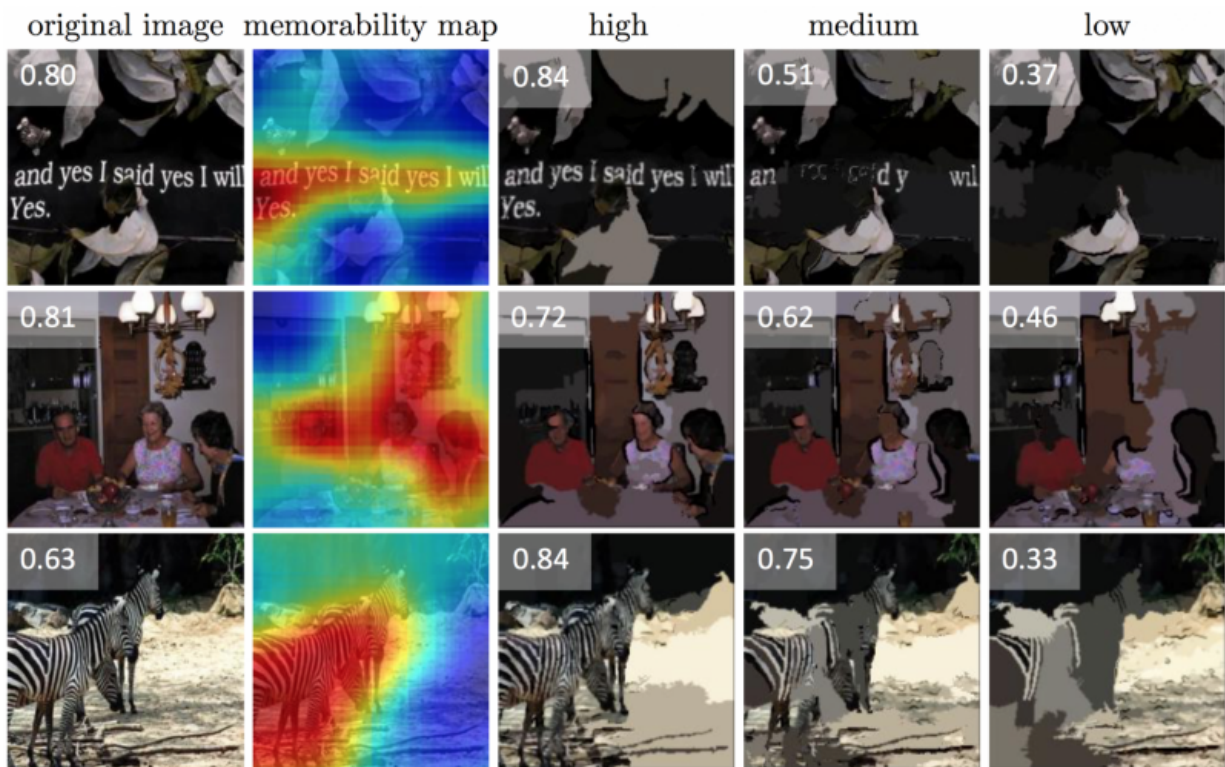# Deep-learning algorithm predicts photos' memorability at "near-human" levels

December 16 2015, by Adam Conner-Simons



For each image, the MemNet algorithm creates a heat map identifying its most memorable and forgettable regions. The image can then be subtly tweaked to increase or decrease its memorability score. Credit: CSAIL

Researchers from MIT's Computer Science and Artificial Intelligence Laboratory (CSAIL) have created an algorithm that can predict how

memorable or forgettable an image is almost as accurately as humans—and they plan to turn it into an app that subtly tweaks photos to make them more memorable.

For each photo, the "MemNet" algorithm—which you can try out online by uploading your own photos—also creates a heat map that identifies exactly which parts of the image are most memorable.

"Understanding memorability can help us make systems to capture the most important information, or, conversely, to store information that humans will most likely forget," says CSAIL graduate student Aditya Khosla, who was lead author on a related paper. "It's like having an instant focus group that tells you how likely it is that someone will remember a visual message."

Team members picture a variety of potential applications, from improving the content of ads and social media posts, to developing more effective teaching resources, to creating your own personal "health-assistant" device to help you remember things.

Part of the project the team has also published the world's largest image-memorability dataset, LaMem. With 60,000 images, each annotated with detailed metadata about qualities such as popularity and emotional impact, LaMem is the team's effort to spur further research on what they say has often been an under-studied topic in computer vision.

The paper was co-written by CSAIL graduate student Akhil Raju, Professor Antonio Torralba, and principal research scientist Aude Oliva, who serves as senior investigator of the work. Khosla will present the paper in Chile this week at the International Conference on Computer Vision.

## How it works

The team previously developed a similar algorithm for facial memorability. What's notable about the new one, besides the fact that it can now perform at near-human levels, is that it uses techniques from "deep-learning," a field of [artificial intelligence](#) that use systems called "[neural networks](#)" to teach computers to sift through massive amounts of data to find patterns all on their own.

Such techniques are what drive Apple's Siri, Google's auto-complete, and Facebook's photo-tagging, and what have spurred these tech giants to spend hundreds of millions of dollars on deep-learning startups.

"While deep-learning has propelled much progress in object recognition and scene understanding, predicting human memory has often been viewed as a higher-level cognitive process that computer scientists will never be able to tackle," Oliva says. "Well, we can, and we did!"

Neural networks work to correlate data without any human guidance on what the underlying causes or correlations might be. They are organized in layers of processing units that each perform random computations on the data in succession. As the network receives more data, it readjusts to produce more accurate predictions.

The team fed its algorithm tens of thousands of images from several different datasets, including LaMem and the scene-oriented SUN and Places (all of which were developed at CSAIL). The images had each received a "memorability score" based on the ability of human subjects to remember them in online experiments.

The team then pitted its algorithm against human subjects by having the model predicting how memorable a group of people would find a new never-before-seen image. It performed 30 percent better than existing algorithms and was within a few percentage points of the average human performance.

For each image, the algorithm produces a heat map showing which parts of the image are most memorable. By emphasizing different regions, they can potentially increase the image's memorability.

"CSAIL researchers have done such manipulations with faces, but I'm impressed that they have been able to extend it to generic images," says Alexei Efros, an associate professor of computer science at the University of California at Berkeley. "While you can somewhat easily change the appearance of a face by, say, making it more 'smiley,' it is significantly harder to generalize about all image types."

## Looking ahead

The research also unexpectedly shed light on the nature of human memory. Khosla says he had wondered whether human subjects would remember everything if they were shown only the most memorable images.

"You might expect that people will acclimate and forget as many things as they did before, but our research suggests otherwise," he says. "This means that we could potentially improve people's memory if we present them with memorable images."

The team next plans to try to update the system to be able to predict the memory of a specific person, as well as to better tailor it for individual "expert industries" such as retail clothing and logo design.

"This sort of research gives us a better understanding of the visual information that people pay attention to," Efros says. "For marketers, movie-makers and other content creators, being able to model your mental state as you look at something is an exciting new direction to explore."

**More information:** memorability.csail.mit.edu/demo.html

memorability.csail.mit.edu/

Research paper: people.csail.mit.edu/khosla/pa … /iccv2015_khosla.pdf

*This story is republished courtesy of MIT News (web.mit.edu/newsoffice/), a popular site that covers news about MIT research, innovation and teaching.*

Provided by Massachusetts Institute of Technology

Citation: Deep-learning algorithm predicts photos' memorability at "near-human" levels (2015, December 16) retrieved 10 April 2024 from https://phys.org/news/2015-12-deep-learning-algorithm-photos-near-human.html