



## UvA-DARE (Digital Academic Repository)

### Nuances in visual recognition

Gavves, E.

**Publication date**  
2014

[Link to publication](#)

**Citation for published version (APA):**  
Gavves, E. (2014). *Nuances in visual recognition*.

#### **General rights**

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

#### **Disclaimer/Complaints regulations**

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

---

## BIBLIOGRAPHY

---

- [1] Z. Akata, F. Perronnin, Z. Harchaoui, and C. Schmid. Label-embedding for attribute-based classification. In *Proc. IEEE Conf. Comput. Vision and Pattern Recognition*, 2013.
- [2] B. Alexe, N. Heess, Y. Teh, and V. Ferrari. Searching for objects driven by context. In *Proc. Neural Inform. and Process. Systems*, 2012.
- [3] R. Arandjelovic and A. Zisserman. Three things everyone should know to improve object retrieval. In *Proc. IEEE Conf. Comput. Vision and Pattern Recognition*, 2012.
- [4] P. Arbelaez, B. Hariharan, C. Gu, S. Gupta, L. Bourdev, and J. Malik. Semantic segmentation using regions and parts. In *Proc. IEEE Conf. Comput. Vision and Pattern Recognition*, 2012.
- [5] P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik. From contours to regions: An empirical evaluation. In *Proc. IEEE Conf. Comput. Vision and Pattern Recognition*, 2009.
- [6] Aristotle. *Categories*.
- [7] Y. S. Avrithis, Y. Kalantidis, G. Toliass, and E. Spyrou. Retrieving landmark and non-landmark images from community photo collections. In *Proc. ACM Int'l. Conf. Multimedia*, 2010.
- [8] H. Azizpour and I. Laptev. Object detection using strongly-supervised deformable part models. In *Proc. Eur. Conf. Comput. Vision*, 2012.
- [9] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool. Speeded-up robust features (surf). *Comput. Vision and Image Understanding*, 2008.
- [10] Y. Bengio, A. Courville, and P. Vincent. Representation learning: A review and new perspectives. *IEEE Trans. Pattern Analysis and Machine Intell.*, 2013.
- [11] T. Berg and P. N. Belhumeur. Poof: Part-based one-vs-one features for fine-grained categorization, face verification, and attribute estimation. In *Proc. IEEE Conf. Comput. Vision and Pattern Recognition*, 2013.
- [12] I. Biederman. Recognition-by-components: A theory of human image understanding. *Psychological Review*, 1987.
- [13] C. M. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006.
- [14] L. Bo, X. Ren, and D. Fox. Kernel descriptors for visual recognition. In *Proc. Neural Inform. and Process. Systems*, 2010.
- [15] L. Bo and C. Sminchisescu. Efficient match kernels between sets of features for visual recognition. In *Proc. Neural Inform. and Process. Systems*, 2009.
- [16] L. Bourdev and J. Malik. Poselets: Body part detectors trained using 3d human pose annotations. In *Proc. Int'l. Conf. Comput. Vision*, 2009.
- [17] Y. Boureau, J. Ponce, and Y. LeCun. A theoretical analysis of feature pooling in visual recognition. In *Proc. Int'l. Conf. Machine Learning*, 2010.
- [18] S. Branson, P. Perona, and S. Belongie. Strong supervision from weak annotation: Interactive training of deformable part models. In *Proc. Int'l. Conf. Comput. Vision*, 2011.
- [19] S. Branson, C. Wah, F. Schroff, B. Babenko, P. Welinder, P. Perona, and S. Belongie. Visual recognition with humans in the loop. In *Proc. Eur. Conf. Comput. Vision*, 2010.
- [20] J. Bucklew. *Introduction to rare event simulation*. Elsevier, 2004.
- [21] J. Carreira, R. Caseiro, J. Batista, and C. Sminchisescu. Semantic Segmentation with Second-Order Pooling. In *Proc. Eur. Conf. Comput. Vision*, 2012.
- [22] J. Carreira, F. Li, and C. Sminchisescu. Object recognition by sequential figure-ground ranking. *Int. J. of Comput. Vision*, 2012.
- [23] J. Carreira and C. Sminchisescu. Cpmc: Automatic object segmentation using constrained parametric min-cuts. *IEEE Trans. Pattern Analysis and Machine Intell.*, 2012.
- [24] C. Carson, S. Belongie, H. Greenspan, and J. Malik. Blobworld: Image segmentation using expectation-maximization and its application to image querying. *IEEE Trans. Pattern Analysis and Machine Intell.*, 2002.

