Yanjia Huang

(+1) 479-306-0263 | yanjia_0812@tamu.edu

Summary

Research Interests: My research primarily lies in **embodied AI** and **dexterous manipulation** of robotics hands. My goal is to **develop intelligent, language-driven robotic agents** capable of **performing complex tasks in real-world** environments, enhancing human productivity and quality of life. Currently, I concentrate on enhancing the capabilities of robots in real-word **grasping** and **navigation** tasks.

Highlight: 2 years of research experience with solid mathematical and theoretical background; 2 years of experience with real robotic platforms: Kinova gen2, Franka and Realman compound robot.

Education

Texas A&M University — *College Station, Texas, United States*

M.S. in Engineering Technology — *Aug. 2024 - Present*

- Research Interest: Robotics, Vision-Language-Model(VLM)

Liverpool John Moores University — *Liverpool, England*

B.Eng.(Hons) in Mechanical Engineering — *Sep. 2020 - Jun. 2024*

- **First Class Honours**
- Major GPA: **3.81/4.0, 72/100**; Overall GPA: 3.52/4.0

University of Shanghai for Science and Technology — *Shanghai, China*

B.Eng. in Mechanical Engineering — *Sep. 2020 - Jun. 2024*

- Major GPA: **87/100**, Ranked **6** out of 42
- Recipient of the Dean's Scholarship

Publications

(*equal contribution)

[1] Zero-shot Object Navigation with Vision-Language Models Reasoning — *ICPR 2024*

Congcong Wen[1*], Yisiyuan Huang[2*], **Yanjia Huang[2*]**, Wenyu Han[2], Shuaihang Yuan[1], Yu-Shen Liu[3] and Yi Fang[1]

[2] A self-healing composite actuator for multifunctional soft robot via photo-welding — *Composites Part B: Engineering*

Mingxia Liu[1 2], Shu Zhu[3], **Yanjia Huang[1]**, Zihui Lin[2], Weiping Liu[1 4], Lili Yang[1] and Dengteng Ge[2]

[3] Can Large Vision Language Models Read Maps like a Human? — *NeurIPS 2025 (under review)*

Shuangyu Xie[1*], Xing Shuo[1*], **Yanjia Huang[1]**, Zezhou Sun[2], Zhengzhong Tu[1]

[4] PANDORA: Diffusion Policy Learning for Dexterous Robotic Piano Playing with Oracle Reward Assessment — *IROS 2025 (under review)*

Yanjia Huang[1*], Renjie Li[1], Zhengzhong Tu[1]

[5] VISTA: Visual Imagination with Scheduler for Task-Aware Navigation — *CoRL 2025 (under review)*

Yanjia Huang[1*], Mingyang Wu[1], Renjie Li[1], Zhengzhong Tu[1]

[6] Branching into Noise: A Monte Carlo Diffusion Framework for Token-Level Long-Horizon Robot Planning — *CoRL 2025 (under review)*

Yanjia Huang[1*],Shuo Liu[2*], Peiran Li[2], Xing Shuo[1], Zhengzhong Tu[1]

[7] Lock, Forecast, Act: End-to-End Object-Centric 3-D Gaussian Splatting and Diffusion Trajectory Prediction for Dynamic Grasping — *ICCV Workshop 2025*

Yanjia Huang[1*],Mingyang Wu[1*], Ruijie Ye[2], Renjie Li[1], Zhengzhong Tu[1]

Research Experience

Reality.Inc — *Berkeley, CA, U.S.A*

ML Agent Team Tech Lead, Scam.AI — *April. 2025 - Present*

- Successfully secured a **$300K angel round**
- As ML Agent Team Lead, spearheading the development of AI-driven SMS phishing attack agents to validate and strengthen our defense capabilities.

Texas A&M University — *College Station, TX, U.S.A*

Research Assistant, Advised by Prof. Zhengzhong Tu — *Aug. 2024 - Present*

- **Research Topics**: Robotics, Embodied Agent, Computer Vision, Vision-Language-Model(VLM)
- Proposed a Benchmark on human-readable map.

Noah's Ark Lab — *Shanghai, China*

Research Intern — *Oct. 2023 - Aug. 2024*

- **Research Topics**: Real-World Embodied AI, Voice Command-Driven Robotics
- Creating a compound robot that can perform complex tasks like pouring water in challenging settings such as offices and can navigate to objects all through voice commands.

New York University — *New York, U.S.A*

Research Assistant at Multimedia and Visual Computing Lab, Advised by Prof. Yi Fang — *May. 2023 - Sept. 2023*

- **Research Topics**: Autonomous Vehicle Navigation, Vision-Based Navigation, Motion and Path Planning
- Proposed a novel Vision Language model named VLTNet (Vision Language model with Tree-of-thoughts NETwork) for Language-driven Zero-Shot Object Navigation (L-ZSON) and it leverages a ToT reasoning framework, enabling multi-path reasoning, self-evaluation, and anticipation of future demands.

Fudan University — *Shanghai, China*

Research Assistant, Advised by Prof. Yanwei Fu — *Sept. 2022 - Present*

- **Research Topics**: Dexterous Robotic Hand Manipulation, Machine Learning in Robotic Control
- Contributed into a pioneering method for approximating the 3D shape of liquid using the 6-DoF pose of source containers and the estimated liquid mask. This approach is innovative in that it models 3D liquid from a single image without relying on temporal information.

Coursework Projects

The Design of a Single Level Standard Spur Gear Reducer — *Shanghai, China*

Designer — *2020-2021*

- Engineered a single-level standard spur gear reducer, establishing transmission solutions and defining gear ratios, resulting in a robust gearbox design.
- Developed V-shaped conveyor belts with strong key linkages and couplings, achieving optimal tension and speed in software simulations.

Design of Chassis for Transportation Automated Guided Vehicle — *Shanghai, China*

Project Leader — *Sept.2022 - Dec.2022*

- Developed a sturdy, durable, and cost-effective chassis for an Automated Guided Vehicle (AGV), incorporating a Navigation System and Travel System, and optimizing material selection.
- Aimed to design an AGV chassis for transporting food and medical equipment in hospitals and campuses, focusing on enhancing the vehicle's dynamic performance and reducing friction losses.

Skills

Programming C, C++, Python, SQL, Linux, MATLAB, ROS
Platforms Kinova, Franka, Realman
Design Solidworks, CAD, ArtCAM, COMSOL, Arduino

Honors & Awards

2021 **First Prize**, RoboMaster University Technical Challenge (RMUT) 2021 Regional Competition (Central Region) — *Jiangsu, China*
2021 **First Prize (Top 5%)**, Internet+ Innovation and Entrepreneurship Competition — *Shanghai, China*
2023 **the Dean's Scholarship**, Sino-British College, University of Shanghai for Science and Technology — *Shanghai, China*