

# 一种轻量化目标检测算法研究

余进<sup>1,2</sup>, 史燕中<sup>3</sup>, 王春华<sup>2</sup>, 赵倩<sup>2</sup>, 吴蔚<sup>2</sup>

(1. 中国航天科工集团第二研究院, 北京 100039;

2. 北京航天长峰科技工业集团有限公司, 北京 100039;

3. 北京航天长峰股份有限公司, 北京 100039)

**摘要:**基于深度卷积神经网络的目标检测算法对硬件的计算性能要求很高,难以部署在一些嵌入式设备和移动终端中,而当前的一些轻量化分类算法没有针对目标检测任务的特点进行网络结构设计。针对这一问题,借鉴深度可分离卷积的思路,通过引入多尺度的特征融合模块,设计了一个针对目标检测任务的轻量化特征提取网络 TinyNet,进而提高了轻量化特征提取网络对不同尺度目标的适应性。结合当前性能较好的 YOLOv3 目标检测框架,用 TinyNet 取代 YOLOv3 的特征提取网络,并利用轻量化模块进一步优化 YOLOv3 的检测子网络,得到一个轻量化的目标检测模型 Tiny-YOLOv3+。实验结果表明,相比于使用其他轻量化特征提取网络设计的 YOLOv3, Tiny-YOLOv3+ 在检测精度有所提升的基础上,大大地降低了原模型的参数量,明显提升了检测速度,有效提高了轻量化检测模型的性能和效率。

**关键词:**深度学习;目标检测;轻量化;深度可分离卷积;YOLOv3

**中图分类号:** TP301.6

**文献标识码:** A

**文章编号:** 1673-629X(2020)11-0042-06

**doi:** 10.3969/j.issn.1673-629X.2020.11.008

## Research of a Lightweight Object Detection Algorithm

YU Jin<sup>1,2</sup>, SHI Yan-zhong<sup>3</sup>, WANG Chun-hua<sup>2</sup>, ZHAO Qian<sup>2</sup>, WU Wei<sup>2</sup>

(1. 2nd Institute of China Aerospace Science & Industry Corp, Beijing 100039, China;

2. Changfeng Science Technology Industry Group Corp, Beijing 100039, China;

3. Beijing Aerospace Changfeng Co., Ltd., Beijing 100039, China)

**Abstract:** Detectors based on deep convolutional neural networks which demand hardware of high performance computing capability are difficult to operate on embedded devices and mobile terminals. In addition, recent networks structure design of lightweight models for classification have no consideration of the feature of object detectors. To address this issue, a lightweight feature extraction network TinyNet for target detection task is designed by using the idea of deep separable convolution as reference and introducing the multi-scale feature fusion module, so as to improve the adaptability of the lightweight feature extraction network to targets of different scales. Combining with the current YOLOv3 target detection framework with better performance, the feature extraction network of YOLOv3 is replaced by TinyNet and the detection sub-network of YOLOv3 is further optimized by the lightweight module to obtain a lightweight target detection model Tiny-YOLOv3+. The experiment shows that compared with YOLOv3, which uses other lightweight features to extract network design, Tiny-YOLOv3+, on the basis of improved detection accuracy, greatly reduces the number of parameters of the original model, significantly improves the detection speed and effectively improves the performance and efficiency of the lightweight detection model.

**Key words:** deep learning; object detection; lightweight; depthwise separable convolutions; YOLOv3

## 0 引言

计算机视觉技术广泛应用于视频侦查、自动驾驶、人脸识别等诸多领域。目标检测作为计算机视觉领域的基础研究方向之一,其任务是识别出图像中的一个或多个目标,给出这些目标的位置信息和类别信息,从

而能够进一步实现后续任务和应用<sup>[1-2]</sup>。

目标检测的研究最早开始于二十世纪六十年代,传统的目标检测算法是基于人工设计特征结合浅层分类器的框架,这类目标检测算法虽然取得了不错的效果,但冗余度高,运行速度慢,鲁棒性不强,不能很好地

收稿日期:2019-12-28

修回日期:2020-04-30

基金项目:国家重点研发计划项目(2018YFC0832000)