## Bibliography

- [25] Y. Chai, V. Lempitsky, and A. Zisserman. Bicos: A bi-level co-segmentation method for image classification. In *Proc. Int'l. Conf. Comput. Vision*, 2011.
- [26] Y. Chai, V. Lempitsky, and A. Zisserman. Symbiotic segmentation and part localization for fine-grained categorization. In *Proc. Int'l. Conf. Comput. Vision*, 2013.
- [27] Y. Chai, E. Rahtu, V. Lempitsky, L. Van Gool, and A. Zisserman. TriCoS: a tri-level class-discriminative co-segmentation method for image classification. In *Proc. Eur. Conf. Comput. Vision*, 2012.
- [28] K. Chatfield, V. Lempitsky, A. Vedaldi, and A. Zisserman. The devil is in the details: an evaluation of recent feature encoding methods. In *Proc. Brit. Mach. Vision Conf.*, 2011.
- [29] O. Chum, J. Matas, and J. Kittler. Locally optimized ransac. *Patt. Recogn.*, 2003.
- [30] O. Chum, J. Philbin, J. Sivic, M. Isard, and A. Zisserman. Total recall: Automatic query expansion with a generative feature model for object retrieval. In *Proc. Int'l. Conf. Comput. Vision*, 2007.
- [31] R. G. Cinbis, J. Verbeek, and C. Schmid. Segmentation driven object detection with fisher vectors. In *Proc. Int'l. Conf. Comput. Vision*, 2013.
- [32] G. Csurka, C. Dance, L. Fan, J. Willamowski, and C. Bray. Visual categorization with bags of keypoints. In *Proc. Eur. Conf. Comput. Vision*, 2004.
- [33] G. Csurka and F. Perronnin. An efficient approach to semantic segmentation. *Int. J. of Comput. Vision*, 2011.
- [34] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *Proc. IEEE Conf. Comput. Vision and Pattern Recognition*, 2005.
- [35] C. Darwin. *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. 1859.
- [36] J. Deng, W. Dong, R. Socher, L. Li, K. Li, and L. Fei-Fei. Imagenet: A large-scale hierarchical image database. In *Proc. IEEE Conf. Comput. Vision and Pattern Recognition*, 2009.
- [37] J. Donahue, Y. Jia, O. Vinyals, J. Hoffman, N. Zhang, E. Tzeng, and T. Darrell. Decaf: A deep convolutional activation feature for generic visual recognition. Technical report, 2013. arXiv:1310.1531.
- [38] K. Duan, D. Parikh, D. Crandall, and K. Grauman. Discovering localized attributes for fine-grained recognition. In *Proc. IEEE Conf. Comput. Vision and Pattern Recognition*, 2012.
- [39] M. Everingham, L. Van Gool, C. K. Williams, J. Winn, and A. Zisserman. The pascal visual object classes (voc) challenge. *Int. J. of Comput. Vision*, 2010.
- [40] R. Farrell, O. Oza, N. Zhang, V. I. Morariu, T. Darrell, and L. S. Davis. Birdlets: Subordinate categorization using volumetric primitives and pose-normalized appearance. In *Proc. Int'l. Conf. Comput. Vision*, 2011.
- [41] P. Felzenszwalb, R. Girshick, D. McAllester, and D. Ramanan. Object detection with discriminatively trained part-based models. *IEEE Trans. Pattern Analysis and Machine Intell.*, 2010.
- [42] P. F. Felzenszwalb and D. P. Huttenlocher. Efficient graph-based image segmentation. *Int. J. of Comput. Vision*, 2004.
- [43] B. Fulkerson, A. Vedaldi, and S. Soatto. Localizing objects with smart dictionaries. In *Proc. Eur. Conf. Comput. Vision*, 2008.
- [44] E. Gavves, B. Fernando, C. G. Snoek, A. W. Smeulders, and T. Tuytelaars. Fine-grained categorization by alignments. In *Proc. Int'l. Conf. Comput. Vision*, 2013.
- [45] E. Gavves, C. Snoek, and A. W. Smeulders. Visual synonyms for landmark image retrieval. *Comput. Vision and Image Understanding*, 2011.
- [46] E. Gavves, C. Snoek, and A. W. Smeulders. Convex reduction of high-dimensional kernels for visual classification. In *Proc. IEEE Conf. Comput. Vision and Pattern Recognition*, 2012.
- [47] E. Gavves and C. G. M. Snoek. Landmark image retrieval using visual synonyms. In *Proc. ACM Int'l. Conf. Multimedia*, 2010.
- [48] R. B. Girshick, J. Donahue, T. Darrell, and J. Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. *CoRR*, 2013.
- [49] P. H. Gosselin, N. Murray, H. Jégou, and F. Perronnin. Inria+Xerox@FGcomp: Boosting the Fisher vector for fine-grained classification. Research Report RR-8431, INRIA, 2013.
- [50] H. Grabner, J. Gall, , and L. V. Gool. What makes a chair a chair? In *Proc. IEEE Conf. Comput. Vision and Pattern Recognition*, 2011.
- [51] K. Grauman and T. Darrell. The pyramid match kernel: Efficient learning with sets of features. *J. of Machine Learn. Resear.*, 2007.

