Jiageng Mao

maojiageng@gmail.com | 323-620-9207 | Homepage | Google Scholar | Github

Professional Summary

Jiageng is a final-year PhD candidate at the University of Southern California, advised by Prof. Yue Wang. His research centers on physical AI. His goal is to bring AI to the real world, by developing algorithms across robotics, computer vision, and language models. He is particularly interested in training large vision-language(-action) models and world(-action) models. He has held research positions at NVIDIA Research and Google DeepMind. His research is generously supported by a Qualcomm Innovation Fellowship and an NVIDIA Graduate Fellowship.

Education

University of Southern California, Ph.D. in Computer Science — Sept. 2023 – Oct. 2026

- NVIDIA Graduate Fellowship, 2025.
- Qualcomm Innovation Fellowship, 2024.

The Chinese University of Hong Kong, M.Phil. in Electronic Engineering — Sept. 2018 – May 2022

- Hong Kong Government Fellowship, 2018.

Zhejiang University, B.E. in Electronic Engineering — Sept. 2014 – May 2018

- National Scholarships, 2015 & 2016.

Selected Awards

- **NVIDIA Graduate Fellowship**. Awarded to 10 recipients worldwide — 2025
- **Best Paper Award, IROS RoboGen Workshop** — 2025
- **Apple AI/ML Scholar (University Nomination)**. 1 out of 3 students nominated by USC — 2025
- **Google PhD Fellowship (University Nomination)**. 1 out of 4 students nominated by USC — 2025
- **Oral Presentation at ICLR 2025**. Top 1.8% acceptance — 2025
- **Oral Presentation at CoRL 2024**. Top 5% acceptance — 2024
- **OpenAI Researcher Access Grant** — 2024
- **Qualcomm Innovation Fellowship**. Awarded to 12 teams worldwide — 2024
- **Oral Presentation at ICCV 2019**. Top 5% acceptance rate — 2019
- **Hong Kong Government Fellowship**. Awarded to 100 candidates worldwide — 2018
- **National Scholarship**. Top 1% Nationwide — 2015 & 2016
- **Top-10 Outstanding University Students**, Zhejiang University — 2016

Work Experience

- **Google DeepMind** — June 2025–Jan. 2026
  Student Researcher
  - Developed PhysWorld, a framework that unifies **video generation** and **reinforcement learning** in real-to-sim environments, enabling zero-shot generalizable robotic manipulation (First-authored paper accepted to ICRA 2026).
- **NVIDIA Research** — May 2024–Aug. 2024
  Research Intern
  - Developed DreamDrive, a framework that seamlessly unifies **video generation** and **neural simulation** to build large-scale 4D world models for robots (First-authored paper accepted to ICRA 2025 and featured at NVIDIA CES 2025).
- **University of Southern California** — Aug. 2023–Present
  Graduate Research Assistant
  - Mentored 20+ interns and junior PhD students in robotics and language model research. Most students submitted their first-authored papers to top-tier conferences during their internships.
  - Co-developed Ψ_0, a **vision-language-action (VLA) model** that leverages large-scale egocentric human video pre-training for universal humanoid loco-manipulation.

- Developed UH-1, the first **Humanoid Foundation Model** that enables generalizable humanoid pose learning from Internet human videos (First-authored paper accepted to IEEE Humanoids 2025).
- Developed GPT-Driver, the first **LLM-based planning and decision-making** method for autonomous vehicles and robotics (First-authored paper accepted to NeurIPS-W 2024; 300+ citations).
- Developed Agent-Driver, an **LLM agent** for complex perception and decision-making in dynamic embodied AI environments (First-authored paper accepted to CoLM 2024 with Top 1% ratings).

- **Waabi Innovation** Sep. 2022–Mar. 2023
  Research Intern
  - Developed a joint perception and prediction method for autonomous vehicles. The method was integrated into Waabi's product with a US patent filed (First-authored paper accepted to ECCV 2024).
- **The Chinese University of Hong Kong** Aug. 2018–Aug. 2022
  Graduate Research Assistant
  - Developed state-of-the-art point cloud and voxel data processing methods for 3D perception. Published six first-authored papers at top-tier venues (1 CVPR, 3 ICCV, 1 NeurIPS, 1 IJCV).

Publications

Summary

Total Citations: **5000+**, H-index: **20**, i10-Index: **23**

20+ papers published in top-tier robotics, vision, and language venues.