作者简介:余进(1993-),男,硕士研究生,研究方向为深度学习、目标检测;史燕中,研究员,硕士研究生导师,研究方向为视频图像处理。

表达多类目标的特征<sup>[3-4]</sup>。随着近年来深度学习技术的发展,基于深度卷积神经网络的目标检测算法的识别准确率取得了巨大的提升。基于卷积神经网络的目标检测算法主要分为两大类,一类是基于候选区域的目标检测框架,如以 Faster R-CNN<sup>[5]</sup>、Mask R-CNN<sup>[6]</sup> 为代表的一系列目标检测框架,这类检测算法准确度较高,但检测速度很难满足实时性要求;另一类是基于直接回归的目标检测框架,如以 YOLO<sup>[7]</sup>、SSD<sup>[8]</sup>、YOLOv3<sup>[9]</sup> 等为代表的一系列模型,由于省去了提取候选区域的过程,检测速度大大提升。

然而,基于深度卷积神经网络的目标检测算法对硬件的计算性能要求很高。在实际应用中,很多嵌入式设备和移动终端并不具备目前深度卷积神经网络所需的计算性能。当前大多数的目标检测算法难以在这类设备上得到有效部署,因而对于模型轻量化的需求显得尤为迫切。在这种情况下,近年来有一系列针对图像分类任务而设计的轻量化卷积神经网络模型被提出,包括 Xception<sup>[10]</sup>、MobileNet<sup>[11]</sup> 系列和 Shuffle-Net<sup>[12]</sup> 系列,虽然在目标检测算法中引入这类轻量化模型可以极大地降低模型的参数量和运算量,但由于这类轻量化模型主要是针对图像分类任务而设计的,所以并未针对目标检测框架进行优化设计。

针对上述问题,该文以当前应用广泛的 YOLOv3 目标检测模型作为检测框架,在研究分析当前一些常见的轻量化分类模型的基础上,针对目标检测的特点设计了一个新型的轻量化特征提取网络 TinyNet,用该模型取代 YOLOv3 的特征提取子网络,并利用轻量化模块进一步优化 YOLOv3 的检测子网络。实验结果表明,相比于使用其他轻量化特征提取网络设计的 YOLOv3,新的网络模型在检测精度有所提升的基础上,大大地降低了原模型的参数量,明显提升了检测速度。

## 1 相关知识

### 1.1 相关定义

Ground Truth (GT):数据的真实标签,对于目标检测问题,Ground Truth 包括图像中物体的类别以及该图像中每个物体的真实边界框。

Intersection over Union (IoU):交并比,即两个区域的交集与这两个区域的并集的比值。IoU 的值代表了两个区域之间的重合程度,IoU 越大表示两个区域的重合度越高。

True Positives (TP):即正样本被正确识别为正样本,在目标检测中表示被正确识别出来的目标。

True Negatives (TN):即负样本被正确识别为负样本,在目标检测中表示背景没有被识别成目标。

False Positive (FP):即负样本被错误识别为正样本,在目标检测中表示背景被识别成目标。

False Negatives (FN):即正样本被错误识别为负样本,在目标检测中表示目标被识别成背景。

Precision:查准率,表示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,定义如下:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (1)$$

Recall:查全率,表示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,定义如下:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (2)$$

### 1.2 YOLOv3 目标检测算法

YOLOv3 是基于直接回归的目标检测算法,在检测速度和精度上都取得了不错的效果。YOLOv3 算法将图像划分为  $S \times S$  的网格,利用整张图片的特征,对每个网格分别预测中心点落在该网格内的目标。对每个网格区域,网络需要预测每个预测框的置信度和四个坐标值,以及预测框的类别概率分布。相比于之前的 YOLO 模型,YOLOv3 使用了 Darknet-19 的改进版本 Darknet-53 作为特征提取网络,DarkNet-19 包含 5 个池化层和 19 个卷积层,并通过  $3 \times 3$  卷积之间的  $1 \times 1$  卷积对特征图通道维数进行压缩。DarkNet-53 的改进主要体现在三个方面,用步长为 2 的卷积层替换池化层,使用跳层连接<sup>[13]</sup> 结构,增加了网络层数。整个检测网络输出三个不同尺度的特征图,采用多尺度的方法对不同大小的目标进行检测,并借鉴了 FPN<sup>[14]</sup> 模型中的多层特征融合的思想,将深层特征上采样后回传给浅层特征以丰富浅层特征的语义信息,有效地提升了小目标的检测精度。其网络结构如图 1 所示。

