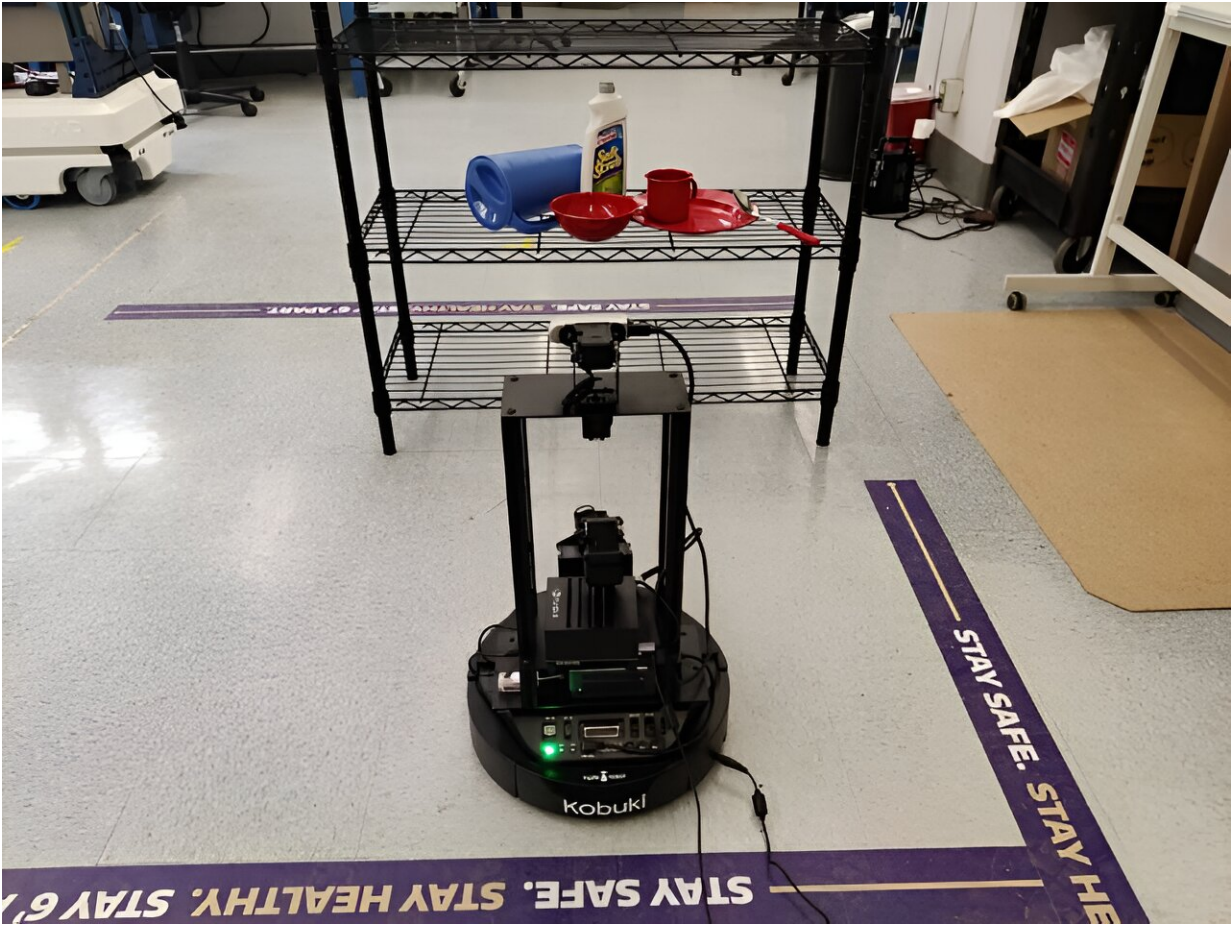


Q&A: Researcher discusses how newly developed method can help robots identify objects in cluttered spaces

February 7 2024, by Sarah McQuate



Researchers at the University of Washington have developed a method that teaches a low-cost robot to identify objects on a cluttered shelf. For the test, the robot (shown here in the center of the photo) was asked to identify all objects on the shelf in front of it. Credit: *IEEE Transactions on Robotics* (2023). DOI:

Imagine a coffee cup sitting on a table. Now, imagine a book partially obscuring the cup. As humans, we still know what the coffee cup is even though we can't see all of it. But a robot might be confused.

