# Large scale visual search

Wu, S.

# Bibliography

[1] Sivic, J., Zisserman, A.: Video Google: A text retrieval approach to object matching in videos. In: Proceedings of the 9th International Conference on Computer Vision. (2003) 1470–1477

[2] Perronnin, F., Dance, C.: Fisher kernels on visual vocabularies for image categorization. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2007) 1–8

[3] Jegou, H., Perronnin, F., Douze, M., Sánchez, J., Perez, P., Schmid, C.: Aggregating local image descriptors into compact codes. IEEE Transactions on Pattern Analysis and Machine Intelligence **34** (2012) 1704–1716

[4] Arandjelovic, R., Zisserman, A.: All about VLAD. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2013) 1578–1585

[5] Delhumeau, J., Gosselin, P.H., Jégou, H., Pérez, P.: Revisiting the VLAD image representation. In: Proceedings of the 21st ACM international conference on Multimedia. (2013) 653–656

[6] Wang, Z., Di, W., Bhardwaj, A., Jagadeesh, V., Piramuthu, R.: Geometric VLAD for large scale image search. arXiv preprint arXiv:1403.3829 (2014)

[7] Jégou, H., Chum, O.: Negative evidences and co-occurences in image retrieval: The benefit of PCA and whitening. In: Proceedings of European Conference on Computer Vision. (2012) 774–787

[8] LeCun, Y., Bottou, L., Bengio, Y., Haffner, P.: Gradient-based learning applied to document recognition. Proceedings of the IEEE **86** (1998) 2278–2324

[9] Deng, J., Dong, W., Socher, R., Li, L.J., Li, K., Fei-Fei, L.: ImageNet: A large-scale hierarchical image database. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2009) 248–255

[10] Lindeberg, T.: Scale-space theory: A basic tool for analyzing structures at different scales. Journal of Applied Statistics **21** (1994) 225–270

[11] Lowe, D.G.: Object recognition from local scale-invariant features. In: Proceedings of the 7th International Conference on Computer Vision. (1999) 1150–1157

[12] Bay, H., Tuytelaars, T., Van Gool, L.: SURF: Speeded up robust features. In: Proceedings of European Conference on Computer Vision. (2006) 404–417

[13] Weickert, J., Romeny, B.T.H., Viergever, M.A.: Efficient and reliable schemes for nonlinear diffusion filtering. IEEE Transactions on Image Processing **7** (1998) 398–410

[14] Lowe, D.G.: Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision **60** (2004) 91–110

[15] Rosin, P.L.: Measuring corner properties. Computer Vision and Image Understanding **73** (1999) 291–307

[16] Leutenegger, S., Chli, M., Siegwart, R.Y.: BRISK: Binary robust invariant scalable keypoints. In: Proceedings of the International Conference on Computer Vision. (2011) 2548–2555

[17] Alahi, A., Ortiz, R., Vandergheynst, P.: FREAK: Fast retina keypoint. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2012) 510–517

[18] Wu, S., Lew, M.S.: RIFF: Retina-inspired invariant fast feature descriptor. In: Proceedings of the ACM International Conference on Multimedia. (2014) 1129–1132

[19] Chum, O., Matas, J.: Fast computation of min-hash signatures for image collections. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2012) 3077–3084

[20] Raginsky, M., Lazebnik, S.: Locality-sensitive binary codes from shift-invariant kernels. In: Advances in Neural Information Processing Systems. (2009) 1509–1517

[21] Gionis, A., Indyk, P., Motwani, R., et al.: Similarity search in high dimensions via hashing. In: International Conference on Very Large Databases. Volume 99. (1999) 518–529

[22] Chum, O., et al.: Large-scale discovery of spatially related images. IEEE Transactions on Pattern Analysis and Machine Intelligence **32** (2010) 371–377

[23] Weiss, Y., Torralba, A., Fergus, R.: Spectral hashing. In: Advances in Neural Information Processing Systems. (2009) 1753–1760