### 1.3 深度可分离卷积

基于深度卷积神经网络的目标检测算法通常对硬件的计算性能要求很高,为了能够让深度卷积神经网络运行在一些嵌入式设备和移动终端中,近年来有一系列针对图像分类任务而设计的轻量化卷积神经网络模型被提出,比如 MobileNet 和 ShuffleNet。这两个轻量化模型均使用到了深度可分离卷积来减少计算量和参数量。

如图 2 所示,在标准卷积中,卷积核通道数与需要进行卷积运算的特征图通道数相同,卷积核需要与对应图像区域中的所有通道进行点乘求和运算。假定输入特征图形状是  $D_f \times D_f \times M$ ,而输出特征图形状是  $D_f \times D_f \times N$ ,其中  $D_f$  是特征图的宽和高,而  $M$  和  $N$  指的是通道数,并假定输入特征图和输出特征图的分辨率是一致的。对于标准卷积核  $D_k \times D_k$ ,其计算量是:

$$D_K \times D_K \times M \times N \times D_F \times D_F。$$

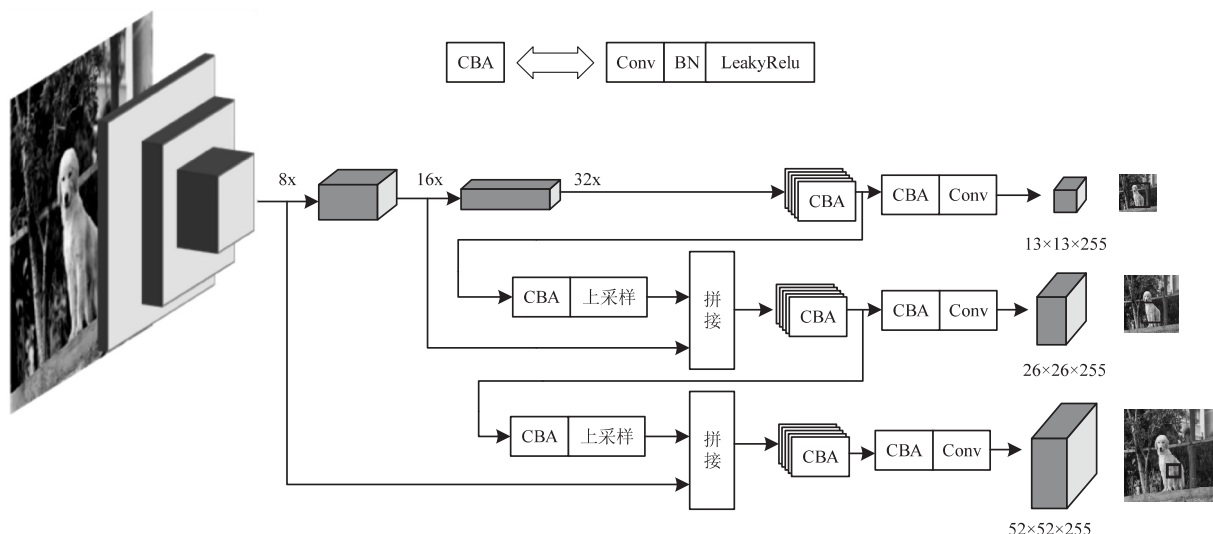


图 1 YOLOv3 网络结构示意图

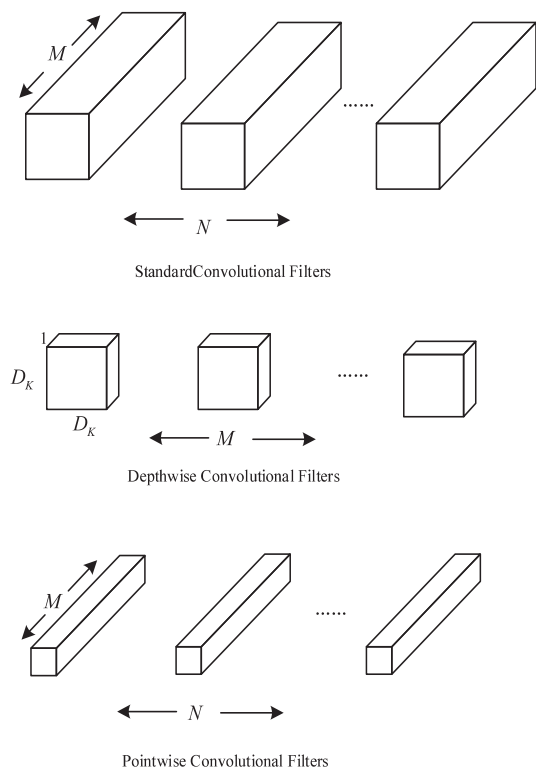


图 2 标准卷积和深度可分离卷积示意图

而深度可分离卷积将标准卷积分解为逐通道卷积 (depthwise convolution) 和点卷积 (pointwise convolution) 两部分,在逐通道卷积中,卷积核只含一个通道,卷积核只与特征图对应图像区域内的单个特征通道进行点乘求和运算,特征图的每个特征通道分别与不同的单层卷积核进行运算,所以卷积核的数量与需要进行卷积运算的特征图通道数相同。

但是在逐通道卷积中,输出特征图的每一个特征通道只与输入特征图的某个特征通道有相关性,没有整合利用输入特征图中的所有特征通道的信息。所以在逐通道卷积之后,采用点卷积进一步对深度可分离

