

DOI:10.16136/j.joel.2023.02.0178

一种轻量化的水下目标检测算法 SG-Det

周华平*, 宋明龙, 孙克雷

(安徽理工大学 计算机科学与工程学院, 安徽 淮南 232000)

摘要: 基于深度学习的目标检测算法在水下进行检测主要存在两个困难: 水下设备的存储和计算能力有限; 水下图像模糊且小生物聚集。这种局限性要求水下目标检测算法要做到轻量且高效, 因此现有的目标检测算法不能完全满足水下目标检测的需求。为此本文在 SSD(single shot MultiBox detector) 的网络框架上进行改进, 设计了一种轻量化的一阶段检测模型 SG-Det。一方面, 借鉴 GhostNet 的思想, 对 ShuffleNetV2 网络进行了重构, 提出一种新的轻量化特征提取网络 SGnet。此网络进一步减少了模型参数量, 使模型大小更适合部署在水下设备。另一方面, 网络主要是利用 6 个不同尺度的特征图检测不同大小的生物, 为此设计了一种引入双分支注意力机制的跨尺度特征融合模块(cross-scale feature fusion module, AFF)。模块首先引入注意力机制对输入的特征在全局通道和局部通道两方面进行加权, 突出有用信息, 从而减少背景等无关信息的干扰。然后选取非线性化程度更高的第 4 层分别增强前 3 层的语义信息, 以较小的代价使前 3 层在识别小物体方面有更好的表现。模型在中国水下机器人大赛提供的水下数据集 UPRC 进行测试, 平均检测精度(mAP)和速度分别达到了 71.75% 和 69 FPS, 且模型参数量仅有 4.91 M。结果表明, 所提出的方法在精度、速度和参数量之间取得了很好的平衡。

关键词: 神经网络; 水下目标检测; 轻量化; 跨尺度特征增强

中图分类号: TP391.41 **文献标识码:** A **文章编号:** 1005-0086(2023)02-0156-10

SG-Det: A lightweight underwater image target detection method

ZHOU Huaping*, SONG Minglong, SUN Kelei

(College of Computer Science and Engineering, Anhui University of Science and Technology, Huainan, Anhui 232001, China)

Abstract: Deep learn-based object detection algorithm has two main difficulties in underwater target detection, which are limited storage and computing capabilities of underwater equipment and fuzzy underwater images and small organisms gathering. This limitation requires the underwater target detection algorithm to be lightweight and efficient, so the existing target detection algorithm can not fully meet the needs of underwater target detection. In this paper, a lightweight one-stage detection model SG-DET is designed to improve the network framework of single shot MultiBox detector (SSD). On the one hand, using the idea of GhostNet for reference, ShuffleNetV2 network is reconstructed, and a new lightweight feature extraction network SGnet is proposed, which further reduce the number of model parameters and made the model size more suitable for deployment in underwater equipment. On the other hand, the network mainly uses six feature maps of different scales to detect organisms of different sizes. For this purpose, a cross-scale feature fusion module (AFF) using two-branch attention mechanism is designed. The AFF module firstly introduces the attention mechanism to weight the input features in both global channel and local channel, so as to highlight useful information. To improve the problem of low detection accuracy caused by small and fuzzy objects, the fourth layer with a higher degree of nonlinearity is selected to enhance the semantic information of the first three layers, so that the first three layers have a better

* E-mail: mlsong97@163.com

收稿日期: 2022-04-19 修订日期: 2022-05-02

基金项目: 安徽省重点研发计划、国际科技合作专项、矿用防爆纯电动运输车辆关键技术的研究与开发(202004b11010029)资助项目

performance in identifying small objects with a small cost. The model is tested on the underwater data set UPRC provided by China Underwater Robot Competition, and the mAP and speed reach 71.75% and 69 FPS respectively, while the number of model parameters is only 4.91 M. The results show that the proposed method achieves a good balance between precision, speed and number of parameters.

Key words: neural networks; underwater object detection; lightweight networks; feature enhancement by cross-scale

0 引言

水下生物是水下生态环境的重要组成部分,水下生态保护组织通常通过人工潜水的方式进行水下拍摄,以便对水下生物的分布和生活习性进行研究。由于水下环境复杂多变、水下图像模糊,研究人员难以准确发现水下生物,因此,用水下机器人代替人眼进行检测将成为一种趋势。

基于深度学习的目标检测算法在通用数据集上保持着较高的准确率,相比于一般的目标检测任务,现在主流的目标检测网络不完全适用于水下生物的检测。其中出现的问题一方面是由于水下采集的部分图像不清晰、水下生物较小且大量聚集,需要设备部署识别精度高的网络;但另一方面,精度高的网络往往是以参数量和速度为代价的,而水下设备一般存储能力有限,并不适合使用大型的深度神经网络。

近年来,人们开始探索一些针对水下生物的目标检测方法,例如 SHI 等^[1]使用 ResNet 和双向特征金字塔网络(bi-directional feature pyramid network, BiFPN)对 Faster-RCNN 进行改进并应用于水下生物的检测;PAN^[2]和 XU 等^[3]改进了基于多尺度特征的 ResNet 方法,结合高层语义特征和低层空间特征对不同尺度的目标进行检测;MATHIAS 等^[4]将视觉特征和高斯混合模型与 YOLOv3(you only look once v3)深度网络相结合,提高了水下目标自动检测方法的检测效率。上述方法都对水下目标的检测精度有明显提升,但对于存储空间和计算能力有限的水下嵌入式设备来说,网络还是过于冗余。为了减少网络的参数量,ZHANG 等^[5]对 YOLOV4 进行改进,用轻量级的 MobileNetV2 作为主干网络,NASEER 等^[6]使用 Inception 网络对 Faster-RCNN 进行轻量化改进。两种方法均在牺牲了部分检测精度的情况下大大