- [52] B. Hariharan, P. Arbelaez, L. D. Bourdev, S. Maji, and J. Malik. Semantic contours from inverse detectors. In *Proc. Int'l. Conf. Comput. Vision*, 2011.
- [53] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2003.
- [54] H. Harzallah, F. Jurie, and C. Schmid. Combining efficient object localization and image classification. In *Proc. IEEE Conf. Comput. Vision and Pattern Recognition*, 2009.
- [55] Y. Huang, Z. Wu, L. Wang, and T. Tan. Feature coding in image classification: A comprehensive study. *IEEE Trans. Pattern Analysis and Machine Intell.*, 2013.
- [56] L. Itti, C. Koch, and E. Niebur. A model of saliency-based visual attention for rapid scene analysis. *IEEE Trans. Pattern Analysis and Machine Intell.*, 1998.
- [57] H. Jegou, M. Douze, and C. Schmid. Hamming embedding and weak geometric consistency for large scale image search. In *Proc. Eur. Conf. Comput. Vision*, 2008.
- [58] H. Jegou, F. Perronnin, M. Douze, J. Sánchez, P. Pérez, and C. Schmid. Aggregating local image descriptors into compact codes. *IEEE Trans. Pattern Analysis and Machine Intell.*, 2012.
- [59] Y. Jia, O. Vinyals, and T. Darrell. Pooling-invariant image feature learning. Technical report, 2013. arXiv:1302.5056.
- [60] Y.-G. Jiang, C.-W. Ngo, and J. Yang. Towards optimal bag-of-features for object categorization and semantic video retrieval. In *Proc. Int'l. Conf. Image and Video Retr.*, 2007.
- [61] F. Jurie and B. Triggs. Creating efficient codebooks for visual recognition. In *Proc. IEEE Conf. Comput. Vision and Pattern Recognition*, 2005.
- [62] I. Kant. *Kritik der reinen Vernunft*. 1781.
- [63] K.E.A. van de Sande, T. Gevers, and C.G.M. Snoek. Evaluating color descriptors for object and scene recognition. *IEEE Trans. Pattern Analysis and Machine Intell.*, 2010.
- [64] A. Khosla, N. Jayadevaprakash, B. Yao, and L. Fei-Fei. Novel dataset for fine-grained image categorization. In *Proc. IEEE Conf. Comput. Vision and Pattern Recognition*, 2011.
- [65] P. Krahenbuhl and V. Koltun. Efficient inference in fully connected crfs with gaussian edge potentials. In *Proc. Neural Inform. and Process. Systems*, 2011.
- [66] A. Krizhevsky, I. Sutskever, and G. Hinton. Imagenet classification with deep convolutional neural networks. *Proc. Neural Inform. and Process. Systems*, 2012.
- [67] Y. Kuo, W. Cheng, H. Lin, and W. Hsu. Unsupervised semantic feature discovery for image object retrieval and tag refinement. 2012.
- [68] L. Ladicky, C. Russell, P. Kohli, and P. Torr. Graph cut based inference with co-occurrence statistics. In *Proc. Eur. Conf. Comput. Vision*, 2010.
- [69] C. Lampert, M. Blaschko, and T. Hofmann. Efficient subwindow search: A branch and bound framework for object localization. *IEEE Trans. Pattern Analysis and Machine Intell.*, 2009.
- [70] C. H. Lampert, H. Nickisch, and S. Harmeling. Attribute-based transfer learning for object categorization with zero/one training example. *IEEE Trans. Pattern Analysis and Machine Intell.*, 2013.
- [71] S. Lazebnik and M. Raginsky. Supervised learning of quantizer codebooks by information loss minimization. *IEEE Trans. Pattern Analysis and Machine Intell.*, 2009.
- [72] S. Lazebnik, C. Schmid, and J. Ponce. Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. In *Proc. IEEE Conf. Comput. Vision and Pattern Recognition*, 2006.
- [73] Q. Le, M. Ranzato, R. Monga, M. Devin, K. Chen, G. Corrado, J. Dean, and A. Ng. Building high-level features using large scale unsupervised learning. In *Proc. Int'l. Conf. Machine Learning*, 2012.
- [74] H. Lee, A. Battle, R. Raina, and A. Y. Ng. Efficient sparse coding algorithms. In *Proc. Neural Inform. and Process. Systems*, 2007.
- [75] A. Levinstein, A. Stere, K. N. Kutulakos, D. J. Fleet, S. J. Dickinson, and K. Siddiqi. Turbopixels: Fast superpixels using geometric flows. *IEEE Trans. Pattern Analysis and Machine Intell.*, 2009.
- [76] D. Li, L. Yang, X. Hua, and H. Zhang. Large-scale robust visual codebook construction. In *Proc. ACM Int'l. Conf. Multimedia*, 2010.
- [77] F. Li, C. Ionescu, and C. Sminchisescu. Random fourier approximations for skewed multiplicative histogram kernels. In *German Assoc. for Patt. Recogn. DAGM*, 2010.
- [78] X. Li, E. Gavves, C. G. M. Snoek, M. Worring, and A. W. Smeulders. Personalizing automated image annotation using cross-entropy. In *Proc. ACM Int'l. Conf. Multimedia*, 2011.

