SiamRPN

**Abstract:** 提出了一种SiamRPN框架, 通过在目标跟踪算法中引入RPN机制, 在兼顾跟踪速度的同时提升了跟踪的精度.

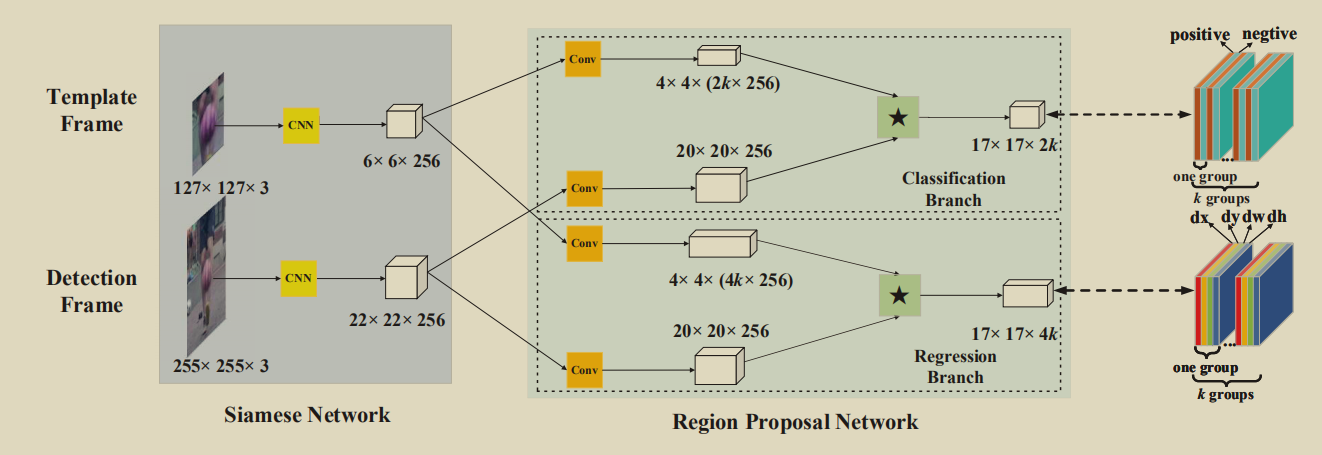


图1 RPN框架图

**Detail:** 1. 基于目标框(bounding-box)预定义几种不同尺寸和宽高比的锚框(假设共k种), 计算锚框与目标框的偏移量(dx,dy,dw,dh)和IOU, 确定分类label(IOU>0.6为目标, IOU<0.3为背景, 其余点不计入loss). 2. 如上图通过孪生网络得到目标(127x127)与搜索区域(255x255)的特征, 同时输入RPN网络得到2k(目标,背景)通道的分类结果和4k通道(dx,dy,dw,dh)的回归结果. 3. 计算loss并反向传播训练网络

**Test:** 得到2k通道的分类结果后, 取最大值坐标, 根据坐标得到对应锚框的dx,dy,dw,dh. 计算得出当前帧的目标框.

DaSiamRPN

**Abstract:** 在SiamRPN算法的训练过程中引入coco,det,yt-bb等新的数据集. 在数据集中加入了难例负样本. 并提供了一种长时跟踪方法.

**Detail:** 1. 引入新的数据集, 极大提升了数据的丰富程度. 2. 加入了难例负样本, 原因是搜索区域由目标框确定, 导致在搜索区域中背景比较单调,在网络中很容易训练, 因此需要加入一些难例负样本提升网络的稳定性. 加入难例负样本的方法是取不同视频的帧组成目标与搜索区域的图像对, 因此在搜索区域中必然存在另一个视频的目标作为背景, 在训练过程可以提升区分背景的能力.

**Test:** (长时跟踪)根据当前帧分类图后, 选取top-k个锚框, 分别计算锚框对应的目标框, 其中top1结果为目标结果, top2-k结果对应相似目标的干扰结果. 以这k个目标框为基础分别在下一帧图像中跟踪, 得到k个目标的分类图, 通过目标2-k分类图对目标1分类图修正, 抑制干扰项的影响.(长时跟踪方式对速度影响较大)

SiamRPN++

**Abstract:** 论证了网络的平移不变性对跟踪器的影响, 打破网络深度的限制, 将更深的网络resnet50应用到目标跟踪中, 提供了新的相关滤波方式depth-wise.

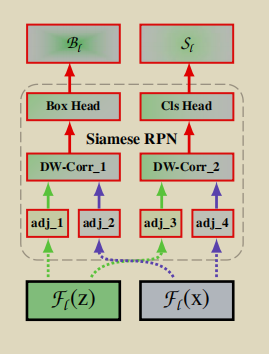
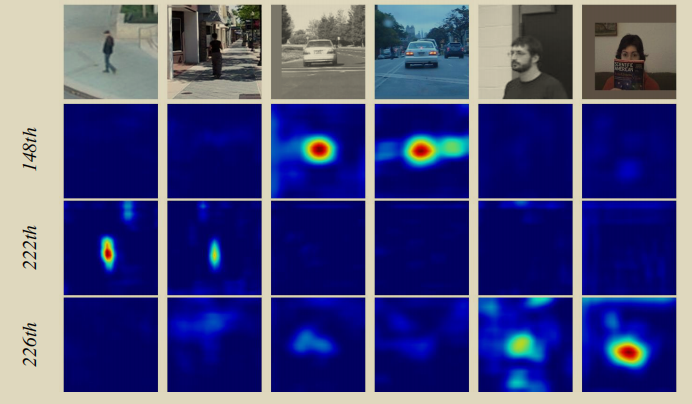
 

图2 depth-wise 图3 Correlation filter result

**Detail:** 1. 平移不变性, 在现有的alex,resnet等主干网络中通常加入了padding, 因此在卷积计算过程中, 边缘值的计算中会引入0, 而远离边缘的值完全通过像素点的值计算得到, 因此在提取模板特征时边界点其实是引入0计算得到, 而当目标位于搜索区域中心时,得到的特征边界点并没有引入0, 即便是相同的图像块,在模板和搜索区域中得到的特征是不同的(即平移会产生变化). 同时, 随着网络深度增加, padding的影响会向内扩散. 为了解决平移不变性,RPN++首先取消了深层卷积的padding. 其次, 在提取模板特征后, 仅采取中间7x7范围的特征作为目标的特征模板,因为越靠近图像中间的特征受padding影响越小. 最后在训练过程中对输入的图像对做随机偏移,进一步缓解平移不变性的破坏. 2. depth-wise, 一种通道对通道的卷积, 如图3, 作者发现在相关滤波过程中,网络输出的不同通道的特征适用于不同目标的匹配, 因此,采用depth-wise这种通道对通道的卷积来计算不同通道的响应程度,最后通过一层卷积来得到最终的相应图.