



整车先进设计制造技术全国重点实验室  
State Key Laboratory of Advanced Design and Manufacturing Technology for Vehicle



湖南大学  
HUNAN UNIVERSITY

# 计算机视觉与机器学习

## YOLO目标检测原理与指标分析

报告人：刘天适



模型衡量指标



mAP值的计算



YOLO输出结果分析

## 机器学习中常见模型衡量指标：

### ① 精确度 precision —— “找的对”

让模型的现有预测结果尽可能**不出错**

( 宁愿漏检，尽可能不使现有的预测有错 )

### ② 召回率 recall —— “找出来”

让模型预测到**所有**想被预测到的样本

( 宁可错，不能漏 )

应用场景：

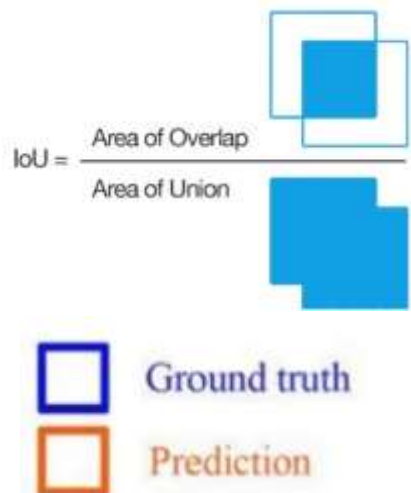
预测地震 —— 不能接受漏报 —— 提升召回率

银行人脸支付 —— 不能接受误检 —— 提升精确度

往往矛盾 一个高另一个低



IOU - 交并比 - 越接近1越好



### ③ F1-Score，定义为精确率和召回率的调和平均数。

值越接近 1，说明模型性能越好。

$$F - score = \frac{2}{1/precision + 1/recall}$$

### ④准确率Accuracy，正确预测的样本数占总预测样本数的比值。

## 评价指标的基础：混淆矩阵 - confusion matrix

对于二分类问题，每一条数据要么预测结果正确，表示为1；要么预测错误，表示为0。

而事物本身也是被分为0（负样本）和1（正样本）两类。

用矩阵来表示即——

		实际	
		1	0
预测	1	TP	FP
	0	FN	TN

T/F: True, False 判断的正误

P/N: Positive, Negative 判断结果的正例、负例。正例即目标样本。

T/F: True, False 判断的正误

P/N: Positive, Negative 判断结果的正例、负例。正例即目标样本

已知条件：班级总人数100人，其中男生80人，女生20人。

目标：找出所有的女生。

结果：从班级中选择了50人，其中20人是女生，还错误的把30名男生挑选出来了。

T: 判断正确

F: 判断错误

P: 判断的结果是女生

N: 判断的结果是男生

TP: 20      判断正确&判断为女生  
TN: 50      判断正确&判断为男生  
FP: 30      判断错误前提下，判断结果为女生（实际为男生）  
FN: 0      判断错误前提下，判断结果为男生（实际为女生）

判断为男生即不选择，即漏检的女生

$$Precision = \frac{TP}{TP + FP}$$
$$Recall = \frac{TP}{TP + FN}$$

TP+FP: 预测结果为目标样本的总数，  
(T+F)P，包括预测正确的和错误的。

TP+FN: 实际目标样本数。  
包括检出和漏检的目标样本。

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

区分accuracy & precision:

accuracy指的是正确预测的样本数占总预测样本数的比值，

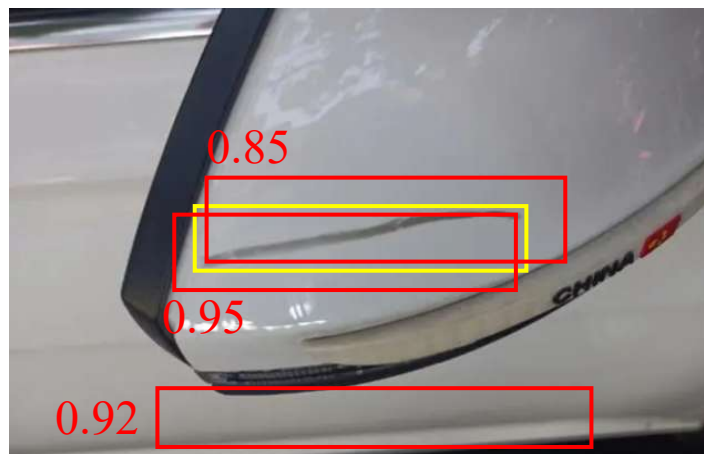
它不考虑预测的样本是正例还是负例。

而precision指的是正确预测的正样本数占有所有预测为正样本的数量的比值。

precision只关注预测为正样本的部分，而accuracy考虑全部样本。



TP(True Positive): 真阳性  
此处TP=1

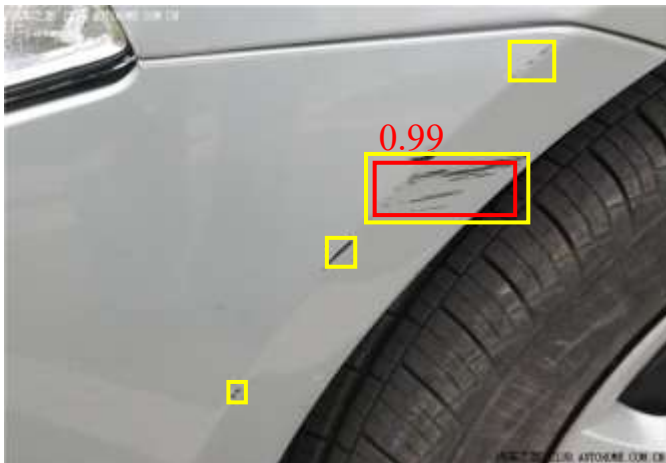


FP(False Positive): 假阳性  
此处FP=2  
(GT: Ground Truth, 真实标注)  
【一个GT只参与一次比较】  
注意0.85, 0.92FP的原因不同!

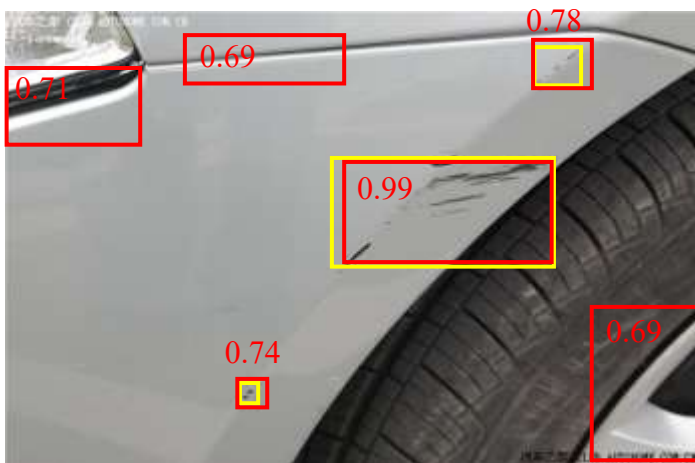


FN(False Negative): 假阴性  
没有检测到的GT的数量, 漏检  
此处FN=1





Precision=1



Recall=1

这样的预测好吗？  
如何统一精度和全度？

引入性能综合评价指标mAP 的概念

性能综合评价指标mAP

所有类别取平均

AP

(Average Precision, 平均精度)  
每个类别

P-R曲线下面积

精确率-召回率曲线  
(P-R曲线)

