Real-Time Temporal Action Localization in Untrimmed Videos by Sub-Action Discovery

Rui Hou[1]
rhou@crcv.ucf.edu

Rahul Sukthankar[2]
sukthankar@google.com

Mubarak Shah[1]
shah@crcv.ucf.edu

[1] Center for Research in
Computer Vision
University of Central Florida
Orlando, USA

[2] Google Research
Mountain View, USA

In this supplementary material, we provide more experimental results and detailed analysis of our proposed approach.

On the THUMOS'14 dataset, the AP (Average Precision) for each class for each method can be found in Table A. Our method outperforms [2, 3] for most of the actions at IOU of $\alpha = 0.5$, improving the overall average performance from 17.1% to 22.0%. Finally, compared to [1] our method outperforms in all but two actions at IOU of $\alpha = 0.5$; improving mAP from 14.4% to 22%.

action	Ours	Yeung et al. [3]	Oneata et al. [1]	Shou et al. [2]
Baseball Pitch	**38.7**	14.6	8.6	14.7
Basketball Dunk	9.1	6.3	1.0	**20.0**
Billiards	**10.5**	9.4	2.6	7.6
Clean And Jerk	19.0	**42.8**	13.3	24.7
Cliff Diving	23.5	15.6	17.7	27.5
Cricket Bowling	**30.9**	10.8	9.5	15.4
Cricket Shot	3.8	3.5	2.6	**13.8**
Diving	**15.5**	10.8	4.6	**17.6**
Frisbee Catch	9.6	10.4	1.2	**15.3**
Golf Swing	13.0	13.8	**22.6**	18.2
Hammer Throw	28.6	28.9	**34.7**	18.8
High Jump	20.3	**33.3**	18.4	19.9
Javelin Throw	**28.1**	20.4	22.0	17.7
Long Jump	**53.6**	39.0	47.6	34.8
Pole Vault	30.9	16.3	19.6	**31.9**
Shotput	**26.3**	16.6	11.9	11.8
Soccer Penalty	**22.7**	8.3	8.7	18.6
Tennis Swing	4.3	5.6	3.0	**19.2**
Throw Discus	32.3	29.5	**36.2**	24.4
Volleyball Spiking	**16.1**	5.2	1.4	4.4
mAP	**22.0**	17.1	14.4	18.8

Table 1: *
Table A. APs for each action in the THUMOS'14 dataset.

References

[1] Dan Oneata, Jakob Verbeek, and Cordelia Schmid. The LEAR submission at thumos 2014, 2014.

[2] Zheng Shou, Dongang Wang, and Shih-Fu Chang. Temporal action localization in untrimmed videos via multi-stage cnns. In *CVPR*, 2016.

[3] Serena Yeung, Olga Russakovsky, Greg Mori, and Li Fei-Fei. End-to-end learning of action detection from frame glimpses in videos. In *CVPR*, 2016.