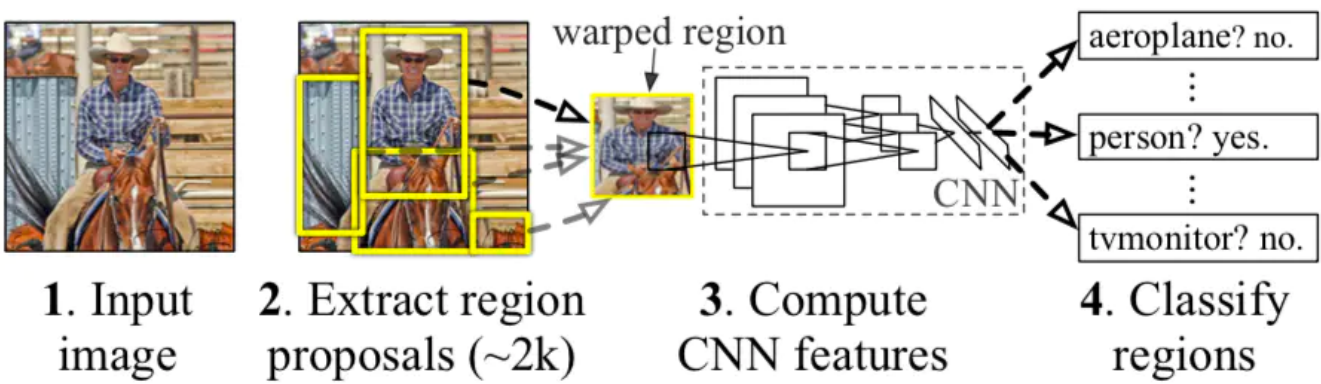
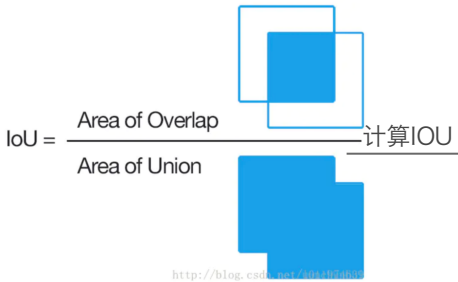


