

基于改进YOLOv4-tiny的分心驾驶行为检测

魏启康, 朱文忠, 江嘉文, 谢鑫煌

(四川轻化工大学计算机科学与工程学院, 四川 宜宾 644000)

摘要:为实现对常见分心驾驶行为(打电话、抽烟、喝水)进行高精度的实时检测,提出基于改进YOLOv4-tiny的分心驾驶行为检测算法。首先针对香烟这类小目标检测精度差的问题,将主干特征提取网络中最后一个跨阶段局部连接(Cross Stage Partial connections, CSP)层的输出特征进行卷积和上采样之后与第二个CSP层的输出特征进行特征融合,增加了一个52×52的预测尺度,提升对小目标的检测能力;其次在特征金字塔中添加高效通道注意力机制(Efficient Channel Attention, ECA)模块,提升模型的检测精度;最后使用k-means聚类算法在自制分心驾驶行为数据集上重新确定先验框,使用迁移学习、余弦退火学习率和标签平滑进行模型训练。结果表明,本文方法的所有类别平均精度(mAP)为98.88%,相较于原YOLOv4-tiny算法提高了2.83%,其中对于检测香烟的平均精度值提高了8.46%,提升了对香烟的检测能力。改进YOLOv4-tiny的分心驾驶行为检测算法具有较好的综合性能,有利于车载系统对驾驶员的分心驾驶行为进行实时检测并提醒,对减少交通安全事故具有一定的现实意义。

关键词:分心驾驶;YOLOv4-tiny;预测尺度;ECA注意力机制

中图分类号:TP391.41

文献标识码:A

引言

根据《中国统计年鉴-2021》公布的数据,2020年我国有超过4万人死于汽车导致的交通事故^[1],有14%~33%的重大交通事故是驾驶员出现分心驾驶行为导致^[2],这说明驾驶员存在分心驾驶行为是导致交通事故发生的一个重要原因。根据国际标准

化组织的定义,分心是驾驶员将注意力转向与正常驾驶无关的活动,从而导致驾驶员操作能力下降的一种现象^[3]。分心根据干扰源可以分为内部分心和外部分心,而内部分心又可以分为由驾驶员引起(如打电话、抽烟、喝水等)和非驾驶员引起(如乘客的干扰等)^[4]。对于内部分心中由驾驶员自身引起的部分分心行为,如果能及时检测并通过报警提

收稿日期:2022-04-16

基金项目:四川省科技创新(苗子工程)培育项目(2022049);企业信息化与物联网测控技术四川省高校重点实验室开放基金(2020WZJ02)

作者简介:魏启康(1996-),男,硕士生,研究方向为深度学习与目标检测,(E-mail)qk_w1996@163.com

通信作者:朱文忠(1971-),男,教授,硕士,研究方向为企业信息化与集成、物联网技术及应用、智能信息处理、最优控制等,(E-mail)zww@suse.edu.cn

示,对保障驾驶员与乘客人身安全和减少交通事故发生率具有重大意义。

目前有多种方法可以实现对驾驶员分心行为的检测。Sahayadhas等^[5]提出了一种监测驾驶员生理信号的检测方法,通过心电图(ECG)和表面肌电图(sEMG)信号判断驾驶员是否具有分心驾驶行为。夏瀚笙等^[6]提出了一种基于人体关键点位置的检测方法,使用Alpha Pose系统获取驾驶员的人体关键点位置识别驾驶员是否具有分心驾驶行为。Hu等^[7]设计了一个轻量级架构的半级连网络,根据识别到的人脸部和手部的形态判断驾驶员是否存在分心驾驶行为。叶绿等^[8]提出了一种结合图像分割与卷积神经网络的分心驾驶行为检测方法。

随着目标检测技术的发展,识别精度和速度都在不断提高。本文提出一种基于轻量级目标检测算法的检测方法,该方法通过目标检测技术识别驾驶员手中是否有香烟、手机、水杯等物品,以判断驾驶员是否具有分心驾驶行为,进而保障驾驶员和乘客人身安全、减少交通事故的发生。

1 理论基础

1.1 目标检测算法

目标检测算法主要可以分为两阶段检测算法和一阶段检测算法。两阶段检测算法主要有R-CNN^[9]、Fast R-CNN^[10]、Faster R-CNN^[11]等,它主要由骨干网络、区域建议模块和检测头构成。区域建议模块可以生成包含检测目标的很多候选区域,然后通过区域建议网络、骨干网络结构和区域改进等方法来调整候选框,最后得到检测结果。因此,两阶段的目标检测算法需要更多的计算资源,依赖于更好的硬件基础。一阶段目标检测算法如YOLO

