DOI:10. 20079/j. issn. 1001-893x. 220901003

融合视觉机制和多尺度特征的小目标检测算法*

武德彬,刘笑楠,刘振宇,杨　娜

(沈阳工业大学 信息科学与工程学院,沈阳 110870)

摘　要:针对 SSD(Single Shot MultiBox Detector)目标检测算法对小目标检测能力不足的问题,提出一种引入视觉机制和多尺度语义信息融合的 VFF-SSD(Vision Feature Fusion SSD)改进算法。为了增大浅层网络的感受野提高特征提取能力,首先在 SSD 浅层特征层中加入视觉机制,然后利用改进 PANet(Path Aggregation Network)多尺度特征融合网络与深层特征增强网络得到新的特征层,旨在增强浅层网络的语义信息并加强深层特征的特征表达能力,最后应用注意力机制模块提高对重要信息的学习能力。实验结果表明,在 PASCAL VOC2007 测试集检测的 mAP(Mean Average Precision)值达到 81. 1%,对数据集中小目标的 mAP 值较原 SSD 提高了 6. 6%。

关键词:小目标检测;深度学习;视觉机制;多尺度语义信息;注意力机制

开放科学(资源服务)标识码(OSID):

中图分类号:TP391. 41　　**文献标志码**:A　　**文章编号**:1001-893X(2024)02-0200-07

A Small Object Detection Algorithm Fusing Vision Mechanism and Multi-scale Features

WU Debin, LIU Xiaonan, LIU Zhenyu, YANG Na

(School of Information Science and Engineering, Shenyang University of Technology, Shenyang 110870, China)

Abstract: For the problem that the Single Shot Multibox Detector (SSD) object detection algorithm has insufficient ability to detect small targets, an improved Vision Feature Fusion SSD (VFF-SSD) algorithm is proposed in which visual mechanism and multi-scale semantic information fusion are adopted. In order to increase the receptive field of the shallow network and improve the feature extraction ability, a visual mechanism is added to the SSD shallow feature layer. Then, new feature layers is obtained by using the improved path aggregation network (PANet) multi-scale feature fusion network and the deep feature enhancement network, to enhance the semantic information of the shallow network and enhance the feature expression ability of the deep features. Finally, the attention mechanism module is applied to improve the learning ability of important information. The experimental results show that the mean average precision (mAP) value detected in the PASCAL VOC2007 test set reaches 81. 1%, and the mAP value for small targets in the data set is 6. 6% higher than that of the original SSD.

Key words: small object detection; deep learning; vision mechanism; multi-scale semantic information; attention mechanism

0　引　言

随着深度学习的快速发展,目标检测技术已经在交通、医疗等各领域取得了显著的成果。近年来基于深度学习的目标检测技术在卷积神经网络的基础上不断发展,整体上可以分为有锚框与无锚框两大类,其中有锚框的目标检测算法包括两阶段目标

*　**收稿日期**:2022-09-01;**修回日期**:2022-10-09
基金项目:辽宁省自然科学基金(20180520022)
通信作者:刘笑楠　　Email:liuxiaonan@ sut. edu. cn

检测算法如 R-CNN[1]、Fast-RCNN[2]、Faster-RCNN[3]和单阶段目标检测算法如 SSD[4]、YOLO[5-6]系列。单阶段目标检测算法相较于两阶段目标检测算法检测速度更快,但检测精度较低。而无锚框的目标检测算法包括 CornerNet[7]等。

在目标检测过程中,图像中的小目标所占像素小,特征不明显,易受背景及噪声影响,导致检测精度不好,如何提高小目标的检测精度一直是目标检测技术的研究热点。虽然 SSD 算法通过多尺度特征图预测在速度与精度上有所提高,但是由于信息传递是单向的,未充分利用各特征层之间关系,其对小目标检测效果依然不理想,为此研究者们对其进行了大量的改进[8-14]。

