

1.目标侦测和目标分割的区别有哪些？

目标侦测属于回归+分类问题，检测识别每个实例，数据标注仅需要标注矩形框，且计算量较小，常用模型有RCNN，YOLO，SSD

目标分割属于像素点分类问题，更强调目标的精确轮廓，数据标注需要详细的像素级标注，计算量大，常用模型有UNet和U2Net，DeepLab，Mask RCNN

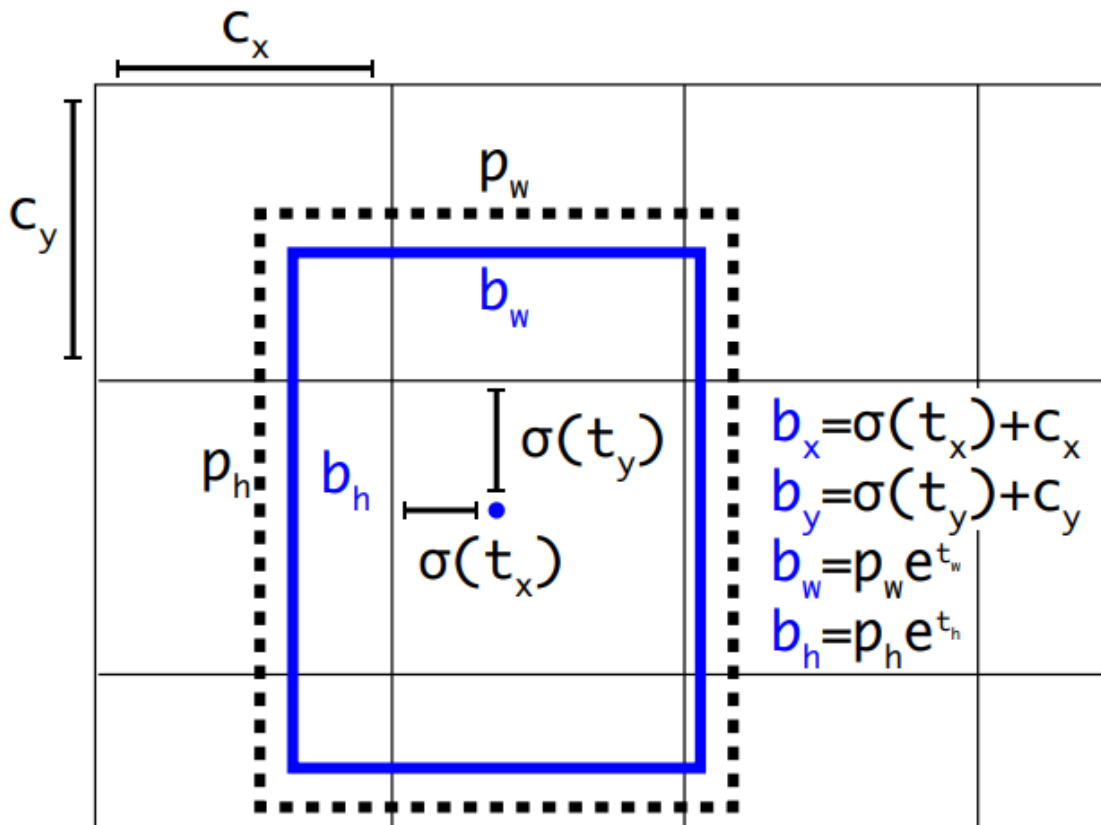
2.传统机器算法和深度学习算法的区别是什么？

传统机器算法需要人工识别特征，算法更加透明，对数据量和算力要求低，适用于结构化数据。

深度学习属于端到端模型，自动学习特征，对数据和算力要求高，适用于图像语音等非结构化数据。

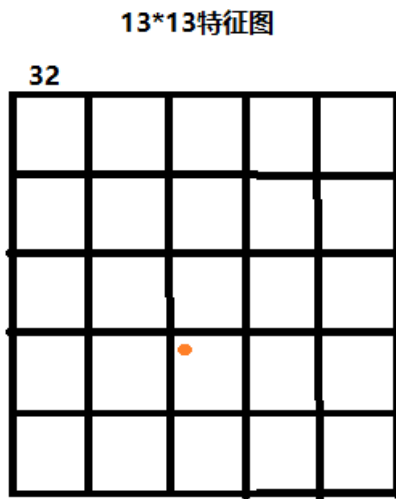
3.yolov3如何确定目标位置

- 首先yolov3会设置3种建议框（anchor box）
- yolov3的侦测头针对每个anchor box预测出预测框（bounding box）的4个值 t_x ， t_y ， t_w ， t_h
- 对预测框进行解码



