

# 基于改进SSD的水下光学图像感兴趣目标检测算法研究

李宝奇 黄海宁\* 刘纪元 刘正君 韦琳哲

(中国科学院声学研究所 北京 100190)

(中国科学院先进水下信息技术重点实验室 北京 100190)

**摘要:** 针对轻量化目标模型SSD-MV2对水下光学图像感兴趣目标检测精度低的问题, 该文提出一种通道可选择的轻量化特征提取模块(SEB)和一种卷积核可变形、通道可选择的特征提取模块(SDB)。与此同时, 利用SEB模块和SDB模块分别重新设计了SSD-MV2的基础网络和附加特征提取网络, 记作SSD-MV2SDB, 并为其选择了合理的基础网络扩张系数和附加特征提取网络SDB模块数量。在水下图像感兴趣目标检测数据集UOI-DET上, SSD-MV2SDB比SSD-MV2检测精度提高3.04%。实验结果表明, SSD-MV2SDB适用于水下图像感兴趣目标检测任务。

**关键词:** 水下光学图像感兴趣目标检测; SSD; MobileNet V2; 可变形卷积; 通道可选择

中图分类号: TN911.73; TP391

文献标识码: A

文章编号: 1009-5896(2022)10-3372-07

DOI: [10.11999/JEIT210761](https://doi.org/10.11999/JEIT210761)

## Underwater Optical Image Interested Object Detection Model Based on Improved SSD

LI Baoqi HUANG Haining LIU Jiyuan LIU Zhengjun

WEI Linzhe

(*Institute of Acoustics, Chinese Academy of Sciences, Beijing 100190, China*)

(*Key Laboratory of Science and Technology on Advanced Underwater Acoustic Signal Processing, Chinese Academy of Sciences, Beijing 100190, China*)

**Abstract:** In order to solve the problem of low detection accuracy of SSD-MV2, a Selective and Efficient Block (SEB) and a Selective and Deformable Block (SDB) are proposed. At the same time, the basic network and additional feature extraction network of SSD-MV2 are redesigned by using SEB and SDB, which is named SSD-MV2SDB, and a set of reasonable expansion coefficient of basic network and number of SDB in additional feature extraction network are selected for SSD-MV2SDB. On UOI-DET, mAP of SSD-MV2SDB is 3.04% higher than that of SSD-MV2. The experimental results show that SSD-MV2SDB is suitable for underwater optical image interested object detection task.

**Key words:** Underwater optical image interested object detection; Single Shot Detection (SSD); MobileNet V2; Deformable convolution; Channel selectable

### 1 引言

我国海洋装备体系正朝着无人化、智能化的方向快速发展。光学成像具有信息丰富、分辨率高和成本低等优点, 是水下近距离环境感知的主要手

段, 广泛用于海底物种监测、无人潜航器导航、定位、避障和目标识别等多种场合<sup>[1,2]</sup>。然而, 水下目标在不同观察距离和视角下通常会呈现不同的特征, 尤其是水下非刚性目标, 增加了水下光学图像目标检测识别的难度。

基于传统特征的水下目标检测与识别方法通过特征描述符(颜色特征、形状特征、纹理特征等)刻画水下目标物体<sup>[3]</sup>, 具有方法简单、实时性好的优点, 但易受目标物体尺寸变化、旋转、被遮挡和拍摄角度变化的影响。随着深度学习<sup>[4-6]</sup>技术的兴起, 研究学者利用深度卷积神经网络(Convolutional Neural Networks, CNN)提高水下光学图像目标分

收稿日期: 2021-07-30; 改回日期: 2021-11-15; 网络出版: 2021-11-19

\*通信作者: 黄海宁 hhn@mail.ioa.ac.cn

基金项目: 国家自然科学基金(11904386), 国家基础科研计划重大项目(JCKY2016206A003), 中国科学院青年创新促进会(2019023)

Foundation Items: The National Natural Science Foundation of China (11904386), The Major Projects of the National Basic Scientific Research Plan (JCKY2016206A003), The Youth Innovation Promotion Association of Chinese Academy of Sciences (2019023)

