

Carnegie Mellon computer searches web 24/7 to analyze images and teach itself common sense

November 20 2013

A computer program called the Never Ending Image Learner (NEIL) is running 24 hours a day at Carnegie Mellon University, searching the Web for images, doing its best to understand them on its own and, as it builds a growing visual database, gathering common sense on a massive scale.

NEIL leverages recent advances in [computer vision](#) that enable computer programs to identify and label objects in [images](#), to characterize scenes and to recognize attributes, such as colors, lighting and materials, all with a minimum of human supervision. In turn, the data it generates will further enhance the ability of computers to understand the visual world.

But NEIL also makes associations between these things to obtain common sense information that people just seem to know without ever saying—that cars often are found on roads, that buildings tend to be vertical and that ducks look sort of like geese. Based on text references, it might seem that the color associated with sheep is black, but people—and NEIL—nevertheless know that sheep typically are white.

"Images are the best way to learn visual properties," said Abhinav Gupta, assistant research professor in Carnegie Mellon's Robotics Institute.

"Images also include a lot of common sense information about the world. People learn this by themselves and, with NEIL, we hope that computers will do so as well."

A computer cluster has been running the NEIL program since late July and already has analyzed three million images, identifying 1,500 types of objects in half a million images and 1,200 types of scenes in hundreds of thousands of images. It has connected the dots to learn 2,500 associations from thousands of instances.

The public can now view NEIL's findings at the project website, <http://www.neil-kb.com>.

The research team, including Xinlei Chen, a Ph.D. student in CMU's Language Technologies Institute, and Abhinav Shrivastava, a Ph.D. student in robotics, will present its findings on Dec. 4 at the IEEE International Conference on Computer Vision in Sydney, Australia.

One motivation for the NEIL project is to create the world's largest visual structured knowledge base, where objects, scenes, actions, attributes and contextual relationships are labeled and catalogued.

"What we have learned in the last 5-10 years of computer vision research is that the more data you have, the better computer vision becomes," Gupta said.

Some projects, such as ImageNet and Visipedia, have tried to compile this structured data with human assistance. But the scale of the Internet is so vast—Facebook alone holds more than 200 billion images—that the only hope to analyze it all is to teach computers to do it largely by themselves.

Shrivastava said NEIL can sometimes make erroneous assumptions that compound mistakes, so people need to be part of the process. A Google Image search, for instance, might convince NEIL that "pink" is just the name of a singer, rather than a color.

"People don't always know how or what to teach computers," he observed. "But humans are good at telling computers when they are wrong."

People also tell NEIL what categories of objects, scenes, etc., to search and analyze. But sometimes, what NEIL finds can surprise even the researchers. It can be anticipated, for instance, that a search for "apple" might return images of fruit as well as laptop computers. But Gupta and his landlubbing team had no idea that a search for F-18 would identify not only images of a fighter jet, but also of F18-class catamarans.

As its search proceeds, NEIL develops subcategories of objects – tricycles can be for kids, for adults and can be motorized, or cars come in a variety of brands and models. And it begins to notice associations – that zebras tend to be found in savannahs, for instance, and that stock trading floors are typically crowded.

NEIL is computationally intensive, the research team noted. The program runs on two clusters of computers that include 200 processing cores.

Provided by Carnegie Mellon University

Citation: Carnegie Mellon computer searches web 24/7 to analyze images and teach itself common sense (2013, November 20) retrieved 26 April 2024 from <https://phys.org/news/2013-11-carnegie-mellon-web-images-common.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.