

REFERENCES

- [1] Yang Cong, Junsong Yuan, and Jiebo Luo. 2012. Towards scalable summarization of consumer videos via sparse dictionary selection. *IEEE Transactions on Multimedia* 14, 1 (2012), 66–75.
- [2] Sandra Eliza Fontes De Avila, Ana Paula Brandão Lopes, Antonio da Luz Jr, and Arnaldo de Albuquerque Araújo. 2011. VSUMM: A mechanism designed to produce static video summaries and a novel evaluation method. *Pattern Recognition Letters* 32, 1 (2011), 56–68.
- [3] Ehsan Elhamifar, Guillermo Sapiro, and Rene Vidal. 2012. See all by looking at a few: Sparse modeling for finding representative objects. In *CVPR*. 1600–1607.
- [4] Christoph Feichtenhofer, Axel Pinz, and Andrew Zisserman. 2016. Convolutional two-stream network fusion for video action recognition. In *CVPR*. 1933–1941.
- [5] Boqing Gong, Wei-Lun Chao, Kristen Grauman, and Fei Sha. 2014. Diverse sequential subset selection for supervised video summarization. In *NIPS*. 2069–2077.
- [6] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2014. Generative adversarial nets. In *Advances in neural information processing systems*. 2672–2680.
- [7] Genliang Guan, Zhiyong Wang, Shaohui Mei, Max Ott, Mingyi He, and David Dagan Feng. 2014. A top-down approach for video summarization. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)* 11, 1 (2014), 4.
- [8] Michael Gygli, Helmut Grabner, Hayko Riemenschneider, and Luc Van Gool. 2014. Creating summaries from user videos. In *ECCV*. Springer, 505–520.
- [9] Michael Gygli, Helmut Grabner, and Luc Van Gool. 2015. Video summarization by learning submodular mixtures of objectives. In *CVPR*. 3090–3098.
- [10] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 770–778.
- [11] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. 2012. Imagenet classification with deep convolutional neural networks. In *NIPS*. 1097–1105.
- [12] Xuelong Li, Zhigang Wang, and Xiaoqiang Lu. 2016. Surveillance video synopsis via scaling down objects. *IEEE Transactions on Image Processing* 25, 2 (2016), 740–755.
- [13] Xuelong Li, Bin Zhao, and Xiaoqiang Lu. 2017. A general framework for edited video and raw video summarization. *IEEE Transactions on Image Processing* 26, 8 (2017), 3652–3664.
- [14] Jonathan Long, Evan Shelhamer, and Trevor Darrell. 2015. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 3431–3440.
- [15] Mingyang Ma, Shaohui Mei, Jingyu Ji, Shuai Wan, Zhiyong Wang, and Dagan Feng. 2017. Exploring the influence of feature representation for dictionary selection based video summarization. In *ICIP*. IEEE, 2911–2915.
- [16] Mingyang Ma, Shaohui Mei, Shuai Wan, Zhiyong Wang, and Dagan Feng. 2019. Video Summarization via Nonlinear Sparse Dictionary Selection. *IEEE Access* 7 (2019), 11763–11774.
- [17] Mingyang Ma, Shaohui Mei, Shuai Wan, Zhiyong Wang, and David Dagan Feng. 2018. Robust video summarization using collaborative representation of adjacent frames. *Multimedia Tools and Applications* (2018), 1–21.
- [18] Behrooz Mahasseni, Michael Lam, and Sinisa Todorovic. 2017. Unsupervised video summarization with adversarial lstm networks. In *CVPR*, Vol. 1.
- [19] Shaohui Mei, Genliang Guan, Zhiyong Wang, Shuai Wan, Mingyi He, and David Dagan Feng. 2015. Video summarization via minimum sparse reconstruction. *Pattern Recognition* 48, 2 (2015), 522–533.
- [20] MerchDope 2019. 37 Mind Blowing YouTube Facts, Figures and Statistics C 2019. <https://merchdope.com/youtube-stats/>.
- [21] Joe Yue-Hei Ng, M. Hausknecht, S. Vijayanarasimhan, O. Vinyals, R. Monga, and G. Toderici. 2015. Beyond short snippets: Deep networks for video classification. In *CVPR*. 4694–4702.
