

Discriminative Structure Learning of Hierarchical Representations for Object Detection Supplementary Material

Paul Schnitzspan¹, Mario Fritz², Stefan Roth¹, and Bernt Schiele¹

¹Department of Computer Science, TU Darmstadt

²UC Berkeley EECS & ICSI

Due to lack of space in the paper we show additional examples of true and false positives of our model in the supplementary material (supplementary to Fig. 2 in the paper).

In Fig. 1 and Fig. 2 we show the highest-scored true positives and the highest-scored false positives of our model with linear kernels (hHOG+hBoW) and the model of Felzenszwalb *et al.* [4] (DPM) for all 20 classes of the PASCAL VOC 2007 object detection challenge. A visual assessment shows that our approach appears to be more flexible in modeling changes in viewpoint, appearance and articulation (see, *e.g.*, the category horse for higher flexibility in viewpoint and articulation, and aeroplane for higher flexibility in appearance). Furthermore, many of the highest-scored false positives of our model are simply due to misaligned bounding boxes containing the object of interest (*e.g.*, aeroplane, bird, bicycle, bus, car, horse, motorbike, person, sofa, train) and thus appear to be more easily interpretable than those from the DPM approach. For classes car, cow, cat, motorbike, sheep, train our model yields sensible false positives (*e.g.*, detecting a bicycle as a motorbike), which again appear to be more interpretable than those of DPM.






































































































































































































PASCAL VOC 2007	5 highest-scored <i>true positives</i>					5 highest-scored <i>false positives</i>				
DPM aeroplane										
Our model aeroplane										
DPM bicycle										
Our model bicycle										
DPM bird										
Our model bird										
DPM boat										
Our model boat										
DPM bottle										
Our model bottle										
DPM bus										
Our model bus										
DPM car										
Our model car										
DPM cat										
Our model cat										
DPM chair										
Our model chair										
DPM cow										
Our model cow										

Figure 1. Comparison of the highest-scored true and false positives of Felzenszwalb *et al.* [4] (DPM) and our model with linear kernels (hHOG+hBoW). Classes: aeroplane, bicycle, bird, boat, bottle, bus, car, cat, chair, cow.

PASCAL VOC 2007	5 highest-scored <i>true positives</i>					5 highest-scored <i>false positives</i>				
DPM diningtable										
Our model diningtable										
DPM dog										
Our model dog										
DPM horse										
Our model horse										
DPM motorbike										
Our model motorbike										
DPM person										
Our model person										
DPM pottedplant										
Our model pottedplant										
DPM sheep										
Our model sheep										
DPM sofa										
Our model sofa										
DPM train										
Our model train										
DPM tvmonitor										
Our model tvmonitor										

Figure 2. Comparison of the highest-scored true and false positives of Felzenszwalb *et al.* [4] (DPM) and our model with linear kernels (hHOG+hBoW). Classes: diningtable, dog, horse, motorbike, person, pottedplant, sheep, sofa, train, tvmonitor.