Robots in warehouses and even around our houses struggle to identify and pick up objects if they are too close together, or if a space is cluttered. This is because robots lack what psychologists call "object unity," or our ability to identify things even when we can't see all of them.

Researchers at the University of Washington have developed a way to teach robots this skill. The method, called THOR for short, allowed a low-cost robot to identify objects—including a mustard bottle, a Pringles can and a tennis ball—on a cluttered shelf. In [a recent paper](#) published in *IEEE Transactions on Robotics*, the team demonstrated that THOR outperformed current state-of-the-art models.

UW News reached out to senior author Ashis Banerjee, UW associate professor in both the industrial & [systems engineering](#) and mechanical engineering departments, for details about how robots identify objects and how THOR works.

How do robots sense their surroundings?

We sense the world around us using vision, sound, smell, taste and touch. Robots sense their surroundings using one or more types of sensors. Robots "see" things using either standard color cameras or more complex stereo or depth cameras. While standard cameras simply record colored and textured images of the surroundings, stereo and depth

cameras also provide information on how far away the objects are, just like our eyes do.

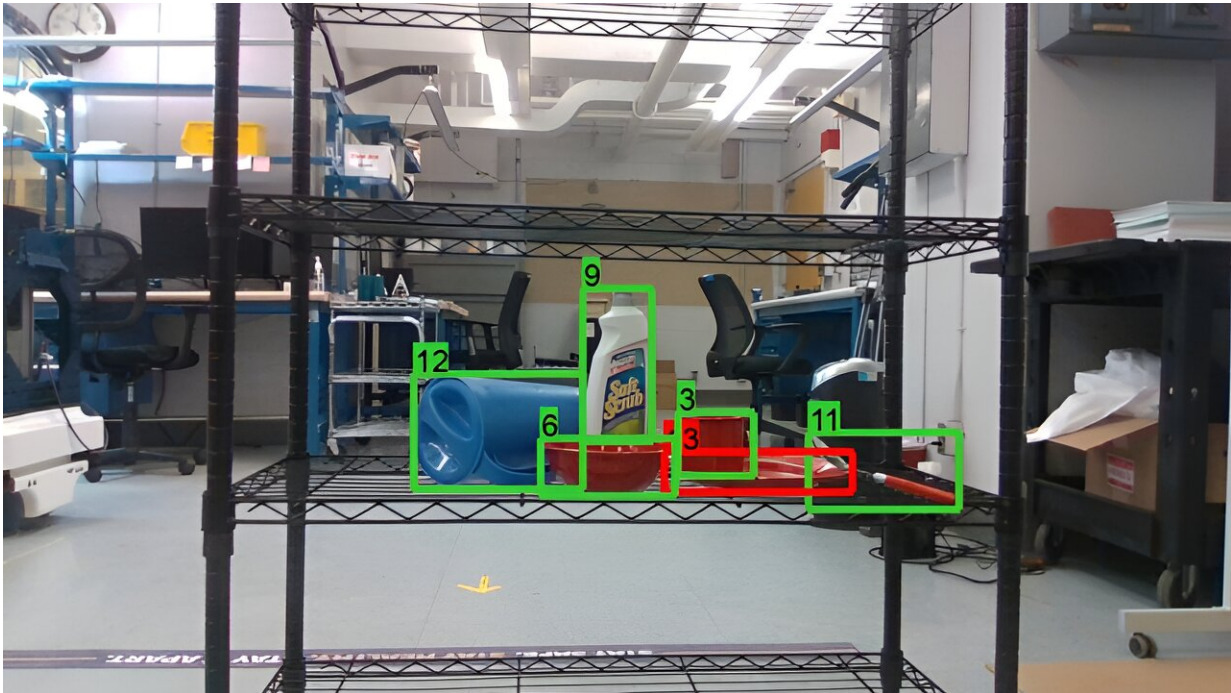
On their own, however, the sensors cannot enable the robots to make "sense" of their surroundings. Robots need a visual perception system, similar to the visual cortex of the human brain, to process images and detect where all the objects are, estimate their orientations, identify what the objects might be and parse any text written on them.