原始 SSD 算法虽然能利用多尺度特征图检测不同大小的目标,但是浅层特征层缺少细节语义信息,检测时会存在误检漏检的问题,现有的改进方法还存在忽略深层特征层传递到浅层特征层造成小目标语义信息损失的问题。为了弥补这种小目标信息弱化的问题,本文提出一种融合视觉机制和改进的 PANet[15](Path Aggregation Network)多尺度特征提取的视觉特征融合 SSD 目标检测算法,简称为 VFF-SSD(Vision Feature Fusion SSD)。该方法旨在强化浅层特征的同时,提高对小目标的检测能力。VFF-SSD 网络在 SSD 基础上引入 RFB-S 视觉机制,使浅层特征层的感受野增大,增强特征提取能力。在前 4 个特征层使用改进 PANet 特征融合网络进行多尺度特征融合,使深层特征层的特征信息能够传递到浅层特征层,以增强浅层特征层的语义信息;利用深层特征增强模块在后 3 个特征层对深层特征层进行特征增强,增强深层特征的细节信息;最后在每一个预测特征层的最后添加 CBAM[16](Convolutional Block Attention Module)注意力机制,以提高关键信息的权重,加强对重要信息的学习能力。

1　SSD 算法

SSD(Single Shot MultiBox Detector)目标检测算法以 VGG16 为主干网络,利用多尺度特征层进行分类和定位,并将 Fc6 和 Fc7 分别转换成卷积层,用于提取浅层特征层。该算法分别在大小为 38×38,19×19,10×10,5×5,3×3,1×1 并命名为 Conv4_3,Fc7,Conv8_2,Conv9_2,Conv10_2 和 Conv11_2 的 6 个不同尺度特征层上进行检测,利用浅层特征层检测小目标,深层特征层检测大目标。整体检测过程为,首先输入图像,经过 SSD 算法对位置与类别进行预测后输出候选框,然后通过 NMS 非极大值抑制得到最终检测的预测框和类别。其结构示意图见图 1。

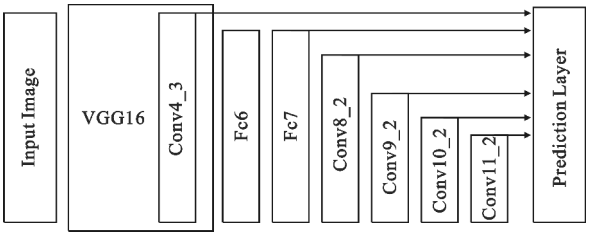

图 1　SSD 结构示意图
Fig. 1 SSD structural diagram

在原始 SSD 算法中,每个输出特征层都是直接输出结果,使得浅层特征层缺乏特征的语义信息,从而导致 SSD 算法在目标检测以及小目标检测任务上检测效果不佳。因此,针对上述问题,本文提出一种改进的融合网络增强 SSD 浅层特征层之间的关系,以提高对目标的检测能力。

2　VFF-SSD 算法

本文所提出的网络模型如图 2 所示。该网络结合了 RFB-S 视觉机制、改进 PANet 多尺度特征融合模块、深层特征增强模块和注意力机制模块,通过 RFB-S 视觉机制提高浅层网络的感受野,再利用 PANet 特征融合模块丰富语义信息,从而增强浅层特征层的特征提取能力,并在此基础上加入深层特征增强模块和注意力机制,加强上下文信息及关键信息的提取,使模型在小目标检测上的性能获得显著提高。本文算法将 7 个特征层命名为 Conv4_3,Conv5_3,Fc7,Conv8_2,Conv9_2,Conv10_2 和 Conv11_2,除 Conv5_3 以外的其余 6 个特征层对应特征图像素大小分别为(38,38),(19,19),(10,10),(5,5),(3,3)和(1,1),从而生成尺度不同的 6 个特征图。首先,在 Conv4_3 和 Fc7 特征层先使用 RFB-S 结构扩大感受野,增强特征提取能力。其次,为了丰富浅层特征层的语义信息,将 Conv4_3,Conv5_3,Fc7,Conv8_2 4 个特征层输入到改进 PANet 多尺度特征融合模块中。然后,将 Conv9_2,Conv10_2,Conv11_2 3 个特征层输入到深层特征增强模块中旨在提高深层特征层特征的表征能力,有助于确定目标的位置信息和分类信息。接着,将 6 个特征层提取到的特征图分别使用 CBAM 注意力机制模块,将关键信息权重加在原特征图中,从而提高对重要信息的学习能力。最后,使用 NMS 非极大值抑制筛选预测框,得到最终的预测结果。