(You Only Look Once)系列^[12-15]、SSD(Single Shot MultiBox Detector)系列^[16-18]等,它主要是采用端到端的检测方法,不需要区域建议寻找目标^[19],将物体的分类和定位都放在一个步骤中进行,从而使得训练过程相对简单,可以更加快速地识别出目标类别和位置,但是准确度稍低于两阶段检测算法。

1.2 YOLOv4-tiny算法

YOLOv4-tiny算法在YOLOv4算法的基础上去掉了空间金字塔池化(Spatial Pyramid Pooling, SPP)模块,减少了模型的计算量,具有更快的目标检测速度^[15]。YOLOv4-tiny算法的参数量几乎是YOLOv4的十分之一,这也导致检测精度方面略低于YOLOv4。YOLOv4-tiny算法采用CSPDarknet53-tiny作为主干特征提取网络,主要完成普通卷积、残差块堆叠和下采样操作。普通卷积使用CBL结构,由卷积层(Convolutional)、批量标准化层(Batch Normalization, BN)和激活函数层(LeakyReLU)组成。LeakyReLU函数如式(1)所示。

$$y_i = \begin{cases} x_i, & x_i \geq 0 \\ \frac{x_i}{a_i}, & x_i < 0 \end{cases} \quad (1)$$

其中: $a_i \in (1, +\infty)$ 是一个常量; x_i 表示输入; y_i 表示输出。

BN层可以有效地避免梯度消失和梯度爆炸的问题,提高模型的泛化能力。残差块堆叠过程采用CSP(Cross Stage Partial)连接结构,加强网络的学习能力^[20],通过最大池化(Maxpool)完成下采样的过程。加强特征提取网络采用特征金字塔结构(Feature Pyramid Networks, FPN),实现低分辨率强语义信息的特征图与高分辨率弱语义信息的特征图进行特征融合^[21],减少计算量的增加,而且提高特征提取的能力。YOLOv4-tiny模型如图1所示。

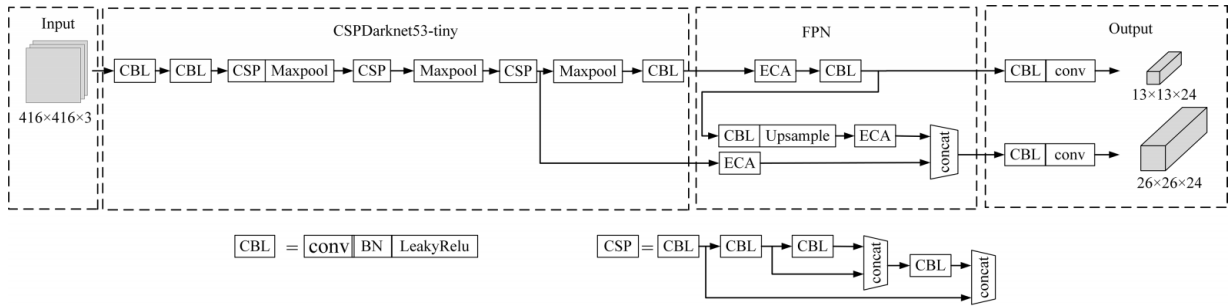


图1 YOLOv4-tiny模型结构

2 构建YOLOv4-tiny-HE网络

2.1 增加预测尺度

由于YOLOv4-tiny网络中只有13x13和26x26两个预测尺度,所以它对大型目标和中型目标的检测效果较好,但对于香烟这类小型目标的检测效果相对较差。YOLOv4-tiny主干特征提取网络中包含了3个CSP模块,第二个CSP模块中包含的位置信息更加精准,细节信息也更多,但是语义信息相对较少;第三个CSP模块中包含了更多的语义信息,但是细节信息较少、位置信息不够精准,这样会导致小目标的位置信息和细节信息的丢失^[22]。本文在原YOLOv4-tiny模型的加强特征提取网络FPN中增加一条与主干特征提取网络中第二个CSP层相连的路径,将第三个CSP层中的深层特征进行卷积和上采样操作之后与第二个CSP层的浅层特征进行特征融合,再经过一个CBL和conv层,最后得到一个52x52的预测尺度。最终改进后的YOLOv4-tiny模型拥有3个预测尺度,分别是13x13、26x26、52x52,从而加

