

DOI:10.19651/j.cnki.emt.2209761

# 基于改进 SSD 的骑行人员佩戴头盔检测<sup>\*</sup>

王 新 冯艺楠

(河南理工大学物理与电子信息学院 焦作 454000)

**摘要:** 为了解决骑行人员佩戴头盔检测任务中目标小、密集、准确率差、检测速度慢、应用困难等问题,本文基于 SSD 网络提出了 EfficientNetV2-SSD 算法。针对原 SSD 网络参数多的问题,使用改进后的轻量级网络 EfficientNetV2 替换 SSD 中的特征提取网络,减少网络参数,提升网络检测速度;针对难检测的小目标,使用自上而下与自下而上的 FPN 金字塔结构,最大程度丰富所有预测特征层信息,提升小目标的检测准确率;针对头盔等被检测的目标特征,重新设计先验框尺寸与比例,提高了小目标检测的准确率,同时加快网络收敛速度,减小网络体积。实验结果表明, EfficientNetV2-SSD 网络对佩戴头盔的检测平均精度均值相比 SSD 网络提高 7.01%,网络体积减少 75%,具有更好的实用性。

**关键词:** 目标检测;头盔检测;小目标;SSD;EfficientNetV2-SSD

**中图分类号:** TP391.4 **文献标识码:** A **国家标准学科分类代码:** 520.60

## Helmet wearing detection for cyclists based on improved SSD

Wang Xin Feng Yinan

(School of Physics &amp; Electronic Information Engineering, Henan Polytechnic University, Jiaozuo 454000, China)

**Abstract:** In order to solve the problems of small target, intensive, poor accuracy, slow detection speed and difficult application in helmet wearing detection task, this paper proposed a EfficientNetV2-SSD algorithm based on SSD network. Aiming at the problem of multiple SSD network parameters, the improved lightweight network EfficientnetV2 is used to replace the feature extraction network in the SSD to reduce network parameters and improve detection speed. For small targets that are difficult to detect, top-down and bottom-up FPN pyramid structures are used to enrich the information of all prediction feature layers to the maximum extent and improve the detection accuracy of small targets. Aiming at the characteristics of the detected target such as helmet, the size and proportion of the prior frame are redesigned to improve the accuracy of small target detection, accelerate the convergence speed of the network and reduce the network volume. The experimental results show that EfficientNetV2-SSD improved the average accuracy of helmet wearing detection by 7.01% and the network volume was reduced by 75% compared with the SSD network, with better practicability.

**Keywords:** target detection; helmet detection; small target; SSD; EfficientNetV2-SSD

## 0 引 言

近年来,电动车由于经济环保,出行方便,正逐步地取代自行车成为人们出行的一种交通工具<sup>[1]</sup>。但是,由于电动车的车速较快,而且驾驶员骑行时缺乏保护,电动车造成的人员伤亡也逐年上升。公安部交管局已在全国开展“一盔一带”安全守护活动,各省也逐渐将骑行电动车佩戴头盔写入交通管理条例。检测电动车骑行人员佩戴头盔对于公共交通与个人安全都有很重要的意义。

传统的头盔目标检测方法主要为基于手工设计特征的视觉检测方法,如 HOG、自适应高斯混合模型等算法,首先从背景中分割出头盔、人等目标,再进行分类。传统方法主要根据目标手工设计特征识别到目标,不仅影响检测精度,而且降低了运算速度<sup>[2]</sup>。深度学习目标检测算法通过模拟实现大脑神经元工作原理,对数据特征进行分析学习,具有强大的目标特征提取与检测能力,是当前头盔目标检测的主流算法<sup>[3]</sup>。

基于深度学习的目标检测网络主要分为二阶段(two-

收稿日期:2022-04-27

<sup>\*</sup> 基金项目:国家重点研发计划(2016YFC0600906)、国家自然科学基金(61403129)项目资助

stage)网络与一阶段(one-stage)网络。二阶段网络如 Fast R-CNN<sup>[4]</sup>,Faster R-CNN<sup>[5]</sup>等,其原理为先在一张图像上生成若干个候选区域(region proposal),再对这些候选区域进行特征提取与分类识别,具有准确率高特点,但实时性差、体积大,难以应用到实际场景中。一阶段网络的出现,弥补了二阶段网络的缺点,如 YOLO(you only look once)<sup>[6]</sup>、SSD(single shot MultiBox detector)<sup>[7]</sup>。这类网络相比于二阶段网络,省去了在图像上生成候选区域这一步骤,将目标区域定位看作一个回归问题,直接在图像上划分候选区域,图像从输入到输出预测目标的位置、类别一步完成,实现网络端到端检测。相比与二阶段网络,一阶段网络精度虽略低于二阶段网络,但减小了网络体积,提升了检测速度。在实际应用场景中达到良好的效果,得到广泛应用。一阶段网络正成为深度学习中目标检测的主流方法。

