

# Tony Jie Wang

Mobile: +1 (267)-403-8674 | E-mail: [tonyw3@seas.upenn.edu](mailto:tonyw3@seas.upenn.edu) | LinkedIn: [Jie Wang's LinkedIn](#)  
Website: [everloom-129.github.io](https://everloom-129.github.io) | GitHub: [github.com/Everloom-129](https://github.com/Everloom-129) | Twitter: [twitter.com/JieWang\\_ZJU](https://twitter.com/JieWang_ZJU)

## EDUCATION

**University of Pennsylvania** (GPA: 3.63/4.0) Philadelphia, PA  
M.S.E Robotics, Advisor: [Prof. Dinesh Jayaraman](#) & [Prof. Kostas Daniilidis](#) Aug 2024 – Jun 2026  
**Zhejiang University** (GPA: 3.83/4.0) Zhejiang, China  
BEng. Electronic and Computer Engineering, Advisor: [Prof. Said Mikki](#) Aug 2020 – Jun 2024  
**University of Illinois at Urbana-Champaign** (GPA: 3.50/4.0) Urbana, IL  
BSc. Computer Engineering, Advisor: Prof. [Deepak Vasisht](#) Aug 2020 – Jun 2024

## INDUSTRY EXPERIENCE

### [DynaRobotics](#)

*DYNA-2 Research Team, Mentor: Dr. Jason Ma* Redwood City, CA  
Research Intern Jun 2025 – Aug 2025

- Prototyped **Agentic Reasoning** project: designed an autoregressive head to enable VLAs talk, verify current progress and maintain task memory using one model. [\[video\]](#)
- Engineered large-scale data pipelines by curating, filtering, and normalizing external datasets into pre-training recipes.
- Developed a subtask prediction module for task progress estimation, achieving 90% accuracy on Napkin Folding.
- Optimized ML infrastructure: conducted code reviews and resolved critical system bottlenecks.
- Deployed and validated models on real-world tasks including *Cloth Stacking* and *Candy Pick-and-Place*.

### [International Digital Economy Academy](#)

*Computer Vision & Robotics Department, Mentor: Dr. Xiaoke Jiang* Shenzhen, China  
Machine Learning Engineering Intern Jun 2024 – Sep 2024

- Diagnosed I/O bottlenecks between SLURM clusters, accelerated data loading throughput by 16.7%
- Adapted Grounding DINO 1.6 as a vision encoder to train VLMs on Motion-X++.
- Trained, Finetuned and Optimized an LLaVA-style VLM for human behavior recognition.
- Contributed to the DINO-X, a SOTA vision foundation model family for open-world understanding.

## PUBLICATIONS

- Song, Y., Le, L., Park, Y., **Wang, J.**, Shi, J., Liu, L., Gu, J., Eaton, E., Jayaraman, D., Daniilidis, K. [OmniGuide: Universal Guidance Fields for Enhancing Generalist Robot Policies](#). Under review.
- Shen, W., Kumar, N., Chintalapudi, S., **Wang, J.**, Watson, C., Hu, E. S., Cao, J., Jayaraman, D., Kaelbling, L. P., Lozano-Pérez, T. [TiPToP: A Modular Open-Vocabulary Planning System for Robotic Manipulation](#). Under review.
- Hu, E. S.\* **Wang, J.\***, Yuan, X.\* Luo, F. Li, M. Lambrechts, G. Rybkin, O. Jayaraman, D. (2025) [AAWR: Real-World Reinforcement Learning of Interactive Perception Behaviors](#), (NeurIPS2025 & ARLET Workshop)
- **Wang, J.\*** Leonard, M. Daniilidis, K. Jayaraman, D. Hu, E. S.(2025) [Evaluating pi0 in the Wild: Strengths, Problems, and the Future of Generalist Robot Policies](#), GRASP Lab Blog.
- RoboArena Team. (2025) [RoboArena: Distributed Real-World Evaluation of Generalist Robot Policies](#), CoRL 2025 Oral
- Shi, J. Zhao, Z. Wang. T. Pedroza, I. Luo, A. **Wang, J.** Ma, J. Jayaraman. D (2025) [ZeroMimic: Distilling Robotic Manipulation Skills from Web Videos](#). ICRA 2025 & Best Paper at CVPR 2025 3DVLM Workshop.
- Zhu, X., Li, Z., Jiang, Y., Xu, J., **Wang, J.**, & Bai, X. "Real-time V2V Communication Network Cooperative Control System through Distributed Database," oral at ICICT2024, London, UK.
- Xiang, X., Lei, Z., **Wang, J.** Zheng, Q., Huang Y. "Visionary Co-Driver: LLMs Enhance Driver Risk Perception with ARHUD", IEEE Transactions on Intelligent Transportation Systems (IEEE-ITS)

## Research Experience

### 2025 Master Thesis: Enhancing VLAs with 3D Representation

Keywords: 3D Foundation Models, Guidance, VLAs

GRASP Lab  
Prof. Daniilidis & Prof. Jayaraman

- Benchmarked VLAs performance on Toy Cube for complex spatial reasoning tasks.
- Visualized the attention map of VLAs to study the correlation between image and behavior.
- Adapted VGGT, DINOv3, SAM-3 as critic, applying Diffusion Forcing to steer behavior with reward.
- Reproduced [DemoDiffusion](#), [DynaGuide](#) on Franka Panda Arm, proofread paper in submission for RSS 2026.