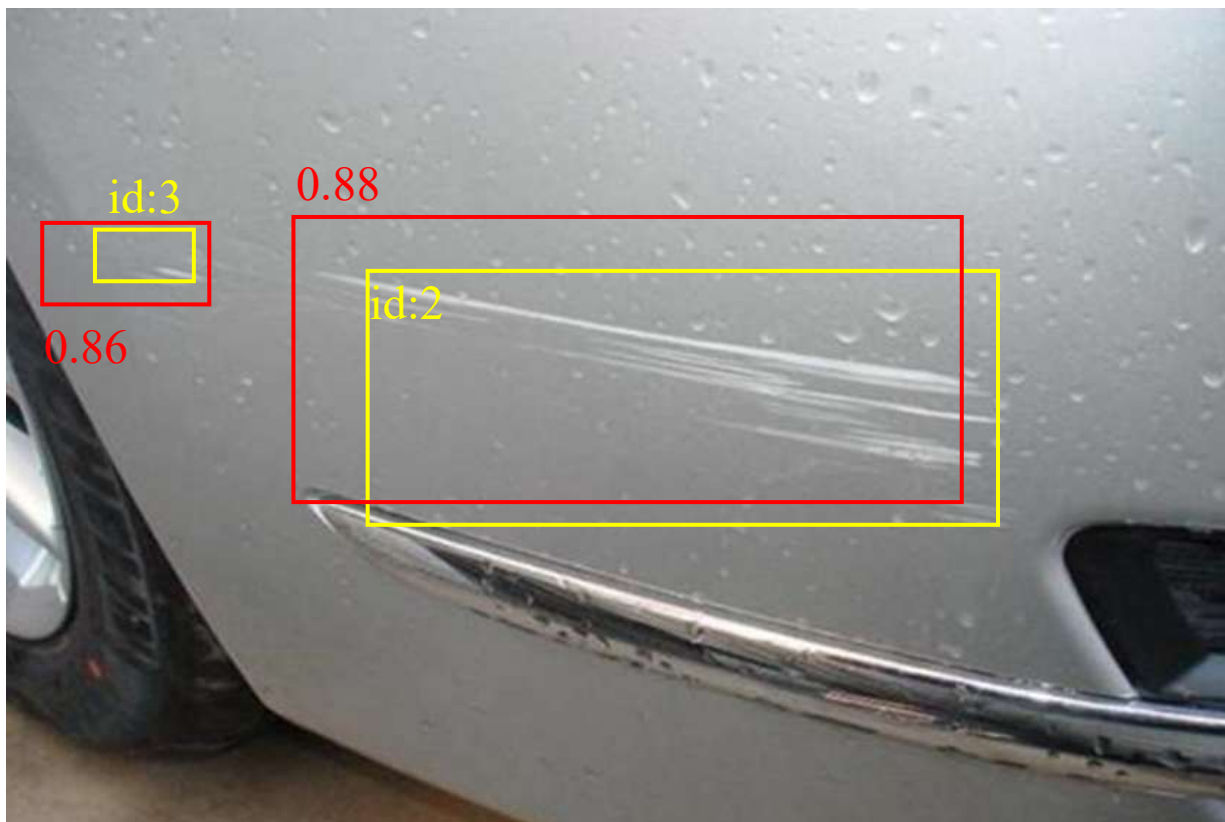




此轮判断以IOU=0.5为基准

GT ID	Confidence	OB(IOU=0.5)
1	0.96	Y
\	0.84	N
\	0.76	N

置信度从大到小排列



GT ID	Confidence	OB(IOU=0.5)
1	0.96	Y
2	0.88	Y
3	0.86	Y
\	0.84	N
\	0.76	N



GT ID	Confidence	OB(IOU=0.5)
1	0.96	Y
2	0.88	Y
3	0.86	Y
4	0.85	Y
\	0.84	N
\	0.76	N
\	0.68	N



GT ID	Confidence	OB(IOU=0.5)
1	0.96	Y
2	0.88	Y
3	0.86	Y
4	0.85	Y
\	0.84	N
5	0.80	Y
\	0.79	N
\	0.76	N
\	0.68	N

正例个数=5

GT ID	Confidence	OB(IOU=0.5)
1	0.96	Y
2	0.88	Y
3	0.86	Y
4	0.85	Y
\	0.84	N
5	0.80	Y
\	0.79	N
\	0.76	N
\	0.68	N

$$P(\text{Precision}) = TP / (TP + FP)$$

FP 错解

$$R(\text{Recall}) = TP / (TP + FN)$$

FN 漏解

$$P = 1 / (1 + 0) = 1.0$$

$$R = 1 / (1 + 4) = 0.2$$

Precision	Recall
1.0	0.2

GT\_num=5

GT ID	Confidence	OB(IOU=0.5)
1	0.96	Y
2	0.88	Y
3	0.86	Y
4	0.85	Y
\	0.84	N
5	0.80	Y
\	0.79	N
\	0.76	N
\	0.68	N

$$P(\text{Precision}) = TP / (TP + FP)$$

$$R(\text{Recall}) = TP / (TP + FN)$$

$$P = 2 / (2 + 0) = 1.0$$

$$R = 2 / (2 + 3) = 0.4$$

Precision	Recall
1.0	0.2
1.0	0.4

GT\_num=5

GT ID	Confidence	OB(IOU=0.5)
1	0.96	Y
2	0.88	Y
3	0.86	Y
4	0.85	Y
\	0.84	N
5	0.80	Y
\	0.79	N
\	0.76	N
\	0.68	N

$$P(\text{Precision}) = TP / (TP + FP)$$

$$R(\text{Recall}) = TP / (TP + FN)$$

$$P = 3 / (3 + 0) = 1.0$$

$$R = 3 / (3 + 2) = 0.6$$

Precision	Recall
1.0	0.2
1.0	0.4
1.0	0.6



GT\_num=5

GT ID	Confidence	OB(IOU=0.5)
1	0.96	Y
2	0.88	Y
3	0.86	Y
4	0.85	Y
\	0.84	N
5	0.80	Y
\	0.79	N
\	0.76	N
\	0.68	N

$$P(\text{Precision}) = TP / (TP + FP)$$

$$R(\text{Recall}) = TP / (TP + FN)$$

$$P = 4 / (4 + 0) = 1.0$$

$$R = 4 / (4 + 1) = 0.8$$

Precision	Recall
1.0	0.2
1.0	0.4
1.0	0.6
1.0	0.8

GT\_num=5

GT ID	Confidence	OB(IOU=0.5)
1	0.96	Y
2	0.88	Y
3	0.86	Y
4	0.85	Y
\	0.84	N
5	0.80	Y
\	0.79	N
\	0.76	N
\	0.68	N

$$P(\text{Precision}) = TP / (TP + FP)$$

FP 错解

$$R(\text{Recall}) = TP / (TP + FN)$$

FN 漏解

$$P = 4 / (4 + 1) = 0.8$$

$$R = 4 / (4 + 1) = 0.8$$

Precision	Recall
1.0	0.2
1.0	0.4
1.0	0.6
1.0	0.8
0.8	0.8

GT\_num=5

GT ID	Confidence	OB(IOU=0.5)
1	0.96	Y
2	0.88	Y
3	0.86	Y
4	0.85	Y
\	0.84	N
5	0.80	Y
\	0.79	N
\	0.76	N
\	0.68	N

$$P(\text{Precision}) = \text{TP} / (\text{TP} + \text{FP})$$

$$R(\text{Recall}) = \text{TP} / (\text{TP} + \text{FN})$$

$$P = 5 / (5 + 1) = 0.83$$

$$R = 5 / (5 + 0) = 1.0$$

Precision	Recall
1.0	0.2
1.0	0.4
1.0	0.6
1.0	0.8
0.8	0.8
0.83	1.0

GT\_num=5

GT ID	Confidence	OB(IOU=0.5)
1	0.96	Y
2	0.88	Y
3	0.86	Y
4	0.85	Y
\	0.84	N
5	0.80	Y
\	0.79	N
\	0.76	N
\	0.68	N