强了对香烟这类小目标的检测能力。

2.2 注意力机制的引入

YOLOv4-tiny算法只能通过神经网络自发地传递特征信息,没有对重点信息进行筛选,从而会忽略掉一些小目标和部分遮挡目标的信息。当引入注意力机制之后,网络可以动态地对感兴趣区域的特征数据进行权重系数加权,从而加大对重点区域的关注,增强了对小目标和部分遮挡目标的识别。本文引入的ECA(Efficient Channel Attention)注意力机制^[23]是对SENet(Squeeze-and-Excitation Networks)^[24]的一种改进,如图2所示。SENet在对特征图降维之后,会对通道注意力的预测产生副作用,导致网络性能的降低,并且捕获所有通道的依赖关系是低效和不必要的。而ECA模块可以在不改变输入特征通道的维度的情况下,通过卷积核为k的一维卷积,完成局部跨通道融合的策略,从而达到了降低模型计算量与复杂度的效果。通过在不同位置添加ECA注意力机制模块,可以加强对于小目标和部分遮挡目标的识别精度,同时减少计算量。

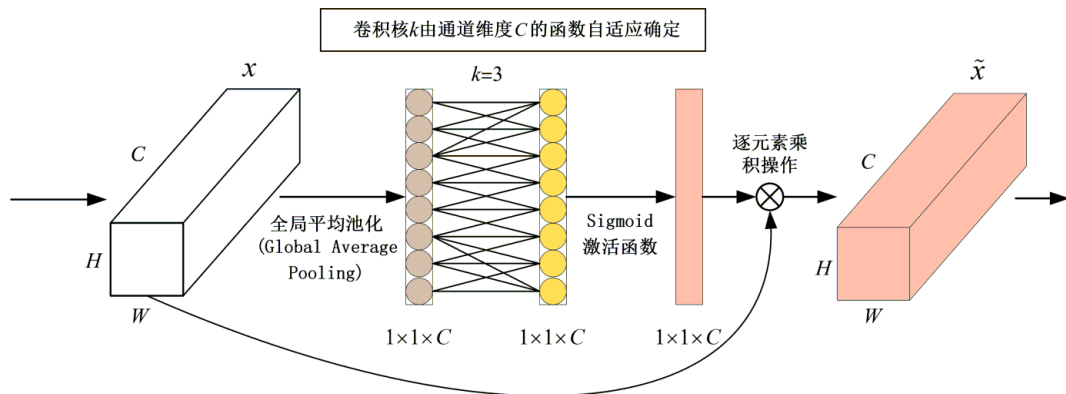


图2 ECA注意力机制模块

图2中, W 、 H 和 C 分别表示宽度、高度和通道维度, x 和 \tilde{x} 分别表示卷积块经过处理前、后的输出, 卷积核 k 的取值是由通道维度 C 的函数自适应确定, 如式(2)所示。

$$k = \left\lfloor \frac{\log_2(C)}{\gamma} + \frac{b}{\gamma} \right\rfloor_{\text{odd}} \quad (2)$$

其中: k 表示一维卷积的内核大小, $\lfloor t \rfloor_{\text{odd}}$ 表示最接近 t

的奇数, C 为通道数维数, $b = 1, \gamma = 2$ 。

对YOLOv4-tiny增加一个 52×52 预测尺度, 在FPN中增加ECA注意力机制模块以后, 可以加强模型对小目标和部分遮挡目标的检测效果, 提升模型的整体精度。改进后的YOLOv4-tiny-HE模型结构如图3所示。

