# PRE-TRAINING AUTO-REGRESSIVE ROBOTIC MODELS WITH 4D REPRESENTATIONS

Tech ID: 33979 / UC Case 2025-109-0

## PATENT STATUS

Patent Pending

## BRIEF DESCRIPTION

Current methods for training robotic policies often struggle with efficiently learning from rich, time-varying visual data, leading to brittle and data-intensive solutions. This innovation, developed by UC Berkeley researchers, addresses this challenge by introducing a robotic system that utilizes four-dimensional (4D) representations estimated directly from videos to pre-train and test an auto-regressive machine learning transformer model. By explicitly encoding space and time in a unified representation, the system allows the transformer model to leverage a much richer context than standard 2D image or 3D point cloud approaches, facilitating the learning of complex, long-horizon tasks and improving the generalization capabilities of the resulting policy. The use of 4D representations significantly enhances the policy's understanding of the dynamic environment and object interactions compared to existing alternatives, enabling more robust and efficient training of robotic systems.

## SUGGESTED USES

» Pre-training foundational models for robotics: Creating general-purpose, auto-regressive robotic policies that can be quickly fine-tuned for various downstream tasks.

» Learning complex manipulation and locomotion skills from video demonstrations.

» Developing robust policies for autonomous navigation in dynamic, real-world environments.

» Improving simulation-to-reality transfer (Sim2Real) by providing a unified representation that bridges the gap between simulated and real-world data.

## ADVANTAGES

» Enhanced contextual understanding: The 4D representation explicitly captures the temporal and spatial evolution of the environment, providing richer context for the transformer model.

» Improved data efficiency: The system requires less labeled data for training effective policies due to the informative nature of the 4D input.

» Greater robustness and generalization: The resulting policy is more capable of handling novel scenarios and dynamic changes in the environment than policies trained on 2D or 3D data alone.

» Scalability for auto-regressive tasks: The architecture is specifically designed to leverage the temporal dependencies inherent in the 4D data for better performance on auto-regressive tasks like sequential decision-making.

## RELATED MATERIALS

### ADDITIONAL TECHNOLOGIES BY THESE INVENTORS

- Humanoid Locomotion As Next Token Prediction
- In-Context Learning Enables Robot Action Prediction in LLMs
- RealWorldPlay: Physical AI In-Situ Revisited

## CONTACT

Laleh Shayesteh
lalehs@berkeley.edu
tel: 510-642-4537.




## INVENTORS

» Darrell, Trevor J.

## OTHER INFORMATION

**KEYWORDS**

robotics, robot, 4D

**CATEGORIZED AS**

» **Computer**
» Software
» **Imaging**
» Software
» **Nanotechnology**
» Other
» **Sensors & Instrumentation**
» Scientific/Research
» **Engineering**
» Robotics and Automation

**RELATED CASES**

2025-109-0

▶ Llarva: Vision-Action Instruction Tuning Enhances Robot Learning

Berkeley
UNIVERSITY OF CALIFORNIA

**University of California, Berkeley Office of Technology Licensing**

2150 Shattuck Avenue, Suite 510, Berkeley,CA 94704

Tel: 510.643.7201 | Fax: 510.642.4566

https://ipira.berkeley.edu/ | otl-feedback@lists.berkeley.edu

© 2025, The Regents of the University of California

Terms of use | Privacy Notice