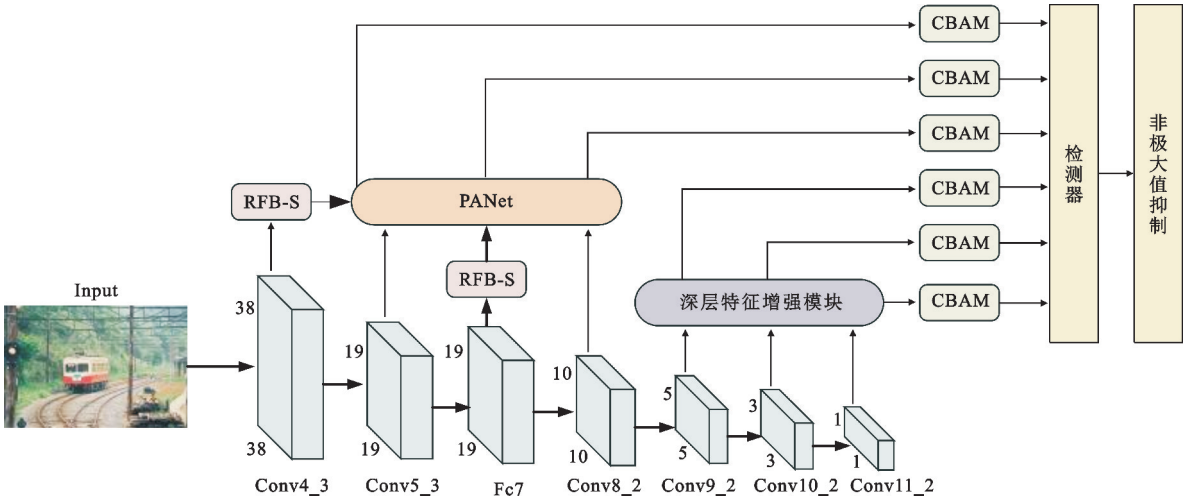

图 2 VFF-SSD 算法网络结构

Fig. 2 VFF-SSD algorithm network structure

2.1 RFB-S 视觉机制

RFB-S 视觉机制结构与 RFB 结构一样都是受启发于人的视觉感知系统,利用群体感受野来模拟人类视网膜,能够突出视网膜中心区域目标的重要性[14]。这种视觉机制结构是利用多分支卷积与空洞卷积来实现的。而在整个网络的浅层网络引入这种视觉机制,可以获取不同尺寸的特征,并且扩大感受野,从而提高网络的检测能力。本文所采用的 RFB-S 结构如图 3 所示。

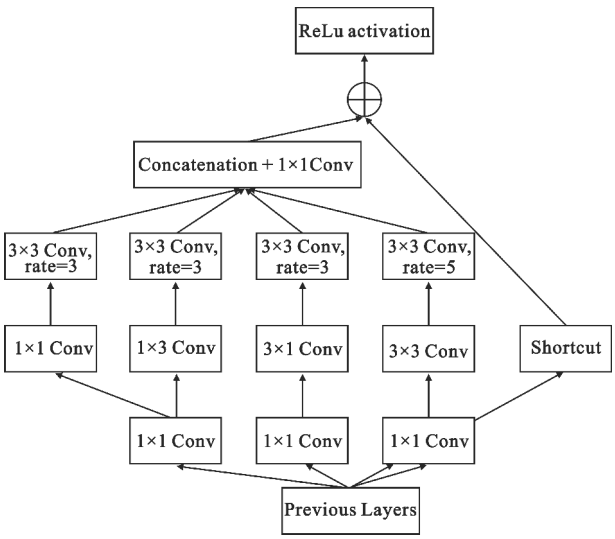

图 3 RFB-S 结构示意图

Fig. 3 RFB-S structural diagram

RFB-S 结构采用的多分支结构中,每个分支都分别采用不同大小的卷积核来得到不同比例的感受野,并且还运用了直连结构。而相对于 RFB 结构,RFB-S 结构分支更多而且卷积核大小更小,并结合空洞卷积增加浅层特征层的感受野,获取更多的上下文信息。RFB-S 结构在没有增加复杂计算量的情况下提高了感受野的范围,使网络在轻量化的同时获得具有高判别性的特征,使整个结构更加接近人类的视网膜模型。