卷积的输出特征图进行运算,使用卷积核大小为  $1 \times 1$  的点卷积可以在尽量降低运算量的情况下将所有的特征通道串联起来,使得每一个输出的特征通道能包含所有输入的特征通道的信息,解决了深度可分离卷积在特征通道方向上信息流通不畅的问题。而对于深度可分离卷积,其计算量由逐通道卷积和点卷积两部分组成:

$$D_K \times D_K \times M \times D_F \times D_F + M \times N \times D_F \times D_F$$

深度可分离卷积与标准卷积的计算量相比,可以得到如下的比值:

$$\frac{D_K \times D_K \times M \times D_F \times D_F + M \times N \times D_F \times D_F}{D_K \times D_K \times M \times N \times D_F \times D_F} = \frac{1}{N} + \frac{1}{D_K^2}$$

通常来说,  $N$  是一个比较大的值,那么在采用  $3 \times 3$  卷积核的情况下,  $D_K = 3$ ,与标准卷积运算量相比,深度可分离卷积降低了 9 倍的运算量,极大地减少了运算量。

#### 1.4 训练方式

当前目标检测模型通常是在预训练模型的基础上微调获得的,具体来说,就是把在 ImageNet 数据集上预训练好的模型作为目标检测模型的特征提取网络,然后保持预训练模型的参数值不变,在目标检测数据集(如 VOC、COCO)上训练模型得到检测子网络的权重。然而,使用预训练模型也给修改模型网络结构造成了不便,因为这意味着当模型的网络结构出现调整后,需要在 ImageNet 数据集进行预训练,由于 ImageNet 数据集的图片数量很大,也就意味着需要花费很长时间来得到预训练模型。ScratchDet<sup>[15]</sup> 中提出的一种从头开始训练的方式,不使用预训练模型直接在目标检测数据集上进行训练。具体而言,在设计

特征提取网络中,保证在每一个卷积层后存在批量归一化层,由于批量归一化层可以使得优化过程更加平滑,使得梯度更加稳定,从而允许更大的搜索空间和更快的收敛速度,同时还降低了网络训练过程中出现收敛性问题的可能。这种方式极大地提高了修改模型网络结构的效率。