- [22] OVP 2011. Open video project. <https://open-video.org/>.
- [23] Danila Potapov, Matthijs Douze, Zaid Harchaoui, and Cordelia Schmid. 2014. Category-specific video summarization. In *ECCV*. 540–555.
- [24] Mrigank Rochan, Linwei Ye, and Yang Wang. 2018. Video Summarization Using Fully Convolutional Sequence Networks. *arXiv preprint arXiv:1805.10538* (2018).
- [25] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, et al. 2015. Imagenet large scale visual recognition challenge. *International journal of computer vision* 115, 3 (2015), 211–252.
- [26] Alan F Smeaton, Paul Over, and Wessel Kraaij. 2006. Evaluation campaigns and TRECVID. In *Proceedings of the 8th ACM international workshop on Multimedia information retrieval*. ACM, 321–330.
- [27] Yale Song, Jordi Vallmitjana, Amanda Stent, and Alejandro Jaimes. 2015. Tvsum: Summarizing web videos using titles. In *CVPR*. 5179–5187.
- [28] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. 2015. Going deeper with convolutions. In *CVPR*. 1–9.
- [29] Du Tran, Lubomir Bourdev, Rob Fergus, Lorenzo Torresani, and Manohar Paluri. 2015. Learning spatiotemporal features with 3d convolutional networks. In *ICCV*. 4489–4497.
- [30] Ba Tu Truong and Svetha Venkatesh. 2007. Video abstraction: A systematic review and classification. *ACM transactions on multimedia computing, communications, and applications* 3, 1 (2007), 3.
- [31] S. Venugopalan, M. Rohrbach, J. Donahue, R. Mooney, T. Darrell, and K. Saenko. 2015. Sequence to Sequence – Video to Text. In *2015 IEEE International Conference on Computer Vision (ICCV)*. 4534–4542.
- [32] Junbo Wang, Wei Wang, Yan Huang, Liang Wang, and Tieniu Tan. 2018. M3: Multimodal memory modelling for video captioning. In *CVPR*. 7512–7520.
- [33] Huawei Wei, Bingbing Ni, Yichao Yan, Huanyu Yu, Xiaokang Yang, and Chen Yao. 2018. Video Summarization via Semantic Attended Networks.. In *AAAI*.
- [34] Caiming Xiong, Stephen Merity, and Richard Socher. 2016. Dynamic memory networks for visual and textual question answering. In *ICML*. 2397–2406.
- [35] Huan Yang, Baoyuan Wang, Stephen Lin, David Wipf, Minyi Guo, and Baining Guo. 2015. Unsupervised extraction of video highlights via robust recurrent auto-encoders. In *ICCV*. 4633–4641.
- [36] Ting Yao, Tao Mei, and Yong Rui. 2016. Highlight detection with pairwise deep ranking for first-person video summarization. In *CVPR*. 982–990.
- [37] Ke Zhang, Wei-Lun Chao, Fei Sha, and Kristen Grauman. 2016. Summary transfer: Exemplar-based subset selection for video summarization. In *CVPR*. 1059–1067.
- [38] Ke Zhang, Wei-Lun Chao, Fei Sha, and Kristen Grauman. 2016. Video summarization with long short-term memory. In *ECCV*. 766–782.
- [39] Ke Zhang, Kristen Grauman, and Fei Sha. 2018. Retrospective Encoders for Video Summarization. In *ECCV*. 383–399.
- [40] Bin Zhao, Xuelong Li, and Xiaoqiang Lu. 2017. Hierarchical recurrent neural network for video summarization. In *Proceedings of the 2017 ACM on Multimedia Conference*. ACM, 863–871.
- [41] Bin Zhao, Xuelong Li, and Xiaoqiang Lu. 2018. HSA-RNN: Hierarchical Structure-Adaptive RNN for Video Summarization. In *CVPR*. 7405–7414.
- [42] Bin Zhao and Eric P Xing. 2014. Quasi real-time summarization for consumer videos. In *CVPR*. 2513–2520.
- [43] Kaiyang Zhou, Yu Qiao, and Tao Xiang. 2018. Deep reinforcement learning for unsupervised video summarization with diversity-representativeness reward. In *AAAI*.