类的准确率<sup>[7,8]</sup>。不过,分类任务是对图像整体的描述,无法确认目标在图像中的准确位置。近年来,基于CNN的目标检测模型在一个框架内实现了目标检测和识别的统一,目标检测模型可分为基于候选区域和基于回归两类,其中基于候选区域的模型主要代表有R-CNN<sup>[9]</sup>,Fast R-CNN<sup>[10]</sup>和Faster R-CNN<sup>[11]</sup>等,基于回归思想的模型主要代表有YOLO (You Only Look Once)<sup>[12]</sup>和SSD(Single Shot Detection)<sup>[13]</sup>,其中SSD模型较好地兼顾了检测精度和检测速度。为了缩短SSD的检测时间,Iandola等人<sup>[14]</sup>提出了轻量化网络SqueezeNet。SqueezeNet主要是利用 $1 \times 1$ 的卷积层对输入特征降维来降低模型的参数和计算量,同时也利用Inception<sup>[15]</sup>结构提高FireModule的特征提取能力。Howard等人<sup>[16]</sup>提出了轻量化的卷积神经网络MobileNet V1。MobileNet V1用深度可分离卷积 (Depthwise Separable Convolution, DSC) 替换标准卷积来减少模型的参数和计算量,计算成本约为标准卷积的 $1/9$ 。基于MobileNet V1的SSD-MV1模型目标检测速度也得到了极大的提升,但检测精度略有降低。为此,Sandler等人<sup>[17]</sup>提出了MobileNet V1的改进版本MobileNet V2。MobileNet V2在深度可分离卷积的基础上引入了ResNet中的shortcut connection结构,并设计了新的特征提取模块逆残差模块 (Inverted Residual Block, IRB)。新模块将原来的先“压缩”后“扩张”调整为先“扩张”后“压缩”。IRB模块利用扩张压缩的计算方式进一步减少计算量,引入了ResNet中的shortcut连接提高模块的特征提取能力。基于MobileNet V2的SSD-MV2较好地兼顾了目标检测精度和检测速度,但对存在一定形变的水下光学图像目标检测精度有待提高。

为了提高卷积神经网络的特征提取能力,Hu等人<sup>[18]</sup>提出了压缩激励(Squeeze and Excitation, SE)特征提取模块。SE模块首先对卷积得到的特征进行压缩操作,得到全局特征,然后对全局特征进行激励操作,得到不同特征的权重,最后乘以对应通道的特征得到最终特征。本质上,SE模块是在特征维度上做选择,这种注意力机制让模型可以更加关注信息量最大的特征,而抑制那些不重要的特征。另外,Dai等人<sup>[19]</sup>提出了可变形卷积网络(Deformable Convolutional Networks, DCN)。DCN加入1个偏移量使其可以自由变形,从而提升卷积神经网络的几何变换建模能力,提高对不规则物体、非刚性目标及复杂环境下目标的检测效果。

为此,本文针对SSD-MV2及其改进算法存在的不足提出了两种特征提取模块:一种是通道可选

择的轻量化特征提取模块(Selective and Efficient Block, SEB);一种是通道可选择和卷积核可变形的特征提取模块(Selective and Deformable Block, SDB),SEB模块在IRB模块基础上引入了通道可选择组件,SDB模块在IRB模块基础上引入了通道可选择组件和卷积核可变形组件。接着,利用SEB模块和SDB模块分别重新设计了目标检测模型SSD的基础网络和附加特征提取网络,新的基础网络和附加特征提取网络对水下光学图像目标具有更好的适用性。最后,在SSD框架内实现对水下光学图像感兴趣目标准确的检测识别。

## 2 基于改进SSD的水下光学图像感兴趣目标检测算法

水下目标具有形态多样的特点,本文通过改进SSD-MV2基础网络和附加特征提取网络特征提取模块来提高水下光学图像目标检测精度。

### 2.1 SEB模块和SDB模块

为了提高目标检测模型对水下光学图像目标的检测精度,本文在IRB模块的基础上提出了两种特征提取模块:SEB模块和SDB模块,如图1所示。SEB在IRB模块<sup>[17]</sup>内引入了通道可选择组件,SDB在IRB模块<sup>[17]</sup>内引入了通道可选择组件和卷积核可变形组件,其中通道可选择组件并不显著增加计算成本<sup>[18]</sup>,卷积核可变形组件计算成本增加显著<sup>[19]</sup>。事实上,通道可选择组件和卷积核可变形组件是相互独立的,而且都是提升网络特征提取能力的有效手段:通道可选择组件利用注意力机制选择最有价值的通道数据来提高网络的性能,卷积核可变形组件利用卷积核形变准确捕捉目标特征来提高网络的性能。鉴于SEB是SDB的简化版本,本文仅对SDB模块进行说明和分析,SDB模块沿用反残差网络结构,即先对通道采取先“扩张”后“压缩”的策略,并由扩张层、可变形卷积组件、通道可选择组件和压缩层组成,其中扩张层负责输入特征通道扩张;可变形卷积组件通过学习改变卷积核形态提取水下感兴趣目标特征;通道可选择组件通过学习权重选择包含重要信息的通道;压缩层负责将特征通道压缩成与输入特征的一致数量。

