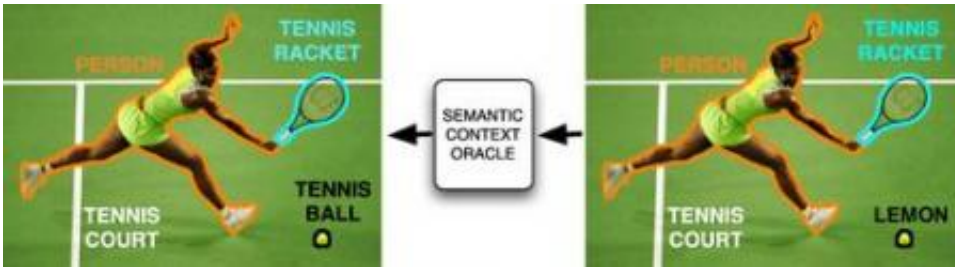# Researchers Give Computers Common Sense

October 17 2007

The computer scientists injected context into an automated image labeling system through a post-processing context check. The approach strives to maximize the contextual agreement among the labeled objects within each picture.

Using a little-known Google Labs widget, computer scientists from UC San Diego and UCLA have brought common sense to an automated image labeling system. The common sense comes as the ability to use context to help identify objects in photographs.

For example, if a conventional automated object identifier has labeled a person, a tennis racket, a tennis court and a lemon in a photo, the new post-processing context check will re-label the lemon as a tennis ball.

"We think our paper is the first to bring external semantic context to the problem of object recognition," said computer science professor Serge Belongie from UC San Diego.

The researchers show that the Google Labs tool called Google Sets can

be used to provide external contextual information to automated object identifiers. The paper will be presented on Thursday 18 October 2007 at ICCV 2007 – the 11th IEEE International Conference on Computer Vision in Rio de Janeiro, Belongie.

Google Sets generates lists of related items or objects from just a few examples. If you type in John, Paul and George, it will return the words Ringo, Beatles and John Lennon. If you type "neon" and "argon" it will give you the rest of the noble gasses.

"In some ways, Google Sets is a proxy for common sense. In our paper, we showed that you can use this common sense to provide contextual information that improves the accuracy of automated image labeling systems," said Belongie.

The image labeling system is a three step process. First, an automated system splits the image up into different regions through the process of image segmentation. In the photo above, image segmentation separates the person, the court, the racket and the yellow sphere.

Next, an automated system provides a ranked list of probable labels for each of these image regions.

Finally, the system adds a dose of context by processing all the different possible combinations of labels within the image and maximizing the contextual agreement among the labeled objects within each picture.

It is during this step that Google Sets can be used as a source of context that helps the system turn a lemon into a tennis ball. In this case, these "semantic context constraints" helped the system disambiguate between visually similar objects.

In another example, the researchers show that an object originally

labeled as a cow is (correctly) re-labeled as a boat when the other objects in the image – sky, tree, building and water – are considered during the post-processing context step. In this case, the semantic context constraints helped to correct an entirely wrong image label. The context information came from co-occurence object information from the training data rather than from Google Sets.

The computer scientists also highlight other advances they bring to automated object identification. First, instead of doing just one image segmentation, the researchers generated a collection of image segmentations and put together a shortlist of stable image segmentations. This increases the accuracy of the segmentation process and provides an implicit shape description for each of the image regions.

Second, the researchers ran their object categorization model on each of the segmentations, rather than on individual pixels. This dramatically reduced the computational demands on the object categorization model.

In addition to Google Sets, the researchers gleaned semantic context information from the co-occurrence of object labels in the training sets.

In the two sets of images that the researchers tested, the categorization results improved considerably with inclusion of context. For one image dataset, the average categorization accuracy increased more than 10 percent using the semantic context provided by Google Sets. In a second dataset, the average categorization accuracy improved by about 2 percent using the semantic context provided by Google Sets. The improvements were higher when the researchers gleaned context information from data on co-occurrence of object labels in the training data set for the object identifier.

Right now, the researchers are exploring ways to extend context beyond the presence of objects in the same image. For example, they want to

make explicit use of absolute and relative geometric relationships between objects in an image – such as "above" or "inside" relationships. This would mean that if a person were sitting on top of an animal, the system would consider the animal to be more likely a horse than a dog.

Source: University of California, San Diego

Citation: Researchers Give Computers Common Sense (2007, October 17) retrieved 20 March 2024 from https://phys.org/news/2007-10-common.html