

### 3.3 目标检测性能指标

目标检测模型的性能指标分为速度指标和准确度指标。速度指标通常用每秒推断帧数 FPS (Frames Per Second) 衡量, 但受硬件影响较大, 详见第 6 章, 本章重点研究准确度指标<sup>[94]</sup>。

目标检测输入图像, 输出图像中各个目标预测框的矩形坐标及各类别预测置信度 Conf。

采用交并比 IOU 衡量预测框和标注框的重合程度, 即预测框的定位是否准确。

$$IOU = \frac{area(B_p \cap B_{gt})}{area(B_p \cup B_{gt})} \quad (3-5)$$

式中,  $B_p$  为预测框,  $B_{gt}$  为标注框。

根据与标注框的关系, 可将某一预测框划分为下式四类中的某一类:

$$TP: \{Conf > P_{thresh} \text{ 且 } IOU > IOU_{thresh}\} \quad (3-6)$$

$$FP: \{Conf > P_{thresh} \text{ 且 } IOU < IOU_{thresh}\} \quad (3-7)$$

$$FN: \{Conf < P_{thresh} \text{ 且 } IOU > IOU_{thresh}\} \quad (3-8)$$

$$TN: \{Conf < P_{thresh} \text{ 且 } IOU < IOU_{thresh}\} \quad (3-9)$$

式中,  $IOU_{thresh}$  为 0-1 之间的常数, 需人工指定。

对于某一特定类别, TP、FP、FN、TN 四种预测框的个数构成混淆矩阵 (Confusion Matrix), 例如波磨类别的混淆矩阵如表 3.1 所示。

表 3.1 混淆矩阵示例 (以波磨类别为例)

预测框个数	预测为波磨	预测为其它类别
真实为波磨	TP 预测框个数	FN 预测框个数
真实为其它类别	FP 预测框个数	TN 预测框个数

进一步定义以下参数:

(1) Precision (查准率) 是指所有预测框中预测正确的比例, 反映了模型“不把背景冤枉为目标”的准确性。

$$Precision = \frac{TP}{TP+FP} = \frac{TP}{\text{所有预测框}} \quad (3-10)$$

(2) Recall (查全率、敏感性、召回率) 是指所有标注框中被正确预测的比例, 反映了模型“不把目标放过为背景”的敏感性。

$$Recall = \frac{TP}{TP+FN} = \frac{TP}{\text{所有标注框}} \quad (3-11)$$

(3) Average Precision (平均精度, 简称 AP): 将  $P_{threshold}$  阈值从 0 到 1 变化, 计算每个  $P_{threshold}$  阈值对应的 Precision 和 Recall, 绘制成某类别的 PR 性能曲线, 其围成的面积为该类别的 AP<sup>[94]</sup>。取所有类别的 AP 和不同的  $IOU_{thresh}$ , 可分别计算  $mAP@0.5$  和  $mAP@0.5:0.95$ 。  $mAP@0.5$  为  $IOU_{thresh}$  取 0.5 时, 各类别 AP 的平均值。  $mAP@0.5:0.95$  为  $IOU_{thresh}$  分别取以 0.05 为步长, 从 0.5 增大至 0.95 的 10 个数时, 各类别 AP 的平均值。如下式所示, 式 (3-14) 中,  $j$  分别取 0.5、0.55、0.6、0.65、0.7、0.75、0.8、0.85、0.9、0.95,  $N$  为类别总数 13。

$$AP = \int_0^1 P_{thresh}(r) dr \quad (3-12)$$

$$mAP@0.5 = \frac{1}{N} \sum_{i=1}^N AP_i(IOU_{thresh} = 0.5) \quad (3-13)$$

$$mAP@0.5:0.95 = \frac{1}{N} \sum_{i=1}^N \sum_j AP_i(IOU_{thresh} = j) \quad (3-14)$$

本文以 Precision、Recall、AP、 $mAP@0.5$ 、 $mAP@0.5:0.95$  作为性能指标。

### 3.4 训练配置

本文计算环境基于 Featurize 云算力平台的 Ubuntu 操作系统和 Pytorch 深度学习框架搭建, CPU 为 Intel Core i5, 显卡为 GeForce RTX 3080, CUDA 版本 11.1。在 Subway-1094 数据集上基于 YOLOV5\_V4.0 的 S 预训练模型迁移学习训练得到基准模型, 训练超参数如表 3.2 所示。

表 3.2 基准模型训练超参数

超参数	值	超参数	值
初始学习率	0.01	优化器	带动量的 SGD
学习率折减系数	0.20	训练轮次	50、300
输入图像长宽尺寸	640	初始动量	0.8
L2 权重衰减系数	0.0005	动量系数	0.937
IOU 阈值	0.20	batch size	8
色调增强	0.015	饱和度增强	0.7
明度增强	0.40	平移增强	0.1
缩放增强	0.5	马赛克拼贴增强	1
旋转、翻转增强	0	剪切、透视增强	0