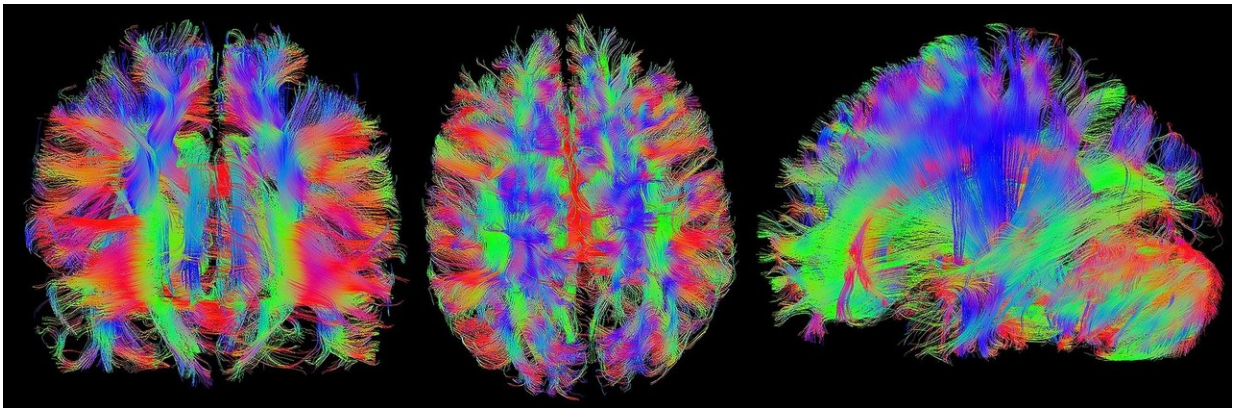# Study shows that artificial neural networks can be used to drive brain activity

May 2 2019

Credit: CC0 Public Domain

MIT neuroscientists have performed the most rigorous testing yet of computational models that mimic the brain's visual cortex.

Using their current best model of the brain's visual neural network, the researchers designed a new way to precisely control individual neurons and populations of neurons in the middle of that network. In an animal study, the team then showed that the information gained from the computational model enabled them to create images that strongly activated specific brain neurons of their choosing.

The findings suggest that the current versions of these models are similar

enough to the brain that they could be used to control brain states in animals. The study also helps to establish the usefulness of these vision models, which have generated vigorous debate over whether they accurately mimic how the visual cortex works, says James DiCarlo, the head of MIT's Department of Brain and Cognitive Sciences, an investigator in the McGovern Institute for Brain Research and the Center for Brains, Minds, and Machines, and the senior author of the study.

"People have questioned whether these models provide understanding of the visual system," he says. "Rather than debate that in an academic sense, we showed that these models are already powerful enough to enable an important new application. Whether you understand how the model works or not, it's already useful in that sense."

MIT postdocs Pouya Bashivan and Kohitij Kar are the lead authors of the paper, which appears in the May 2 online edition of *Science*.

Neural control

Over the past several years, DiCarlo and others have developed models of the visual system based on artificial neural networks. Each network starts out with an arbitrary architecture consisting of model neurons, or nodes, that can be connected to each other with different strengths, also called weights.

The researchers then train the models on a library of more than 1 million images. As the researchers show the model each image, along with a label for the most prominent object in the image, such as an airplane or a chair, the model learns to recognize objects by changing the strengths of its connections.

It's difficult to determine exactly how the model achieves this kind of recognition, but DiCarlo and his colleagues have previously shown that

the "neurons" within these models produce activity patterns very similar to those seen in the animal visual cortex in response to the same images.

In the new study, the researchers wanted to test whether their models could perform some tasks that previously have not been demonstrated. In particular, they wanted to see if the models could be used to control neural activity in the visual cortex of animals.

"So far, what has been done with these models is predicting what the neural responses would be to other stimuli that they have not seen before," Bashivan says. "The main difference here is that we are going one step further and using the models to drive the neurons into desired states."

To achieve this, the researchers first created a one-to-one map of neurons in the brain's visual area V4 to nodes in the computational model. They did this by showing images to animals and to the models, and comparing their responses to the same images. There are millions of neurons in area V4, but for this study, the researchers created maps for subpopulations of five to 40 neurons at a time.

"Once each neuron has an assignment, the model allows you to make predictions about that neuron," DiCarlo says.

The researchers then set out to see if they could use those predictions to control the activity of individual neurons in the visual cortex. The first type of control, which they called "stretching," involves showing an image that will drive the activity of a specific neuron far beyond the activity usually elicited by "natural" images similar to those used to train the neural networks.

The researchers found that when they showed animals these "synthetic" images, which are created by the models and do not resemble natural

objects, the target neurons did respond as expected. On average, the neurons showed about 40 percent more activity in response to these images than when they were shown natural images like those used to train the model. This kind of control has never been reported before.

In a similar set of experiments, the researchers attempted to generate images that would drive one neuron maximally while also keeping the activity in nearby neurons very low, a more difficult task. For most of the neurons they tested, the researchers were able to enhance the activity of the target neuron with little increase in the surrounding neurons.

"A common trend in neuroscience is that experimental data collection and computational modeling are executed somewhat independently, resulting in very little model validation, and thus no measurable progress. Our efforts bring back to life this 'closed loop' approach, engaging model predictions and neural measurements that are critical to the success of building and testing models that will most resemble the brain," Kar says.

Measuring accuracy

The researchers also showed that they could use the model to predict how neurons of area V4 would respond to synthetic images. Most previous tests of these models have used the same type of naturalistic images that were used to train the model. The MIT team found that the models were about 54 percent accurate at predicting how the brain would respond to the synthetic images, compared to nearly 90 percent accuracy when the natural images are used.

"In a sense, we're quantifying how accurate these models are at making predictions outside the domain where they were trained," Bashivan says. "Ideally the model should be able to predict accurately no matter what the input is."

The researchers now hope to improve the models' accuracy by allowing them to incorporate the new information they learn from seeing the synthetic images, which was not done in this study.

This kind of control could be useful for neuroscientists who want to study how different neurons interact with each other, and how they might be connected, the researchers say. Farther in the future, this approach could potentially be useful for treating mood disorders such as depression. The researchers are now working on extending their model to the inferotemporal cortex, which feeds into the amygdala, which is involved in processing emotions.

"If we had a good model of the neurons that are engaged in experiencing emotions or causing various kinds of disorders, then we could use that model to drive the neurons in a way that would help to ameliorate those disorders," Bashivan says.

More information: P. Bashivan el al., "Neural population control via deep image synthesis," *Science* (2019). science.sciencemag.org/cgi/doi … 1126/science.aav9436

Provided by Massachusetts Institute of Technology