

Supplementary Material

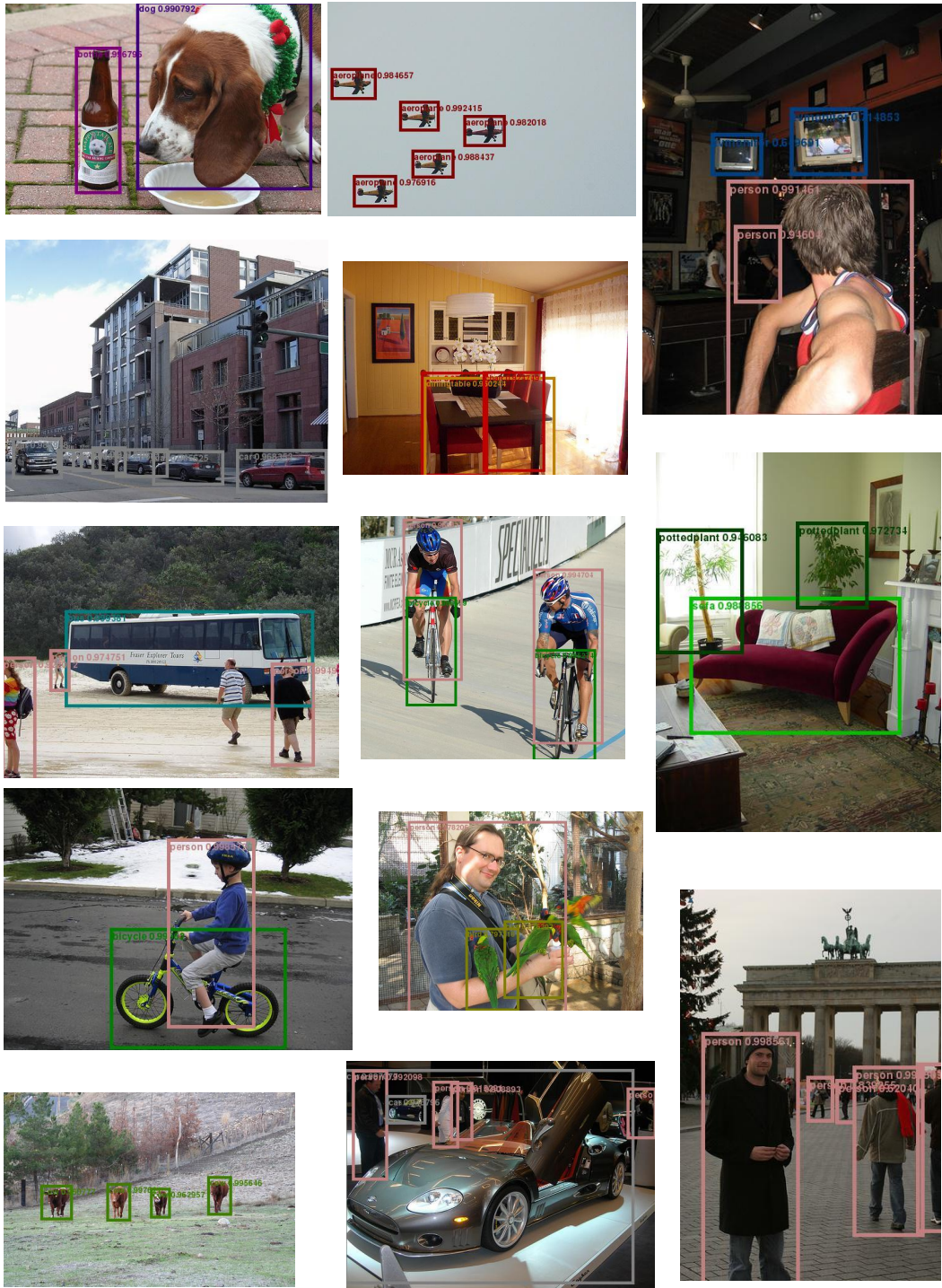


Figure 1: **Qualitative results for the tasks of object detection.** The results are obtained by the BlitzNet512 trained on VOC07 and VOC12 train-val augmented with extra segmentation masks.



Figure 2: **Improved and failure cases of detection by BlitzNet300 comparing to SSD300.** Each pair of images corresponds to the results of detection by SSD300 (left) and BlitzNet300 (right). The cases of improved detection are presented on the top part of the figure and the cases where both methods still fail are placed below the dashed line. It's clear that our pipeline provides more accurate detections in presence of small objects, complicated scenes and objects consisting of several parts with different appearance. The failure cases indicate that modern pipelines still struggle to handle ambiguous big objects (top left), intraclass variability (top right), misleading context (bottom right) and highly occluded objects (bottom left)