

# List of Publications

- [1] Datta, Deepanwita, Himanshu Mittal, and Sanjay K. Singh. “Seeing the bigger picture: A novel statistical approach for enhancing image annotation by employing community detection.” *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Banff Center, Banff, Canada, October 5-8, 2017*. Oct 2017.
- [2] Datta, Deepanwita, Prasad Suresh Nakhate, Sanjay K. Singh, and C. Ravindranath Chowdary. “Why stop there?: A novel Hill Climbing based approach towards multimodal classification.” *14th ACS/IEEE International Conference on Computer Systems and Applications (AICCSA) 2017, Hammamet, Tunisia, October 30-November 3, 2017*. Nov 2017.
- [3] Datta, Deepanwita, Sanjay Kumar Singh, and C. Ravindranath Chowdary. “Bridging the gap: effect of text query reformulation in multimodal retrieval.” *Multimedia Tools Appl.* 76 (2017): 22871–22888.
- [4] Datta, Deepanwita, Shubham Varma, C. Ravindranath Chowdary, and Sanjay Kumar Singh. “Multimodal Retrieval using Mutual Information based Textual

Query Reformulation.” *Expert Syst. Appl.* 68 (2017): 81–92.

# Bibliography

- [1] Agichtein, E., Brill, E., and Dumais, S. (2006). Improving web search ranking by incorporating user behavior information. In *Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval*, SIGIR '06, pages 19–26, New York, NY, USA. ACM.
- [2] Aizawa, A. (2003). An information-theoretic perspective of tf-idf measures. *Information Processing & Management*, 39(1):45 – 65.
- [3] Aslandogan, Y. A. and Yu, C. T. (1999). Techniques and systems for image and video retrieval. *IEEE Transactions on Knowledge and Data Engineering*, 11(1):56–63.
- [4] Atrey, P. K., Hossain, M. A., El Saddik, A., and Kankanhalli, M. S. (2010a). Multimodal fusion for multimedia analysis: a survey. *Multimedia Systems*, 16(6):345–379.
- [5] Atrey, P. K., Hossain, M. A., El Saddik, A., and Kankanhalli, M. S. (2010b). Multimodal fusion for multimedia analysis: a survey. *Multimedia systems*, 16(6):345–379.

- 
- [6] Baeza-Yates, R., Ribeiro-Neto, B., et al. (1999). *Modern information retrieval*, volume 463. ACM press New York.
- [7] Bay, H., Ess, A., Tuytelaars, T., and Van Gool, L. (2008). Speeded-up robust features (SURF). *Computer vision and image understanding*, 110(3):346–359.
- [8] Belkin, N., Cool, C., Kelly, D., Kim, G., Kim, J., Lee, H., Muresan, G., Tang, M., and Yuan, X. (2003). Query Length in Interactive Information Retrieval. *SIGIR Forum (ACM Special Interest Group on Information Retrieval)*, (SPEC. ISS.):205–212.
- [9] Bendersky, M. and Croft, W. B. (2008). Discovering key concepts in verbose queries. In *Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval*, pages 491–498. ACM.
- [10] Bohlin, L., Edler, D., Lancichinetti, A., and Rosvall, M. (2014). *Community Detection and Visualization of Networks with the Map Equation Framework*, pages 3–34. Springer International Publishing, Cham.
- [11] Buffoni, D., Tollari, S., and Gallinari, P. (2012). A Learning to Rank framework applied to text-image retrieval. *Multimedia Tools and Applications*, 60(1):161–180.
- [12] Caicedo, J. C., Moreno, J. G., Niño, E. A., and González, F. A. (2010). Combining visual features and text data for medical image retrieval using latent semantic kernels. In *Proceedings of the international conference on Multimedia information retrieval*, pages 359–366. ACM.

