

文章编号: 2095-2163(2020)06-0199-04

中图分类号: TP391.4

文献标志码: A

YOLO 算法在交通监控目标检测中的研究与应用

陈小娥

(福建船政交通职业学院 信息工程系, 福州 350007)

摘要: 为了促进智能交通的快速建设与发展,实现机器自动准确快速对交通监控抓拍的静态图像进行目标检测显得尤为重要。近年来,基于深度学习的目标检测研究发展十分迅速,各种检测算法及其改进算法不断涌现。主要采用 YOLOv3 算法对交通监控中各种车辆和行人等目标进行检测,通过实验分析了算法在交通监控目标中的准确率和召回率等性能指标,对同类应用场景的目标检测提供了参考。

关键词: 目标检测; 交通监控; YOLOv3; 深度学习

Research and Application of YOLO in Traffic Monitoring Target Detection

CHEN Xiaoe

(Department of Information Engineering, Fujian Chuanzheng Communications College, Fuzhou 350007, China)

[Abstract] In order to promote the rapid construction and development of intelligent transportation, it is very important to realize automatic, accurate and fast target detection of static images captured by traffic monitoring. In recent years, the research of target detection based on deep learning develops rapidly, various detection algorithms and their improved algorithms are emerging. yolov3 algorithm is used to detect various vehicles and pedestrians in traffic monitoring. The accuracy and recall rate of the algorithm in traffic monitoring target are analyzed by experiments, which provides a reference for target detection in similar application scenarios.

[Key words] object detection; traffic monitoring; YOLOv3; deep learning

0 引言

随着我国社会经济的快速发展以及城市化进程的加快,道路上车辆和行人数量不断增多,交通问题也随之变得日益严重。道路交通目标种类繁多,对不同车辆和行人进行快速检测和识别是计算机研究者近年来面临的挑战性任务之一,也是智能交通监控系统和自动驾驶技术等应用的关键。

目标检测是计算机视觉技术的一项重要研究课题,是指通过相关算法来实现客观世界中对象的分类和定位。主要分为传统目标检测和现阶段基于卷积神经网络的目标检测。传统目标检测主要通过手工提取特征,然后针对特定检测对象设计和训练分类器,该方法很大程度上依赖于提取的特征的准确性,然而手工提取的特征往往对目标的表达能力不足,鲁棒性低。现阶段基于卷积神经网络的目标检测,又分为两大类,包括 Two-Stage 和 One-Stage 的两种不同的方法和路线。其中,Two-Stage 方法主要通过利用 RPN 网络来对候选区域进行推荐,其中具有代表性的是 R-CNN、Faster RCNN 系列方法,如 2014 年 R-CNN 框架^[1]被提出,通过大容量卷积神经网络来提取特征,而后接连出现了一系列改进

框架,如 SPP-Net^[2]、Fast-R-CNN^[3]、Faster R-CNN^[4]、R-FCN^[5]等算法。而 One Stages 主要通过直接回归目标的位置来进行目标的检测,比较典型的有 YOLO 和 SSD 系列方法,能够对目标检测的速度进行进一步的提升,如 2015 年 Joseph Redmon 等^[6]提出的 YOLO,用一个卷积神经网络实现定位分类一体化,是一个不基于区域算法的代表。接着,刘维等^[7]借用了区域目标检测的思想,将提取候选框的网络和目标分类的网络进行合并,提升了准确率。2017 年和 2018 年,YOLOv2^[8]和 YOLOv3^[9]版本相继被提出,它们在保持高效检测速度的基础上,引入了 anchor box 和多尺度等思想,进一步改进了网络结构,特别是 YOLOv3 引进了 ResNet 思想,进一步提高了网络的检测精度和速度。

本文采用基于 YOLOv3 深度学习网络,对道路交通中的行人、自行车、小汽车、公共汽车、卡车等目标进行检测。首先,对采集的数据集进行预处理,然后利用 YOLOv3 对交通监控对象进行目标检测,最后进行实验结果分析。

1 YOLOv3 网络结构

YOLOv3 采用 Darknet-53 作为骨干网络,进一

基金项目:福建省教育厅中青年骨干教师科研项目(JZ180353)。

作者简介:陈小娥(1984-),女,硕士,副教授,主要研究方向:人工智能。

收稿日期:2020-04-08

步地提高了目标检测的定位和分类精度。Darknet-53 包含 53 个卷积层,大量采用 3×3 和 1×1 的卷积核。为了训练 53 层的深度网络,借鉴了 ResNet 的思想,在卷积层之间构建残差模块,设置跳跃连

