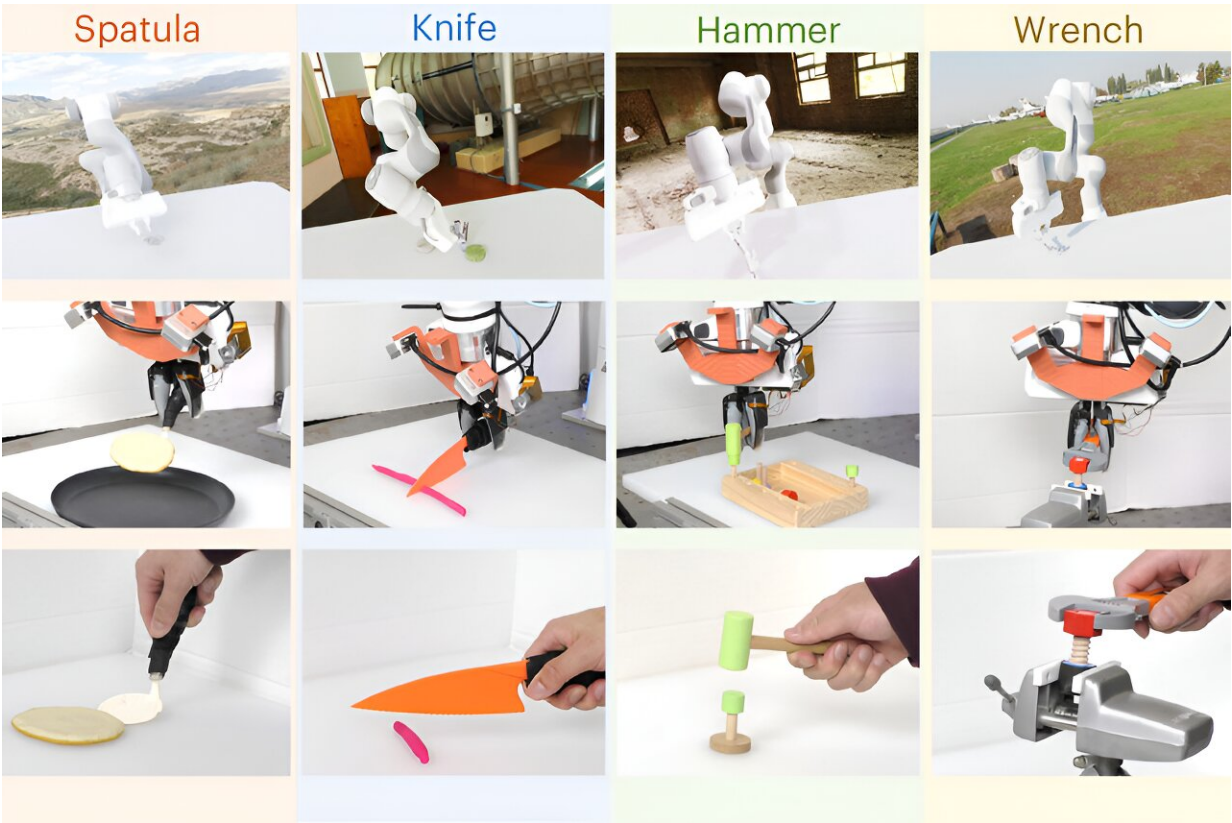


New technique combines data from different sources for more effective multipurpose robots

June 3 2024, by Adam Zewe



Three different data domains — simulation (top), robot tele-operation (middle) and human demos (bottom) — allow a robot to learn to use different tools.
Credit: Massachusetts Institute of Technology

Let's say you want to train a robot so it understands how to use tools and can then quickly learn to make repairs around your house with a hammer, wrench, and screwdriver. To do that, you would need an enormous amount of data demonstrating tool use.

Existing robotic datasets vary widely in modality—some include color images while others are composed of tactile imprints, for instance. Data could also be collected in [different domains](#), like simulation or human demos. And each dataset may capture a unique task and environment.

It is difficult to efficiently incorporate data from so many sources in one machine-learning model, so many methods use just one type of data to train a [robot](#). But robots trained this way, with a relatively small amount of task-specific data, are often unable to perform new tasks in unfamiliar environments.

In an effort to train better multipurpose robots, MIT researchers developed a technique to combine multiple sources of data across domains, modalities, and tasks using a type of generative AI known as [diffusion](#) models.

They train a separate diffusion model to learn a strategy, or policy, for completing one task using one specific dataset. Then they combine the policies learned by the diffusion models into a general policy that enables a robot to perform multiple tasks in various settings.

