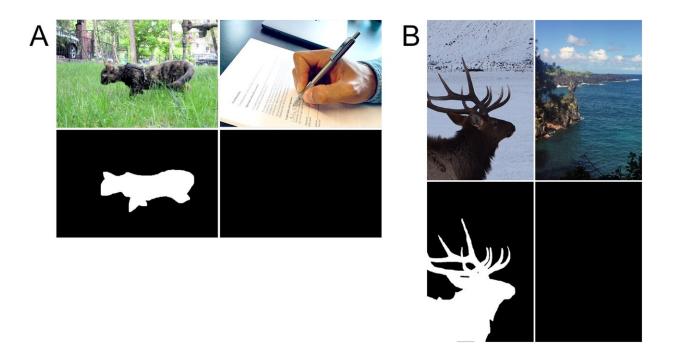# Vision scientists disprove 60-year-old perception theory

October 24 2019

A: Sample images representative of those from the Potter dataset with target present (top left), no target present (top right), and corresponding ground truth masks (bottom row). In the experiment, participants were asked if they saw a particular target (here, "a walking cat"). Only those trials where the question was asked after the stimulus presentation were used. The bottom panels show the hand-drawn ground truth masks; no-target images have no ground-truth. B: Sample images, representative of those in the Thorpe dataset with target present (top left), no target present (top right), and corresponding ground truth masks (bottom row). In the experiment participants were asked if there is an animal in the image. The bottom panels show the hand-drawn ground truth masks; no-target images have no ground-truth. Credit: *PLOS ONE* (2019). DOI: 10.1371/journal.pone.0224306

Vision researchers at York University have disproved a long-standing theory of how the human vision system processes images, using computational models and human experiments.

A team led by John Tsotsos, professor in the Department of Electrical Engineering and Computer Science at the Lassonde School of Engineering, found that the human brain does not select interesting portions of an image to process preferentially, as the highly influential 1958 theory of Donald Broadbent proposed.

For psychologist Broadbent, the interesting image portions are those that have relevance to why you are looking at a scene in the first place, or are novel items that immediately grab our attention. Broadbent's Theory of Early Selection, which has a modern counterpart in the Saliency Map Theory of Christof Koch and Shimon Ullman published in 1985, claims that these interesting regions are processed by the brain one at a time, in order of their salience, which is a numerical score of how interesting a region is. There are now hundreds of saliency algorithms, rooted in the work of Koch and Ullman, to accomplish such a ranking.

Tsotsos' team found however, that salience is not needed at all for the simple task of quickly deciding what an image depicts. Moreover, none of the current algorithms within artificial intelligence (AI) for this task come close to matching human performance, which is remarkably good. On the other hand, salience computation does play a primary role in determining where humans move their eyes, and it is eye movement that selects portions of a scene to process next.

"Our study looks at this for vision and tests the leading algorithms that compute the saliency measure and asks the question 'are those algorithms performing at the same level as humans do on these images'? For

example, if the task is to determine if there is a cat in a scene, does the saliency [algorithm](#) pick out the cat correctly? The study showed that these algorithms are far from doing as well as humans," said Tsotsos.

To further test existing algorithms, the team conducted additional experiments with 17 subjects from 25 to 34 years old. In one of the replicated experiments, participants were shown 2000 colour images. The subjects were not familiar with the images and viewed each image with and without animals, only once. The images were then manipulated in such a way so that only the most central parts of the retina that have the highest resolution would see what was in the image and see nothing in the periphery. Participants were asked to look at the centre of each photograph for 20 seconds before it disappeared. The participants were able to correctly identify if an animal was present in the picture or not.

Tsotsos says this finding has important ramifications for our understanding of human vision and human visual processing especially for diagnosing vision pathologies, such as aspects of autism.

"When you want to diagnose issues in [vision](#), you're basing on it how the healthy visual processing system should work. What we've done with this study is added a piece of the puzzle to how the 'healthy' system works which then would change how you compare an anomaly in order to be able to diagnose it."

Tsotsos adds that this piece of the puzzle could also be useful in building new models and improving current ones for autonomous driving or security applications.

This study, Rapid Visual Categorization is not Guided by Early Salience-Based Selection is published today in *PLOS ONE*.

  More information: John K. Tsotsos et al, Rapid visual categorization

is not guided by early salience-based selection, *PLOS ONE* (2019).

Provided by York University