对于一个任意的输入特征 $\mathbf{D} \in \Phi^{H \times H \times M}$ ,其中 $H \times H$ 为输入特征的尺寸, $M$ 为输入特征的通道数。输入特征 $\mathbf{D}$ 进入SDB模块的两个支路网络:下侧支路负责水下感兴趣目标特征提取和选择;上侧支路保持输入特征 $\mathbf{D}$ 不变,并最后与下侧支路网络的输出特征相加。对于下侧支路网络,输入特征 $\mathbf{D}$ 首先经过扩张层,其输出特征的数学表达为

$$\mathbf{D}_{\text{ex}} = F_{\text{ex}}(\mathbf{D}), \mathbf{D} \in \Phi^{H \times H \times M} \quad (1)$$

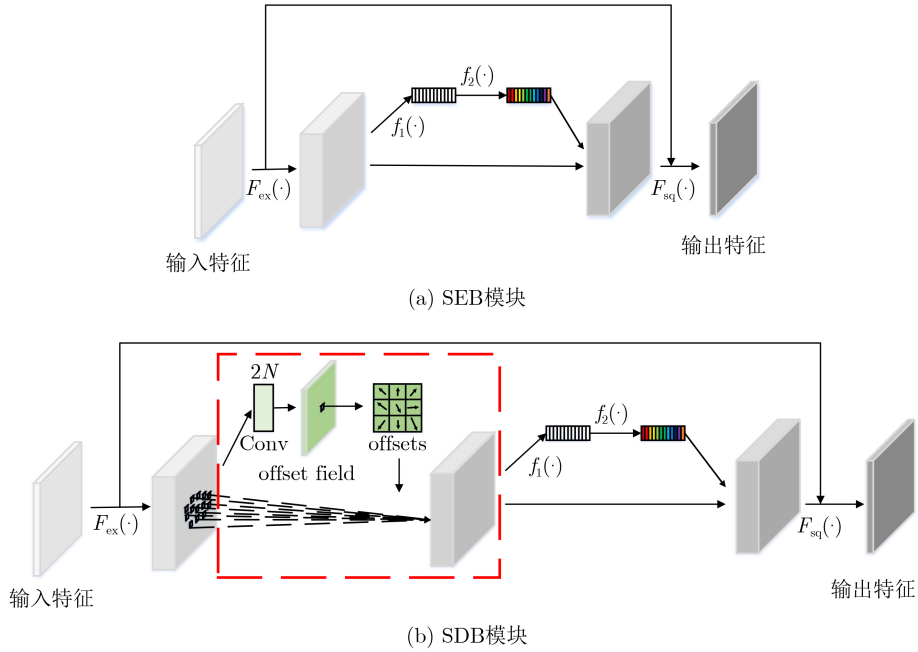


图1 SEB模块和SDB模块

其中,  $D$ 为原始输入特征,  $D_{\text{ex}}$ 为经过扩张层后的特征, 扩张层的卷积核尺寸为 $1 \times 1$ , 卷积核的数量为输入特征通道的 $k$ 倍, 即 $k \times M$ 。

随后, 输出特征 $D_{\text{ex}}$ 送入可变形卷积组件, 其输出特征的数学表达式为

$$D_{\text{de}}(p_0) = \sum_{p_n \in R} W(p_n) \cdot D_{\text{ex}}(p_0 + p_n + \Delta p_n),$$

$$D_{\text{ex}} \in \Phi^{H \times H \times (k \times M)} \quad (2)$$

其中,  $D_{\text{de}}(p_0)$ 表示以 $p_0$ 为中心的可变形卷积组件输出特征,  $W(p_n)$ 表示卷积核权重,  $p_n$ 是对 $R$ 中所列位置的枚举,  $R = \{(-1, -1), (-1, 0), \dots, (0, 1), (1, 1)\}$ ,  $\Delta p_n$ 是 $p_n$ 的偏移量。在实际计算过程中, 并不是真正改变卷积核的形状, 而是对输入特征重新整合, 变相地实现卷积核的形变。也就是说, 在输入图像的偏置特征图上利用标准卷积求得最终的输出特征。

通常 $\Delta p_n$ 为小数, 非整数的坐标( $p = p_0 + p_n + \Delta p_n$ )无法在图像上使用的, 具体实现通过双线性插值寻找距离这个对应坐标最近的4个特征点来计算该点的值

$$D_{\text{ex}}(p) = \sum_q G(q, p) \cdot D_{\text{ex}}(q) \quad (3)$$

其中,  $q$ 表示距离 $p$ 最近的4个特征点,  $G(q, p)$ 表示对应的系数, 由偏移量(offsets)计算得到。可变形卷积组件为了保留输入图像的边沿特征, 需要对其进行零填充(zero\_padding)。理论上, 对一个填充后不小于卷积核尺寸的特征图, 可变形卷积组件依然能够计算。只不过, 当填充后的特征图中有效特

