Conclusion

We've proposed a method to learn discriminative classifiers directly from keyword search returns along with a novel iterative refinement technique to simultaneously improve both the classifier and the representation of training examples.

Advantages of our approach:

- allows direct specification of categories of interest
- flexible to choice of kernel and features

On both binary classification and image re-ranking our approach outperforms some of the best unsupervised techniques and performs comparable to state-of-the-art supervised techniques on a variety of datasets.

In the future we are interested in considering how prior knowledge about a category's expected sparsity might be captured in order to boost accuracy.

References

1. R. Bunescu and R. Mooney. Multiple Instance Learning for Sparse Positive Bags. In *ICML*, 2007
2. R. Fergus, L. Fei-Fei, P. Perona, and A. Zisserman. Learning Object Categories from Google's Image Search. In *ICCV*, 2005.
3. F. Schroff, A. Criminisi, and A. Zisserman. Harvesting Image Databases from the Web. In *ICCV*, 2007
4. T. Berg and D. Forsyth. Animals on the Web. In *CVPR*, 2006.
5. T. Dietterich, R. Lathrop, and T. Lozano-Perez. Solving the multiple instance problem with axis-parallel rectangles. *Artificial Intelligence*, 89(1-2):31–71,1997
6. Li-Jia Li, Gang Wang and Li Fei-Fei. OPTIMOL: automatic Object Picture collecTion via Incremental MOdel Learning. In *CVPR,* 2007.