

REFERENCES

- [1] Mark Baillie and Joemon M Jose. Audio-based event detection for sports video. In *International Conference on Image and Video Retrieval*, pages 300–309. Springer, 2003.
- [2] Marc-André Carbonneau, Alexandre J Raymond, Eric Granger, and Ghyslain Gagnon. Real-time visual play-break detection in sport events using a context descriptor. In *2015 IEEE International Symposium on Circuits and Systems (IS-CAS)*, pages 2808–2811. IEEE, 2015.
- [3] François Chollet. Xception: Deep learning with depthwise separable convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1251–1258, 2017.
- [4] Samuel F de Sousa, Arnaldo de A Araújo, and David Menotti. An overview of automatic event detection in soccer matches. In *2011 IEEE Workshop on Applications of Computer Vision (WACV)*, pages 31–38. IEEE, 2011.
- [5] Carl Doersch. Tutorial on variational autoencoders. *arXiv preprint arXiv:1606.05908*, 2016.
- [6] Ahmet Ekin and M Tekalp. Generic play-break event detection for summarization and hierarchical sports video analysis. In *2003 International Conference on Multimedia and Expo. ICME'03. Proceedings (Cat. No. 03TH8698)*, volume 1, pages I–169. IEEE, 2003.
- [7] Ross Girshick. Fast r-cnn. In *Proceedings of the IEEE international conference on computer vision*, pages 1440–1448, 2015.
- [8] Ross Girshick, Jeff Donahue, Trevor Darrell, and Jitendra Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 580–587, 2014.
- [9] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE transactions on pattern analysis and machine intelligence*, 37(9):1904–1916, 2015.
- [10] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

- [11] Ali Karimi, Ramin Toosi, and Mohammad Ali Akhaee. Soccer event detection using deep learning. *arXiv preprint arXiv:2102.04331*, 2021.
- [12] Abdullah Khan, Beatrice Lazzerini, Gaetano Calabrese, and Luciano Serafini. Soccer event detection. In *4th international conference on image processing and pattern recognition (IPPR 2018)*, pages 119–129. AIRCC Publishing Corporation, 2018.
- [13] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25:1097–1105, 2012.
- [14] Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.
- [15] Baoxin Li and M Ibrahim Sezan. Event detection and summarization in sports video. In *Proceedings IEEE Workshop on Content-Based Access of Image and Video Libraries (CBAIVL 2001)*, pages 132–138. IEEE, 2001.
- [16] Yongqiang Li, S Mohammad Mavadati, Mohammad H Mahoor, Yongping Zhao, and Qiang Ji. Measuring the intensity of spontaneous facial action units with dynamic bayesian network. *Pattern Recognition*, 48(11):3417–3427, 2015.
- [17] Tsung-Yi Lin, Piotr Dollár, Ross Girshick, Kaiming He, Bharath Hariharan, and Serge Belongie. Feature pyramid networks for object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2117–2125, 2017.
- [18] Li Liu, Ling Shao, Xuelong Li, and Ke Lu. Learning spatio-temporal representations for action recognition: A genetic programming approach. *IEEE transactions on cybernetics*, 46(1):158–170, 2015.
- [19] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, and Alexander C Berg. Ssd: Single shot multibox detector. In *European conference on computer vision*, pages 21–37. Springer, 2016.
- [20] Li Lu, Fengpei Ge, Qingwei Zhao, and Yonghong Yan. A svm-based audio event detection system. In *2010 International Conference on Electrical and Control Engineering*, pages 292–295. IEEE, 2010.
- [21] Mahdyar Ravanbakhsh, Moin Nabi, Enver Sangineto, Lucio Marcenaro, Carlo Regazzoni, and Nicu Sebe. Abnormal event detection in videos using generative adversarial nets. In *2017 IEEE International Conference on Image Processing (ICIP)*, pages 1577–1581. IEEE, 2017.
- [22] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 779–788, 2016.

- [23] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster r-cnn: Towards real-time object detection with region proposal networks. *Advances in neural information processing systems*, 28, 2015.
- [24] A Salunke, Sonal Deshmukh, and Ratnadeep Deshmukh. Event detection in sports video game -a review. 01 2010.
- [25] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.
- [26] Nitish Srivastava, Elman Mansimov, and Ruslan Salakhudinov. Unsupervised learning of video representations using lstms. In *International conference on machine learning*, pages 843–852. PMLR, 2015.
- [27] Christian Szegedy, Sergey Ioffe, Vincent Vanhoucke, and Alexander A Alemi. Inception-v4, inception-resnet and the impact of residual connections on learning. In *Thirty-first AAAI conference on artificial intelligence*, 2017.
- [28] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1–9, 2015.
- [29] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna. Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2818–2826, 2016.
- [30] Grigorios Tsagkatakis, Mustafa Jaber, and Panagiotis Tsakalides. Goal!! event detection in sports video. *Electronic Imaging*, 2017(16):15–20, 2017.
- [31] Gül Varol, Ivan Laptev, and Cordelia Schmid. Long-term temporal convolutions for action recognition. *IEEE transactions on pattern analysis and machine intelligence*, 40(6):1510–1517, 2017.
- [32] Jinjun Wang, Changsheng Xu, Engsiong Chng, Xinguo Yu, and Qi Tian. Event detection based on non-broadcast sports video. In *2004 International Conference on Image Processing, 2004. ICIP'04.*, volume 3, pages 1637–1640. IEEE, 2004.
- [33] Lexing Xie, Shih-Fu Chang, Ajay Divakaran, and Huifang Sun. Structure analysis of soccer video with hidden markov models. In *2002 IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 4, pages IV–4096. IEEE, 2002.
- [34] Saining Xie, Ross Girshick, Piotr Dollár, Zhuowen Tu, and Kaiming He. Aggregated residual transformations for deep neural networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1492–1500, 2017.

- [35] Changsheng Xu, Jinjun Wang, Kongwah Wan, Yiqun Li, and Lingyu Duan. Live sports event detection based on broadcast video and web-casting text. In *Proceedings of the 14th ACM international conference on Multimedia*, pages 221–230, 2006.
- [36] Gu Xu, Yu-Fei Ma, Hong-Jiang Zhang, and Shiqiang Yang. A hmm based semantic analysis framework for sports game event detection. In *Proceedings 2003 International Conference on Image Processing (Cat. No. 03CH37429)*, volume 1, pages I–25. IEEE, 2003.
- [37] Min Xu, Namunu C Maddage, Changsheng Xu, Mohan Kankanhalli, and Qi Tian. Creating audio keywords for event detection in soccer video. In *2003 International Conference on Multimedia and Expo. ICME'03. Proceedings (Cat. No. 03TH8698)*, volume 2, pages II–281. IEEE, 2003.
- [38] Hossam M Zawbaa, Nashwa El-Bendary, Aboul Ella Hassanien, and Tai-hoon Kim. Event detection based approach for soccer video summarization using machine learning. *International Journal of Multimedia and Ubiquitous Engineering*, 7(2):63–80, 2012.