## Bibliography

- [79] Z. Li, E. Gavves, K. E. A. van de Sande, C. G. M. Snoek, and A. W. Smeulders. Codemaps segment, classify and search objects locally. In *Proc. Int'l. Conf. Comput. Vision*, 2013.
- [80] J. Liu, A. Kanazawa, D. Jacobs, and P. Belhumeur. Dog breed classification using part localization. In *Proc. Eur. Conf. Comput. Vision*, 2012.
- [81] D. Lowe. Distinctive image features from scale-invariant keypoints. *Int. J. of Comput. Vision*, 2004.
- [82] S. Maji and A. Berg. Max-margin additive classifiers for detection. In *Proc. IEEE Conf. Comput. Vision and Pattern Recognition*, 2009.
- [83] S. Maji, A. Berg, and J. Malik. Classification using intersection kernel support vector machines is efficient. In *Proc. IEEE Conf. Comput. Vision and Pattern Recognition*, 2008.
- [84] S. Maji, J. Kannala, E. Rahtu, M. Blaschko, and A. Vedaldi. Fine-grained visual classification of aircraft. Technical report, 2013.
- [85] P. Mallapragada, R. Jin, and A. Jain. Online visual vocabulary pruning using pairwise constraints. In *Proc. IEEE Conf. Comput. Vision and Pattern Recognition*, 2010.
- [86] S. Manén, M. Guillaumin, and L. Van Gool. Prime object proposals with randomized prim's algorithm. In *Proc. Int'l. Conf. Comput. Vision*, 2013.
- [87] S. Mannor, D. Peleg, and R. Rubinstein. The cross entropy method for classification. In *Proc. Int'l. Conf. Machine Learning*, 2005.
- [88] J. Matas, O. Chum, U. Martin, and T. Pajdla. Robust wide baseline stereo from maximally stable extremal regions. In *Proc. Brit. Mach. Vision Conf.*, 2002.
- [89] I. Menache, S. Mannor, and N. Shimkin. Basis function adaptation in temporal difference reinforcement learning. *Annals of Operations Research*, 2005.
- [90] T. Mensink, E. Gavves, and C. G. M. Snoek. Co-occurrence statistics for zero-shot classification. *Proc. IEEE Conf. Comput. Vision and Pattern Recognition*, 2014.
- [91] K. Mikolajczyk and C. Schmid. Scale & affine invariant interest point detectors. *Int. J. of Comput. Vision*, 2004.
- [92] K. Mikolajczyk and C. Schmid. A performance evaluation of local descriptors. *IEEE Trans. Pattern Analysis and Machine Intell.*, 2005.
- [93] K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir, and L. Van Gool. A comparison of affine region detectors. *Int. J. of Comput. Vision*, 2005.
- [94] M. Muja and D. G. Lowe. Fast approximate nearest neighbors with automatic algorithm configuration. In *Int. Conf. Comp. Vision, Theory and Applic.*, 2009.
- [95] M. Nilsback and A. Zisserman. Automated flower classification over a large number of classes. In *ICVGIP*, 2008.
- [96] D. Nister and H. Stewenius. Scalable recognition with a vocabulary tree. In *Proc. IEEE Conf. Comput. Vision and Pattern Recognition*, 2006.
- [97] D. B. P. Yadollahpour and G. Shakhnarovich. Discriminative re-ranking of diverse segmentations. In *Proc. IEEE Conf. Comput. Vision and Pattern Recognition*, 2013.
- [98] S. J. Pan and Q. Yang. A survey on transfer learning. *IEEE Trans. Knowl. and Data Engineer.*, 22(10):1345–1359, October 2010.
- [99] D. Parikh and K. Grauman. Interactive discovery of task-specific nameable attributes. In *Proc. IEEE Conf. Comput. Vision and Pattern Recognition*, 2011.
- [100] O. M. Parkhi, A. Vedaldi, A. Zisserman, and C. V. Jawahar. Cats and dogs. In *Proc. IEEE Conf. Comput. Vision and Pattern Recognition*, 2012.
- [101] M. Perdoch, O. Chum, and J. Matas. Efficient representation of local geometry for large scale object retrieval. In *Proc. IEEE Conf. Comput. Vision and Pattern Recognition*, 2009.
- [102] F. Perronnin and C. R. Dance. Fisher kernels on visual vocabularies for image categorization. In *Proc. IEEE Conf. Comput. Vision and Pattern Recognition*, 2007.
- [103] F. Perronnin, J. Sánchez, and Y. Liu. Large-scale image categorization with explicit data embedding. In *Proc. IEEE Conf. Comput. Vision and Pattern Recognition*, 2010.
- [104] F. Perronnin, J. Sánchez, and T. Mensink. Improving the fisher kernel for large-scale image classification. In *Proc. Eur. Conf. Comput. Vision*, 2010.
- [105] J. Philbin, O. Chum, M. Isard, J. Sivic, and A. Zisserman. Object retrieval with large vocabularies and fast spatial matching. In *Proc. IEEE Conf. Comput. Vision and Pattern Recognition*, 2007.