压缩了模型参数,从而在水下摄像机上实现了对生物的自动检测。综合国内外的研究现状发现,当下关于轻量化水下目标检测网络的研究依然较少,目前的一系列算法仍存在改进空间。

为了更有效的进行水下目标检测,本文设计了一种轻量化检测网络 SG-Det,在算法精度和复杂度之间取得了很好的平衡。本文的贡献有以下几点:

1) 更少的参数量:借鉴 GhostNet^[7]思想对 ShuffleNetV2^[8]网络结构进行改进,设计了一种更加轻量化的特征提取网络 SGnet,既获得了更快的检测速度,也大大减少了模型参数的数量。2) 更轻量的特征融合:提出一种融合注意力的跨尺度特征融合模块(cross-scale feature fusion module, AFF),以较小的代价增强浅层特征的语义信息,使其有利于水下小目标的检测。

1 相关工作

1.1 SSD 及其改进

SSD(single shot MultiBox detector)^[9]使用 VGG 作为网络的 Backbone 进行特征提取,选取 6 个不同尺度的特征层直接回归检测框,对大尺寸特征图使用小的回归框,既保证了较快的检测速度,又确保了检测精度。这种方法存在的一个缺陷是尺寸大的特征层卷积次数较少,这意味着学到的可用信息较少,因此 SSD 对小目标的检测效果较差。后期人们为了提高小目标的检测精度对网络进行了多次改进,如在 DSSD^[10]中,把 Backbone 改成了 Residual-101,添加了反卷积模块,使前后特征层信息复用,以参数量的增加和检测速度的降低为代价大大提高了小目标检测准确率;FSSD^[11]中设计了一个轻量级的特征融合模块,使用不同尺度的特征图相互增强,在速度稍有下降的同时提高了 SSD 的检测精度。SSD 网络结构见图 1。

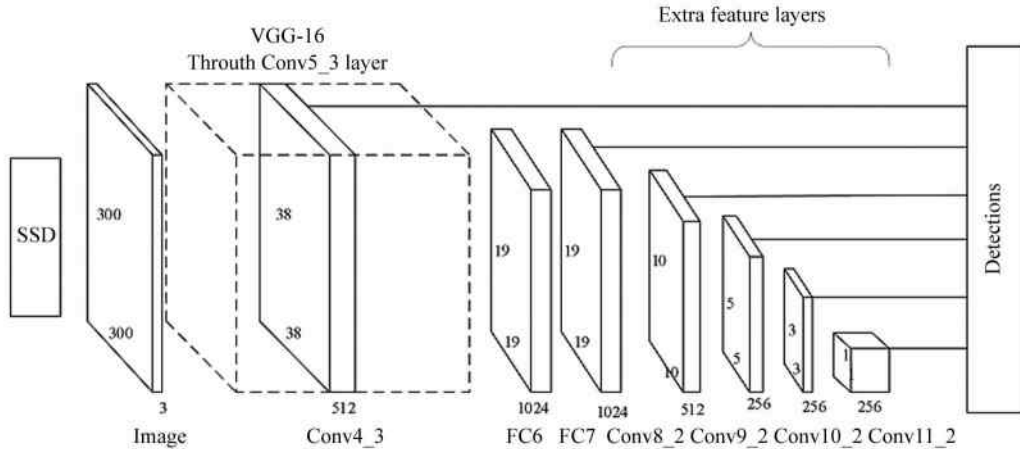


图 1 SSD 网络结构图

Fig. 1 SSD network structure

1.2 轻量化网络

模型轻量化是神经网络的另一个发展方向,例如 MobileNetV1^[12] 和 MobileNetV2^[13] 通过使用深度可分离卷积来代替传统卷积,ShuffleNet^[14] 使用分组卷积和通道重排来进行模型压缩。GhostNet 使用了另一种思想,其作者通过观察发现图片经过卷积操作之后,存在大量非常相似的特征图,因此考虑只用传统卷积方式生成部分特征图,其余的特征图不再通过卷积生成,而是通过廉价的线性运算生成。对于传统卷积给定输入数据 $X \in R^{c \times h \times w}$ (c, h, w 分别是输入数据的维度、高度和宽度)、输出数据 $Y \in R^{h' \times w' \times n}$ (n 是特征图通道数, h' 和 w' 分别是输出数据的高度和宽度)、卷积核大小为 $k \times k$, 则所需要的浮点运算数 (floating-point operations, FLOPs) 为 $n \times h' \times w' \times c \times k \times k$ 。对于 Ghost 模块,同样给定上面的数据,假设 m 个特征图是由原始卷积生成,每个原始模块变幻的数量为 s ,每个线性运算的平均内核大小为 $d \times d$,则线性变换的 FLOPs 仅为 $(s-1) \times m \times h' \times w' \times d \times d$ 。经过式(1)分析可以得出,模型的参数压缩比为 S ,因此使用 Ghost 模块网络的参数和计算成本显著降低。

$$r_c = \frac{n \times c \times k \times k}{\frac{n}{s} \times c \times k \times k + (s-1) \times \frac{n}{s} \times d \times d} \approx \frac{s \times c}{s + c - 1} \approx s. \quad (1)$$

2 SG-Det 结构设计

SG-Det 算法主要由轻量级提取特征的网络 SG-net、AFF 和分类及回归子网络构成,网络总体框架如图 2 所示。SGnet 借鉴了 ShufflenetV2 和 Ghost-

Net 的思想,可以在保证特征提取能力的同时大大减少模型参数量。AFF 模块是一个引入双分支注意力机制的轻量化特征融合模块,可以在不同尺度的特征层上进行 3 次跨尺度语义增强,通过增加少量的参数提高了模型对小目标的检测能力。在分类和回归子网络中,选取 6 个不同尺度的特征层作为额外层,每个层设置大小和数量均不相同的检测框,总共选取了 8732 个候选框进行下一步的筛选,最后选取合适的框生成检测结果。