$$P(\text{Precision}) = TP / (TP + FP)$$

$$R(\text{Recall}) = TP / (TP + FN)$$

$$P = 5 / (5 + 2) = 0.71$$

$$R = 5 / (5 + 0) = 1.0$$

Precision	Recall
1.0	0.2
1.0	0.4
1.0	0.6
1.0	0.8
0.8	0.8
0.83	1.0
0.71	1.0

GT\_num=5

GT ID	Confidence	OB(IOU=0.5)
1	0.96	Y
2	0.88	Y
3	0.86	Y
4	0.85	Y
\	0.84	N
5	0.80	Y
\	0.79	N
\	0.76	N
\	0.68	N

$$P(\text{Precision}) = TP / (TP + FP)$$

$$R(\text{Recall}) = TP / (TP + FN)$$

$$P = 5 / (5 + 3) = 0.63$$

$$R = 5 / (5 + 0) = 1.0$$

Precision	Recall
1.0	0.2
1.0	0.4
1.0	0.6
1.0	0.8
0.8	0.8
0.83	1.0
0.71	1.0
0.63	1.0

GT\_num=5

GT ID	Confidence	OB(IOUS=0.5)
1	0.96	Y
2	0.88	Y
3	0.86	Y
4	0.85	Y
\	0.84	N
5	0.80	Y
\	0.79	N
\	0.76	N
\	0.68	N

$$P(\text{Precision}) = TP / (TP + FP)$$

$$R(\text{Recall}) = TP / (TP + FN)$$

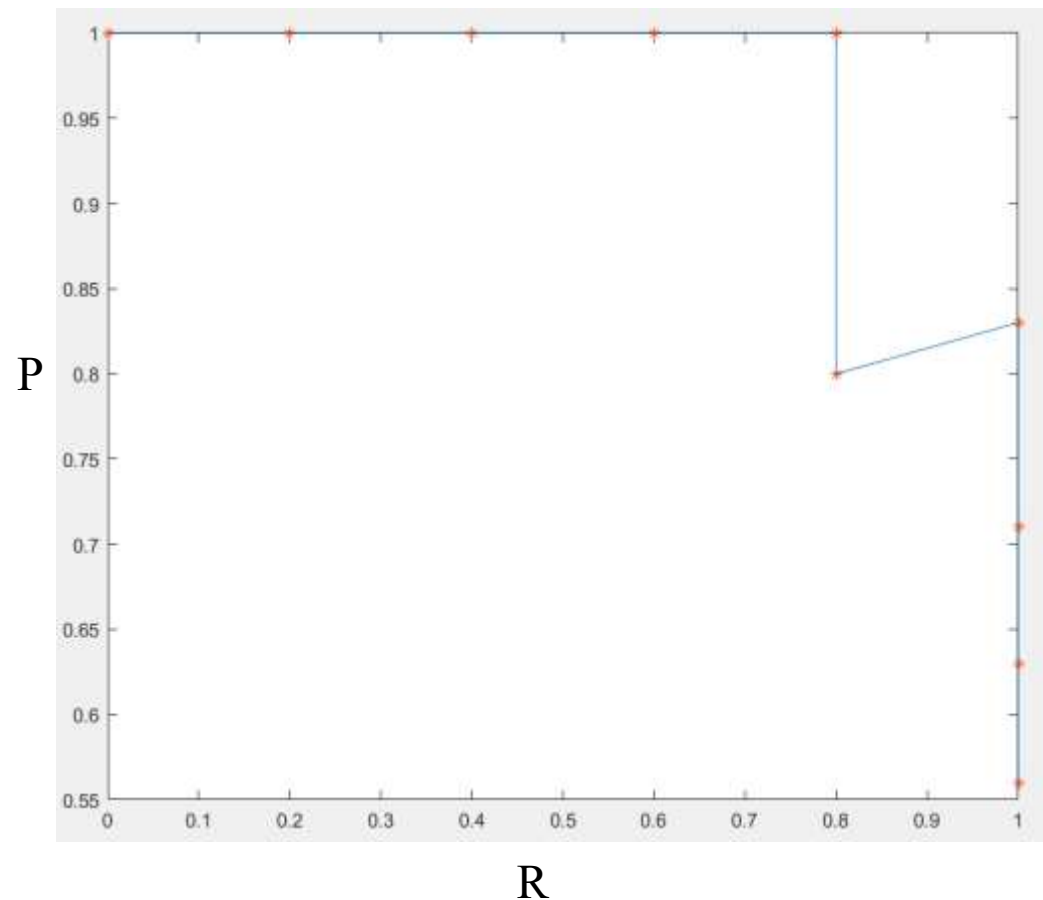
$$P = 5 / (5 + 4) = 0.56$$

$$R = 5 / (5 + 0) = 1.0$$

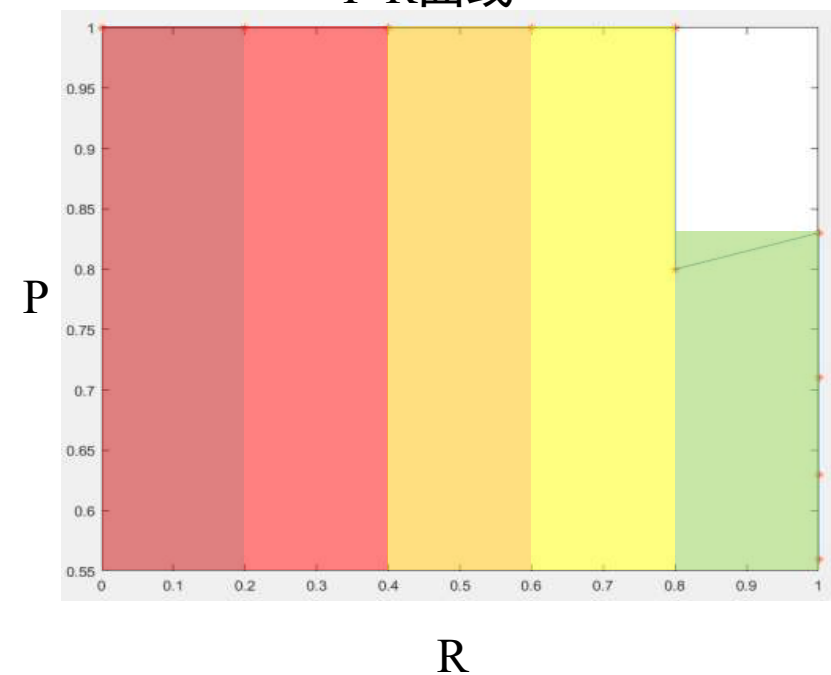
Precision	Recall
1.0	0.2
1.0	0.4
1.0	0.6
1.0	0.8
0.8	0.8
0.83	1.0
0.71	1.0
0.63	1.0
0.56	1.0

Precision	Recall
1.0	0.0
1.0	0.2
1.0	0.4
1.0	0.6
1.0	0.8
0.8	0.8
0.83	1.0
0.71	1.0
0.63	1.0
0.56	1.0

