

# Haotian Liu

---

## CONTACT INFORMATION

UW–Madison, CS Department  
Madison, WI 53706 USA

E-mail: [liuhaotian.cn@gmail.com](mailto:liuhaotian.cn@gmail.com)  
WWW: [hliu.cc](http://hliu.cc)

## RESEARCH INTERESTS

I am generally interested in computer vision and machine learning. My recent focus is on **building steerable large models**. The first baby is [LLaVA](#).

## EDUCATION

**University of Wisconsin–Madison**, Madison, WI, USA

Ph.D., Computer Science, 2021-2024  
Doctoral Minor in Quantitative Biology

- Advisor: Prof. Yong Jae Lee

**University of California, Davis**, Davis, CA, USA

Ph.D. Student, Computer Science (transferred), 2019-2021

- Advisor: Prof. Yong Jae Lee

**Zhejiang University, Chu Kochen Honors College**, Zhejiang, China

Bachelor of Computer Science and Technology (Honors), June, 2019

Dual Degree of English Language and Literature, June, 2019

- Advisors: Prof. Xiaogang Jin and Prof. Fei Wu
- GPA: **3.94**/4; Ranking: **Top 5%**

## PUBLICATIONS

**Haotian Liu**, Chunyuan Li, Yuheng Li, Yong Jae Lee. Improved Baselines with Visual Instruction Tuning. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024.

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. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024.

Thao Nguyen, Utkarsh Ojha, Yuheng Li, **Haotian Liu**, Yong Jae Lee. Edit One for All: Interactive Batch Image Editing. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024.

**Haotian Liu\***, Chunyuan Li\*, Qingyang Wu, Yong Jae Lee. Visual Instruction Tuning. *Neural Information Processing Systems (NeurIPS)*, 2023. (\*equal contribution) ([oral presentation](#), 0.6% acceptance rate)

Chunyuan Li\*, Cliff Wong\*, Sheng Zhang\*, Naoto Usuyama, **Haotian Liu**, Jianwei Yang, Tristan Naumann, Hoifung Poon, Jianfeng Gao. LLaVA-Med: Training a Large Language-and-Vision Assistant for Biomedicine in One Day. *Neural Information Processing Systems (NeurIPS)*, Datasets and Benchmarks Track, 2023. ([spotlight presentation](#))

Yadong Lu, Chunyuan Li, **Haotian Liu**, Jianwei Yang, Jianfeng Gao, Yelong Shen. An Empirical Study of Scaling Instruct-Tuned Large Multimodal Models. *Neural Information Processing Systems (NeurIPS)*, 2023, Workshop on Instruction Tuning and Instruction Following.

**Haotian Liu**, Kilho Son, Jianwei Yang, Ce Liu, Jianfeng Gao, Yong Jae Lee\*, Chunyuan Li\*. Learning Customized Visual Models with Retrieval-Augmented Knowledge. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023. (\*equal advising) ([highlight](#), 2.5% acceptance rate)

Yuheng Li, **Haotian Liu**, Qingyang Wu, Fangzhou Mu, Jianwei Yang, Jianfeng Gao, Chunyuan Li\*, Yong Jae Lee\*. GLIGEN: Open-Set Grounded Text-to-Image Generation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023. (\*equal advising)

Chunyuan Li\*, **Haotian Liu\***, Liunian Harold Li, Pengchuan Zhang, Jyoti Aneja, Jianwei Yang, Ping Jin, Yong Jae Lee, Houdong Hu, Zicheng Liu, and Jianfeng Gao. ELEVATER: A Benchmark and Toolkit for Evaluating Language-Augmented Visual Models. In *Neural Information Processing Systems (NeurIPS)*, Datasets and Benchmarks Track, 2022. (\*equal contribution)

**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.

Xueyan Zou, **Haotian Liu**, and Yong Jae Lee. End-to-end Instance Edge Detection. *arXiv* 2022.

**Haotian Liu\***, Rafael A. Rivera Soto\*, Fanyi Xiao, and Yong Jae Lee. YolactEdge: Real-time Instance Segmentation on the Edge. In *Proceedings of the International Conference on Robotics and Automation (ICRA)*, Xi'an, China, June 2021. (\*equal contribution)

Fanyi Xiao, **Haotian Liu**, and Yong Jae Lee. Identity from here, Pose from there: Self-supervised Disentanglement and Generation of Objects using Unlabeled Videos. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, Seoul, Korea, November 2019.

Yimeng Chen, **Haotian Liu**, and Lei Shi. Operation strategy of public building: Implications from trade-off between carbon emission and occupant satisfaction. *Journal of Cleaner Production (JCLP)* Vol. 205, pp. 629-644, December 2018.

## EXPERIENCE

**Microsoft Research**, Redmond, WA, USA

Research Intern (Part-time)

**September, 2022 - June, 2023**

Research project: REACT and LLaVA (Visual Instruction Tuning).

Research Intern (Full-time)

**May, 2022 - August, 2022**

Research project: retrieval-augmented vision-language models (REACT, CVPR 2023 highlight).

Research Intern (Part-time)

**March, 2022 - May, 2022**

Research project: benchmarking and studying vision-language models like CLIP.

**Cruise LLC**, San Francisco, CA, USA

**March, 2021 - November, 2021**

Research Intern

Research project: an efficient transformer backbone to improve point cloud understanding.

**DawnLight Technologies Inc.**, Palo Alto, CA, USA

**June, 2020 - November, 2020**

Research Intern

Research project: an approach of generating realistic human body animation in the scene.

## PROFESSIONAL ACTIVITIES

### • Conference Reviewer:

- Computer Vision and Pattern Recognition (CVPR), 2021-23
- International Conference on Computer Vision (ICCV), 2021, 2023
- European Conference on Computer Vision (ECCV), 2022
- International Conference on Machine Learning (ICML), 2023
- Neural Information Processing Systems (NeurIPS), 2022-23
- Empirical Methods in Natural Language Processing (EMNLP), 2022
- Annual Meeting of the Association for Computational Linguistics (ACL), 2023
- AAAI Conference on Artificial Intelligence (AAAI), 2023
- International Conference on 3D Vision (3DV), 2022

- Asian Conference on Computer Vision (ACCV), 2022
- **Journal Reviewer:**
  - International Journal of Computer Vision (IJCV), 2023
- **Workshop Organizer:**
  - The 2nd Workshop on Computer Vision in the Wild, CVPR 2023
  - The 1st Workshop on Computer Vision in the Wild, ECCV 2022
- **Workshop Program Committee:**
  - Workshop on Attention and Transformers in Vision, CVPR 2022
  - Workshop on Foundation Models in Vision and Language, ICDM 2022
- **Teaching Assistant:**
  - CS 839 Deep Learning for Visual Recognition, UW Madison, Spring 2022
  - ECS 174 Computer Vision, UC Davis, Spring 2020

#### HONORS AND AWARDS

- a16z Open Source AI Grant, 2023
- Zhejiang University, Outstanding Graduates (with honors), 2019
- UC Davis, Outstanding Research Performance, 2018
- Zhejiang University, Student Award for Research and Innovation, 2018
- Zhejiang University, First-Class Scholarship for Outstanding Students, 2016
- Zhejiang University, Scholarship for Outstanding Merits for 3 successive years, 2016-2018

#### SKILLS

Programming: Python, C/C++, MATLAB, JavaScript, Swift  
 Misc: PyTorch, Tensorflow, OpenGL, L<sup>A</sup>T<sub>E</sub>X, Markdown  
 OS: LINUX, macOS, Windows