

Prominent Video Action Understanding Challenges 2013-2020

Matthew Hutchinson (hutchinson@alum.mit.edu)

September 17, 2020

Workshop	Year	Conf.	Problem	Dataset(s)	Metric(s)	#Teams
THUMOS [7]	2013	ICCV	AR	UCF101	average accuracy	17
			SAL/D	UCF101-24	ROC AUC sIoU@0.2	n/a
THUMOS [8]	2014	ECCV	AR	UCF101+	mAP	11
			TAL/D	UCF101-20	mAP tIoU@{0.1,0.2,0.3,0.4,0.5}	3
THUMOS [5, 6]	2015	CVPR	AR	UCF101+1	mAP	11
			TAL/D	UCF101-20	mAP tIoU@{0.1,0.2,0.3,0.4,0.5}	1
ActivityNet [2]	2016	CVPR	AR	ActivityNet 1.3	mAP, Top-1 accuracy, Top-3 accuracy	26
			TAL/D	ActivityNet 1.3	mAP-50, mAP-75, average-mAP	6
ActivityNet [4]	2017	CVPR	AR	ActivityNet 1.3	Top-1 error	n/a
			AR	Kinetics-400	average(Top-1 error, Top-5 error)	31
			TAP	ActivityNet 1.3	AR-AN AUC	17
			TAL/D	ActivityNet 1.3	mAP tIoU@0.5:0.05:0.95	17
ActivityNet [3]	2018	CVPR	TAP	ActivityNet 1.3	AR-AN AUC	55
			TAL/D	ActivityNet 1.3	mAP tIoU@0.5:0.05:0.95	43
			AR	Kinetics-600	average(Top-1 error, Top-5 error)	13
			SAL/D	AVA	frame-mAP sIoU@0.5	23
			AR	MiT (full-track)	average(Top-1 acc, Top-5 acc)	29
			AR	MiT (mini-track)	average(Top-1 acc, Top-5 acc)	12
ActivityNet [10, 9]	2019	CVPR	TAP	ActivityNet 1.3	AR-AN AUC	72
			TAL/D	ActivityNet 1.3	mAP tIoU@0.5:0.05:0.95	23
			AR	Kinetics-700	average(Top-1 error, Top-5 error)	15
			SAL/D	AVA	frame mAP sIoU@0.5	32
			AR	EPIC-KITCHENS-55	micro-avg Top-1,5 acc, macro-AP,AR	39
			AP	EPIC-KITCHENS-55	micro-avg Top-1,5 acc, macro-AP,AR	19
			TAL/D	VIRAT	$P_{rate}@miss_{FA}$	42
Multi-modal [1]	2019	ICCV	AR	Multi-MiT	mAP	10
			TAL/D	HACS Segments	mAP tIoU@0.5:0.05:0.95	5
ActivityNet [11]	2020	CVPR	TAL/D	ActivityNet 1.3	mAP tIoU@0.5:0.05:0.95	n/a
			AR	Kinetics-700	average(Top-1 error, Top-5 error)	n/a
			SAL/D	AVA	frame mAP sIoU@0.5	n/a
			TAL/D	VIRAT	$P_{rate}@miss_{FA}$	11
			TAL/D	HACS Segments	mAP tIoU@0.5:0.05:0.95	22
			TAL/D	HACS Clips+Seg.	mAP tIoU@0.5:0.05:0.95	13

AR = Action Recognition

TAP = Temporal Action Proposal

TAL/D = Temporal Action Localization/Detection

SAL/D = Spatiotemporal Action Localization/Detection

Research was sponsored by the United States Air Force Research Laboratory and the United States Air Force Artificial Intelligence Accelerator and was accomplished under Cooperative Agreement Number FA8750-19-2-1000. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the United States Air Force or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein.

References

- [1] Workshop on multi-modal video analysis and moments in time challenge. 2019.
- [2] Fabian Caba Heilbron, Victor Escorcia, Bernard Ghanem, and Juan Carlos Nibbles. Activitynet: A large-scale video benchmark for human activity understanding. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2015.
- [3] Bernard Ghanem, Juan Carlos Nibbles, Cees Snoek, Fabian Caba Heilbron, Humam Alwassel, Victor Escorcia, Ranjay Krishna, Shyamal Buch, and Cuong Duc Dao. The activitynet large-scale activity recognition challenge 2018 summary, 2018.
- [4] Bernard Ghanem, Juan Carlos Nibbles, Cees Snoek, Fabian Caba Heilbron, Humam Alwassel, Ranjay Khrisna, Victor Escorcia, Kenji Hata, and Shyamal Buch. Activitynet challenge 2017 summary, 2017.
- [5] A. Gorban, H. Idrees, Y.-G. Jiang, A. Roshan Zamir, I. Laptev, M. Shah, and R. Sukthankar. THUMOS challenge: Action recognition with a large number of classes. <http://www.thumos.info/>, 2015.

- [6] Haroon Idrees, Amir R. Zamir, Yu-Gang Jiang, Alex Gorban, Ivan Laptev, Rahul Sukthankar, and Mubarak Shah. The thumos challenge on action recognition for videos “in the wild”. *Computer Vision and Image Understanding*, 155:1–23, Feb 2017.
- [7] Y.-G. Jiang, J. Liu, A. Roshan Zamir, I. Laptev, M. Piccardi, M. Shah, and R. Sukthankar. THUMOS challenge: Action recognition with a large number of classes, 2013.
- [8] Y.-G. Jiang, J. Liu, A. Roshan Zamir, G. Toderici, I. Laptev, M. Shah, and R. Sukthankar. THUMOS challenge: Action recognition with a large number of classes, 2014.
- [9] Yooyoung Lee, Jon Fiscus, Afzal Godil, Andrew Delgado, Jim Golden, Lukas Diduch, and Maxime Hubert. Summary of the 2019 activity detection in extended videos prize challenge. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV) Workshops*, March 2020.
- [10] Cees Snoek, Juan Carlos Niebles, Bernard Ghanem, Fabian Caba Heilbron, Humam Alwassel, Victor Escorcia, Ranjay Krishna, Shyamal Buch, and Frost Xu. International challenge on activity recognition, 2019.
- [11] Cees Snoek, Juan Carlos Niebles, Bernard Ghanem, Fabian Caba Heilbron, Humam Alwassel, Victor Escorcia, Ranjay Krishna, Shyamal Buch, and Frost Xu. International challenge on activity recognition, 2020.