

Robots That Learn From Demonstration

Hashim Ismail

The State of Imitation Learning

- Vision-Language-Action (VLA) models
- Foundation models for robotics (e.g., $\pi 0.5$ -style systems)
- These models can make robots learn to:
 - Making coffee
 - Folding laundry
 - Multi-step kitchen tasks
 - Use tools
 -



Video from: Black, Kevin, Noah Brown, Danny Driess, Adnan Esmail, Michael Equi, Chelsea Finn, Niccolo Fusai et al. " $\pi 0$ ": A Vision-Language-Action Flow Model for General Robot Control." *arXiv preprint arXiv:2410.24164* (2024).

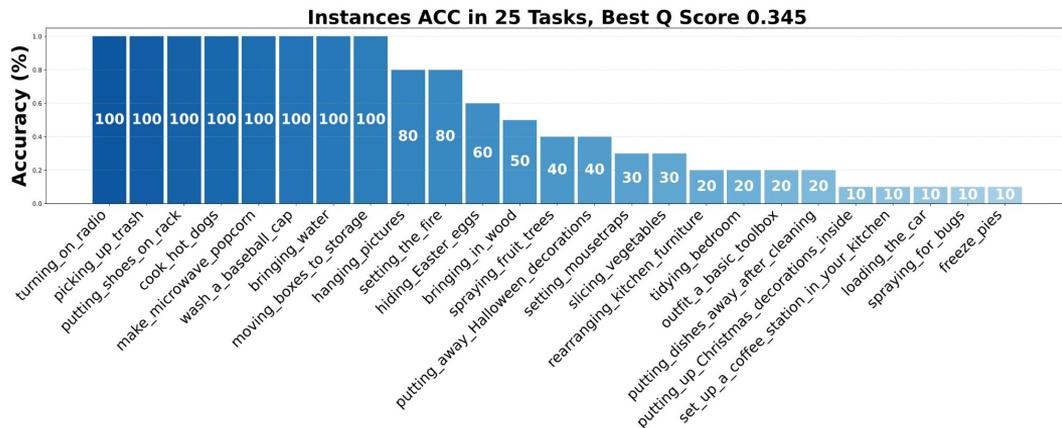
The Bottleneck: Data & Hard Tasks

On the other side, most task requires demonstration data to be:

- Clean
- Structured
- Short-horizon
- Massive

And even that, many models fail in:

- Contact-rich settings
- Deformable objects
- Long-horizon tasks
- Out-of-distribution scenarios



Performance of $\pi 0.5$ model on BEHAVIOR-1K tasks, a benchmark focus on long-horizon tasks. The model can reach higher than 80% of success rate on only 10 out of 50 tasks.

Table from Bai, Junjie, Yu-Wei Chao, Qizhi Chen, Jinwei Gu, Moo Jin Kim, Zhaoshuo Li, Xuan Li et al. "Openpi Comet: Competition Solution For 2025 BEHAVIOR Challenge." *arXiv preprint arXiv:2512.10071* (2025).

Let's try it yourself

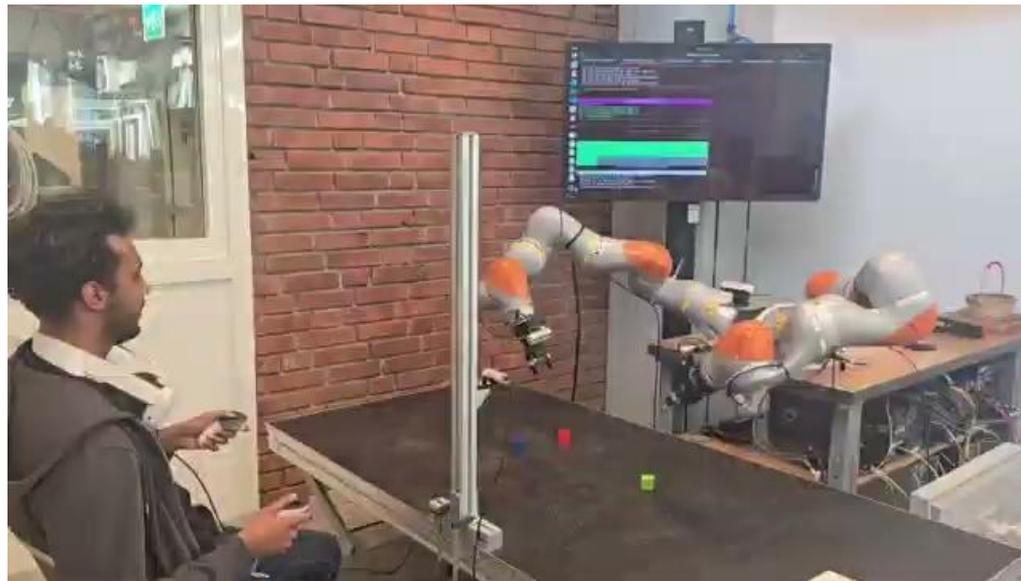
This project aim to: Build Hard Tasks. Collect Real Data. Train Real Models. Run Real Robot

We have:

- Dual arm Kuka robots that can move in command, and be teleoperated with Meta Quest 3

You have the chance to:

1. Study modern imitation learning methods
2. Design challenging manipulation tasks
3. Collect demonstrations on this real robot setup
4. Contribute to a growing research dataset



Video showing people collect data with VR based teleoperation system on our Dual arm Kuka robots

Models You Can Explore & Tasks You Can Design

Models You Can Implement / Benchmark

1. The OG: Behavior Cloning (BC)
2. Diffusion-based Models
 - a. Diffusion Policy
 - b. 3D Diffusion Policy
3. Action Chunk Transformer
4. Vision-Language-Action Models
 - a. π family (π_0 , $\pi_{0.5}$, $\pi_{0.6}$)
 - b. OpenVLA
 - c. Octo
5. Normalizing Flows Policy
6.

Example Task Categories You can try

1. Precision & Contact
 - a. Peg-in-hole
 - b. Tool alignment
2. Deformable Objects
 - a. Cable routing
 - b. Cloth manipulation
 - c. Food handling
3. Long-Horizon
 - a. Multi-step assembly
4. Evaluation
 - a. Success rate
 - b. Generalization to new objects
 - c. Robustness to perturbation

Skills & Expectations

Core Requirements (Ideal)

- Strong understanding of Deep Learning
- Comfortable with: Python, especially PyTorch
- Neural network training pipelines
- Debugging training instability

Robotics Stack (Preferred)

Robotics Stack

- ROS2 (Node, Topics / Services, Sensor streams)
- Basic robot kinematics
- Working with real hardware (safety-aware)

Bonus Knowledge

- Experience with:
 - Transformers
 - Diffusion models
 - Computer vision
 - Dataset management
- Linux fluency
- GPU training workflows

Don't Have All the Skills Yet?

That's okay. If you have serious interest, strong fundamentals, and motivation to learn fast, We encourage you give it a shot.

Thank You

Contact hashim.ismail@cs.lth.se

or

jialong.li@cs.lth.se

If you have any questions