2.1 特征提取网络 SGnet

2.1.1 扩张块

特征提取网络是目标检测的基础,也是检测网络最重要的部分。受 Xception^[15] 的启发,本文在第一个特征提取块之前设计了一种经济高效的特征扩张块 (Expand Block)。扩张块结构如图 3(a) 所示,网络可以在不增加太多计算开销的情况下,增加首个卷积层的通道数,有效地提高特征表达能力,实现了特征的初步提取。其中用 Ghost 模块代替 1×1 卷积,可以进一步减少参数和计算量。

2.1.2 提取块

设计轻量化网络要考虑计算资源和能力,因此不能设计过深的网络。用 Ghost 模块代替 ShufflenetV2 分组卷积中的逐点卷积,构造了特征提取块 (Extract Block) 来代替传统卷积。在图 3(b) 所示的 Extract BlockA 模块中,首先对输入特征进行一个通道分割 (channel split),将通道平分成相等的两部分;然后保持其中一个分支不做变化,另一个分支利用 Ghost 模块进行通道扩张,然后进行 3×3 的深度卷积,再利用一个 Ghost 模块进行通道压缩;接着将两个分支进行通道拼接,这样做可以保持运算前后通

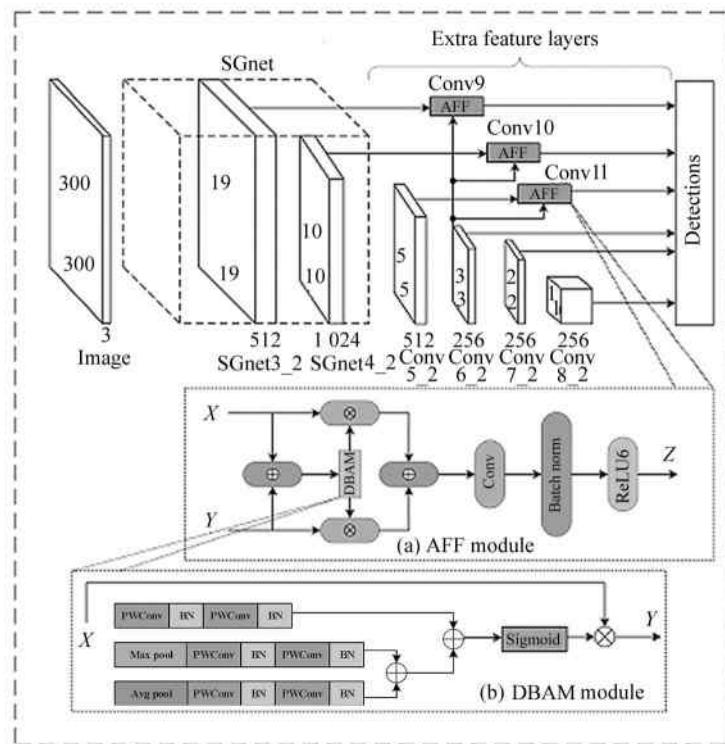


图2 SG-Det 网络整体结构

Fig. 2 Overall structure of SG-Det network

道数相同,使内存访问更为高效。最后通过一个通道混洗(channel shuffle),在不增加计算量的情况下对通道进行充分的混合。图3(c)的 Extract BlockB 与 Extract BlockA 相比,深度卷积的步长(stride)设

为2,特征图就被下采样到原来的1/2,在增大卷积核感受野同时过滤有干扰的噪声。同时 Extract BlockB 不需要进行 channel split 操作,拼接后通道数扩大一倍,减少特征信息的丢失。