- [106] J. Philbin, O. Chum, M. Isard, J. Sivic, and A. Zisserman. Lost in quantization: Improving particular object retrieval in large scale image databases. In *Proc. IEEE Conf. Comput. Vision and Pattern Recognition*, 2008.
- [107] A. Rahimi and B. Recht. Random features for large-scale kernel machines. In *Proc. Neural Inform. and Process. Systems*, 2007.
- [108] R. Raina, A. Battle, H. Lee, B. Packer, and A. Y. Ng. Self-taught learning: transfer learning from unlabeled data. In *ICML*, 2007.
- [109] X. Ren and J. Malik. Learning a classification model for segmentation. In *Proc. Int'l. Conf. Comput. Vision*, 2003.
- [110] S. Romberg, L. G. Pueyo, R. Lienhart, and R. van Zwol. Scalable logo recognition in real-world images. In *Proc. ACM Int'l. Conf. Multim. Retr.*, 2011.
- [111] E. Rosch, C. B. Mervis, W. D. Gray, D. M. Johnson, and P. Boyes-Braem. Basic objects in natural categories. *Cognitive psychology*, 1976.
- [112] C. Rother, V. Kolmogorov, and A. Blake. Grabcut: Interactive foreground extraction using iterated graph cuts. *ACM Trans. Graphics*, 2004.
- [113] R. Rubinfeld and D. Kroese. *The cross-entropy method: a unified approach to combinatorial optimization, Monte-Carlo simulation and machine learning*. Springer, 2004.
- [114] J. Sanchez, F. Perronnin, and Z. Akata. Fisher vectors for fine-grained visual categorization. In *Proc. IEEE Conf. Comput. Vision and Pattern Recognition*, 2011.
- [115] G. Schindler, M. Brown, and R. Szeliski. City-scale location recognition. In *Proc. IEEE Conf. Comput. Vision and Pattern Recognition*, 2007.
- [116] B. Schölkopf and A. J. Smola. *Learning with kernels : support vector machines, regularization, optimization, and beyond*. MIT Press.
- [117] M. Schultz and T. Joachims. Learning a distance metric from relative comparisons. In *Proc. Neural Inform. and Process. Systems*, 2004.
- [118] S. Shalev-Shwartz, Y. Singer, and N. Srebro. Pegasos: Primal estimated sub-gradient solver for svm. In *Proc. Int'l. Conf. Machine Learning*, 2007.
- [119] G. Sharma, F. Jurie, and C. Schmid. Discriminative spatial saliency for image classification. In *Proc. IEEE Conf. Comput. Vision and Pattern Recognition*, 2012.
- [120] M. Shi, X. Sun, D. Tao, and C. Xu. Exploiting visual word co-occurrence for image retrieval. In *Proc. ACM Int'l. Conf. Multimedia*, 2012.
- [121] J. Shotton, M. Johnson, and R. Cipolla. Semantic texton forests for image categorization and segmentation. In *Proc. IEEE Conf. Comput. Vision and Pattern Recognition*, 2008.
- [122] J. Shotton, J. Winn, C. Rother, and A. Criminisi. Textonboost for image understanding: Multi-class object recognition and segmentation by jointly modeling texture, layout, and context. *Int. J. of Comput. Vision*, 2009.
- [123] K. Simonyan, A. Vedaldi, and A. Zisserman. Deep fisher networks for large-scale image classification. In *Proc. Neural Inform. and Process. Systems*, 2013.
- [124] J. Sivic and A. Zisserman. Video Google: A text retrieval approach to object matching in videos. In *Proc. Int'l. Conf. Comput. Vision*, 2003.
- [125] A. F. Smeaton, P. Over, and W. Kraaij. Evaluation campaigns and trecvid. In *Proc. ACM SIGMM Int. Workshop on Multim. Infor. Retr.*, 2006.
- [126] M. J. Swain and D. H. Ballard. Color indexing. *Int. J. of Comput. Vision*, 1991.
- [127] W. Tang, R. Cai, Z. Li, and L. Zhang. Contextual synonym dictionary for visual object retrieval. In *Proc. ACM Int'l. Conf. Multimedia*, 2011.
- [128] R. Tao, E. Gavves, C. G. M. Snoek, and A. W. M. Smeulders. Locality in generic instance search from one example. *Proc. IEEE Conf. Comput. Vision and Pattern Recognition*, 2014.
- [129] G. Toulas, Y. Kalantidis, and Y. Avrithis. Sycmity: feature selection by symmetry for large scale image retrieval. In *Proc. ACM Int'l. Conf. Multimedia*, 2012.
- [130] P. Turcot and D. G. Lowe. Better matching with fewer features: The selection of useful features in large database recognition problems. In *Proc. Int'l. Conf. Comput. Vision*, 2009.
- [131] T. Tuytelaars, M. Fritz, K. Saenko, and T. Darrell. The nbnn kernel. In *Proc. Int'l. Conf. Comput. Vision*, 2011.
- [132] T. Tuytelaars and K. Mikolajczyk. Local invariant feature detectors: A survey. *Found. and Trends in Comp. Graph. and Vision*, 2008.