SSD 网络在一阶段网络中率先利用不同尺度特征图进行预测,利用了上下文的语义信息,对于检测不同尺度的目标有良好的效果,但网络体积很大,难以应用在移动端。因此,出现了许多基于 SSD 改进的目标检测网络。刘寒迪等<sup>[8]</sup>使用了轻量级特征提取网络 MobileNet<sup>[9]</sup>替换了 SSD 网络中特征提取网络,改进后的网络参数量大大减少,速度

得到提升,适用于实时性要求高的场合。刘琛等<sup>[10]</sup>在 SSD 网络中引入视觉机制完成对摩托车头盔的检测。孙备等<sup>[11]</sup>通过对 SSD 进行改进,在小目标检测上取得不错的效果。但以上方法并未提到在复杂密集场景下的目标检测效果。

本文中道路环境下的头盔检测存在密集、小目标等问题,同时检测场景对网络检测的实时性、网络的体积有一定的要求。SSD 网络体积大,对密集、小目标的检测效果差,无法满足要求。因此,本文对 SSD 算法的特征提取网络、特征融合网络、先验框机制进行了设计与改进,提出了 EfficientV2-SSD 网络,满足了道路环境下骑行人员佩戴头盔检测的任务要求。

## 1 SSD 算法模型及其原理

SSD 目标检测算法选择 VGG16<sup>[12]</sup>网络作为特征提取网络,将 VGG16 中两个卷积层的输出作为预测特征图。同时,在 VGG16 后增加 4 个输出尺度不同的卷积层,4 个卷积层的输出也作为预测特征图。对 6 个尺度不同的特征图中包含的目标位置、类别信息回归预测,完成对特征信息的提取与检测。其整体网络结构如图 1 所示。

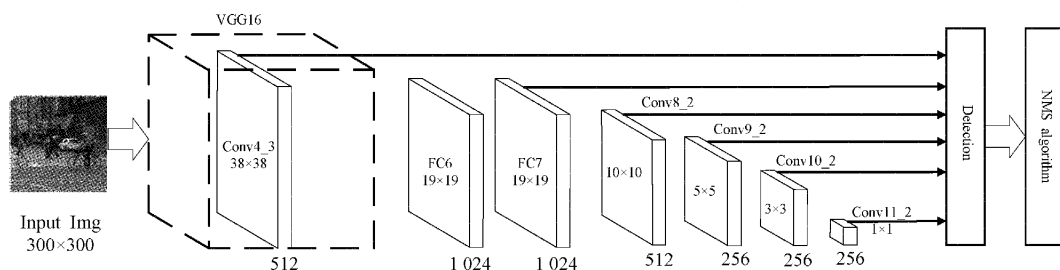


图 1 SSD 网络结构图

SSD 算法采用的先验框机制与 Faster R-CNN 类似,在输出的特征图上生成不同大小的先验框,先验框与图中的真实目标进行匹配,得到预测框及置信度,之后经过非极大值抑制层(non-maximum suppression,NMS)剔除掉置信度低的先验框,最终输出与真实目标匹配度最高先验框作为最终预测结果。

SSD 也存在许多问题。首先 SSD 网络的特征提取网络 VGG16 体积过大,降低了 SSD 网络的检测速度。其次 SSD 网络的多尺度检测机制只是将不同尺度的特征图一起进行检测,不同尺度特征图之间的信息没有得到充分利用,造成小目标的检测精度不高。因此不能直接用于检测道路环境下骑行人员佩戴头盔的任务。本文将会针对以上问题,在 SSD 算法的基础上进行改进和优化。

## 2 改进的 SSD 算法

### 2.1 特征提取网络的改进

针对 SSD 特征提取网络 VGG16 存在体积大、计算量大、梯度消失、特征提取能力低等问题,本文使用了改进的

EfficientNetV2<sup>[13]</sup>作为特征提取网络。EfficientNetV2 具有更小的网络体积、更快的检测速度、高精度的优点。EfficientNetV2 的网络结构如表 1 所示,选取 EfficientNetV2 网络的浅层卷积层 Stage3 与 Stage4 的输出作为预测特征层,提高对小目标的检测效果,获得更多的目标位置信息。网络的 Stage5~Stage7 不再使用,减小网络参数;其次替换 EfficientNetV2 网络中的注意力模块,提高网络准确率与降低网络参数。

