候选区域ROI：

1. 滑动窗口。
2. 选择性搜索SS。可以大大减少窗口数量。
3. RPN

## RCNN

**RCNN**是把SS算法得到的2000+个ROI缩放为固定大小，然后作为输入。2000+个ROI是有很多部分是重叠的，重叠区域的特征会被多次计算。

**Fast-RCNN**是把整张图作为输入，在CNN网络的特征图上计算（候选区域算法如SS）ROI（特征图块），不会重复提取特征， ROI 池化层(不同于一般的池化层，原理类似)将特征图块转换为固定的大小，并馈送到全连接层进行分类(sofamax)和定位.

**Faster-RCNN** 和fast-RCNN一样，只是采用RPN网络提取 ROI 。RPN是浅层的CNN带有全连接层，用前一个CNN提取的特征作为输入，用到了anchor，输出是候选区域与是否有目标得分,选择目标得分高的候选区域。

**RFPN**结构和faster-RNN一样，ROIpooling被改进成基于位置敏感的。

分类需要特征具有平移不变性，检测（目标位置）要求对目标的平移做出准确响应。卷积操作是具有平移不变性的，但是插入ROI pooling之后，后面的网络结构就不再具备平移不变性。而RFCN的position sensitive score map吧目标的位置信息也融合到ROI pooling.

### YOLO

YOLO v1的优点：

(1) 检测物体非常快：因为没有复杂的检测流程，只需要将图像输入到神经网络就可以得到检测结果，YOLO可以非常快的完成物体检测任务。标准版本的YOLO在Titan X 的 GPU 上能达到45 FPS。更快的Fast YOLO检测速度可以达到155 FPS。而且，YOLO的mAP是之前其他实时物体检测系统的两倍以上。

(2) YOLO可以很好的避免背景错误，产生false positives：不像其他物体检测系统使用了滑窗或region proposal，分类器只能得到图像的局部信息。YOLO在训练和测试时都能够看到一整张图像的信息，因此YOLO在检测物体时能很好的利用上下文信息，从而不容易在背景上预测出错误的物体信息。和Fast-R-CNN相比，YOLO的背景错误不到Fast-R-CNN的一半。

(3) YOLO可以学到物体的泛化特征：当YOLO在自然图像上做训练，在艺术作品上做测试时，YOLO表现的性能比DPM、R-CNN等之前的物体检测系统要好很多。因为YOLO可以学习到高度泛化的特征，从而迁移到其他领域。

尽管YOLO有这些优点，它也有一些缺点：

(1) YOLO的物体检测精度低于其他state-of-the-art的物体检测系统。

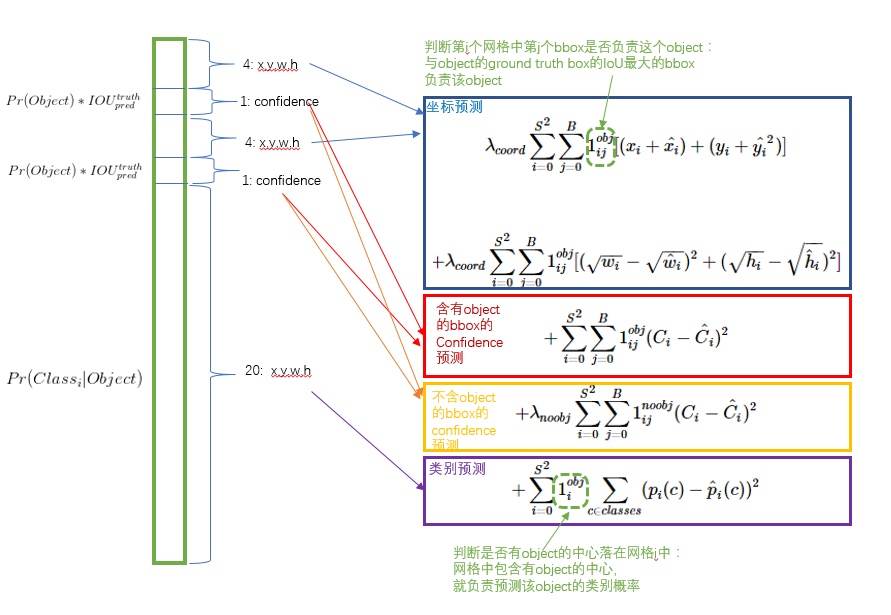
(2) YOLO容易产生物体的定位错误。

(3) YOLO对小物体的检测效果不好（尤其是密集的小物体，因为一个栅格只能预测2个物体）。

1. 网络结：24卷积层，2全连接层
2. 训练

训练集：imageNet1000 和 Pascal VOC 20

损失函数：



Yolo2改进：

1. Batch normalization: [https://www.cnblogs.com/zhengmingli/p/8031690.html](%20https:/www.cnblogs.com/zhengmingli/p/8031690.html)
2. 高分辨率分类器。
3. Anchor box（实现对训练集Kmeans 聚类得到合适大下的boxs）
4. **Multi-Scale Training。**
5. 细粒度特征(fine grain features)。（类似Resnet）
6. 网络结构为Darknet19.

Yolo3改进：

1. 提出了一种联合训练方法，能够容许同时使用目标检测数据集和分类数据集。使用有标记的检测数据集精确定位，使用分类数据增加类别和鲁棒性
2. 模型比之前的模型复杂了不少，可以通过改变模型结构的大小来权衡速度与精度
3. 多尺度预测