

# YIMING CHEN

---

## Education

### Georgia Institute of Technology

*Master's Degree in Computer Science*

- GPA: 4.00/4.00
- Specialization: Machine Learning
- Advisor: James Hays

Atlanta, GA

2024 - 2026 (*expected*)

### New York University

*Bachelor's Degree in Computer Science and Mathematics*

- GPA: 3.73/4.00
- Advisor: Gizem Kayar

Manhattan, NY

2020 - 2024

## Publications

1. **Chen, Y.**, Singh, S., Chatterjee, A., Raj, A., Hays, J., Yang, Y., & Baral, C. (2026). Chimera: Compositional Image Generation using Part-based Concepting. *Under review at ICLR 2026*.
2. Yang, M., **Chen, Y.**, Pei, H., Agarwal, S., Vasudevan, A. B., & Hays, J. (2025). Clink! Chop! Thud! - Learning Object Sounds from Real-World Interactions. *ICCV 2025*.
3. **Chen, Y.**, Liu, Y., & Kayar-Ceylan, G. (2025). CSG-based ML-supported 3D translation of sketches into game assets for game designers. *The Visual Computer*, 1-13.
4. **Chen, Y.** (2020). A Device for Compressing Garbage in Dustbin. *China National Intellectual Property Administration*, Patent No. CN210823859U.

## Presentations

1. **Chen, Y.**, Liu, Y., "CSG-based ML-supported 3D translation of sketches into game assets for game designers", New York University Undergraduate Research Conference, NY, May 2024

## Research Experience

### Benchmarking World Models with Physical Property Understanding

*Research Assistant, Advised by Prof. James Hays*

Aug 2025 - Present

- Developing a novel benchmark of physically impossible videos for evaluating the physical property understanding and reasoning capabilities of modern world models.
- Generating diverse, counterfactual scenarios using a structured prompt taxonomy with both high-fidelity simulations and state-of-the-art Text-to-Video models.

### Egocentric-to-3D Interaction Reconstruction

*Research Assistant, Advised by Prof. Yalong Yang*

Mar 2025 - Present

- Implementing a 3D reconstruction pipeline using MegaSam to map egocentric interaction data into semantically segmented 3D spaces.
- Developing a novel view synthesis framework that decouples the observer from the user, allowing third parties to visualize first-person tasks from arbitrary external viewpoints.

### Learning Object Sounds from Real-World Interactions

*Research Assistant, Advised by Prof. James Hays*

Sep 2024 - Oct 2025

- Proposed the sounding object detection task to identify source objects from interaction sounds.
- Developed a multimodal, object-aware framework using a slot attention encoder and an automatic annotation pipeline to learn from egocentric videos.
- Improved SOTA on sounding object detection by 11.8% on Epic Kitchens and 9.8% on Ego4D.

### Part-based Compositional Image Generation

*Research Assistant, Collaboration with Google DeepMind*

Aug 2024 - Present

- Constructed a novel training dataset of 37k images based on a taxonomy of 464 unique pairs to enable fine-grained compositional generation.
- Proposed the PartEval metric to systematically assess the fidelity and compositional accuracy of part-based generation pipelines.
- Developed a diffusion prior model with part-conditional guidance that outperforms baselines by 14% in compositional accuracy and 21% in visual quality.

## Gaussian Visibility Field for Uncertainty Mapping

*Research Assistant, Collaboration with NVIDIA*

Aug 2024 - Dec 2024

- Developed novel neural field representations by applying Bayesian deep learning principles to map and quantify uncertainty in 3D scenes for autonomous systems.
- Engineered the system for real-time performance by optimizing the neural field's computational efficiency through advanced GPU parallelization and CUDA techniques.

## Unity ML-Agents with Walking Robot using Curriculum Training

*Research Assistant, Advised by Prof. Jeremy Curuksu*

Mar 2024 - May 2024

- Investigated sample-efficient training for embodied agents by integrating curriculum learning with advanced deep RL algorithms (PPO) to accelerate policy convergence.
- Enhanced the agent's navigation capabilities by implementing a ray-based perception system for robust and efficient obstacle avoidance in complex environments.

## CSG-based ML-supported 3D Translation of Sketches into Game Assets

*Research Assistant, Advised by Prof. Gizem Kayar-Ceylan*

May 2023 - Jan 2025

- Developed a computer vision pipeline to extract geometric primitives, training a CNN shape classifier that achieved 95.74% accuracy on the test set.
- Engineered a real-time web application using CSG Boolean operations to reconstruct sketches into textured 3D models in 31-33 seconds.

## Amazon Web Services - Annapurna Labs

*Machine Learning Engineer Internship*

Cupertino, CA

May 2025 - Aug 2025

- Designed and implemented pipeline parallel processing in NxDI stack for large language models, improving distributed training efficiency by 35%.
- Implemented Sequence Pipeline Parallel to create a prompt parallel architecture for window context encoding, enhancing vLLM inference performance by 2.3x.
- Collaborated with the inference team to optimize distributed tensor and pipeline parallelism, documenting 40% latency improvements using NeuronTransferEngine.

## T.C.L. Industries Holdings

*Machine Learning Engineer Internship*

Shenzhen, China

Jun 2024 - Aug 2024

- Led the development of predictive analytics models using LSTM, increased accuracy from 40% to 90%.
- Developed APIs for the model, making the model efficient to use internally and reducing costs by 35%.
- Integrated the model with Microsoft 365, enabled others to fine-tune different specific prediction models.

## Awards

- **Dean's Undergraduate Research Fund**, New York University 2023.09
- **Registration Support Scholarship**, New York University 2022.09
- **Student Relief Scholarship II**, New York University 2022.02
- **Student Relief Scholarship**, New York University 2021.09

## Teaching

*Teaching Assistant*

Georgia Institute of Technology

Computer Vision (CS 6476 A) - Head TA

Fall 2025

Computer Vision (CS 6476 A)

Spring 2025

*Teaching Assistant*

New York University

Computer Systems Organization (CSCI-UA 201)

Spring 2024

Computer Systems Organization (CSCI-UA 201)

Fall 2023

Great Ideas in Math (CORE-UA 110)

Spring 2023

Algebra, Trigonometry, and Functions (MATH-UA 9)

Fall 2022