

Yuan (Paul) Zhuang

(267) 788-6834 | pauyuanzhuang@gmail.com | [Website](#) | [Scholar](#) | [GitHub](#) | [LinkedIn](#)

Research focus: large language models, test-time training, efficiency, reinforcement learning, embodied AI, Robotics

Education

University of Connecticut, Ph.D. in Computer Science and Engineering, Storrs, CT	Expected Dec 2026
University of Pennsylvania, M.S. in Data Science, Philadelphia, PA (GPA: 3.87/4.00)	May 2024
Zhejiang University, B.Eng. in Ocean Engineering and Technology, Hangzhou, China (GPA: 3.65/4.00)	Jun 2022

Industry Research Experience

Microsoft, Microsoft Research Lab, Research Intern, Redmond, WA, US	May 2026–Aug 2026
- Developed adaptive early-exit methods and selective memory updates for long-context test-time training LLMs with linear attention structure. - Improved long-context retrieval while reducing estimated prefill compute by up to $2.2\times$ across two different models families. - Explored depth-feature and KV-cache approaches for improving LLM inference efficiency.	
JD.com, Research Intern, Embodied Technology Group, Beijing, China	May–Aug 2025
Experimented with π, OpenVLA, and NVIDIA GR00T vision-language-action models (VLAs) for parcel sorting and palletizing, studying multitask robustness and generalization across task and environment variations.	

Selected Research & Publications

LD-MoLE: Learnable Dynamic Routing for Mixture of LoRA Experts, First author	ICLR 2026 [Paper] [Code]
- Developed a fully differentiable router with analytical sparsity control to learn experts allocation for mixture-of-experts. - Achieved superior performance across eight benchmarks on Llama-3.2-3B and Qwen3-1.7B while maintaining model sparsity.	
Low-Rank Adaptation for Critic Learning in Off-Policy Reinforcement Learning, First author	Preprint, 2026 [Paper] [Code]
- Introduced LoRA as structural regularization for off-policy critics by freezing dense bases and optimizing low-rank updates. - Evaluated SAC and FastTD3 on 13 DMC-Hard and IsaacLab tasks; raised large-critic mean return from 719.7 to 777.2 while only training 5.20M vs. 34.82M critic parameters in large-scale reinforcement learning networks.	
YOLO-MARL: You Only LLM Once for Multi-Agent Reinforcement Learning, First author	IROS 2025 [Paper] [Code]
- Used one LLM interaction per environment to generate executable planning and reward-shaping functions in Python code; decentralized MARL policies train and execute without further LLM inference. - Evaluated QMIX, MADDPG, and MAPPO across multiple benchmarks; improved mean return by up to 105% and doubled convergence speed.	
Artificial Foveated Perception for Mitigating Shortcut Learning in Robotic Foundation Models, Second author; manuscript under review	Preprint, 2026 [Paper]
Proposed a masked visual-embedding method designed to reduce sensitivity to task-irrelevant backgrounds; evaluated VLA robustness under environment changes.	
GNN-Based Swarm Modeling and Decentralized Control, Research Assistant, Penn GRASP	May–Sep 2023
Built GNNs to infer swarm interactions and leaders and implemented decentralized multi-robot controllers.	

Technical Skills

Programming/Systems: Python, C/C++, PyTorch, JAX, CUDA/Triton, Linux, distributed training
ML/Robotics: LLMs, foundation models, MoE, PEFT/LoRA, test-time training, reinforcement learning, VLA models, ROS

Honors

Pratt & Whitney Advanced Systems Engineering Fellowship (2024, 2025); UConn Predoctoral Fellowship (2026)
Outstanding Graduate, Zhejiang University (2022); Second Prize, National Underwater Robotics Competition (2021)

Additional Projects

Image-to-Image Translation: Diversity-Fidelity Trade-off
Implemented BicycleGAN and designed a Soft-IntroVAE architecture to generate higher-quality images. [Code]
Neural Style Transfer
Applied multiple neural networks to produce stylized videos: [Original], [Leon], and [Johnson].