表 1 EfficientNetV2 网络结构

阶段	操作	步距	通道数	层数
0	Conv3 * 3	2	24	1
1	Fused-MBConv, k3 * 3	1	24	2
2	Fused-MBConv, k3 * 3	2	48	4
3	Fused-MBConv, k3 * 3	2	64	4
4	MBConv, k3 * 3, SE0.25	2	128	6
5	MBConv, k3 * 3, SE0.25	1	160	9
6	MBConv, k3 * 3, SE0.25	2	272	15
7	Conv1 * 1 & Pooling & FC	—	1 280	1

EfficientNetV2 网络主要由 MBConv 卷积块与 Fused-MBConv 块堆叠而成。MBConv 块使用了 Depthwise 卷积,Depthwise 在卷积过程中一个卷积核只对一个通道进行卷积操作,之后再使用  $1 \times 1$  的卷积来改变通道数,大大减少了网络的参数。Fused-MBConv 块将 MBConv 块中的 Depthwise 卷积替换成普通  $3 \times 3$  卷积,能更好的利用移动端或服务端加速器,加快网络的运算速度。MBConv 块与 Fused-MBConv 块结构中的残差结构也增加了梯度路径,防止梯度爆炸。结构分别为图 2(a)、(b)所示。

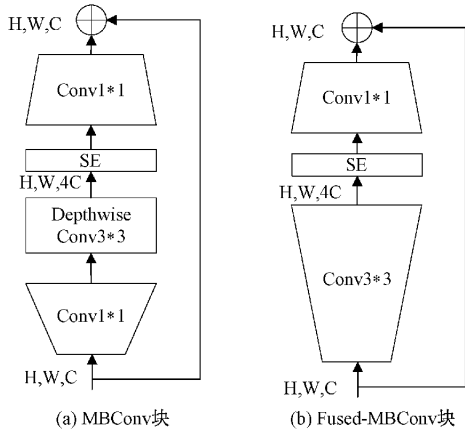


图 2 MBConv 块与 Fused-MBConv 块结构

由图 2 看出,两个卷积块都使用了注意力机制 SE (squeeze-and-excitation)<sup>[14]</sup>。注意力的作用是使网络关注到通道中所含的目标位置信息与语义信息。SE 注意力机制获取特征层中所有通道的权重,如图 3(a)所示。但 SE 机制对通道注意力机制的预测带来了副作用,捕获所有通道的依赖关系是低效并且是不必要的。本文引入 ECA<sup>[15]</sup> (efficient channel attention)模块,在全局平均池化后使用了 1D 卷积来代替全连接层,只捕获一定数量的通道信息,如图 3(b)所示。相比于 SE,ECA 模块在提高了网络性能的同时,也减少了运算量。

2.2 特征融合网络的改进

SSD 网络的多尺度检测将不同尺度的特征图在经过卷积后直接输出进行检测,这样的多尺度检测机制没有将不同尺度特征图的语义信息与位置信息进行融合,特征图中缺乏用于预测的关键信息,造成了网络在预测多尺度目标时鲁棒性不强。

PANet<sup>[16]</sup>很好的解决了 SSD 多尺度检测存在的问题,并极大程度地利用了不同尺度特征图的信息。PANet 首先对特征提取网络输出的深层特征图进行二倍上采样,使其与上一级的浅层特征图尺寸相同,之后进行融合。在完成一次自上而下的特征图融合操作后,浅层特征图得到了深层特征图的语义信息,但是深层特征图依旧缺乏位置信息,因此 PANet 在自上而下的融合操作后又进行了一次自下而上的操作,将包含位置信息丰富的浅层特征图与包含位置信息少的深层特征图融合,丰富深层特征图的位置信

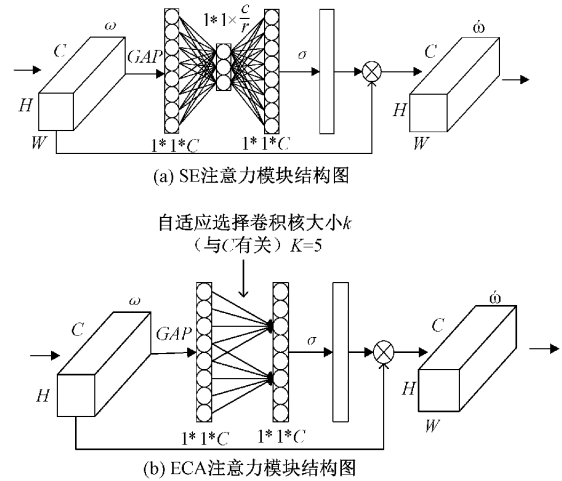


图 3 SE 模块与 ECA 模块结构图

息。改进后的 FPN 金字塔网络使输出的各个尺度特征图的位置信息与语义信息都得到了最大程度的丰富,在 FPN 网络的基础上,进一步提高了网络的鲁棒性。改进的特征融合网络如图 4 所示。

