Deep Neural Networks for Object Detection

Christian Szegedy Alexander Toshev Dumitru Erhan
Google, Inc.
fszegedy, toshev, dumitrug@google.com

Abstract

Deep Neural Networks (DNNs) have recently shown outstanding performance on

In this way, only through a few dozen DNN-regressions we can achieve state-of-art bounding box localization.

In this paper, we demonstrate that DNN-based regression is capable of learning features which are not only good for classification, but also capture *strong geometric information*. We use the

The complete algorithm is outlined in Algorithm 1.

Algorithm 1: Overall algorithm: multi-scale DNN-based localization and subsequent refinement. The above algorithm is applied for each object class separately.

Input: x input image of size; networks

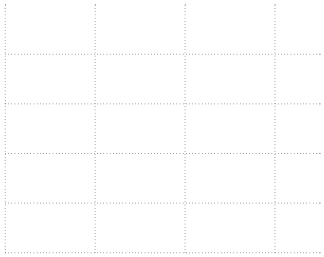

Figure 4: Precision recall curves of DetectorNet after the first stage and after the refinement.

0:5 to discard boxes. After the initial training, we performed two rounds of hard negative mining on the training set. This added two million examples to our original training set and has cut down the ratio of false positives.

The second approach is the 3-layer compositional model by [19] which can be considered a deep architecture. As a co-winner of VOC2011 this approach has shown excellent performance. Finally, we compare against the DPM by [9] and [11].

Although our comparison is somewhat unfair, as we trained on the larger VOC2012 training set, we show state-of-the art performance on most of the models: we outperform on 8 classes and perform on par on other 1. Note that it might be possible to tune the sliding window to perform on par with DetectorNet, however the sheer amount of network evaluations makes that approach infeasible while DetectorNet requires only (#windows #mask types) 120 crops per class to be evaluated. On

References

[1] Narendra Ahuja and Sinisa Todorovic. Learning the taxonomy and models of categories present in arbitrary images. In *International Conference on Computer Vision*, 2007.

[2] Yoshua Bengio. Learning deep architectures for ai. *Foundations and Trends R in Machine Learning*, 2(1):1–127, 2009.

[3] Dan Ciresan, Alessandro Giusti, Juergen Schmidhuber, et al. Deep neural networks segment neuronal membranes in electron microscopy images. In