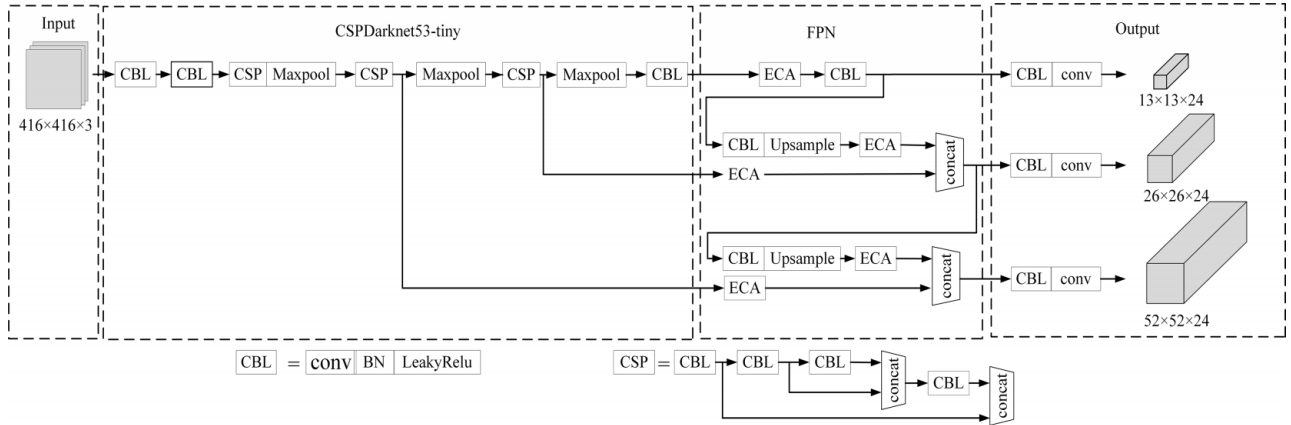


图3 YOLOv4-tiny-HE模型结构

2.3 先验框聚类

原始YOLOv4-tiny算法中只有两个预测尺度, 而本文中增加了一个 52×52 的预测尺度, 因此需要增加相应的先验框。为了使YOLOv4-tiny-HE算法更加准确地预测出目标的位置, 本文使用k-means算法针对自制的分心驾驶行为数据集进行重新聚类, 从而得到更加具有代表性、更加精准的先验框。最后在自制的分心驾驶数据集上通过k-means算法重新聚类出的先验框分别是 (16×29) 、 (33×20) 、 (35×109) 、 (35×43) 、 (52×72) 、 (67×115) 、 (94×198) 、 (130×128) 以及 (159×197) 。

2.4 学习率余弦退火衰减

模型训练过程中对于学习率的调整是非常关键的, 学习率余弦退火衰减^[25]可以使学习率在下降的过程中周期性的重启, 学习率先急速下降, 再陡然升高, 不断的重复这个过程。这样做是为了跳过当前的局部最优点, 寻找新的最优解, 从而实现逼

近全局最优解的目的。本文采用PyTorch深度学习框架中内置的余弦退火算法来优化学习率。

2.5 标签平滑

深度学习中采用独热编码(One-hot)标签来计算交叉熵损失, 真实标签位置的One-hot标签为1, 反之为0。训练过程中仅考虑率样本中One-hot标签为1的损失, 忽略One-hot值为0的损失, 但这使得模型易出现过拟合的现象, 泛化能力差。标签平滑^[26]是一种正则化技术, 可以防止模型在训练过程中过分的相信真实标签, 从而防止发生过拟合的现象, 提高模型的泛化能力。标签平滑如式(3)所示。

$$y' = y \times (1 - \varepsilon) + \frac{\varepsilon}{k} \quad (3)$$

其中: y' 表示标签平滑后的数据标签; y 表示数据的One-hot标签; ε 表示平滑因子; k 表示类别个数。

2.6 迁移学习

迁移学习^[27]是指将一个模型在其他任务中学习到的知识应用于另一个新的模型训练当中。通过迁移学习的策略可以有效地减少新模型对于大量数据的依赖和训练时间,提高新模型的学习效率和准确率。因此本文采用迁移学习的训练策略,可以加快模型训练过程中的收敛速度,使模型训练过程变得相对快速和简单。

3 实验与结果分析

为了验证改进之后的YOLOv4-tiny-HE网络模型在分心驾驶行为检测时的有效性,本实验使用未改进的YOLOv4-tiny和改进之后的YOLOv4-tiny-HE算法在自制的分心驾驶行为数据集上进行训练和测试,比较两者的性能。实验平台配置见表1。

表1 计算机配置

名称	配置
操作系统	Windows 11
CPU	11th Gen Intel® Core™ i5-11400H @2.70 GHz 2.69 GHz
GPU	NVIDIA GeForce RTX 3060
RAM	16 G
显存	6 G
深度学习框架	PyTorch 1.10.1
加速环境	CUDA 11.3 CUDNN 8.2