[24] Shao, J., Wu, F., Ouyang, C., Zhang, X.: Sparse spectral hashing. Pattern Recognition Letters **33** (2012) 271–277

[25] Zhang, D., Wang, J., Cai, D., Lu, J.: Self-taught hashing for fast similarity search. In: Proceedings of the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval. (2010) 18–25

[26] Liu, W., Wang, J., Kumar, S., Chang, S.F.: Hashing with graphs. In: Proceedings of the 28th International Conference on Machine Learning. (2011) 1–8

[27] Irie, G., Li, Z., Wu, X.M., Chang, S.F.: Locally linear hashing for extracting non-linear manifolds. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2014) 2115–2122

[28] Li, X., Lin, G., Shen, C., Van Den Hengel, A., Dick, A.R.: Learning hash functions using column generation. In: Proceedings of the International Conference on Machine Learning. (2013) 142–150

[29] Liu, W., Wang, J., Ji, R., Jiang, Y.G., Chang, S.F.: Supervised hashing with kernels. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2012) 2074–2081

[30] Norouzi, M., Blei, D.M.: Minimal loss hashing for compact binary codes. In: Proceedings of the 28th International Conference on Machine Learning. (2011) 353–360

[31] Huang, L.K., Yang, Q., Zheng, W.S.: Online hashing. In: Proceedings of the 23rd International Joint Conference on Artificial Intelligence. (2013) 1422–1428

[32] Norouzi, M., Fleet, D.J., Salakhutdinov, R.R.: Hamming distance metric learning. In: Advances in Neural Information Processing Systems. (2012) 1061–1069

[33] Wang, J., Kumar, S., Chang, S.F.: Sequential projection learning for hashing with compact codes. In: Proceedings of the 27th International Conference on Machine Learning. (2010) 1127–1134

[34] Wang, J., Kumar, S., Chang, S.F.: Semi-supervised hashing for large-scale search. IEEE Transactions on Pattern Analysis and Machine Intelligence **34** (2012) 2393–2406

[35] Simo-Serra, E., Trulls, E., Ferraz, L., Kokkinos, I., Fua, P., Moreno-Noguer, F.: Discriminative learning of deep convolutional feature point descriptors. In: Proceedings of the International Conference on Computer Vision. (2015) 118–126

[36] Kumar, B., Carneiro, G., Reid, I.: Learning local image descriptors with deep siamese and triplet convolutional networks by minimising global loss functions. arXiv preprint arXiv:1512.09272 (2015)

[37] Zagoruyko, S., Komodakis, N.: Learning to compare image patches via convolutional neural networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2015) 4353–4361

[38] Han, X., Leung, T., Jia, Y., Sukthankar, R., Berg, A.C.: Matchnet: Unifying feature and metric learning for patch-based matching. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2015) 3279–3286

[39] Mikolajczyk, K., Leibe, B., Schiele, B.: Multiple object class detection with a generative model. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Volume 1. (2006) 26–36

[40] Nister, D., Stewenius, H.: Scalable recognition with a vocabulary tree. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Volume 2. (2006) 2161–2168

[41] Zeiler, M.D.: Hierarchical convolutional deep learning in computer vision. PhD thesis, NEW YORK UNIVERSITY (2013)

[42] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A.: Going deeper with convolutions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2015) 1–9

[43] Oquab, M., Bottou, L., Laptev, I., Sivic, J.: Is object localization for free?-weakly-supervised learning with convolutional neural networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2015) 685–694

[44] Boureau, Y.L., Ponce, J., LeCun, Y.: A theoretical analysis of feature pooling in visual recognition. In: Proceedings of the 27th International Conference on Machine Learning. (2010) 111–118

[45] Scherer, D., Müller, A., Behnke, S.: Evaluation of pooling operations in convolutional architectures for object recognition. In: International Conference on Artificial Neural Networks. (2010) 92–101