P-R曲线



P-R曲线

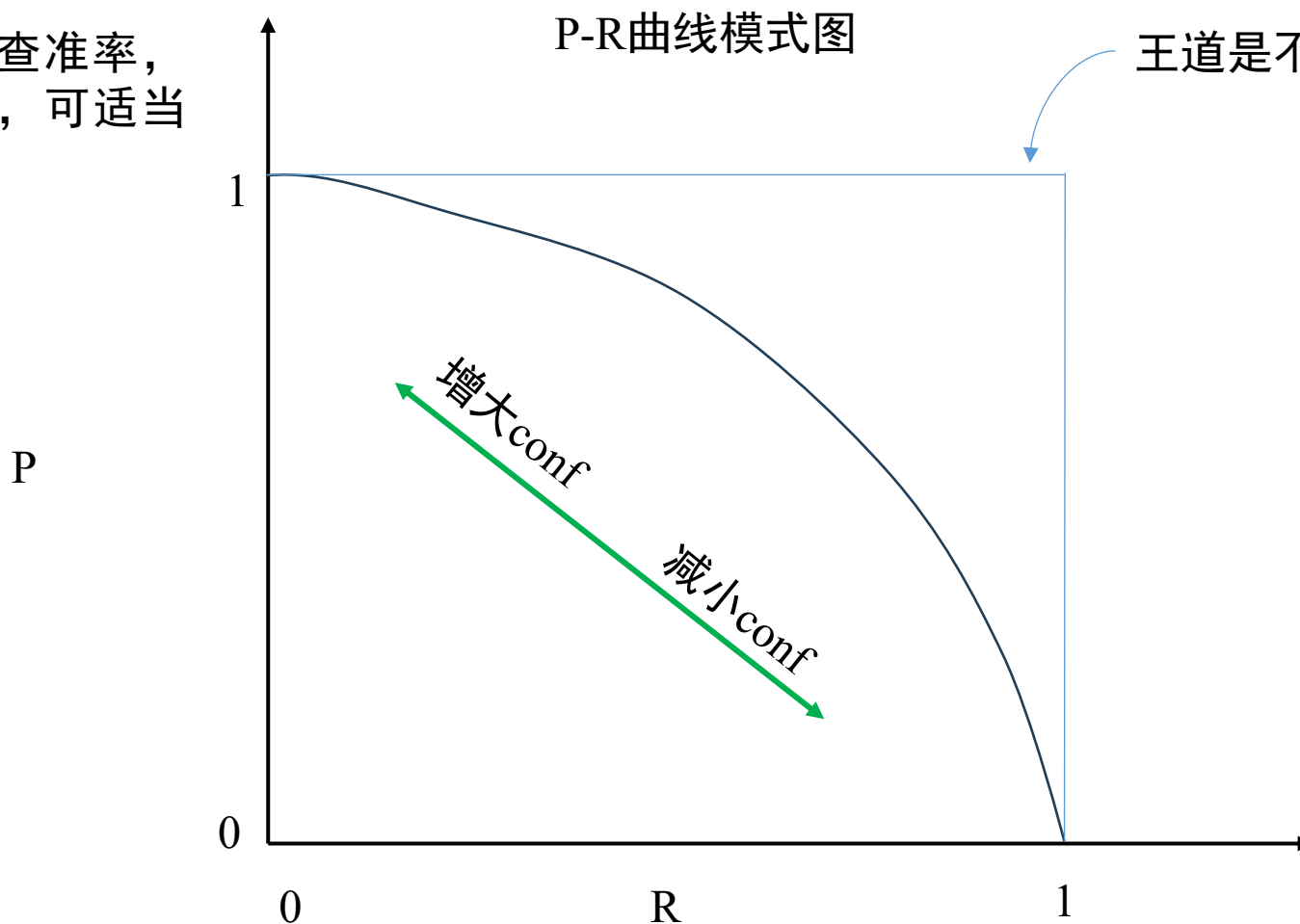


$$S = (0.8 - 0) * 1.0 + (1.0 - 0.8) * 0.83 = 0.966$$

AP  
mean  
mAP



若业务需较高的查准率，  
允许一定的漏判，可适当  
增大conf阈值



王道是不断接近这条曲线！

若业务需较高的召回率，允许一定的误判，  
可适当减少conf阈值

AP: 对于某一个类的PR曲线所围住的面积

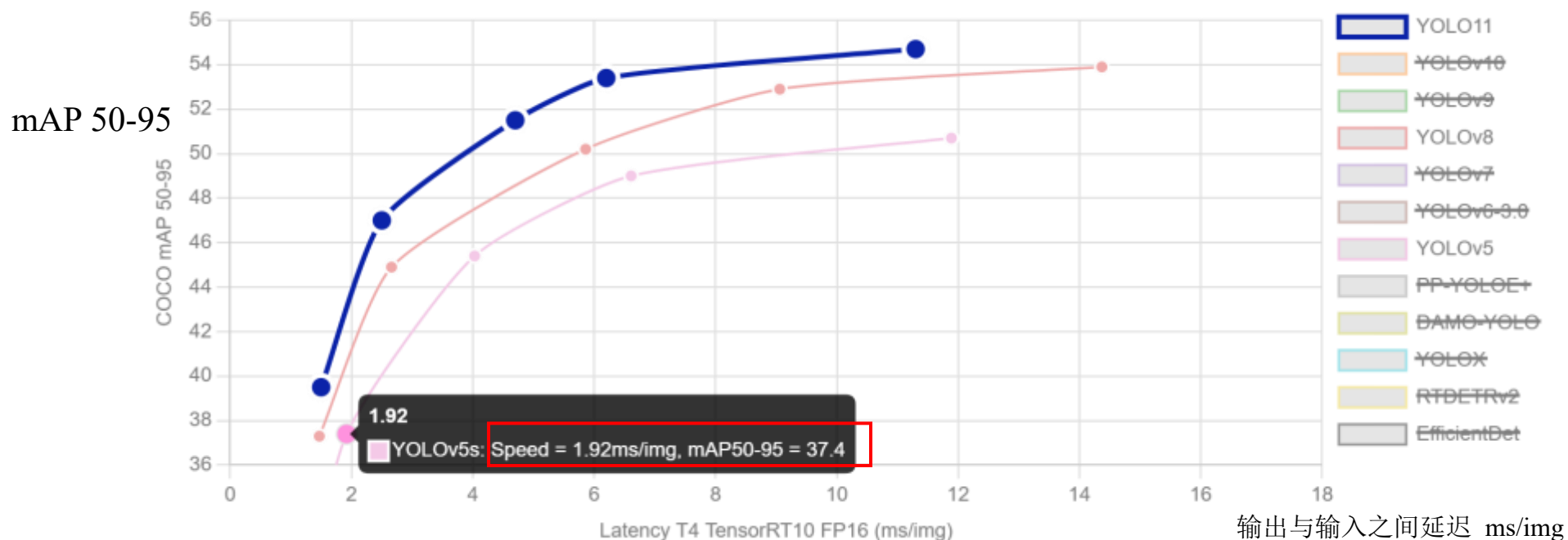
mAP: 各个类的AP均值 (一般情况mAP指的是mAP50, IOU=0.5)

