# LEI ZHOU

## EDUCATION

**National University of Singapore** **Jan 2021 - Present**

Ph.D. in Mechanical Engineering (Advisor: Prof. Marcelo H. Ang Jr.)

- GPA: 4.33/5.00
- Relevant Coursework: Machine Vision; Deep Learning for Robotics; Digital Human

**Huazhong University of Science and Technology** **Sep 2014 - Jun 2018**

B.E. in Mechanical Engineering

- GPA: 3.80/4.00

## CORE SKILLS

- Programming: Python, PyTorch, ROS.
- Computer Vision: Human Pose Estimation, 3D Reconstruction, Object Pose Estimation.
- Robotics: Robotic Grasping, Reinforcement Learning.
- Language: Mandarin (native), English (fluent).

## WORK EXPERIENCE

**Intern** **May 2024 – Present**

Microsoft Research Asia – Beijing (Mentor: Jiaolong Yang)

- Contributing to **Embodied AI** research focusing on **Vision-Language-Action (VLA) model pre-training** and **data scaling**.
- **Led** development of a high-fidelity **camera motion estimation** and **egoview hand mesh reconstruction pipeline** for daily monocular egocentric videos. Achieved **1.4 cm Absolute Trajectory Error (ATE)** by leveraging a spatio-temporal transformer for camera-space reconstruction and transforming results to world space using egoview camera motion.
- **Designed** a robust **multiview hand mesh reconstruction system** for the **EgoExo4D dataset** for high-quality 3D hand pose annotation regeneration, validating over **70%** of frames across **10 million samples** to support large-scale VLA model training.
- **Engineered** a **Transformer-based reinforcement learning model** for dexterous robotic grasping, **achieving a 91.2% success rate on seen objects** (88.3% on new categories) in simulation, and **developed a real-sim-real pipeline** integrating **NeRF-based object reconstruction**, improving real-world grasping performance.

**Research Assistant** **Jan 2021 – May 2024**

Advanced Robotics Centre, National University of Singapore

- **Implemented Instant-NGP** for **3D workspace geometry reconstruction**, enhancing **detection and localization** capabilities for vertical farming robots, improving spatial awareness and task accuracy.
- Optimized control systems for **Franka Emika Panda** and **Kinova Movo**, streamlining integration with industrial applications.
- **Created synthetic datasets** and **trained instance-level object pose estimation algorithms**, facilitating **robotic grasp generation** and **visual affordance detection** in real-world applications, **increasing task efficiency by 25%**.

## RESEARCH & PROJECT HIGHLIGHTS

**Hand Mesh Reconstruction from Egocentric Videos** **Nov 2024 – Present**

- Developed a **framework** for recovering **world-space hand motion** from daily monocular egocentric videos.
- Utilized **SLAM** and **depth estimation** methods to achieve **1.4 cm** Absolute Trajectory Error (ATE) in camera motion.
- Implemented a **spatio-temporal transformer-based approach** to reconstruct camera-space hand meshes from RGB image sequences.
- Transformed camera-space hand meshes to world space for **accurate hand motion tracking**.

**Reinforcement Learning for Vision-Based Dexterous Robotic Grasping** **Aug 2024 – Nov 2024**

- Created a **universal Transformer-based model** via **offline distillation** from individually trained RL policies on **3,200 objects**.
- In simulation, it achieves up to **a 91.2%** success rate on seen objects (**88.3%** on new categories) in a **state-based** setting and **88.9%** (**86.8%**) in a **vision-based** setting.
- Designed a **real-sim-real pipeline** integrating **NeRF-based** object reconstruction into RL environments. Captured real-world objects as meshes, trained grasping policies in simulation, and transferred them back for real-world testing.
- Achieved state-of-the-art performance, with a paper accepted by **CVPR 2025**.

**Robust Multiview Hand Mesh Reconstruction** **May 2024 – Aug 2024**

- Tackled **severe occlusions** during human manipulation and **occasional camera time misalignment** in the EgoExo4D dataset by developing a **robust multiview hand reconstruction** pipeline.
- **Applied triangulation-based optimization** for multiview consistency, significantly improving reconstruction accuracy.
- **Incorporated smoothing techniques** to ensure temporal coherence across video frames, yielding high-precision 3D keypoint annotations despite challenging occlusions.
- Validated **70%+** of frames (10 million total) for **4D hand pose** estimation and **VLA** model pretraining.

**Diffusion-based Multi-Hands Robotic Grasp Generation** **Nov 2023 - Mar 2024**

- Developed a **diffusion-based model** to generate grasp poses for multiple robotic dexterous hands.
- Introduced **visual affordance detection** and **open-vocabulary** analysis to filter functional grasp candidates.
- Achieved **44.73%** overall success rate on the **MultiDex Dataset**, improving generalization across multiple hand types, with a paper accepted by **ISRR 2024.**

**Dynamic Scene Reconstruction for Robotic Grasping** **Jun 2023 - Sep 2023**

- Utilized **SDF-based methods** to reconstruct novel objects from multi-view images.
- Designed a **dynamic scene reconstruction pipeline** that completed object point clouds in real time (**9.2 FPS**) by leveraging tracked object poses.
- Achieved **state-of-the-art performance** on the **GraspNet-1Billion benchmark** for robotic grasping tasks, with a paper accepted by **ICRA 2024**.

**Category-Level Object Pose Estimation** **Jan 2022 - Mar 2023**

- Developed a network that **completes partial point clouds** from depth camera data and **reconstructs objects** in canonical space by **deforming a shape prior**.
- Designed a robust **3D shape-matching module** to align reconstructed objects with observed partial point clouds for **pose and size estimation**.
- Achieved **state-of-the-art performance**, with a paper accepted by **IROS 2023**.

## SELECTED PUBLICATIONS


- Wenbo Wang, Fangyun Wei, **Lei Zhou**, Xi Chen, Lin Luo, Xiaohan Yi, Yizhong Zhang, Yaobo Liang, Chang Xu, Yan Lu, Jiaolong Yang, and Baining Guo, "UniGraspTransformer: Simplified Policy Distillation for Scalable Dexterous Robotic Grasping", CVPR 2025
- **Lei Zhou**, Haozhe Wang, Zhengshen Zhang, Zhiyang Liu, Francis EH Tay, and Marcelo H. Ang Jr., "You Only Scan Once: A Dynamic Scene Reconstruction Pipeline for 6-DoF Robotic Grasping of Novel Objects", ICRA 2024
- Zhengning Zhou, **Lei Zhou**, Shengxin Sun, and Marcelo H. Ang Jr., "A Robust and Efficient Robotic Packing Pipeline with Dissipativity-Based Adaptive Impedance-Force Control", IROS 2024
- Zhiyang Liu, Ruiteng Zhao, **Lei Zhou**, Chengran Yuan, Yuwei Wu, Sheng Guo, Zhengshen Zhang, and Marcelo H. Ang Jr., "3D Affordance Keypoint Detection for Robotic Manipulation", IROS 2024
- Zhengshen Zhang, **Lei Zhou**, Chenchen Liu, Chengran Yuan, Sheng Guo, Ruiteng Zhao, Marcelo H. Ang Jr., and Francis EH Tay, "DexGrasp-Diffusion: Diffusion-Based Unified Functional Grasp Synthesis Method for Multi-Dexterous Robotic Hands", ISRR 2024
- **Lei Zhou**, Zhiyang Liu, Runze Gan, Haozhe Wang, and Marcelo H. Ang Jr., "DR-Pose: A Two-stage Deformation-and-Registration Pipeline for Category-level 6D Object Pose Estimation", IROS 2023