[46] Girshick, R., Donahue, J., Darrell, T., Malik, J.: Rich feature hierarchies for accurate object detection and semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2014) 580–587

[47] Lin, M., Chen, Q., Yan, S.: Network in network. arXiv preprint arXiv:1312.4400 (2013)

[48] Krizhevsky, A., Sutskever, I., Hinton, G.E.: ImageNet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems. (2012) 1097–1105

[49] Sermanet, P., Eigen, D., Zhang, X., Mathieu, M., Fergus, R., LeCun, Y.: Overfeat: Integrated recognition, localization and detection using convolutional networks. arXiv preprint arXiv:1312.6229 (2013)

[50] He, K., Zhang, X., Ren, S., Sun, J.: Spatial pyramid pooling in deep convolutional networks for visual recognition. In: Proceedings of European Conference on Computer Vision. (2014) 346–361

[51] Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014)

[52] Uijlings, J.R., van de Sande, K.E., Gevers, T., Smeulders, A.W.: Selective search for object recognition. International Journal of Computer Vision **104** (2013) 154–171

[53] Hariharan, B., Arbeláez, P., Girshick, R., Malik, J.: Simultaneous detection and segmentation. In: Proceedings of European Conference on Computer Vision. (2014) 297–312

[54] Zhu, Y., Urtasun, R., Salakhutdinov, R., Fidler, S.: segdeepm: Exploiting segmentation and context in deep neural networks for object detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2015) 4703–4711

[55] Zhang, Y., Sohn, K., Villegas, R., Pan, G., Lee, H.: Improving object detection with deep convolutional networks via bayesian optimization and

structured prediction. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2015) 249–258

[56] Girshick, R.: Fast R-CNN. In: Proceedings of the IEEE International Conference on Computer Vision. (2015) 1440–1448

[57] Ren, S., He, K., Girshick, R., Sun, J.: Faster R-CNN: Towards real-time object detection with region proposal networks. In: Advances in Neural Information Processing Systems. (2015) 91–99

[58] Gong, Y., Wang, L., Guo, R., Lazebnik, S.: Multi-scale orderless pooling of deep convolutional activation features. In: Proceedings of European Conference on Computer Vision. (2014) 392–407

[59] Liu, Y., Guo, Y., Wu, S., Lew, M.S.: Deepindex for accurate and efficient image retrieval. In: Proceedings of the 5th ACM on International Conference on Multimedia Retrieval. (2015) 43–50

[60] Sharif Razavian, A., Azizpour, H., Sullivan, J., Carlsson, S.: Cnn features off-the-shelf: an astounding baseline for recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. (2014) 806–813

[61] Wan, J., Wang, D., Hoi, S.C.H., Wu, P., Zhu, J., Zhang, Y., Li, J.: Deep learning for content-based image retrieval: A comprehensive study. In: Proceedings of the 22nd ACM International Conference on Multimedia, ACM (2014) 157–166

[62] Sun, S., Zhou, W., Li, H., Tian, Q.: Search by detection: Object-level feature for image retrieval. In: Proceedings of International Conference on Internet Multimedia Computing and Service. (2014) 46

[63] Babenko, A., Slesarev, A., Chigorin, A., Lempitsky, V.: Neural codes for image retrieval. In: Proceedings of European Conference on Computer Vision. (2014) 584–599

[64] Razavian, A.S., Sullivan, J., Maki, A., Carlsson, S.: A baseline for visual instance retrieval with deep convolutional networks. arXiv preprint arXiv:1412.6574 (2014)

[65] Babenko, A., Lempitsky, V.: Aggregating local deep features for image retrieval. In: Proceedings of the IEEE International Conference on Computer Vision. (2015) 1269–1277

[66] Zheng, S., Jayasumana, S., Romera-Paredes, B., Vineet, V., Su, Z., Du, D., Huang, C., Torr, P.H.: Conditional random fields as recurrent neural networks. In: Proceedings of the IEEE International Conference on Computer Vision. (2015) 1529–1537

