

New program picks out targets in a crowd quickly and efficiently

February 22 2019



Credit: CC0 Public Domain

It can be harder for computers to find Waldo, an elusive character that hides within crowds in a popular children's book series, than it is for humans.

Now, an A*STAR researcher and her colleagues have developed a biologically-inspired program that could enable computers to identify

real-life Waldos and other targets more efficiently.

Computer image analysis is routinely used in medicine, security, and rescue. Speed is often critical in these efforts, says Mengmi Zhang, a computer scientist at A*STAR's Institute for Infocomm Research, who led the study. She cites the use of computers to help find victims of natural disasters, such as earthquakes.

But these efforts are often hampered because computers lack human intuition. A person can quickly spot a dog in a crowded space, for instance, even if they have never seen that particular dog before. A computer, by contrast, needs to be trained using thousands of images of different dogs, and even then, they can falter when looking for a new dog whose image they have not encountered previously.

This weakness could be particularly problematic when scanning for weapons, says Zhang. A computer trained to look for knives and guns, might overlook another sharp object. "If there is one sharp metal stick which has not been seen in the training set, it doesn't mean the passenger should be able to take it on board the airplane," says Zhang.

Current computer searches also tend to be slow because the computer must scan every part of an image in sequence, paying equal attention to each part. Humans, however, rapidly shift their attention between several different locations in an image to find their target. Zhang and her colleagues' wanted to understand how humans do this so efficiently. They presented 45 people with crowded images and asked them to hunt for a target, say, a sheep. They monitored how the subjects' eyes darted around the scene, fixating briefly on different locations in the image. They found that, on average, people could locate the sheep in around 640 milliseconds. This corresponded to switching the location of their gaze, on average, just over two and a half times.

The team then developed a [computer model](#) to implement this more human-like search strategy in the hunt for a dog. Rather than looking for a target that was identical to an image of a dog given beforehand, the [model](#) was trained to look for something that had similar features to the example image. This enabled the model to generalize from a single dog image, to the "general concept of a dog," and quickly pick out other [dogs](#) it had not seen before, explains Zhang.

The researchers tested how effective the new computer visual search model was by measuring the number of times the computer had to fixate on different locations in a scene before finding its target. "What surprises us is that by using our method, computers can search images as fast as humans, even when searching for objects they've never seen before," says Zhang. The [computer](#) was even as good as humans at finding Waldo.

The team is now programming their model with a better understanding of context. For example, humans naturally understand that a cup is more likely to be sitting on a table than floating in the air. Once implemented, this should improve the model's efficiency even further, says Zhang, adding, "Waldo cannot hide anymore."

More information: Mengmi Zhang et al. Finding any Waldo with zero-shot invariant and efficient visual search, *Nature Communications* (2018). [DOI: 10.1038/s41467-018-06217-x](https://doi.org/10.1038/s41467-018-06217-x)

Provided by Agency for Science, Technology and Research (A*STAR), Singapore

Citation: New program picks out targets in a crowd quickly and efficiently (2019, February 22) retrieved 20 September 2024 from

<https://phys.org/news/2019-02-crowd-quickly-efficiently.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.