## 2 算法设计

### 2.1 轻量化主干网络的设计

该文借鉴当前一些轻量化分类模型中的深度可分离卷积设计了一个针对目标检测任务的轻量化主干网络 TinyNet,该网络采用模块化设计方式,通过将轻量化模块有序地连接最终形成模型的整体结构。设计的轻量化模块如图 3 所示,其中左图表示输入和输出特征图分辨率相同的模块,右图表示降采样模块。两种模块都包含两个分支,一个分支包含卷积核大小为  $3 \times 3$  的深度可分离卷积,另外一个分支包含卷积核大小为  $5 \times 5$  的深度可分离卷积。由于使用到了不同大小的卷积核分支,实现了多尺度的感受野融合,进而提高了轻量化主干网络对不同尺度目标的适应性。在左图所示的模块中,由上一模块输入的特征图在特征通道方向上被平均拆分为两部分,再分别经由这两个不同大小的卷积核分支进行卷积运算之后拼接到一起,然后使用一个  $1 \times 1$  的卷积核对不同通道的特征进行融合,最后与跳层连接传送来的输入特征进行元素级加法运算并输出。而在右图所表示的降采样模块中,每个分支直接对所有上一模块输入的特征通道进行卷积运算。

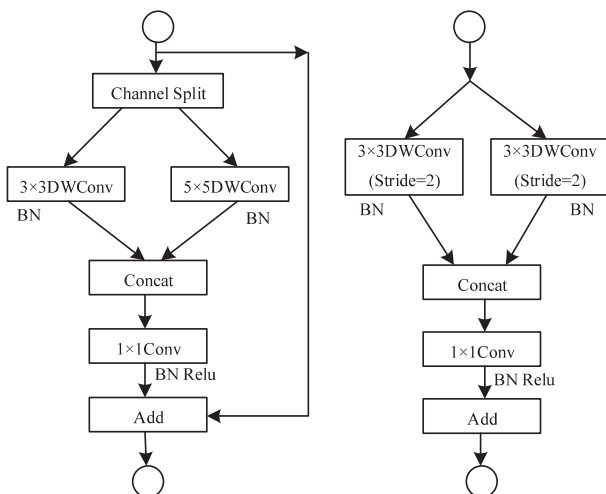


图 3 TinyNet 的基本模块

网络的整体结构如表 1 所示,网络结构设计思想与 VGG 类似,堆叠的使用卷积操作,只不过这里堆叠的是该文设计的轻量化模块。

按照特征图分辨率大小的不同,网络被分为了不同的阶段(Stage)。在第一阶段中,特征图进行了 1 次

降采样,分辨率缩减为原图的  $1/2$ ;在第二阶段中,特征图又进行了 1 次降采样,分辨率缩减为原图的  $1/4$ ,依此类推,整个网络一共被划分为 5 个阶段,最后一个阶段输出的特征图的分辨率已经缩减为原图的  $1/32$ 。在前两个阶段中,仅仅使用到了降采样单元,其中第一阶段使用到的是标准卷积的降采样单元,而第二阶段使用到的是该文设计的轻量化降采样单元。在第三、四、五个阶段,除了使用到了一个轻量化降采样模块,还分别堆叠了 4 个、5 个和 3 个步长为 1 的轻量化模块。

表 1 TinyNet 网络结构

层名称	输出维度	卷积核	步长	重复次数
Image	224×224×3	—	—	—
Stage1	112×112×54	3×3	2	1
Stage2	56×56×108	3×3 and 5×5	2	1
Stage3	28×28×216	3×3 and 5×5	2	1
	28×28×216	3×3 and 5×5	1	4
Stage4	14×14×432	3×3 and 5×5	2	1
	14×14×432	3×3 and 5×5	1	5
Stage5	7×7×864	3×3 and 5×5	2	1
	7×7×864	3×3 and 5×5	1	3
Globalpool	1×1×1 000	—	global	1

