modeling. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 1, page 3, 2017. 2, 7

[30] Fadime Sener and Angela Yao. Unsupervised learning and segmentation of complex activities from video. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018. 2, 6, 7, 8

[31] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014. 5

[32] Sebastian Stein and Stephen J McKenna. Combining embedded accelerometers with computer vision for recognizing food preparation activities. In *ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 729–738. ACM, 2013. 5

[33] Simon Thorpe, Denis Fize, and Catherine Marlot. Speed of processing in the human visual system. *Nature*, 381(6582):520, 1996. 5

[34] Teun Adrianus Van Dijk, Walter Kintsch, and Teun Adrianus Van Dijk. *Strategies of discourse comprehension*. Academic Press New York, 1933. 2

[35] Subhashini Venugopalan, Marcus Rohrbach, Jeffrey Donahue, Raymond Mooney, Trevor Darrell, and Kate Saenko. Sequence to sequence-video to text. In *IEEE International Conference on Computer Vision (ICCV)*, pages 4534–4542, 2015. 1

[36] Subhashini Venugopalan, Huijuan Xu, Jeff Donahue, Marcus Rohrbach, Raymond Mooney, and Kate Saenko. Translating videos to natural language using deep recurrent neural networks. *arXiv preprint arXiv:1412.4729*, 2014. 1

[37] Carl Vondrick, Hamed Pirsiavash, and Antonio Torralba. Anticipating visual representations from unlabeled video. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 98–106, 2016. 4

[38] Heng Wang and Cordelia Schmid. Action recognition with improved trajectories. In *IEEE International Conference on Computer Vision (ICCV)*, pages 3551–3558, 2013. 3

[39] Li Yao, Atousa Torabi, Kyunghyun Cho, Nicolas Ballas, Christopher Pal, Hugo Larochelle, and Aaron Courville. Describing videos by exploiting temporal structure. In *IEEE International Conference on Computer Vision (ICCV)*, pages 4507–4515, 2015. 1

[40] Jeffrey M Zacks and Khena M Swallow. Event segmentation. *Current Directions in Psychological Science*, 16(2):80–84, 2007. 2

[41] Jeffrey M Zacks and Barbara Tversky. Event structure in perception and conception. *Psychological bulletin*, 127(1):3, 2001. 2

[42] Jeffrey M Zacks, Barbara Tversky, and Gowri Iyer. Perceiving, remembering, and communicating structure in events. *Journal of Experimental Psychology: General*, 130(1):29, 2001. 1, 2