Zichen "Charles" Zhang

(+1) 651-600-5947 | charlesz@allenai.org | zcczhang.github.io | zcczhang | @ZCCZHANG | Google Scholar

Education

Macalester College, St. Pual, MN *Sep. 2019 - Dec. 2022*

GPA 3.95/4.0 **MAJORS GPA** 4.0/4.0 *Mathematics, C.S.*

- Charles J. Turck Presidential Honor Scholarship (Four-year scholarship); School Dean's List (2019 - 2022)

Research Interest

My research lies at the intersection of **vision**, **learning**, and **robotics**. My ambition is to build a unified, general-purpose agent capable of understanding and interacting within the multimodal world. Such an agent should be capable of mobile manipulation with reasoning and planning using multi-sensory inputs, assisting and communicating with humans in natural ways, and lifelong autonomous and adaptive learning.

Publications / Preprints

* indicates equal contribution, † indicates equal advising

Video2Sim2Real: Articulated Scenes from Large-Scale Videos for Universal Mobile Manipulator
***Z. Zhang**, et al.*
To Be Submitted to Robotics: Science and Systems (RSS), 2025.

PoliFormer: On-Policy RL with Transformers Results in Masterful Navigators
*K. Zeng, **Z. Zhang**, K. Ehsani, R. Hendrix, J. Salvador, A. Herrasti, R. Girshick, A. Kembhavi, L. Weihs*
Conference on Robot Learning (CoRL), 2024. **Outstanding Paper Award** [website] [pdf]

UNIFIED-IO 2: Scaling Autoregressive Multimodal Model with Vision, Language, Audio, and Action
J. Lu, C. Clark*, S. Lee*, **Z. Zhang*** (Leading Authors), S. Khosla, R. Marten, D. Hoiem, and A. Kembhavi*
IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2024. **Highlight** (2.8%) [website] [pdf]

Universal Visual Decomposer: Long-Horizon Manipulation Made Easy
***Z. Zhang**, Y. Li*, O. Bastani, A. Gupta, D. Jayaraman, Y. Ma†, and L. Weihs†*
IEEE International Conference on Robotics and Automation (ICRA), 2024. **Best Paper Award in Robot Vision Finalist**; Also **Best Paper** at CoRL LEAP workshop and **oral** (6/112) at NeurIPS FMDM workshop, 2023. [website] [pdf]

Open X-Embodiment: Robotic Learning Datasets and RT-X Models
Open X-Embodiment Collaboration
IEEE International Conference on Robotics and Automation (ICRA), 2024. **Best Paper Award** [website] [pdf]

When Learning Is Out of Reach, Reset: Generalization in Autonomous Visuomotor Reinforcement Learning
***Z. Zhang** and L. Weihs.*
Preprint, arXiv:2303.17600 (2024). Also lightning talk at CoRL OOD workshop, 2023. [website] [pdf]

VIMA: Robot Manipulation with Multimodal Prompts
Y. Jiang, A. Gupta, **Z. Zhang***, G. Wang*, Y. Dou, Y. Chen, L. Fei-Fei, A. Anandkumar, Y. Zhu†, and L. Fan†*
International Conference on Machine Learning (ICML), 2023; Also **oral** at NeurIPS FMDM workshop 2022. [website] [pdf]

Characterization of Rectifiable Measures Carried by Lipschitz Curves
***Z. Zhang**, Y. Wu, and L. Naples*
The Joint Mathematics Meetings (JMM) Contributed Paper Session, AMS-PME Poster Session, 2022. [pdf]

Research Experience

Predoctoral Young Investigator (PYI) *Seattle, WA*

PRIOR team, Allen Institute for AI (AI2) *Dec. 2022 - Present*

- Leading the ongoing real2sim2real project, which creates digital twins from web-scale indoor videos, and employs mobile manipulators in these reconstructed physical simulators for training general manipulation agents that can be deployed in the wild.
- Co-led the model architecture development and the initial large-scale training, resulting in the recipe for the project *PoliFormer*.
- Co-led UNIFIED-IO 2, the first large multimodal generation model unifying over 200 understanding and generation tasks (and augmentations) across vision, language, audio, video, and action, trained from scratch.
- Led the project *Universal Visual Decomposer (UVD)*, with the collaboration of people in Upenn and UW.