### 2.2 轻量化目标检测网络的设计

该文借鉴了 YOLOv3 的检测网络结构,通过将设计的轻量化主干网络 TinyNet 替换 YOLOv3 中原有的 Darknet-53 网络,得到了轻量化模型 Tiny-YOLOv3。保持了 YOLOv3 目标检测框架与 Darknet-53 网络的衔接方法,即分别将 TinyNet 中分辨率为 8 倍、16 倍、32 倍降采样特征层的最后一层作为输出分支,并与回传的深层特征进行融合后再进行输出。为浅层特征提供了语义特征,从而提高了物体识别能力。然后,进一步使用轻量化模块对 YOLOv3 检测子网络中的标准卷积单元进行替换得到一个端到端的轻量化目标检测模型,完成了轻量化模型 Tiny-YOLOv3+ 的设计,替换后的检测子网络结构如图 4 所示。

## 3 实验测试

### 3.1 实验环境与实验参数

该文保证所有实验的运行环境保持一致,硬件环境为 NVIDIA 1080Ti 显卡和 Intel i7-7700H 处理器,16G DDR4 内存;软件环境为 Ubuntu16.04 操作系统;实验框架为 MXNet 开源框架。考虑到实验硬件和软件的限制,并结合从头开始训练的特点,所有实验的网络输入分辨率为  $416 \times 416$ ;批量大小设为 32,初始学习率设为 0.01,学习率调整策略采用 cosine;使用 SGD+

