Haoyang He

hehaoyang@cmu.edu | (412) 352-0092 | https://purenothingness24.github.io

Skills: Multimodal Learning, Computer Vision, Visual Generation, Embodiment, Reinforcement Learning.
Programming Languages: Python, C, C++, Objective-C, Matlab, HTML/CSS.
Tools: PyTorch, Unix, IsaacSim, MuJuCo, ROS2, Xcode, Android Studio.

EDUCATION

Carnegie Mellon University (CMU) *Pittsburgh, PA*

- M.S. in Robotic Systems Development (GPA 4.08/4). Dec 2025
- M.S. in Electrical and Computer Engineering (GPA 3.92/4). May 2023
- B.S. in Electrical and Computer Engineering (3-year graduation). May 2022
- CMU Innovation Scholar.

Relevant coursework:

11-777 Multimodal Machine Learning	16-726 Learning-Based Image Synthesis	16-825 Learning for 3D Vision
11-785 Introduction to Deep Learning	16-785 Integrated Intelligence in Robotics	16-831 Introduction to Robot Learning
16-720 Computer Vision	16-824 Visual Learning & Recognition	16-833 Robot Localization & Mapping

RESEARCH PROJECTS

The Autonomous Wheelchair Project *Pittsburgh, PA*

Research Assistant at CMU AIRLab Aug 2024 - Present

- Developing language-grounded cost maps for safe human-interactive autonomous wheelchair systems.

SceneGaussian: Unconstrained Generation of 3D Gaussian Splatting Scenes *Pittsburgh, PA*

Course Project of CMU 16-825 Learning for 3D Vision Mar 2024 - May 2024

- Proposed a novel method for unconstrained text-to-3D scene generation by updating 3D Gaussian Splatting with novel view generated by Stable Diffusion RGB Inpainting and its depth estimation to improve intrinsic spatial understanding of the model, a key limitation in unconstrained 3D scene generation.
- Achieved similar performance as state-of-the-art methods LucidDreamer and Text2Room with more potentials for high efficiency to achieve high-quality on-the-fly unconstrained 3D scene generation.
- Project report available on Google Drive.

Natural Dexterous Piano Playing at Scale With Video Hand Priors *Pittsburgh, PA*

Course Project of CMU 16-831 Introduction to Robot Learning Mar 2024 - May 2024

- Proposed a novel method to learn dexterous robot piano playing from YouTube videos by generating fingering labels with state-of-the-art hand pose estimation and music note transcription, enabling large-scale training data generation for zero-shot piano playing at scale.
- Model learned to play a challenging 14-minute piano piece with reasonable performance, compared to 30-second piano pieces by previous methods.
- Project report available on Google Drive.

Multimodal Analysis of Embodied Instruction Following on ALFRED *Pittsburgh, PA*

Course Project of CMU 11-777 Multimodal Machine Learning Feb 2024 - May 2024

- Conducted modality analysis of the ALFRED embodied instruction following (EIF) task with 16 modality ablations.
- Concluded that the ALFRED task is insufficient to benchmark EIF performance, as baseline method with ground truth perception outperforms state-of-the-art methods based on foundation models by 40%.
- Project report available on Google Drive.

Stable PFGM++ *Remote*

Independent Researcher at Stability AI Sep 2023 - Dec 2023

- Researched Poisson Flow Generative Model in latent space for text-to-image generation.

A Survey on Offline Model-Based Reinforcement Learning
Pittsburgh, PA

Course Project of CMU 18-831 Foundations of Reinforcement Learning Mar 2023 - May 2023

- Analyzed literature on offline model-based reinforcement learning, focused on adaptation to offline environment.
- Project report published on arXiv: https://arxiv.org/abs/2305.03360.

GAS-NeXt: Few-Shot Cross-Lingual Font Generator
Pittsburgh, PA

Course Project of CMU 11-785 Introduction to Deep Learning Sep 2022 - Jan 2023

- Proposed a novel method for few-shot cross-lingual font style generation based on FTransGAN and AGIS-Net, outperforming previous state-of-the-art methods in generating Chinese fonts from English fonts and vice versa.
- Cited by published paper in 2024 as state-of-the-art convolution-based Chinese font generation.
- Project report published on arXiv: https://arxiv.org/abs/2212.02886.

Automatic Eye-in-Hand Calibration Using EKF
Pittsburgh, PA

Course Project of CMU 16-833 Robot Localization and Mapping Feb 2022 - May 2022

- Proposed a novel fully-automatic industrial robotic arm calibration method based on Extended Kalman Filter.
- Project report published on arXiv: https://arxiv.org/abs/2211.06881.

WORK EXPERIENCE

CMU School of Computer Science
Pittsburgh, PA

Teaching Assistant Jan 2024 - Present

- Course staff for courses 10-423/623 Generative AI (Spring & Fall) and 11-777 Multimodal Machine Learning (Fall)).
- Created homework for Multimodal Foundation Models in the newly created Generative AI course.
- Taught 3 out of 5 recitations for Generative AI to a class of 150+ students and held office hours.
- Mentored 9 students in multimodal generative AI course projects.

Peking University Yuanpei Young Scholar Program
Remote

Research Mentor Dec 2023 - Jul 2024

- Designed and taught curriculum of computer vision and visual generation for advanced high school students.
- Mentored 11 research projects in computer vision and visual generation, 3 published in the program journal.

Zoo Capital
Shanghai, China

Tech Analyst Jun 2023 - Dec 2023

- Evaluated technical potentials of AI and Robotics startups for investment decisions.
- Tracked promising new research and evaluate potentials for the industry applications.

Apple
Cupertino, CA

Software Engineer Intern May 2022 - Aug 2022

- Built a system-level development tool for Camera App and API developments, used by colleagues.

Bilibili
Shanghai, China

Algorithm Engineer Intern Jun 2020 - Feb 2021

- Joined newly created Computer Vision group as 4-month full-time and 4-month part-time intern.
- Trained generic object tracking model based on YOLOv3, released in Jan-2021 version B-Cut app.
- Built data labeling SDK for facial interpretation, hand recognition, and body segmentation, used by colleagues.
- Assisted building Android and iOS apps for model efficiency testing on mobile devices, used by colleagues.
- Attended and presented in weekly novel computer vision paper reviews.

ZhenFund
Shanghai, China

Investment Intern Jun 2019 - Aug 2019

- Interviewed 50+ front-edge tech start-ups and evaluated potentials for early-stage investments.