Selected Publications

- Ψ_0: An Open Foundation Model Towards Universal Humanoid Loco-Manipulation
  Songlin Wei*, Hongyi Jing*, Boqian Li*, Zhenyu Zhao*, **Jiageng Mao**, Zhenhao Ni, Sicheng He, Jie Liu, Xiawei Liu, Kaidi Kang, Sheng Zang, Weiduo Yuan, Marco Pavone, Di Huang, Yue Wang
  In submission to Robotics: Science and Systems (**RSS**), 2026
- Humanoid Everyday: A Comprehensive Robotic Dataset for Open-World Humanoid Manipulation
  Zhenyu Zhao*, Hongyi Jing*, Xiawei Liu, **Jiageng Mao**†, Abha Jha, Hanwen Yang, Rong Xue, Sergey Zakharor, Vitor Guizilini, Yue Wang† (†: advising roles)
  International Conference on Robotics and Automation (**ICRA**), 2026
- Robot Learning from a Physical World Model
  Jiageng Mao, Sicheng He, Hao-Ning Wu, Yang You, Shuyang Sun, Zhicheng Wang, Yanan Bao, Huizhong Chen, Leonidas Guibas, Vitor Guizilini, Howard Zhou, Yue Wang
  International Conference on Robotics and Automation (**ICRA**), 2026
- Robot Learning from Any Images
  Jiageng Mao*, Siheng Zhao*, Wei Chow, Zeyu Shangguan, Tianheng Shi, Rong Xue, Yuxi Zheng, Yijia Weng, Yang You, Daniel Seita, Leonidas Guibas, Sergey Zakharov, Vitor Campagnolo Guizilini, Yue Wang
  Conference on Robot Learning (**CoRL**), 2025 (Oral presentation at ICCV Digital Twins Workshop)
- Universal Humanoid Robot Pose Learning from Internet Human Videos
  Jiageng Mao*, Siheng Zhao*, Siqi Song*, Tianheng Shi, Junjie Ye, Mingtong Zhang, Haoran Geng, Jitendra Malik, Vitor Guizilini, Yue Wang
  IEEE International Conference on Humanoid Robots (**HUMANOIDS**), 2025 (Oral presentation)
- DreamDrive: Generative 4D Scene Modeling from Street View Images
  Jiageng Mao, Boyi Li, Boris Ivanovic, Yuxiao Chen, Yan Wang, Yurong You, Chaowei Xiao, Danfei Xu, Marco Pavone, Yue Wang
  International Conference on Robotics and Automation (**ICRA**), 2025 (Oral, featured in NVIDIA's keynote talk at CES 2025)
- PhysBench: Benchmarking and Enhancing Vision-Language Models for Physical World Understanding
  Wei Chow*, **Jiageng Mao***, Boyi Li, Daniel Seita, Vitor Guizilini, Yue Wang
  International Conference on Learning Representations (**ICLR**), 2025 (Oral presentation, Top 1.8%)
- A Language Agent for Autonomous Driving
  Jiageng Mao*, Junjie Ye*, Yuxi Qian, Marco Pavone, Yue Wang
  Conference on Language Modeling (**CoLM**), 2024 (Reviewer ratings 7-7-8-9, Top 1% out of 1000+ submissions)
- RAM: Retrieval-Based Affordance Transfer for Generalizable Zero-Shot Robotic Manipulation
  Yuxuan Kuang*, Junjie Ye*, Haoran Geng*, **Jiageng Mao**, Congyue Deng, Leonidas Guibas, He Wang, Yue Wang
  Conference on Robot Learning (**CoRL**), 2024 (Oral presentation, Top 5%)
- DeTra: A Unified Model for Object Detection and Trajectory Forecasting

Jiageng Mao*, Sergio Casas*, Ben Agro*, Alexander Cui, Thomas Gilles, Raquel Urtasun
European Conference on Computer Vision (**ECCV**), 2024

- Driving Everywhere with Large Language Model Policy Adaptation
Boyi Li, Yue Wang, **Jiageng Mao**, Boris Ivanovic, Sushant Veer, Karen Leung, Marco Pavone
IEEE Conference on Computer Vision and Pattern Recognition (**CVPR**), 2024 (Featured in NVIDIA GTC 2024 & DriveLabs)
- GPT-Driver: Learning to Drive with GPT
Jiageng Mao, Yuxi Qian, Hang Zhao, Yue Wang
Neural Information Processing Systems Workshop (**NeurIPSW**), 2023
- 3D Object Detection for Autonomous Driving: A Comprehensive Survey
Jiageng Mao, Shaoshuai Shi, Xiaogang Wang, Hongsheng Li
International Journal of Computer Vision (**IJCV**), 2023
- One Million Scenes for Autonomous Driving: ONCE Dataset
Jiageng Mao, Minzhe Niu, Chenhan Jiang, Hanxue Liang, Jingheng Chen, Xiaodan Liang, Yamin Li, Chaoqiang Ye, Wei Zhang, Zhenguo Li, Jie Yu, Hang Xu, Chunjing Xu
Neural Information Processing Systems Track on Datasets and Benchmarks (**NeurIPS**), 2021
- Voxel Transformer for 3D Object Detection
Jiageng Mao, Yujing Xue, Minzhe Niu, Haoyue Bai, Jiashi Feng, Xiaodan Liang, Hang Xu, Chunjing Xu
International Conference on Computer Vision (**ICCV**), 2021 (Selected into Stanford CS 231n)
- Pyramid R-CNN: Towards Better Performance and Adaptability for 3D Object Detection
Jiageng Mao, Minzhe Niu, Haoyue Bai, Xiaodan Liang, Hang Xu, Chunjing Xu
International Conference on Computer Vision (**ICCV**), 2021 (Ranking 1st on Waymo Perception Leaderboard (2021.3))
- Interpolated Convolutional Networks for 3D Point Cloud Understanding
Jiageng Mao, Xiaogang Wang, Hongsheng Li
International Conference on Computer Vision (**ICCV**), 2019 (Oral presentation, Top 4%)

