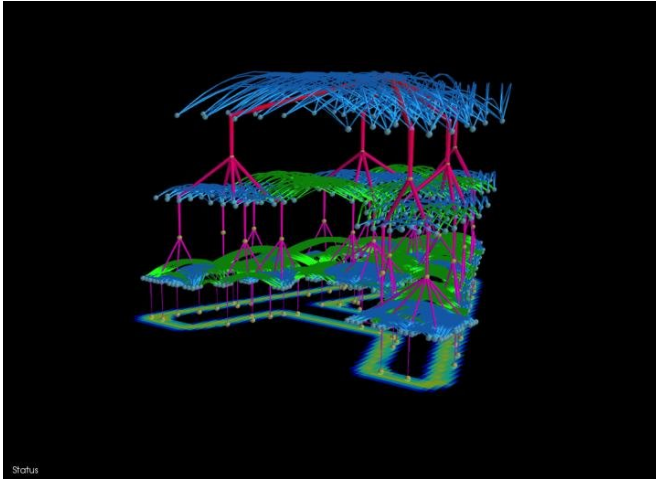


# Vicarious AI team reveals how it defeated CAPTCHA

27 October 2017, by Bob Yirka



A representation of the letter A. This material relates to a paper that appeared in the 27 Oct.2017, issue of *Science*, published by AAAS. The paper, by D. George at Vicarious AI in Union City, CA, and colleagues was titled, "A generative vision model that trains with high data efficiency and breaks text-based CAPTCHAs." Credit: Vicarious AI

A group of researchers at Vicarious AI has revealed for the first time the new and innovative method they used to defeat CAPTCHA. In their paper published in the journal *Science*, the team describes their neural network and how it was used to crack the software that was created to prevent bots from accessing websites.

Back in 2000, groups hosting websites and users alike had grown quite irritated with bots creating havoc on websites. To prevent such access, a system called the "Completely Automated Public Turing test to tell Computers and Humans Apart" (CAPTCHA) was developed. It required users to type in text that had been distorted—a task deemed easy enough for humans, but impossible for bots. Unfortunately, as it turned out, defeating CAPTCHA was possible, as several research

teams showed. They used [neural networks](#) to learn what a CAPTCHA was and then to foil systems that used them. But such methods required the system to crunch thousands or millions of examples to become reasonably proficient at cracking CAPTCHA. But then, four years ago, a team at Vicarious AI announced that they had come up with a modified neural network that could crack CAPTCHA after studying just a few examples. The company did not publish its work, however, because they realized it would allow bot makers to run rampant again. In recent times, CAPTCHA has changed—quite often, images are used instead of text, requiring users to identify something unique in them. Because of that, Vicarious AI has decided it is now safe to reveal its novel CAPTCHA-cracking technology.

The system is called a recursive cortical network—a name that offers insight into how it works. In a traditional neural network, nodes are created to hold new information—a network is built from the nodes and it is used to judge how to deal with new data. This is how it learns. The team at Vicarious AI used a neural network, too, but they added something new—recursion. Recursion is a [software](#) technique whereby data is used to learn something new. As that new process is learned, the results go back into the software, as well. This process is used over and over until a solution is reached. It is a technique that has long been used to solve mazes. By applying recursion to a neural network, the researchers found they were able to reduce the learning curve of their software dramatically—it needed just five training steps, for example, to crack Google's reCAPTCHA 67 percent of the time.

**More information:** D. George et al. A generative vision model that trains with high data efficiency and breaks text-based CAPTCHAs, *Science* (2017). [DOI: 10.1126/science.aag2612](https://doi.org/10.1126/science.aag2612)

**Abstract**

Learning from few examples and generalizing to dramatically different situations are capabilities of human visual intelligence that are yet to be matched by leading machine learning models. By drawing inspiration from systems neuroscience, we introduce a probabilistic generative model for vision in which message-passing based inference handles recognition, segmentation and reasoning in a unified way. The model demonstrates excellent generalization and occlusion-reasoning capabilities, and outperforms deep neural networks on a challenging scene text recognition benchmark while being 300-fold more data efficient. In addition, the model fundamentally breaks the defense of modern text-based CAPTCHAs by generatively segmenting characters without CAPTCHA-specific heuristics. Our model emphasizes aspects like data efficiency and compositionality that may be important in the path toward general artificial intelligence.

© 2017 Tech Xplore

APA citation: Vicarious AI team reveals how it defeated CAPTCHA (2017, October 27) retrieved 27 May 2022 from <https://techxplore.com/news/2017-10-vicarious-ai-team-reveals-defeated.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.*