# 测试报告

## 对方小组人员说明

第三组：李航、赵凌云、朱卫明、王子谦

## 数据集准备

使用VOC07+12数据集进行训练，实现对以下20类别进行目标检测任务。20类别如下：aeroplane, bicycle, bird, boat, bottle, bus, car, cat, chair, cow, diningtable, dog, horse, motorbike, person, pottedplant, sheep, sofa, train, tvmonitor。

例如VOC2007格式的数据集如下：

VOCdevkit

--VOC2007

----Annotations

------xml files (eg. 000001.xml)

----JPEGImages

------image files (eg. 000001.jpg)

----ImageSets

------Main

--------train.txt

--------test.txt

--------trainval.txt

## 测试环境

显卡：NVIDIA 4060显卡、Win11系统、环境：python3.8 + pytorch1.14

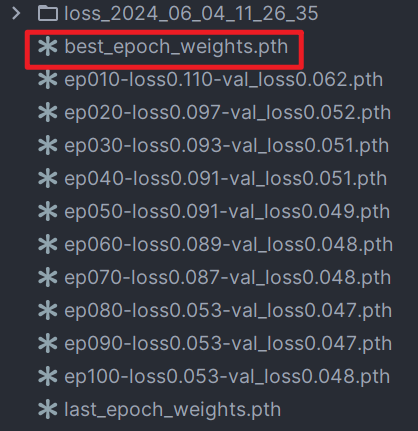
## 模型准备

从零开始训练一个yolo5s模型，其中超参数设置如下：

|  |
| --- |
| # anchors\_path代表先验框对应的txt文件。  # anchors\_mask用于帮助代码找到对应的先验框。 anchors\_path= 'model\_data/yolo\_anchors.txt', anchors\_mask= [[6, 7, 8], [3, 4, 5], [0, 1, 2]], # 输入图片的大小，必须为32的倍数。 input\_shape=[640, 640], backbone='cspdarknet', # 所使用的YoloV5的版本。s、m、l、x # 在除cspdarknet的其它主干中仅影响panet的大小 phi='s', # 只有得分大于置信度的预测框会被保留下来 confidence=0.5, # 非极大抑制所用到的nms\_iou大小 nms\_iou=0.3, # 该变量用于控制是否使用letterbox\_image对输入图像进行不失真的resize， # 在多次测试后，发现关闭letterbox\_image直接resize的效果更好 letterbox\_image=True, cuda=True,  epoch=100,  mosaic = True  mosaic\_prob = 0.5  mixup = True  mixup\_prob = 0.5  special\_aug\_ratio = 0.7  # label\_smoothing 标签平滑。一般0.01以下。如0.01、0.005。  label\_smoothing = 0  UnFreeze\_Epoch = 100  Unfreeze\_batch\_size = 8  # Freeze\_Train 是否进行冻结训练  # 默认先冻结主干训练后解冻训练。  Freeze\_Train = True  # 冻结阶段训练参数  Init\_Epoch = 0  Freeze\_Epoch = 50  Freeze\_batch\_size = 16  # 其它训练参数：学习率、优化器、学习率下降有关  # Init\_lr 模型的最大学习率  # Min\_lr 模型的最小学习率，默认为最大学习率的0.01  Init\_lr = 1e-2  Min\_lr = Init\_lr \* 0.01  optimizer\_type = "sgd"  momentum = 0.937  weight\_decay = 5e-4  # lr\_decay\_type 使用到的学习率下降方式，可选的有step、cos  lr\_decay\_type = "cos"  eval\_flag = True  eval\_period = 10  # num\_workers 用于设置是否使用多线程读取数据  num\_workers = 4 |