[67] Liu, Z., Li, X., Luo, P., Loy, C.C., Tang, X.: Semantic image segmentation via deep parsing network. In: Proceedings of the IEEE International Conference on Computer Vision. (2015) 1377–1385

[68] Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2015) 3431–3440

[69] Chen, L.C., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A.L.: Semantic image segmentation with deep convolutional nets and fully connected crfs. arXiv preprint arXiv:1412.7062 (2014)

[70] Lin, G., Shen, C., Reid, I., et al.: Efficient piecewise training of deep structured models for semantic segmentation. arXiv preprint arXiv:1504.01013 (2015)

[71] Toshev, A., Szegedy, C.: Deeppose: Human pose estimation via deep neural networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2014) 1653–1660

[72] Li, H., Lin, Z., Shen, X., Brandt, J., Hua, G.: A convolutional neural network cascade for face detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2015) 5325–5334

[73] Chen, X., Yuille, A.L.: Articulated pose estimation by a graphical model with image dependent pairwise relations. In: Advances in Neural Information Processing Systems. (2014) 1736–1744

[74] Tompson, J.J., Jain, A., LeCun, Y., Bregler, C.: Joint training of a convolutional network and a graphical model for human pose estimation. In: Advances in Neural Information Processing Systems. (2014) 1799–1807

[75] Tompson, J., Goroshin, R., Jain, A., LeCun, Y., Bregler, C.: Efficient object localization using convolutional networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2015) 648–656

[76] Fan, X., Zheng, K., Lin, Y., Wang, S.: Combining local appearance and holistic view: Dual-source deep neural networks for human pose estimation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2015) 1347–1355

[77] Lazebnik, S., Schmid, C., Ponce, J.: Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Volume 2. (2006) 2169–2178

[78] Ramesh, B., Xiang, C., Lee, T.H.: Shape classification using invariant features and contextual information in the bag-of-words model. Pattern Recognition **48** (2015) 894–906

[79] Rothganger, F., Lazebnik, S., Schmid, C., Ponce, J.: 3D object modeling and recognition using local affine-invariant image descriptors and multi-view spatial constraints. International Journal of Computer Vision **66** (2006) 231–259

[80] Yang, H., Shao, L., Zheng, F., Wang, L., Song, Z.: Recent advances and trends in visual tracking: A review. Neurocomputing **74** (2011) 3823–3831

[81] Gauglitz, S., Höllerer, T., Turk, M.: Evaluation of interest point detectors and feature descriptors for visual tracking. International Journal of Computer Vision **94** (2011) 335–360

[82] Thomee, B., Lew, M.S.: Interactive search in image retrieval: a survey. International Journal of Multimedia Information Retrieval **1** (2012) 71–86

[83] Jiang, Y.G., Bhattacharya, S., Chang, S.F., Shah, M.: High-level event recognition in unconstrained videos. International Journal of Multimedia Information Retrieval **2** (2013) 73–101

[84] Guo, Y., Bennamoun, M., Sohel, F., Lu, M., Wan, J., Kwok, N.M.: A comprehensive performance evaluation of 3D local feature descriptors. International Journal of Computer Vision (2015) 1–24

[85] Buoncompagni, S., Maio, D., Maltoni, D., Papi, S.: Saliency-based keypoint selection for fast object detection and matching. Pattern Recognition Letters (2015)

[86] Lin, W.C., Tsai, C.F., Chen, Z.Y., Ke, S.W.: Keypoint selection for efficient bag-of-words feature generation and effective image classification. Information Sciences **329** (2016) 33–51

[87] Wu, S., Lew, M.S.: Evaluation of salient point methods. In: Proceedings of the 21st ACM International Conference on Multimedia. (2013) 685–688

[88] Wu, S., Lew, M.S.: Salient features for visual word based image copy detection. In: Proceedings of International Conference on Multimedia Retrieval. (2014) 475–478