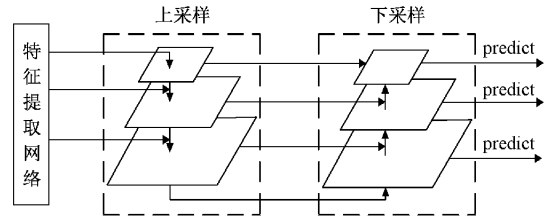


图 4 改进的特征融合网络

本文主要检测的是头盔等小目标且较密集,因此网络进行预测时需要更多的目标位置信息。选取网络输出尺度大的前 3 个特征层构建特征融合网络,进一步增强了浅层特征图的位置信息,同时丰富了语义信息,增加网络的特征表达能力,提升了对头盔小目标及密集目标的检测能力。

2.3 先验框的选择

SSD 网络的先验框机制是预先设计好先验框的尺度 (scale)与宽高比 (ratio)。网络根据预设的先验框与宽高比在特征图上生成用于预测的先验框。

SSD 网络预先设计的先验框尺寸为:[21, 45, 99, 153, 207, 261, 315],宽高比为: $\alpha_r \in \{1, 1/2, 1/3, 2, 3\}$ 。对于 SSD 的 6 个输出特征图,每个特征图上的先验框比例分别为:

$$[1, 2], [1, 2, 3], [1, 2, 3], [1, 2, 3], [1, 2], [1, 2]$$

SSD 网络每个特征图的每个区域生成的先验框个数分别为:[4, 6, 6, 6, 4, 4]。输出 6 个预测特征图尺寸为:[38, 19, 10, 5, 3, 1],SSD 网络最终用来预测的先验框个数为由式(1)得共 8 732 个。

$$A = \sum_{i=1}^6 s_i^2 \times r_i \tag{1}$$

式中:  $A$  为先验框总数,  $S_i$  为第  $i$  层预测特征图尺寸,  $r_i$  为第  $i$  层分配先验框个数,  $i \in [1, 6]$ , 表示第  $i$  层特征图。

传统 SSD 先验框机制不符合本文的数据集检测。首先先验框的比例不适用匹配本文的小目标, 其次 SSD 网络生成的先验框数量太多, 增加网络参数量, 降低检测速度。因此, 需要对传统 SSD 进行改进。

本文使用了 K-means 算法对数据集中的目标进行聚类。本文对输出的 6 个特征层进行预测, 因此选取聚类中心 K 个数为 6。K-means 聚类结果如图 5 所示。从图片左下角至右上角依次为聚类 0~聚类 5, 聚类后生成的先验框尺寸从大到小依次为:

[16, 24], [38, 55], [56, 115], [83, 178], [94, 247], [183, 249]

从聚类生成的先验框尺寸来看, 与 SSD 网络预设先验框尺寸基本相同, 因此本文的预设先验框尺寸与 SSD 网络相同。

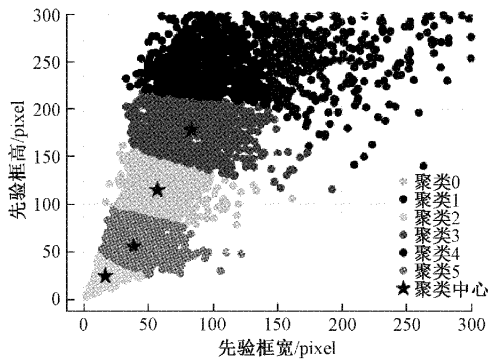


图 5 K-means 聚类结果图

从聚类生成的先验框宽高比来看, 除了第一组先验框的宽高比为 1, 剩余的宽高比都接近 1:2。基于聚类结果, 重新设计的先验框宽高比剔除原 SSD 网络 3、1/3 比例。对第一个特征图只保留  $\alpha_y = 1$  的先验框, 对每个预测特征层分配的先验框比例为:

[1], [1, 2], [1, 2], [1, 2], [1, 2], [1, 2]

本文预测特征图尺寸与原 SSD 算法相同。每个区域生成的先验框个数为 [2, 4, 4, 4, 4, 4], 最终用来预测的先验框个数由式(1)得共 4 872 个, 相比原 SSD 网络减少了近一半。

