

何佳琦,周思艺,唐晓萌,等.基于改进的轻量版SOLOv2红鳍东方鲀实例分割方法[J].华中农业大学学报,2023,42(3):71-79.
DOI:10.13300/j.cnki.hnlkxb.2023.03.009

基于改进的轻量版SOLOv2红鳍东方鲀实例分割方法

何佳琦¹,周思艺¹,唐晓萌¹,胡显辉¹,王魏¹,蔡克卫²

1.大连海洋大学信息工程学院,大连116023;2.大连民族大学机电工程学院,大连116650

摘要 为解决红鳍东方鲀养殖密度不均导致图像分割精度低和小目标分割效果差的问题,提出一种改进的轻量版SOLOv2实例分割方法。首先进行可变形卷积(deformable convolutional networks, DCN)网络结构的优化调整,通过在卷积核上增加偏移参数,调整卷积的感受野,使感受野与物体的实际形状更加贴近;再在残差模块最后一层引入无参数注意力机制SimAM,捕捉图像中更多的局部信息,获得不同尺度的目标特征,优化模型对小目标分割的性能。试验结果显示,改进后的轻量版SOLOv2模型较原有模型平均分割精度提高了3.7个百分点,对小目标的分割精度提升了1.4个百分点,同时加入DCN和SimAM注意力模块后,模型的分割精度提高到65.2%。结果表明,改进后的SOLOv2模型可以提高边界处的细节感知能力,强化模型对小目标鱼群特征的提取能力,可用于高密度场景下的精准实例分割,实现红鳍东方鲀鱼群目标精准像素级分割。

关键词 实例分割;分割精度;注意力机制;小目标;精准化养殖

中图分类号 TP391 **文献标识码** A **文章编号** 1000-2421(2023)03-0071-09

精准化养殖是促进海洋渔业转型升级、发展海洋经济、贯彻海洋强国战略的重要保障。鱼类精准实例分割是实现精准化养殖的重要前提^[1],是鱼类体长、体质量以及生长状态评估的核心基础。近年来深度学习技术^[2]的迅猛发展,使计算机视觉技术上升了一个新的台阶,与传统视觉方法相比,基于深度学习的计算机视觉方法可实现图像特征自动提取,无需人工干预,具有准确度更高、效率更快等优点^[3-4]。基于深度卷积神经网络(convolutional neural network, CNN)^[5-8]的实例分割技术是一种比目标检测技术更加精确的检测方法,不但可以获得待测目标的位置和类别,还能获得每个目标的位置和类别,对每个分割结果作不同的实例输出。这种方法可以有效解决图像中的噪声和不均匀性问题,目前已有许多学者将深度学习应用于各个领域中。He等^[9]提出Mask R-CNN,使用特征金字塔^[10](feature pyramid network, FPN)和ResNet结合作为特征提取器,实现了对每个区域进行扫描,将输出的前景或背景进行分类。但Mask R-CNN属于双阶段实例分割网络,对于图像检测和分割整体耗时较长。Bolya等^[11]提出YOLACT实例分割网络,减少了分割时间,但分

割精度较Mask R-CNN稍差。由于以上模型不能很好地平衡精度和检测速度,Wang等^[12]提出了兼顾速度和准确性的SOLOv2模型,直接按位置分割对象,动态分割图像中的每个实例,无需求助于边界框检测,降低了推理开销,既保证了检测速度,又提高了检测精度。吴逢斌等^[13]在SOLOv2中融合通道分离注意力机制,降低了漏检率,提高了模型检测效率。花玮等^[14]优化SOLOv2掩膜特征分支,引入影响因子优化损失函数,根据遮挡程度自适应调整损失函数的权重,减少遮挡问题对服装图像分割带来的影响。

SOLOv2因其检测速度快、精度高的特点已成功应用于以上工业领域,但并未见到应用于水产养殖领域的相关研究。水下养殖环境鱼群图像相似性高、水下背景干扰大,模型的分割效果仍有很大改进空间。因此,本研究基于SOLOv2框架,提出改进轻量版SOLOv2图像分割的方法,融合多尺度特征,设计满足目标多尺寸要求的特征提取主干网络,引入无参数注意力机制SimAM,设计专用实例分割器,建立适用于高密度场景下的精准实例分割模型,以期实现养殖鱼类像素级分割,为鱼类实例分割提供新途径。

收稿日期:2022-09-29

基金项目:设施渔业教育部重点实验室(大连海洋大学)开放课题(202314);辽宁省教育厅科学研究项目(JL202015)