征点较少时, 双线性插值重构的特征图并不会产生更多的有用信息。

然后, 输出特征 $D_{\text{de}}$ 送入通道选择组件, 其输出特征的数学表达式为

$$D_{\text{se}} = s \cdot D_{\text{de}} \quad (4)$$

$$s = \text{softmax}(f_2(f_1(P_g(D_{\text{de}})))) \quad (5)$$

其中,  $D_{\text{se}}$ 为通道选择后的通道特征,  $s$ 为通道的选择系数,  $s \in \Phi^{1 \times (k \times M)}$ ;  $P_g()$ 为全局池化函数, 输出特征维度为 $\Phi^{1 \times (k \times M)}$ ;  $f_1$ 为第1全连接层, 输出特征维度为 $\Phi^{1 \times d}$ , 其中 $d = 16$ ;  $f_2$ 为第2全连接层, 输出特征维度为 $\Phi^{1 \times (k \times M)}$ ;  $\text{softmax}()$ 为归一化指数函数。

接着, 对 $D_{\text{se}}$ 进行通道压缩, 数学表达式为

$$D' = F_{\text{sq}}(D_{\text{se}}), D_{\text{se}} \in \Phi^{H \times H \times (k \times M)} \quad (6)$$

其中,  $D'$ 为通道压缩后的特征。

通过上面的计算, 最后可以得到SDB模块的输出特征数学表达式为

$$D = D + D' \quad (7)$$

其中,  $D$ 为SDB模块的输出特征,  $D \in \Phi^{H \times H \times M}$ , 特征图尺寸为 $H \times H$ , 通道数为 $M$ 。

## 2.2 SSD\_MV2SDB模型

SEB在IRB的基础上引入了通道可选择组件, 通道可选择组件主要由两个全连接层组成, 计算成本相对较低。SDB在IRB的基础上引入了通道可选择组件和可变形卷积组件, 可变形卷积组件新增双线性插值计算和两个卷积层计算成本相对较高。为



了兼顾SSD-MV2SDB模型的性能，基础网络采用了轻量化的SEB模块，附加特征提取网络采用目标适应性更强的SDB模块。改进的SSD水下目标检测模型结构如图2所示，记作SSD-MV2SDB，包括基础网络、特征提取网络、候选框生成和卷积预测4个部分。基础网络MobileNet V2SEB与MobileNet V2的网络结构保持一致，利用SEB模块替换IRB模块实现。特征提取网络一共提取6个尺度的特征，MobileNet V2SEB中的第14层Conv14和第19层Conv19作为特征提取网络的第1特征层和第2特征层，输入特征图尺寸为 $38 \times 38$ 和 $19 \times 19$ ，输出特征图尺寸为 $19 \times 19$ 和 $10 \times 10$ ；Conv20\_1, Conv20\_2, Conv20\_3和Conv20\_4作为特征提取网络的第3、第4、第5和第6尺度特征层，输入特征图尺寸为 $10 \times 10$ ,  $5 \times 5$ ,  $3 \times 3$ 和 $2 \times 2$ ，输出特征图尺寸为 $5 \times 5$ ,  $3 \times 3$ ,  $2 \times 2$ 和 $1 \times 1$ 。Conv20\_1, Conv20\_2, Conv20\_3和Conv20\_4 4个卷积层也称作附加特征提取网络，与基础网络中的Conv14和Conv19两个尺度的特征层相比，附加特征提取网络中的特征区分度更高，对水下感性目标的检测识别更有帮助。因此，本文利用SDB模块替换附加特征提取网络中的IRB模块，但考虑最后1个特征提取层Conv20\_4的输入特征图尺寸为 $2 \times 2$ ，有效特征点少，因此Conv20\_4依然采用SEB模块，即Conv20\_1, Conv20\_2和Conv20\_3采用SDB模块，Conv20\_4采用SEB模块。候选框生成部分根据预先定义的尺度(scales)和纵横比(aspect ratios)从上述6个尺度的特征层中提取数量和大小不同的候选框；卷积预测部分则是对候选框内目标的类型和位置进行判断，并利用非极大值抑制算法对候选框内目标进行优化。

SSD-MV2SDB基础网络和附加特征提取网络中模块参数选取参考了SSD-MV2模块的设计原

则：基础网络中模块采用扩张系数大于1、压缩系数小于1(先扩张后压缩)的计算方式，附加特征提取网络模块采用扩张系数小于1，压缩系数大于1(先压缩后扩张)的计算方式，为了便于研究依然称SEB模块和SDB模块的第1个卷积层和最后1个卷积层为扩张层和压缩层。附加特征提取网络模块采用扩张系数小于1、压缩系数大于1的计算方式主要是为了保持与SSD-MV2具有相同的通道数，另外降低网络的计算成本。具体地，本文SSD-MV2SDB附加特征提取网络Conv20\_1, Conv20\_2和Conv20\_3的初始通道数分别为1280, 512和256，扩张层的卷积核尺寸为 $1 \times 1$ ，扩张系数分别为0.2, 0.25和0.5，即扩张后的通道数分别为256, 128和128，可变形组件和通道可选择组件并不影响输入特征的尺寸和通道数，压缩层的卷积核尺寸为 $1 \times 1$ ，压缩系数分别为5, 4和2，最终的输出通道数分别为1280, 512和256。