本文重新设计的先验框对不同目标分配了合适的宽高比, 网络训练过程中更容易匹配到目标, 加快网络收敛速度, 提高了网络的准确率。同时数量相比于原 SSD 网络减少了近一半, 大大减少了计算量。

### 2.4 改进的 SSD 算法模型

本文对 SSD 的特征提取网络、多尺度检测网络、先验框机制进行了改进。首先使用 EfficientNetV2-ECA 作为特征提取网络, 减小网络体积并提高检测速度; 之后将 SSD 网络中的卷积下采样替换为倒残差卷积 (inverted residual)<sup>[17]</sup>, 倒残差结构相比于 MBConv 卷积块少了在 Depthwise 卷积后的 SE 注意力模块, 在损失极小准确率的同时起到减小网络参数的作用, 结构如图 6 所示; 使用自上而下与自下而上的特征金字塔结构, 增强网络特征表达能力, 提高对小目标, 密集目标的检测能力; 最后使用重新设计的先验框匹配目标并进行预测, 提升了网络的检测精度和收敛速度。改进后的 SSD 算法模型如图 7 所示。

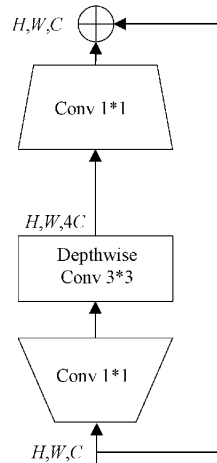


图 6 倒残差结构网络图

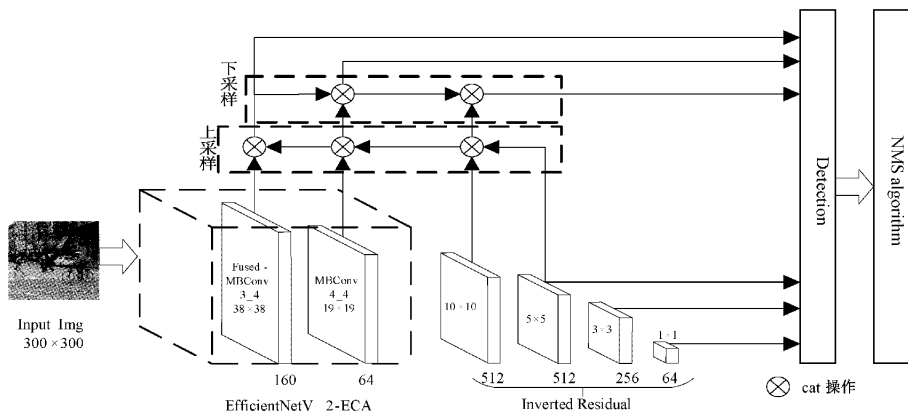


图 7 EfficientNetV2-SSD 网络结构图

### 3 实验结果分析

#### 3.1 实验环境

本文使用的实验平台为 Windows10 操作系统,硬件环境为 Intel XeonW-2265 CPU,NVIDIA Quadro RTX4000。编程语言为 python3.7,采用 pytorch 深度学习框架,版本为 1.7.1。本文实验均在此环境下完成。

#### 3.2 数据集制作与处理

本文使用了网络爬取与实地拍摄相结合的方法制作了数据集。剔除掉相似图片与不合格图片后,整个数据集共有 2 336 图片。使用 LabelImg 标记软件对数据集图片中的目标进行标注,标注类别有 3 个: rider, helmet, non-helmet, 分别表示骑行电动车的人,头部佩戴头盔与未佩戴头盔。其中 rider 类目标有 3 190 个, helmet 类目标有 1 897 个, non-helmet 类目标有 2 067 个。标注的格式和文件分别为 VOC 格式和 xml 文件。采用随机缩放、随机反转、随机色域扭曲等方式实现在线数据增强,扩大数据集样本数量,提升算法的准确度和泛化能力。图 8 为数据集图片示例。



图 8 数据集图片示例

#### 3.3 模型训练

数据集按照 9 : 1 划分为训练集与测试集后进行训

练。训练过程中学习率采用动态调整策略,每训练 3 个 epoch,学习率减少原来的 0.94。训练分为两个阶段,第一个阶段为 150 个 epoch,初始学习率为 0.000 5。第二个阶段也为 150 个 epoch,初始学习率为 0.000 1,训练策略与第一阶段相同。模型训练的 loss 值收敛曲线如图 9 所示。从图中可以看出,两种 loss 曲线下降趋势基本一致,在迭代了 250 个 epoch 后,loss 值趋于平稳,在迭代了 280 个 epoch 后,loss 值几乎不再变化,模型达到收敛。