2.2 改进 PANet 多尺度特征融合模块

针对原始 SSD 算法浅层特征层仅包含单层语义信息,导致没有足够的全局信息的问题,本文使用改进 PANet 多尺度特征融合模块将浅层特征层的位置信息与经过多次卷积的深层特征层的细节语义信息进行特征融合,使浅层特征层能够通过反向路径获取更多的语义信息,从而在确定小目标位置的基础上提高对小目标的分类能力。改进 PANet 模块结构如图 4 所示,先将 4 个原始特征图分别经过卷积核大小为 3 的卷积,进行初步的特征提取,将卷积后的 Conv8_2 进行上采样和 3×3 卷积向浅层特征层 Conv4_3 进行反向传递,并依次与卷积后的 Fc7 和 Conv4_3 进行 Concat 通道拼接;再使用 1×1 卷积进行平滑特征,最终得到 Conv4_3 的特征层。与原始 PANet 不同的是,为了加快浅层信息的传播效率,将 Conv4_3 最终的特征图进行空洞卷积下采样,达到扩大感受野提取细节信息的作用,并向 Conv8_2 层传递信息,在 Fc7 与 Conv8_2 特征层进行 Concat 特征融合,并进行 1×1 卷积平滑生成对应的最终特征图。使用改进 PANet 多尺度特征融合模块可以将深层特征层的语义信息传递到浅层特征层,并且将浅层特征层的信息有效传递到深层特征层,达到特征反复提取的目的。通过增强小目标的细节特征表达能力,增加了浅层特征层对小目标检测的优势。

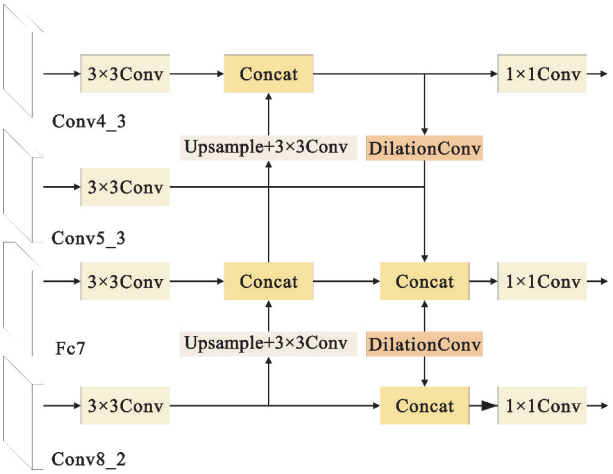

图 4　PANet 多尺度特征融合模块

Fig. 4 PANet multi-scale feature fusion module

2.3　深层特征增强模块

为了增强深层特征的细节信息,提高检测的准确性,本文采用了深层特征增强模块,如图 5 所示。首先,由于 Conv11_2,Conv10_2,Conv9_2 都包含丰富的细节语义信息,所以先使用 3×3 卷积提取局部上下文信息。其次,将卷积后的 Conv11_2 上采样并卷积与卷积后的 Conv10_2 进行特征融合,将融合后的图像继续进行上采样与卷积,然后与卷积后的 Conv9_2 进行特征融合。最后,将 Conv10_2 和 Conv9_2 两层得到的特征图经过 1×1 卷积平滑。使用深层特征增强模块旨在获得更全面的上下文信息,增强特征之间的关系。这种结构解决了局部模糊的问题并有利于对目标更好地分类。