## Bibliography

- [133] T. Tuytelaars and L. Van Gool. Wide baseline stereo matching based on local, affinely invariant regions. In *Proc. Brit. Mach. Vision Conf.*, 2000.
- [134] T. Tuytelaars and L. Van Gool. Matching widely separated views based on affine invariant regions. *Int. J. of Comput. Vision*, 2004.
- [135] J. Uijlings, A. W. Smeulders, and R. Scha. The visual extent of an object. *Int. J. of Comput. Vision*, 2012.
- [136] J. R. Uijlings, K. E. A. van de Sande, T. Gevers, and A. W. Smeulders. Selective search for object recognition. *Int. J. of Comput. Vision*, 2013.
- [137] J. R. Uijlings, A. W. Smeulders, and R. J. H. Scha. What is the spatial extent of an object? In *Proc. IEEE Conf. Comput. Vision and Pattern Recognition*, 2009.
- [138] K. van de Sande, J. Uijlings, T. Gevers, and A. W. Smeulders. Segmentation as selective search for object recognition. In *Proc. Int'l. Conf. Comput. Vision*, 2011.
- [139] J. van Gemert, C. Snoek, C. Veenman, A. W. Smeulders, and J. Geusebroek. Comparing compact codebooks for visual categorization. *Comput. Vision and Image Understanding*, 2010.
- [140] J. C. van Gemert, J.-M. Geusebroek, C. J. Veenman, and A. W. Smeulders. Kernel codebooks for scene categorization. In *Proc. Eur. Conf. Comput. Vision*, 2008.
- [141] J. C. van Gemert, C. J. Veenman, A. W. Smeulders, and J.-M. Geusebroek. Visual word ambiguity. *IEEE Trans. Pattern Analysis and Machine Intell.*, 2010.
- [142] A. Vedaldi and B. Fulkerson. Vlfeat: An open and portable library of computer vision algorithms. In *Proc. ACM Int'l. Conf. Multimedia*, 2010.
- [143] A. Vedaldi, V. Gulshan, M. Varma, and A. Zisserman. Multiple kernels for object detection. In *Proc. IEEE Conf. Comput. Vision and Pattern Recognition*, 2009.
- [144] A. Vedaldi and A. Zisserman. Efficient additive kernels via explicit feature maps. *IEEE Trans. Pattern Analysis and Machine Intell.*, 2012.
- [145] A. Vezhnevets, V. Ferrari, and J. M. Buhmann. Weakly supervised structured output learning for semantic segmentation. In *Proc. IEEE Conf. Comput. Vision and Pattern Recognition*, 2012.
- [146] S. Vijayanarasimhan and K. Grauman. Efficient region search for object detection. In *Proc. IEEE Conf. Comput. Vision and Pattern Recognition*, 2011.
