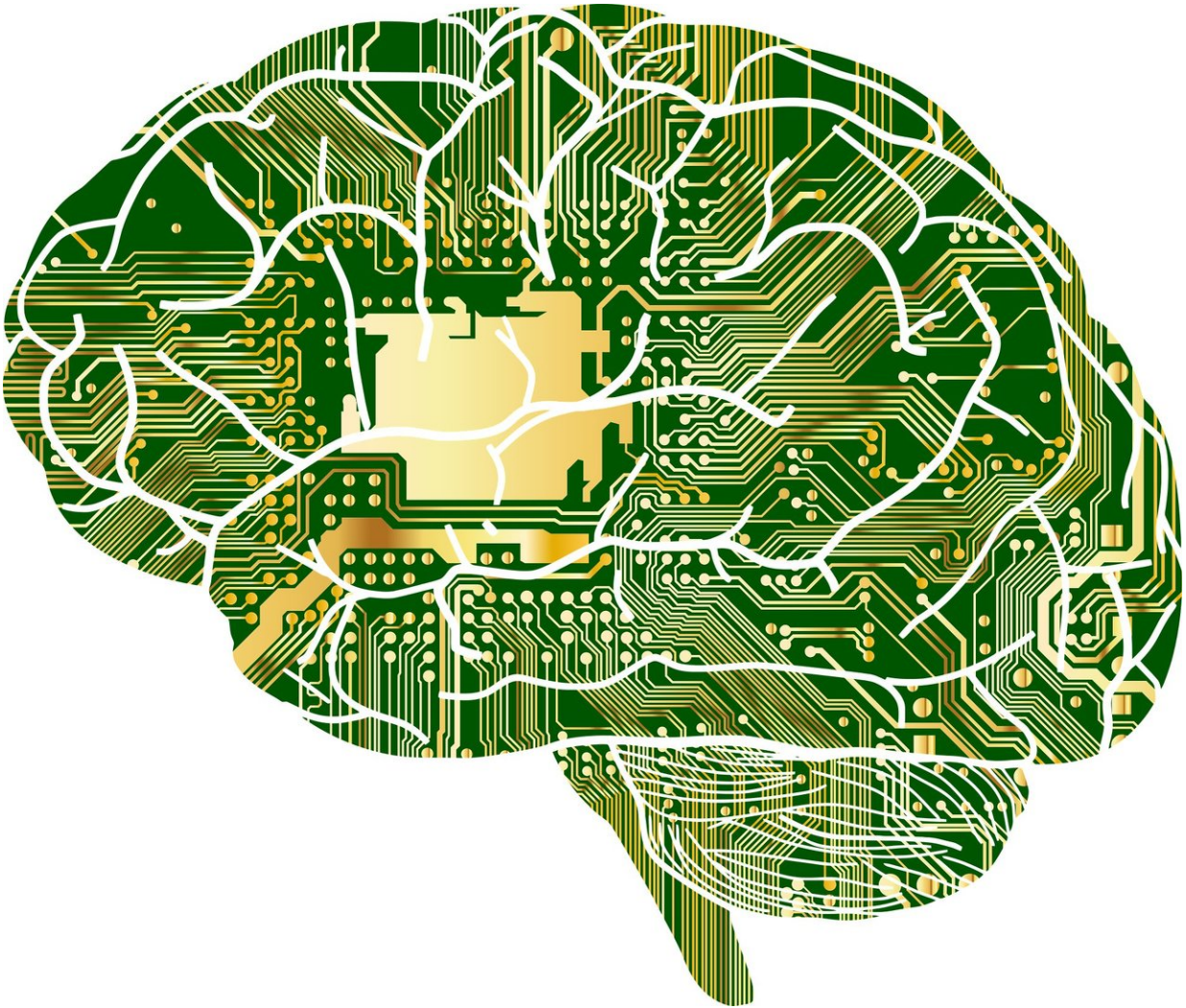


Deep learning comes full circle

May 7 2018, by Nathan Collins



Credit: CC0 Public Domain

For years, the people developing artificial intelligence drew inspiration

from what was known about the human brain, and it has enjoyed a lot of success as a result. Now, AI is starting to return the favor.

Although not explicitly designed to do so, certain artificial intelligence systems seem to mimic our brains' inner workings more closely than previously thought, suggesting that both AI and our minds have converged on the same approach to solving problems. If so, simply watching AI at work could help researchers unlock some of the deepest mysteries of the [brain](#).

"There's a real connection there," said Daniel Yamins, assistant professor of psychology. Now, Yamins, who is also a faculty scholar of the Stanford Neurosciences Institute and a member of Stanford Bio-X, and his lab are building on that connection to produce better theories of the brain – how it perceives the world, how it shifts efficiently from one task to the next and perhaps, one day, how it thinks.

A vision problem for AI

Artificial intelligence has been borrowing from the brain since its early days, when computer scientists and psychologists developed algorithms called neural networks that loosely mimicked the brain. Those algorithms were frequently criticized for being biologically implausible – the "neurons" in neural networks were, after all, gross simplifications of the real neurons that make up the brain. But computer scientists didn't care about biological plausibility. They just wanted systems that worked, so they extended neural network models in whatever way made the algorithm best able to carry out certain tasks, culminating in what is now called [deep learning](#).