- 
- [13] Carpineto, C. and Romano, G. (2012). A Survey of Automatic Query Expansion in Information Retrieval. *ACM Comput. Surv.*, 44(1):1:1–1:50.
- [14] Carroll, S. and Swain, M. (1993). Explicit and implicit negative feedback. *Studies in second language acquisition*, 15(3):357–386.
- [15] Cheng, Y.-C. and Chen, S.-Y. (2003). Image classification using color, texture and regions. *Image and Vision Computing*, 21(9):759 – 776.
- [16] Church, K. W. and Hanks, P. (1990). Word association norms, mutual information, and lexicography. *Comput. Linguist.*, 16(1):22–29.
- [17] Clinchant, S., Ah-Pine, J., and Csurka, G. (2011). Semantic combination of textual and visual information in multimedia retrieval. In *Proceedings of the 1st acm international conference on multimedia retrieval*, page 44. ACM.
- [18] Das, N., Reddy, J. M., Sarkar, R., Basu, S., Kundu, M., Nasipuri, M., and Basu, D. K. (2012). A statistical–topological feature combination for recognition of handwritten numerals. *Applied Soft Computing*, 12(8):2486 – 2495.
- [19] Datta, D., Varma, S., C., R. C., and Singh, S. K. (2017a). Multimodal retrieval using mutual information based textual query reformulation. *Expert Systems with Applications*, 68(Supplement C):81 – 92.
- [20] Datta, D., Varma, S., Chowdary, C. R., and Singh, S. K. (2017b). Multimodal Retrieval using Mutual Information based Textual Query Reformulation. *Expert Systems with Applications*, 68:81 – 92.

- 
- [21] Datta, R., Joshi, D., Li, J., and Wang, J. Z. (2008). Image retrieval: Ideas, influences, and trends of the new age. *ACM Computing Surveys (CSUR)*, 40(2):5.
- [22] de Boer, M., Schutte, K., and Kraaij, W. (2015). Knowledge based query expansion in complex multimedia event detection. *Multimedia Tools and Applications*, pages 1–19.
- [23] Del Bimbo, A. (1999). *Visual Information Retrieval*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.
- [24] Depeursinge, A. and Müller, H. (2010). Fusion techniques for combining textual and visual information retrieval. In *ImageCLEF*, pages 95–114. Springer.
- [25] Deselaers, T., Keysers, D., and Ney, H. (2008). Features for image retrieval: an experimental comparison. *Information Retrieval*, 11(2):77–107.
- [26] Diamant, I., Klang, E., Amitai, M., Konen, E., Goldberger, J., and Greenspan, H. (2016). TASK DRIVEN DICTIONARY LEARNING BASED ON MUTUAL INFORMATION FOR MEDICAL IMAGE CLASSIFICATION. *IEEE Transactions on Biomedical Engineering*.
- [27] Faloutsos, C. and Lin, K.-I. (1995). FastMap: A Fast Algorithm for Indexing, Data-mining and Visualization of Traditional and Multimedia Datasets. In *Proceedings of the 1995 ACM SIGMOD International Conference on Management of Data*, SIGMOD '95, pages 163–174, New York, NY, USA. ACM.

- 
- [28] Faro, A., Giordano, D., Pino, C., and Spampinato, C. (2010). Visual attention for implicit relevance feedback in a content based image retrieval. In *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications*, pages 73–76. ACM.
- [29] Feng, L. and Bhanu, B. (2016a). Semantic Concept Co-Occurrence Patterns for Image Annotation and Retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(4):785–799.
- [30] Feng, L. and Bhanu, B. (2016b). Semantic concept co-occurrence patterns for image annotation and retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(4):785–799.
- [31] Feng, S. L., Manmatha, R., and Lavrenko, V. (2004). Multiple bernoulli relevance models for image and video annotation. In *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004.*, volume 2, pages II–1002–II–1009 Vol.2.
- [32] Forman, G. (2003). An extensive empirical study of feature selection metrics for text classification. *Journal of machine learning research*, 3(Mar):1289–1305.
- [33] Gehler, P. and Nowozin, S. (2009). On feature combination for multiclass object classification. In *2009 IEEE 12th International Conference on Computer Vision*, pages 221–228.
- [34] Gilbert, A., Piras, L., Wang, J., Yan, F., Dellandrea, E., Gaizauskas, R., Villegas, M., and Mikolajczyk, K. (2015). Overview of the ImageCLEF 2015 Scalable