- [147] C. Wah, S. Branson, P. Perona, and S. Belongie. Multiclass recognition and part localization with humans in the loop. In *Proc. Int'l. Conf. Comput. Vision*, 2011.
- [148] C. Wah, S. Branson, P. Welinder, P. Perona, and S. Belongie. The caltech-ucsd birds-200-2011 dataset. Technical report, 2011.
- [149] L. Wang, L. Zhou, and C. Shen. A fast algorithm for creating a compact and discriminative visual codebook. In *Proc. Eur. Conf. Comput. Vision*, 2008.
- [150] M. Wang, F. Sha, and M. Jordan. Unsupervised kernel dimension reduction. In *Proc. Neural Inform. and Process. Systems*, 2010.
- [151] J. Weston, S. Bengio, and N. Usunier. Large scale image annotation: Learning to rank with joint word-image embeddings. In *Proc. Eur. Conf. Machine Learning*, 2010.
- [152] C. Williams and M. Seeger. The effect of the input density distribution on kernel-based classifiers. In *Proc. Int'l. Conf. Machine Learning*, 2000.
- [153] J. Winn, A. Criminisi, and T. Minka. Object categorization by learned universal visual dictionary. In *Proc. Int'l. Conf. Comput. Vision*, 2005.
- [154] Z. Wu, Q. Ke, M. Isard, and J. Sun. Bundling features for large scale partial-duplicate web image search. 2009.
- [155] L. Xie, Q. Tian, B. Yan, and S. Zhang. Hierarchical part matching for fine-grained visual categorization. In *Proc. Int'l. Conf. Comput. Vision*, 2013.
- [156] J. Yang, K. Yu, Y. Gong, and T. Huang. Linear spatial pyramid matching using sparse coding for image classification. In *Proc. IEEE Conf. Comput. Vision and Pattern Recognition*, 2009.
- [157] S. Yang, L. Bo, J. Wang, and L. Shapiro. Unsupervised template learning for fine-grained object recognition. In *Proc. Neural Inform. and Process. Systems*, 2012.
- [158] B. Yao, G. Bradski, and L. Fei-Fei. A codebook-free and annotation-free approach for fine-grained image categorization. In *Proc. IEEE Conf. Comput. Vision and Pattern Recognition*, 2012.
- [159] B. Yao, A. Khosla, and L. Fei-Fei. Combining randomization and discrimination for fine-grained image categorization. In *Proc. IEEE Conf. Comput. Vision and Pattern Recognition*, 2011.