接。YOLOv3 通过借鉴特征金字塔网络,设计多尺度特征提取结构,提升了小目标的检测效果。YOLOv3 模型结构如图 1 所示。

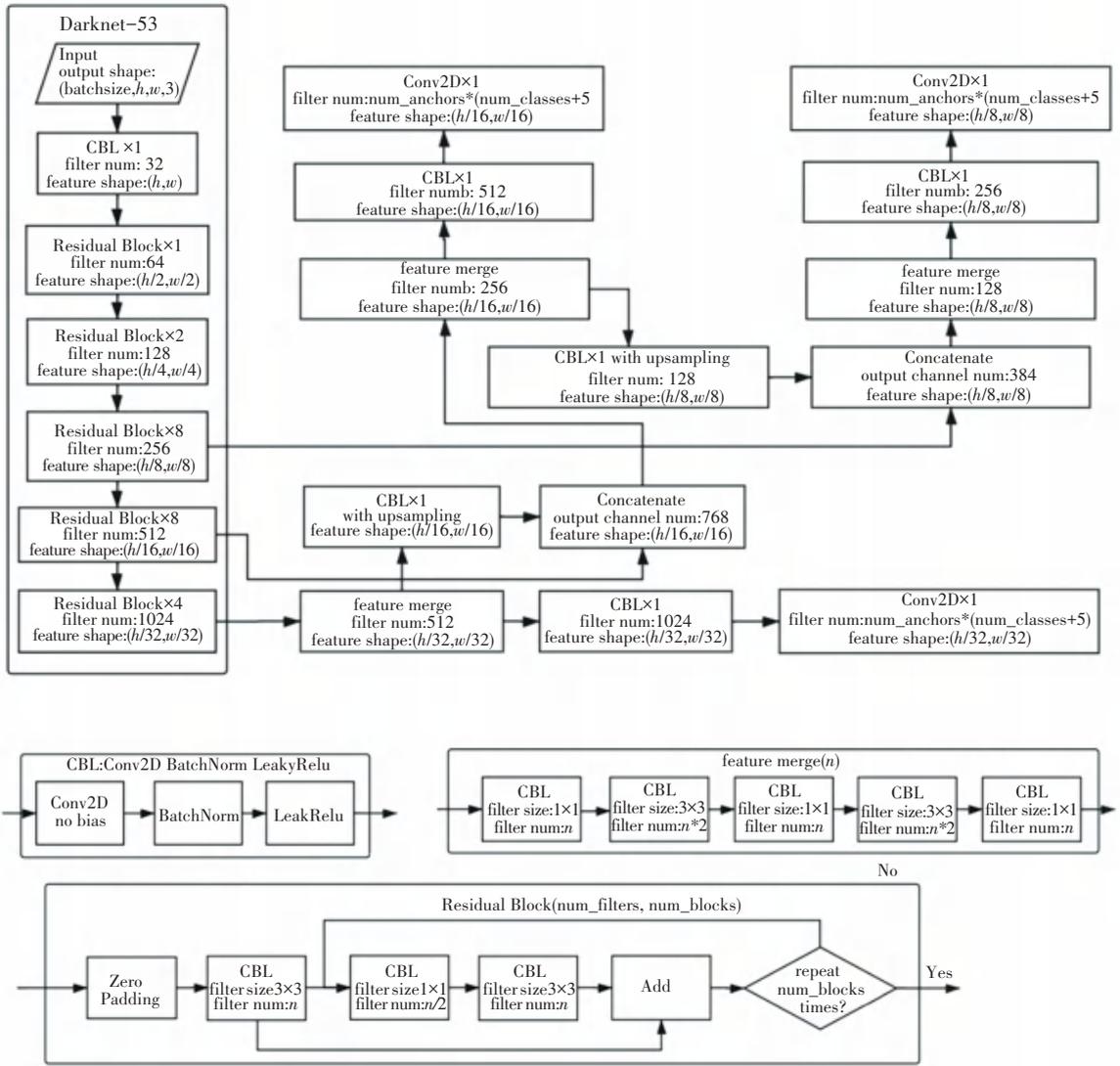


图 1 YOLOv3 模型结构

Fig. 1 Yolov3 model structure

2 实验

2.1 实验环境及步骤

本实验采用的硬件配置参数如下: Intel (R) Core(TM) i7-8750H CPU@ 2.20 GHz 处理器、16 GB 内存、6G NVIDIA GeForce GTX 1060 GPU。