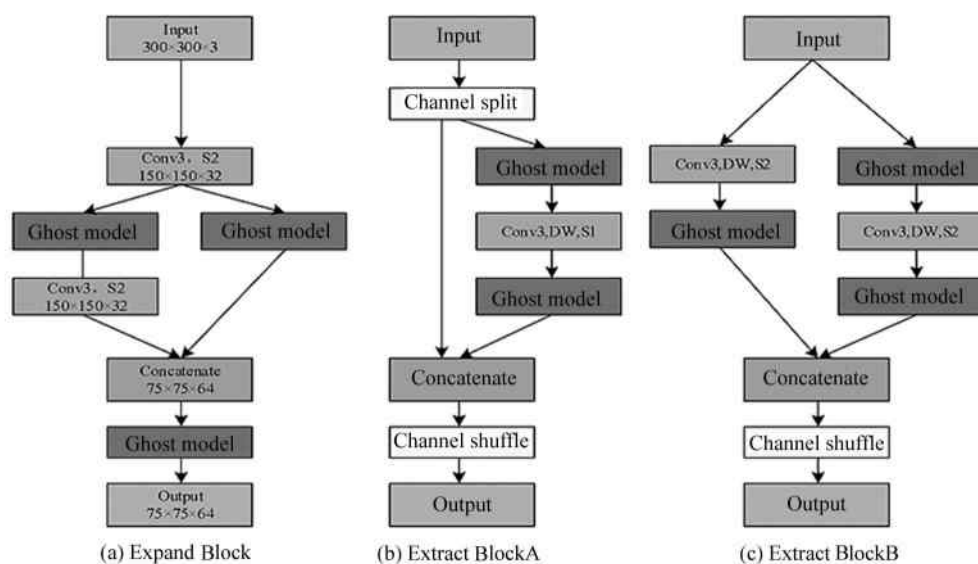


图3 SGnet 子模块

Fig. 3 SGnet unit

2.1.3 SGnet 网络整体结构

根据上文设计的 SGnet unit,构建了如表1所示

的骨干网络 SGnet。SGnet 网络整体结构分为4个阶段:Stage0 是由 Expand Block 构成以进行特征的

初步提取,这个阶段将特征图下采样到 75×75 ,特征图维度增加到 64,可以有效增强特征表达能力; Stage1 由上文提出的若干 Extract BlockA 和 Extract BlockB 堆叠而成(其中 stride 为 1 表示使用 Extract BlockA, stride 为 2 表示使用 Extract BlockB, Repeat 代表使用该模块的次数),首先用一个 Extract BlockB 模块,将特征图下采样为 38×38 的尺寸,然后连续使用 3 个 Extract BlockA 模块进行特征提取,然后将提取后的特征图送入 Stage2; 同理 Stage2 和 Stage3 分别使用 1 个 Extract BlockB、7 个 Extract BlockA 和 1 个 Extract BlockB、3 个 Extract BlockA 进行特征提取。在每个 Stage 中,输入 unit 的特征的一半维度直接与下一个 unit 相连,这种特征重用可以在保持高效率的同时提高检测精度。

表 1 SGnet 网络整体结构
Tab.1 SGnet network framework

Layer	Output size	Stride	Repeat	Output channel
Image	300×300	—	—	3
Stage0	75×75	2	1	64
Stage1	38×38	2	1	256
	38×38	1	3	256
Stage2	19×19	2	1	512
	19×19	1	7	512
Stage3	10×10	2	1	1 024
	10×10	1	3	1 024
Global avgpool	1×1	—	—	—
FC	—	—	—	1 000

2.2 跨尺度特征增强

2.2.1 双分支注意力模块

在水下图像中,由于各种水下生物的颜色与背景色相似导致检测精度的降低,所以在 AFF 模块中设计了如图 2(b)所示的双分支注意力模块(dual-branch attention module, DBAM)来缓解这一情况。该模块通过在全局特征和局部特征通道上加权来放大有用的物体信息,消除无用的背景干扰等信息。DBAM 模块利用两个不同的分支来获得通道注意力权重。一个分支分别使用全局最大池化和全局平均池化来研究全局特征的通道信息,而另一个分支使用逐点卷积来提取局部特征的通道信息。最后将两个分支融合在一起,以此来更好地组合不同尺度的特征。

经过 DBAM 得到特征 $Y \in R^{C \times H \times W}$ 的过程可以解释为式(2)。DBAM(X)表示经过 DBAM 得到的注意力权重,其中 X 是 CNN 中某一层的输出, C 为

通道数, $H \times W$ 是特征图大小。G(X)和 L(X)分别表示全局特征和局部特征通道上下文,可表示为式(3)和(4):

$$Y = X \otimes DBAM(X) = X \otimes (\sigma(G(X) \oplus L(X))), \quad (2)$$

$$G(X) = MLP(Avgpool(X)) \oplus MLP(Maxpool(X)), \quad (3)$$

$$L(X) = B(PwConv_2(\delta(B(PwConv_1(X))))). \quad (4)$$

式中, Avgpool(X)表示全局平均池化, Maxpool(X)表示全局最大池化, B 表示批量归一化, δ 表示 ReLU 函数, σ 表示 Sigmoid 函数, $\oplus$ 表示广播机制的加法运算, $\otimes$ 表示逐元素乘法运算。

2.2.2 AFF

网络需要在浅层特征进行小物体的检测,因此设计了如图 2(a)所示的融合双分支注意力机制的 AFF,它可以使浅层特征获得更多高层特征的语义信息。实验结果表明: Conv7 和 Conv8 完全丢失了小物体的细节,所以模型使用了 Conv6 之前的特征图进行特征增强。图 4 所示是将 SGnet3_2 和 Conv6_2 进行跨尺度特征增强的过程,首先需要将 Conv6_2 特征图上采样到与 SGnet3_2 相同大小;其次,将上采样的输出和 SGnet3_2 分别依次通过一个 3×3 的卷积层、ReLU 激活函数层和 L2 正则化层后对两个特征图进行求和输出后传递给 DBAM 模块,通过得到相应的注意力权重后对原来特征图进行调整;最后,将两个特征相加后使用 $3 \times 3 \times 512$ 的卷积核保证特征的可分辨性,并经过 ReLU 层后生成新的特征图 Conv9。同理通过另外两个 AFF 模块生成了增强后的 Conv10 和 Conv11。在 SG-Det 网络的分类和回归子网络中将 Conv6_2、Conv7_2、Conv8_2 与特征增强后的 Conv9 (19×19)、Conv10 (10×10)、Conv11 (5×5)一块发送到预测模块进行预测。

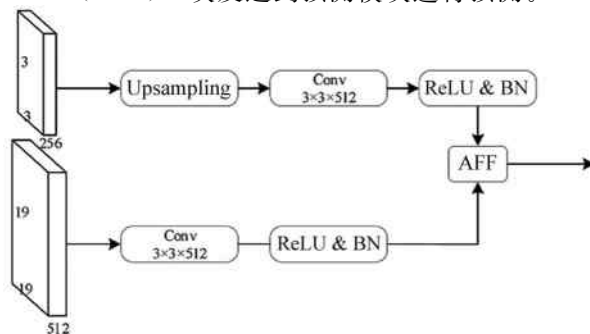


图 4 SGnet3_2 和 Conv6_2 跨尺度特征增强

Fig. 4 Conv6_2 features enhanced SGnet3_2

3 实验结果分析

3.1 数据集来源及处理

本文共使用两个水下生物数据集。第一个是UPRC水下数据集(2020年全国水下机器人目标抓取大赛提供),UPRC共有6185张图片,又通过翻转、裁剪等方式将数据集扩充到11500张,图片中包含海参、海胆、扇贝、海星、水草等5类水下生物,但因为水草标注过少(仅114)而不参与预测。第二个