3.1 数据集构建

目前在分心驾驶的研究中尚未有公开且成熟的专用数据集,虽然State Farm数据集公开,但仅授权用于竞赛。本文使用的驾驶员分心驾驶行为数据集是召集了18名志愿者,封闭路段内每名志愿者在车内录制一段模拟分心驾驶(打电话、喝水、抽烟)视频,然后通过视频帧生成图片。为了提高模型的泛化能力和鲁棒性,本文对图像数据集进行了水平和垂直旋转、调节亮度、镜面翻转、加入高斯噪声的数据增强操作,增加模型训练的数据量。最终

分心驾驶数据集包含8106幅图像,分心驾驶数据集采用Pascal VOC2007标准格式,使用LabelImg图片标志工具对分心驾驶行为数据集中的手机、香烟、水瓶用矩形框标记,标注后会生成与图片对应的XML文件并存储。最后将整个分心驾驶行为数据集分为训练集、测试集与验证集3个部分,占比分别为80%、10%和10%。

3.2 性能评价指标

为了能够准确评估改进网络模型的检测效果,本次实验中使用准确率(Precision)、召回率(Recall)、F1分数、mAP(Mean Average Precision)以及检测速度作为评价指标。Precision和Recall计算如式(4)~(5)所示,F1分数是Precision和Recall的调和平均数,如式(6)所示。mAP是指所有类别AP的均值,数值越大表示模型的识别准确率越高,计算如式(8)所示;AP的计算如式(7)所示。检测速度是指每秒检测图像的帧数,数值越高表示检测速度越快。

$$Precision = \frac{TP}{TP + FP} \quad (4)$$

$$Recall = \frac{TP}{TP + FN} \quad (5)$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (6)$$

$$AP = \int_0^1 P(R) dR \quad (7)$$

$$mAP = \frac{\sum_{i=1}^k AP_i}{k} \quad (8)$$

其中:TP、FP和FN分别表示预测为正类的正样本、预测为正类的负样本和预测为负类的正样本,k表示类别数,P(R)表示由Precision和Recall构成的曲线,AP表示Precision-Recall曲线下的面积,AP_i表示第i类别的AP值。

3.3 模型训练

实验均在以下参数设置中进行,输入图片为 416×416 ;训练分为两个阶段,Epoch设置为200轮;前100轮冻结阶段的batch_size设置为16,learning_rate设置为0.001,此阶段模型的主干被冻结。后100轮解冻阶段的batch_size设置为8;learning_rate设置为0.0001,此阶段模型的主干被解冻;label_smoothing设置为0.01;训练过程Loss曲线如图4所示。随着Epoch的不断增加,Loss损失值不断降低,训练进行到125轮后损失值趋于稳定,YOLOv4-tiny-HE算法和原始YOLOv4-tiny算法的损失值分别收敛于1.1和0.9左右,这将有利于提高检测的准确度。

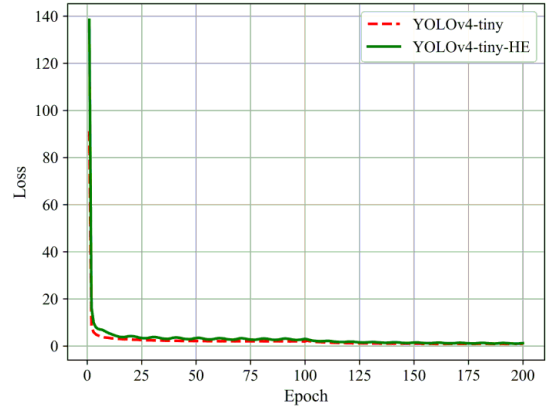


图4 训练损失曲线

YOLOv4-tiny 和 YOLOv4-tiny-HE 的 mAP 值对比如图5所示。从图5可以看出改进之后的 mAP 值从96.05%提高到98.88%,提高了2.83%,其中对于检测香烟的 AP 值提高了8.46%,这表明增加一个 52×52 的预测尺度和添加ECA注意力机制模块,能够有效提高小目标的检测能力。

