Haochen Zhang

haochen4@andrew.cmu.edu | 412-589-4898 | haochenz11.github.io | haochenz11 | haochenz11

Research Interests

Vision-Language-Action Models, Semantic Understanding, Policy Steering, Post-Training, VLA Safety

Education

Carnegie Mellon University *Aug 2025 - Present*
PhD in Robotics, Advised by Prof. Yonatan Bisk

Carnegie Mellon University *Sept 2023 - Aug 2025*
MS in Robotics, Advised by Profs. Wenshan Wang, Ji Zhang

University of Toronto *Sept 2018 - May 2023*
BASc in Engineering Science, ECE, Advised by Prof. Scott Sanner

Technologies

Python/PyTorch, OpenCV, C/C++, Simulators (Habitat, MuJoCo), ROS, Linux, Docker, Git

Relevant Experience

Research Assistant (PhD), Carnegie Mellon University *Pittsburgh, PA*
CLAW Lab, Advised by Prof. Yonatan Bisk *Aug 2025 - Present*

- *[In Progress]* Improving linguistic generalization of Vision-Language-Action (VLA) models
  - Leveraging representations of ontological linguistics in LLMs to mitigate forgetting of language semantics in state-of-the-art VLAs (pi0.5, OpenVLA-OFT)
  - Investigating ability to steer a VLA policy linguistically to learn semantic object abstractions and relations through post-training
  - Evaluation done in simulation (Libero, SimplerEnv) and on real robot manipulator (Franka Panda)
- Mentoring projects: 1) Situated Perspective Taking: grounding objects in 3D scenes based on contextual information, 2) Hierarchical Agentic Scene Memory: long-context agent scene memory for embodied retrieval

Research Intern, Honda Research Institute *San Jose, CA*
Systems Engineering Team *May 2025 - Aug 2025*

- Developed FAST-EQA, a method for embodied question answering in home environments, achieving state-of-the-art question answering accuracy while reducing inference time of each exploration step
- Implemented an exploration algorithm using geometrically-guided frontiers and regional semantic maps in Habitat simulator with a chain-of-thought reasoning module

Research Assistant (MS), Carnegie Mellon University *Pittsburgh, PA*
AirLab/Zhang Lab, Advised by Profs. Wenshan Wang, Ji Zhang *Aug 2023 - Aug 2025*

- *[IROS 2025]* SORT3D: Real-time object-centric grounding and reasoning for navigation in 3D scenes
  - Implemented a full system pipeline for referential grounding and navigation with real-time semantic mapping, VLM captioning, and LLM reasoning modules using ROS
  - Evaluated in RViz simulator environments and validated on real-world ground vehicles with LiDAR and 360-degree RGB camera data in indoor environments
- *[ICRA 2025]* IRef-VLA: Benchmarking 3D object-centric referential grounding with imperfect statements
  - Implemented data creation pipeline: processed 3D scan data (meshes, pointclouds), generated spatial scene graphs and language statements; evaluated multimodal transformer baselines
- Led the CMU VLA Challenge by designing the questions, infrastructure, and evaluation procedure and communicating with participants—gaining industry sponsorship and participation from 20+ teams worldwide
- *[IROS 2024, IROS 2025]* Organized and hosted the AI Meets Autonomy Workshop with over 100 attendees and invited speakers from MIT/BIGAI/Meta/Ai2, etc.

R&D Intern, Qualcomm *Toronto, ON*
Machine Learning Systems Team *May 2021 - Aug 2022*

- Contributed to embedded AI compiler pipeline for converting models (PyTorch/ONNX) to a proprietary deployment format
- Implemented finetuning methods to improve model performance after quantization and compression

Talks/Presentations

Object-Centric Grounding for Deployable and Interactive Vision-Language Navigation Agents *2025*
Master's Thesis Defense, Carnegie Mellon University

Multi-aspect Natural Language Preference-based Retrieval *2023*
Undergraduate Thesis Talk, University of Toronto

CMU VLA Challenge: Details, Results, and Lessons *2024, 2025*
AI Meets Autonomy Workshop, IROS 2024, IROS 2025

Relevant Publications

H. Zhang*, N. Savaliya*, F. Siddiqui, E. Sachdeva. 2025. "FAST-EQA: Efficient Embodied Question Answering with Global and Local Region Relevancy". Accepted to IEEE/CVF Winter Conference on Applications of Computer Vision (**WACV**), March 2026

N. Zantout*, **H. Zhang***, P. Kachana, J. Qiu, G. Chen, J. Zhang, and W. Wang. 2025. "SORT3D: Spatial Object-centric Reasoning Toolbox for Zero-Shot 3D Grounding Using Large Language Models". IEEE/RSJ International Conference on Intelligent Robots and Systems (**IROS**), October 2025

H. Zhang*, N. Zantout*, P. Kachana, Z. Wu, J. Zhang, and W. Wang. 2025. "IRef-VLA: A Benchmark for Interactive Referential Grounding with Imperfect Language in 3D Scenes". IEEE International Conference on Robotics & Automation (**ICRA**), May 2025

H. Zhang, N. Zantout, P. Kachana, J. Zhang, and W. Wang. 2024. "VLA-3D: A Dataset for 3D Semantic Scene Understanding and Navigation". In **RSS Workshop** for Semantic Reasoning and Goal Understanding in Robotics, July 2024

H. Zhang*, A. Korikov*, [8 other authors], and S. Sanner. 2023. "Recipe-MPR: A Test Collection for Evaluating Multi-aspect Preference-based Natural Language Retrieval". ACM **SIGIR** Conference on Research and Development in Information Retrieval, July 23–27, 2023

Teaching and Mentorship

Graduate Teaching Assistant, Carnegie Mellon University *Jan 2025 - Present*
16-831 Introduction to Robot Learning, Taught by Prof. Guanya Shi

Capstone Mentor, Carnegie Mellon University *Jan 2025 - Present*
Master's in Computational Data Science

Graduate Student Mentor, Carnegie Mellon University *Oct 2023 - May 2025*
Pathways to AI Research (PAIR) Program

Awards

Vector Scholarship in Artificial Intelligence *2023*
IEEE Toronto Student Scholarship *2022-2023*
Bruno D. Stefano Scholarship *2021-2022*
Laura Chizuko Fujino Scholarship in Engineering Science *2020-2021*
ESROP Research Fellowship *2020*
Dean's Merit Award Entrance Scholarship *2018*