是自制的海洋生物数据集UWD数据集,共有4254张图片,又通过翻转、裁剪等方式将数据集扩充到7300张,数据集中包含鱼、水母、企鹅、鲨鱼、海星、黄貂鱼、海鹦等7种海洋生物。其中UPRC数据集由水下设备拍摄,如图5(a)所示,大部分图片存在模糊、失真等现象。UWD数据集是在水族馆拍摄的清晰图片,如图5(b)所示,图片质量较高。通过在2种不同清晰度的数据集上进行实验,验证SG-Det的鲁棒性。

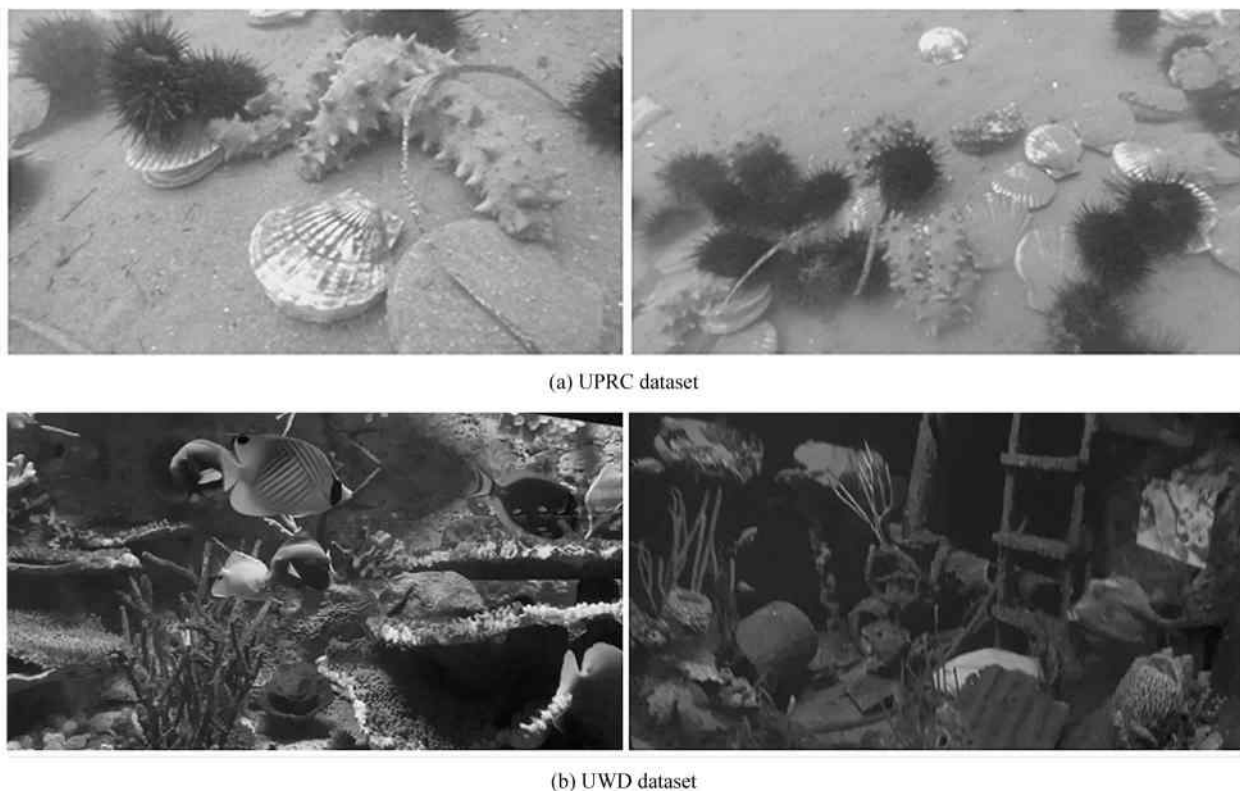


图5 两个水下生物数据集

Fig. 5 Two underwater biological dataset

3.2 实验环境及评价指标

实验操作系统为Ubuntu18.04,显卡为NVIDIA GeForce RTX 2080Ti,CUDA为10.0版本。使用深度学习框架Pytorch1.3.1对SG-Det进行训练、测试以及可视化分析,首先在PASCALVOC上进行预训练,初始学习率设置为0.001,批处理大小为32,每经过50个Epoch将学习率降为原来的10%。目标检测模型的性能优劣可以从检测精度和速度两个方面进行评价。检测速度以帧每秒(FPS)为单位,即模型每秒能够检测到的图像数。对于单类检测,本文以精准率(Precision)、召回率(Recall)作为检测精度的评价指标。这些被定义为式(5)和

式(6):

$$Precision = \frac{TP}{TP + FP}, \quad (5)$$

$$Recall = \frac{TP}{TP + FN}, \quad (6)$$

式中, TP 、 FP 、 FN 分别为准确检测到、错误检测到、未检测到的对象数。以精准率和查全率所包围的图形区域作为这类对象的平均检测精度(AP)。

在检测种类较多的情况下,检测精度一般以平均精度均值(mAP)作为评价指标,其中 N 为数据集中对象类别的个数。

$$mAP = \frac{1}{N} \sum AP. \quad (7)$$

由于模型在考虑精度的同时,还要兼顾模型的参数量,因此计算量(FLOPs)和参数量(Parameters)也作为评价的重要指标。

3.3 实验结果

3.3.1 消融实验

本文进行了消融实验来验证不同子模块或手段的有效性。表2显示了在UPRC上进行消融实验的结果。其中‘√’表示使用了相应的子模块。在列出的模型中,Model4使用了SSD算法的主要框架,Model1将SSD的主干网络替换为SGnet,其余网络结构与SSD一致;Model2是将AFF模块加入到Model1的模型;Model3是将DBAM模块加入到Model2的模型。

