

2019 年春-第 2 周

本周基于 RGB 图像的 SSN 网络的训练结果在 iou 为 0.5 的条件下 MAP 为 20.15，符合作者给定的结果，下面记录 SSN 实现过程中的细节。

mAP@0.5IoU (%)	RGB	Flow	RGB+Flow
BNInception	16.18	22.50	27.36
BNInception (Kinetics Pretrained)	21.31	27.93	32.50
InceptionV3	18.28	23.30	28.00 (29.8*)
InceptionV3 (Kinetics Pretrained)	22.12	30.51	33.15 (34.3*)

Detection Performance on thumos14-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+										
IoU thresh	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	Average
-----+-----+-----+-----+-----+-----+-----+-----+-----+										
mean AP	0.4867	0.4401	0.3781	0.2927	0.2015	0.1229	0.0699	0.0305	0.0059	0.2253
-----+-----+-----+-----+-----+-----+-----+-----+-----+										

具体细节：

TAG:

第一步：生成一系列滑动窗口

```
local_boxes = [(i, i + t_span) for i in np.arange(0, duration, step)]
```

生成 t_span 间距的 proposal

t_span 从 2^0 到 2^7 循环

```
其中 span = t_span * time_step 注意 time_step = 1  
step = int(np.ceil(span * (1 - overlap))) 注意 overlap = 0.7
```

然后对这些 proposal 添加一个必须大于 1 秒的约束筛选掉一部分

第二步：训练二分类网络

网络结构：BNInception 输出为 1000 维，在这个基础上添加一个 fc，把输出变成 2 维，分别代表 proposal 对前景和背景的得分

训练：fg: bg 为 1:3

测试：fg: bg 为 1:1

第三步：用二分类网络对第一步的 proposal 生成得分

第四步：通过 tag 算法生成输入 STPP 的 proposal

代码角度：

对 proposal 的得分进行 softmax 归一化使得分在 (0, 1)，然后进行多维高斯滤波。

thresh=[0.01, 0.05, 0.1, .15, 0.25, .4, .5, .6, .7, .8, .9, .95,]

返回所有归一化得分大于 thresh 的 proposal，此时这些 proposal 之间还有重复

tol_lst = [0.05, .1, .2, .3, .4, .5, .6, 0.8, 1.0] 信号下降的快慢

找出正负样本交替处，因为一开始的 proposal 近似为等距采样，所以会因为与 GT 的 iou 而有正负标签的过渡

proposal: up[0 73 124 278 759]; down[66 95 268 748 874] 从两个列表中组合，最后应用 0.9 的 nms

论文角度：

从一个种子盆地开始，并连续吸收随后的盆地，直到盆地持续时间超过总持续时间（即从第一个盆地开始到最后一个盆地结束）的部分下降一定的阈值 τ 。直观来说就是水填满了后在满的基础上下降 τ 。

然后将吸收的盆地和它们之间的空白区域分组以形成单个 proposal。

将每个盆地视为种子并执行分组程序以获得一组表示的提议 $G'(\tau, \gamma)$ 。用步长为 0.05 统一采样 τ , $\gamma \in (0, 1)$ 。这两个阈值的组合导致多组 regions。

最后将非最大抑制应用于具有 IoU 阈值 0.95 的并集，过滤掉高度重叠的 proposal。