

doi:10.11835/j.issn.1000-582X.2020.016

改进 YOLOv3 算法的车辆信息检测

冯加明, 储茂祥, 杨永辉, 巩荣芬

(辽宁科技大学 电子与信息工程学院, 辽宁 鞍山 114051)

摘要: 车辆信息检测是车型识别在智慧交通领域中的首要任务。针对现有的车辆信息检测技术在检测速度、精度以及稳定性方面存在的问题, 提出了基于 YOLOv3 的深度学习目标检测算法——YOLOv3-fass。该算法以 DarkNet-53 网络结构为基础, 删减了部分残差结构, 降低了卷积层的通道数, 添加了 1 条下采样支路和 3 个尺度跳连结构, 增加了一个检测尺度, 并通过 K-均值聚类与手动调节相结合的方法计算出 12 组锚框值。最后通过迁移学习机制对 YOLOv3-fass 算法进行微调。在自研的车辆数据集上, YOLOv3-fass 算法与 YOLOv3、YOLOv3-tiny、YOLOv3-spp 算法以及具有 ResNet50 和 DenseNet201 经典网络结构的算法做了对比实验, 结果表明 YOLOv3-fass 算法能够更精准、高效、稳定地检测到车辆信息。

关键词: 深度学习; YOLOv3 算法; DarkNet-53 网络结构; 车辆信息检测; 目标检测

中图分类号: TP391.4

文献标志码: A

文章编号: 1000-582X(2021)12-071-09

Vehicle information detection based on improved YOLOv3 algorithm

FENG Jiaming, CHU Maoxiang, YANG Yonghui, GONG Rongfen

(School of Electronics and Information Engineering, University of Science and Technology Liaoning, Anshan 114051, Liaoning, P. R. China)

Abstract: Vehicle information detection is the primary task of vehicle type identification in the field of intelligent transportation. Based on deep learning YOLOv3 (You Only Look Once Version 3) model, a new YOLOv3-fass object detection algorithm was proposed to address some problems existing in vehicle information detection technology such as detection speed, accuracy and stability. In this improved algorithm, based on DarkNet-53 network structure, some residual structures were deleted, and a number of channels of convolutional layer were reduced; a down-sampling branch, three scale-hopping connection structures, and one detection scale were added; and twelve groups of anchor frame values were calculated through the means of K-means clustering algorithm combined with manual setting. Finally, YOLOv3-fass algorithm was fine-tuned through the migration learning mechanism of multi-stage pre-training. YOLOv3-fass algorithm was compared with YOLOv3, YOLOv3-tiny, YOLOv3-spp and two algorithms with ResNet50 and DenseNet201 on the vehicle data set. The experimental results show that YOLOv3-fass algorithm can detect vehicle information more accurately, efficiently and stably.

Keywords: deep learning; YOLOv3 algorithm; DarkNet-53 network structure; vehicle information detection; object detection

收稿日期: 2020-03-15 **网络出版日期:** 2020-06-19

基金项目: 国家自然科学基金资助项目(71771112); 辽宁省自然科学基金资助项目(20180550067)。

Supported by National Natural Science Foundation of China (71771112) and Natural Science Foundation of Liaoning Province (20180550067).

作者简介: 冯加明(1994—), 男, 辽宁科技大学硕士研究生, 主要从事模式识别、图像处理等方向的研究, (E-mail) feng_jia_ming@163.com。

近年来,随着现代科技的飞速发展,汽车保有量逐年增多,汽车套牌的犯罪率也逐渐增高。为了减轻对交通管理部门的压力,人们开始对车辆进行识别来确认车辆的型号,然后通过车辆的型号与车牌的匹配关系来确认车辆是否套牌,从而减少了人力管控资源。基于此背景,车辆识别技术应运而生。在车辆识别的方法中最简单方法便是通过图像的方式,即通过交通摄像头拍摄交通图像,然后通过模式识别的方法识别交通图像中的车辆。而要识别交通图像中的车辆,首要任务是检测出图像中的车辆信息,主要包括车牌、车标、车脸。