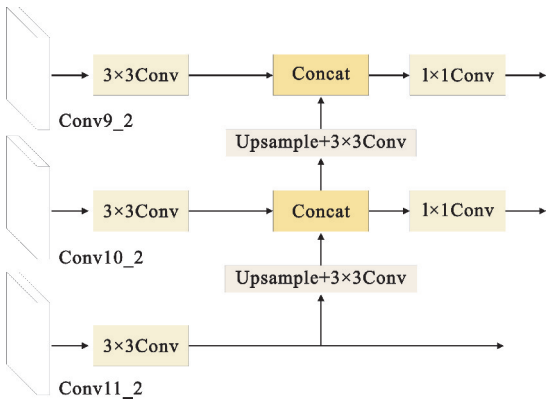

图 5　深层特征增强模块

Fig. 5 Deep feature enhancement module

2.4　注意力机制

本文所使用的 CBAM 注意力机制主要分为通道和空间注意力两部分,将每一层经过平滑卷积后的特征图先后经过上述两部分,如图 6 所示。

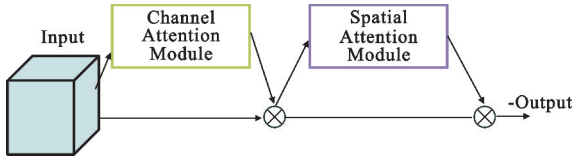

图 6　CBAM 注意力机制结构示意图

Fig. 6 Schematic diagram of CBAM attention mechanism structure

通道注意力模块如图 7 所示,将输入的特征图同时进行全局平均池化与全局最大池化,再将得到的结果送入一个共享网络,并将两个特征图按照像素求和合并,经过激活函数产生一个权重结果,最后将权重结果和输入图像相乘得到缩放后的新特征图。

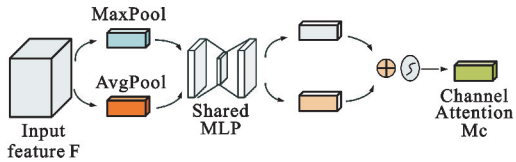

图 7　通道注意力结构示意图

Fig. 7 Schematic diagram of channel attention structure

空间注意力模块如图 8 所示。首先,通过平均池化和最大池化减少通道数。其次,将得到的两个特征图拼接融合在一起,然后再经过激活函数得到新的权重结果。最后,将这个权重结果与通道注意力生成的新结果相乘,得到最后缩放的特征图。

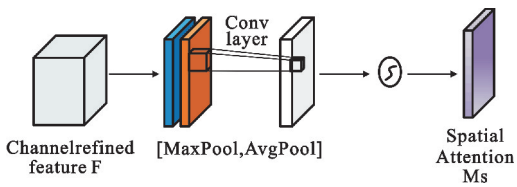

图 8　空间注意力结构示意图

Fig. 8 Schematic diagram of spatial attention structure

3　实验与分析

3.1　实验环境、实验数据集及评价指标

为测试本文算法的性能,在编译环境为 torch-1.8.0、torchvision-0.9.0、python3.8,Windows10 操作系统,显卡为 NVIDIA GeForce RTX3060 的条件下进行实验。

实验采用 PASCAL VOC 开放数据集中的 PASCAL VOC2007 和 PASCAL VOC2012 数据集的训练集进行模型训练。两个训练数据集均包括 20 个类别,共 16 651 张图片。使用 PASCAL VOC2007test 中 4 952 张图片进行模型测试,具体类别如表 1 所示。

表 1 数据集类别
Tab. 1 Dataset categories

目录	类别
Person	人
Animal	狗,猫,牛,鸟,马,羊
Vehicle	小轿车,公交车,自行车,船,火车,摩托车,飞机
Household	沙发,椅子,杯子,植物,电视,餐桌

实验使用的评价标准是所有类别的平均精度值(Mean Average Precision,mAP)和单类别平均精度(Average Precision,AP),其定义式如下:

$$P_{\mathrm{A}}=\int_0^1 P(R)\,\mathrm{d}R \tag{1}$$

$$P_{\mathrm{mA}}=\frac{\sum_{i=1}^{N} P_{\mathrm{Ai}}}{N} \tag{2}$$

式(1)和(2)中:P 表示准确率;R 表示召回率;N 表示数据集总类别数。

为了说明本文方法小目标检测的性能,实验将目标面积小于 32 pixel×32 pixel 的物体归为小目标,大于 96 pixel×96 pixel 的归为大目标,介于两者之间的归为中等目标,采用 IOU=0.5∶0.05∶0.95 10 个阈值(0.5~0.95,以 0.05 为步长)的 P_{mAs} 检测小目标的平均精度。

