Zixian Ma

zixianma@cs.washington.edu | https://zixianma.github.io

EDUCATION

University of Washington
PhD in Computer Science and Engineering *9/2023 – now*
Research Areas: Multi-modal foundation models, human-AI interaction

Stanford University
BS with Honors and MS in Computer Science; Minor in Biology; GPA: 3.97 *9/2018 – 6/2022*
Courses: Artificial Intelligence: Principles and Techniques, Machine Learning, Natural Language Understanding, From Languages to Information, Convolutional Neural Network for Visual Recognition, Reinforcement Learning, Machine Learning with Graphs, Introduction to Human-Computer Interaction Design, Web Applications, Virtual People

AWARDS

TwoSigma PhD Fellowship Winner *2026*
Qualcomm Innovation Fellowship Winner *2025*
The Firestone Medal for Excellence in Undergraduate Research (top 10% of Honors Theses across all disciplines) *2022*
The Ben Wegbreit Prize for Undergraduate CS Research (The CS Department's Best Honors Thesis award) *2022*

RESEARCH EXPERIENCE

UW Raivn Lab with Prof. Ranjay Krishna, Prof. Ali Farhadi *Sep 2023 – now*
Stanford Vision and Learning Lab with Prof. Li Fei-Fei *Apr 2020 – Jun 2022*
Stanford Human-Computer Interaction Group with Prof. Michael Bernstein *Apr 2022 – Sep 2022*

INDUSTRY EXPERIENCE

AI2 Asta Team (Scientific Discovery), with Prof. Daniel Weld and Joseph Chang *Jun 2026 – Sep 2026*
AI2 PRIOR Team (Computer Vision), with Prof. Ranjay Krishna *Jun 2025 – Sep 2025*
Salesforce AI Research – Research Intern with Shelby Heinecke and Juan Carlos Niebles *Jun 2024 – Sep 2024*
Google Research – Research Intern *Aug 2022 – Nov 2022*
Meta – Software Engineer @ Facebook Feed Ranking *Dec 2022 – May 2023*
Facebook (currently Meta) – Software Engineering Intern *Jun 2021 – Aug 2021*

PUBLICATIONS

VFIG: Vectorizing Complex Figures in SVG with Vision-Language Models *2026*
Qijia He*, Xunmei Liu*, Hammaad Memon*, Ziang Li,* **Zixian Ma***, Jaemin Cho, Jason Ren, Daniel S Weld, Ranjay
NeurIPS 2026 Datasets and Benchmarks

JobBench: Aligning Agent Work With Human Will *2026*
Yuetai Li, Yichen Feng, Zhangchen Xu, **Zixian Ma**, Kaiyuan Zheng, Fengqing Jiang, Xinghua Sun, Rulin Shao, Zichen Chen, Yue Huang, Xinyang Han, Brian Lee, Kayla Xu, Shenglai Zeng, Hang Hua, Xiangliang Zhang, Basel Alomair, Ranjay Krishna, Luke Zettlemoyer, Pang Wei Koh, Bhaskar Ramasubramanian, Luyao Niu, Xiang Yue, Radha Poovendran
NeurIPS 2026 Datasets and Benchmarks

MolmoWeb: Open Visual Web Agent and Open Data for the Open Web *2026*
Tanmay Gupta*, Piper Wolters*, **Zixian Ma***, Peter Sushko*, Rock Yuren Pang, Yue Yang, Jason Ren, Harsh Trivedi, Taira Anderson, Winson Han, Ranjay Krishna
ECCV 2026

MolmoPoint: Better Pointing for VLMs with Grounding Tokens *2026*
Christopher Clark, Yue Yang, Jae Sung Park, **Zixian Ma**, Jieyu Zhang, Rohun Tripathi, Mohammadreza Salehi, Sangho Lee, Taira Anderson, Winson Han, Ranjay Krishna
ECCV 2026 | Oral Presentation

Molmo2: Open Weights and Data for Vision-Language Models with Video Understanding and Grounding *2026*
Christopher Clark*, Jieyu Zhang*, **Zixian Ma***, Jae Sung Park, Mohammadreza Salehi, Rohun Tripathi, Sangho Lee, Zhongzheng Ren, Chris Dongjoo Kim, Yinuo Yang, Vincent Shao, Yue Yang, Weikai Huang, Ziqi Gao, Taira Anderson, Jianrui Zhang, Jitesh Jain, George Stoica, Winson Han, Ali Farhadi, Ranjay Krishna
CVPR 2026 | Oral Presentation & Best Paper Candidate