Momentum 的优化方法,并设置动量为 0.9;最大循环迭代次数设为 100 K。

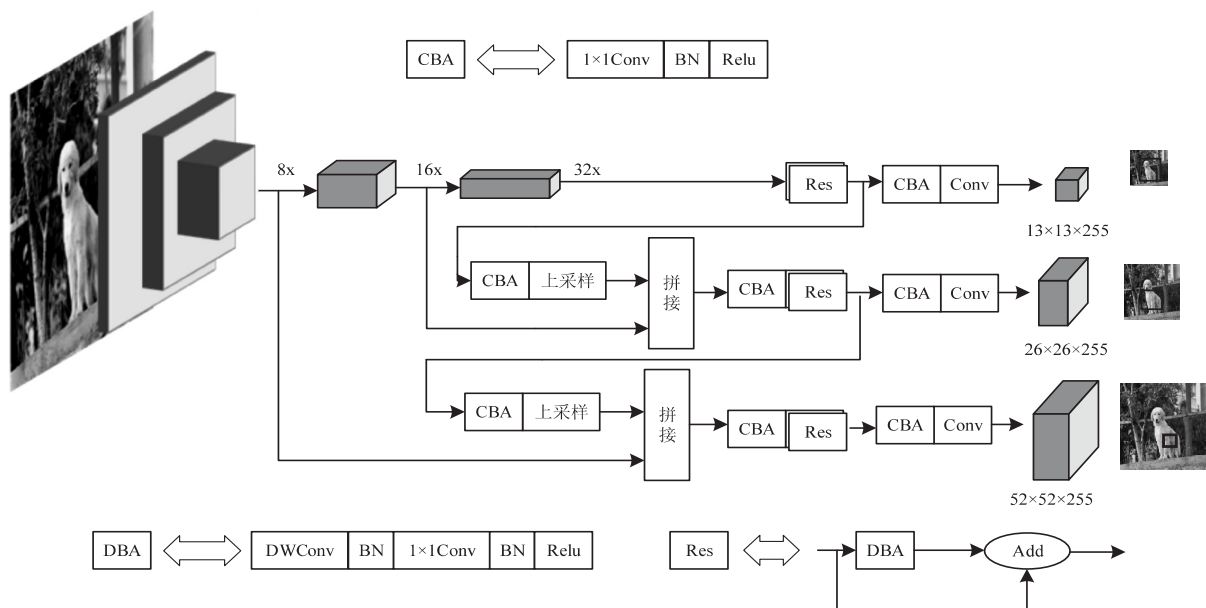


图 4 轻量化模型 Tiny-YOLOv3+结构示意图

### 3.2 实验数据集

该文使用 PASCAL VOC 数据集和 BDD100K 数据集分别进行实验。其中 PASCAL VOC 数据集是一个包含 Person、Bird、Cat、Cow、Dog、Horse、Sheep、Aeroplane、Bicycle、Boat、Bus、Car、Motorbike、Train、Bottle、Chair、Table、Plant、Sofa、Monitor 等 20 个类别物体的目标检测数据集,该文使用 PASCAL VOC2007 和 PASCAL VOC2012 的 trainval 部分的 16 551 张图片作为训练集,使用 PASCAL VOC2007 的 test 部分的 4 952 张图片作为测试集。BDD100K 数据集是伯克利大学发布的当前规模最大、最多样化的公开驾驶数据集。BDD100K 包含 10 万张用于道路目标检测的图片,包含 Bus、Light、Sign、Person、Bike、Truck、Motor、Car、Train、Rider 等 10 个类别。其中训练集包含 7 万张标注图片,验证集包含 1 万张标注图片,测试集包含 2 万张图片。

### 3.3 评价指标

采用 PASCAL VOC2012 的评价标准 mAP (mean Average Precision) 作为模型检测性能的评价指标,计算方法如下:首先针对每个类别,按照目标检测模型预测结果的置信度对样本进行排序,再逐个样本地将置信度作为样本被分为正负样本的阈值,累计计算查准率 (Precision) 和查全率 (Recall),然后绘制 P-R (Precision and Recall) 曲线,将 PR 曲线上的所有召回率相同的点分为一组,共分成  $n$  组,记第  $i$  组的最大精度值为  $P_i$ ,并记该类别在测试集中的所有实例个数为  $N$ ,则 AP 的计算公式为:

$$AP = \sum_{i=1}^n \frac{P_i}{N} \quad (3)$$

最后计算不同类别 AP 值的均值即为 mAP。mAP 的值越高,说明检测模型的检测效果越好。

### 3.4 实验结果

该文从模型精度、参数量和运行速度三个方面对比了直接使用轻量化分类模型 (MobileNet 和 ShuffleNet v2) 作为 YOLOv3 特征提取子网络的目标检测模型,和该文设计的轻量化目标检测模型 Tiny-YOLOv3 以及 Tiny-YOLOv3+,实验结果如表 2 和表 3 所示。

表 2 PASCAL VOC 数据集实验结果对比

模型	数据集	mAP/%	参数量	FPS
YOLOv3-MobileNet	VOC	72.2	24.2 M	136
YOLOv3-ShuffleNet	VOC	69.8	23.3 M	155
Tiny-YOLOv3	VOC	73.9	19.2 M	154
Tiny-YOLOv3+	VOC	73.0	8.8 M	162

表 2 表示的是不同模型在 PASCAL VOC 数据集上的实验结果,通过对比 Tiny-YOLOv3 和 YOLOv3-MobileNet、YOLOv3-ShuffleNet 的结果可以发现,使用该设计的 TinyNet 作为 YOLOv3 的特征提取网络,与使用 MobileNet 相比,mAP 值提高了 1.7 个百分点,参数量降低了 21%,推理速度提高了 13%;与使用 ShuffleNet v2 相比,在推理速度几乎一致的情况下,mAP 值提高了 4.1 个百分点,参数量降低了 17%。这说明该文针对目标检测任务所设计的轻量化特征提取网络 TinyNet 优于现有的轻量化特征提取网络。

值得注意的是,与 ShuffleNet v2 在图像分类任务上的效果要优于 MobileNet 不同 (ImageNet 数据集

Top1 准确率 74.9% VS 70.6%), ShuffleNet v2 在 YOLOv3 目标检测框架下的实验结果不如使用 MobileNet 的实验结果,这也说明将针对分类任务的轻量化模型直接作为目标检测框架的特征提取网络是值得商榷的。

通过对比 Tiny-YOLOv3 和 Tiny-YOLOv3+可以发现,在将 YOLOv3 的检测子网络进行轻量化处理后,在检测精度略微下降的情况下,网络权重减少了 54%,并且推理速度得到明显提升。最后通过对比 Tiny-YOLOv3+与 YOLOv3-MobileNet 的实验结果可知,Tiny-YOLOv3+在精度略微提升的情况下,参数量大大降低,推理速度提升明显,对比 Tiny-YOLOv3+与 YOLOv3-ShuffleNet 的实验结果可知,Tiny-YOLOv3+在推理速度略微提升的情况下,参数量大大降低,明显提升了检测精度。