3.2 实验超参数设置及分析

本文使用 SGD 优化器进行 120 000 次迭代,Batch size 设为 16,共训练 116 个周期。初始学习率设为 0.001,权重衰减参数设为 0.000 5。在迭代 80 000 次学习率下降为 0.000 1,在迭代 100 000 次学习率下降为 0.000 01。将 VFF-SSD 算法与 SSD 算法的训练损失进行比较,结果如图 9 所示,可见 VFF-SSD 算法的训练损失低于 SSD 算法,证明了本文改进算法的有效性。

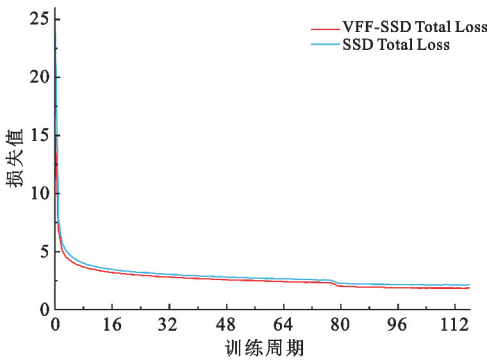

图 9 训练损失图
Fig. 9 Training loss graph

3.3 实验结果及分析

3.3.1 VOC2007 测试集目标检测性能检测对比

为了说明本文方法目标检测性能,将本文方法应用于 PASCAL VOC2007test 数据集并将实现结果与近年流行的目标检测算法的检测精度进行比较,实验结果如表 2 所示。从表 2 可以看出,本文的改进算法相对于 Faster RCNN[3] 在两种骨干网络下取得了更好的检测结果,分别提高了 7.9% 和 4.7%;相对于同为单阶段目标检测算法的 YOLOv1[5] 和 YOLOv2[6] 有显著提升,分别提高了 17.7% 和 7.4%;相较于原始 SSD 算法检测精度提高 3.9%,而与其他研究者们对原始 SSD 算法改进的 DSSD[17]、RSSD[18]、文献[9]、文献[10]也有一定的优势,说明 VFF-SSD 方法在输入尺寸为 300 pixel 的情况下,检测精度具有明显的优势,较其他输入分辨率较大的目标检测方法检测效果也有一定的提高。

表 2 在 PASCAL VOC2007 测试集本方法与其他方法对比(IOU=0.5)
Tab. 2 Comparison between this method and other methods on the PASCAL VOC2007 test set(IOU=0.5)

方法	骨干网络	输入尺寸/pixel	mAP/%
Faster RCNN[3]	VGG16	1 000×600	73.2
Faster RCNN[3]	ResNet-101	1 000×600	76.4
SSD[4]	VGG16	300×300	77.2
YOLOv1[5]	GoogleNet	448×448	63.4
YOLOv2[6]	Darknet-19	352×352	73.7
DSSD[17]	ResNet-101	321×321	78.6
RSSD[18]	VGG16	300×300	78.5
文献[9]	VGG16	300×300	78.0
文献[10]	VGG16	300×300	78.0
VFF-SSD	VGG16	300×300	81.1

3.3.2 小目标检测性能评估

为了进一步说明本文方法对小目标检测的性能,实验使用相应的评价指标 P_{mAs} 对实验结果进行评估,并与两种主干网络下的两阶段目标检测算法 Faster RCNN、两种主干网络下的 SSD 算法进行比较,实验结果如表 3 所示。由表 3 可以发现,本文提出的 VFF-SSD 算法对小目标的检测精度达到 17.2%,远高于其他算法,相较于原始 SSD 算法有显著提升,提高 6.6%,证明了本文改进算法对小目标检测有明显的优势。虽然 VFF-SSD 算法对浅层特征进行反复特征提取导致算法计算量有一定增加,但是该算法对小目标的检测性能提升显著,整体看该算法对小目标效果良好,可以应用到实际场景中。