SAGE: Training Smart Any-Horizon Agents for Long Video Reasoning with Reinforcement Learning *2026*
Jitesh Jain, Jialuo Li, **Zixian Ma**, Jieyu Zhang, Chris Dongjoo Kim, Sangho Lee, Rohun Tripathi, Tanmay Gupta, Christopher Clark, Humphrey Shi
CVPR 2026

Unfolding Spatial Cognition: Evaluating Multimodal Models on Visual Simulations *2026*
Linjie Li, Mahtab Bigverdi, Jiawei Gu, **Zixian Ma**, Yinuo Yang, Ziang Li, Yejin Choi, Ranjay Krishna
ICLR 2026

LATTE: Learning to Think with Vision Specialists *2025*
Zixian Ma, Jianguo Zhang, Zhiwei Liu, Jieyu Zhang, Juntao Tan, Manli Shu, Juan Carlos Niebles, Shelby Heinecke, Huan Wang, Caiming Xiong, Ranjay Krishna, Silvio Savarese
EMNLP 2025 | Oral Presentation

Synthetic Visual Genome *2025*
Jae Sung Park, **Zixian Ma**, Linjie Li, Chenhao Zheng, Cheng-Yu Hsieh, Ximing Lu, Khyathi Chandu, Quan Kong, Norimasa Kobori, Ali Farhadi, Yejin Choi, Ranjay Krishna
CVPR 2025

Coarse Correspondence Elicit 3D Spacetime Understanding in Multimodal Language Model *2025*
Benlin Liu*, Yuhao Dong*, Yiqin Wang*, **Zixian Ma**, Yansong Tang, Luming Tang, Yongming Rao, Wei-Chiu Ma, Ranjay Krishna
CVPR 2025

m&m's: A Benchmark to Evaluate Tool-Use for multi-step multi-modal Tasks *2024*
Zixian Ma, Weikai Huang, Jieyu Zhang, Tanmay Gupta, Ranjay Krishna
ECCV 2024 | Synthetic Data for Computer Vision Workshop

Task Me Anything *2024*
Jieyu Zhang, Weikai Huang*, **Zixian Ma***, Oscar Michel, Dong He, Tanmay Gupta, Wei-Chiu Ma, Ali Farhadi, Aniruddha Kembhavi, Ranjay Krishna
NeurIPS 2024 Datasets and Benchmarks | Oral Presentation @ Video-Language Models Workshop

NaturalBench: Evaluating Vision-Language Models on Natural Adversarial Samples *2024*
Baiqi Li*, Zhiqiu Lin*, Wenxuan Peng*, Jean de Dieu Nyandwi*, Daniel Jiang, **Zixian Ma**, Simran Khanuja, Ranjay Krishna, Graham Neubig, Deva Ramanan
NeurIPS 2024 Datasets and Benchmarks

CREPE: Can Foundation Vision-Language Models Reason Compositionally? *2023*
Zixian Ma*, Jerry Hong*, Mustafa Omer Gul*, Mona Gandhi, Irena Gao, Ranjay Krishna
CVPR 2023 | Highlight

SugarCREPE: Fixing Hackable Benchmarks for Vision-Language Compositionality *2023*
Cheng-Yu Hsieh*, Jieyu Zhang*, **Zixian Ma**, Aniruddha Kembhavi, Ranjay Krishna
NeurIPS 2023 Datasets and Benchmarks

ELIGN: Expectation Alignment as a Multi-Agent Intrinsic Reward *2022*
Zixian Ma, Rose E. Wang, Li Fei-Fei, Michael Bernstein, Ranjay Krishna
NeurIPS 2022

Model Sketching: Centering Concepts in Early-Stage Machine Learning Model Design *2022*
Michelle Lam, **Zixian Ma,** Anne Li, Izequiel Freitas, Dakuo Wang, James Landay, Michael Bernstein
CHI 2023

OpenAttack: An Open-source Textual Adversarial Attack Toolkit *2021*
Guoyang Zeng, Fanchao Qi, Qianrui Zhou, Tingji Zhang, **Zixian Ma**, Bairu Hou, Yuan Zang, Zhiyuan Liu, Maosong Sun
The Joint Conference of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th
ACL 2021: System Demonstrations

SKILLS

Technical
Python, Pytorch, TensorFlow, C++, C, R, HTML&CSS, JavaScript, React, SQL
Language
English, Chinese (Mandarin and Cantonese)