表 3 BDD100K 数据集实验结果对比

模型	数据集	mAP/%	参数量	FPS
YOLOv3-MobileNet	BDD	43.2	24.1 M	79
YOLOv3-ShuffleNet	BDD	41.8	24.0 M	82
Tiny-YOLOv3	BDD	44.7	19.1 M	81
Tiny-YOLOv3+	BDD	44.1	8.7 M	85

表 3 表示的是不同模型在 BDD100K 数据集上的实验结果,对比 Tiny-YOLOv3+与 YOLOv3-MobileNet、YOLOv3-ShuffleNet 的实验结果可以发现,在推理速度略微提升的情况下,Tiny-YOLOv3+同样得到了比 YOLOv3-MobileNet 和 YOLOv3-ShuffleNet 更优的实验精度和更少的参数量。在两个数据集上不同的实验结果均说明该文针对目标检测任务设计的轻量化特征提取网络 TinyNet 优于现有的轻量化特征提取网络。

#### 4 结束语

针对当前轻量化模型没有针对目标检测任务的特点进行网络结构设计的问题,借鉴深度可分离卷积的思路,通过引入多尺度的特征融合模块,设计了一个针对目标检测任务的轻量化模型 Tiny-YOLOv3+。相比于使用其他轻量化特征提取网络设计的 YOLOv3,设计的轻量化目标检测模型在实验精度略微提升的情况下,大大降低了模型的参数量,明显提升了检测速度。实验的训练和推理过程仍旧是在高性能计算机上进行,下一步将考虑将检测模型移植到嵌入式设备上进一步进行研究。

#### 参考文献:

- [1] 寇大磊,权冀川,张仲伟. 基于深度学习的目标检测框架进展研究[J]. 计算机工程与应用,2019,55(11):25-34.
- [2] 黄莉芝. 基于深度卷积神经网络的目标检测算法研究[D]. 成都:西南交通大学,2018.
- [3] 陈鹏飞. 基于多尺度特征融合的实时目标检测[D]. 西安:西安电子科技大学,2018.
- [4] 李旭冬,叶茂,李涛. 基于卷积神经网络的目标检测研究综述[J]. 计算机应用研究,2017,34(10):2881-2886.
- [5] REN Shaoqing, HE Kaiming, ROSS G, 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] HE Kaiming, GEORGIA G, PIOTR D, et al. Mask R-CNN[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(2): 386-397.
- [7] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]//IEEE conference on computer vision and pattern recognition. Las Vegas, NV; IEEE, 2016: 779-788.
- [8] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[C]//European conference on computer vision. Cham; Springer, 2016: 21-37.
- [9] 张富凯,杨峰,李策. 基于改进 YOLOv3 的快速车辆检测方法[J]. 计算机工程与应用,2019,55(2):12-20.
- [10] CHOLLET F. Xception: deep learning with depthwise separable convolutions[C]//IEEE conference on computer vision and pattern recognition. Honolulu, HI; IEEE, 2017: 1800-1807.
- [11] 孔英会,朱成诚,车麟麟. 复杂背景下基于 MobileNets 的花卉识别与模型剪枝[J]. 科学技术与工程,2018,18(19):84-88.
- [12] ZHANG X, ZHOU X, LIN M, et al. ShuffleNet: an extremely efficient convolutional neural network for mobile devices[C]//IEEE conference on computer vision and pattern recognition. Salt Lake City, UT; IEEE, 2018: 6848-6856.
- [13] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//IEEE conference on computer vision and pattern recognition. Las Vegas, NV; IEEE, 2016: 770-778.
- [14] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]//IEEE conference on computer vision and pattern recognition. Honolulu, HI; IEEE, 2017: 936-944.
- [15] ZHU Rui, ZHANG Shifeng, WANG Xiaobo, et al. Scratch-Det: training single-shot object detectors from scratch[C]//IEEE conference on computer vision and pattern recognition. Long Beach, CA; IEEE, 2019: 2263-2272.