R-CNN: Regions with CNN features

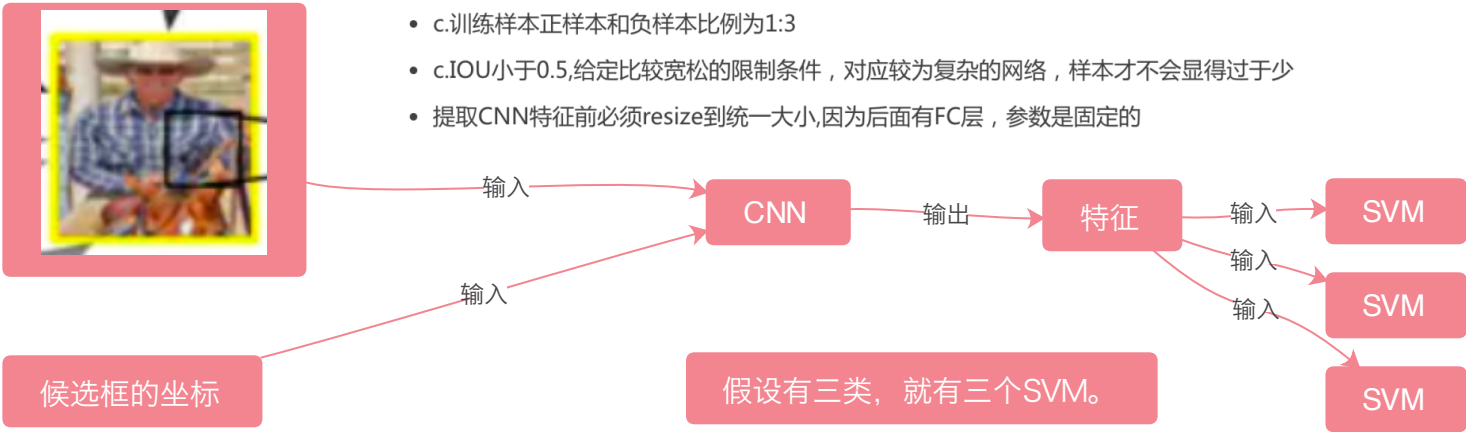


在 2. extract region proposals处使用



selective search (SS) 的作用：生成候选框，没有预测框。

- a.在数据集ILSVRC2012 dataset上预训练好的分类网络(AlexNet)直接拿来用，将大约2k个图片一个个送到分类网络（backbone），只拿了特征提取层，扔掉了最后的1*1000的分类层。
- b.在真实数据集上finetune微调，主要更新后两层全连接dense层
- c.训练样本正样本和负样本比例为1:3
- c.IOU小于0.5,给定比较宽松的限制条件，对应较为复杂的网络，样本才不会显得过于少
- 提取CNN特征前必须resize到统一大小,因为后面有FC层，参数是固定的



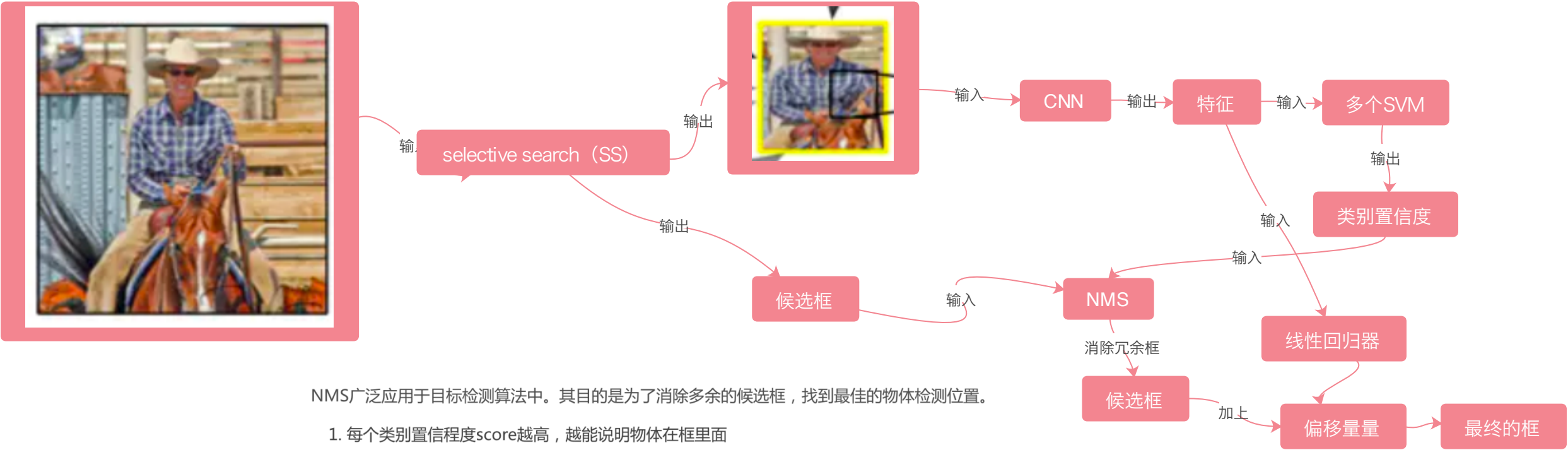
训练是分阶段的，不是端到端，线性回归器是修正SS的框

IOU大于0.6 的候选框与真实框的差距（偏移量）



- 给每个样本做分类
- 正样本：Groud True
- 负样本：IOU < 0.3[others ignored]
- Q：和AlexNet网络训练相比，svm分类时IOU的阈值threshold为什么不一样？A：训练CNN需要更多数据样本，更多数据样本不一定质量高，svm分类时需要质量高数据样本
- Q：为什么用SVM,为什么不直接用AlexNet最后的softmax？A：AlexNet专注于分类，不好用少数样本来做定位。数据多，质量差；数据少，质量高，数据少时用SVM相对好一些

RCNN 测试阶段



NMS广泛应用于目标检测算法中。其目的是为了消除多余的候选框，找到最佳的物体检测位置。

1. 每个类别置信程度score越高，越能说明物体在框里面
2. 取score值最高的一个框，将其他的框和他的iou越大，越可能是同一个物体，所以去掉其他和该框iou高于某一阈值候选框