

论文题目

BSN: Boundary-Sensitive Network for Temporal Action Proposal Generation

论文作者

Tianwei Lin, Xu Zhao, Haisheng Su, Chongjing Wang, Ming Yang

论文代码地址

<https://arxiv.org/pdf/1806.02964>

<https://github.com/wzmsltw/BSN-boundary-sensitive-network>

论文观点

高质量的时序动作提名应该具备几点特质

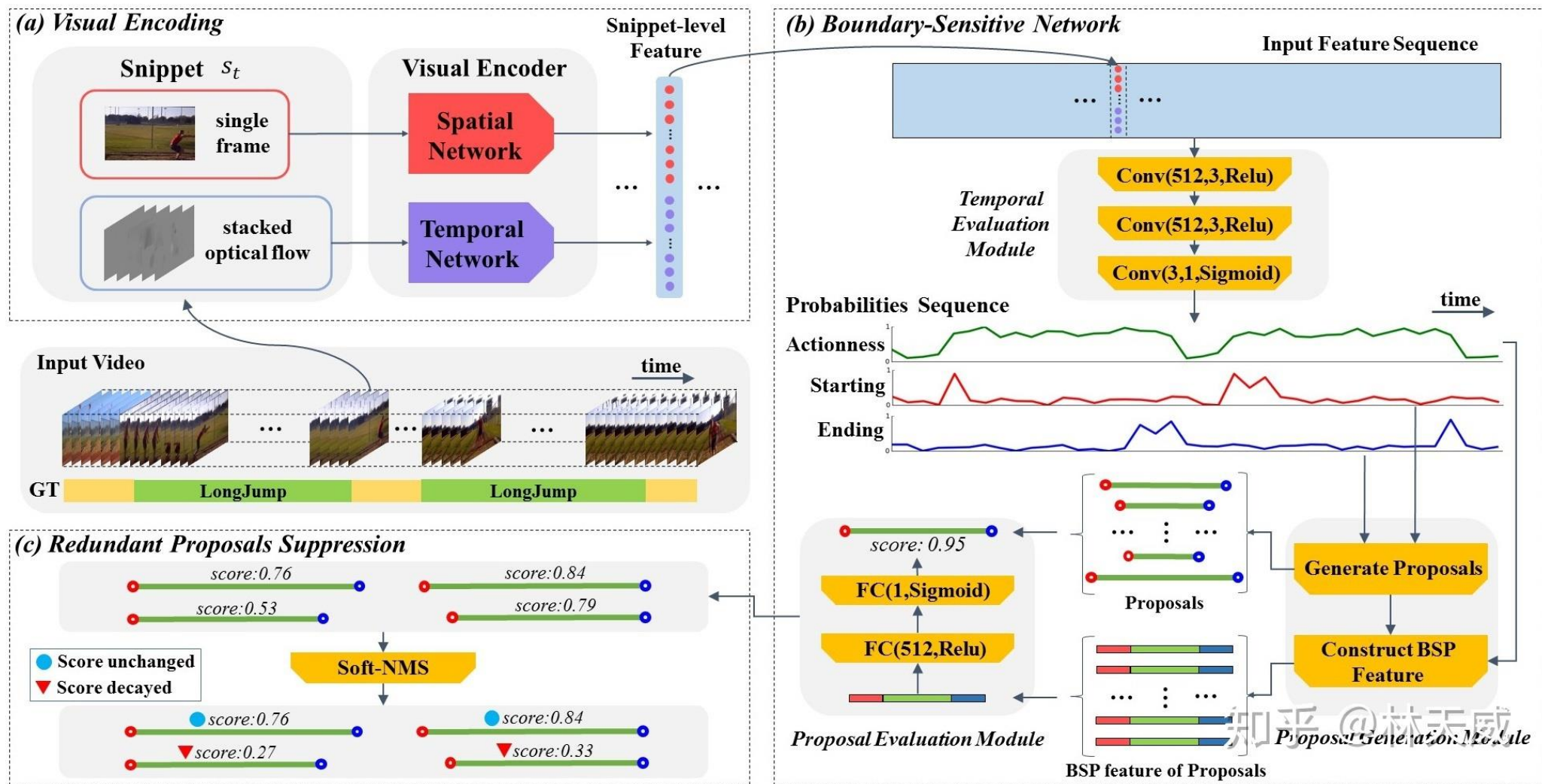
- 灵活的时序长度

- 精确的时序边界

- 可靠的置信度分数