目前,车辆信息检测的方法众多,可总结为两类,即传统学习方法和深度学习方法。传统学习方法主要通过研究车牌、车标、车脸的特点,提取相应的特征,然后通过滑动窗口的方式对车辆信息进行检测。Zakaria 等^[1]使用梯度直方图特征和支持向量机相结合的方法进行车辆检测;张小琴等^[2]利用车辆的对称特征检测车脸,然后通过提取车脸的梯度直方图特征结合支持向量机的方法识别车辆。上述传统学习方法能够检测出图像中的目标,但是提取车辆信息的特征比较简单,算法稳定性容易受到环境变化的影响。随着近几年深度学习的兴起,RCNN^[3]、Fast R-CNN^[4]、Faster R-CNN^[5]、SSD^[6]以及 YOLO 系列^[7-9]等基于深度学习方法的检测算法相继出现。Huang 等^[10]采用 Faster-RCNN 算法与 VGG-16、ResNet-50 网络结构相结合的方法识别道路监控中的车标;Chen 等^[11]使用 SSD 目标检测算法实现了交通监控视频中的车辆检测以及车辆计数;He 等^[12]通过卷积神经网络的快速深度学习算法检测视频中的车辆,并与雷达相结合来提高车辆检测的性能;桑军等^[13]提出了一种 Faster-RCNN 算法与 ZF、VGG-16、ResNet-101 网络结构相结合的车型识别策略;Yang 等^[14]提出了一种 Faster-RCNN 与三级级联卷积神经网络相结合的车辆零件检测方法。上述深度学习方法能够自动提取图像中更适合车辆信息的特征,很好地解决了算法对环境的影响,但是对于小尺寸的目标检测效果差,且检测速度较慢,不能满足实时检测的需求。

随着 YOLOv3 算法的出现,检测速度以及小目标检测性能都有很大提升。Krittayanawach 等^[15]通过压缩 YOLOv3 算法模型结构进行车辆检测;Chen 等^[16]受到 YOLOv3 算法的启发,提出了一种级联特征金字塔网络的浅层模型用于交通流量中的车辆检测。目前,YOLOv3 算法在智慧交通领域中主要应用于交通监控中车辆的检测,而应用于车辆信息检测的研究并不多见。笔者将改进的 YOLOv3-fass 算法应用于多类车辆信息的检测,旨在更快、更稳定、更精准地实现图像中多类车辆信息的同时检测。

1 YOLOv3 算法的原理

YOLOv3 算法采用 DarkNet-53 的网络结构提取图像中车辆信息的特征,并采用多尺度的检测方式实现图像中目标的检测。数据集中的图像归一化到 416×416 尺寸,送入算法网络结构中进行车辆信息的检测。DarkNet-53 网络结构中采用了大量含有 1×1 和 3×3 卷积核的残差网络结构。残差网络结构如图 1 所示。

残差网络结构的残差函数为

$$F(x) = H(x) - x, \quad (1)$$

式中: $F(x)$ 表示残差网络结构的残差函数; $H(x)$ 表示残差网络结构的输出值; x 表示残差网络结构的输入值。

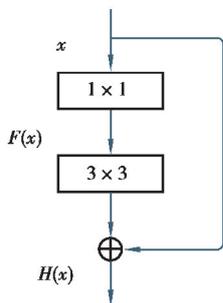


图 1 残差网络结构图

Fig. 1 Schematic diagram of residual network

YOLOv3 算法网络结构的输入图像经过 5 次步长为 2 的卷积层进行下采样,分别得到 208×208 , 104×104 , 52×52 , 26×26 , 13×13 共 5 个尺度的特征图。YOLOv3 算法将 13×13 , 26×26 , 52×52 尺度上的特征图输入到检测网络中进行多尺度检测。YOLOv3 算法将 9 组锚框平均分配在检测网络的 3 个尺度上,每个尺度的每个锚点分配 3 组锚框,共计生成 10 647 个锚框。YOLOv3 算法使用 Adam 优化器,首先采用回归的方式对锚框进行类别、置信度和预测框的预测;然后根据每个预测框的得分使用 NMS 算法选出最终的预测框。最后,根据特征图与原图的关系将预测框映射到原图上,完成图像中车辆信息的定位。

置信度的计算公式为

$$C = P_r(\text{object}) \times I, \quad (2)$$

$$P_r(\text{object}) = \begin{cases} 0 & \text{无目标,} \\ 1 & \text{有目标.} \end{cases} \quad (3)$$