[89] Fischler, M.A., Bolles, R.C.: Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Communications of the ACM **24** (1981) 381–395

[90] Yu, G., Morel, J.M.: A fully affine invariant image comparison method. In: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing. (2009) 1597–1600

[91] Moravec, H.P.: Towards automatic visual obstacle avoidance. In: Proceedings of the 5th International Joint Conference on Artificial Intelligence. (1977) 584

[92] Harris, C., Stephens, M.: A combined corner and edge detector. In: Alvey Vision Conference. Volume 15. (1988) 50

[93] Beaudet, P.R.: Rotationally invariant image operators. In: International Joint Conference on Pattern Recognition. Volume 579. (1978) 583

[94] Mikolajczyk, K., Schmid, C.: Indexing based on scale invariant interest points. In: Proceedings of the 8th International Conference on Computer Vision. Volume 1. (2001) 525–531

[95] Mikolajczyk, K., Schmid, C.: Scale & affine invariant interest point detectors. International Journal of Computer Vision **60** (2004) 63–86

[96] Rosten, E., Drummond, T.: Machine learning for high-speed corner detection. In: Proceedings of European Conference on Computer Vision. (2006) 430–443

[97] Mikolajczyk, K., Schmid, C.: An affine invariant interest point detector. In: Proceedings of European Conference on Computer Vision. (2002) 128–142

[98] Matas, J., Chum, O., Urban, M., Pajdla, T.: Robust wide-baseline stereo from maximally stable extremal regions. Image and Vision Computing **22** (2004) 761–767

[99] Tuytelaars, T., Van Gool, L.J.: Content-based image retrieval based on local affinely invariant regions. In: Visual Information and Information Systems. (1999) 493–500

[100] Tuytelaars, T., Van Gool, L.J.: Wide baseline stereo matching based on local, affinely invariant regions. In: Proceedings of the British Machine Vision Conference. (2000) 412–422

[101] Ke, Y., Sukthankar, R.: PCA-SIFT: A more distinctive representation for local image descriptors. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Volume 2. (2004) 506–513

[102] Van De Sande, K.E., Gevers, T., Snoek, C.G.: Evaluating color descriptors for object and scene recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence **32** (2010) 1582–1596

[103] Li, B., Xiao, R., Li, Z., Cai, R., Lu, B.L., Zhang, L.: Rank-SIFT: Learning to rank repeatable local interest points. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2011) 1737–1744

[104] Arandjelović, R., Zisserman, A.: Three things everyone should know to improve object retrieval. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2012) 2911–2918

[105] Dong, J., Soatto, S.: Domain-size pooling in local descriptors: DSP-SIFT. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2015) 5097–5106

[106] Alcantarilla, P.F., Bartoli, A., Davison, A.J.: KAZE features. In: Proceedings of European Conference on Computer Vision. (2012) 214–227

[107] Schmid, C., Mohr, R., Bauckhage, C.: Evaluation of interest point detectors. International Journal of Computer Vision **37** (2000) 151–172

[108] Mikolajczyk, K., Schmid, C.: A performance evaluation of local descriptors. IEEE Transactions on Pattern Analysis and Machine Intelligence **27** (2005) 1615–1630

[109] Miksik, O., Mikolajczyk, K.: Evaluation of local detectors and descriptors for fast feature matching. In: Proceedings of the 21st International Conference on Pattern Recognition. (2012) 2681–2684

[110] Heinly, J., Dunn, E., Frahm, J.M.: Comparative evaluation of binary features. In: Proceedings of European Conference on Computer Vision. (2012) 759–773

[111] Figat, J., Kornuta, T., Kasprzak, W.: Performance evaluation of binary descriptors of local features. In: Computer Vision and Graphics. (2014) 187–194

[112] Moreels, P., Perona, P.: Evaluation of features detectors and descriptors based on 3d objects. International Journal of Computer Vision **73** (2007) 263–284