Then came a surprise. In 2012, AI researchers showed that a deep learning neural network could learn to identify objects in pictures as well as a human being, which got neuroscientists wondering: How did deep

learning do it?

The same way the brain does, as it turns out. In 2014, Yamins and colleagues showed that a deep learning system that had learned to identify objects in pictures – nearly as well as humans could – did so in a way that closely mimicked the way the brain processes vision. In fact, the computations the deep learning system performed matched activity in the brain's vision-processing circuits substantially better than any other model of those circuits.

Around the same time, other teams made similar observations about parts of the brain's vision– and movement-processing circuits, suggesting that given the same kind of problem, deep learning and the brain had evolved similar ways of coming up with a solution. More recently, Yamins and colleagues have demonstrated similar observations in the brain's auditory system.

On one hand, that's not a big surprise. Although the technical details differ, deep learning's conceptual organization is borrowed directly from what neuroscientists already knew about the organization of neurons in the brain.

But the success of Yamins and colleagues' approach and others like it depends equally as much on another, more subtle choice. Rather than try to get the deep learning system to directly match what the brain does at the level of individual neurons, as many researchers had done, Yamins and colleagues simply gave their deep learning system the same problem: Identify objects in pictures. Only after it had solved that problem did the researchers compare how deep learning and the brain arrived at their solutions – and only then did it become clear that their methods were essentially the same.

"The correspondence between the models and the visual system is not

entirely a coincidence, because one directly inspired the other," said Daniel Bear, a postdoctoral researcher in Yamins' group, "but it's still remarkable that it's as good a correspondence as it is."

One likely reason for that, Bear said, is natural selection and evolution. "Basically, object recognition was a very evolutionarily important task" for animals to solve – and solve well, if they wanted to tell the difference between something they could eat and something that could eat them. Perhaps trying to do that as well as humans and other animals do – except with a computer – led researchers to find essentially the same solution.

Seek what the brain seeks

Whatever the underlying reason, insights gleaned from the 2014 study led to what Yamins calls goal-directed models of the brain: Rather than try to model neural activity in the brain directly, instead train artificial intelligence to solve problems the brain needs to solve, then use the resulting AI system as a model of the brain. Since 2014, Yamins and collaborators have been refining the original goal-directed model of the brain's vision circuits and extending the work in new directions, including understanding the neural circuits that process inputs from rodents' whiskers.

In perhaps the most ambitious project, Yamins and postdoctoral fellow Nick Haber are investigating how infants learn about the world around them through play. Their infants – actually relatively simple computer simulations – are motivated only by curiosity. They explore their worlds by moving around and interacting with objects, learning as they go to predict what happens when they hit balls or simply turn their heads. At the same time, the model learns to predict what parts of the world it doesn't understand, then tries to figure those out.

While the computer simulation begins life – so to speak – knowing essentially nothing about the world, it eventually figures out how to categorize different objects and even how to smash two or three of them together. Although direct comparisons with babies' neural activity might be premature, the model could help researchers better understand how infants use play to learn about their environments, Haber said.

On the other end of the spectrum, models inspired by artificial intelligence could help solve a puzzle about the physical layout of the brain, said Eshed Margalit, a graduate student in neurosciences. As the vision circuits in infants' brains develop, they form specific patches – physical clusters of neurons – that respond to different kinds of objects. For example, humans and other primates all form a face patch that is active almost exclusively when they look at faces.

Exactly why the brain forms those patches, Margalit said, isn't clear. The brain doesn't need a face patch to recognize faces, for example. But by building on AI models like Yamins' that already solve object recognition tasks, "we can now try to [model](#) that spatial structure and ask questions about why the brain is laid out this way and what advantages it might give an organism," Margalit said.

Closing the loop

There are other issues to tackle as well, notably how artificial intelligence systems learn. Right now, AI needs much more training – and much more explicit training – than humans do in order to perform as well on tasks like object recognition, although how humans succeed with so little data remains unclear.

A second issue is how to go beyond models of vision and other sensory systems. "Once you have a sensory impression of the world, you want to make decisions based on it," Yamins said. "We're trying to make models

of decision making, learning to make decisions and how you interface between sensory systems, decision making and memory." Yamins is starting to address those ideas with Kevin Feiglis, a graduate student in physics, who is building AI models that can learn to solve many different kinds of problems and switch between tasks as needed, something very few AI systems are able to do.

In the long run, Yamins and the other members of his group said all of those advances could feed into more capable [artificial intelligence systems](#), just as earlier neuroscience research helped foster the development of deep learning. "I think people in artificial intelligence are realizing there are certain very good next goals for cognitively inspired [artificial intelligence](#)," Haber said, including systems like his that learn by actively exploring their worlds. "People are playing with these ideas."

Provided by Stanford University

Citation: Deep learning comes full circle (2018, May 7) retrieved 3 May 2024 from <https://phys.org/news/2018-05-deep-full-circle.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>
--