

Supplemental Material: Enhancing Single-Frame Supervision for Better Temporal Action Localization

Changjian Chen, Jiashu Chen, Weikai Yang, Haoze Wang, Johannes Knittel,
Xibin Zhao, Steffen Koch, Thomas Ertl, and Shixia Liu

APPENDIX A: QUANTITATIVE EVALUATION ON ACTION LOCALIZATION

Table 1 presents a comparison of the mAP@Hit between our method and SF-Net on two benchmark datasets. The results demonstrate that our method outperforms SF-Net across both datasets, which aligns with the findings obtained from the mAP@IoU metric.

Table 1: Performance comparison between our method and SF-Net in terms of the mAP@Hit (in %) on two benchmark datasets.

Annotations	2%	5%	10%	20%
SF-Net	72.73	73.50	74.69	77.38
+ our method	(+2.72)	(+1.97)	(+1.84)	(+0.39)
SF-Net	70.01	71.53	72.85	76.99

(a) THUMOS14

Annotations	2%	5%	10%	20%
SF-Net	74.41	76.76	77.70	79.11
+ our method	(+0.23)	(+0.23)	(+0.70)	(+0.94)
SF-Net	74.18	76.53	77.00	78.17

(b) BEOID