训练集和测试集的划分：9:1

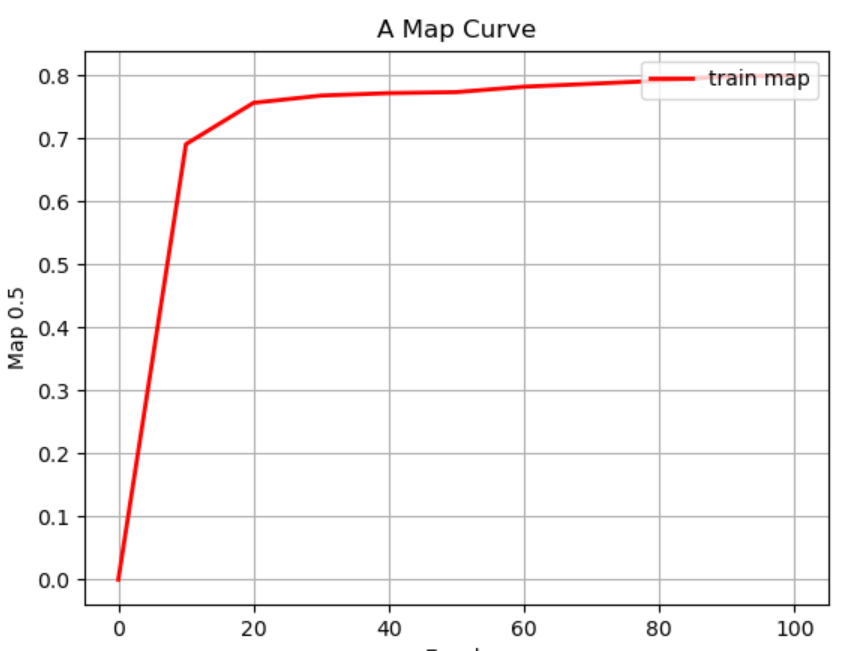
训练时间：4h

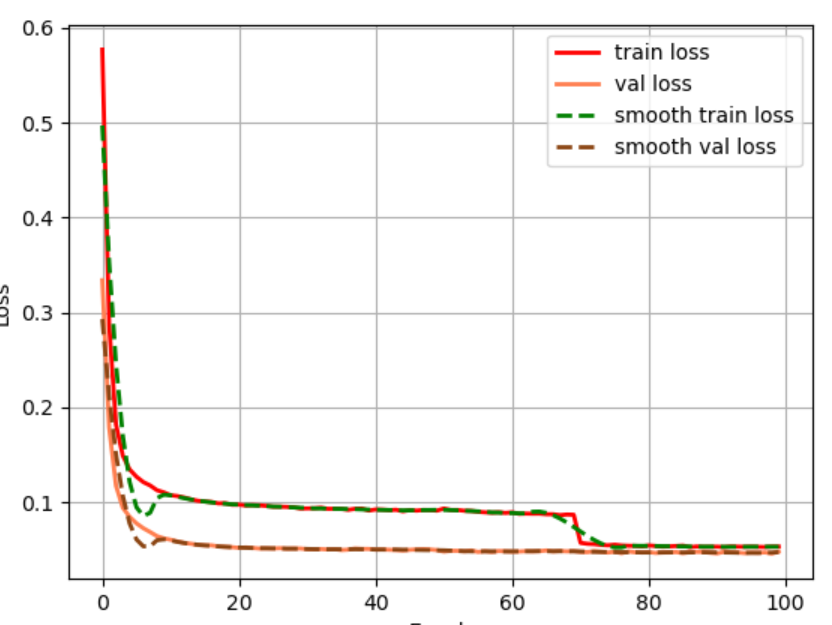
将效果最好的模型进行保存进行进一步测试

## 评估模型

评估指标1：fps

|  |  |  |
| --- | --- | --- |
|  | 单张图片测试时间 | fps |
| 测试1 | 0.016501s | 60.6 |
| 测试2 | 0.015683s | 63.7 |
| 测试3 | 0.015876s | 62.9 |
| 平均测试 | 0.016020s | 62.4 |

评估指标2：map

评估指标3：loss

## 结论

最终测试的模型符合需求，实时帧率至少都可以达到60帧以上，同时map基本收敛于0.8附近，符合要求。