

在VOC数据集上训练并测试目标检测模型Faster R-CNN和YOLO V3
任务2

神经网络和深度学习

-期中作业-

马超

2024 年 5 月 31 日

目录

1	引言	1
2	VOC2007数据集介绍	1
3	模型介绍	2
3.1	Faster R-CNN简介	2
3.2	YOLO V3简介	2
4	实验结果	3
4.1	Faster R-CNN Proposal Box可视化对比	3
4.2	YOLO V3检测结果可视化	3
4.3	Tensorboard可视化	4
5	github链接及模型权重下载地址	4

1 引言

本实验旨在学习使用现成的目标检测框架，在VOC数据集上训练和测试目标检测模型Faster R-CNN和YOLO V3。实验要求包括模型训练过程中的可视化、模型测试结果的对比和分析、实验设置等内容。

2 VOC2007数据集介绍

VOC2007 (Visual Object Classes) 数据集是一个常用的目标检测和图像分类数据集，用于评估计算机视觉算法的性能。该数据集包含了20个常见的物体类别，如人、猫、狗、汽车等，并且每个类别都有大量的图像样本。VOC2007数据集是PASCAL VOC项目的一部分，该项目旨在推动计算机视觉领域的发展。

主要特点

- **类别标注：**VOC2007数据集包含20个不同的物体类别，每个类别都有相应的标签，如人（person）、猫（cat）、狗（dog）、汽车（car）等。
- **图像样本：**数据集中包含了大量的图像样本，用于训练和测试目标检测和图像分类模型。每个图像样本都有与之相关联的标注信息，包括物体的位置（bounding box）和类别标签。
- **训练集和测试集：**VOC2007数据集分为训练集和测试集两部分，用于模型的训练和评估。训练集用于训练模型的参数，而测试集则用于评估模型在新样本上的性能。
- **挑战性任务：**除了目标检测和图像分类任务外，VOC2007数据集还包含了一些挑战性的任务，如目标分割（semantic segmentation）和边界框精确度评估（bounding box precision evaluation）等。

总的来说，VOC2007数据集是一个广泛应用于目标检测和图像分类领域的基准数据集，它提供了丰富的图像样本和相应的标注信息，可以帮助研究人员和工程师评估和改进他们的算法和模型。

3 模型介绍

3.1 Faster R-CNN简介

Faster R-CNN是最先进的目标检测模型之一，建立在R-CNN和Fast R-CNN成功的基础上。它引入了区域建议网络（RPN），高效生成区域建议，显著加快了检测过程。

Faster R-CNN的架构主要包括三个部分：

- **卷积神经网络（CNN）骨干网：**通常使用VGG16或ResNet等深度卷积神经网络，从输入图像中提取特征图。
- **区域建议网络（RPN）：**该网络利用骨干网提取的特征图生成一组候选对象建议框，这些矩形区域可能包含对象。
- **检测网络：**该网络对RPN生成的建议框进行分类，将每个建议框分类为对象类别或背景，并调整边界框坐标。

3.2 YOLO V3简介

YOLO V3 (You Only Look Once Version 3) 是一个最先进的实时目标检测模型。与Faster R-CNN等传统方法不同，YOLO V3在单个神经网络中直接进行边界框预测和类别预测，因此具有极高的检测速度和精度。

YOLO V3的架构主要包括以下几个部分：

- **骨干网络（Backbone）：**通常使用Darknet-53，一个包含53层卷积层的深度卷积神经网络，用于特征提取。
- **检测头（Detection Heads）：**YOLO V3在三个不同尺度上进行检测，分别从特征图的不同层提取信息。这些尺度帮助模型检测不同大小的对象。
- **锚框（Anchor Boxes）：**在特征图上的每个位置，YOLO V3使用预定义的锚框来预测对象的边界框和类别。