式中: C 表示置信度; $P_r(\text{object})$ 表示预测框中检测到目标的概率; I 表示预测框与真值框区域的交并比。

在 NMS 算法中预测框中目标的得分计算公式为

$$S = C \times P_r(\text{logo} | \text{object}), \tag{4}$$

式中: S 表示预测框的得分; $P_r(\text{logo} | \text{object})$ 表示车标类别的条件概率。

2 YOLOv3 算法的改进

在对 YOLOv3 算法的研究中发现了一系列问题,诸如 YOLOv3 算法对于小尺寸车辆信息的检测稳定性差,参与运算的参数量大等。针对以上问题,笔者基于 YOLOv3 算法提出了更适合车辆信息检测的 YOLOv3-fass 算法。YOLOv3-fass 算法的网络结构如图 2 所示。网络结构主要包括两部分:主网络、检测网络,其中主网络由网络主干和网络支路组成。主网络含有输入模块、卷积层模块、下采样模块、节点 1~5 模块、残差 1、残差 2、残差 5 模块以及池化模块;检测网络含有 4 个输入,检测层(1)~(4)模块以及卷积层 4 模块、卷积层 7 模块、2 个卷积层 5 模块。其中输入模块表示网络的输入图像,卷积层模块表示对输入图像做卷积的卷积层,下采样模块表示步长为 2 的卷积层,节点 1~5 模块表示 2 个下采样模块融合后的特征图,残差 1 表示含有 1 个残差网络结构,残差 2 表示含有 2 个残差网络结构,残差 5 表示含有 5 个残差网络结构,池化模块表示步长为 2 的最大池化层,检测层(1)~(4)模块表示检测网络的 4 个输出检测层,卷积层 4 模块表示 4 个卷积层,卷积层 7 模块表示 7 个卷积层,卷积层 5 模块表示 5 个卷积层。

