applications.

Hong Kong University of Science and Technology

Research Postgraduate Program

September 2022 - present

Hong Kong SAR, China

- Advisor: Prof. Jun Zhang (IEEE Fellow)
- Worked on object tracking and neural compression.
- Worked on multi-agent perception system.
- Proposed a generalized relation modeling formulation for Transformer-based tracking.
- Proposed a content-aware masked image modeling for bidirectional prior interaction.

Huazhong University of Science and Technology

Final Year Project for Bachelor Degree

March 2022 - June 2022

Wuhan, China

- Advisor: Prof. Peng Yang
- Worked on the intersection of tracking and efficient deep learning.
- Proposed a slimmable tracker with hierarchical weight sharing.
- Awarded as outstanding graduation thesis.

University at Buffalo

Summer Research Intern

July 2021 - March 2022

New York, United States (Remote)

- Advisor: Prof. Junsong Yuan (IEEE Fellow)
- Worked on object tracking.
- Proposed Attention in Attention to facilitate correspondence learning.

The University of Hong Kong, MMLab

HKU CS Summer Research Internship Programme

July 2021 - August 2021

Hong Kong SAR, China (Remote)

- Advisor: Prof. Ping Luo
- Worked on the intersection of tracking and visual grounding.
- Completed the research project successfully with full stipend award.

EDUCATION

Hong Kong University of Science and Technology

Ph.D. in Electronic and Computer Engineering

Advised by Prof. Jun Zhang (IEEE Fellow)

2022 - 2026 (expected)

Huazhong University of Science and Technology

B.Eng. in Electronic Information Engineering

Advanced Class (Elite Program for Information Science, 30/400)

2018 - 2022

GPA: 3.9/4.0

Rank: 1/30

HONORS AND AWARDS

- NeurIPS Top Reviewer 2024
- Full Postgraduate Scholarship 2022-2026
- RedBird PhD Scholarship 2022-2023

- Outstanding Graduate 2022
- Outstanding Graduation Thesis 2022
- Outstanding Undergraduate in Terms of Academic Performance (**Top 1%**) 2019
- National Scholarship (**Top 2%**) 2019

ACADEMIC SERVICES

Conference Reviewer

ICLR, ICML, NeurIPS, RSS, CVPR, ICCV, ECCV, AAAI, AISTATS, ICPR

Journal Reviewer

TMLR, TPAMI, TMM, TCSVT, IMAVIS, PR

Teaching Assistant

COMP 5214: Advanced Deep Learning Architectures (for postgraduate/undergraduate)

ELEC 3100: Signal Processing and Communications (for undergraduate)

CODING LANGUAGE

Python, PyTorch, HTML, \LaTeX

MISC

During my undergraduate, I built and launched a personal blog on my own.

So far, I have posted about 177,000 words of notes and already received 150,000 views from 100,000 unique visitors.

I am a crazy fan of One Piece comics (not its animation).

I own 14 T-shirts with One Piece characters, which allows me to change for 2 weeks without repeating.