- 解码后得到预测框的宽高 b_w ， b_h ，和中心点坐标 b_x 和 b_y
- 对同种类别使用NMS进行筛选，保留置信度高的框，最终得到精确的目标位置

4.yolov3如何反算出中心点位置



中心点偏移计算:

$$\frac{cx}{feature_size} = idx + offset_x$$

$$\frac{cy}{feature_size} = idx + offset_y$$

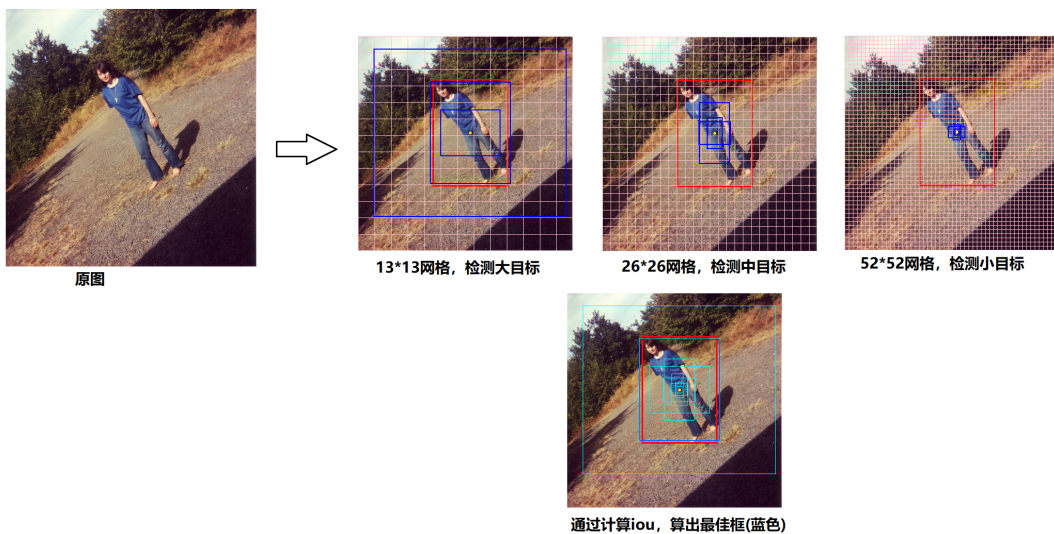
反算:

$$cx = (index_x + offset_x) * feature_size$$

$$cy = (index_y + offset_y) * feature_size$$

5.yolov3建议框的个数和尺寸怎么确定

yolov3建议框共9种 (3种尺寸, 每种尺寸有3个形状)。一般需要对数据集做k-means聚类得到9种建议框

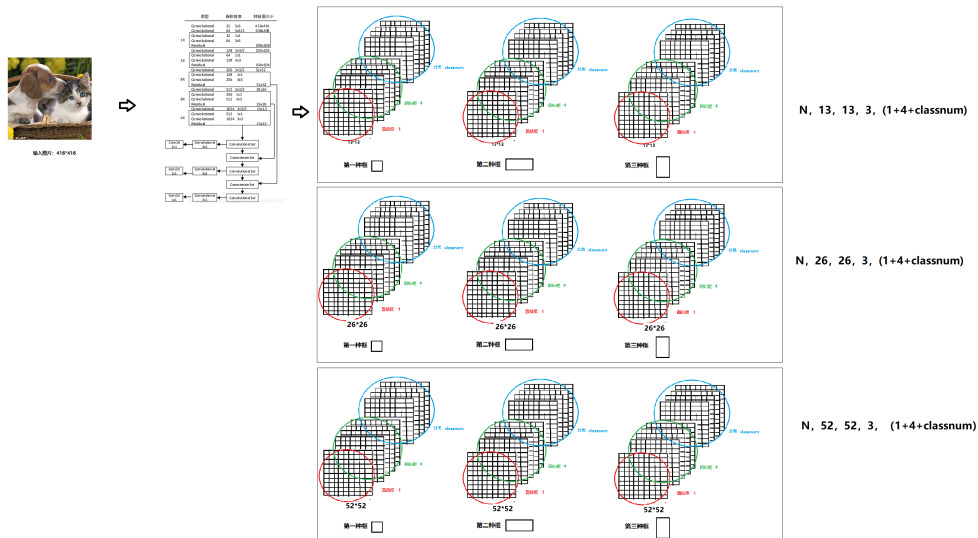


6.yolov3输出结果N45HW是什么意思

yolov3的输出是(N, 3*15, H, W), 其中N为批次, 15表示: (1个置信度, 4个偏移量, 10个类别的概率)。H, W有3种尺寸: 13*13, 26*26, 52*52

因此yolov3最终输出形式为

((N, 45, 13, 13), (N, 45, 26, 26), (N, 45, 52, 52))



7.验证手写数字识别卷积的空间依赖性

原始测试集测试得分：

```
Test Score: 0.9828916139240507
Test Score: 0.9835838607594937
Test Score: 0.9835838607594937
Test Score: 0.9835838607594937
Test Score: 0.9835838607594937
Test Score: 0.9835838607594937
```

通过transforms变换数据集

```
test_transforms = transforms.Compose([
    transforms.RandomHorizontalFlip(p=0.5), # 随机水平翻转
    transforms.RandomVerticalFlip(p=0.5), # 随机垂直翻转
    transforms.RandomRotation(degrees=30), # 随机旋转
    transforms.RandomAffine(degrees=0, translate=(0.1, 0.1)), # 随机平移
    transforms.Resize(28),
    transforms.ToTensor()
])
```

测试得分：

```
Test Score: 0.4365110759493671
Test Score: 0.43829113924050633
Test Score: 0.4410601265822785
Test Score: 0.4411590189873418
Test Score: 0.44966376582278483
```

手写效果测试：

	预测的结果是：3
	预测的结果是：7
	预测的结果是：2

结论：每个卷积核只连接输入特征图的局部区域，识别特定模式的存在。卷积对特征的空间位置是敏感的。