

Haoyu Chen

tonychen54816@gmail.com | +1-347-925-6125 | <http://tonychy.com> | [linkedin.com/in/haoyuchen050400](https://www.linkedin.com/in/haoyuchen050400)

Education

Harvard University

Master of Science in Computational Science and Engineering

Relevant Coursework: Computer Graphics & Mathematical Modeling

Cambridge, MA

Sep. 2025 – May 2027 (expected)

Carnegie Mellon University

Master of Science in Electrical and Computer Engineering, Applied Program

Relevant Coursework: Intro to DL & Pattern Recognition, Big Data Science, Inference Statistics & Applied Machine Learning

Pittsburgh, PA

Sep. 2022 – Dec. 2023

Rensselaer Polytechnic Institute

Bachelor of Science in Computer Science and Mathematics (Minor: Economics)

Relevant Coursework: Introduction to Algorithm, Data Structure, Machine Learning From Data, Linear Algebra

Troy, NY

Sep. 2018 – May 2022

Publications

- **Haoyu Chen**, Kaichen Zhou, Hang Hua, Kaile Zhang, Jingwen Qian, Wufei Ma, Haonan Chen, Chunjiang Liu, Yizhou Zhao, Xiaoyuan Wang, Weiyue Li, Alan Yuille, Paul Pu Liang, Yilun Du, “MemoBench: Benchmarking World Modeling in Dynamically Changing Environments,” *European Conference on Computer Vision (ECCV)*, 2026. Accepted. [arXiv:2606.27537](https://arxiv.org/abs/2606.27537).
- Yizhou Zhao, **Haoyu Chen**, Chunjiang Liu, Zhenyang Li, Charles Herrmann, Junhwa Hur, Yinxiao Li, Ming-Hsuan Yang, Bhiksha Raj, Min Xu, “Toward Material-Agnostic System Identification from Videos,” *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2025.
- Wufei Ma, **Haoyu Chen**, Guofeng Zhang, Celso Mde Melo, Jieneng Chen, Alan Yuille, “3DSRBench: A Comprehensive 3D Spatial Reasoning Benchmark,” *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2025.
- Yizhou Zhao, Chunjiang Liu, **Haoyu Chen**, Bhiksha Raj, Min Xu, Tadas Baltrusaitis, Mitch Rundle, HsiangTao Wu, Kamaran Ghasedi, “Total Editing: Head Avatar with Editable Appearance, Motion, and Lighting,” *ICCV Workshop on E2E3D*, 2025.
- Dachuan Zhao, Weiyue Li, Zhenda Shen, Yushu Qiu, Bowen Xu, **Haoyu Chen**, Yongchao Chen, “Bias Is a Subspace, Not a Coordinate: A Geometric Rethinking of Post-hoc Debiasing in Vision-Language Models,” *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2026.
- Chunjiang Liu, Xiaoyuan Wang, Qingran Lin, Albert Xiao, **Haoyu Chen**, Shizheng Wen, Hao Zhang, Lu Qi, Ming-Hsuan Yang, Laszlo A Jeni, Min Xu, Yizhou Zhao, “MOSIV: Multi-Object System Identification from Videos,” *International Conference on Learning Representations (ICLR)*, 2026.

Research Experience

World Model in Dynamically Changing Environments (Accepted to ECCV 2026)

Boston, MA

Supervised by Prof. Yilun Du at Harvard University

Nov. 2025 – Mar. 2026

- Introduced **MemoBench**, a diagnostic benchmark for evaluating visual memory and world modeling in dynamically changing environments.
- Curated 360 ground-truth video clips across synthetic and real-world scenes under a disappear-and-reappear paradigm, where objects undergo physical changes while out of view.
- Designed an evaluation suite combining automated metrics and VQA-based assessment to analyze memory consistency and state-update capabilities of state-of-the-art video generation models.

4D Physical Plausible Reconstruction (Accepted at ICCV 2025)

Pittsburgh, PA

Supervised by Prof. Min Xu at Carnegie Mellon University

May 2024 – Mar. 2025

- Introduced the first material-agnostic system identification framework from videos, eliminating the need for predefined material priors.
- Utilized **MLS-MPM** accompany with Neural Networks to estimate the intrinsic physical properties.
- Reached state-of-the-art performance among all recent geometry reconstruction models and future state prediction methods across diverse materials.