实验的主要步骤如下:

(1) 数据集的采集。实验中数据集一部分来自网络搜集到的交通监控视频,一部分通过爬虫程序直接爬取网络上交通监控目标图像。

(2) 数据预处理。数据预处理包括对采集到的

监控视频分解出视频帧、对视频帧和直接爬取的目标图像进行清洗等,以及对它们进行标注。

(3) 采用 Yolov3 对测试数据集进行目标检测,并进行结果分析。

2.2 数据预处理

实验中的数据集包含提取的视频帧和爬取的监控目标图像,以及扩充后的图像,共 4 490 幅。

提取视频帧的核心代码如下所示,按照每秒取 1 帧图像进行保存, crop_vertices 存放待保存的图像的裁剪范围,可以根据实际情况设定,提取的帧也可

不裁剪。

表 2 目标像素大小分布

Tab. 2 Target pixel size distribution

目标尺寸	占总目标数的比重/%
(0,32 * 32]	42.27
(32 * 32,96 * 96]	44.63
>96 * 96	13.10

```

for vertex, video_file_name in zip(crop_vertices, os.listdir('.')):
    capture = cv2.VideoCapture(video_file_name)
    fps = capture.get(cv2.CAP_PROP_FPS)
    collect_freq = fps // 1 #每秒取1张图片
    frame_index = 0
    if capture.isOpened():
        while True:
            success, frame = capture.read()
            if success:
                frame_cropped = frame[vertex[0],1:vertex[1],vertex[0],0:vertex[1],0]
                if frame_index % collect_freq == 0:
                    cv2.imwrite('./splited/{frame_id}.jpg', frame_cropped)
                    frame_id += 1
            else:
                break
            frame_index += 1

```

图 2 提取视频帧核心代码

Fig. 2 Core code of extracting video frame

提取的视频帧和直接爬取的图片,经过一定的筛选,过滤到没有目标的样本,使用 labelme 进行标注。数据集的样本示例如图 3 所示,上方两图是交通监控视频提取的帧,下方两图是网上直接爬取的交通监控图像。表 1 中主要包含了数据集所有图片中小汽车、自行车、公共汽车、卡车和行人的目标数量。表 2 中显示了这些目标像素大小的分布情况,尺寸介于(0,32 * 32]之间的属于小物体,(32 * 32,96 * 96]属于中物体,大于 96 * 96 的属于大物体。



图 3 数据集的样本示例图

Fig. 3 Sample pictures of data set

表 1 数据集中目标数

Tab. 1 Target number in dataset

目标	car	bicycle	bus	truck	person
数量	21 140	3 900	880	1 430	6 950

2.3 结果分析

实验过程中,使用原 Yolov3 权重文件^[9],其训练参数中批次大小 batch_size 为 64,初始学习率为 0.001。对标注好的数据集进行测试。YOLOv3 算法对交通监控目标检测的结果如下图所示。



图 4 检测结果示例图

Fig. 4 Some pictures of test result

通过实验结果定量评估 YOLOv3 算法对于交通监控目标检测识别的效果。从准确率 Precision 和召回率 Recall 对算法进行评价。具体地,

$$precision = \frac{TP}{TP + FP}, recall = \frac{TP}{TP + FN},$$

其中, TP 表示本身是正样本,分类器也认为它是正样本;FP 表示本身是负样本,分类器认为它是正样本;FN 表示本身是正样本,分类器认为它是负样本;Precision 表示的是分类器认为是正类并且确实是正类的部分占有所有分类器认为是正类的比例;Recall 表示的是分类器认为是正类并且确实是正类的部分占有所有确实是正类的比例。IOU 交并比衡量监测框和标签框的重合程度。表 3 中显示了对应不同交并比阈值的平均精度,AP 表示 Average Precision,第 1 个 AP 对应的 IoU = 0.5; 0.95,第 2 个

(下转第 205 页)