

Mu Cai

CONTACT INFORMATION

1210 W Dayton St
UW–Madison, CS Department
Madison, WI 53706 USA

Phone: (510) 280-4206
E-mail: mucai@cs.wisc.edu
WWW: <https://pages.cs.wisc.edu/~mucai/>

RESEARCH INTERESTS

My research interest lies in the intersection of deep learning and computer vision. I am especially interested in multimodal generative models, visual prompting, video and 3D understanding.

EDUCATION

University of Wisconsin–Madison, Madison, WI, USA

Ph.D. Student, Computer Sciences, 2020-present

- Advisor: Prof. Yong Jae Lee
- GPA: 4.0/4.0

Xi'an Jiaotong University (XJTU), Xi'an, China

Bachelor in Electrical Engineering, July, 2020

Minor in Artificial Intelligence, July, 2020

- GPA: 4.05/4.3 (93.54/100, rank 1/350)
- Outstanding Student Pacesetter (highest honor, 10 out of ~4,000 undergraduates)
- Advisor: Prof. Badong Chen

PUBLICATIONS & MANUSCRIPTS

Jianrui Zhang*, **Mu Cai***, Tengyang Xie, Yong Jae Lee. CounterCurate: Enhancing Physical and Semantic Visio-Linguistic Compositional Reasoning via Counterfactual Examples. *arXiv* 2024. (*equal contribution)

Mu Cai, Haotian Liu, Siva Karthik Mustikovela, Gregory P. Meyer, Yuning Chai, Dennis Park, Yong Jae Lee. Making Large Multimodal Models Understand Arbitrary Visual Prompts. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024.

Mu Cai*, Zeyi Huang*, Yuheng Li, Haohan Wang, Yong Jae Lee. Leveraging Large Language Models for Scalable Vector Graphics-Driven Image Understanding. *arXiv* 2023. (*equal contribution)

Yuexiang Zhai, Shengbang Tong, Xiao Li, **Mu Cai**, Qing Qu, Yong Jae Lee, Yi Ma. Investigating the catastrophic forgetting in multimodal large language models. *Conference on Parsimony and Learning* (Proceedings Track, oral), 2023.

Zeyi Huang, Andy Zhou, Zijian Ling, **Mu Cai**, Haohan Wang, Yong Jae Lee. A Sentence Speaks a Thousand Images: Domain Generalization through Distilling CLIP with Language Guidance. In *IEEE/CVF International Conference on Computer Vision (ICCV)*, 2023.

Mu Cai, and Yixuan Li. Out-of-distribution Detection via Frequency-regularized Generative Models. In *IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, 2023, **Spotlight**

Haotian Liu, **Mu Cai**, and Yong Jae Lee. Masked Discrimination for Self-Supervised Learning on Point Clouds”. In *Proceedings of the European Conference on Computer Vision (ECCV)*, 2022.

Xuefeng Du, Zhaoning Wang, **Mu Cai**, and Yixuan Li. VOS: Learning What You Don’t Know by Virtual Outlier Synthesis”. In *International Conference on Learning Representations (ICLR)*, 2022.

Mu Cai, Hong Zhang, Huijuan Huang, Qichuan Geng, Yixuan Li, Gao Huang. Frequency Domain Image Translation: More Photo-realistic, Better Identity-preserving. In *IEEE/CVF International Conference on Computer Vision (ICCV)*, 2021.

Liting Sun*, **Mu Cai***, Wei Zhan, and Masayoshi Tomizuka. A Game-Theoretic Policy-Aware Interaction Strategy with Validation on Real Traffic Data. In *IEEE/RSJ Intelligent Robots and Systems Conference (IROS)*, 2020. (*equal contribution)

EXPERIENCE

Cruise LLC, Sunnyvale, CA, USA

Research Intern

May, 2023 - December, 2023

Research project: Submit a paper that *lets large multimodal models understand arbitrary visual prompts* to a top vision venue.

Supervised by Prof Yong Jae Lee, Dr. Dennis Park, and Dr. Siva Karthik Mustikovela

QCraft Inc, San Jose, CA, USA

Research Intern

May, 2022 - March, 2023

Research project: Submit a paper about *self-supervised learning for LiDAR point clouds* to a venue.

Supervised by Dr. Xiaodong Yang

Kuaishou(Kwai) Technology, Beijing, China

Research Intern

June, 2020 - November, 2020

Research project: image translation

Supervised by Prof. Gao Huang and Hong Zhang

SenseTime Research, Beijing, China

Research Intern

December, 2019 - June, 2020

Supervised by Prof. Jifeng Dai

University of California, Berkeley, Berkeley, CA, USA

Visiting Student Researcher

July, 2019 - December, 2019

Research project: autonomous driving motion planning

Supervised by Prof. Masayoshi Tomizuka

PROFESSIONAL ACTIVITIES

• Conference Reviewer:

- Special Interest Group on Computer Graphics and Interactive Techniques (SIGGRAPH), 2024
- International Conference on Machine Learning (ICML), 2024
- International Conference on Learning Representations (ICLR), 2023
- Computer Vision and Pattern Recognition (CVPR), 2023-24
- International Conference on Computer Vision (ICCV), 2023
- European Conference on Computer Vision (ECCV), 2022, 2024
- Neural Information Processing Systems (NeurIPS), 2023
- AAAI Conference on Artificial Intelligence (AAAI), 2023-24
- International Conference on 3D Vision (3DV), 2022
- The International Joint Conference on Artificial Intelligence (IJCAI), 2023
- IEEE CVF Winter Conference on Applications of Computer Vision (WACV), 2023-24

• Journal Reviewer:

- IEEE Transactions on Multimedia (TMM), 2022-23

- **Teaching Assistant:**

- CS639 Deep Learning for Computer Vision, UW Madison, Spring 2023
- CS400 Programming III (Java), UW Madison, Fall 2020, Spring 2021, Fall 2021

HONORS AND
AWARDS

- Qualcomm Innovation Fellowship Finalist, 2024
- UW-Madison, Computer Sciences Department Summer Research Award, 2021
- XJTU, First Prize, Huawei Large-scale Image Classification Competition, 2020
- XJTU, Outstanding Graduate, 2020
- XJTU, Thanksgiving Chinese Scholarship for Modern and Contemporary Scientists, 2019
- XJTU, Gold Prize, "Tengfei Cup" Student Technology Innovation Competition, 2018-2019
- XJTU, Ultra-High Voltage (UHV) Scholarship (Top 1%, highest honor in School of EE), 2018
- XJTU, First Prize, XJTU Mathematical Contest in Modeling, 2018
- XJTU, First Prize, The Chinese Mathematics Competitions, 2017

SKILLS

- Programming: Python, C/C++, Matlab, Java, SQL
- Misc: PyTorch, L^AT_EX, Markdown
- OS: Linux, macOS, Windows