表 3　算法性能比较

Tab. 3 Algorithm performance comparison

算法	骨干网络	P_{mAs}/%
Faster RCNN	VGG16	8.4
	ResNet-101	10.7
SSD	VGG16	10.6
	ResNet-50	9.5
VFF-SSD	VGG16	17.2

3.3.3　单类别实验结果及分析

实验还将数据集中 20 个类别检测精度与 4 种流行算法以及传统 SSD 算法进行了单类别比较,结果如表 4 所示。从表 4 中可以发现,VFF-SSD 算法有 18 个类别物体的检测精度超过其他对比算法,尤其是像杯子、植物、飞机等这种在图片中所占像素比低的小目标,而其他类别检测精度与几种算法最优的检测精度相差甚微;与传统 SSD 算法相比,VFF-SSD 算法 20 个类别检测精度均超过原始 SSD 算法。根据表中数据,本文改进算法对远景小目标有良好的检测精度,对近景目标检测精度有待加强。综上,VFF-SSD 算法对小目标检测有显著的优势。

表 4　单类别精度比较

Tab. 4 Single-category accuracy comparison

种类	精度/%					
	Faster RCNN	SSD	DSSD	文献[9]	文献[10]	VFF-SSD
飞机	76.5	77.7	81.9	80.5	79.5	**85.7**
自行车	79.0	84.0	84.9	85.0	85.4	**89.4**
鸟	70.9	76.3	**80.5**	76.6	77.9	79.0
船	65.5	71.3	68.4	68.1	70.4	**72.4**
杯子	52.1	48.6	53.9	52.2	50.1	55.9
公交车	83.1	85.3	85.6	85.4	85.3	**88.6**
小轿车	84.7	86.3	86.2	86.3	86.1	**89.0**
猫	86.4	87.3	88.9	86.2	88.4	**90.5**
椅子	52.0	57.3	61.1	61.2	60.4	**64.3**
牛	81.9	81.8	83.5	83.5	80.9	**90.6**
餐桌	65.7	77.9	78.7	75.9	79.1	**79.5**
狗	84.8	85.2	86.7	84.8	85.8	**91.3**
马	84.6	87.0	88.7	87.3	87.8	**91.3**
摩托车	77.5	83.7	86.7	86.2	85.5	**87.8**
人	76.7	79.2	79.7	79.1	79.4	**81.8**
植物	38.8	53.0	51.7	53.5	53.6	**54.8**
羊	73.6	76.3	78.0	**79.0**	78.1	77.6
沙发	73.9	79.8	80.9	81.4	79.7	**82.6**
火车	53.0	88.3	87.2	87.2	87.1	**90.1**
电视	72.6	76.2	79.4	78.4	78.1	**80.7**
精度	73.2	77.2	78.6	78.0	78.0	**81.1**

3.3.4　定性结果分析

相应的检测结果可视化,如图 10 所示。从图 10 可以明显看出,本文提出的改进算法相较于 SSD 算法可以有效地解决小目标漏检的问题。通过对比发现,原始 SSD 算法未检测出的牛和羊、错检的羊、密集的小车、密集重叠的人,经过 VFF-SSD 算法都可以被检测出来,更进一步证明增强浅层语义信息与加强深层特征融合能够提高模型对小目标的检测能力。

(a) SSD算法　　(b) VFF-SSD算法

图 10　定性结果分析

Fig. 10 Qualitative result analysis

4　结束语

针对原始 SSD 算法对小目标检测精度不高以及漏检错检的问题,本文提出一种利用 RFB-S 视觉机制增强浅层网络感受野并结合改进 PANet 加强浅层网络语义信息提取,利用深层特征增强模块进行深层网络特征增强,同时使用 CBAM 注意力机制加强对关键信息学习的 VFF-SSD 改进算法。在 PASCAL VOC2007test 数据集上使用两种评价指标进行评估,分别得到检测精度有效提高,尤其对小目标的 mAP 达到 17.2%,较原 SSD 算法提高了 6.6%。

下一步工作拟解决小目标漏检错检问题以及提高小目标检测能力问题,研究如何在不减少精度的情况下将模型进行轻量化。

参考文献:

[1] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos:IEEE,2014:580-587.

