

# Yifan Yin

📍 Baltimore, MD 📩 yyin34@jhu.edu 🌐 yifanyin11.github.io 💬 linkedin.com/in/yifan-yin11

## Research Interests

- Robot learning (language grounding, physical reasoning)
- 3D world modeling and representation learning
- Human–robot collaboration

## Education

### Ph.D. in Computer Science

*Johns Hopkins University (JHU)*

*Baltimore, MD*

*Aug 2024 – Present*

- Research areas: Embodied AI, human–robot interaction.

### M.S.E. in Robotics

*Johns Hopkins University (JHU)*

*Baltimore, MD*

*Aug 2021 – May 2023*

- Thesis Focus: Perception systems, manipulation.

### B.Eng. in Mechanical Design & Automation

*Sichuan University (SCU)*

*Chengdu, China*

*Sep 2017 – Jun 2021*

- Dean’s List; Academic Star

## Experience

### Research Intern

*Social Cognitive AI (SCAI) Lab, JHU*

*Baltimore, MD*

*Jul 2024 – Aug 2024*

- Developed *PartInstruct*, the first benchmark for training & evaluating part-level manipulation models.
- Built a dataset of 513 objects across 14 categories and 1,302 tasks in 16 classes with detailed part annotations.
- Generated 10,000+ expert demos with task instructions, part-based skill chains, and 3D annotations.
- Designed a comprehensive test suite to assess generalization across novel states, objects, and tasks.
- Benchmarked SOTA VLAs and VLM hierarchical planners; revealed challenges in part grounding, 3D part-level action prediction, and long-horizon fine-grained manipulation.

### Computer Vision Intern

*PediaMetrix Inc.*

*Rockville, MD*

*Jul 2023 – Jul 2024*

- Built a classifier for skull abnormality types; achieved 94.6% sensitivity and 99.3% specificity.
- Prototyped smartphone-based 3D cranial reconstruction using CV, DL, and shape modeling.

## Publications (\* equal contribution)

### Preprints & Under Review

- J. Chen, T. Zhang, X. Xuan, J. He, **Y. Yin**, H. Shi, S. Ye, R. Yuan, X. Li, T. Shu, A. Yuille. *Real-Time Generative DeOcclusion for Visual Geometry and Robotics*. Submitted to *CVPR*, 2026.
- L. Ying, X. Li, S. Aarya, Y. Fang, **Y. Yin**, J. X. Liu, S. Tellex, J. B. Tenenbaum, T. Shu. *Pragmatic Embodied Spoken Instruction Following in Human-Robot Collaboration with Theory of Mind*. arXiv:2409.10849, 2025. Preprint.

### Peer-reviewed Conference Papers

- **Y. Yin\***, Z. Han\*, S. Aarya, S. Xu, J. Wang, J. Peng, A. Wang, A. Yuille, T. Shu. *PartInstruct: Part-level Instruction Following for Fine-grained Robot Manipulation*. *Robotics: Science and Systems (RSS)*, 2025.
- **Y. Yin**, Y. Wang, Y. Zhang, R. H. Taylor, B. P. Vagvolgyi. *Applications of Uncalibrated Image-Based Visual Servoing in Micro- and Macroscale Robotics*. *IEEE CASE*, 2023.

## In-Progress Projects

---

### Generative 3D World Models as 3D Grounded Belief Models for Embodied Agents

- Developed a generative 3D world model that predicts unseen scene geometry from partial observations in an explicit and actionable 3D representation.
- Enabled key embodied capabilities: spatially consistent scene expansion, object permanence, semantically informed future prediction, and fast online inference for real-time operation.
- Integrated the model with an embodied agent to continually update world beliefs and replan future motion, supporting closed-loop navigation. Demonstrated advances on challenging embodied tasks.

### Real-Time Generative DeOcclusion for Visual Geometry and Robotics

- Proposed a fast, general-purpose de-occlusion pipeline that selects informative viewpoints, performs 2D amodal completion, and fuses results via depth-aware 3D integration from RGB and point clouds.
- Produced spatially coherent, occlusion-aware reconstructions without extra supervision or heavy optimization; improved full-scene and occluded-region metrics across indoor, outdoor, and in-the-wild datasets.
- Applied to 3D scene understanding to yield substantial gains in object-level perception; compatible with downstream robotics tasks.

## Skills

---

**Programming:** Python, C/C++, MATLAB, C#, SQL

**ML/Robotics:** PyTorch, ROS/ROS2, SLAM, Sensor Fusion, Kinematics, Visual Servoing, Motion Planning

**Computer Vision:** 3D Scene Understanding, 3D Perception, Neural Rendering, 3D Reconstruction

**Tools:** Git, Slurm, AI2-THOR, Isaac Sim, MuJoCo, PyBullet, Habitat, Unity3D

## Service

---

- **Workshop & Tutorial Organizers**
  - RSS 2025 Workshop on Continual Robot Learning from Humans
- **Reviewer**
  - ICLR 2025 Workshop WRL
  - NeurIPS 2025 Workshop LAW

## Media Coverage

---

- “Helping robots become a *part* of our world,” *JHU WSE News*, September 2025.