4 实验结果

4.1 Faster R-CNN Proposal Box可视化对比

以下是训练好的Faster R-CNN模型第一阶段产生的proposal box和最终的预测结果。

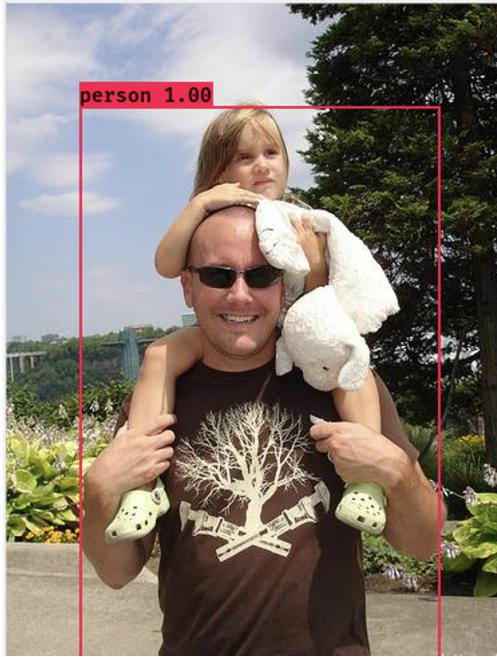


图 1: 1

4.2 YOLO V3检测结果可视化

选取三张不在VOC数据集内包含有VOC中类别物体的图像，比较在这三张图片上的模型检测结果，包括bounding box、类别标签和得分。

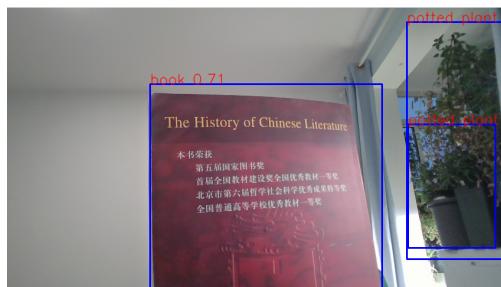


图 2: 1



图 3: 2

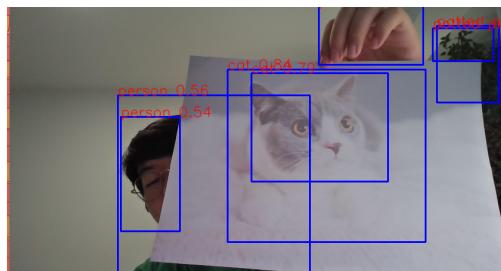
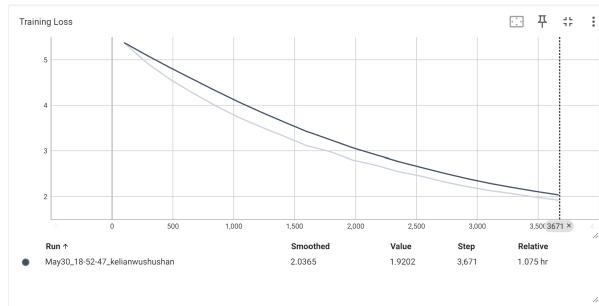
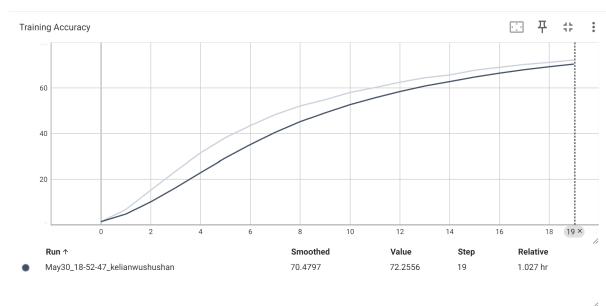


图 4: 3

4.3 Tensorboard可视化



(a) 训练集和验证集的损失曲线



(b) 训练集和验证集的准确率曲线

5 github链接及模型权重下载地址

repo的public github链接:

<https://github.com/Ethereal-Redolent/hw/tree/master>

模型权重的下载地址:

https://pan.baidu.com/s/1gwq605m9oq820Mtd_QS7Lw?pwd=y3gf