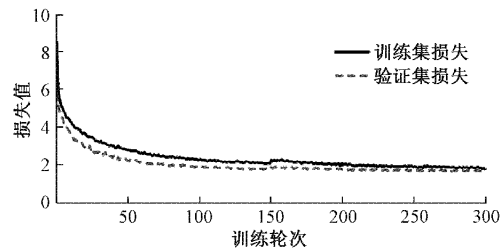


图 9 本文模型 loss 值收敛曲线

#### 3.4 评价指标与实验结果分析

本文选择评估网络性能的指标为:平均精度 AP (average precision),平均精度均值 MAP (mean average precision),每秒检测图片帧数 FPS (frames per second) 以及模型大小 Model size (单位为 MB)。

在数据集中随机划分的 234 张测试集上对改进的网络进行测试,得出 EfficientNetV2-SSD 网络 MAP 达到了 87.45%,helmet 与 rider 两种目标类别的 AP 分别达到了 94%与 91%,non-helmet 类目标 AP 略低,为 77%。

为验证本文算法的有效性,选取二阶段网络 Faster R-CNN,一阶段网络 Yolov4、SSD、MobileNetv2-SSD 与本文算法进行对比,对比结果如表 2 所示。

表 2 不同算法性能对比

算法	AP			MAP%	FPS	模型大小/ MB
	rider	helmet	non-helmet			
Faster R-CNN	88.86	73.56	76.57	79.70	16	301.05
Yolov4	90.34	93.55	83.54	89.14	35	243.96
SSD	83.53	86.69	71.00	80.41	45	99.76
MobileNetv2-SSD	74.36	47.3	53.36	52.53	72	23.16
EfficientNetV2-SSD	90.91	94.07	77.36	87.45	50	24.48

由表 2 可以看出,Yolov4 的准确率最高,但是 FPS 较低,网络体积大,难以应用移动端;SSD 网络的 FPS 与相比于 Yolov4 有提升,但是准确率只有 81.41%,而且网络体积大,移植到移动端困难;MobileNetv2-SSD 网络的检测速度最快,体积最小,但准确率过低,仅有 52.53%;二阶段网络 Faster R-CNN 体积最大,FPS 最低,且准确率表现一般,只有 79.70%,无法满足本文检测场景要求。本文提出的 EfficientNetV2-SSD 网络,有 87.45%的准确率,仅比最高的

Yolov4 低 1.59%。而且 24.48 M 的网络体积与 50 的 FPS 都优于 Yolov4 与 SSD 网络,虽检测速度不如 MobileNetv2-SSD,但满足实时检测的要求,同时准确率远远高于 MobileNetv2-SSD 的 52.23%。综上所述,本文提出的 EfficientNetV2-SSD 网络综合性能更强,能够满足实际场景的实时性,可移植移动端的要求,保证了检测结果的准确。

#### 3.5 消融实验

为了验证本文对 SSD 网络改进的有效性,设计了消融实

验,将每次改进后的网络效果与改进前对比,观察每一步改进后的网络性能。共分为 4 步,具体对比情况如表 3 所示。

表 3 不同模块的消融结果比较

对比组	EfficientNetV2	ECA	改进特征融合	优化 Anchor	MAP%	FPS	模型大小/MB
1	×	×	×	×	81.41	45	99.76
2	√	×	×	×	85.39	55	27.31
3	√	√	×	×	85.56	52	24.08
4	√	√	√	×	86.08	51	26.11
5	√	√	√	√	87.45	50	24.48

由表 3 可以看出,组 1 中的未改进的原 SSD 网络在头盔检测中的准确度、速度都有提升的空间,网络体积也过大。远远不能满足应用移动端要求。在组 2 中使用了 EfficientV2 作为主干网络后,网络检测速度达到最快的 55 FPS,但是网络体积与准确率还不理想,在组 3 中使用 ECA 模块替换了 EfficientV2 中的 SE 模块后,网络检测速度虽然略微减慢,但准确率与网络体积都得到优化。经过组 4、组 5 改进后,网络准确率最终提高到 87.45%,网络体积最终也减小到 24.48 M,尽管检测速度略微减小,但 FPS

仍保持在 50 帧/s,满足实时检测要求。从消融实验来看,本文算法的改进满足了头盔检测场景的实时性要求,小体积网络确保了能够移植到移动设备。同时保证了检测的准确率。证明了本文改进的 EfficientNetV2-SSD 网络的合理性。

