

Researchers expand ability of robots to learn from videos

June 20 2023



A team from Carnegie Mellon University's Robotics Institute used affordances to teach robots how to interact with objects. Credit: Carnegie Mellon University

New work from Carnegie Mellon University has enabled robots to learn household chores by watching videos of people performing everyday tasks in their homes.

The research could help improve the utility of robots in the home, allowing them to assist people with tasks like cooking and cleaning. Two robots successfully learned 12 tasks including opening a drawer, oven door and lid; taking a pot off the stove; and picking up a telephone, vegetable or can of soup.

"The [robot](#) can learn where and how humans interact with different objects through watching videos," said Deepak Pathak, an assistant professor in the Robotics Institute at CMU's School of Computer Science. "From this knowledge, we can train a model that enables two robots to complete similar tasks in varied environments."

Current methods of training robots require either the manual demonstration of tasks by humans or extensive training in a simulated environment. Both are time consuming and prone to failure. Past research by Pathak and his students demonstrated a novel method in which robots learn from observing humans complete tasks. However, [WHIRL](#), short for In-the-Wild Human Imitating Robot Learning, required the [human](#) to complete the task in the same environment as the robot.

Pathak's latest work, Vision-Robotics Bridge, or [VRB](#) for short, builds on and improves WHIRL. The new model eliminates the necessity of human demonstrations as well as the need for the robot to operate within an identical environment. Like WHIRL, the robot still requires practice to master a task. The team's research showed it can learn a new task in as little as 25 minutes.

"We were able to take robots around campus and do all sorts of tasks," said Shikhar Bahl, a Ph.D. student in robotics. "Robots can use this model to curiously explore the world around them. Instead of just flailing its arms, a robot can be more direct with how it interacts."

To teach the robot how to interact with an object, the team applied the concept of affordances. Affordances have their roots in psychology and refer to what an environment offers an individual. The concept has been extended to design and [human-computer interaction](#) to refer to potential actions perceived by an individual.

For VRB, affordances define where and how a robot might interact with an object based on human behavior. For example, as a robot watches a human open a drawer, it identifies the contact points—the handle—and the direction of the drawer's movement—straight out from the starting location. After watching several videos of humans opening drawers, the robot can determine how to open any drawer.

The team used videos from [large datasets](#) such as Ego4D and Epic Kitchens. [Ego4D](#) has nearly 4,000 hours of egocentric videos of daily activities from across the world. Researchers at CMU [helped collect some of these videos](#). [Epic Kitchens](#) features similar videos capturing cooking, cleaning and other kitchen tasks. Both datasets are intended to help train computer vision models.

"We are using these datasets in a new and different way," Bahl said. "This work could enable robots to learn from the vast amount of internet and YouTube videos available."

More information is available on the [project's website](#) and in a [paper](#) presented in June at the Conference on Vision and Pattern Recognition.

Provided by Carnegie Mellon University

Citation: Researchers expand ability of robots to learn from videos (2023, June 20) retrieved 10 September 2024 from <https://techxplore.com/news/2023-06-ability-robots-videos.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.