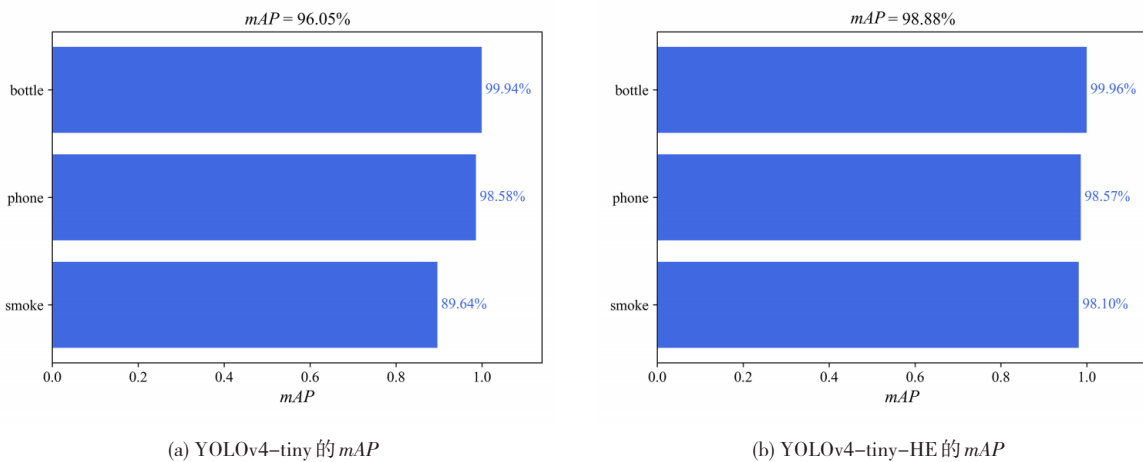


图5 YOLOv4-tiny改进前后 mAP 对比图

3.4 检测结果对比

随机选取测试集中的图片分别用YOLOv4-tiny和YOLOv4-tiny-HE模型进行测试,测试的结果如图6所示,从左往右分别是对香烟、手机和水瓶的检测,两个算法都可以准确地检测到目标。水瓶的置信度都是1.00;手机的置信度分别为0.98和0.99,在

对手机和水瓶这类中大型物品的检测置信度都呈现一个比较高的水平。香烟的置信度分别为0.80和0.98,可以看出YOLOv4-tiny-HE模型对香烟这类小物品的置信度有一个明显的提升,达到了对模型改进的预期效果。



图6 改进前后结果对比

3.5 消融实验

为了进行充分的对比,本文进行了消融实验,改进1表示仅添加ECA注意力机制,改进2表示仅增加52×52预测尺度。分别对YOLOv4-tiny模型与仅添加ECA注意力机制的YOLOv4-tiny-E模型、仅

增加52×52预测尺度的YOLOv4-tiny-H模型以及同时增加52×52预测尺度和加入ECA注意力机制的YOLOv4-tiny-HE模型进行对比实验,其中Precision、Recall和F1值是由宏平均(Macro-averaging)的方式计算得出,结果见表2。

表2 消融实验

模型	改进1	改进2	mAP/%	Precision/%	Recall/%	F1/%
YOLOv4-tiny	×	✓	96.05	96.42	94.58	95.33
YOLOv4-tiny-E	✓	×	96.62	96.99	96.22	96.67
YOLOv4-tiny-H	×	✓	98.01	97.71	95.86	96.67
YOLOv4-tiny-HE	✓	✓	98.88	98.87	96.82	97.67

从表2中可以看出,与YOLOv4-tiny模型相比,当仅加入ECA注意力机制或者仅增加一个52×52的预测尺度时,模型的mAP、Precision、Recall和F1分数均有一个小幅度的增长,证明两种改进方法都会对模型起到促进作用。当两种改进方法同时都添加到YOLOv4-tiny模型中时,模型的mAP、Precision、Recall和F1分数均达到了最高值,此时YOLOv4-

tiny-HE模型的性能为最佳。

3.6 对比实验

通过与其他主流的目标检测模型在自制分心驾驶行为数据集上的训练结果进行对比,验证YOLOv4-tiny-HE模型的综合性能,试验采用的评价指标包括mAP、模型大小和检测速度,结果见表3。

表3 不同模型性能对比

模型	<i>mAP</i> /%	模型大小/MB	检测速度/FPS
Faster-RCNN-ResNet50	94.06	108	9.5
SSD-MobileNetv2	93.99	15.3	100
EfficientDet	93.66	15.0	23.95
YOLOv4-tiny	96.05	22.4	155
YOLOv4-tiny-HE	98.88	24.2	125