- Image Annotation, Localization and Sentence Generation task. In *CLEF2015 Working Notes*, CEUR Workshop Proceedings, Toulouse, France. CEUR-WS.org.
- [35] González, E., Bianconi, F., and Fernández, A. (2016). An investigation on the use of local multi-resolution patterns for image classification. *Information Sciences*, 361–362:1 – 13.
- [36] Grineva, M., Grinev, M., and Lizorkin, D. (2009). Extracting key terms from noisy and multitheme documents. In *Proceedings of the 18th international conference on World wide web*, pages 661–670. ACM.
- [37] Gudivada, V. N. and Raghavan, V. V. (1995). Content based image retrieval systems. *Computer*, 28(9):18–22.
- [38] Guillaumin, M., Verbeek, J., and Schmid, C. (2010). Multimodal semi-supervised learning for image classification. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 902–909.
- [39] Hasan, K. S. and Ng, V. (2014). Automatic Keyphrase Extraction: A Survey of the State of the Art. In *ACL (1)*, pages 1262–1273.
- [40] Hearst, M. A. (2009). *Search User Interfaces*. Cambridge University Press, New York, NY, USA, 1st edition.
- [41] Hochmair, H. H. and Fu, Z. J. (2006). User Interface Design for Semantic Query Expansion in Geodata Repositories. *Angewandte Geoinformatik*, 18(18):(pp. 236–245).

- 
- [42] Hoi, S. C. and Wu, P. (2011). SIRE: A Social Image Retrieval Engine. In *Proceedings of the 19th ACM International Conference on Multimedia, MM '11*, pages 817–818, New York, NY, USA. ACM.
- [43] Hou, J., Gao, H., Xia, Q., and Qi, N. (2016). Feature combination and the knn framework in object classification. *IEEE Transactions on Neural Networks and Learning Systems*, 27(6):1368–1378.
- [44] Huang, J., Liu, Z., Wang, Y., Chen, Y., and Wong, E. K. (1999). Integration of multimodal features for video scene classification based on hmm. In *1999 IEEE Third Workshop on Multimedia Signal Processing (Cat. No.99TH8451)*, pages 53–58.
- [45] Jeon, J., Lavrenko, V., and Manmatha, R. (2003). Automatic Image Annotation and Retrieval Using Cross-media Relevance Models. In *Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Informaion Retrieval, SIGIR '03*, pages 119–126, New York, NY, USA. ACM.
- [46] Ji, P., Gao, X., and Hu, X. (2017). Automatic image annotation by combining generative and discriminant models. *Neurocomputing*, 236:48 – 55. Good Practices in Multimedia Modeling.
- [47] Joint, M., Moellic, P.-A., Hede, P., and Adam, P. (2004). PIRIA: a general tool for indexing, search, and retrieval of multimedia content. In *Electronic Imaging 2004*, pages 116–125. International Society for Optics and Photonics.

- 
- [48] Kiliç, D. and Alpkocak, A. (2011). An expansion and reranking approach for annotation-based image retrieval from web. *Expert Systems with Applications*, 38(10):13121 – 13127.
- [49] Lahat, D., Adali, T., and Jutten, C. (2015). Multimodal data fusion: an overview of methods, challenges, and prospects. *Proceedings of the IEEE*, 103(9):1449–1477.
- [50] Lancaster, F. W. (1972). Vocabulary control for information retrieval.
- [51] Lavrenko, V., Manmatha, R., and Jeon, J. (2003). A model for learning the semantics of pictures. In *in NIPS*. MIT Press.
- [52] Lew, M. S., Sebe, N., Djeraba, C., and Jain, R. (2006). Content-based Multimedia Information Retrieval: State of the Art and Challenges. *ACM Trans. Multimedia Comput. Commun. Appl.*, 2(1):1–19.
- [53] Li, X., Uricchio, T., Ballan, L., Bertini, M., Snoek, C. G. M., and Bimbo, A. D. (2016). Socializing the Semantic Gap: A Comparative Survey on Image Tag Assignment, Refinement, and Retrieval. *ACM Comput. Surv.*, 49(1):14:1–14:39.
- [54] Li, Z., Zhang, B., and Wang, Y. (2011). Document dependent fusion in multimodal music retrieval. In *Proceedings of the 19th ACM international conference on Multimedia*, pages 1105–1108. ACM.

