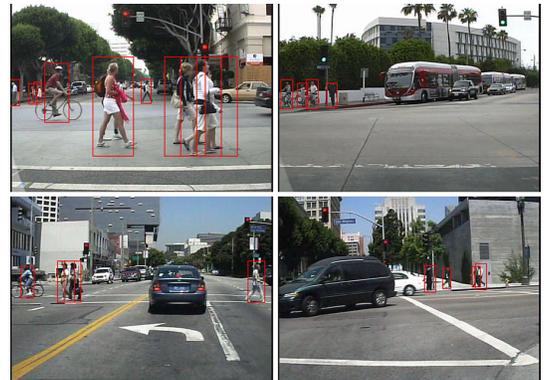


February 08, 2016

New Algorithm Improves Speed and Accuracy of Pedestrian Detection

What if computers could recognize objects as well as the human brain could? Electrical engineers at the University of California, San Diego have taken an important step toward that goal by developing a pedestrian detection system that performs in near real-time (2–4 frames per second) and with higher accuracy (close to half the error) compared to existing systems. The technology, which incorporates deep learning models, could be used in “smart” vehicles, robotics and image and video search systems.

“We’re aiming to build computer vision systems that will help computers better understand the world around them,” said Nuno Vasconcelos, electrical engineering professor at the UC San Diego Jacobs School of Engineering who directed the research. A big goal is real-time vision, he says, especially for pedestrian detection systems in self-driving cars. Vasconcelos is a faculty affiliate of the Center for Visual Computing and the Contextual Robotics Institute, both at UC San Diego.



Pedestrian detection system developed in the Statistical Visual Computing Lab at UC San Diego.

The new pedestrian detection algorithm developed by Vasconcelos and his team combines a traditional computer vision classification architecture, known as cascade detection, with deep learning models.

Pedestrian detection systems typically break down an image into small windows that are processed by a classifier that signals the presence or absence of a pedestrian. This approach is challenging because pedestrians appear in different sizes—depending on distance to the camera—and locations within an image. Typically, millions of windows must be inspected by video frame at speeds ranging from 5–30 frames per second.

In cascade detection, the detector operates throughout a series of stages. In the first stages, the algorithm quickly identifies and discards windows that it can easily recognize as not containing a person (such as the sky). The next stages process the windows that are harder for the algorithm to classify, such as those containing a tree, which the algorithm could recognize as having person-like features (shape, color, contours, etc.). In the final stages, the algorithm must distinguish between a pedestrian and very similar objects. However, because the final stages only process a few windows, the overall complexity is low.

Traditional cascade detection relies on “weak learners,” which are simple classifiers, to do the job at each stage. The first stages use a small number of weak learners to reject the easy windows, while the later stages rely on larger numbers of weak learners to process the harder windows. While this method is fast, it isn’t powerful enough when it reaches the final stages. That’s because the weak learners used in all stages of the cascade are identical. So even though there are more classifiers in the last stages, they’re not necessarily capable of performing highly complex classification.

Deep learning models

To address this problem, Vasconcelos and his team developed a novel algorithm that incorporates deep learning models in the final stages of a cascaded detector. Deep learning models are better suited for complex pattern recognition, which they can perform after being trained with hundreds or thousands of examples—in this case, images that either have or don’t have a person. However, deep learning models are too complex for real-time implementation. While they work well for the final cascade stages, they are too complex to be used in the early ones.

The solution is a new cascade architecture that combines classifiers from different families: simple classifiers (weak learners) in the early stages complex classifiers (deep learning models) in the later stages. This is not trivial to accomplish, noted Vasconcelos, since the algorithm used to learn the cascade has to find the combination of weak learners that achieves the optimal trade-off between detection accuracy and complexity for each cascade stage. Accordingly, Vasconcelos and his team introduced a new mathematical formulation for this problem, which resulted in a new algorithm for cascade design.

“No previous algorithms have been capable of optimizing the trade-off between detection accuracy and speed for cascades with stages of such different complexities. In fact, these are the first cascades to include stages of deep learning. The results we’re obtaining with this new

algorithm are substantially better for real-time, accurate pedestrian detection,” said Vasconcelos.

The algorithm currently only works for binary detection tasks, such as pedestrian detection, but the researchers are aiming to extend the cascade technology to detect many objects simultaneously.

“One approach to this problem is to train, for example, five different detectors to recognize five different objects. But we want to train just one detector to do this. Developing that algorithm is the next challenge,” said Vasconcelos.

Click [here](#) to watch a video of the pedestrian detection system in action.

The work, titled “Learning Complexity-Aware Cascades for Deep Pedestrian Detection,” was presented Dec. 15, 2015 at the International Conference on Computer Vision in Santiago, Chile. The research team also included Zhaowei Cai of UC San Diego and Mohammed Saberian of Yahoo Labs. The work was supported by awards from the National Science Foundation and funding from Northrop Grumman.

MEDIA CONTACT

Liesel Labios, 858-246-1124, llabios@ucsd.edu

UC San Diego’s [Studio Ten 300](#) offers radio and television connections for media interviews with our faculty, which can be coordinated via studio@ucsd.edu. To connect with a UC San Diego faculty expert on relevant issues and trending news stories, visit <https://ucsdnews.ucsd.edu/media-resources/faculty-experts>.