从表3中可以看出, YOLOv4-tiny-HE模型的*mAP*值相较于Faster-RCNN-ResNet50、SSD-MobileNetv2、EfficientDet、YOLOv4-tiny分别提高了4.82%、4.89%、5.22%、2.83%, 优于其他算法。Faster-RCNN-ResNet50的模型最大并且检测速度最低, 不满足实时检测的要求。相较于SSD-MobileNetv2和EfficientDet, YOLOv4-tiny-HE模型较大, 但是*mAP*和检测速度均更高。相较于YOLOv4-tiny, YOLOv4-tiny-HE模型略大、检测速度略低, 但是*mAP*值高于YOLOv4-tiny且仍满足实时检测的要求, 所以YOLOv4-tiny-HE模型拥有较好的综合性能。

4 结束语

本文提出一种改进YOLOv4-tiny的分心驾驶行

为检测算法, 实现了对驾驶员自身部分分心驾驶行为的检测。通过增加一个52×52的预测尺度和在加强特征提取网络的不同地方添加ECA注意力机制模块提高了模型的精确度, 尤其提高了对香烟这类小目标的检测精度。通过构建分心驾驶行为数据集并使用翻转、高斯噪声等方法对数据进行增强操作, 提高了模型在实际检测场景中的泛化能力。使用k-means算法重新聚类先验框, 采用迁移学习、余弦退火学习率和标签平滑的策略进行模型训练, 实验结果表明, YOLOv4-tiny-HE模型的*mAP*值比YOLOv4-tiny模型提高了2.83%, 其中对于香烟这类小目标检测的*AP*值提高了8.46%, 具有较好的综合性能。因此YOLOv4-tiny-HE模型有利于车载系统对驾驶员的分心驾驶行为进行实时检测并提醒, 对减少交通安全事故具有一定的现实意义。

参考文献:

- [1] 中华人民共和国统计局. 中国统计年鉴—2021[M]. 北京: 国家统计局出版社, 2021.
- [2] PARKER D, REASON J, MANSTEAD A, et al. Driving errors, driving violations and accident involvement[J]. Ergonomics, 1995, 38(5): 1036-1048.
- [3] 秦斌斌, 彭良康, 卢向明, 等. 司机分心驾驶检测研究进展[J]. 计算机应用, 2021, 41(8): 2330-2337.
- [4] 葛慧敏, 郑明强, 吕能超, 等. 驾驶分心综述[J]. 交通运输工程学报, 2021, 21(2): 38-55.
- [5] SAHAYADHAS A, SUNDARAJ K, MURUGAPPAN M, et al. A physiological measures-based method for detecting inattention in drivers using machine learning approach[J]. Biocybernetics and Biomedical Engineering, 2015, 35(3): 198-205.
- [6] 夏瀚笙, 沈焯, 胡委. 基于人体关键点的分心驾驶行为识别[J]. 计算机技术与发展, 2019, 29(7): 1-5.
- [7] HU J, LIU W, KANG J, et al. Semi-cascade network for driver's distraction recognition[J]. Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering, 2019, 233(9): 2323-2332.
- [8] 叶绿, 陈斌, SUGIANTO S, 等. 基于图像分割的驾驶员分心行为识别研究[J]. 浙江科技学院学报, 2020, 32(3): 209-215.