- 
- [55] Lienhart, R., Romberg, S., and Hörsler, E. (2009). Multilayer pLSA for multimodal image retrieval. In *Proceedings of the ACM International Conference on Image and Video Retrieval*, page 9. ACM.
- [56] Lisin, D. A., Mattar, M. A., Blaschko, M. B., Learned-Miller, E. G., and Benfield, M. C. (2005). Combining local and global image features for object class recognition. In *Computer Vision and Pattern Recognition-Workshops, 2005. CVPR Workshops. IEEE Computer Society Conference on*, pages 47–47. IEEE.
- [57] Liu, Z., Huang, W., Zheng, Y., and Sun, M. (2010). Automatic keyphrase extraction via topic decomposition. In *Proceedings of the 2010 conference on empirical methods in natural language processing*, pages 366–376. Association for Computational Linguistics.
- [58] Magalhães, J. and Rüger, S. (2010). An information-theoretic framework for semantic-multimedia retrieval. *ACM Transactions on Information Systems (TOIS)*, 28(4):19.
- [59] Manning, C. D., Raghavan, P., Schütze, H., et al. (2008a). *Introduction to Information Retrieval*, volume 1. Cambridge university press Cambridge.
- [60] Manning, C. D., Raghavan, P., and Schütze, H. (2008b). Scoring, term weighting, and the vector space model. In *Introduction to Information Retrieval*, pages 100–123. Cambridge University Press. Cambridge Books Online.
- [61] Manning, D. (1995). Introduction. In *Introduction to Industrial Minerals*, pages 1–16. Springer.

- 
- [62] Mihalcea, R. and Tarau, P. (2004). TextRank: Bringing Order into Texts. In Lin, D. and Wu, D., editors, *Proceedings of EMNLP 2004*, pages 404–411, Barcelona, Spain. Association for Computational Linguistics.
- [63] Mitra, M., Singhal, A., and Buckley, C. (1998). Improving automatic query expansion. In *Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval*, pages 206–214. ACM.
- [64] Moulin, C., Largeton, C., Ducottet, C., Géry, M., and Barat, C. (2014). Fisher linear discriminant analysis for text-image combination in multimedia information retrieval. *Pattern Recognition*, 47(1):260–269.
- [65] Myoupo, D., Popescu, A., Le Borgne, H., and Moëllic, P.-A. (2010). *Multi-modal Image Retrieval over a Large Database*, pages 177–184. Springer Berlin Heidelberg, Berlin, Heidelberg.
- [66] Nguyen, C.-T., Zhan, D.-C., and Zhou, Z.-H. (2013). Multi-Modal Image Annotation with Multi-Instance Multi-Label LDA.
- [67] Oh, S., McCloskey, S., Kim, I., Vahdat, A., Cannons, K. J., Hajimirsadeghi, H., Mori, G., Perera, A. G. A., Pandey, M., and Corso, J. J. (2014). Multimedia event detection with multimodal feature fusion and temporal concept localization. *Machine Vision and Applications*, 25(1):49–69.
- [68] Ojala, T., Pietikäinen, M., and Harwood, D. (1996). A comparative study of texture measures with classification based on featured distributions. *Pattern Recognition*, 29(1):51 – 59.