工业应用广泛

mAP95: IOU取0.95时的mAP

mAP50-95:  $(mAP50 + mAP55 + \dots + mAP90 + mAP95) / 10$

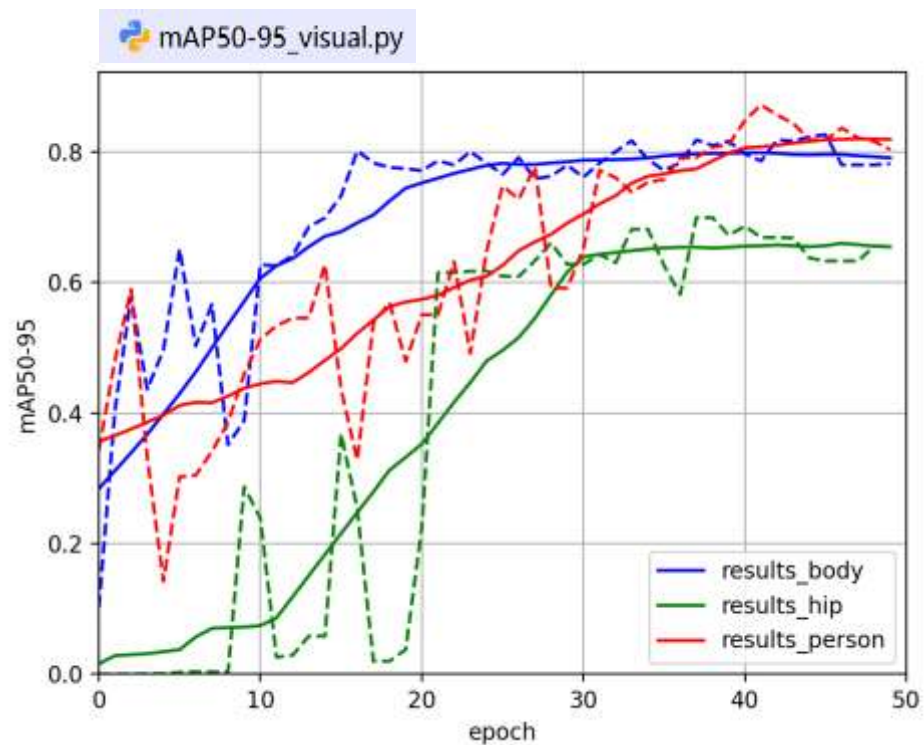
学术应用广泛



# YOLO11 mAP可视化



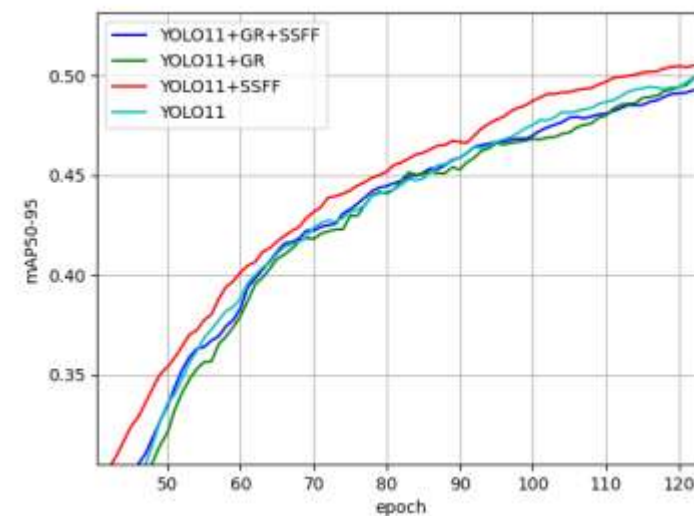
整车先进设计制造技术全国重点实验室  
State Key Laboratory of Advanced Design and Manufacturing Technology for Vehicle



results

epoch	metrics/mAP50(B)	metrics/mAP50-95(B)
1	0.33436	0.10197
2	0.75506	0.40397
3	0.90087	0.57863
4	0.7614	0.43601
5	0.82016	0.49673

- YOLO11
- YOLO11+GR
- YOLO11+GR+SSFF
- YOLO11+SSFF



# YOLO11输出结果分析



整车先进设计制造技术全国重点实验室  
State Key Laboratory of Advanced Design and Manufacturing Technology for Vehicle



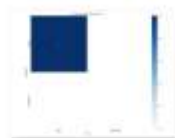
weights



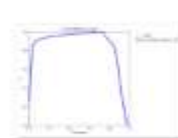
args



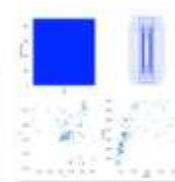
confusion\_matrix



confusion\_matrix\_normalized



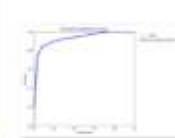
F1\_curve



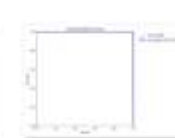
labels



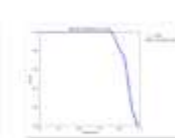
labels\_correlation



P\_curve



PR\_curve



R\_curve



results



results



train\_batch0



train\_batch1



train\_batch2



train\_batch480



train\_batch481



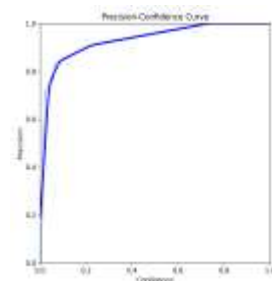
train\_batch482



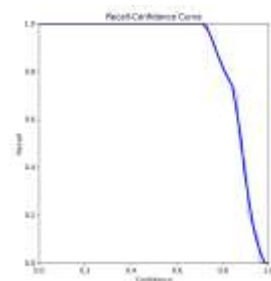
val\_batch0\_labels



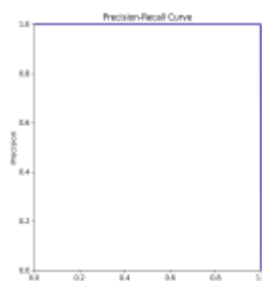
val\_batch0\_predicted



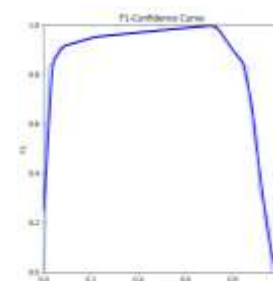
“P\_curve”



“R\_curve”



“PR\_curve”



“F1\_curve”

# YOLO11输出结果分析

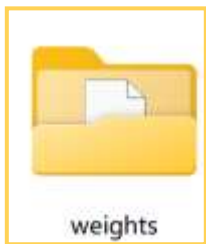


整车先进设计制造技术全国重点实验室  
State Key Laboratory of Advanced Design and Manufacturing Technology for Vehicle