Why is it hard for robots to identify objects in cluttered spaces?

There are two main challenges here. First, there are likely a large number of objects of varying shapes and sizes. This makes it difficult for the robot's perception system to distinguish between the different object types. Second, when several objects are located close to each other, they obstruct the views of other objects. Robots have trouble recognizing objects when they don't have a full view of the object.

Are there any types of objects that are especially hard to identify in cluttered spaces?

A lot of that depends on what objects are present. For example, it is challenging to recognize smaller objects if there are a variety of sizes present. It is also more challenging to differentiate between objects with similar or identical shapes, such as different kinds of balls, or boxes. Additional challenges occur with soft or squishy objects that can change shape as the robot collects images from different vantage points in the room.



Green boxes shown here surround the objects that the robot correctly identified. Red boxes surround incorrectly identified items. Credit: *IEEE Transactions on Robotics* (2023). DOI: 10.1109/TRO.2023.3343994

So how does THOR work and why is it better than previous attempts to solve this problem?

THOR is really the brainchild of lead author Ekta Samani, who completed this research as a UW doctoral student. The core of THOR is that it allows the robot to mimic how we as humans know that partially visible objects aren't broken or entirely new objects.

THOR does this by using the shape of objects in a scene to create a 3D representation of each object. From there it uses topology, an area of mathematics that studies the connectivity between different parts of objects, to assign each object to a "most likely" object class. It does this

by comparing its 3D representation to a library of stored representations.

THOR does not rely on training machine learning models with images of cluttered rooms. It just needs images of each of the different objects by themselves. THOR does not require the robot to have specialized and expensive sensors or processors, and it also works well with commodity cameras.

This means that THOR is very easy to build, and is, more importantly, readily useful for completely new spaces with diverse backgrounds, lighting conditions, object arrangements and degree of clutter. It also works better than the existing 3D shape-based recognition methods because its 3D representation of the objects is more detailed, which helps identify the objects in real time.

How could THOR be used?

THOR could be used with any indoor service robot, regardless of whether the robot operates in someone's home, an office, a store, a warehouse facility or a manufacturing plant. In fact, our experimental evaluation shows that THOR is equally effective for warehouse, lounge and family room-type spaces.

While THOR performs significantly better than the other existing methods for all kinds of objects in these cluttered spaces, it does the best at identifying kitchen-style objects, such as a mug or a pitcher, that typically have distinctive but regular shapes and moderate size variations.

What's next?

There are several additional problems that need to be addressed, and we

are working on some of them. For example, right now, THOR considers only the shape of the objects, but future versions could also pay attention to other aspects of appearance, such as color, texture or text labels. It is also worth looking into how THOR could be used to deal with squishy or damaged objects, which have shapes that are different from their expected configurations.

Also, some spaces may be so cluttered that certain objects might not be visible at all. In these scenarios, a robot needs to be able to decide to move around to "see" the objects better, or if allowed, move around some of the objects to get better views of the obstructed objects.

Last but not least, the robot needs to be able to deal with objects it hasn't seen before. In these scenarios, the [robot](#) should be able to place these objects into a "miscellaneous" or "unknown" object category, and then seek help from a human to correctly identify these objects.

More information: Ekta U. Samani et al, Persistent Homology Meets Object Unity: Object Recognition in Clutter, *IEEE Transactions on Robotics* (2023). [DOI: 10.1109/TRO.2023.3343994](https://doi.org/10.1109/TRO.2023.3343994)

Provided by University of Washington

Citation: Q&A: Researcher discusses how newly developed method can help robots identify objects in cluttered spaces (2024, February 7) retrieved 1 March 2024 from <https://techxplore.com/news/2024-02-qa-discusses-newly-method-robots.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.