Research
Remote

NVIDIA & Stanford Vision and Learning Lab (SVL) & UT Robot Perception and Learning Lab *May. 2022 - Oct. 2022*

- Joined the research with people from NVIDIA and Stanford Vision & Learning (SVL), supervised by Linxi "Jim" Fan and Yuke Zhu.
- Developed and improved the multi-modal prompt-driven robotic manipulation tasks suite, VIMABench. Take part in a variety of foundation model implementations, initial experiments, and detection modules for object-centric manipulation for VIMA.
- Participated in developing skill primitives in MineDojo (precursor of Voyager); collaborated with the project for few-shot imitation learning using contrastive learning for analogy making.

Research (Intern)
Seattle, WA

PRIOR team, Allen Institute for AI (AI2) *Sep. 2021 - Dec. 2022*

- Led the project for Autonomous RL (aka. reset-free RL) supervised by *Luca Weihs*. After identifying the main problem was near-irreversible (NI) states, we built the agent towards only request intervention (reset) when necessary, while nailing generalizations.
- Built mobile manipulation benchmark Stretch-P&P in AI2-THOR; Contributed to the distributed RL framework *AllenAct*.

Summer Research
St. Paul, MN

Advised by Professor Lisa Naples, Macalester College *Jun. 2021 - Aug. 2021*

- Led to develop and prove theorems and lemmas to extend the characterization of geometric measures that are carried by rectifiable curves in the dyadic cube system.

Research (Intern)
Beijing, China

R & D Department, Thorough Images (Now Thorough Future) *Jan. 2021 - May 2021*

- Led the project of automated scoring systems for human epidermal growth factor receptor 2 (HER-2) after immunohistochemical (IHC) staining mentored by co-CEO & CTO *Shuhao Wang.*

Summer Research
St. Paul, MN

Advised by Prof. Esra Kadioglu Urtis, Macalester College *Jun. 2020 - Aug. 2020*

- Led the development of Q-learning-based and graph-based algorithms with simulations for UAVs coverage.

Service

Conference/Workshop Reviewer

RA-L 2024, NeurIPS 2024, CVPR 2024, ICRA 2024, NeurIPS 2023, CoRL 2023

Talks

05/2024	"Universal Visual Decomposer" and "Unified-IO 2" at Tsinghua University hosted by Huaping Liu
01/2024	"Unified-IO 2" at Google Deepmind with Jiasen Lu, hosted by Fei Xia and Brian Ichter.
04/2023	"Autonomous Visuomotor Reinforcement Learning" at Ranjay Krishna group.

Teaching

Macalester College	Comp 484 Intro to Artificial Intelligence, MATH 378 Complex Analysis, STAT/COMP 112 Intro to Data Science, COMP 394 Topics Course: Reinforcement Learning, and COMP 128 Data Structure

References

Alphabetical order

Abhishek Gupta, *Assistant Professor, University of Washington, abhgupta@cs.washington.edu*
Aniruddha Kembhavi, *Senior Director of Computer Vision, Allen Institute for AI, anik@allenai.org*
Jiasen Lu, *Senior Research Scientist, Allen Institute for AI (now Research Scientist at Apple), jiasenl@allenai.org*
Linxi "Jim" Fan, *Senior Research Scientist, NVIDIA, linxif@nvidia.com*
Luca Weihs, *Research Manager, Allen Institute for AI, lucaw@allenai.org*
Shuhao Wang, *CTO, co-CEO, Thorough Future, eric.wang@thorough.ai*
Susan Fox, *DeWitt Wallace Professor and Department Chair, Macalester College, fox@macalester.edu*
Yuke Zhu, *Assistant Professor, UT Austin, yukez@cs.utexas.edu*