### 3 仿真实验

为了验证SSD-MV2SDB模型的有效性以及基础网络SEB模块扩张系数和附加特征提取网络SDB模块数量对模型性能的影响，实验以mAP，参数大小和平均检测时间作为定量评价指标。设计实验1，以SSD-MV2(基础网络和附加特征提取网络的特征提取模块为IRB)为参考，比较分析不同轻量化目标检测模型之间的性能差异。设计实验2，以基础网络SEB模块扩张系数为研究对象，比较分析不同扩张系数对SSD-MV2SDB模型性能的影响。设计实验3，以附加特征提取网络SDB模块数量为研究对象，比较分析不同数量的SDB对SSD-MV2SDB模型性能的影响。

#### 3.1 实验数据集

为了更好地检验SSD-MV2SDB模型对水下光

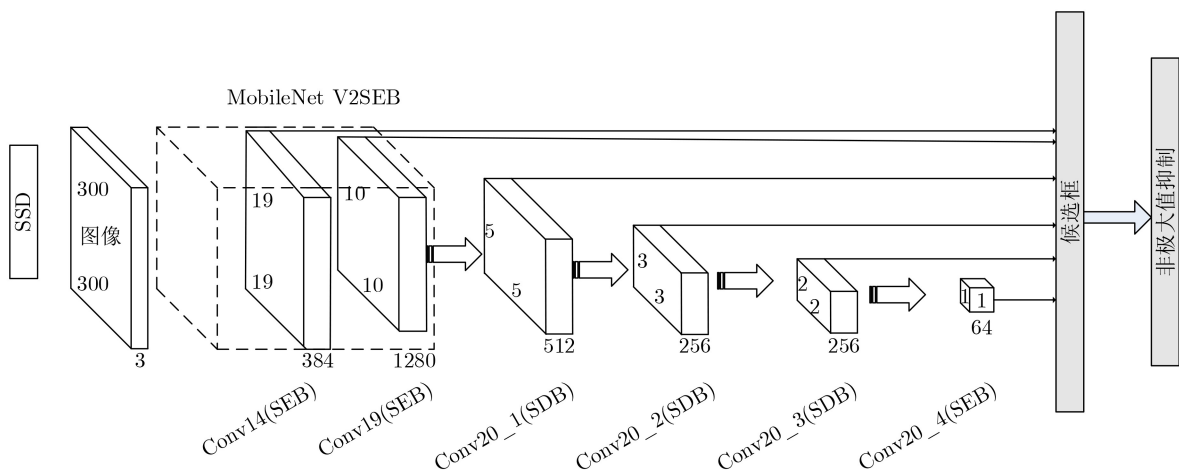


图2 SSD-MV2SDB目标检测模型

学图像感兴趣目标的检测识别性能, 本文建立了一个水下光学图像感兴趣目标检测数据集(Underwater Optical Interest DEtection Target, UOI-DET)。UOI-DET共计1135幅图像, 包括方框、渔网、蛙人、UUV和球体5类目标, 其中1035幅用于模型训练, 100幅图像用于模型测试, 如表1所示。图像采集地点为千岛湖, 采集时间为2020年11月, 目标与水下图像采集设备距离在3~10 m。

### 3.2 实验1: 目标检测算法的性能比较

本实验比较分析SSD-MV2, SSD-MV2SEB, SSD-MV2IRBD与本文目标检测方法SSD-MV2SDB在数据集UOI-DET上的性能差异。SSD-MV2的基础网络的特征提取模块为IRB, 附加特征提取网络的特征提取模块为IRB, 其中基础网络IRB模块的扩张系数等于4; SSD-MV2SEB的基础网络的特征提取模块为SEB, 附加特征提取网络的特征提取模块为SEB, 其中基础网络SEB模块的扩张系数等于4; SSD-MV2IRBD的基础网络的特征提取模块为SEB, 附加特征提取网络为IRBD模块, 其中基础网络SEB模块的扩张系数等于4, IRBD表示在IRB模块中引入了可变形卷积模块<sup>[19]</sup>; SSD-MV2SDB的基础网络的特征提取模块为SEB, 附加特征提取网络Conv20\_1, Conv20\_2和Conv20\_3为SDB模块, Conv20\_4为SEB模块。分别记录目标检测模型在迭代500次时对UOI-DET测试数据集的mAP、参数大小和平均检测时间, 实验结果如表2所示。

