

# Zhangyun Tan

Email: ztan12@u.rochester.edu/[Homepage](#)

## EDUCATION

---

### University of Rochester

*B.S. in Computer Science; Honor Distinction in research*

- Relevant Courses: Computer Vision, Machine Learning, Deep Learning, Natural Language Processing

**Rochester, NY**

*05/2027 (expected)*

### University of California, Berkeley

*Exchange in EECS*

- Relevant Courses: Convex Optimization

**Berkeley, CA**

*01/2024-05/2024*

## EXPERIENCE

---

### University of Rochester | Research Assistant advised by [Prof. Chenliang Xu](#)

**01/2025-present**

- *Multimodal Learning*: Contributing to projects on Multimodal learning especially in visual understanding
- *Efficient LLMs*: Investigated methods for improving the efficiency of training and inference in LLMs.

### Johns Hopkins University | Research Assistant advised by [Prof. Alan L. Yuille](#)

**06/2025-12/2025**

- *Spatial Reasoning in VLMs*: Built benchmark evaluation tasks for spatial reasoning in vision-language models.

### DeepBlue AI | Research Intern

**07/2023 - 01/2024**

- *Embodied Agents*: Integrated LLMs into a virtual town simulation and collaborated with Unity engineers to support context-aware agent behavior in 3D environments.
- *3D Reconstruction*: Implemented NeRF-based methods for converting 2D images into 3D reconstructions and generating high-fidelity virtual environments.

## PUBLICATIONS

- 
- *Y. Tang\*, P. Liu\*, **Zhangyun Tan\***, et al. MMPerspective: Do MLLMs Understand Perspective? A Comprehensive Benchmark for Perspective Perception, Reasoning, and Robustness (**NeurIPS 2025**)*
  - ***Zhangyun Tan**, Z. Zhang, S. Liang, C. Xu. Can VLMs Truly Forget? Benchmarking Training-Free Visual Concept Unlearning (submitted to **COLM 2026**)*

- *Y. Tang, J. Bi, P. Liu, Z. Pan, **Zhangyun Tan**, et al. Video-LMM Post-Training: A Deep Dive into Video Reasoning with Large Multimodal Models (submitted to COLM 2026)*

## SELECTED PROJECTS

---

### Efficient Multi-Agent Systems

08/2025

- *Engineered multi-agent workflows for task decomposition and coordination, addressing the inefficiency of single-agent reasoning on complex tasks*
- *quantified performance-cost trade-offs across multiple agent configurations.*
- *Delivered consistent gains over local and sparse-attention baselines under tight KV cache budgets, reaching performance close to full-cache models with only 20% cache usage.*

### Accelerated Gradient Descent with Mirror Descent

05/2024

- *Established and validated theoretical convergence rates for mirror descent and hybrid mirror-gradient methods, enhancing optimization efficiency in convex settings.*
- *Developed and implemented accelerated gradient descent algorithms in Python, outperforming baseline optimizers in empirical benchmarks.*

### Embodied Gaming Agent

09/2023 - 01/2024

- *Embodied Decision-Making: Constructed agent pipeline for sequential decision-making in a Minecraft simulation, addressing the challenge of long-horizon planning in interactive environments; improved task completion over smaller-model baselines.*
- *Multimodal Grounding: Fused visual observations, environment states, and language instructions into the agent decision loop, enabling more adaptive behavior across diverse in-game scenarios.*

### Generative Agents for Urban Simulation

08/2023 - 12/2023

- *Urban Generative Agents: Integrated a Chinese large language model into a map-based simulation pipeline to support language-driven, context-aware agent interaction in urban scenarios.*
- *3D Simulation Deployment: Coordinated with Unity engineers to deploy generative agents in a 3D simulation of downtown Shanghai, enabling interactive agent behavior in dynamic city environments.*

## SKILLS

---

*Python, R, C/C++, Java Script, MATLAB, PyTorch, TensorFlow, scikit-learn, NumPy, pandas*