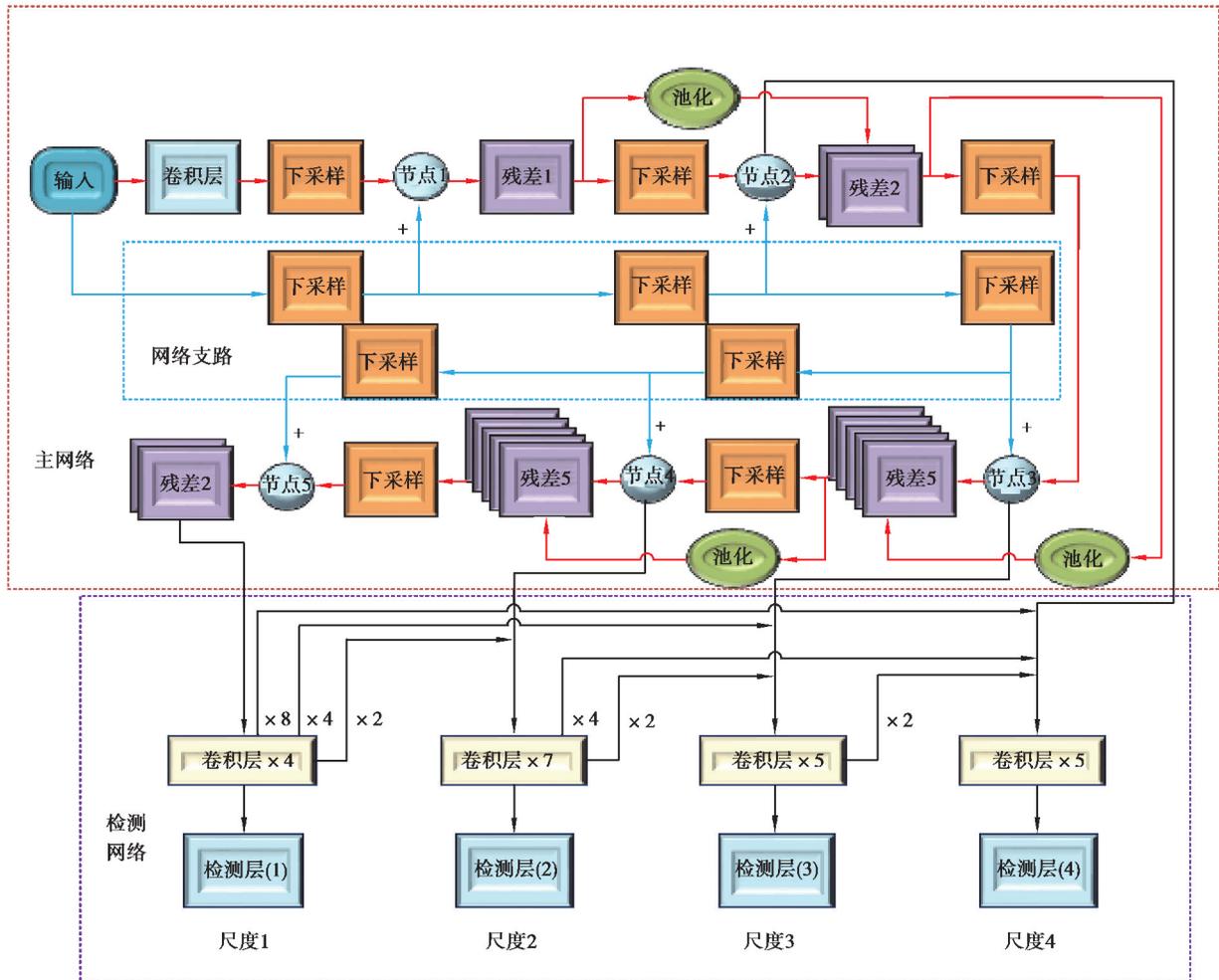


图 2 YOLOv3-fass 算法的网络结构图

Fig. 2 Network structure of YOLOv3-fass algorithm

以下为 YOLOv3-fass 算法的说明。

1) 简化网络结构。YOLOv3 算法含有大量参与运算的参数,一方面会导致车辆信息检测速度的下降,另一方面当数据集样本数量较少时算法容易产生过拟合。为了降低算法的参数数量,笔者将网络结构中的残差网络结构和滤波器的数量进行了适度的调整,分别把网络主干中 2,4,8,16,32 倍下采样后的残差网络结构的数量调整为 1,2,5,5,2 个,同时把 YOLOv3-fass 算法的网络结构中所有卷积层的通道数降低 50%,从而降低了网络结构的复杂程度,在提高车辆信息检测效率的同时也避免了算法产生过拟合现象。

2) 添加网络支路。DarkNet-53 网络结构仅采用一个下采样模块作为下采样层,在特征图传递过程中图像特征信息损失较多。为了以上问题,在网络结构中添加了一条含有 5 个下采样模块的网络支路,使该支路上的 5 个下采样模块与网络主干中相同尺度特征层的下采样模块具有相同的通道数量,并将网络主干与网络支路上的相同尺度的下采样模块进行特征融合,节点 1~5 模块便是融合之后的特征图。该设计通过添加网络支路,在特征图传递过程中减少了信息的损失,同时融合了特征图的深、浅层特征,也增强了网络结构的表征能力。

3) 添加尺度跳连结构。为了提高网络结构中浅层特征的复用率并保证相邻尺度特征图的融合,YOLOv3-fass 算法在网络主干的 208×208 , 104×104 , 52×52 , 26×26 相邻两个尺度之间添加了含有池化模块的 3 个尺度跳连结构,将下采样模块的输出作为尺度跳连结构的输入,并将 3 个尺度跳连结构的输出分别与残差 2 模块中的第 1 个残差网络结构的输出、残差 5 模块中的第 3 个残差网络结构的输出进行特征图的融合。网络结构融合了浅层特征,增强了表征能力,也起到了防止梯度弥散的作用。

4) 增加检测尺度。为了更精准地检测小尺寸的车辆信息,在 YOLOv3 网络结构的基础上增加了 1 个检测网络的检测尺度。改进的检测网络如图 2 的检测网络部分所示,首先将检测网络的 3 尺度检测改为 4 尺度检测,其次参考文献[17]中的方法,在检测网络中增加了密集连接结构,即各检测尺度之间采用 2,4,8 倍上采样进行连接,最后将主网络的第 2 个残差 2 模块、节点 4、节点 3、节点 2 模块分别与上采样的特征图融合之后输入到检测网络中。通过增加检测尺度使锚框的数量增加至原来的 4 倍之多,提高了车辆信息的检出率;同时将融合特征图作为检测网络的输入信息以及增加密集连接结构,提高了检测网络输入特征图的可靠性。

5) K-均值聚类与手动调节相结合。在本文算法中,检测网络包含了 4 个检测尺度,每个检测尺度分配 3 组锚框值,共需 12 组锚框值,并采用 K-均值方法聚类出 12 组锚框值。由于车辆数据集中标注的车辆信息的形状、车脸及车牌均为长方形且数量占比大,车标的形状多数近似正方形且在数据集中数量占比小,仅采用 K-均值方法聚类的 12 组锚框值准确度不够。因此,笔者采用 K-均值方法对车辆信息数据集进行锚框的聚类,然后通过手动调节的方式对聚类出的 12 组锚框值进行适当的调整。因为车标尺寸较小,所以只需要调整负责小目标检测的检测尺度上的锚框值,即调整前 9 组锚框值。将每 3 组中的长宽之差绝对值最小的一组锚框值调整到长宽之差绝对值小于 5 的范围内。最终,确定的 12 组锚框值为:(7,7),(19,9),(24,17),(33,29),(46,15),(66,20),(50,50),(94,26),(113,47),(203,76),(270,113),(312,180)。通过 K-Means 聚类与手动调节相结合的方法进行锚框的聚类,提高了 YOLOv3-fass 算法检测的准确度。

6) 采用迁移学习机制。针对数据集样本较少的情况,采用迁移学习机制对网络模型进行微调训练。为了加快网络的收敛速度,首先在 Pascal voc2012 数据集上训练 YOLOv3-fass 算法模型,生成新的模型权重文件 W1,然后把 W1 文件作为预训练权重文件,在自研的车辆数据集上微调 YOLOv3-fass 算法模型,得到最终的权重文件。研究中采用迁移学习机制对 YOLOv3-fass 算法进行微调,一方面在训练过程中能够使模型更快地收敛,另一方面,能够避免算法因数据集样本较少而产生过拟合。