本文与原 SSD 算法进行了检测效果图对比。4 张图在数据集中随机选取,对比结果如图 10 所示。图中共有 3 列图片,其中最左边为原图,第二列为 SSD 算法检测效果图,最后一列为本文提出的 EfficientNetV2-SSD 算法检测效果图。

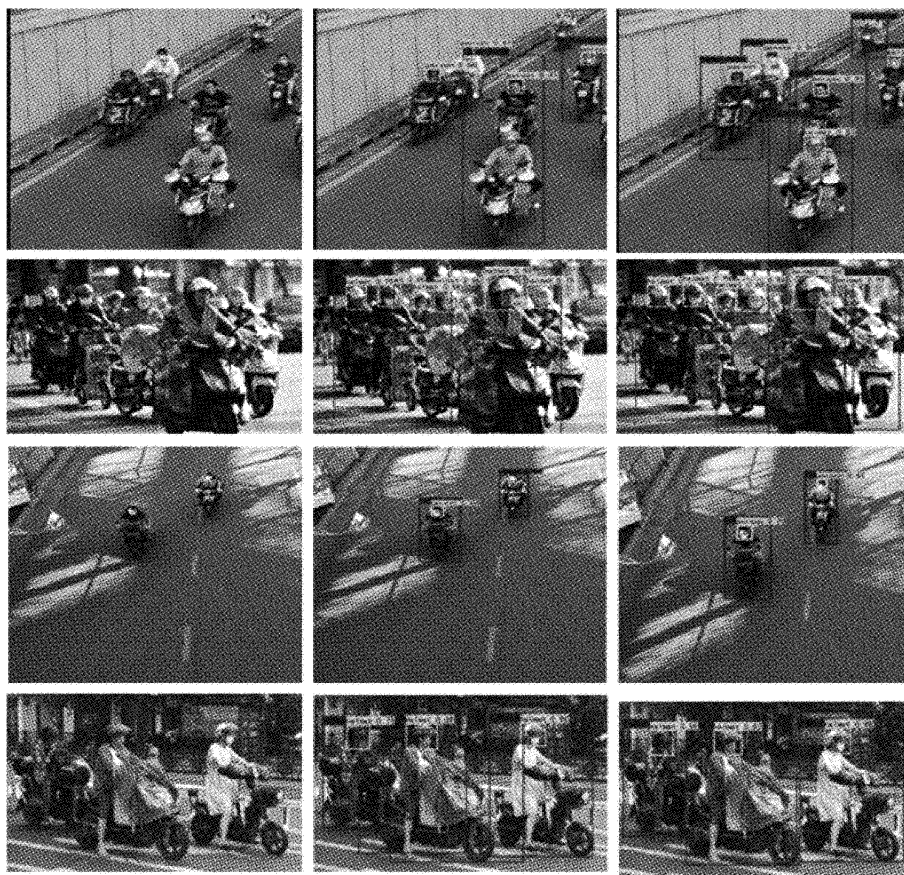


图 10 不同算法下检测效果对比

图中共有 3 类目标框,分别为 rider 类、helmet 类、non-helmet 类。前两张图片均为密集场景,可以看出,SSD 算法产生了漏检的情况。第一张图目标较小且密集,SSD 算

法不仅漏检了 helmet 这种小目标,rider 这种大目标也出现了漏检,同时错误地将图中间位置的两个 rider 目标预测为一个目标。第二张图片场景较第一张图片更为密集,

但目标较大,只漏检了场景中被遮挡的目标,如图片中央的 rider 目标。在前两张图片中,EfficientNetV2-SSD 算法与 SSD 算法相比,在密集场景中几乎不存在漏检情况,能够正确地预测到目标的位置,不存在误检的情况。相比于 SSD 算法置信度也有不同程度的提高;第三张图小目标检测场景中,SSD 算法置信度低,同时误检了一个 non-helmet 类目标。EfficientNetV2-SSD 算法不仅提高了对小目标的置信度,同时不存在漏检误检的情况;第四张图简单检测场景,SSD 算法能够正确地预测到目标的位置,不存在漏检的情况,但是预测框不能准确地定位到目标位置。从图中可以看出,SSD 算法预测框比目标大,同时预测了一个无目标的预测框,出现误检。EfficientNetV2-SSD 算法能够准确地定位到目标,预测框不存在冗余的情况,无误检目标。实验结果显示,本文提出的 EfficientNetV2-SSD 算法更适合复杂道路场景下的头盔检测任务。