- [9] GIRSHICK R,DONAHUE J,DARRELL T,et al.Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,Columbus,USA,June 23-28,2014:580-587.
- [10] GIRSHICK R.Fast R-CNN[C]//Proceedings of the IEEE International Conference on Computer Vision,Santiago,Chile,December 7-13,2015: 1440-1448.
- [11] REN S,HE K,GIRSHICK R,et al.Faster R-CNN:towards real-time object detection with region proposal networks[J].IEEE Transactions on Pattern Analysis & Machine Intelligence,2017,39(6):1137-1149.
- [12] REDMON J,DIVVALA S,GIRSHICK R,et al.You only look once:unified,real-time object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,Las Vegas,USA,June 27-30,2016:779-788.
- [13] REDMON J,FARHADI A. YOLO9000: better, faster, stronger[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,Honolulu,USA,July 21-26,2017:7263-7271.
- [14] ZHANG H,QIN L,LI J,et al.Real-time detection method for small traffic signs based on Yolov3[J].IEEE Access,2020,8:64145-64156.
- [15] DU S,ZHANG P,ZHANG B,et al.Weak and occluded vehicle detection in complex infrared environment based on improved YOLOv4[J]. IEEE Access,2021,9:25671-25680.
- [16] LIU W,ANGUELOV D,ERHAN D,et al.SSD:single shot multibox detector[C]//Computer Vision-ECCV 2016:14th European Conference, Amsterdam,the Netherlands,October 10-14,2016:21-37.
- [17] 邝先验,杨江波,张建华.基于改进的DSSD算法的行人检测[J].中国仪器仪表,2021(5):21-27.
- [18] 高洁,王战红,刘纲.基于FSSD的微光烟雾检测方法[J].电子测量技术,2021,44(5):123-128.
- [19] 肖体刚,蔡乐才,汤科元,等.改进SSD的安全帽佩戴检测方法[J].四川轻化工大学学报(自然科学版),2020,33(4):68-76.
- [20] 汤科元,刘川莉,蔡乐才,等.基于YOLO轻量化网络的交通标志检测算法[J].四川轻化工大学学报(自然科学版),2021,34(5):62-70.
- [21] LIN T,DOLLÁR P,GIRSHICK R,et al.Feature pyramid networks for object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,Honolulu,USA,July 21-26,2017:2117-2125.
- [22] 胡嘉沛,李震,黄河清,等.采用改进YOLOv4-Tiny模型的柑橘木虱识别[J].农业工程学报,2021,37(17):197-203.
- [23] WANG Q,WU B,ZHU P,et al.ECA-net:efficient channel attention for deep convolutional neural networks[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR),Seattle,USA,June 13-19,2020:11534-11542.
- [24] HU J,SHEN L,SUN G.Squeeze-and-excitation networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR),Salt Lake City,USA,June 18-23,2018:7132-7141.
- [25] 蔡舒婷,闫子砚,师利中.基于改进YOLOv4的航空发动机损伤检测方法[J].现代制造工程,2023,509(2):99-108.
- [26] SZEGEDY C,VANHOUCKE V,IOFFE S,et al.Rethinking the inception architecture for computer vision[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR),Las Vegas,USA,June 27-30,2016:2818-2826.
- [27] ZHUANG F,QI Z,DUAN K,et al.A comprehensive survey on transfer learning[J].Proceedings of the IEEE,2020,109(1):43-76.

引用格式:

中文:魏启康,朱文忠,江嘉文,等.基于改进YOLOv4-tiny的分心驾驶行为检测[J].四川轻化工大学学报(自然科学版),2023,36(2):

67-76.

英文:WEI Q K,ZHU W Z,JIANG J W,et al.The distracted driving behavior detection based on improved Yolov4-tiny[J].Journal of Sichuan University of Science & Engineering (Natural Science Edition),2023,36(2):67-76.

The Distracted Driving Behavior Detection Based on Improved Yolov4-tiny

WEI Qikang, ZHU Wenzhong, JIANG Jiawen, XIE Xinhuang

(School of Computer Science and Engineering, Sichuan University of Science & Engineering,
Yibin 644000, China)

Abstract: To achieve high-precision real-time detection of common distracted driving behaviors (such as calling, smoking, and drinking water), a distracted driving behavior detection algorithm based on improved YOLOv4-tiny has been proposed. First of all, to solve the problem of poor detection accuracy of small targets such as cigarettes, the output features of the last Cross Stage Partial connections (CSP) layer in the backbone feature extraction network are convolved and upsampled. After which the feature fusion is performed with the output features of the second CSP layer, and a prediction scale of 52×52 is added to improve the detection ability of small targets. Secondly, the Efficient Channel Attention (ECA) attention mechanism module is added to the feature pyramid to improve the detection accuracy of the model. Finally, the k-means clustering algorithm is used to redetermine bounding box priors on the self-made distracted driving behavior data set, and transfer learning, cosine annealing learning rate, and label smoothing are used for model training. The results show that the mean average precision (*mAP*) of the method proposed in the present study is 98.88%, which is 2.83% higher than the original YOLOv4-tiny algorithm, and the average precision for cigarette detection is increased by 8.46%, which improves the detection ability of cigarettes. Therefore, the distracted driving behavior detection algorithm based on improved YOLOv4-tiny show good comprehensive performance, which is conducive to the on-board system to detect and remind the driver's distracted driving behavior in real-time and has certain practical significance for reducing the occurrence of traffic safety accidents.

Key words: distracted driving; YOLOv4-tiny; prediction scale; ECA attention mechanism