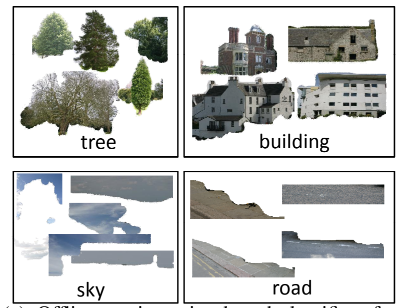
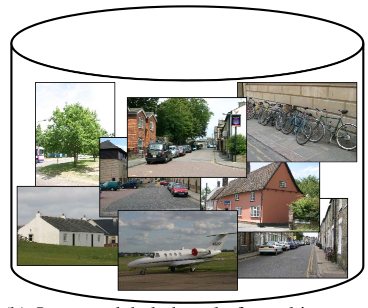
Supplementary file for: Object-Graphs for Context-Aware Category Discovery

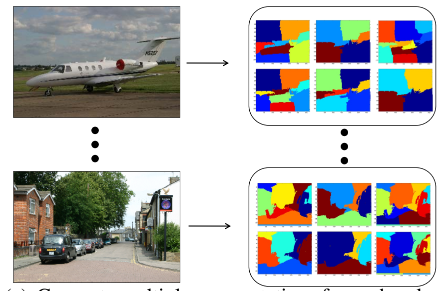
Yong Jae Lee and Kristen Grauman



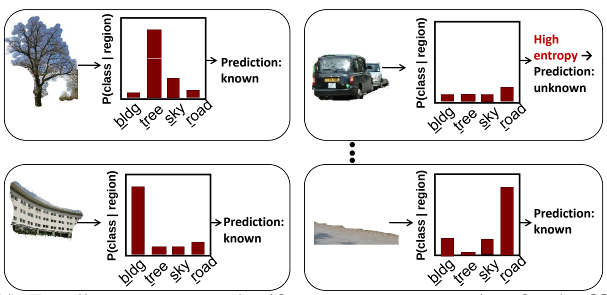
(a) Offline: train region-based classifiers for N "known" categories using labeled data.



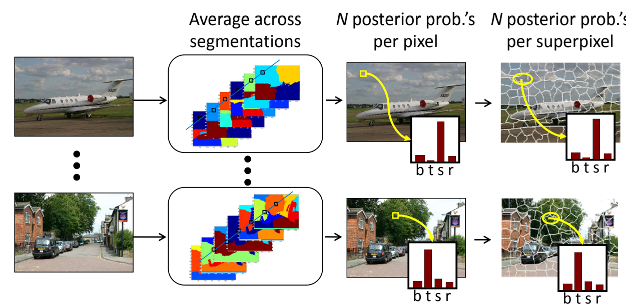
(b) Input: unlabeled pool of novel images.



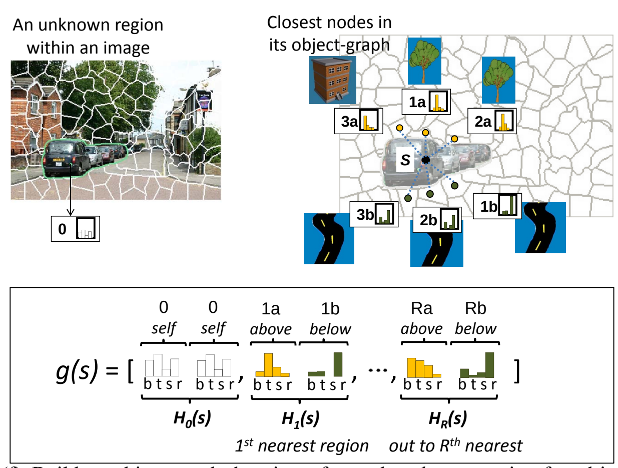
(c) Compute multiple segmentations for each unlabeled image



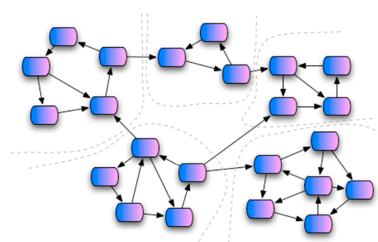
(d) For all segments, use classifiers to compute posteriors for the N "known" categories. Deem each segment as either *known* or *unknown* based on resulting entropy score.



(e) Map the per-region posteriors to per-pixel posteriors by averaging values pixel-wise across all segmentations computed for a given image. Superpixel regions then assigned posteriors using average of their member pixels.



(f) Build an object-graph descriptor for each *unknown* region found in the unlabeled pool.



$$K(s_m, s_n) = K_{\chi^2}\left(\mathrm{a}(s_m), \mathrm{a}(s_n)\right) + K_{\chi^2}\left(\mathrm{g}(s_m), \mathrm{g}(s_n)\right)$$

(g) Compute affinities between all pairs of unknown regions based on both their appearance and object-graph similarity. Cluster using those affinities. Each node here represents an unknown region.

