Multiview RGB-D Dataset for Object Instance Detection (Supplementary)

Figure 1. Remaining sparse reconstructions of our scenes using the state-of-the-art incremental Structure from Motion (SfM) algorithm of COLMAP (left column) and their corresponding dense point clouds (middle column). The rightmost column shows the per 3D point annotation that is provided by the dataset.

Figure 2. More failure examples of Selective search (upper row) where the multi-view approach is successful (lower row). The blue bounding boxes are the ground truth, the green are the proposals with the highest overlap with the ground truth, and the red are the failure cases.
Figure 3. Qualitative detection results for all four baselines. From left to right: HMP Scene Folds, Turntable, Turntable Background, and CNN Scene Folds. Blue bounding boxes are ground truth, green are correct detections, and red are false detections. False detections can occur by misclassification or by insufficient overlap with the ground truth. Proposals that are classified as background are not shown for clarity. The labels shown in the images are: 1: coca cola, 2: coffee mate, 3: honey bunches, 4: hunts sauce, 5: mahatma rice, 6: nature valley 1(yellow), 7: nature valley 2, 8: palmoilive orange, 9: pop secret, 10: pringles bbq, 11: red bull.