从表2可以发现, SSD-MV2SDB的检测精度比SSD-MV2, SSD-MV2SEB和SSD-MV2IRBD分别高3.04%, 2.19%和1.31%, 模型参数分别多4.7 MB,

表1 水下图像目标检测数据集组成(幅)

目标	训练	测试
方框	203	19
渔网	221	13
蛙人	214	26
UUV	194	22
球体	203	20
总计	1035	100

3.9 MB和0.1 MB, 运算时间分别高6.66 ms, 3.85 ms和0.34 ms。通道可选择组件和卷积核可变形组件对提高水下光学图像感兴趣目标的检测识别精度都有帮助, 不过均增加了一定的计算成本, 并且可变形组件的计算成本更高。从检测识别精度角度考虑, SSD-MV2SDB更适合基于水下光学图像感兴趣目标检测任务。

为了更直观地说明SSD-MV2SDB对水下光学图像感兴趣目标的检测效果, 利用SSD-MV2SDB模型对3种类型的水下感兴趣目标进行检测, 结果如图3所示。从图3可以看出, SSD-MV2SDB算法对水下感兴趣目标能实现准确检测识别。

### 3.3 实验2: 基础网络SEB模块扩张系数对SSD-MV2SDB性能的影响

本实验比较SSD-MV2SDB基础网络SEB模块不同扩张系数对SSD-MV2SDB性能的影响。SSD-MV2SDB附加特征提取网络Conv20\_1, Conv20\_2和Conv20\_3为SDB模块, Conv20\_4为SEB模块, 基础网络的SEB模块扩张系数分别为2, 4, 6和8。记录模型迭代500次时模型对UOI-DET测试数据集的mAP、平均检测时间和参数大小, 实验结果如表3所示。

从表3可以看出, SSD-MV2SDB的检测精度随基础网络SEB模块扩张系数的增加逐渐增加, 当扩张系数等于8时, SSD-MV2SDB的检测精度已经达到97.76%。与此同时, SSD-MV2SDB模型参数随基础网络扩张系数增加也不断增大, 当扩张系数等于8时, SSD-MV2SDB的模型参数已经达到20.4 MB。事实上, SSD-MV2SDB的检测时间并没有随扩张系数的增加有明显的变化, 主要是由于SEB模块沿用了IRB模块的深度可分离计算方式。增大扩张系数能提高SSD-MV2SDB的检测精度, 但模型参数的增加也比较明显, 扩张系数为8的SSD-MV2SDB比扩张系数为2的SSD-MV2SDB参数增加8.3 MB。另外, 与扩张系数等于4时的SSD-MV2SDB相比, 扩张系数等于6或8的SSD-MV2SDB的检测精度并没有明显增加, 仅增加0.05%和0.48%。综合考虑检测精度、参数大小和运算时间3个评价指标, 扩张系数等于4的SSD-MV2SDB效果最好。

表2 目标检测模型性能比较

模型	基础网络	附加特征提取网络	检测精度(%)	参数大小(MB)	检测时间(ms)
SSD-MV2	IRB	IRB	94.24	10.2	7.20
SSD-MV2SEB	SEB	SEB	95.09	11.0	10.01
SSD-MV2IRBD	SEB	IRBD	95.97	14.8	13.52
SSD-MV2SDB	SEB	SDB	97.28	14.9	13.86

### 3.4 实验3：附加特征提取网络SDB模块数量对SSD-MV2SDB性能的影响

本实验比较附加特征提取网络中SDB数量对SSD-MV2SDB性能的影响。基础网络SEB模块扩张系数等于4，附加特征提取网络中SDB模块的数量分别为0, 1, 2和3，其中0表示附加特征提取网络均为SEB模块的SSD-MV2SEB模型，1表示附加特征提取网络仅Conv20\_1为SDB模块，2表示附加特征提取网络Conv20\_1和Conv20\_2为SDB模块，3表示附加特征提取网络Conv20\_1, Conv20\_2和Conv20\_3为SDB模块。记录模型迭代500次时对UOI-DET测试数据集的mAP、平均检测时间和参数大小，实验结果如表4所示。

从表4可以看出，与附加特征提取网络均为SEB模块的SSD-MV2SEB相比，SSD-MV2SDB随附加特征提取网络SDB模块数量的增加(1, 2和3)检测精度逐渐增大，分别增加0.99%, 2.00%和2.19%，可见结合了通道可选择组件和卷积核可变形组件的SDB模块对水下感兴趣目标具有很好的适用性。不过，模型参数和检测时间也存在一定的增加，主要是由于SDB模块中的可变形卷积组件采用标准卷积计算方式。考虑模型的计算成本，本文仅在附加特征提取网络中使用SDB模块。另外，附加特征提取网络中SDB模块的数量等于3。