(h) Output: discovered objects

Figure 1. Overview of the proposed method. Read from (a) to (h) in order.

1. Details on dataset splits tested in our experiments

For each dataset, we form multiple splits of known/unknown classes, for multiple settings of both the number of knowns (N) and the number of true unknowns present (denoted U). For MSRC-v2, we create two sets for each of three different split sizes: $U = [5, 10, 15]$, $N = [16, 11, 6]$, forming six total variations. Similarly, for the PASCAL, we create two sets each for three sizes: $U = [5, 10, 15]$, $N = [15, 10, 5]$. For the smaller Corel set we create a single split with $U = 2$ and $N = 5$. For MSRC-v0, we create a single split with $U = 8$ and $N = 13$. For the Corel and MSRC-v0 we choose the split manually, selecting as unknown those categories that we think could benefit most from object-level context. However, for the MSRC-v2 and PASCAL we select all 12 splits randomly. See Table 1 for a detailed breakdown of the category names in each split.

	Unknown Categories	**Known Categories**
MSRC-v2 set 1	*building, tree, cow, airplane, bicycle*	*grass, sheep, sky, water, face, car, flower, sign, bird, book, chair, road, cat, dog, body, boat*
MSRC-v2 set 2	*grass, sky, water, road, dog*	*building, tree, cow, sheep, airplane, face, car, bicycle, flower, sign, bird, book, chair, cat, body, boat*
MSRC-v2 set 3	*tree, sheep, sky, airplane, water, bicycle, bird, chair, road, boat*	*building, grass, cow, face, car, flower, sign, book, cat, dog, body*
MSRC-v2 set 4	*tree, cow, sheep, face, flower, sign, bird, chair, cat, body*	*building, grass, sky, airplane, water, car, bicycle, book, road, dog, boat*
MSRC-v2 set 5	*building, cow, sheep, sky, airplane, face, bicycle, sign, bird, book, chair, road, cat, body, boat*	*grass, tree, water, car, flower, dog*
MSRC-v2 set 6	*grass, sheep, sky, airplane, water, face, flower, sign, bird, book, chair, road, dog, body, boat*	*building, tree, cow, car, bicycle, cat*
PASCAL VOC 2008 set 1	*airplane, car, cow, motorbike, tv/monitor*	*bicycle, bird, boat, bottle, bus, cat, chair, diningtable, dog, horse, person, pottedplant, sheep, sofa, train*
PASCAL VOC 2008 set 2	*bicycle, bird, chair, sofa, train*	*airplane, boat, bottle, bus, car, cat, cow, diningtable, dog, horse, motorbike, person, pottedplant, sheep, tvmonitor*
PASCAL VOC 2008 set 3	*airplane, boat, bottle, bus, chair, diningtable, motorbike, sofa, train, tv/monitor*	*bicycle, bird, car, cat, cow, dog, horse, person, pottedplant, sheep*
PASCAL VOC 2008 set 4	*airplane, bicycle, bottle, bus, dog, motorbike, person, sheep, sofa, train*	*bird, boat, car, cat, chair, cow, diningtable, horse, pottedplant, tv/monitor*
PASCAL VOC 2008 set 5	*boat, bottle, bus, car, cat, chair, cow, diningtable, dog, horse, motorbike, pottedplant, sofa, train, tv/monitor*	*airplane, bicycle, bird, person, sheep*
PASCAL VOC 2008 set 6	*airplane, bicycle, bird, boat, bottle, bus, cat, cow, dog, horse, motorbike, person, pottedplant, train, tv/monitor*	*car, chair, diningtable, sheep, sofa*
MSRC-v0	*building, tree, cow, sheep, car, bicycle, sign, window*	*grass, sky, mountain, airplane, water, flower, bird, chair, road, body, leaf, chimney, door*
Corel	*rhino/hippo, polar bear*	*water, snow, vegetation, ground, sky*

Table 1. Breakdown of category names for each dataset.

2. Results

This section contains additional supporting results, as referenced in the main paper.

To ensure that the improvement over [1] is not a result of stronger appearance features, we repeated the experiment using the same features for all methods, letting $a(s)$ be a SIFT bag of words as in [1]. Figure 2 shows the results. Our full model significantly outperforms the baselines, even though it performs slightly worse than when using TH, CH, and pHOG for appearance features (Figure 5 in main paper). Note that the two appearance-based methods (black and red curves) show even closer results.