weights:  
训练结果  
可直接使用的权重文件

best.pt 部署

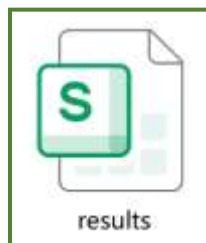
last.pt 继续训练



weights



args



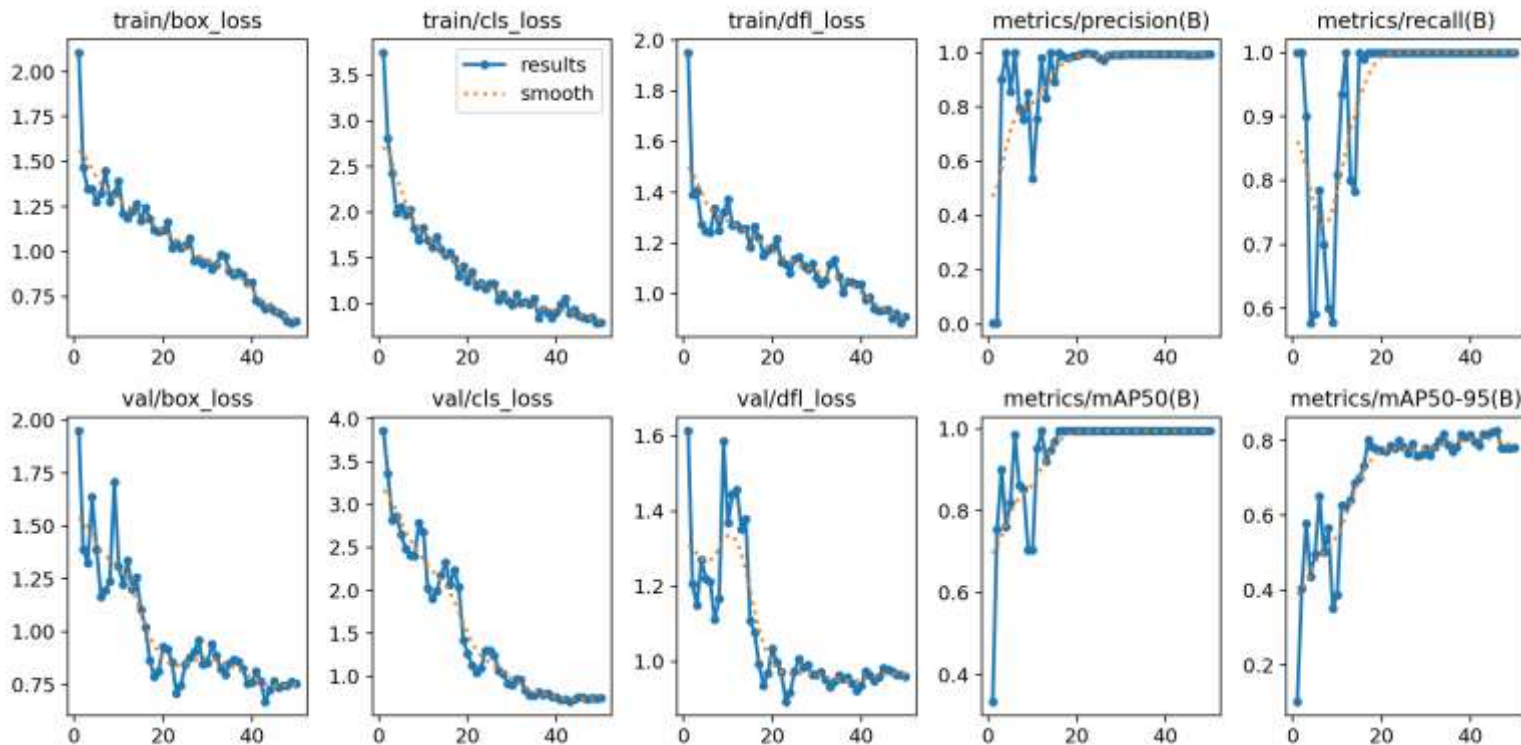
results



results

results.csv:  
对训练过程中loss和准确率  
等参数变化的文字记录

args:  
存储训练参数的  
配置文件

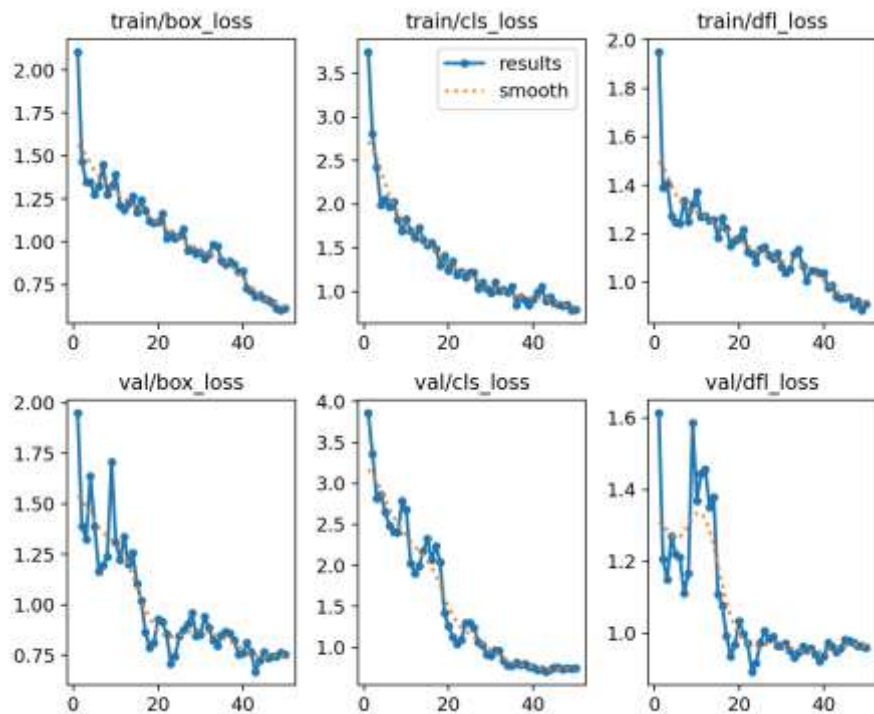


loss: 损失函数  
随周期进行不断降低

P值, R值, mAP值。  
随周期进行不断趋于最高



## “损失函数”



train/测试轮      val/验证轮

1. **box\_loss**(边界框回归损失)——用于优化预测边界框与真实边界框之间的差异。

2. **class\_loss**(分类损失)——用于优化对目标类别的预测。分类损失确保模型能够正确地识别出图像中的对象属于哪个类别。

3. **dfi\_loss**(Distribution Focal Loss)——解决目标检测中的类别不平衡问题，并提高模型在处理小目标和困难样本时的性能。

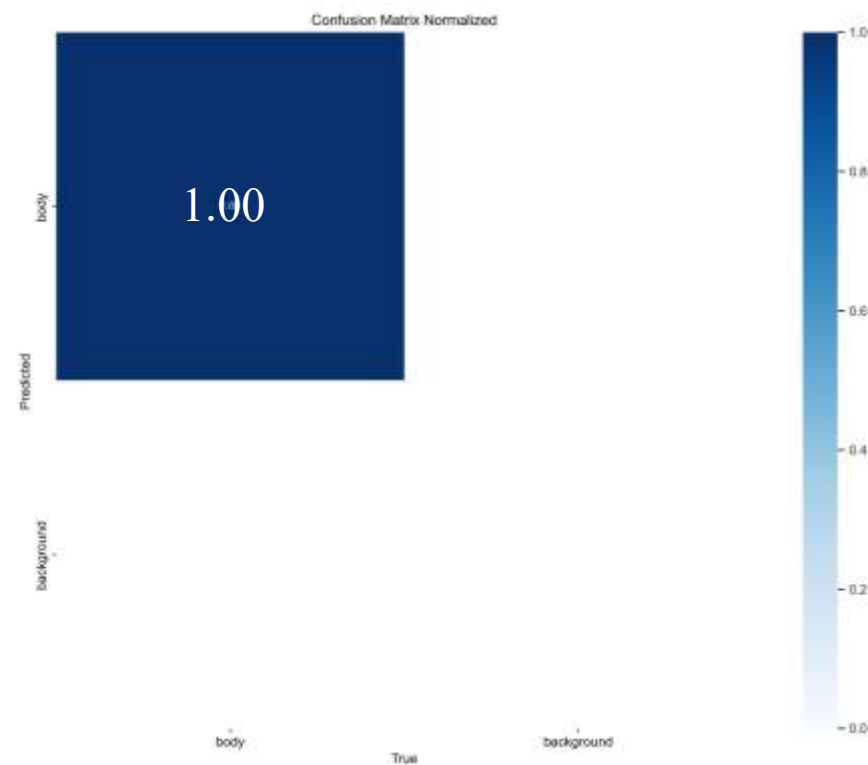
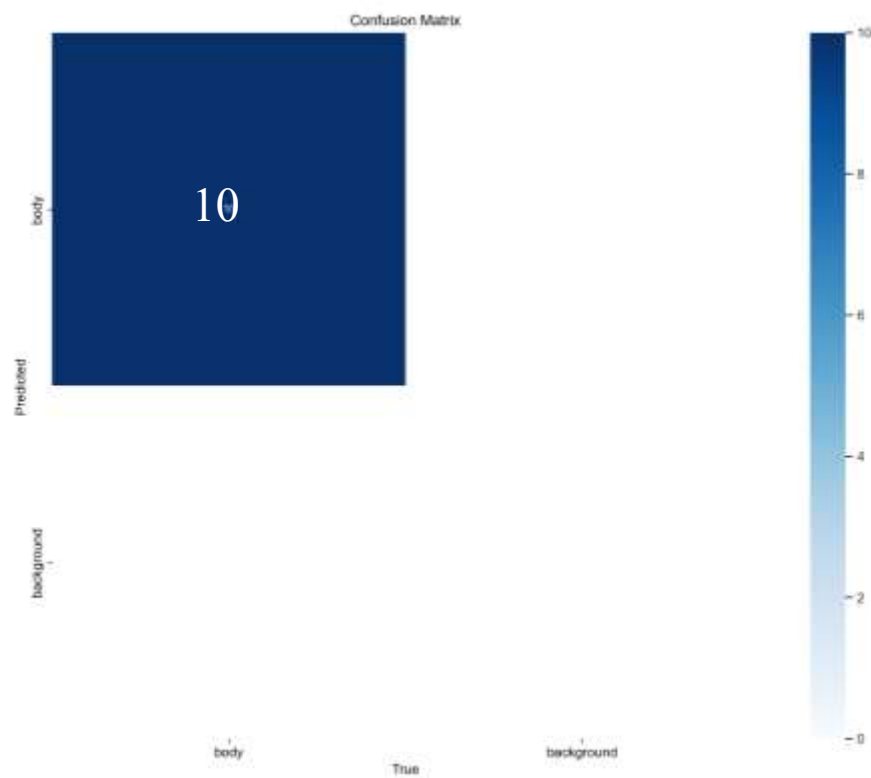
[读懂损失函数的奥秘\\_yolov11损失函数-CSDN博客](#)

