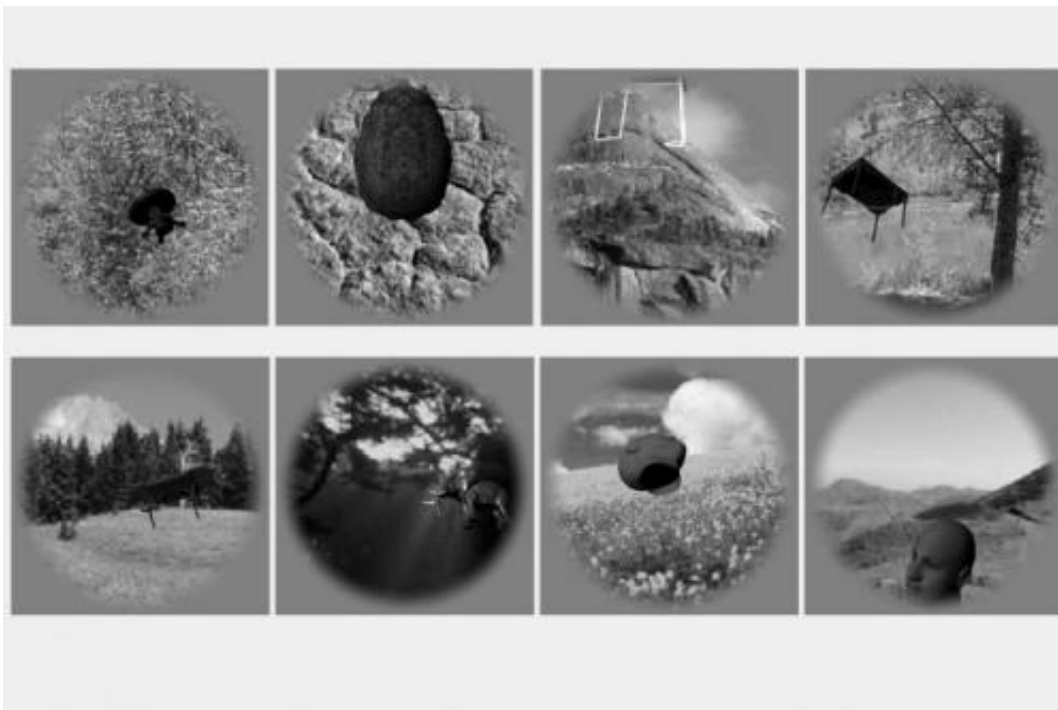


Newest computer neural networks can identify visual objects as well as the primate brain

December 18 2014, by Anne Trafton



A team of MIT neuroscientists has found that some computer programs can identify the objects in these images just as well as the primate brain. Credit: Image courtesy of the researchers

For decades, neuroscientists have been trying to design computer networks that can mimic visual skills such as recognizing objects, which the human brain does very accurately and quickly.

Until now, no computer model has been able to match the [primate brain](#) at visual [object recognition](#) during a brief glance. However, a new study from MIT neuroscientists has found that one of the latest generation of these so-called "[deep neural networks](#)" matches the primate brain.

Because these networks are based on neuroscientists' current understanding of how the brain performs object recognition, the success of the latest networks suggest that neuroscientists have a fairly accurate grasp of how object recognition works, says James DiCarlo, a professor of neuroscience and head of MIT's Department of Brain and Cognitive Sciences and the senior author of a paper describing the study in the Dec. 11 issue of the journal *PLoS Computational Biology*.

"The fact that the models predict the neural responses and the distances of objects in neural population space shows that these models encapsulate our current best understanding as to what is going on in this previously mysterious portion of the brain," says DiCarlo, who is also a member of MIT's McGovern Institute for Brain Research.

This improved understanding of how the primate brain works could lead to better artificial intelligence and, someday, new ways to repair visual dysfunction, adds Charles Cadieu, a postdoc at the McGovern Institute and the paper's lead author.

Other authors are graduate students Ha Hong and Diego Ardila, research scientist Daniel Yamins, former MIT graduate student Nicolas Pinto, former MIT undergraduate Ethan Solomon, and research affiliate Najib Majaj.

Inspired by the brain

Scientists began building neural networks in the 1970s in hopes of mimicking the brain's ability to process visual information, recognize

speech, and understand language.

For vision-based neural networks, scientists were inspired by the hierarchical representation of visual information in the brain. As visual input flows from the retina into [primary visual cortex](#) and then inferotemporal (IT) cortex, it is processed at each level and becomes more specific until objects can be identified.

To mimic this, neural network designers create several layers of computation in their models. Each level performs a mathematical operation, such as a linear dot product. At each level, the representations of the visual object become more and more complex, and unneeded information, such as an object's location or movement, is cast aside.

"Each individual element is typically a very simple mathematical expression," Cadieu says. "But when you combine thousands and millions of these things together, you get very complicated transformations from the raw signals into representations that are very good for object recognition."

For this study, the researchers first measured the brain's object recognition ability. Led by Hong and Majaj, they implanted arrays of electrodes in the IT cortex as well as in area V4, a part of the visual system that feeds into the IT cortex. This allowed them to see the neural representation—the population of neurons that respond—for every object that the animals looked at.

The researchers could then compare this with representations created by the deep neural networks, which consist of a matrix of numbers produced by each computational element in the system. Each image produces a different array of numbers. The accuracy of the model is determined by whether it groups similar objects into similar clusters within the representation.

"Through each of these computational transformations, through each of these layers of networks, certain objects or images get closer together, while others get further apart," Cadieu says.

The best network was one that was developed by researchers at New York University, which classified objects as well as the macaque brain.

More processing power

Two major factors account for the recent success of this type of [neural network](#), Cadieu says. One is a significant leap in the availability of computational processing power. Researchers have been taking advantage of graphical processing units (GPUs), which are small chips designed for high performance in processing the huge amount of visual content needed for video games. "That is allowing people to push the envelope in terms of computation by buying these relatively inexpensive graphics cards," Cadieu says.

The second factor is that researchers now have access to large datasets to feed the algorithms to "train" them. These datasets contain millions of images, and each one is annotated by humans with different levels of identification. For example, a photo of a dog would be labeled as animal, canine, domesticated dog, and the breed of dog.

At first, neural networks are not good at identifying these images, but as they see more and more images, and find out when they were wrong, they refine their calculations until they become much more accurate at identifying objects.

Cadieu says that researchers don't know much about what exactly allows these networks to distinguish different objects.

"That's a pro and a con," he says. "It's very good in that we don't have to

really know what the things are that distinguish those objects. But the big con is that it's very hard to inspect those networks, to look inside and see what they really did. Now that people can see that these things are working well, they'll work more to understand what's happening inside of them."

DiCarlo's lab now plans to try to generate models that can mimic other aspects of visual processing, including tracking motion and recognizing three-dimensional forms. They also hope to create models that include the feedback projections seen in the human visual system. Current networks only model the "feedforward" projections from the retina to the IT cortex, but there are 10 times as many connections that go from IT cortex back to the rest of the system.

Provided by Massachusetts Institute of Technology

Citation: Newest computer neural networks can identify visual objects as well as the primate brain (2014, December 18) retrieved 27 April 2024 from <https://phys.org/news/2014-12-neural-networks-visual-primate-brain.html>

<p>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.</p>
--