

Jiageng Mao

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

Education

University of Southern California , Ph.D. in Computer Science	Sept 2023 – now
• 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

Publications

Summary

Total Citations: 4369, H-index: 20, i10-Index: 23

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

Selected Publications

- 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*, Sisi 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 (NeurIPS-W), 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
- Point2Seq: Detecting 3D Objects as Sequences
Yujing Xue*, **Jiageng Mao***, Minzhe Niu, Hang Xu, Michael Bi Mi, Wei Zhang, Xiaogang Wang, Xinchao Wang
IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2022
- 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%)

Working Experiences

- | | |
|---|---------------------|
| • Google DeepMind
Student Researcher | June 2025–Jan. 2026 |
| • Developed PhysWorld, a robot learning framework that unifies video generation and physics simulation , enabling scalable digital twin generation and robotic manipulation. The first-authored paper was submitted to ICRA 2026. | |
| • NVIDIA Research
Research Intern | May 2024–Aug. 2024 |
| • Developed DreamDrive, a framework that seamlessly unifies video generation , 4D reconstruction , and neural simulation for large-scale 4D scene generation. The first-authored paper was accepted to ICRA 2025. | |
| • University of Southern California
Graduate Research Assistant | Aug. 2023–Now |
| • Developed UH-1, the first humanoid robot learning framework that enables learning from massive human videos. The first-authored paper was accepted to IEEE Humanoids 2025. | |
| • Developed GPT-Driver, the first LLM-based planning and decision-making method for autonomous vehicles and robotics. The paper was accepted to NeurIPS-W 2024 and received 300+ citations. | |
| • Developed Agent-Driver, an LLM agent for complex perception and decision-making in dynamic embodied AI environments. The first-authored paper was accepted to CoLM 2024 with Top 1% ratings (7, 7, 8, 9). | |
| • Waabi Innovation
Research Intern | Sep. 2022–Mar. 2023 |
| • Developed a joint perception and prediction method for autonomous vehicles. The method was integrated into Waabi's product with a filed US patent. The first-authored paper was accepted to ECCV 2024. | |
| • The Chinese University of Hong Kong
Graduate Research Assistant | Aug. 2018–Aug. 2022 |

- Developed state-of-the-art point cloud and voxel data processing methods for 3D perception. **6 first-authored papers** were accepted at top-tier venues (1 CVPR, 3 ICCV, 1 NeurIPS, 1 IJCV).

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

• Google DeepMind (Host: Howard Zhou & Jie Tan): <i>Robot Learning from Texts, Images, and Videos.</i>	2025/10
• Stanford University (Host: Dr. Yang You): <i>Modeling, Understanding, and Interacting with the Physical World.</i>	2025/07
• ICRA 2025 Oral Talk: <i>Generative 4D Scene Modeling from Street View Images.</i>	2025/05
• Stanford University (Host: Prof. Mac Schwager): <i>Generative 4D Scene Modeling from Street View Images.</i>	2025/03
• ICRA 2025 HCRL Workshop: <i>Universal Humanoid Robot Pose Learning from Internet Human Videos.</i>	2025/05
• CMU (Host: Prof. Ding Zhao): <i>Universal Humanoid Robot Pose Learning from Internet Human Videos.</i>	2024/11
• Qualcomm: <i>A Language Agent for Autonomous Driving.</i>	2024/02
• Cornell University (Host: Prof. Kilian Weinberger), <i>A Language Agent for Autonomous Driving.</i>	2023/10
• ICCV 2019 Oral Talk, <i>Interpolated Convolutional Networks for 3D Point Cloud Understanding.</i>	2019/11

Mentorship

Graduate Students

• Emily Yueling Jia (Publications: ICCV'25 (student first author))	2024-2025
• Rong Xue (Publications: ICRA'26 submission (student first author))	2024-2025
• Zhiyuan Gao (Publications: NeurIPS'25 (student first author))	2024-2025
• Tianheng Shi (Publications: IEEE Humanoids'25, CoRL'25)	2024-2025
• Yujing Xue (Publications: ICCV'21, CVPR'22 (both as student first author))	2020-2021

Undergraduate Students

• James Yuxi Qian (Honorable Mention for CRA URA program. Publications: NeurIPS-W'23, CoLM'24)	2023-2024
• Wei Chou (Publications: ICLR'25 Oral (student first author))	2024-2025
• Siqi Song (Publications: IEEE Humanoids'25 (student first author))	2024-2025
• Chuye Hong (Publications: IEEE Humanoids'25)	2024-2025