Activity-Net 2017 Challenge Report

Xiyang Dai	Bharat Singh	Larry S. Davis
UMD	UMD	UMD
xdai@umiacs.umd.edu	bharat@umiacs.umd.edu	lsd@umiacs.umd.edu

In this year's competition, we use Temporal Context Network (TCN) for precise temporal localization of human activities. To improve performance for the metric used for evaluating proposals (i.e. area under the Average Recall vs. Average Number of Proposals per Video (AR-AN) curve with 100 proposals), we study the influence of two major hyper-parameters: the IOU threshold for NMS and the number of proposals needed from each level of the pyramid. Table 1 shows the impact of NMS IOU threshold on final proposal performance. In contrast to the localization task, we found high IOUs lead to better results on the proposal task. Figure 1 shows the impact of proposals sampled from different anchor lengths (we use percentage of video length to build a temporal pyramid instead of fixed number of frames used in previous works). Table 2 shows our performance on the testing server. After tuning the NMS threshold and proposal anchor length used in TCN, we obtain an AUC score of 61.56 on the testing server, which is ranked third on the ActivityNet 2017 Challenge for the proposal task.

For maximizing recall using a fixed number of proposals, it is important to change the NMS threshold. Since we are allowed to generate 100 proposals (even for 1-2 activities per video), changing NMS is important to improve recall at higher overlap thresholds. Note that if recall is the only evaluation metric, for different numbers of proposals (like 5,10,50,500 etc.) one should pick a different NMS threshold. For example, if recall at only 5 proposals is measured, an NMS threshold of 0.85 would be very bad because most of the proposals would be on the same activity interval.

Please check out the full paper[1] for further details on TCN.

---

[1] https://arxiv.org/pdf/1708.02349.pdf

NMS	0.3	0.45	0.6	0.75	0.8	0.85	0.9
AUC	42.52	47.14	55.44	57.94	58.22	58.54	57.98

Table 1: Impact of NMS IOU threshold on the proposal performance on the ActivityNet validation set
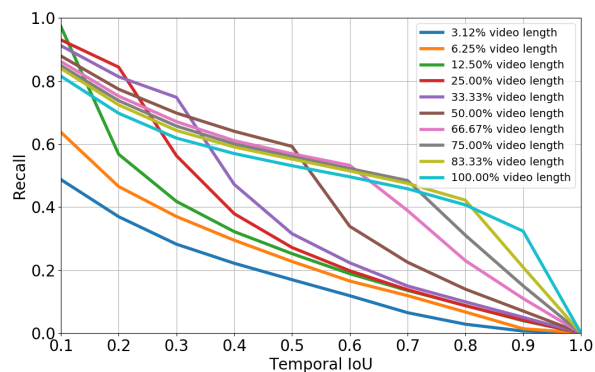

Figure 1: Impact of different anchor lengths on proposal performance on ActivityNet validation set

NMS threshold	AUC on Validation	AUC on Test
baseline	47.14	49.03
fine-tune NMS	58.54	59.89
final	59.58	61.56

Table 2: Our final proposal performance on ActivityNet 2017 Challenge