[113] Mukherjee, D., Wu, Q.J., Wang, G.: A comparative experimental study of image feature detectors and descriptors. Machine Vision and Applications **26** (2015) 443–466

[114] Perd'och, M., Chum, O., Matas, J.: Efficient representation of local geometry for large scale object retrieval. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2009) 9–16

[115] Agrawal, M., Konolige, K., Blas, M.R.: Censure: Center surround extremas for realtime feature detection and matching. In: Proceedings of European Conference on Computer Vision. (2008) 102–115

[116] Shi, J., Tomasi, C.: Good features to track. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (1994) 593–600

[117] Calonder, M., Lepetit, V., Strecha, C., Fua, P.: BRIEF: Binary robust independent elementary features. In: Proceedings of European Conference on Computer Vision. (2010) 778–792

[118] Rublee, E., Rabaud, V., Konolige, K., Bradski, G.: ORB: an efficient alternative to SIFT or SURF. In: Proceedings of the IEEE International Conference on Computer Vision. (2011) 2564–2571

[119] Mair, E., Hager, G.D., Burschka, D., Suppa, M., Hirzinger, G.: Adaptive and generic corner detection based on the accelerated segment test. In: Proceedings of European Conference on Computer Vision. (2010) 183–196

[120] Hartley, R., Zisserman, A.: Multiple view geometry in computer vision. Robotica **23** (2005) 271–271

[121] Fischer, P., Dosovitskiy, A., Brox, T.: Descriptor matching with convolutional neural networks: a comparison to SIFT. arXiv preprint arXiv:1405.5769 (2014)

[122] Philbin, J., Chum, O., Isard, M., Sivic, J., Zisserman, A.: Object retrieval with large vocabularies and fast spatial matching. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2007) 1–8

[123] Trzcinski, T., Christoudias, M., Fua, P., Lepetit, V.: Boosting binary keypoint descriptors. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2013) 2874–2881

[124] Levi, G., Hassner, T.: LATCH: learned arrangements of three patch codes. In: Proceedings of the IEEE Winter Conference on Applications of Computer Vision. (2016) 1–9

[125] Tola, E., Lepetit, V., Fua, P.: Daisy: An efficient dense descriptor applied to wide-baseline stereo. IEEE Transactions on Pattern Analysis and Machine Intelligence **32** (2010) 815–830

[126] Fei-Fei, L., Perona, P.: A bayesian hierarchical model for learning natural scene categories. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Volume 2. (2005) 524–531

[127] Jégou, H., Douze, M., Schmid, C., Pérez, P.: Aggregating local descriptors into a compact image representation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2010) 3304–3311

[128] Sánchez, J., Perronnin, F., Mensink, T., Verbeek, J.: Image classification with the fisher vector: Theory and practice. International Journal of Computer Vision **105** (2013) 222–245

[129] Thomee, B., Bakker, E.M., Lew, M.S.: TOP-SURF: a visual words toolkit. In: Proceedings of the International Conference on Multimedia. (2010) 1473–1476

[130] Jégou, H., Douze, M., Schmid, C.: Improving bag-of-features for large scale image search. International Journal of Computer Vision **87** (2010) 316–336

[131] Grana, C., Borghesani, D., Manfredi, M., Cucchiara, R.: A fast approach for integrating orb descriptors in the bag of words model. In: IS&T/SPIE Electronic Imaging. Volume 8667. (2013) 866709–1–866709–8

[132] O'Hara, S., Draper, B., et al.: Are you using the right approximate nearest neighbor algorithm? In: IEEE Workshop on Applications of Computer Vision. (2013) 9–14

[133] Salton, G., McGill, M.J.: Introduction to modern information retrieval. London: Library Association Publishing (1986)

[134] Everingham, M., Van Gool, L., Williams, C.K., Winn, J., Zisserman, A.: The pascal visual object classes (voc) challenge. International Journal of Computer Vision **88** (2010) 303–338

