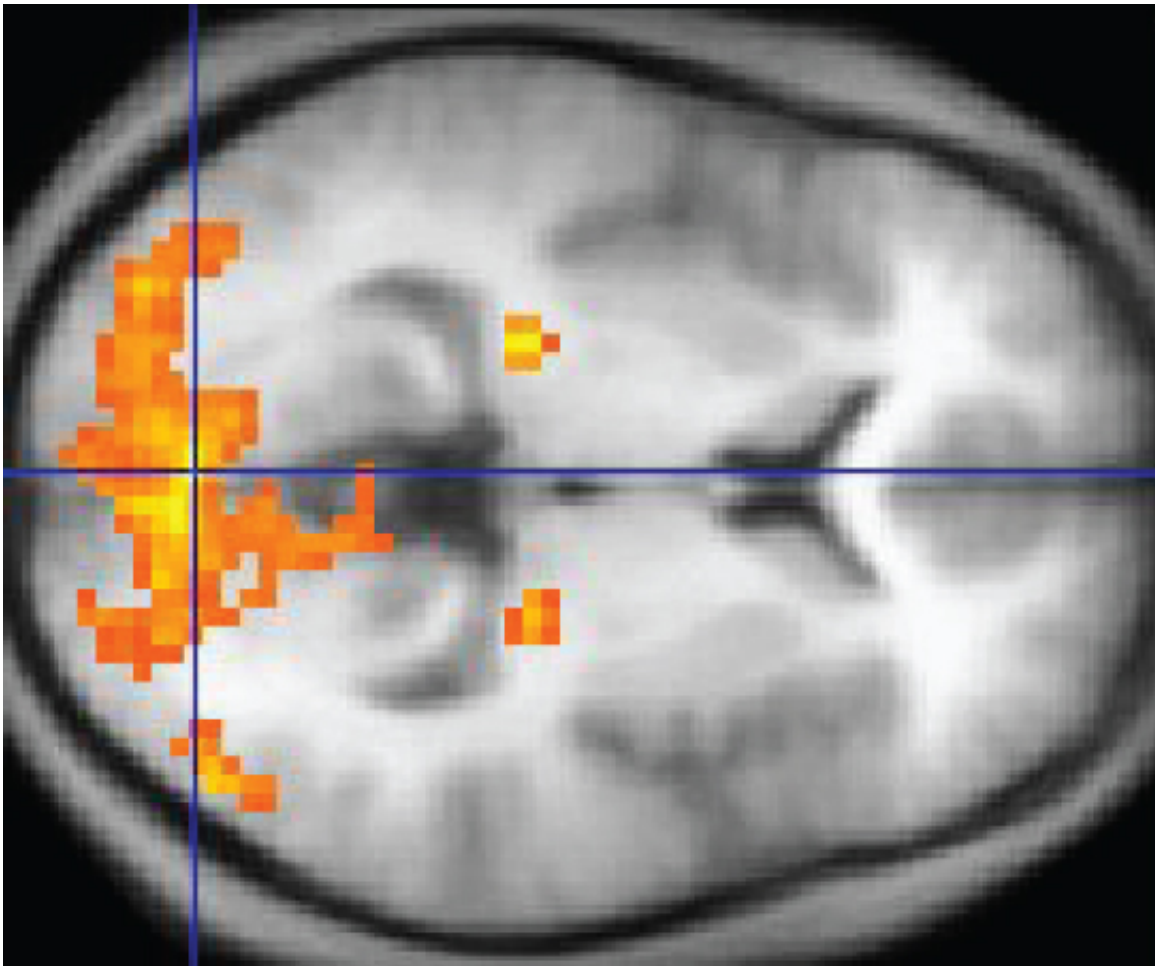# Dataset bridges human vision and machine learning

May 6 2019

An fMRI image with yellow areas showing increased activity. Credit: Wikipedia/ CC BY 3.0

Neuroscientists and computer vision scientists say a new dataset of

unprecedented size—comprising brain scans of four volunteers who each viewed 5,000 images—will help researchers better understand how the brain processes images.

Researchers at Carnegie Mellon University and Fordham University, reporting today in the journal *Scientific Data*, said acquiring functional magnetic resonance imaging (fMRI) scans at this scale presented unique challenges.

Each volunteer participated in 20 or more hours of MRI scanning, challenging both their perseverance and the experimenters' ability to coordinate across scanning sessions. The extreme design decision to run the same individuals over so many sessions was necessary for disentangling the neural responses associated with individual images.

The resulting dataset, dubbed BOLD5000, allows cognitive neuroscientists to better leverage the deep learning models that have dramatically improved artificial vision systems. Originally inspired by the architecture of the human visual system, deep learning may be further improved by pursuing new insights into how human vision works and by having studies of human vision better reflect modern computer vision methods. To that end, BOLD5000 measured neural activity arising from viewing images taken from two popular computer vision datasets: ImageNet and COCO.

"The intertwining of brain science and computer science means that scientific discoveries can flow in both directions," said co-author Michael J. Tarr, the Kavčic-Moura Professor of Cognitive and Brain Science and head of CMU's Department of Psychology. "Future studies of vision that employ the BOLD5000 dataset should help neuroscientists better understand the organization of knowledge in the human brain. As we learn more about the neural basis of visual recognition, we will also be better positioned to contribute to advances in artificial vision."

Lead author Nadine Chang, a Ph.D. student in CMU's Robotics Institute who specializes in computer vision, suggested that computer vision scientists are looking to neuroscience to help innovate in the rapidly advancing area of artificial vision—reinforcing the two-way nature of this research.

"Computer-vision scientists and visual neuroscientists essentially have the same end goal: to understand how to process and interpret visual information," Chang said.

Improving computer vision was an important part of the BOLD5000 project from its onset. Senior author Elissa Aminoff, then a post-doctoral fellow in CMU's Psychology Department and now an assistant professor of psychology at Fordham, initiated this research direction with co-author Abhinav Gupta, an associate professor in the Robotics Institute.

Among the challenges faced in connecting biological and computer vision is that the majority of human neuroimaging studies include very few stimulus images—often 100 or less—which typically are simplified to depict only single objects against a neutral background. In contrast, BOLD5000 includes more than 5,000 real-world, complex images of scenes, single objects and interacting objects.

The group views BOLD5000 as only the first step toward leveraging modern computer vision models to study biological vision.

"Frankly, the BOLD5000 dataset is still way too small," Tarr said, suggesting that a reasonable fMRI dataset would require at least 50,000 stimulus images and many more volunteers to make headway in light of the fact that the class of deep neural nets used to analyze visual imagery are trained on millions of images. To this end, the research team hopes their ability to generate a dataset of 5,000 brain scans will pave the way

for larger collaborative efforts between human vision and computer vision scientists.

So far, the field's response has been positive. The publicly available BOLD5000 dataset has already been downloaded more than 2,500 times.

More information: *Scientific Data*, DOI: 10.1038/s41597-019-0052-3

Provided by Carnegie Mellon University