

# LLARVA: VISION-ACTION INSTRUCTION TUNING ENHANCES ROBOT LEARNING

Tech ID: 33981 / UC Case 2025-111-0

## PATENT STATUS

Patent Pending

## BRIEF DESCRIPTION

Bridging the gap between a language model's next-word prediction and physical robot control, researchers at UC Berkeley have developed LLARVA (Large Language model for Robotic Vision and Action). This model utilizes a novel vision-action instruction tuning method that allows a robotic device to handle various tasks and environments without task-specific fine-tuning.

## SUGGESTED USES

»

General-Purpose Robot Assistants: Equipping service robots with the ability to follow natural language instructions for varied household tasks like "clear the table" or "stack the boxes."

»

Multi-Robot Industrial Automation: Implementing a single, unified model that can control different robot models (e.g., Franka or UR5) across diverse manufacturing scene configurations.

»

Rapid Task Deployment: Enabling robots in warehouse environments to switch between novel manipulation tasks instantly via simple text-based prompts and visual context.

»

Enhanced Teleoperation: Providing operators with predictive visual traces that show the robot's intended path, improving the precision of remote control in complex environments.

»

Robot Skill Acquisition: Serving as a foundation model for learning complex, long-horizon manipulation sequences through instruction-based "waypoint" prediction.

## ADVANTAGES

»

Superior Generalization: The model can adapt to different robot configurations and environments because it is trained on diverse, large-scale datasets rather than specialized niche data.

»

Scalable Data Efficiency: By leveraging the Open X-Embodiment dataset, the system utilizes millions of existing trajectories, reducing the need for expensive, manual real-world demonstrations.

»

Zero-Shot Task Execution: LLARVA can often perform new tasks correctly the first time by reasoning through the structured language prompts that define the robot type and control mode.

»

Improved Spatial Awareness: The auxiliary task of waypoint prediction provides the robot with better fine-grained localization, leading to higher success rates in contact-rich tasks like stacking.

## RELATED MATERIALS

### ADDITIONAL TECHNOLOGIES BY THESE INVENTORS

- ▶ [Pre-Training Auto-Regressive Robotic Models With 4D Representations](#)
- ▶ [Humanoid Locomotion As Next Token Prediction](#)
- ▶ [In-Context Learning Enables Robot Action Prediction in LLMs](#)

## CONTACT

Laleh Shayesteh  
[lalehs@berkeley.edu](mailto:lalehs@berkeley.edu)  
tel: 510-642-4537.



## INVENTORS

» Darrell, Trevor J.

## OTHER INFORMATION

### CATEGORIZED AS

» [Computer](#)

» [Software](#)

» [Research Tools](#)

» [Other](#)

» [Engineering](#)

» [Robotics and Automation](#)

### RELATED CASES

2025-111-0