首先去定位时序动作片段的边界(开始节点和结束节点),再将边界节点直接结合成时序提名,最后基于所设计的 proposal-level 的 feature 来对每个 proposal 的置信度进行评估

论文做法



1. 视频特征提取

采用了 two-stream network 双流网络来提取特征,每个用作特征提取的视频帧单元称为一个 snippet,包含图像帧以及光流
拼接 spatial network 和 temporal network 的最后一层输出作为该 snippet 对应的特征

2. 边界敏感网络 (生成时序动作提名)

- 时序评估模块

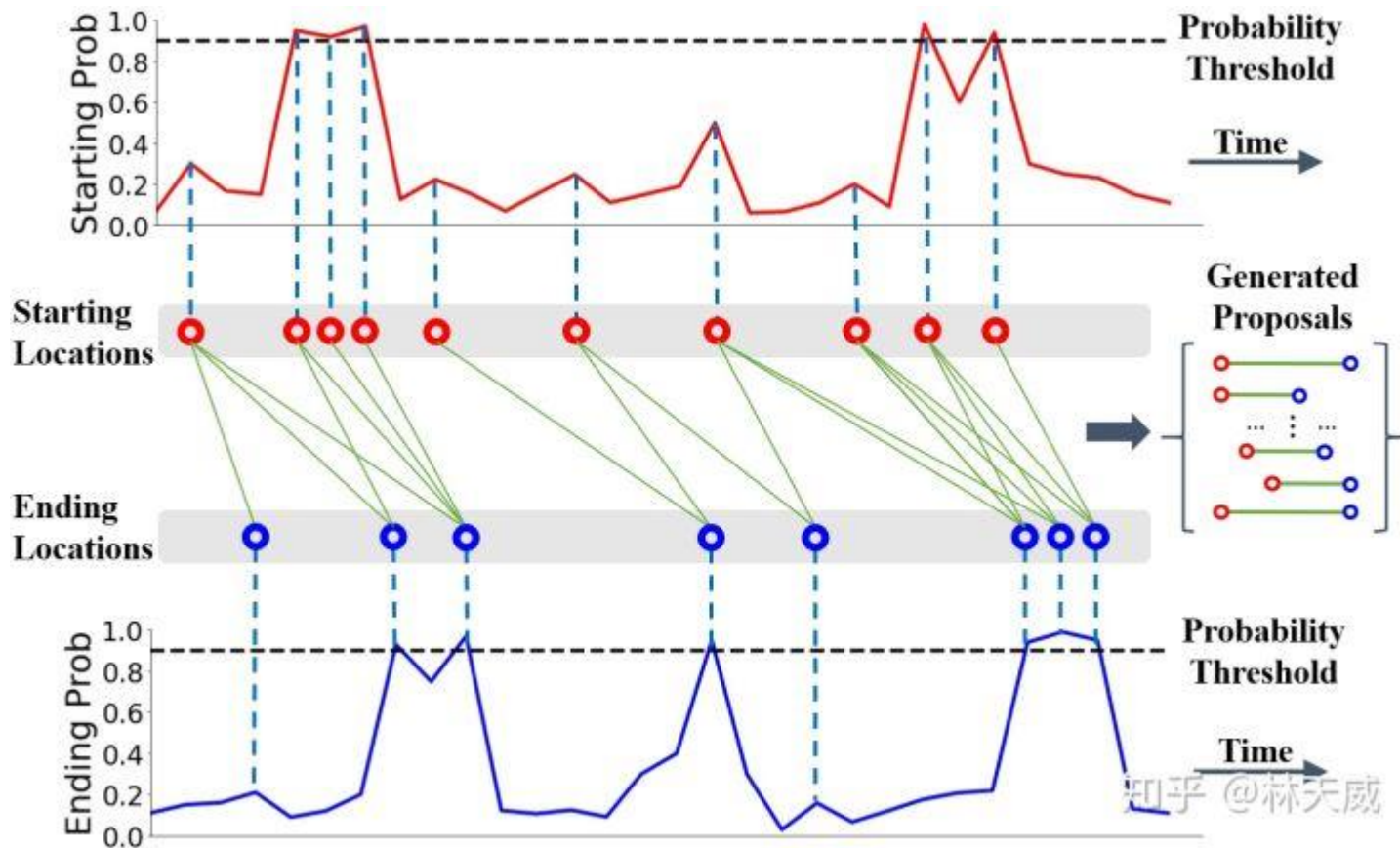
提取的图像特征序列→时序评估模块(采用 3 层时序卷积层来对视频序列中每个位置上动作开始的概率,动作结束的概率和动作进行的概率同时进行建模)→动作开始概率序列,动作结束概率序列和动作概率序列