表2所示的实验结果表明,Model1与Model4相比,模型在精度相差不大的情况下,使用SGnet模块后模型大小降为原来的1/9,速度提升了将近一倍,网络的参数和计算成本显著降低。Model2与Model1相比,因为进行了多尺度的特征融合,速度下降了10 FPS,但参数量仅有少量的增加且精度提升了3%。表3显示了在使用AFF模块前后模型各类精度的变化,表4显示了在使用AFF模块前后召回率的变化,结果表明海参和海胆的准确率和召回率得到了明显的提升,其他目标也有不同程度的提升。图6是使用AFF模块前后的检测效果对比,明显看出在使用了AFF模块后,小海参的漏检和模糊的海胆的漏检现象都有很大的改善,且准确率明显提升。实验结果充分说明了使用AFF模块进行跨尺度特征融合可以有效提升小目标的检测性能。Model3与Model2相比,在参数量和速度相差不大的情况下进一步提升了检测精度。综上所述,本文所提出的方法是有效的,以小的成本实现了大的收益。

表2 UPRC上的消融实验结果

Tab. 2 Test results of the ablation experiment on the UPRC

Model	Methods			$mAP/\%$	Parameters /M	Speed /FPS
	SGnet	AFF	DBAM			
Model1	√	—	—	68.95	3.75	81
Model2	√	√	—	71.55	4.87	71
Model3	√	√	√	71.75	4.91	69
Model4	—	—	—	68.75	34.3	46

表3 模型有无AFF模块的单类精度比较

Tab. 3 Accuracy comparison of model with AFF module or not

Methods	AP				$mAP/\%$
	Holothurian	Echinus	Scallop	Starfish	
With AFF	0.554 0	0.764 1	0.499 5	0.738 9	71.55
Without AFF	0.393 5	0.720 8	0.436 5	0.700 1	68.95

表4 模型有无AFF模块的召回率比较

Tab. 4 Comparison of recall rate of model with and without AFF module

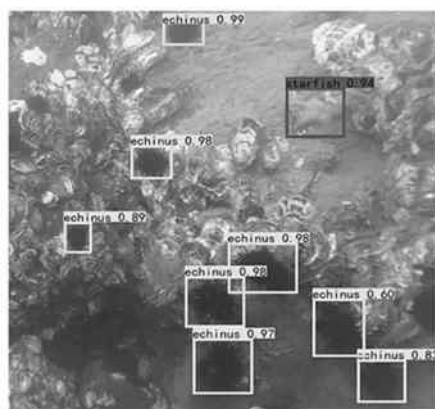
Methods	Holothurian	Echinus	Scallop	Starfish
With AFF	0.554 0	0.764 1	0.499 5	0.738 9
Without AFF	0.393 5	0.720 8	0.436 5	0.700 1

3.3.2 SG-Det与其他目标检测算法在水下数据集上的比较

为了验证SG-Det算法的有效性,在PASCAL VOC上进行预训练后的SG-Det模型分别转移到水下数据集UPRC和UWD上进行最终训练,总epoch设置为200(freezing epoch设置为30,non-freezing epoch设置为170),其他训练参数保持不变,训练过程的损失曲线图如图7所示。对比使用6种检测的网络结构与本文提出的检测器,分别为SSD、DSSD、YOLOV4、YOLOV4-tiny、SSD-ShuffleNetV2、PeleNet,所有模型都属于单阶段检测器,其中YOLOV4-



(a) Without AFF



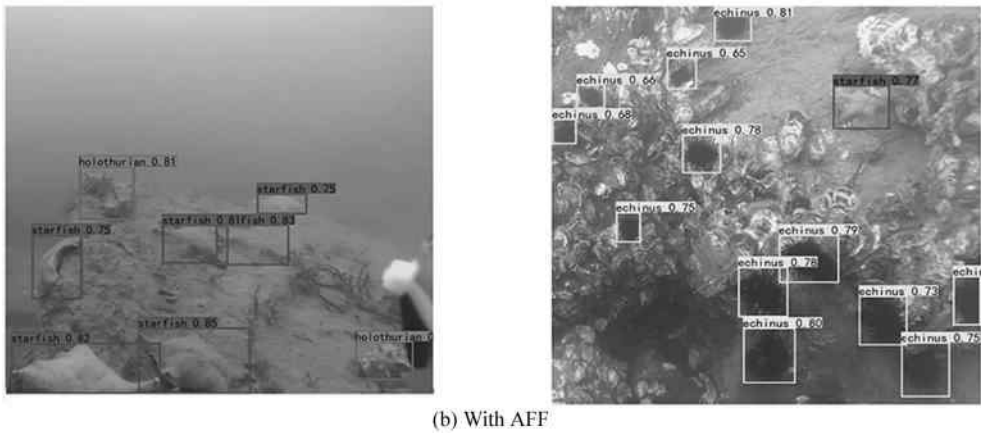


图 6 模型有无 AFF 模块的检测效果对比

Fig. 6 Comparison of the detection effect with and without AFF module in the model

tiny、SSD-ShuffleNetV2、PeeleNet 又是轻量级的检测器。

表 5 和表 6 的数据显示,算法在两个数据集上与其他算法对比均有较好的表现。其中 SG-Det 算法在 UWD 上 mAP 能达到 87.3%,而且检测速度能达到 67 FPS,完全满足实时检测的要求。UPRC 数据集中的图片大多存在模糊和失真等现象,所以算法在 UPRC 数据集上检测效果要差于在 UWD 数据集上的检测效果,图 7 显示了算法在两个数据集上损失函数变化,显而易见算法在 UPRC 数据集上损失较高。尽管如此,算法在 UPRC 上的 mAP 依然维持在较高的水平,达到了 71.75%,在所有模型中排名第二,仅次于 DSSD。由此可见 SG-Det 相比于其他算法更能适应不同的检测环境,具有更好的鲁棒性。