- [160] M. Zeiler. *Hierarchical Convolutional Deep Learning in Computer Vision*. PhD thesis, 2013.
- [161] M. D. Zeiler and R. Fergus. Visualizing and understanding convolutional networks. *CoRR*, 2013.
- [162] J. Zhang, M. Marszałek, S. Lazebnik, and C. Schmid. Local features and kernels for classification of texture and object categories: A comprehensive study. *Int. J. of Comput. Vision*, 2007.
- [163] L. Zhang, C. Chen, J. Bu, Z. Chen, S. Tan, and X. He. Discriminative codeword selection for image representation. In *Proc. ACM Int'l. Conf. Multimedia*, 2010.
- [164] N. Zhang, R. Farrell, and T. Darrell. Pose pooling kernels for sub-category recognition. In *Proc. IEEE Conf. Comput. Vision and Pattern Recognition*, 2012.
- [165] N. Zhang, R. Farrell, F. Iandola, and T. Darrell. Deformable part descriptors for fine-grained recognition and attribute prediction. In *Proc. Int'l. Conf. Comput. Vision*, 2013.
- [166] S. Zhang, Q. Tian, G. Hua, Q. Huang, and S. Li. Descriptive visual words and visual phrases for image applications. In *Proc. ACM Int'l. Conf. Multimedia*, 2009.
- [167] W. Zhou, H. Li, Y. Lu, and Q. Tian. Large scale image search with geometric coding. In *Proc. ACM Int'l. Conf. Multimedia*, 2011.
- [168] W. Zhou, Y. Lu, H. Li, and Q. Tian. Scalar quantization for large scale image search. In *Proc. ACM Int'l. Conf. Multimedia*, 2012.
- [169] X. Zhou, K. Yu, T. Zhang, and T. S. Huang. Image classification using super-vector coding of local image descriptors. In *Proc. Eur. Conf. Comput. Vision*, 2010.