- 提名生成模块

概率序列→提名生成模块(生成候选时序动作提名,并对每个动作提名生成对应的特征描述)

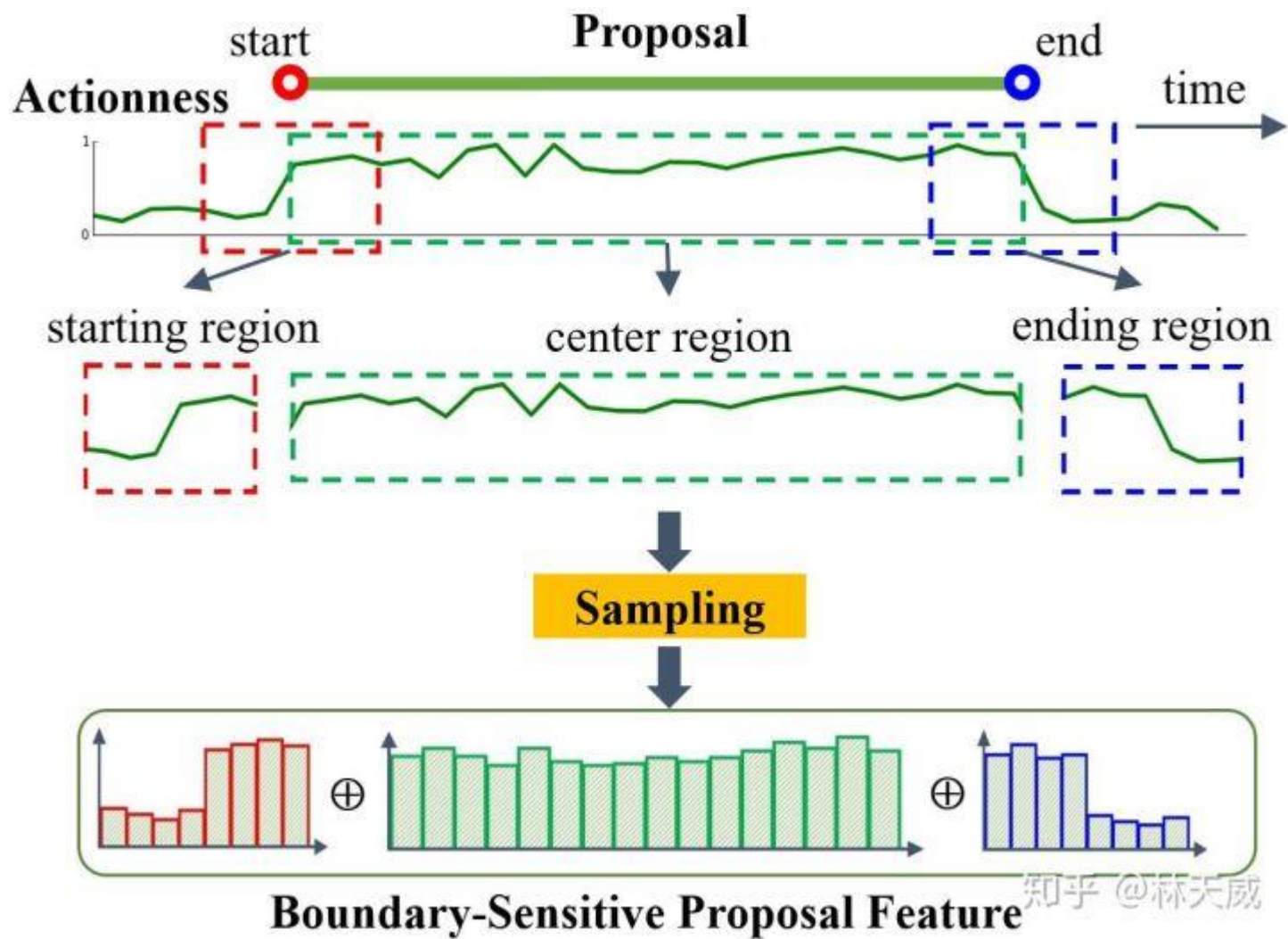
生成候选时序动作提名→选择动作开始和动作结束概率序列中满足以下两点条件之一的时间节点作为候选时序边界节点