### 3.5 讨论

实验从mAP、平均检测时间和参数大小3个方面比较了本文水下光学图像感兴趣目标检测方法SSD-

MV2SDB与经典轻量化目标检测算法(SSD-MV2)和最新算法(SSD-MV2SEB和SSD-MV2IRBD)性能上的差异，也进一步分析了基础网络的扩张系数和附加特征提取网络SDB模块数量的选取如何影响SSD-MV2SDB的性能，同时也验证了SSD-MV2SDB模型对水下光学图像感兴趣目标的适用性。

SSD-MV2较好地兼顾了目标检测精度和检测速度，但对存在一定形变的水下光学图像目标检测精度有待提高。SSD-MV2SEB在基础网络和附加特征提取网络中利用SEB模块替换IRB模块，在不显著增加计算成本的前提下，通过通道可选择组件提高了模型对水下感兴趣目标的检测识别能力。SSD-MV2IRBD在SSD-MV2SEB的基础上，在附加特征提取网络中利用IRBD替换SEB模块，利用可变形卷积组件提高了模型对水下感兴趣目标的检测识别能力。可以确定，单独的通道可选择组件和单独的卷积核可变形组件对提高水下感兴趣目标的检测识别能力均是有帮助的。SSD-MV2SDB在SSD-MV2SEB基础上，在附加特征提取网络中利用SDB模块替换SEB模块，通过结合通道可选择组件和卷积核可变形组件进一步提高了模型对水下感兴趣目标的检测识别能力。融合了通道可选择组件和卷积核可变形组件的SDB模块对水下感兴趣目标具有更好的特征提取能力。但为了平衡目标检测模型的检测精度与计算成本，本文仅附加特征提取网络前3个特征层使用了SDB模块，而基础网络和附加特征提取网络的最后1个卷积层使用SEB模块。

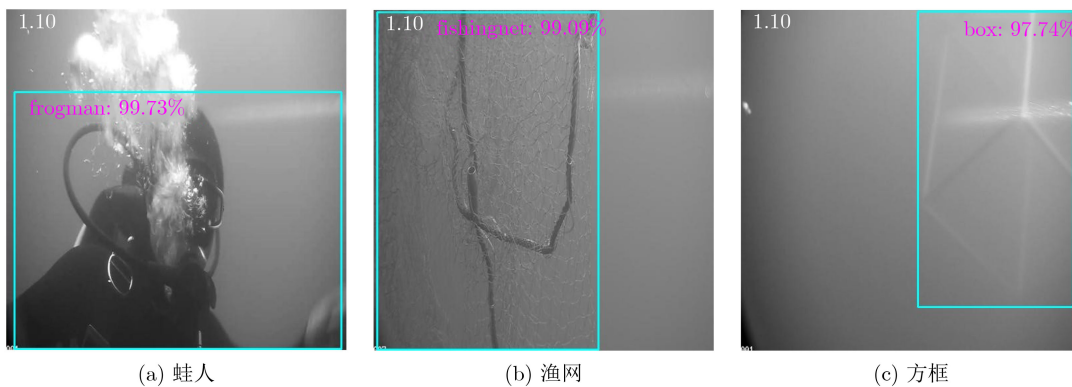


图3 SSD-MV2SDB对水下光学图像感兴趣目标的检测效果图

表3 基础网络扩张系数对SSD-MV2SDB性能的影响

扩张系数	检测精度(%)	参数大小(MB)	检测时间(ms)
2	95.03	12.1	13.66
4	97.28	14.9	13.86
6	97.33	17.7	13.90
8	97.76	20.4	14.12

表4 附加特征提取网络SDB模块数量对SSD-MV2SDB性能的影响

模块数量	检测精度(%)	参数大小(MB)	检测时间(ms)
0	95.09	11.0	10.01
1	96.08	13.5	11.11
2	97.09	14.2	12.53
3	97.28	14.9	13.86



## 4 结束语

水下光学图像感兴趣目标检测具有重要的理论研究和实际应用价值。本文提出了SEB和SDB两种特征提取模块，并利用SEB和SDB重新设计了SSD的基础网络和附加特征提取网络，经理论分析和仿真实验证明了改进SSD模型对水下光学图像感兴趣目标的有效性。

对于水下光学图像感兴趣目标检测任务，下一步的研究重点包括：(1) 研究适合捕获水下感兴趣目标特征的附加特征提取网络；(2) 研究更适合水下感兴趣目标的特征提取模块。