3 实验说明

3.1 实验准备

首先为迁移学习准备了 Pascal voc2012 标准数据集,然后准备了自研的车辆数据集,用于微调本文算法。车辆数据集中的图片均来自校园和街道,共包含 2 000 张图像,图例如图 3 中(a)~(h)所示。使用 LabelImg 工具在车辆数据集中标记了 2 300 个车标,2 360 个车牌,2 520 个车脸。将数据集按 8:1:1 的比例

分成训练集、验证集和测试集,分别用于模型的训练、验证以及测试过程。



图 3 车辆数据集图例

Fig. 3 Legend of vehicle dataset

实验中使用的服务器设备的型号为超微 6027AX-TRF,内存为 32 G,硬盘为 2 T,显卡为 TITAN-XP,操作系统为 Ubuntu 16.04。在实验中采用旋转角度、饱和度、曝光度和色调共 4 种数据增强策略,具体参数配置如表 1 所示。在实验中,对本文算法以下指标进行了比较:交并比、查准率、平均精度均值、平均精度、召回率、单张图片检测时间以及模型文件大小(Ubuntu 16.04 环境下)。

表 1 参数配置

Table 1 Parameter configuration

批次数量	批次划分	动量	正则项	角度	饱和度	曝光	色调	阈值(置信度)	学习率
64	16	0.9	0.001	0	1.5	1.5	0.1	0.5	0.001

3.2 实验结果分析

分别使用 YOLOv3 算法和 YOLOv3-fass 算法迭代 10 000 次的模型对实例图像中的车辆信息做了检测试验。图 4 为 YOLOv3-fass 算法和 YOLOv3 算法在实例图像上的车辆信息检测对比图,其中图 4 中的(a)~(d)为 YOLOv3-fass 算法的车辆信息检测图,图 4 中的(e)~(h)为 YOLOv3 算法的车辆信息检测图,图中绿色框为车脸框,红色框为车标框,黄色框为车牌框。可以看出,图 4 中的(e)(g)(h)出现了车辆信息严重漏检的情况;图 4 中的(a)(e)(f)(h)出现了车辆信息错检的情况;在图 4(a)中,YOLOv3-fass 算法能够将图像中远处的所有车脸及部分车牌成功地检测出来,而(e)中,YOLOv3 算法仅仅检测出远处的一个车牌,漏检情况较为严重。通过本实验中的两组对比图,可以发现 YOLOv3-fass 算法能够检测出更多的车辆信息,而且更适合检测图像中的小尺寸车辆信息。

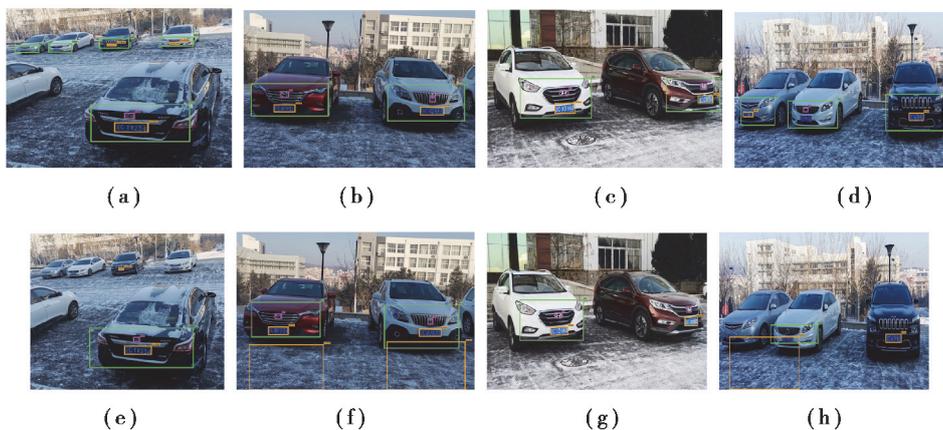


图 4 YOLOv3-fass 算法与 YOLOv3 算法车辆信息检测对比图

Fig. 4 Comparison of vehicle information detection of YOLOv3-fass algorithm and YOLOv3 algorithm

