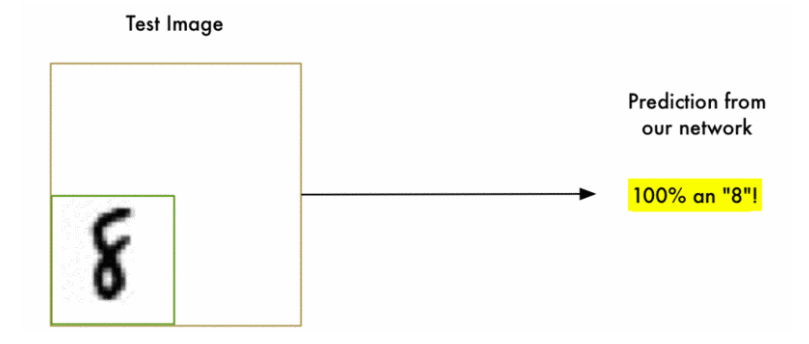
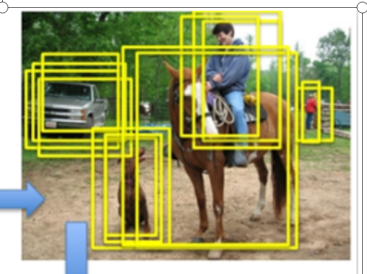
### **滑动窗口法**

简单来说，就是那个框，在图片上移动，就看框里面有没有你要找的目标，有就是了。

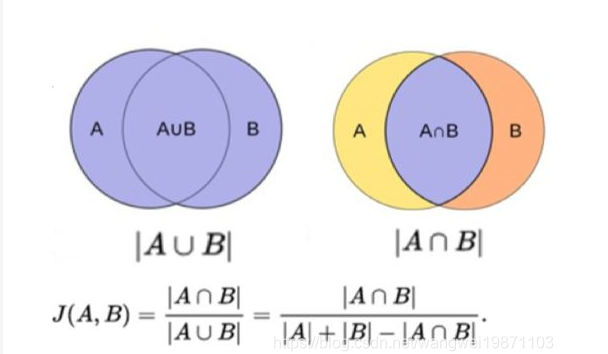
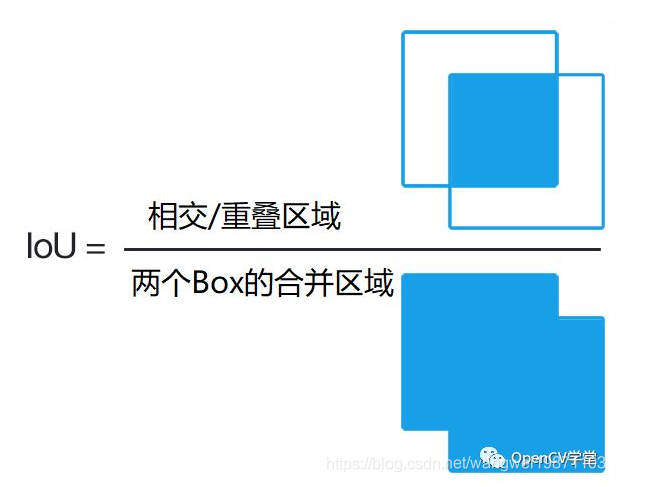


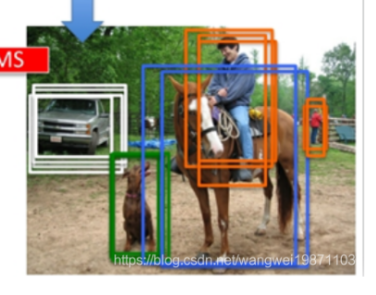
其实就是用不同大小的图，在图片上从左到右，从上到下的搜索，找到你要的目标，因为目标的大小不一样，所以很多时候你看你需要很多个尺寸的框来搜索，所以基本是属于穷举法啦，找到适合的为止。





## **交并比(IoU)**

简单理解为交集和并集的比值，交集和并集的概念应该知道吧，然后用在窗口上，也就是两个窗口相交的区域和并起来的区域的比值，反应的是两个窗口的重合度。  
喜欢看公式的：  
  
喜欢直观的：  


简单理解就是把不是最大的给扔了，具体就是说把很多重叠的框按照分类分数排个序，取出最高的那个，把一些和他交并比大的给抛弃了。

剩下的一些继续做非极大抑制,直到全部处理完，最后留下的既是分类分数高的，又是框比较合适的窗口

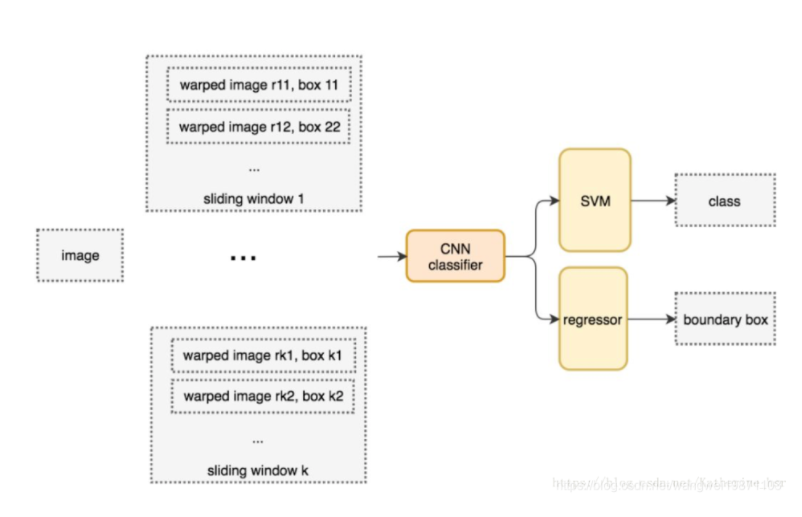


1.我们把要找的目标先训练好一个分类器，比如上面的狗，人，骑车，马，训练一个4分类器。比如用SVM或者CNN训练一个分类器，输入一个固定大小图片，输出类别概率。

2.用很多不同大小的窗口进行移动，把图片分成很多个小窗口，然后把每个窗口喂给分类器，如果给出较高的分类分数，那就说明这个窗口有物体，那就留着，这样处理完所有窗口。

3.处理完的窗口当然会有很多重叠的啦，我们当然希望处理一些比较像我们标注的窗口啦，然后我们就使用 非极大抑制，留下合适的框，丢弃其他的框。

4.得到各种窗口和对应的类别，此时可以拿标注的数据框来进行回归，让窗口更加精确点，当然不回归也可以，可能不是很精确啦。



**缺点：**

1.如果考虑到有不同大小的物体，那需要不同的尺寸的框来进行处理，基本是穷举很多尺寸了，不然效果可能不好，所以效率非常低下。

2.很多个窗口的情况下，执行比较复杂的分类器，可能又会很耗时。

3.如果分类器的输入图片尺寸是固定的，还要把不同的窗口大小缩放到固定尺寸，又要耗时费力，可能还会对分类产生影响。