# Mo Xu

Madison, WI | xmo@cs.wisc.edu | Website | LinkedIn | GitHub | Google Scholar

**Research Interests:** Robot Learning, Dexterous Manipulation, Reinforcement Learning, Sim-to-Real Transfer, Embodied AI, Vision-Language-Action Models, World Models

## Education

**University of Wisconsin - Madison** — Madison, WI
*Ph.D. in Computer Sciences* — *2024 – Present*

**University of Michigan - Ann Arbor** — Ann Arbor, MI
*Master in Electrical and Computer Engineering* — *2022 – 2023*

## Selected Publications

**Mo Xu**, et al. "Grasp2Twist: Learning Bimanual Dexterous Jar Opening by Reinforcement Learning." **Under review at ICRA 2027**; accepted to an **IROS 2026 Workshop**. [Project]
Jingxin Du, **Mo Xu**, et al. "Toward Real-time High-fidelity Immersive Robotic Telepresence via Online 3D Gaussian Splatting." **RA-L 2026**; presented at **IROS 2026**. [Paper]

## Research & Industry Experience

**Bimanual Dexterous Manipulation for Real-World Jar Opening** — 2025 – 2026
*University of Wisconsin–Madison | NSF HAND ERC*

- Developed a unified **PPO** policy for **bimanual dexterous manipulation**, learning grasp acquisition, grasp-to-twist transition, and sustained lid twisting without explicit task-phase input; achieved **88% success (53/60 trials)** across six real-world household containers.
- Introduced a geometric **hand–object enclosure representation** with a continuous measure to guide grasp formation and a binary indicator to encode the grasp-to-twist transition within a unified policy.
- Designed a three-stage curriculum for contact-reconfiguration exploration and sim-to-real adaptation, enabling emergent **finger gaiting** to sustain lid rotation beyond the motion range of a fixed grasp.
- Scaled asymmetric PPO to **4,096 GPU-parallel environments** in **mjlab/MuJoCo** and deployed the policy zero-shot on dual Franka arms with TESOLLO dexterous hands and RGB-D object state estimation.

**Non-Verbal Human Behavior Understanding for Interactive VLA Systems** — 2026
*University of Wisconsin–Madison | NSF HAND ERC*

- Prototyped a hierarchical **VLM–VLA** system for non-verbal HRI that maps visual observations of human behaviors and human–object interactions to context-aware robot instructions.
- Integrated fast and reasoning **Qwen VLMs** with $\pi_{0.5}$ execution, enabling high-level behavior interpretation and planning to drive vision-conditioned low-level robot actions.

**Immersive Robotic Telepresence via Online 3D Gaussian Splatting** — 2025 – 2026
*University of Wisconsin–Madison*

- Engineered and deployed a real-world **VR robotic telepresence** system integrating Stretch 3, RealSense RGB-D sensing, Meta Quest/OpenXR, and a remote online 3D reconstruction and rendering pipeline.
- Implemented a **ZeroMQ-based RGB-D streaming pipeline** for external RealSense cameras, bypassing ROS 2 image topics for low-latency camera transport between the robot and remote host.
- Integrated **OpenXR/Meta Quest teleoperation** with the Stretch ROS 2 stack for bidirectional robot control and immersive live remote exploration.

**ABB** — Shanghai, China
*Electronics Engineering Intern* — *Jun. 2021 – Aug. 2021*

## Technical Skills

**Learning & Control:** PPO, Asymmetric Actor–Critic, Domain Randomization, Curriculum Learning, VLM/VLA Integration
**Robotics & Simulation:** MuJoCo, mjlab, Isaac Sim/Isaac Lab, RSL-RL, ROS/ROS 2, MoveIt, Franka Panda, TESOLLO DG-3F-B, Stretch 2/3, RGB-D Perception
**Programming & Systems:** Python, C/C++, PyTorch, CUDA/TensorRT, Docker, Linux, Git, ZeroMQ, OpenXR