Weakly-Supervised Learning For Lab Automation

Pittsburgh, PA

Supervised by Prof. Min Xu at Carnegie Mellon University

Feb. 2024 – Mar. 2025

- Fine-tuned the **YOLOv5** implementation pre-trained on the COCO to our dataset, to segment regions of pipettes in real-time video streams, enhancing the precision of the subsequent anomaly detection.
- Employed a **lightweight transformer encoder** for dynamic anomaly scoring, integrated with a binary cross-entropy (BCE) loss function to optimize detection accuracy.
- Optimized the video anomaly detection system, achieving an AUC of 98.79%, which outperformed existing state-of-the-art methods by over 11%.

Head Avatar with Editable Appearance, Motion, and Lighting (Accepted at ICCV 2025 Workshop)

Pittsburgh, PA

Supervised by Prof. Min Xu at Carnegie Mellon University

July 2024 – Nov. 2024

- Introduced "Total-Editing," a unified framework combining face reenactment and portrait relighting to generate 3D portraits with specified appearance, motion, and lighting conditions.
- Developed a tri-plane decoder and deformation field to improve perceptual quality, spatiotemporal coherence, and realistic shading effects.
- Enabled flexible applications, such as animatable portraits with customizable backgrounds, through a multi-source formulation.

3D spatial reasoning benchmark (Accepted at ICCV 2025)

Baltimore, MD

Supervised by Prof. Alan Yuille at Johns Hopkins University

July 2024 – Nov. 2024

- Developed a comprehensive benchmark for 3D spatial reasoning, comprising 2,772 manually annotated visual question-answer pairs across 12 distinct question types.
- Co-designed and implemented a novel FlipEval strategy to evaluate 3D spatial reasoning capabilities.
- Uncovered key limitations in large multi-modal models (LMMs) regarding 3D awareness, particularly in height, orientation, and location-based reasoning.

Synthetic Dataset Generation For Video Datasets

Baltimore, MD

Supervised by Prof. Alan Yuille at Johns Hopkins University

May 2024 – Aug. 2024

- Utilized Ctrl-adapter to generate diverse sythetic in-door data based on projectaria datasets.
- Applied sythetic data for several robot learning tasks and enhanced the performance with 4.2 units lower in FID and 3 units lower in KID compared with the benchmark: HSSD-200 datasets.

Predicting crop yield based on remote-sensing data and Deep Learning Models

Pittsburgh, PA

Project leader, Supervised by Prof. Marios Savvides at Carnegie Mellon University

Sept. 2023 – Dec. 2023

- Developed a script for extracting and preprocessing remote-sensing data from Google Earth Engine, applying advanced feature engineering techniques for deep learning model inputs.
- Engineered and optimized a specialized **Convolutional Neural Network (CNN)** for US Corn yield prediction, achieving a 12% reduction in RMSE for the 2022 forecast compared to **Ridge Regression** and **Decision Tree** models, and achieving 5% reduction compared to **Long short-term memory Network (LSTM)** models.

Pipeline for real-time analyzing professional players' in-game performance

Pittsburgh, PA

Summer Research Intern, Supervised by Prof. Priya Narasimhan at Carnegie Mellon University

May 2023 – Aug. 2023

- Developed complex algorithms for emotion detection, gaze tracking, and posture analysis using **CV2**, **FER**, **DLIB**, and **MediaPipe** libraries.
- Implemented a Python-based keyboard and mouse logger to enhance data collection and capture detailed user input and interaction patterns.
- Performed data cleaning, feature selection, and analysis with Python libraries, uncovering meaningful trends and correlations, such as the impact of individual behavioral patterns on in-game performance in FPS games.

Learning Ethical Principles using Lexicographic Preference Models

Troy, NY

Supervised by Prof. Lirong Xia at Rensselaer Polytechnic Institute

Mar. 2021 - May 2022

- Designed a Python-based lexicographic preference model to optimize survival decisions in simulated scenarios.
- Corporated to optimize the model performance and improved prediction accuracy by 18% compared to the **Decision Tree** model and 14% compared to 1&2 hidden layer **Deep Neural Network (DNN)**.

Skills

Programming Languages: Proficient in Python; Knowledgeable in C/C++, MATLAB, Java, R

Tools & Technologies: Skilled in PyTorch, Sklearn; Familiar with MySQL, Pandas, Azure, AWS

Soft Skills: Project Management, Communication, Leadership