按照表1中的数据配置训练参数,将YOLOv3-fass算法与YOLOv3、YOLOv3-spp、YOLOv3-tiny算法做了对比实验。为了验证YOLOv3-fass算法的网络结构对算法性能的影响,将YOLOv3算法的DarkNet-53网络结构换成DenseNet201、ResNet50经典网络结构,并做了实验对比。在实验中,以上6种算法在车辆数据集的训练集上进行多尺度训练并迭代10 000次,按照以下的方式保存模型文件:迭代1 000次以内每迭代100次保存一次模型文件,迭代1 000次以后每迭代500次保存一次模型文件,共计保存了28个模型文件。最后在车辆数据集的测试集上对保存的6种模型文件做了测试,并将相应的指标数据记录在表2中。可以看出,YOLOv3-fass算法的单张图片检测时间相比YOLOv3算法缩短了50%,同时交并比、召回率及平均精度均值指标都有不同程度的提高;YOLOv3-fass算法的单张图片的检测速度仅慢于YOLOv3-tiny算法,但是,其交并比、召回率、平均精度均值指标比YOLOv3-tiny算法分别提高了9.26%,9.00%,6.62%;YOLOv3-fass算法与DenseNet 201、Resnet 50、YOLOv3-spp相比,单张图片的检测时间分别快17.9,8.5,11.1 ms,且其交并比、召回率、平均精度均值指标也都有大幅的提升;YOLOv3-fass算法的模型文件内存仅比YOLOv3-tiny算法的模型文件大12.7 MB,且远小于其他4种算法的模型文件,降低算法网络的参数量能够提升检测速度,也能够避免算法产生过拟合现象。综上所述,YOLOv3-fass算法不仅在车辆信息的检测速度上有了大幅的提高,而且其交并比、召回率及平均精度均值也都有小幅提高。

表2 6种算法性能指标对比表

Table 2 Multi-performance indices of six algorithms

方 法	交并比/ %	召回率/ %	平均精度 均值/%	单张图检测 时间/ms	模型文件 大小/MB
YOLOv3	82.43	94	90.75	20.4	246.3
YOLOv3-tiny	75.22	86	84.26	3.2	34.7
DenseNet 201	73.55	87	87.25	28.1	82.8
ResNet 50	82.47	92	90.65	18.7	168.7
YOLOv3-spp	82.92	93	90.48	21.3	250.6
YOLOv3-fass	84.48	95	90.88	10.2	47.4

如图5、图6所示,实验验证了6种算法的车标平均精度以及3类车辆信息平均精度均值的稳定性,并绘制了车标平均精度和车辆信息平均精度均值的稳定性对比图。

图5为6种算法的车标平均精度稳定性对比图,从图中可以看出,YOLOv3-fass算法的车标平均精度提升速度最快,且在迭代2 000次之后车标平均精度曲线趋于稳定,相比之下ResNet 50的车标平均精度曲线在迭代3 500次之后也能达到相对稳定的态势,但其曲线的平滑度以及车标平均精度值均弱于YOLOv3-fass算法;YOLOv3和YOLOv3-spp算法的车标平均精度在训练过程中出现了多次的严重下跌,而且其提升速度和车标平均精度大小均弱于YOLOv3-fass算法;YOLOv3-tiny,DenseNet 201的车标平均精度曲线均严重偏离其他算法。综上所述YOLOv3-fass算法的车标平均精度的稳定性更好,更适合检测车辆数据集中小尺寸车辆信息。

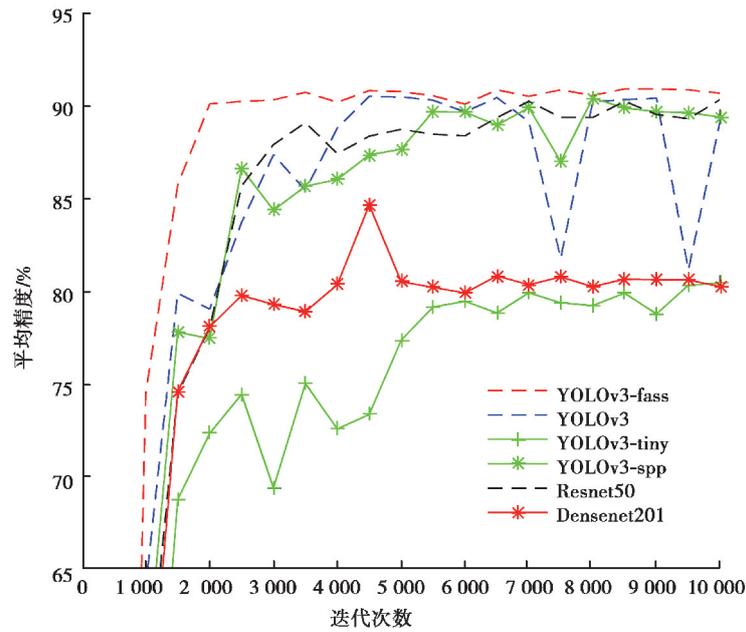


图 5 车标平均精度稳定性对比图

Fig. 5 Stability of average accuracy of vehicle logo with six algorithms

图 6 为 6 种算法的车辆信息平均精度均值稳定性对比图,从图中可以看出, YOLOv3-tiny 算法与 DenseNet 201 算法的车辆信息平均精度均值曲线位于其他 4 种算法的下方且严重偏离,其稳定性差; YOLOv3 和 YOLOv3-spp 算法的车辆信息平均精度均值曲线随着迭代次数的增加均出现了不同程度的震荡,而且位于 YOLOv3-fass 算法的下方; ResNet 50 的车辆信息平均精度均值曲线的稳定性以及平均精度均值均弱于 YOLOv3-fass 算法。综上所述,在车辆数据集上随着迭代次数的增加 YOLOv3-fass 算法的车辆信息的平均精度均值更加稳定。