## 参考文献

- [1] YE H C H, LIN Chuhan, KANG Liwei, *et al.* Lightweight deep neural network for joint learning of underwater object detection and color conversion[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2021, 6: 1–15. doi: [10.1109/TNNLS.2021.3072414](https://doi.org/10.1109/TNNLS.2021.3072414).
  - [2] HMUE P M and PUMRIN S. Image enhancement and quality assessment methods in turbid water: A review article[C]. *IEEE International Conference on Consumer Electronics*, Bangkok, Thailand, 2019: 59–63.
  - [3] CHEN Bin, LI Rong, BAI Wanjian, *et al.* Research on recognition method of optical detection image of underwater robot for submarine cable[C]. *2019 IEEE 3rd Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)*, Chongqing, China, 2019: 1973–1976.
  - [4] LECUN Y, BENGIO Y, and HINTON G. Deep learning[J]. *Nature*, 2015, 521(7553): 436–444. doi: [10.1038/nature14539](https://doi.org/10.1038/nature14539).
  - [5] ZHANG J X, YORDANOV B, GAUNT A, *et al.* A deep learning model for predicting next-generation sequencing depth from DNA sequence[J]. *Nature Communications*, 2021, 12: 4387. doi: [10.1038/s41467-021-24497-8](https://doi.org/10.1038/s41467-021-24497-8).
  - [6] WANG Shiqiang. Efficient deep learning[J]. *Nature Computational Science*, 2021, 1(3): 181–182. doi: [10.1038/s43588-021-00042-x](https://doi.org/10.1038/s43588-021-00042-x).
  - [7] LAGEMANN C, LAGEMANN K, MUKHERJEE S, *et al.* Deep recurrent optical flow learning for particle image velocimetry data[J]. *Nature Machine Intelligence*, 2021, 3(7): 641–651. doi: [10.1038/s42256-021-00369-0](https://doi.org/10.1038/s42256-021-00369-0).
  - [8] LI Sichun, JIN Xin, YAO Sibing, *et al.* Underwater small target recognition based on convolutional neural network[C]. *Global Oceans 2020: Singapore – U. S. Gulf Coast*, Biloxi, USA, 2020: 1–7.
  - [9] GIRSHICK R, DONAHUE J, DARRELL T, *et al.* Rich feature hierarchies for accurate object detection and semantic segmentation[C]. *The IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, USA, 2014: 580–587.
  - [10] GIRSHICK R. Fast R-CNN[C]. *IEEE International Conference on Computer Vision (ICCV)*, Santiago, Chile, 2015: 1440–1448.
  - [11] REN Shaoqing, HE Kaiming, 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. doi: [10.1109/TPAMI.2016.2577031](https://doi.org/10.1109/TPAMI.2016.2577031).
  - [12] 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 (CVPR)*, Las Vegas, USA, 2016: 779–788.
  - [13] SRITHAR S, PRIYADHARSINI M, SHARMILA F M, *et al.* Yolov3 Supervised machine learning framework for real-time object detection and localization[J]. *Journal of Physics: Conference Series*, 2021, 1916: 012032. doi: [10.1088/1742-6596/1916/1/012032](https://doi.org/10.1088/1742-6596/1916/1/012032).
  - [14] IANDOLA F N, MOSKEWICZ N W, ASHRAF K, *et al.* SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <1 MB model size[EB/OL]. <https://arxiv.org/abs/1602.07360v1>, 2016.
  - [15] SZEGEDY C, LIU Wei, JIA Yangqing, *et al.* Going deeper with convolutions[C]. *IEEE Conference on Computer Vision and Pattern Recognition*, Boston, USA, 2015: 1–9.
  - [16] HOWARD A G, ZHU Menglong, CHEN Bo, *et al.* MobileNets: Efficient convolutional neural networks for mobile vision applications[EB/OL]. <https://arxiv.org/abs/1704.04861>, 2017.
  - [17] SANDLER M, HOWARD A, ZHU Menglong, *et al.* MobileNetV2: Inverted residuals and linear bottlenecks[C]. *The IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, USA, 2018: 4510–4520.
  - [18] HU Jie, SHEN Li, ALBANIE S, *et al.* Squeeze-and-excitation networks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 42(8): 2011–2023. doi: [10.1109/TPAMI.2019.2913372](https://doi.org/10.1109/TPAMI.2019.2913372).
  - [19] DAI Jifeng, QI Haozhi, XIONG Yuwen, *et al.* Deformable convolutional networks[C]. *IEEE International Conference on Computer Vision*, Venice, Italy, 2017: 764–773.
- 李宝奇：男，特别研究助理，研究方向为水声信号处理、目标检测/识别和跟踪、深度学习理论。
- 黄海宁：男，研究员，研究方向为水声信号与信息处理、目标探测、水声通信与网络等。
- 刘纪元：男，研究员，研究方向为水声信号处理、数字信号处理和水声成像与图像处理等。
- 刘正君：女，助理研究员，研究方向为水声信号处理等。
- 韦琳哲：男，助理研究员，研究方向为水声信号处理等。

责任编辑：余蓉