[135] Griffin, G., Holub, A., Perona, P.: Caltech-256 object category dataset. Technical Report 7694, California Institute of Technology (2007)

[136] Huiskes, M.J., Thomee, B., Lew, M.S.: New trends and ideas in visual concept detection: the MIR flickr retrieval evaluation initiative. In: Proceedings of the International Conference on Multimedia Information Retrieval. (2010) 527–536

[137] Gong, Y., Lazebnik, S.: Iterative quantization: A procrustean approach to learning binary codes. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2011) 817–824

[138] Andoni, A., Indyk, P.: Near-optimal hashing algorithms for approximate nearest neighbor in high dimensions. In: Proceedings of the 47th Annual IEEE Symposium on Foundations of Computer Science. (2006) 459–468

[139] Heo, J.P., Lee, Y., He, J., Chang, S.F., Yoon, S.E.: Spherical hashing. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2012) 2957–2964

[140] Jin, Z., Li, C., Lin, Y., Cai, D.: Density sensitive hashing. IEEE Transactions on Cybernetics, **44** (2014) 1362–1371

[141] Lai, H., Pan, Y., Liu, Y., Yan, S.: Simultaneous feature learning and hash coding with deep neural networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2015) 3270–3278

[142] Zhao, F., Huang, Y., Wang, L., Tan, T.: Deep semantic ranking based hashing for multi-label image retrieval. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2015) 1556–1564

[143] Oquab, M., Bottou, L., Laptev, I., Sivic, J.: Learning and transferring mid-level image representations using convolutional neural networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2014) 1717–1724

[144] Donahue, J., Jia, Y., Vinyals, O., Hoffman, J., Zhang, N., Tzeng, E., Darrell, T.: Decaf: A deep convolutional activation feature for generic visual recognition. arXiv preprint arXiv:1310.1531 (2013)

[145] Sicre, R., Jurie, F.: Discriminative part model for visual recognition. Computer Vision and Image Understanding **141** (2015) 28–37

[146] Iandola, F., Moskewicz, M., Karayev, S., Girshick, R., Darrell, T., Keutzer, K.: Densenet: Implementing efficient convnet descriptor pyramids. arXiv preprint arXiv:1404.1869 (2014)

[147] Azizpour, H., Razavian, A., Sullivan, J., Maki, A., Carlsson, S.: From generic to specific deep representations for visual recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. (2015) 36–45

[148] Zheng, L., Wang, S., Tian, L., He, F., Liu, Z., Tian, Q.: Query-adaptive late fusion for image search and person re-identification. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2015) 1741–1750

[149] Ng, J., Yang, F., Davis, L.: Exploiting local features from deep networks for image retrieval. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. (2015) 53–61

[150] Tolias, G., Sicre, R., Jégou, H.: Particular object retrieval with integral max-pooling of cnn activations. arXiv preprint arXiv:1511.05879 (2015)

[151] Nandakumar, K., Chen, Y., Dass, S.C., Jain, A.K.: Likelihood ratio-based biometric score fusion. IEEE Transactions on Pattern Analysis and Machine Intelligence **30** (2008) 342–347

[152] Zhang, S., Yang, M., Cour, T., Yu, K., Metaxas, D.N.: Query specific fusion for image retrieval. In: Proceedings of European Conference on Computer Vision. (2012) 660–673

[153] Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama, S., Darrell, T.: Caffe: Convolutional architecture for fast feature embedding. In: Proceedings of the ACM International Conference on Multimedia. (2014) 675–678

[154] Jégou, H., Douze, M., Schmid, C.: Hamming embedding and weak geometric consistency for large scale image search. In: Proceedings of European Conference on Computer Vision. (2008) 304–317

[155] Oliva, A., Torralba, A.: Modeling the shape of the scene: A holistic representation of the spatial envelope. International Journal of Computer Vision **42** (2001) 145–175