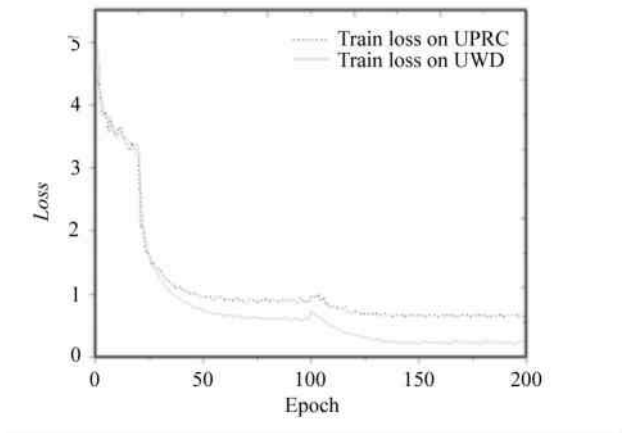


图 7 SG-Det 训练损失函数曲线图

Fig. 7 Chart of SG-Det training loss

水下设备拍摄的图像大多类似于 UPRC 数据集

中的图片,所以重点比较 SG-Det 在 UPRC 数据集上的表现。根据表 6 和图 8 的结果可以看出,本文提出的模型与 DSSD 比,虽然 mAP 较低,但 Parameters 减少到 4.91 M,仅为 DSSD 的 1/6,且速度提升近 2 倍,更适合部署在水下设备上进行实时的目标检测。SG-Det 与检测效果最好的轻量级模型 PeeleNet 相比,尽管 PeeleNet 在 UWD 数据集上的 mAP 与 SG-Det 几乎相同,但在 UPRC 数据集上,两者在速度相差不多的情况下,SG-Det 在平均精度和参数量方面均有更好的表现,可以看出本文提出的算法比 PeeleNet 更适合在水下进行目标检测任务。综上所述,本文的方法在保持检测模型足够轻巧的同时在精度和速度之间取得了平衡,更适合在水下进行实时目标检测。SG-Det 算法在两个数据集上的检测效果图如图 9 所示。

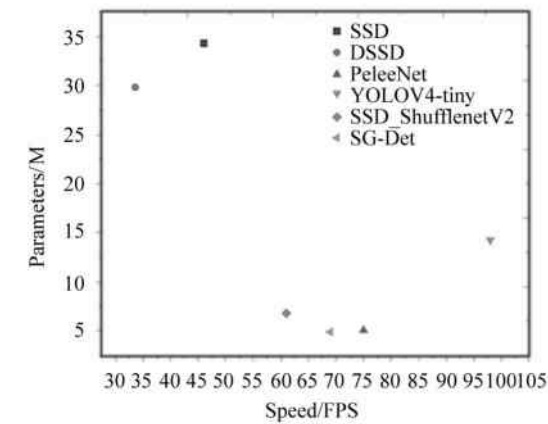


图 8 模型速度测试

Fig. 8 Model speed test

表 5 各算法在 UWD 上的检测结果比较

Tab. 5 Comparison of detection results of each algorithm on UWD

Methods	Backbone	Input size	<i>mAP</i> /%	Parameters/M	Speed/FPS
YOLOV4	CSPDarknet53	416×416	88.3	66.3	36
YOLOV4-tiny	CSPDarknet53Tiny	416×416	83.4	7.8	115
SSD	VGG16	300×300	80.6	36.9	48
DSSD	Residual-101	300×300	94.9	33.4	34
SSD-ShufflenetV2	ShufflenetV2	300×300	85.8	5.8	63
PeeleNet	PeeleNet	300×300	87.5	5.36	77
Ours	SGnet	300×300	87.3	5.43	67

表 6 各算法在 UPRC 上的检测结果比较

Tab. 6 Comparison of detection results of each algorithm on UPRC

Network	Holothurian	Echinus	Scallop	Starfish	<i>mAP</i> /%	FLOPs/M
SSD	0.62	0.79	0.59	0.77	69.25	40 485
DSSD	0.72	0.90	0.71	0.88	80.28	39 475
YoloV4-tiny	0.59	0.82	0.60	0.80	70.36	3 450
SSD-ShuffleNetV2	0.61	0.80	0.57	0.82	70.13	805
PeeleNet	0.62	0.82	0.58	0.83	70.25	508
Ours	0.65	0.81	0.63	0.78	71.75	483