概率高于一个阈值或者该时间节点的概率高于前一时刻以及后一时刻的概率



为每个候选时序动作提名生成对应的特征描述(Boundary-Sensitive Proposal (BSP) feature)

每个提名, 取其本身的时序区间作为 center region,再取开始节点和结束节点附近的一段区间作为 starting region 以及 ending region,每个区域都在动作概率序列上的对应位置采样 N 个点,拼接后则得到一个非常简短的提名特征



- 提名评估模块

生成候选时序动作提名+提名特征→提名评估模块(MLP 多层感知机,对每个提名的置信度分数进行估计,置信度分数越高说明该时序提名与真值的重叠 IoU 应当越高)

3. 非极大化抑制

采用了 soft-nms 算法来通过降低分数的方式来抑制重叠的结果

最终效果

Table 1: Comparison between our method with other state-of-the-art proposal generation methods on validation set of ActivityNet-1.3 in terms of AR@AN and AUC.

Method	Zhao et al. [24]	Dai et al. [42]	Yao et al. [43]	Lin et al. [39]	BSN
AR@100 (val)	63.52	-	-	73.01	74.16
AUC (val)	53.02	59.58	63.12	64.40	66.17
AUC (test)	-	61.56	64.18	64.80	66.26

Table 2: Comparison between our method with other state-of-the-art proposal generation methods on THUMOS14 in terms of AR@AN.

Feature	Method	@50	@100	@200	@500	@1000
C3D	DAPs [5]	13.56	23.83	33.96	49.29	57.64
	SCNN-prop [7]	17.22	26.17	37.01	51.57	58.20
	SST [3]	19.90	28.36	37.90	51.58	60.27
	TURN [6]	19.63	27.96	38.34	53.52	60.75
	BSN + Greedy-NMS	27.19	35.38	43.61	53.77	59.50
C3D	BSN + Soft-NMS	29.58	37.38	45.55	54.67	59.48
2-Stream	TAG [24]	18.55	29.00	39.61	-	-
	Flow	21.86	31.89	43.02	57.63	64.17
2-Stream	BSN + Greedy-NMS	35.41	43.55	52.23	61.35	65.10
2-Stream	BSN + Soft-NMS	37.46	46.06	53.21	60.54	64.52