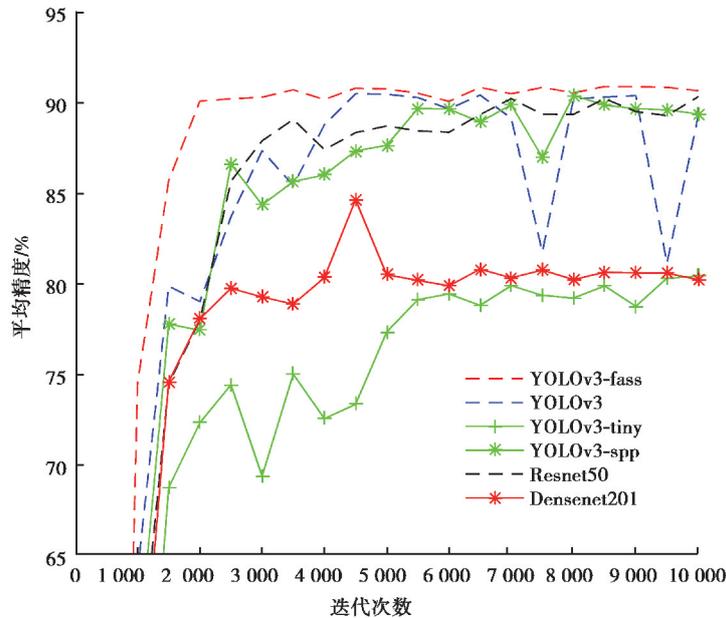


图 6 车辆信息平均精度均值稳定性对比图

Fig. 6 Stability of vehicle information mean average accuracy with six algorithms

通过以上实验对比,可以得出结论,YOLOv3-fass 算法更适合车辆信息的检测。为了进一步确认在何种条件下 YOLOv3-fass 算法能够取得更佳的实验效果,进行了以下实验:首先 YOLOv3-fass 算法在车辆数据集上迭代 10 000 次并保存实验模型;然后选择 320×320 , 352×352 , 384×384 , 416×416 , 448×448 , 480×480 , 512×512 , 544×544 , 576×576 , 608×608 共 10 个尺寸作为输入图像的尺寸,选择 0.5 和 0.75 作为阈值(NMS),进行 10 次测试实验,并记录在不同输入尺寸、不同阈值(NMS)时实验的交并比和查准率,表 3 为实验测试数据,其中阈值(NMS)为 0.50 时对应交并比和查准率各为 1 时的数据,阈值(NMS)为 0.75 时对应交并比和查准率各为 2 时的数据。通过分析数据可以得出,当输入图像尺寸为 448×448 时,2 种阈值(NMS)情况下算法的交并比均能够达到最优值,并且在阈值(NMS)为 0.75 时,查准率更高。因此在输入图像尺寸为 448×448 、阈值(NMS)为 0.75 时,YOLOv3-fass 算法的车辆信息检测效果更好。

表 3 YOLOv3-fass 算法的多条件性能指标

Table 3 Multi-conditional performance indices of YOLOv3-fass Algorithm

尺寸	交并比 1/%	查准率 1/%	交并比 2/%	查准率 2/%
320×320	82.77	98	83.76	99
352×352	83.34	98	84.47	99
384×384	84.36	99	84.92	99
416×416	84.48	98	85.19	99
448×448	84.83	98	85.55	99
480×480	84.49	98	85.14	99
512×512	84.41	98	85.51	99
544×544	84.08	98	84.91	98
576×576	83.97	98	84.76	99
608×608	83.39	98	84.24	99

4 结 论

针对研究中发现的一系列问题,笔者对 YOLOv3 算法进行了改进,并提出了 YOLOv3-fass 算法。经过实验验证,改进后算法的稳定性更好,且能够更精准地检测出图像中的车辆信息,相比原始 YOLOv3 算法,单张图像的检测时间降低了 50%。另外对比了在不同输入尺寸、不同阈值时 YOLOv3-fass 算法的交并比、查准率性能指标,最终得出结论,当图像尺寸为 448×448 、阈值(NMS)为 0.75 时,YOLOv3-fass 算法能够取得更佳的检测效果。但是,本文算法在检测车辆信息时仍存在一定的缺陷,如图像中的车辆信息处于复杂场景中时,易出现错检的情况。今后将针对发现的问题进行相应的改进,提高算法在复杂场景中车辆信息检测的性能。