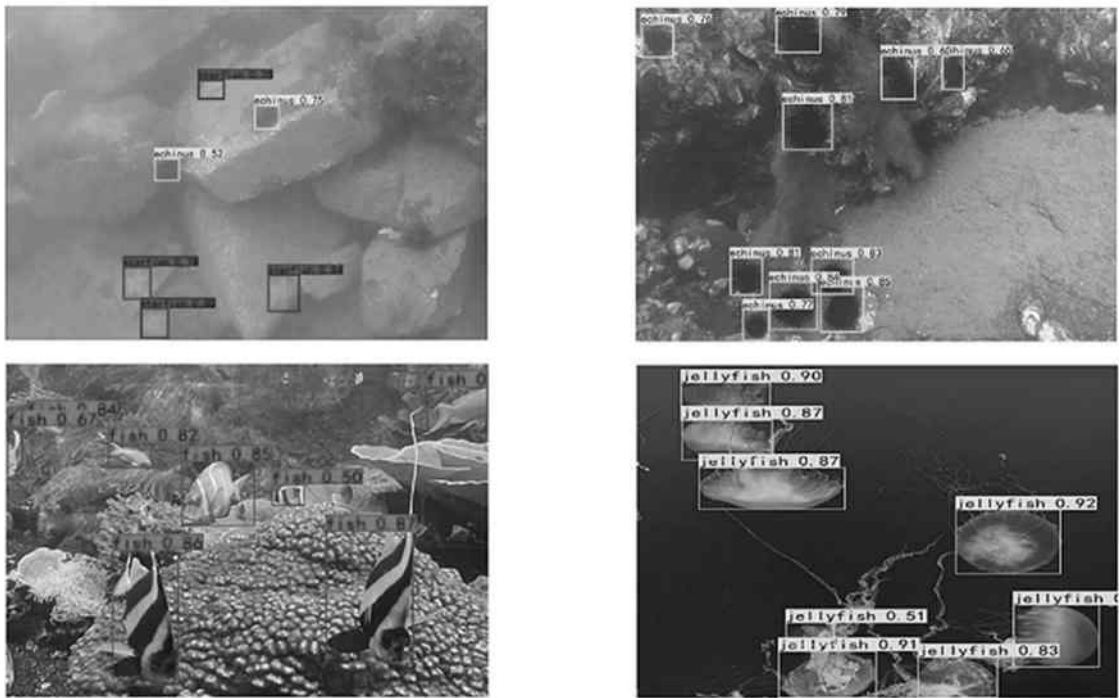


图 9 水下生物检测效果图

Fig. 9 Underwater biological detection renderings

4 结 论

SG-Det 使用 SGnet 作为特征提取网络,同时结合了注意力模块和跨尺度特征增强的思想,形成了

一种新的目标检测结构。SGnet 在保持特征提取能力的基础上大大减少了模型参数;通过注意力模块,将输入图像卷积得到的特征在通道上进行加权,突出有用信息,剔除背景等无用信息干扰;又通过多尺

度特征增强模块,将浅层和深层网络提取的特征融合复用,保证了特征利用的最大化,丰富了浅层特征的语义信息,提高了模型对小目标的检测能力。但改进后的模型仍然存在不足,对水下采集的低质量图像的检测结果还有待提高。因此,对算法进一步改进,将水下图像增强与目标检测相结合是下一步研究的重点和主要方向。

参考文献:

- [1] SHI P F, XU X W, NI J, et al. Underwater biological detection algorithm based on improved Faster-RCNN[J]. *Water*, 2021, 13(17): 2420.
- [2] PAN T S, HUANG H C, LEE J C, et al. Multi-scale ResNet for real-time underwater object detection[J]. *Signal Image and Video Processing*, 2021, 15(5): 941-9.
- [3] XU F Q, WANG H B, PENG J J, et al. Scale-aware feature pyramid architecture for marine object detection[J]. *Neural Computing & Applications*, 2021, 33(8): 3637-3653.
- [4] MATHIAS A, DHANALAKSHMI S, KUMAR R, et al. Deep neural network driven automated underwater object detection[J]. *Computers Materials & Continua*, 2022, 70(3): 5251-67.
- [5] ZHANG M H, XU S B, SONG W, et al. Lightweight underwater object detection based on YOLO v4 and multi-scale attentional feature fusion[J]. *Remote Sensing*, 2021, 22(13): 4706.
- [6] NASEER A, BARO E N, KHAN S D, et al. Automatic detection of nephrops norvegicus burrows from underwater imagery using deep learning[J]. *Computers Materials & Continua*, 2022, 70(3): 5321-44.
- [7] HAN K, WANG Y, TIAN Q, et al. GhostNet: more features from cheap operations[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13-19, 2020, Seattle, WA, USA. New York: IEEE, 2020: 1580-1589.
- [8] MA N, ZHANG X, ZHENG H T, et al. Shufflenet v2: Practical guidelines for efficient cnn architecture design[C]//Proceedings of the European Conference on Computer Vision (ECCV), September 8-14, 2018, Munich, Germany. Berlin, Heidelberg: Springer, 2018: 116-131.
- [9] LIU W, ANGUELOV D, ERHAN D, et al. SSD: singleshot-multiboxdetector[C]//Proceedings of the European Conference on Computer Vision, October 11-14, 2016, Amsterdam, USA. Berlin, Heidelberg: Springer, 2016: 21-37.
- [10] FU C, LIU W, RANGA A, et al. DSSD: deconvolutional single shot detector[EB/OL]. (2017-01-23) [2022-04-19]. <https://arxiv.org/abs/1701.06659>.
- [11] LI Z, ZHOU F. FSSD: feature fusion single shot multibox detector[EB/OL]. (2018-05-17) [2022-04-19]. <https://arxiv.org/abs/1712.00960>.
- [12] HOWARD A, ZHU M, CHEN B, et al. MobileNets: efficient convolutional neural networks for mobile vision applications[EB/OL]. (2017-04-17) [2022-04-19]. <https://arxiv.org/abs/1704.04861>.
- [13] SANDLER M, HOWARD A, ZHU M, et al. MobileNetV2: inverted residuals and linear bottlenecks[C]//MICCAI 2018: Medical Image Computing and Computer Assisted Intervention, September 11-14, Amsterdam, USA. Berlin, Heidelberg: Springer, 2018: 421-429.
- [14] XIANG Y Z, XIN Y Z, MENG X L, et al. Shufflenet: an extremely efficient convolutional neural network for mobile devices[C]//2018 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 18-22, 2018, Salt Lake City, UT, USA. New York: IEEE, 2018: 6848-6856.
- [15] CHOLLET F. Xception: deep learning with depthwise separable convolutions[C]//2017 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 17355762.

作者简介:

周华平 (1979—),女,教授,博士生导师,主要研究方向为深度学习、大数据。