In simulations and real-world experiments, this training approach enabled a robot to perform multiple tool-use tasks and adapt to new tasks it did not see during training. The method, known as Policy Composition (PoCo), led to a 20% improvement in task performance when compared to baseline techniques.

"Addressing heterogeneity in robotic datasets is like a chicken-egg

problem. If we want to use a lot of data to train general robot policies, then we first need deployable robots to get all this data. I think that leveraging all the heterogeneous data available, similar to what researchers have done with ChatGPT, is an important step for the robotics field," says Lirui Wang, an [electrical engineering](#) and computer science (EECS) graduate student and lead author of a [paper](#) on PoCo posted to the *arXiv* preprint server.

Wang's co-authors include Jialiang Zhao, a mechanical engineering graduate student; Yilun Du, an EECS graduate student; Edward Adelson, the John and Dorothy Wilson Professor of Vision Science in the Department of Brain and Cognitive Sciences and a member of the Computer Science and Artificial Intelligence Laboratory (CSAIL); and senior author Russ Tedrake, the Toyota Professor of EECS, Aeronautics and Astronautics, and Mechanical Engineering, and a member of CSAIL.

The research will be presented at the [Robotics: Science and Systems Conference](#), held in Delft, Netherlands, July 15–19.

Combining disparate datasets

A robotic policy is a [machine-learning model](#) that takes inputs and uses them to perform an action. One way to think about a policy is as a strategy. In the case of a robotic arm, that strategy might be a trajectory, or a series of poses that move the arm so it picks up a hammer and uses it to pound a nail.

Datasets used to learn robotic policies are typically small and focused on one particular task and environment, like packing items into boxes in a warehouse.

"Every single robotic warehouse is generating terabytes of data, but it

only belongs to that specific robot installation working on those packages. It is not ideal if you want to use all of these data to train a general machine," Wang says.

The MIT researchers developed a technique that can take a series of smaller datasets, like those gathered from many robotic warehouses, learn separate policies from each one, and combine the policies in a way that enables a robot to generalize to many tasks.

They represent each policy using a type of generative AI model known as a diffusion model. Diffusion models, often used for image generation, learn to create new data samples that resemble samples in a training dataset by iteratively refining their output.

But rather than teaching a diffusion model to generate images, the researchers teach it to generate a trajectory for a robot. They do this by adding noise to the trajectories in a training dataset. The diffusion model gradually removes the noise and refines its output into a trajectory.

This technique, known as [Diffusion Policy](#), was previously introduced by researchers at MIT, Columbia University, and the Toyota Research Institute. PoCo builds off this Diffusion Policy work.

The team trains each diffusion model with a different type of dataset, such as one with human video demonstrations and another gleaned from teleoperation of a robotic arm.

Then the researchers perform a weighted combination of the individual policies learned by all the diffusion models, iteratively refining the output so the combined policy satisfies the objectives of each individual policy.

Greater than the sum of its parts

"One of the benefits of this approach is that we can combine policies to get the best of both worlds. For instance, a policy trained on real-world data might be able to achieve more dexterity, while a policy trained on simulation might be able to achieve more generalization," Wang says.

Because the policies are trained separately, one could mix and match diffusion policies to achieve better results for a certain task. A user could also add data in a new modality or domain by training an additional Diffusion Policy with that [dataset](#), rather than starting the entire process from scratch.

The researchers tested PoCo in simulation and on real robotic arms that performed a variety of tools tasks, such as using a hammer to pound a nail and flipping an object with a spatula. PoCo led to a 20% improvement in task performance compared to baseline methods.

"The striking thing was that when we finished tuning and visualized it, we can clearly see that the composed trajectory looks much better than either one of them individually," Wang says.

In the future, the researchers want to apply this technique to long-horizon tasks where a robot would pick up one tool, use it, then switch to another tool. They also want to incorporate larger robotics datasets to improve performance.

"We will need all three kinds of data to succeed for robotics: internet data, simulation data, and real robot data. How to combine them effectively will be the million-dollar question. PoCo is a solid step on the right track," says Jim Fan, senior research scientist at NVIDIA and leader of the AI Agents Initiative, who was not involved with this work.

More information: Lirui Wang et al, PoCo: Policy Composition from and for Heterogeneous Robot Learning, *arXiv* (2024). [DOI: 10.48550/arxiv.2402.02511](https://doi.org/10.48550/arxiv.2402.02511)

This story is republished courtesy of MIT News (web.mit.edu/newsoffice/), a popular site that covers news about MIT research, innovation and teaching.

Provided by Massachusetts Institute of Technology

Citation: New technique combines data from different sources for more effective multipurpose robots (2024, June 3) retrieved 25 June 2024 from <https://techxplore.com/news/2024-06-technique-combines-sources-effective-multipurpose.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.