# YOLO11输出结果分析



整车先进设计制造技术全国重点实验室  
State Key Laboratory of Advanced Design and Manufacturing Technology for Vehicle

## “混淆矩阵”



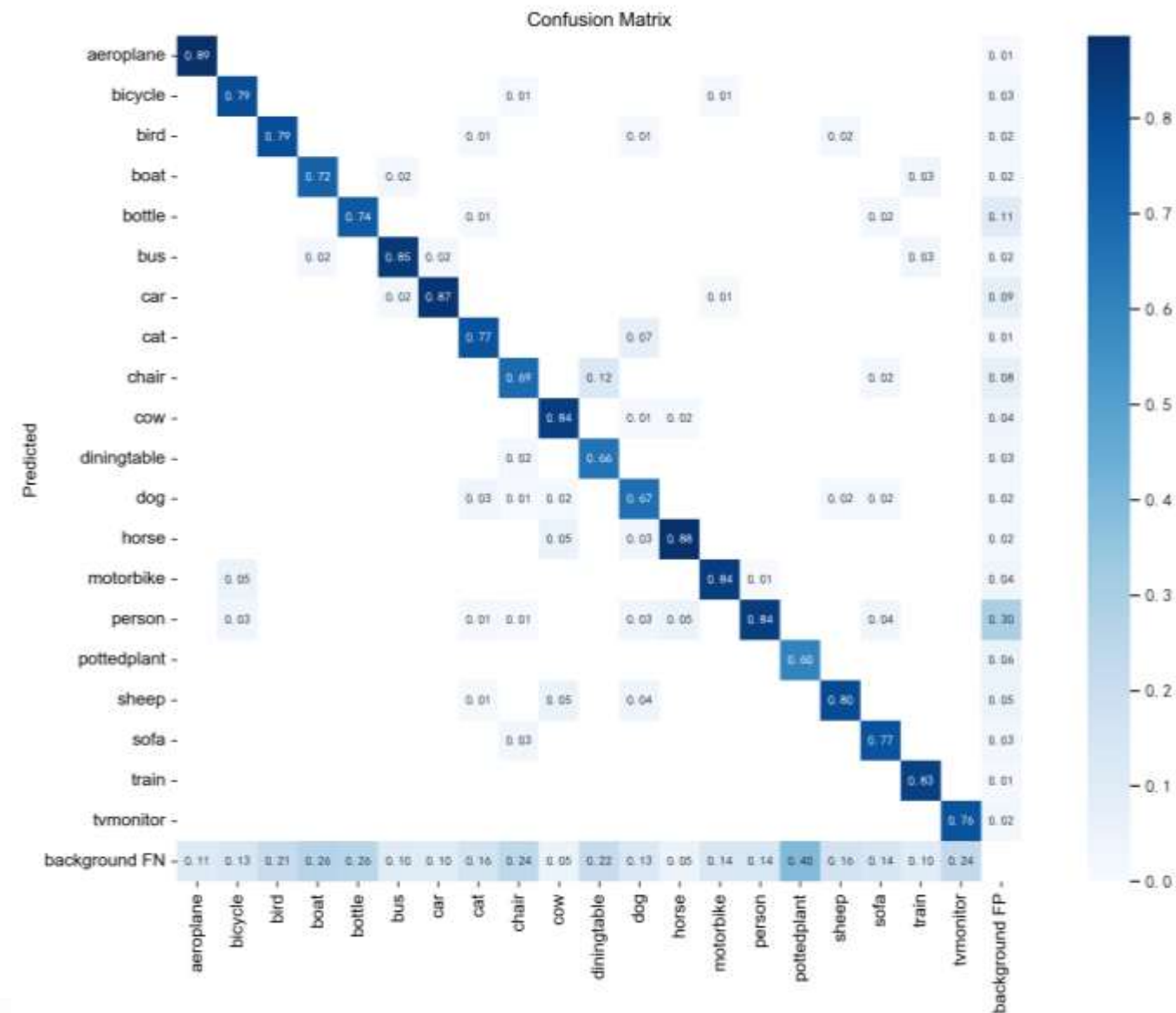


# YOLO11输出结果分析

## “混淆矩阵”

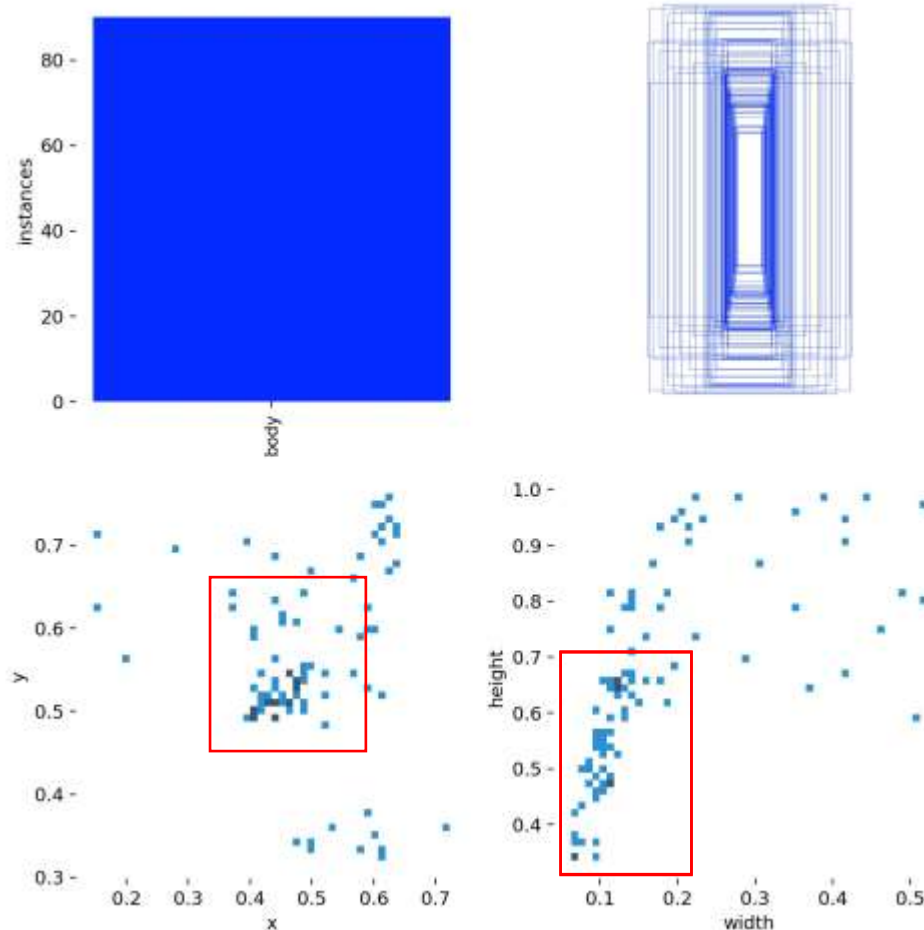


整车先进设计制造技术全国重点实验室  
State Key Laboratory of Advanced Design and Manufacturing Technology for Vehicle