### 3.6 移动设备端部署

在部署移动端方面,本文拟采用 NVIDIA 推出的高性能嵌入式 AI 平台 Jetson nano 搭载 EfficientNetV2-SSD 算法,将 EfficientNetV2-SSD 算法与训练好的权值文件下载到 Jetson nano 上。通过 USB 连接摄像头进行图像采集,最后通过远程桌面连接 Jetson nano,查看预测结果。系统整体框图大致如图 11 所示。

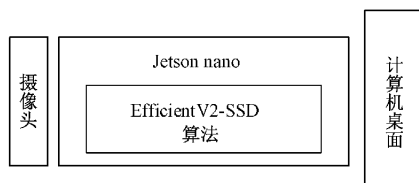


图 11 头盔检测系统框图

在实际场景中,通过摄像头采集道路电动车骑乘人员视频,将采集到视频送入 Jetson nano 嵌入式平台,EfficientNetV2-SSD 算法首先获取视频的每一帧,将每一帧图像转化成网络能够识别的格式,之后进行预测分类。计算机通过远程桌面连接 Jetson nano 嵌入式平台,能够实时查看预测结果。

## 4 结 论

为了应对现实场景中对头盔检测实时性,能够应用在移动设备的要求;同时面对头盔检测存在小目标及密集场景等问题,本文以 SSD 网络为基础提出了一种 EfficientNetV2-SSD 检测网络。首先,替换掉参数量大,检测准确率低的 VGG16 网络,改用轻量级网络 EfficientNetV2,用来减小网络体积。同时提高网络准确率。之后,针对 SSD 网络多尺度融合机制缺少语义信息,提出了一种改进的 FPN 网络,使网络的浅层位置信息与深层的语义信息都能得到丰富,提升了网络的鲁棒性。最后,针对本文数据集的特点,重新设计了 SSD 网络的先验

框机制,提高网络对头盔等小目标的检测能力。实验表明,本文算法相比于原网络,准确率,FPS 都得到提升,网络体积也大大减小。

### 参考文献

- [1] 市场规模近千亿电动两轮车行业大起底[N]. 国际金融报,2021-07-12(12).
- [2] 周敏新,张方舟,龚声蓉.基于新型特征融合的安全帽佩戴检测方法[J]. 计算机工程与设计,2021,42(11):3181-3187.
- [3] 周晓彦,王珂,李凌燕.基于深度学习的目标检测算法综述[J]. 电子测量技术,2017,40(11):89-93.
- [4] GIRSHICK R. Fast R-CNN [C]. 2015 IEEE International Conference on Computer Vision, 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[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [6] 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, 2016: 779-788.
- [7] LIU W, ANGELOV D, ERHAN D, et al. SSD: Single shot multibox detector [C]. European Conference on Computer Vision, 2016: 21-37.
- [8] 刘寒迪,赵德群,陈星辉,等.基于改进 SSD 的航拍施工车辆检测识别系统设计[J]. 国外电子测量技术,2020,39(7):127-132.
- [9] HOWARD A G, ZHU M, CHEN B, et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications [J]. ArXiv Preprint, 2017, ArXiv:1704.04861.
- [10] 刘琛,王江涛,王明阳.引入视觉机制的 SSD 网络在摩托车头盔佩戴检测中的应用[J]. 电子测量与仪器学报,2021,35(3):144-151.
- [11] 孙备,左震,吴鹏,童小钟,等.面向无人艇环境感知的改进型 SSD 目标检测方法[J]. 仪器仪表学报,2021, 42(9):52-61.
- [12] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J]. ArXiv Preprint,2014, ArXiv:1409.1556.
- [13] TAN M, LE Q. Efficientnetv2: Smaller models and faster training [C]. International Conference on Machine Learning. PMLR, 2021: 10096-10106.
- [14] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]. Proceedings of the IEEE Conference on

- Computer Vision and Pattern Recognition, 2018: 7132-7141.
- [15] WANG Q, WU B, ZHU P, et al. ECA-Net: Efficient channel attention for deep convolutional neural networks[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020:11531-11539.
- [16] LIU S, QI L, QIN H, et al. Path aggregation network for instance segmentation[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 8759-8768.
- [17] SANDLER M, HOWARD A, ZHU M, et al. Mobilenetv2: Inverted residuals and linear bottlenecks[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 4510-4520.

#### 作者简介

王新(通信作者),教授,博士生导师,主要研究方向为信号处理,故障诊断。

E-mail:wangxin@hpu.edu.cn

冯艺楠,研究生,主要研究方向为信号处理,目标检测。

E-mail:1481357554@qq.com