- 
- [69] Olsson, J. O. S. and Oard, D. W. (2006). Combining feature selectors for text classification. In *Proceedings of the 15th ACM International Conference on Information and Knowledge Management, CIKM '06*, pages 798–799, New York, NY, USA. ACM.
- [70] Orman, G. K., Labatut, V., and Cherifi, H. (2011). On Accuracy of Community Structure Discovery Algorithms. *CoRR*, abs/1112.4134.
- [71] Palchowdhury, S., Pal, S., and Mitra, M. (2011). Using negative information in search. In *Emerging Applications of Information Technology (EAIT), 2011 Second International Conference on*, pages 53–56. IEEE.
- [72] Pavlidis, T. (2008). Limitations of content-based image retrieval. In *Invited Plenary Talk at the 19th Internat. Conf. on Pattern Recognition, Tampa, Florida, December*, pages 8–11.
- [73] Porter, M. F. (1980). An algorithm for suffix stripping. *Program*, 14(3):130–137.
- [74] Rocchio, J. J. (1971). Relevance feedback in information retrieval. *THE SMART RETRIEVAL SYSTEM: Experiments in Automatic Document Processing*, pages 313–323.
- [75] Rosvall, M. and Bergstrom, C. T. (2008). Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences*, 105(4):1118–1123.

- 
- [76] Russell, B. C., Torralba, A., Murphy, K. P., and Freeman, W. T. (2008). LabelMe: A Database and Web-Based Tool for Image Annotation. *International Journal of Computer Vision*, 77(1):157–173.
- [77] Ruthven, I. (2003). Re-examining the potential effectiveness of interactive query expansion. In *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval*, pages 213–220. ACM.
- [78] Ruthven, I. and Lalmas, M. (2003). A survey on the use of relevance feedback for information access systems. *The Knowledge Engineering Review*, 18(02):95–145.
- [79] Salton, G. and McGill, M. J. (1986). Introduction to modern information retrieval.
- [80] Salton, G., Wong, A., and Yang, C. S. (1975). A Vector Space Model for Automatic Indexing. *Commun. ACM*, 18(11):613–620.
- [81] Santos, R. L., McCreadie, R., Macdonald, C., and Ounis, I. (2010). University of Glasgow at TREC 2010: Experiments with Terrier in Blog and Web Tracks. In *TREC*.
- [82] Schaeffer, S. E. (2007). Graph clustering. *Computer Science Review*, 1(1):27–64.

- 
- [83] Snoek, C. G., Worring, M., and Smeulders, A. W. (2005). Early versus late fusion in semantic video analysis. In *Proceedings of the 13th annual ACM international conference on Multimedia*, pages 399–402. ACM.
- [84] Song, G., Wang, S., and Tian, Q. (2015). Fusing feature and similarity for multimodal search. In *Signal and Information Processing (ChinaSIP), 2015 IEEE China Summit and International Conference on*, pages 787–791. IEEE.
- [85] Sparck Jones, K. (1988). A statistical interpretation of term specificity and its application in retrieval. In *Document retrieval systems*, pages 132–142. Taylor Graham Publishing.
- [86] Stehling, R. O., Nascimento, M. A., and Falcão, A. X. (2002). A compact and efficient image retrieval approach based on border/interior pixel classification. In *Proceedings of the eleventh international conference on Information and knowledge management*, pages 102–109. ACM.
- [87] Uricchio, T., Ballan, L., Seidenari, L., and Bimbo, A. D. (2017). Automatic image annotation via label transfer in the semantic space. *Pattern Recognition*, 71:144 – 157.
- [88] Ushiku, Y., Muraoka, H., Inaba, S., Fujisawa, T., Yasumoto, K., Gunji, N., Higuchi, T., Hara, Y., Harada, T., and Kuniyoshi, Y. (2012). Isi at imageclef 2012: Scalable system for image annotation. In *CLEF (Online Working Notes/Labs/-Workshop)*.

