

First impressions: Computer model behaves like humans on visual categorization task

April 2 2007



Both humans and a computer model developed in Poggio's lab correctly categorize these images when they are presented for just 50 milliseconds followed by a mask that shuts down cognitive feedback in the human subjects. These results support the view that rapid or immediate object recognition occurs in one feed-forward sweep through the ventral stream of the visual cortex. Credit: Images courtesy Thomas Serre, McGovern Institute for Brain Research at MIT

Computers can usually out-compute the human brain, but there are some tasks, such as visual object recognition, that the brain performs easily yet are very challenging for computers. The brain has a much more sophisticated and swift visual processing system than even the most advanced artificial vision system, giving us an uncanny ability to extract

salient information after just a glimpse that is presumably too fleeting for conscious thought. To explore this phenomenon, neuroscientists have long used rapid categorization tasks, in which subjects indicate whether an object from a specific class (such as an animal) is present or not in the image.

Now, in a new MIT study, a computer model designed to mimic the way the brain itself processes visual information performs as well as humans do on rapid categorization tasks. The model even tends to make similar errors as humans, possibly because it so closely follows the organization of the brain's visual system.

"We created a model that takes into account a host of quantitative anatomical and physiological data about visual cortex and tries to simulate what happens in the first 100 milliseconds or so after we see an object," explained senior author Tomaso Poggio of the McGovern Institute for Brain Research at MIT. "This is the first time a model has been able to reproduce human behavior on that kind of task." The study, issued on line in advance of the April 10, 2007 *Proceedings of the National Academy of Sciences*, stems from a collaboration between computational neuroscientists in Poggio's lab and Aude Oliva, a cognitive neuroscientist in the MIT Department of Brain and Cognitive Sciences.

This new study supports a long-held hypothesis that rapid categorization happens without any feedback from cognitive or other areas of the brain. The results also indicate that the model can help neuroscientists make predictions and drive new experiments to explore brain mechanisms involved in human visual perception, cognition, and behavior. Deciphering the relative contribution of feed-forward and feedback processing may eventually help explain neuropsychological disorders such as autism and schizophrenia. The model also bridges the gap between the world of artificial intelligence (AI) and neuroscience

because it may lead to better artificial vision systems and augmented sensory prostheses.

Rapid Categorization

During normal everyday vision, the eye moves around a scene, giving the brain time to focus attention on relevant cues, such as a snake curled in the path. Evolutionarily speaking, however, survival often depends on extracting vital information in one glance, so that we jump out of danger's way before we even realize what we've seen.

Cognitive neuroscientists have studied this phenomenon using a rapid categorization task during which subjects are asked to say whether a specific object (such as an animal) is present or not. In this task, subjects see an image flashed on a screen that is quickly replaced with an erasing mask (pink noise), which is presumed to shut down cognitive feedback. After just a 50 milliseconds glimpse of an image, less than the time it takes to flash two video frames, people can still accurately report an object's category, even though they are barely aware of what they have seen.

In parallel, computational neuroscientists have traced the flow of information from the retina through increasingly complex visual areas (V1, V2, V4) to the highest purely visual region, the inferotemporal cortex (IT), and on to higher areas such as prefrontal cortex (PFC) where object categorization is represented. The Poggio lab replicated the hypothetical computations the brain performs as information speeds forward through the visual pathway. They recently demonstrated that this biologically inspired model, which matches a number of different physiological data, can also learn to recognize objects from real-world examples and identify relevant objects in complex scenes. (See <http://www.physorg.com/news90006040.html>.) That and other studies from the lab demonstrated that the information processing that occurs

during one feed-forward pass through the visual cortex is sufficient for robust object recognition.

The model is thus an appropriate vehicle for testing the behavioral study's no-feedback-necessary theory, while the animal/no animal behavioral test makes a good reality check for the model.

Glimpsing an Animal – or Not

To proceed, Serre "trained" the model on only a few hundred animal and non-animal images, a paltry number compared to human visual experience. "This is a very hard task for any artificial vision system," Serre explained. "Animals are extremely varied in shape and size. Snakes, butterflies, and elephants have little in common, and the animals in the image may be lying, standing, flying, or leaping."

The team organized images in different subcategories from full views of an animal head to far views, using single as well as groups of animals. As preliminary model simulations predicted, the task became harder as the relative size of the animal decreased and the amount of clutter (the background scene) increased.

Importantly, the results showed no significant difference between humans and the model. Both had a similar pattern of performance, with well above 90% accuracy for the close views dropping to 74% for distant views. The 16% drop in performance for distant views represents a limitation of the one feed-forward sweep in dealing with clutter, Serre suggested. With more time for cognitive feedback, people would outperform the model because they could focus attention on the target and ignore the clutter.

"We have not solved vision yet," Poggio cautioned, "but this model of immediate recognition may provide the skeleton of a theory of vision.

The huge task in front of us is to incorporate into the model the effects of attention and top-down beliefs." The team is now exploring what happens after the first feed-forward sweep, during the next 200-300 milliseconds of object recognition.

The Poggio lab plans to include feedback loops in the model by modeling the widespread anatomical backprojections in cortex, while Oliva is designing behavioral studies that can test if the enhanced model matches the performance of humans who have more time to examine a scene.

For cognitive neuroscientists, these results add to the convergence of evidence about the feed-forward hypothesis for rapid categorization. "There could be other mechanisms involved, but this a big step forward in understanding how humans see," said Oliva. "For me, it's putting light in the black box and gives direction to design new experiments, for instance to explore perception in clutter."

Source: McGovern Institute for Brain Research

Citation: First impressions: Computer model behaves like humans on visual categorization task (2007, April 2) retrieved 4 June 2023 from <https://phys.org/news/2007-04-humans-visual-categorization-task.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.