### 2025 [AAWR](#): Real-World Reinforcement Learning of Active Perception Behaviors

Keywords: Active Perception, RL, VLA

GRASP Lab,  
Prof. Dinesh Jayaraman

- Proposed **Asymmetric-AWR**, a novel RL framework that leverages Vision Foundation Models as privileged visual critics during training to scaffold active perception policies, significantly outperforming BC and standard AWR baselines.
- Built hardware and perception stack for Koch Arm, implementing RealSense multi-camera calibration and DepthPro 3D object tracking for reward labeling, blind pick and place got 89% success rate using online RL.
- Refactored DROID infra as [EVA](#), supporting GELLO and 3DSpaceMouse interfaces, accelerating data collection speed by ~20s per trial with better control compared to VR Controller.
- AAWR outperform all baselines in VLA Active Perception in Franka Robot Arm with only 50 demonstrations.

## 2025 Evaluation of Generalist Policies in the Wild

Keywords: *Model Evaluation, Foundation Models, VLA*

GRASP Lab

Prof. Daniilidis & Prof. Jayraman

- Led comprehensive observation study on frontier VLA models across 300+ real-world trials.
- Conducted A/B Comparison of evaluation pool and distributed scoring system.
- Co-organized the RoboArena Workshop at CoRL 2025

## 2024 Zero-shot Manipulation using VLMs as Planner

Keywords: *Foundation Models, Planning, Hybrid Control*

GRASP Lab

Advisor: Prof. Dinesh Jayraman

- Reproduced [ReKep](#) on a Mobile Franka Arm Platform with DROID, conducted experiment for [ZeroMimic](#) as a baseline.
- Reproduced [SPHINX](#) for hybrid salient point control using transformer and diffusion policy.
- Built a [video processing tool](#) as shared infra, upgraded the physical lighting system for the lab.

## 2023 Visionary Co-Driver: LLMs Enhance Driver Risk Perception with ARHUD

Keywords: *Autonomous Driving, Foundation Models, HCI*

Zhejiang University

Advisor: Prof. Wei Xiang

- Engineered an LLM-driven risk assessment framework for human-machine co-driving; integrated the system to interpret monocular video streams and generate eye-tracking, context-aware risk warnings via AR-HUD.
- Designed the Human-Computer Interface incorporating eye-tracking to monitor driver attention
- Conducted a high-fidelity user study (N=41) in a simulated cockpit; statistics of behavioral data demonstrate significantly improved risk awareness and reaction times.

---

## PROJECT EXPERIENCE

### 2024 Undergrad Senior Design: Intelligent Pour-Over Coffee Machine [\[Thesis\]](#)

Haining, China

- Built a pour-over coffee machine by reverse engineering.
- Guaranteed food-grade quality coffee for drink.
- Customizable coffee recipe according to user preferences.

### 2023 CS438: Communication Networks Wireless Project [\[Report\]](#)

Urbana, IL

- Developed an open-source Python tool for comprehensive wireless network analysis, focusing on Wi-Fi access points roaming mechanisms and signal strength heatmap generation.
- Designed a procedure-oriented data pipeline architecture including coordinate construction, data collection, data preprocessing, heatmap generation, and individual AP analysis.

### 2021 Sim & Real Experiment on Baidu Apollo D-kit Autonomous Vehicle

Haining, China

- Reproduced Apollo7.0 simulation experiments on CARLA, testing across 10 maps.
- Assembled the LiDAR, Radar and Camera module for Baidu Apollo D-kit Autonomous Vehicle.
- Constructed the high-resolution LiDAR map of ZJU international campus via Baidu Apollo D-kit Vehicle.

---

## ACADEMIC CONTEST

### 2023 Shell Eco-marathon Autonomous Programming Competition

ROS, Carla, Python,

- Developed path planning, perception, and control modules for simulation autonomous vehicles using the Robot Operating System (ROS) stack provided by the competition.
- Utilized the CARLA simulator with the Unreal Engine to test our vehicle in a simulated environment, with the goal of achieving the most efficient path planning according to the competition's ranking criteria.

### 2022 International Mathematical Contest in Modeling (Honorable Mention)

MATLAB, Python

- Addressed the issue of water scarcity in the Colorado River in the United States by constructing a mathematical model for water-to-electricity supply using dynamic programming and Monte Carlo algorithm.
- Used SPSS time series analysis tools to predict water demand for each state, generated a demand matrix for the water system. By sensitivity analysis, demonstrated strategies for addressing conditions such as rapid depletion of water resources, involvement of renewable energy technologies, and application of conservation measures.

### 2021 DJI RoboMaster University Robotics Competition

Python, PyTorch, C/C++

- Assisted in developing and optimizing vision algorithms for object detection via YOLOv4, contributing to auto-aiming and shooting system for real-time target engagement, improving hit accuracy by 20%.
- Supported the integration of vision systems on Jetson Nano with embedded platforms on DJI manifold2, achieving a 30 FPS during live competition on the industrial camera.

---

## SERVICE

Reviewer for RSS2025, NeurIPS2025, ICLR2026

GRASP Lab Student Faculty Industry (SFI) Committee Member at UPenn

2025

Member, Robotics Entrepreneur Club (PERC) at UPenn

2025

Teaching Assistant in CIS5800: Machine Perception

2025

Teaching Assistant in MEAM5200: Introduction to Robotics

2025

Teaching Assistant in PHIL 206: Early Modern Philosophy

2024

Head Teaching Assistant in MATH213: Introduction to Discrete Mathematics

2023

Founder & President, [PhiloCoffee Club](#) at ZJU

2023