

1. WBF难点

1.1 标签映射

在项目中使用多个模型的时候，会遇到每个模型的标签不一致的问题，所以在加载模型的时候要记录每个模型的标签，同时要把所有标签去重汇总到一个列表。每遍历一个模型的推理结果时，由于模型的类别输出时index，我们通过该index获得该模型的类别标签，然后通过该模型的类别标签转换成汇总列表对应的index，以达到标签映射的目的。（当某组只有一个box 不进入WBF）

1.2 给类别分组

WBF的工作原理：

- 对不同模型推理得到的box，score根据图片的尺寸进行归一化
- 把这些归一化的数据传给WBF
- WBF会根据IOU匹配，把需要融合的框归到一类（需要融合是根据位置来融合的，一个左上角的框不可能和一个右下角不相交的框融合）
- WBF根据融合公式融合对box和score加权融合，这个权重可以自定义，默认是1:1
 - 此时多个IOU匹配得到的框融合成一个，得到最优的一个解
 - 有两个模型每个模型有一个输出结果，这两个结果box不相交，这种情况传进WBF，对这些框进行分类，因为IOU不匹配，所以得到的是两类，但是每类只有1个框，此时要根据权重比对该box的score减分（如果权重比为1:1 只有一个框，框的坐标不变，score减半）
 - 三个模型每个模型一个框，两个模型的结果IOU匹配，另外一个不匹配，此时分成两类。第一个类只有两个框，但是有三个模型作为输入，假如三个模型的权重比为（1:1:1）此时 这类框返回的分数只能是两个框的分数和的 2/3（因为三个模型作为输入，这个类别只有两个框），同理另一个框只有1/3
- 对WBF得到的结果复原

2. ResNet消误检难点

2.1训练ResNet模型

训练图片前处理

1.训练的resnet模型是对目标检测的结果进行过滤的，所以用来训练的图片只能是目标检测的目标区域，不能包含背景信息，我们对数据库的图片进行目标截取。此时我们在resnet中添加了一类（normal类），我们把项目上收集到的误识别的目标（模型识别到缺陷的类别，但不是缺陷的目标）放到normal类中，同时把一些截取得到的背景信息加入normal类中，（这样做的目的当模型存在误识别的时候，把误识别的区域作为resnet的输入，resnet输出类别时输出的是normal，此时检测到的类别和resnet的类别不一样，此时达到消误检的目的）

2.2resnet的使用

在得到目标检测模型的推理box后，截取box区域的图像作为resnet的输入，此时resnet会得到Top-k个结果。当目标检测的结果在前N个结果内（ $N \leq k$ ）时，则说明匹配成功，输出该类，否则过滤掉。

3.OCR识别过滤目标

3.1OCR识别

paddle的ocr识别三个模型

- 文字位置识别
- 文字内容识别
- 文字方向识别
- 占用（300M显存左右） yolox（1000M显存） resnet（不到50M）

3.2OCR识别的应用

- 检测图片获取文字位置矩形
- 便利每个推理box的结果 当是指定类别，且文字框不为空时，计算误检box与文字框的IOU（交集面积/缺陷类别面积）如果大于阈值，则判断为误检，过滤该box