混淆矩阵的每一列代表了预测类别，每一行是真实类别。  
矩阵中 $A_{ij}$ 的含义：第 $j$ 个类别被预测为第 $i$ 个类别的概率。  
方便看出机器是否将两个不同的类混淆。

“ labels ”



“ labels ”图是用来描述数据分布的，其中：

图①：描述每个类别对应的样例个数。

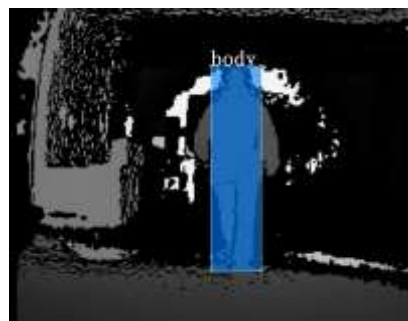
图②：描述边界框的分布。

图③：描述样本中心点相对于整幅图分布。

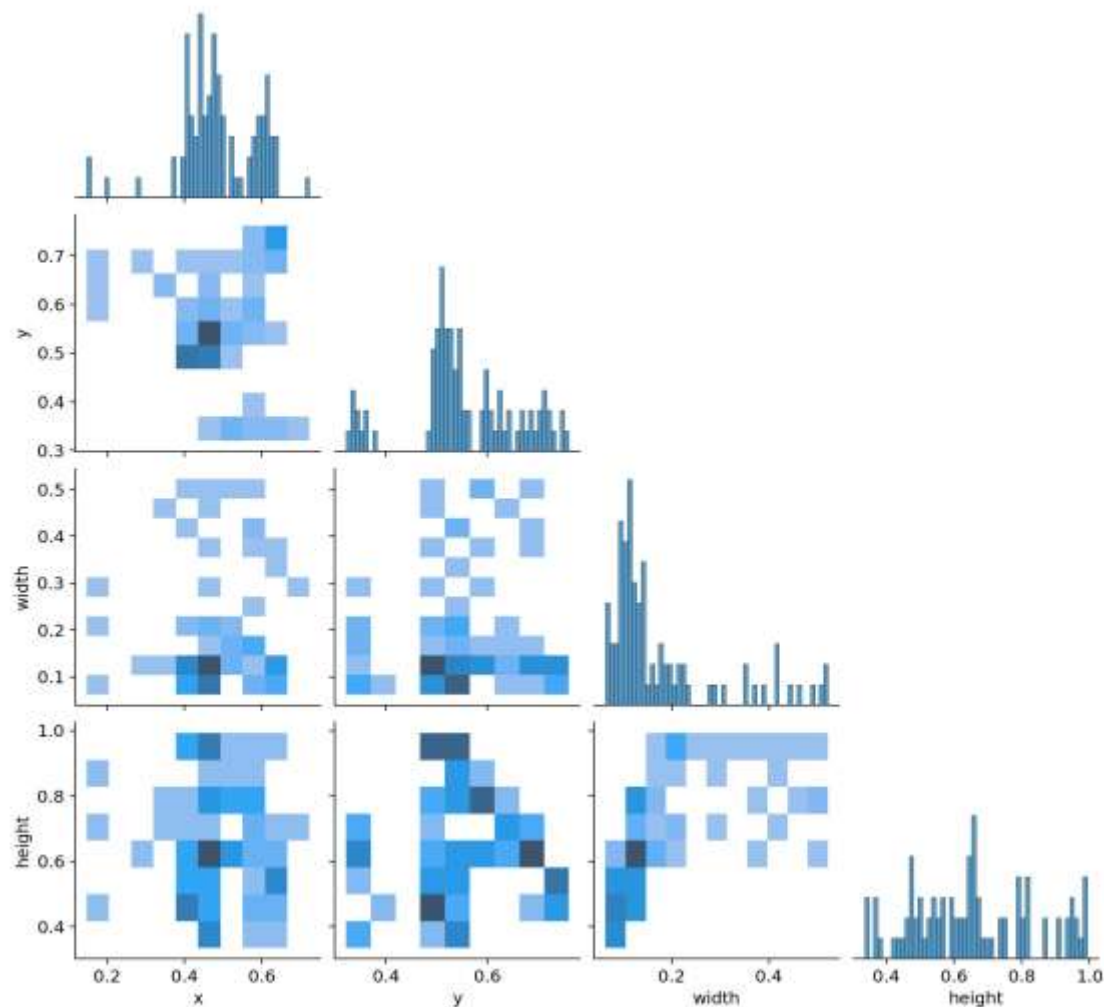
因为我们训练的数据集是人体，人基本直立行走，所以可以看到图幅中心(0.5,0.5)密度较大。

图④：描述目标的宽高相对于整幅图分布。

人基本直立行走，矩形框的宽很小，所以左侧密度大。高几乎都占图幅的0.4以上。



“labels\_correlogram”



中心点坐标x, y, 框的width, height 4个变量相关关系图

每一行的最后一幅图代表的是x, y, 宽和高的分布情况：  
最上面的图（1, 1）表明中心点横坐标x的分布。

可以看到大部分集中在整幅图的中心位置；

（2, 2）图表明中心点纵坐标y的分布。

可以看到大部分集中在整幅图的中心位置；

（3, 3）图表明框的宽的分布情况。

可以看到大部分框的宽小于整幅图的宽。

（4, 4）图表明框的高的分布情况。

其他的图即是寻找这4个变量间的相关关系。

## 数据增强相关参数



args

- `mosaic: 1.0` : 马赛克数据增强的概率。
- `mixup: 0.0` : 混合数据增强的概率。
- `copy_paste: 0.0` : 复制粘贴数据增强的概率。
- `copy_paste_mode: flip` : 复制粘贴数据增强的模式。
- `auto_augment: randaugment` : 自动数据增强的类型。
- `erasing: 0.4` : 随机擦除的概率。
- `crop_fraction: 1.0` : 裁剪数据增强的裁剪比例。
- `hsv_h: 0.015` : HSV 数据增强的 H 通道变化范围。
- `hsv_s: 0.7` : HSV 数据增强的 S 通道变化范围。
- `hsv_v: 0.4` : HSV 数据增强的 V 通道变化范围。
- `degrees: 0.0` : 旋转数据增强的角度范围。
- `translate: 0.1` : 平移数据增强的比例。
- `scale: 0.5` : 缩放数据增强的比例。
- `shear: 0.0` : 剪切数据增强的角度范围。
- `perspective: 0.0` : 透视数据增强的概率。
- `flipud: 0.0` : 上下翻转数据增强的概率。
- `fliplr: 0.5` : 左右翻转数据增强的概率。
- `bgr: 0.0` : BGR 数据增强的概率。