[156] Zheng, L., Wang, S., Liu, Z., Tian, Q.: Packing and padding: Coupled multi-index for accurate image retrieval. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2014) 1939–1946

[157] Spyromitros-Xioufis, E., Papadopoulos, S., Kompatsiaris, I.Y., Tsoumakas, G., Vlahavas, I.: A comprehensive study over vlad and product quantization in large-scale image retrieval. IEEE Transactions on Multimedia **16** (2014) 1713–1728

[158] Jégou, H., Zisserman, A.: Triangulation embedding and democratic aggregation for image search. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2014) 3310–3317

[159] Csurka, G., Dance, C., Fan, L., Willamowski, J., Bray, C.: Visual categorization with bags of keypoints. In: Proceedings of European Conference on Computer Vision Workshop on Statistical Learning in Computer Vision. Volume 1. (2004) 1–2

[160] Boser, B.E., Guyon, I.M., Vapnik, V.N.: A training algorithm for optimal margin classifiers. In: Proceedings of the 5th Annual Workshop on Computational Learning Theory. (1992) 144–152

[161] Lin, Y., Lv, F., Zhu, S., Yang, M., Cour, T., Yu, K., Cao, L., Huang, T.: Large-scale image classification: fast feature extraction and svm training. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2011) 1689–1696

[162] Perronnin, F., Sánchez, J., Mensink, T.: Improving the fisher kernel for large-scale image classification. In: Proceedings of European Conference on Computer Vision. (2010) 143–156

[163] Jaakkola, T.S., Haussler, D., et al.: Exploiting generative models in discriminative classifiers. Advances in Neural Information Processing Systems (1999) 487–493

[164] Krizhevsky, A., Hinton, G.: Learning multiple layers of features from tiny images. Master's thesis, Department of Computer Science, University of Toronto (2009)

[165] He, K., Zhang, X., Ren, S., Sun, J.: Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In: Proceedings of the IEEE International Conference on Computer Vision. (2015) 1026–1034

[166] Zeiler, M.D., Fergus, R.: Visualizing and understanding convolutional networks. In: Proceedings of European Conference on Computer Vision. (2014) 818–833

[167] Lee, C.Y., Xie, S., Gallagher, P., Zhang, Z., Tu, Z.: Deeply-supervised nets. In: Proceedings of the 18th International Conference on Artificial Intelligence and Statistics. Volume 2. (2015) 6

[168] Bromley, J., Bentz, J.W., Bottou, L., Guyon, I., LeCun, Y., Moore, C., Säckinger, E., Shah, R.: Signature verification using a "siamese" time delay neural network. International Journal of Pattern Recognition and Artificial Intelligence **7** (1993) 669–688

[169] Hoffer, E., Ailon, N.: Deep metric learning using triplet network. In: International Workshop on Similarity-Based Pattern Recognition. (2015) 84–92

[170] Sun, Y., Wang, X., Tang, X.: Deep convolutional network cascade for facial point detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2013) 3476–3483

[171] Yang, S., Ramanan, D.: Multi-scale recognition with dag-cnns. In: Proceedings of the IEEE International Conference on Computer Vision. (2015) 1215–1223

[172] Ma, L., Lu, J., Feng, J., Zhou, J.: Multiple feature fusion via weighted entropy for visual tracking. In: Proceedings of the IEEE International Conference on Computer Vision. (2015) 3128–3136

[173] Netzer, Y., Wang, T., Coates, A., Bissacco, A., Wu, B., Ng, A.Y.: Reading digits in natural images with unsupervised feature learning. In: Advances in Neural Information Processing Systems Workshop on Deep Learning and Unsupervised Feature Learning. (2011)

[174] Ouyang, W., Luo, P., Zeng, X., Qiu, S., Tian, Y., Li, H., Yang, S., Wang, Z., Xiong, Y., Qian, C., et al.: Deepid-net: multi-stage and deformable deep convolutional neural networks for object detection. arXiv preprint arXiv:1409.3505 (2014)