

probability, a remarkable achievement considering that this dataset was *designed* to be challenging for machines.

Acknowledgements. We are grateful for financial support from the UKIERI, EU Project AXES ICT-269980 and ERC grant VisRec no. 228180.

References

- [1] American kennel club. <http://www.akc.org/>.
- [2] The cat fanciers association inc. <http://www.cfa.org/Client/home.aspx>.
- [3] Cats in sinks. <http://catsinsinks.com/>.
- [4] Catster. <http://www.catster.com/>.
- [5] Dogster. <http://www.dogster.com/>.
- [6] Flickr!. <http://www.flickr.com/>.
- [7] Google images. <http://images.google.com/>.
- [8] The international cat association. <http://www.tica.org/>.
- [9] My cat space. <http://www.mycatspace.com/>.
- [10] My dog space. <http://www.mydogspace.com/>.
- [11] Petfinder. <http://www.petfinder.com/index.html>.
- [12] World canine organisation. <http://www.fci.be/>.
- [13] P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik. From contours to regions: An empirical evaluation. In *Proc. CVPR*, 2009.
- [14] S. Branson, C. Wah, F. Schroff, B. Babenko, P. Welinder, P. Perona, and S. Belongie. Visual recognition with humans in the loop. In *Proc. ECCV*, 2010.
- [15] Y. Chai, V. Lempitsky, and A. Zisserman. Bicos: A bi-level co-segmentation method for image classification. In *Proc. ICCV*, 2011.
- [16] G. Csurka, C. R. Dance, L. Dan, J. Willamowski, and C. Bray. Visual categorization with bags of keypoints. In *Proc. ECCV Workshop on Stat. Learn. in Comp. Vision*, 2004.
- [17] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *Proc. CVPR*, 2005.
- [18] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. ImageNet: A Large-Scale Hierarchical Image Database. In *Proc. CVPR*, 2009.
- [19] J. Elson, J. Douceur, J. Howell, and J. J. Saul. Asirra: A CAPTCHA that exploits interest-aligned manual image categorization. In *Conf. on Computer and Communications Security (CCS)*, 2007.
- [20] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman. The PASCAL Visual Object Classes Challenge 2011 (VOC2011) Results. <http://www.pascal-network.org/challenges/VOC/voc2011/workshop/index.html>.
- [21] R.-E. Fan, K.-W. Chang, C.-J. Hsieh, X.-R. Wang, and C.-J. Lin. LIBLINEAR: A library for large linear classification. *Journal of Machine Learning Research*, 9, 2008.
- [22] L. Fei-Fei, R. Fergus, and P. Perona. A Bayesian approach to unsupervised one-shot learning of object categories. In *Proc. ICCV*, 2003.
- [23] P. F. Felzenszwalb, R. B. Grishick, D. McAllester, and D. Ramanan. Object detection with discriminatively trained part based models. *PAMI*, 2009.
- [24] F. Fleuret and D. Geman. Stationary features and cat detection. *Journal of Machine Learning Research*, 9, 2008.
- [25] P. Golle. Machine learning attacks against the asirra captcha. In *15th ACM Conference on Computer and Communications Security (CCS)*, 2008.
- [26] G. Griffin, A. Holub, and P. Perona. Caltech-256 object category dataset. Technical report, California Institute of Technology, 2007.
- [27] A. Khosla, N. Jayadevaprakash, B. Yao, and F. F. Li. Novel dataset for fine-grained image categorization. In *First Workshop on Fine-Grained Visual Categorization, CVPR*, 2011.
- [28] C. Lampert, H. Nickisch, and S. Harmeling. Learning to detect unseen object classes by between-class attribute transfer. In *Proc. CVPR*, 2009.
- [29] I. Laptev. Improvements of object detection using boosted histograms. In *Proc. BMVC*, 2006.
- [30] S. Lazebnik, C. Schmid, and J. Ponce. Beyond bag of features: Spatial pyramid matching for recognizing natural scene categories. In *Proc. CVPR*, 2006.
- [31] D. G. Lowe. Object recognition from local scale-invariant features. In *Proc. ICCV*, 1999.
- [32] M.-E. Nilsback and A. Zisserman. A visual vocabulary for flower classification. In *Proc. CVPR*, 2006.
- [33] M.-E. Nilsback and A. Zisserman. Automated flower classification over a large number of classes. In *Proc. ICVGIP*, 2008.
- [34] O. Parkhi, A. Vedaldi, C. V. Jawahar, and A. Zisserman. The truth about cats and dogs. In *Proc. ICCV*, 2011.
- [35] O. Parkhi, A. Vedaldi, A. Zisserman, and C. V. Jawahar. The Oxford-IIIT PET Dataset. <http://www.robots.ox.ac.uk/~vgg/data/pets/index.html>, 2012.
- [36] C. Rother, V. Kolmogorov, and A. Blake. “grabcut” — interactive foreground extraction using iterated graph cuts. In *ACM Trans. on Graphics*, 2004.
- [37] B. Schölkopf and A. J. Smola. *Learning with Kernels*. MIT Press, 2002.
- [38] J. Sivic and A. Zisserman. Video Google: A text retrieval approach to object matching in videos. In *Proc. ICCV*, 2003.
- [39] M. Varma and D. Ray. Learning the discriminative power-invariance trade-off. In *Proc. ICCV*, 2007.
- [40] A. Vedaldi and B. Fulkerson. VLFeat library. <http://www.vlfeat.org/>, 2008.
- [41] A. Vedaldi, V. Gulshan, M. Varma, and A. Zisserman. Multiple kernels for object detection. In *Proc. ICCV*, 2009.
- [42] A. Vedaldi and A. Zisserman. Structured output regression for detection with partial occlusion. In *Proc. NIPS*, 2009.
- [43] P. Welinder, S. Branson, T. Mita, C. Wah, and F. Schroff. Caltech-ucsd birds 200. Technical report, Caltech-UCSD, 2010.
- [44] J. Zhang, M. Marszalek, S. Lazebnik, and C. Schmid. Local features and kernels for classification of texture and object categories: A comprehensive study. *IJCV*, 2007.
- [45] W. Zhang, J. Sun, and X. Tang. Cat head detection - how to effectively exploit shape and texture features. In *Proc. ECCV*, 2008.