[2] GIRSHICK R. Fast R-CNN[C]//Proceedings of 2015 IEEE International Conference on Computer Vision. Los Alamitos:IEEE,2015:1440-1448.

[3] REN S Q,HE K M,GIRSHICK R,et al. Faster R-CNN: towards real-time object detection with region proposal networks[C]//Proceedings of the 28th International Conference on Neural Information Processing Systems. Cambridge:MIT Press,2015:91-99.

[4] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[C]//Proceedings of 2016 European Conference on Computer Vision. Heidelberg: Springer,2016:21-37.

[5] REDMON J,DIVVALA S,GIRSHICK R,et al. You only look once: unified, real-time object detection[C]// Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE, 2016:779-788.

[6] REDMON J, FARHADI A. YOLO9000: better, faster, stronger[C]//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE,2017:7263-7271.

[7] LAW H,DENG J. CornerNet:detecting objects as paired keypoints[C]//Proceedings of 2018 European Conference on Computer Vision. Munich: IEEE, 2018: 765-781.

[8] 杨艳红,钟宝江,徐云龙. 改进的 SSD 算法在智慧交通中的应用[J]. 电讯技术,2022,62(2):259-265.

[9] 叶召元,郑建立. 基于自动驾驶场景的目标检测算法 DFSSD[J]. 计算机工程与应用,2020,56(16):139-147.

[10] 谭红臣,李淑华,刘彬,等. 特征增强的 SSD 算法及其在目标检测中的应用[J]. 计算机辅助设计与图形学学报,2019,31(4):573-579.

[11] LIN T Y,DOLLÁR P,GIRSHICK R,et al. Feature pyramid networks for object detection[C]//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos:IEEE,2017:2117-2125.

[12] LIM J S, ASTRID M, YOON H J, et al. Small object detection using context and attention[C]//Proceedings of 2021 International Conference on Artificial Intelligence in Information and Communication. Jeju Island:IEEE,2021:181-186.

[13] LIU S,HUANG D. Receptive field block net for accurate and fast object detection[C]//Proceedings of 2018 European Conference on Computer Vision. Cham: Springer,2018:385-400.

[14] YU F, KOLTUN V. Multi-scale context aggregation by dilated convolutions[EB/OL]. (2015-11-23)[2022-08-31]. https://arxiv. org/abs/1511. 07122.

[15] LIU S,QI L,QIN H,et al. Path aggregation network for instance segmentation[C]//Proceedings of 2018 IEEE Computer Vision and Pattern Recognition. Los Alamitos: IEEE,2018:8759-8768.

[16] WOO S,PARK J,LEE J Y,et al. CBAM:Convolutional block attention module[C]//Proceedings of 2018 European Conference on Computer Vision. Cham: Springer,2018:3-19.

[17] FU C Y,LIU W,RANGA A,et al. DSSD:deconvolutional single shotdetector[EB/OL]. (2020-04-08)[2022-08-31]. https:arxiv. org/abs/1701. 06659.

[18] JEONG J,PARK H,KWAK N. Enhancement of SSD by concatenating feature maps for object detection[EB/OL]. (2017-05-26)[2022-08-31]. https://arxiv. 53yu. com/abs/1705. 09587.

作者简介:

武德彬 男,1998 年生于辽宁大连,2020 年获学士学位,现为硕士研究生,主要研究方向为计算机视觉、深度学习。

刘笑楠 女,1979 年生于辽宁沈阳,2014 年获博士学位,现为副教授,主要研究方向为图像处理、计算机视觉。

刘振宇 男,1973 年生于辽宁沈阳,2002 年获博士学位,现为教授,主要研究方向为视觉伺服技术、计算机视觉。

杨 娜 女,1999 年生于河北保定,2021 年获学士学位,现为硕士研究生,主要研究方向为图像处理、计算机视觉。