参考文献:

- [1] Zakaria Y, Ali M A, Abd El Munim H E, et al. A novel vehicle detection system[C]//2018 13th International Conference on Computer Engineering and Systems (ICCES). December 18-19, 2018, Cairo, Egypt. IEEE, 2018: 127-131.
- [2] 张小琴, 赵池航, 沙月进, 等. 基于 HOG 特征及支持向量机的车辆品牌识别方法[J]. 东南大学学报(自然科学版), 2013, 43(S2): 410-413.
Zhang X Q, Zhao C H, Sha Y J, et al. Vehicle brand recognition based on HOG feature and support vector machine[J]. Journal of Southeast University (Natural Science Edition), 2013, 43(S2): 410-413. (in Chinese)

- [3] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]// 2014 IEEE Conference on Computer Vision and Pattern Recognition. June 23-28, 2014, Columbus, OH, USA. IEEE, 2014: 580-587.
- [4] Girshick R. Fast R-CNN[C]// 2015 IEEE International Conference on Computer Vision (ICCV). December 7-13, 2015, Santiago, Chile. IEEE, 2015: 1440-1448.
- [5] Ren S Q, He K M, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[C]// IEEE Transactions on Pattern Analysis and Machine Intelligence. IEEE, : 1137-1149.
- [6] Liu W, Anguelov D, Erhan D, et al. SSD: single shot MultiBox detector[M]// Computer Vision - ECCV 2016. Cham: Springer International Publishing, 2016: 21-37.
- [7] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection[C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). June 27-30, 2016, Las Vegas, NV, USA. IEEE, 2016: 779-788.
- [8] Redmon J, Farhadi A. YOLO9000: better, faster, stronger[C]// 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). July 21-26, 2017, Honolulu, HI, USA. IEEE, 2017: 6517-6525.
- [9] Redmon J, Farhadi A. YOLOv3: an incremental improvement[OL]. [2018-04-08]. [https:// arxiv.org/abs/1804.02767](https://arxiv.org/abs/1804.02767).
- [10] Huang Z J, Fu M X, Ni K L, et al. Recognition of vehicle-logo based on faster-RCNN[M]// Lecture Notes in Electrical Engineering. Singapore: Springer Singapore, 2018: 75-83.
- [11] Chen Q Q, Huang N, Zhou J M, et al. An SSD algorithm based on vehicle counting method[C]// 2018 37th Chinese Control Conference (CCC). July 25-27, 2018, Wuhan. IEEE, 2018: 7673-7677.
- [12] He Y, Li L Q. A novel multi-source vehicle detection algorithm based on deep learning[C]// 2018 14th IEEE International Conference on Signal Processing (ICSP). August 12-16, 2018, Beijing, China. IEEE, 2018: 979-982.
- [13] 桑军, 郭沛, 项志立, 等. Faster-RCNN 的车型识别分析[J]. 重庆大学学报, 2017, 40(7): 32-36.
Sang J, Guo P, Xiang Z L, et al. Vehicle detection based on faster-RCNN[J]. Journal of Chongqing University, 2017, 40(7): 32-36. (in Chinese)
- [14] Yang L Q, Sang N, Gao C X. Vehicle parts detection based on Faster - RCNN with location constraints of vehicle parts feature point[C]// MIPPR 2017: Pattern Recognition and Computer Vision. October 28-29, 2017, Xiangyang, China. SPIE, 2018: 106091J.
- [15] Krittanayawach N, Vateekul P. Robust compression technique for YOLOv3 on real-time vehicle detection[C]// 2019 11th International Conference on Information Technology and Electrical Engineering (ICITEE). October 10-11, 2019, Pattaya, Thailand. IEEE, 2019: 1-6.
- [16] Chen P Y, Hsieh J W, Gochoo M, et al. Smaller object detection for real-time embedded traffic flow estimation using fish-eye cameras[C]// 2019 IEEE International Conference on Image Processing (ICIP). September 22-25, 2019, Taipei, Taiwan, China. IEEE, 2019: 2956-2960.
- [17] 戴伟聪, 金龙旭, 李国宁, 等. 遥感图像中飞机的改进 YOLOv3 实时检测算法[J]. 光电工程, 2018, 45(12): 84-92.
Dai W C, Jin L X, Li G N, et al. Real-time airplane detection algorithm in remote-sensing images based on improved YOLOv3[J]. Opto-Electronic Engineering, 2018, 45(12): 84-92.(in Chinese)

(编辑 张 苹)