- 
- [89] Vedaldi, A. and Fulkerson, B. (2010). Vlfeat: An Open and Portable Library of Computer Vision Algorithms. In *Proceedings of the 18th ACM International Conference on Multimedia*, MM '10, pages 1469–1472, New York, NY, USA. ACM.
- [90] Villegas, M. and Paredes, R. (2012a). A k-NN Approach for Scalable Image Annotation Using General Web Data. In *Big Data Meets Computer Vision: First International Workshop on Large Scale Visual Recognition and Retrieval, held in conjunction with NIPS*, volume 2012, pages 1–5.
- [91] Villegas, M. and Paredes, R. (2012b). Overview of the ImageCLEF 2012 Scalable Web Image Annotation Task. In *CLEF 2012 Evaluation Labs and Workshop, Online Working Notes, Rome, Italy, September 17-20, 2012*.
- [92] Wang, S., Pan, P., Lu, Y., and Xie, L. (2015). Improving cross-modal and multi-modal retrieval combining content and semantics similarities with probabilistic model. *Multimedia Tools and Applications*, 74(6):2009–2032.
- [93] Wang, X.-J. and Zhang, L. (2009). *Annotation-based Image Retrieval*, pages 85–88. Springer US, Boston, MA.
- [94] Welling, M. (2005). Fisher linear discriminant analysis. *Department of Computer Science, University of Toronto*, 3.
- [95] Witten, I. H., Paynter, G. W., Frank, E., Gutwin, C., and Nevill-Manning, C. G. (1999). KEA: Practical Automatic Keyphrase Extraction. In *Proceedings of the Fourth ACM Conference on Digital Libraries*, DL '99, pages 254–255, New York, NY, USA. ACM.

- 
- [96] Xu, G., Li, J., Xu, C., and Wang, Q. (2015). Content-Based Image Retrieval Improved by Incorporating Semantic Annotation via Query Expansion.
- [97] Xu, Y., Jones, G. J., and Wang, B. (2009). Query dependent pseudo-relevance feedback based on wikipedia. In *Proceedings of the 32Nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '09, pages 59–66, New York, NY, USA. ACM.
- [98] Yagcioglu, S., Erdem, E., Erdem, A., and Cakıcı, R. (2015). A Distributed Representation Based Query Expansion Approach for Image Captioning. *Volume 2: Short Papers*, page 106.
- [99] Yang, K., Hua, X. S., Wang, M., and Zhang, H. J. (2011). Tag Tagging: Towards More Descriptive Keywords of Image Content. *IEEE Transactions on Multimedia*, 13(4):662–673.
- [100] Yildizer, E., Balci, A. M., Hassan, M., and Alhajj, R. (2012a). Efficient content-based image retrieval using Multiple Support Vector Machines Ensemble. *Expert Systems with Applications*, 39(3):2385 – 2396.
- [101] Yildizer, E., Balci, A. M., Hassan, M., and Alhajj, R. (2012b). Efficient content-based image retrieval using Multiple Support Vector Machines Ensemble. *Expert Systems with Applications*, 39(3):2385 – 2396.
- [102] Yoo, H.-W., Park, H.-S., and Jang, D.-S. (2005). Expert system for color image retrieval. *Expert Systems with Applications*, 28(2):347 – 357.

- 
- [103] Zellhöfer, D. (2012). A permeable expert search strategy approach to multimodal retrieval. In *Proceedings of the 4th Information Interaction in Context Symposium*, pages 62–71. ACM.
- [104] Zhai, C. (2001). Notes on the Lemur TFIDF model. *Unpublished report*.
- [105] Zhang, J., Gao, Y., Feng, S., Yuan, Y., and Lee, C. H. (2016). Automatic image region annotation through segmentation based visual semantic analysis and discriminative classification. In *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1956–1960.
- [106] Zhao, R. and Grosky, W. I. (2002). Narrowing the semantic gap-improved text-based web document retrieval using visual features. *Multimedia, IEEE Transactions on*, 4(2):189–200.
- [107] Zhou, X. S. and Huang, T. S. (2002). Unifying Keywords and Visual Contents in Image Retrieval. *IEEE MultiMedia*, 9(2):23–33.