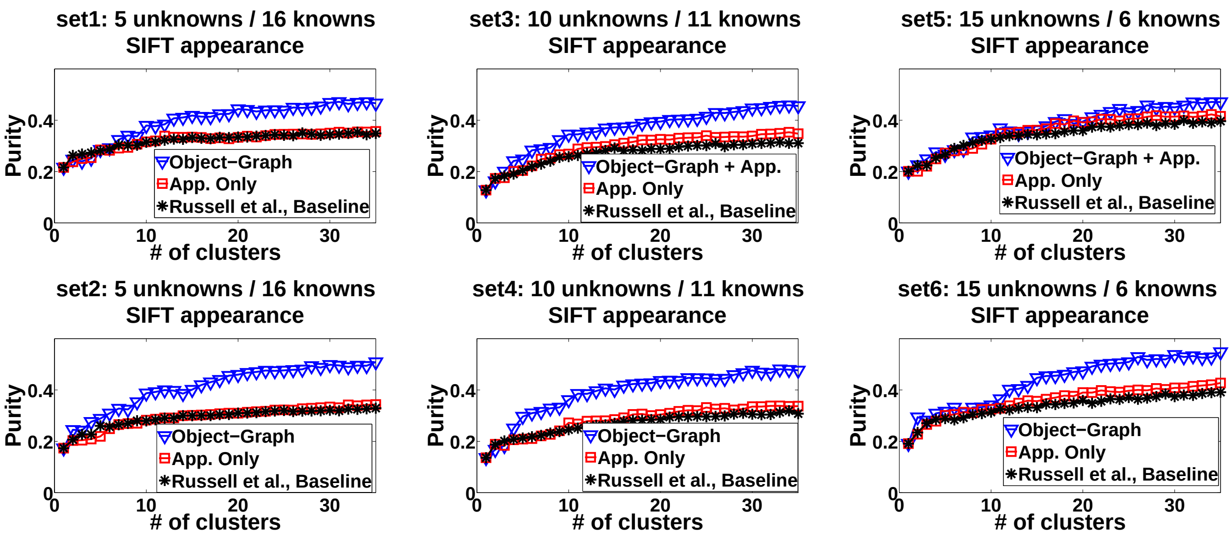
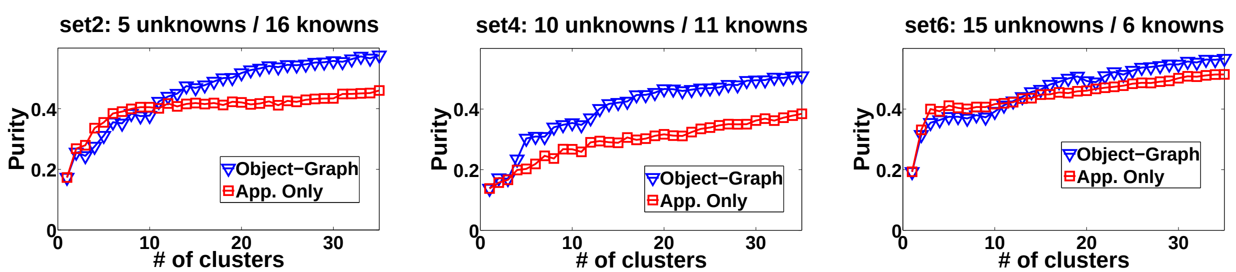
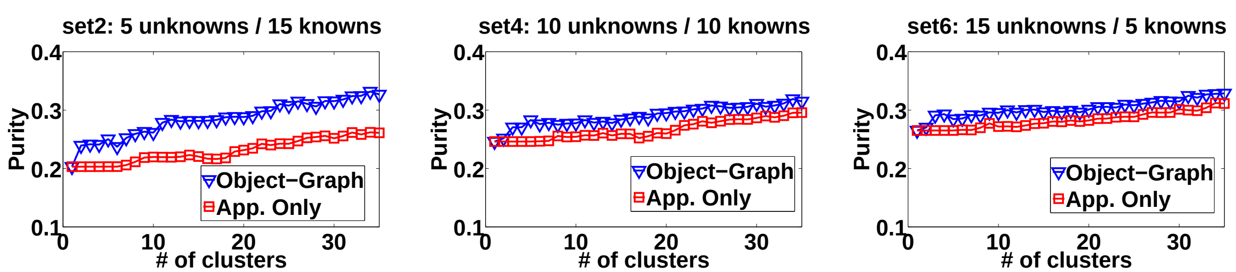


Figure 2. MSRC-v2 discovery accuracy with SIFT features.

We also show additional discovery results on different splits of the MSRC-v2 and PASCAL datasets in Figure 3 (top and bottom rows, respectively). Our model significantly outperforms the appearance-only baseline. For the PASCAL, the improvement over the appearance-only baseline is not as great compared to that on MSRC-v2, especially as the number of unknown categories increases—this can be attributed to the difficulty of the dataset where prediction even of the trained "knowns" is weaker and less reliable.



(a) MSRC-v2



(b) PASCAL VOC 08

Figure 3. MSRC-v2 and PASCAL discovery accuracy results.

References

[1] B. Russell, A. Efros, J. Sivic, W. Freeman, and A. Zisserman. Using Multiple Segmentations to Discover Objects and their Extent in Image Collections. In *CVPR*, 2006.