New computer program aims to teach itself everything about anything

June 12 2014, by Michelle Ma

This image shows some of the many variations that the new program has learned for three different concepts. Credit: University of Washington

(Phys.org) —In today's digitally driven world, access to information appears limitless. But when you have something specific in mind that you don't know, like the name of that niche kitchen tool you saw at a friend's house, it can be surprisingly hard to sift through the volume of information online and know how to search for it. Or, the opposite problem can occur – we can look up anything on the Internet, but how can we be sure we are finding everything about the topic without spending hours in front of the computer?

Computer scientists from the University of Washington and the Allen Institute for Artificial Intelligence in Seattle have created the first fully automated computer program that teaches everything there is to know about any visual concept. Called Learning Everything about Anything, or LEVAN, the program searches millions of books and images on the Web to learn all possible variations of a concept, then displays the results to users as a comprehensive, browsable list of images, helping them explore and understand topics quickly in great detail.

"It is all about discovering associations between textual and visual data," said Ali Farhadi, a UW assistant professor of computer science and engineering. "The program learns to tightly couple rich sets of phrases with pixels in images. This means that it can recognize instances of specific concepts when it sees them."

The research team will present the project and a related paper this month at the Computer Vision and Pattern Recognition annual conference in Columbus, Ohio.

The program learns which terms are relevant by looking at the content of the images found on the Web and identifying characteristic patterns across them using object recognition algorithms. It's different from online image libraries because it draws upon a rich set of phrases to understand and tag photos by their content and pixel arrangements, not simply by words displayed in captions.

Users can browse the existing library of roughly 175 concepts. Existing concepts range from "airline" to "window," and include "beautiful," "breakfast," "shiny," "cancer," "innovation," "skateboarding," "robot," and the researchers' first-ever input, "horse."

If the concept you're looking for doesn't exist, you can submit any search term and the program will automatically begin generating an exhaustive

list of subcategory images that relate to that concept. For example, a search for "dog" brings up the obvious collection of subcategories: Photos of "Chihuahua dog," "black dog," "swimming dog," "scruffy dog," "greyhound dog." But also "dog nose," "dog bowl," "sad dog," "ugliest dog," "hot dog" and even "down dog," as in the yoga pose.

The technique works by searching the text from millions of books written in English and available on Google Books, scouring for every occurrence of the concept in the entire digital library. Then, an algorithm filters out words that aren't visual. For example, with the concept "horse," the algorithm would keep phrases such as "jumping horse," "eating horse" and "barrel horse," but would exclude non-visual phrases such as "my horse" and "last horse."

Once it has learned which phrases are relevant, the program does an image search on the Web, looking for uniformity in appearance among the photos retrieved. When the program is trained to find relevant images of, say, "jumping horse," it then recognizes all images associated with this phrase.

"Major information resources such as dictionaries and encyclopedias are moving toward the direction of showing users visual information because it is easier to comprehend and much faster to browse through concepts. However, they have limited coverage as they are often manually curated. The new program needs no human supervision, and thus can automatically learn the visual knowledge for any concept," said Santosh Divvala, a research scientist at the Allen Institute for Artificial Intelligence and an affiliate scientist at UW in computer science and engineering.

The research team also includes Carlos Guestrin, a UW professor of computer science and engineering. The researchers launched the program in March with only a handful of concepts and have watched it

grow since then to tag more than 13 million images with 65,000 different phrases.

Right now, the program is limited in how fast it can learn about a concept because of the computational power it takes to process each query, up to 12 hours for some broad concepts. The researchers are working on increasing the processing speed and capabilities.

The team wants the open-source program to be both an educational tool as well as an information bank for researchers in the computer vision community. The team also hopes to offer a smartphone app that can run the program to automatically parse out and categorize photos.

Provided by University of Washington