Teaching

- Teaching Assistant, CSCI-677 Robot Perception. 2025 Spring
- Teaching Assistant, CSCI-670 Advanced Computer Vision. 2024 Fall
- Teaching Assistant, ENGR-597 Signals and Systems. 2019, 2020, 2021 Spring
- Teaching Assistant, ENGR-197 Multivariable Calculus. 2019, 2020, 2021 Fall

Academic Services

Reviewer

- CVPR, ICCV, ECCV, IJCV, T-PAMI, ICRA, IROS, CoRL, T-RO, NeurIPS, ICML, ICLR, etc.

Workshop Organizer

- CVPR Vision and Language for Robotics Workshop, 2024.
- RSS MultiSensory Robotics with MultiModal Abilities Workshop, 2025.

Program Committee Member

- NeurIPS Foundation Models for Decision Making Workshop, 2023.

Invited Talks

- **Nanyang Technological University**: *Vision, Robots, and the Physical World*. 2026/02
- **Stanford University** (Host: Prof. Gordon Wetzstein): *Vision, Robots, and the Physical World*. 2026/02
- **Peking University** (Host: Prof. Zhouchen Lin): *Vision, Robots, and the Physical World*. 2026/01
- **Google DeepMind** (Host: Howard Zhou & Jie Tan): *Robot Learning from Texts, Images, and Videos*. 2025/10
- **Stanford University** (Host: Dr. Yang You): *Modeling, Understanding, and Interacting with the Physical World*. 2025/07
- **ICRA 2025** Oral Talk: *Generative 4D Scene Modeling from Street View Images*. 2025/05
- **Stanford University** (Host: Prof. Mac Schwager): *Generative 4D Scene Modeling from Street View Images*. 2025/03
- **ICRA 2025** HCRL Workshop: *Universal Humanoid Robot Pose Learning from Internet Human Videos*. 2025/05
- **CMU** (Host: Prof. Ding Zhao): *Universal Humanoid Robot Pose Learning from Internet Human Videos*. 2024/11
- **Qualcomm**: *A Language Agent for Autonomous Driving*. 2024/02
- **Cornell University** (Host: Prof. Kilian Weinberger), *A Language Agent for Autonomous Driving*. 2023/10

- **ICCV 2019** Oral Talk, *Interpolated Convolutional Networks for 3D Point Cloud Understanding.* 2019/11

Mentorship

Graduate Students

- Weiduo Yuan (Publications: IROS'26 submission (student first author)) 2026
- Hongyi Jing (Publications: ICRA'26 (student first author)) 2025
- Emily Yueting Jia (Publications: ICCV'25 (student first author)) 2025
- Rong Xue (Publications: ICRA'26 (student first author)) 2024
- Zhiyuan Gao (Publications: NeurIPS'25 (student first author)) 2024
- Tianheng Shi (Publications: IEEE Humanoids'25, CoRL'25) 2024
- Yujing Xue (Publications: ICCV'21, CVPR'22 (both as student first author)) 2021

Undergraduate Students

- Jiajun Xu (Publications: ECCV'26 submission (student first author)) 2026
- Yanru Wu (Publications: IROS'26 submission (student first author)) 2026
- Zhenyu Zhao (Publications: ICRA'26 (student first author)) 2025
- Wei Chow (Publications: ICLR'25 Oral (student first author)) 2024
- Siqi Song (Publications: IEEE Humanoids'25 (student first author)) 2024
- Chuye Hong (Publications: IEEE Humanoids'25) 2024
- James Yuxi Qian (Honorable Mention for CRA URA program. Publications: NeurIPS-W'23, CoLM'24) 2023

Skills

Python, C++, CUDA, PyTorch, Isaac Gym/Lab, Docker, Large-scale Distributed Training (Slurm, FSDP, DeepSpeed)