何佳琦, E-mail:994751040@qq.com

通信作者:王魏, E-mail:ww_wangwei@dlou.edu.cn

1 材料与方法

1.1 数据采集

实例分割模型需要大量的图像进行训练。本研究在大连海洋大学信息工程学院的实践基地,以室温环境下通过人工培育的红鳍东方鲀(*Takifugu rubripes*)作为试验对象,利用固定的单目摄像头侧拍进行实时录像,采集鱼类视频样本,经过抽帧处理制作鱼群数据集,存储格式为JPG,图像的大小为1 080

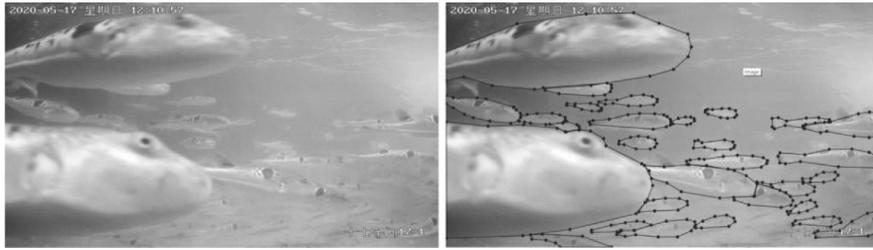


图1 数据集标注示意图

Fig.1 Dataset annotation legend

1.2 红鳍东方鲀实例分割模型设计

1)改进轻量版SOLOv2网络结构。SOLOv2官方发布的模型分为标准版SOLOv2模型和轻量版SOLOv2模型,轻量版SOLOv2模型是将标准版SOLOv2预测头中的卷积层减少为3个,是结构中网络层数最少、模型尺寸最小的模型。因此,选取轻量版SOLOv2模型为本研究的主体网络结构。

利用单目摄像头进行水下侧拍,采集鱼类图像样本。根据图像样本分析水下图像统计信息;结合可变形卷积(DCN)^[15]优化网络结构,利用多尺度融合方法设计感受野与多目标尺寸匹配的主干网络,提高目标特征提取能力,为分割器提供高质量目标特征。通过加入无参数注意力模块SimAM^[16]对样本关注的目标加入更多的注意力资源,通过学习全局的特征信息,可为红鳍东方鲀获取更多所需特征的细节信息,从而抑制其他无用信息。轻量版SOLOv2整体网络结构如图2所示,在实例分割时,小目标的特征提取能力尤为重要,通过加入SimAM模块,用有限的资源获取更关键的目标信息,优化模型对小目标分割的性能,提高模型分割的准确性^[17]。

图像进入主干网络ResNet50,通过在残差模块中加入可变形卷积(DCN)和无参数注意力模块SimAM(具体操作如图2红框内标记所示),进行特征提取,捕捉图像中更多的局部信息,提高特征提取质量,再通过特征金字塔对图片进行多尺度变化增强,能较好地让各个不同尺度的特征都具有较强的语义

像素 \times 1 920像素。随机选取200张图像作为数据集,图像包括鱼群在水下的各种真实活动情况。通过标注软件Labelme进行数据集的标注,标注后的文件保存为json格式,数据集的标注样例如图1所示。为了增加鱼群样本的多样性以及提高模型的分割性能,采用颜色扰动、随机剪裁、多尺度训练、随机翻转、图片归一化以及填充黑边等数据增强方法将数据集增广到1 000张。最后将数据集按照7:2:1的比例划分为训练集、测试集和验证集。

信息,最后输入到检测头中进行动态实例分割,通过位置预测分割对象。

2)基于无参数注意力模块SimAM的ResNet50骨干网络。养殖鱼类精准实例分割的首要任务是提取目标有效特征,特征提取通过主干网络完成。本研究基于水下图像数据,归纳养殖鱼类目标特点,利用注意力机制方法,设计具有与多种目标尺寸匹配感受野的实例分割器,增强特征提取能力,为分割模型提供准确特征。

SOLOv2的Backbone是残差网络ResNet50,有Conv Block和Identity Block 2个基本的模块,结构图如图3所示。

改进后的红鳍东方鲀实例分割模型是在轻量版SOLOv2模型的主干网络中加入无参数注意力机制SimAM。通过在2个残差模块的最后1层卷积中加入SimAM注意力模块,lambda取0.000 1,防止除零,强化模型对小目标鱼群特征的提取能力,提高边界处的细节感知能力,在不增加模型参数量的同时,增强了模型对红鳍东方鲀小目标识别精度,改进后残差模块结构如图4所示。

SimAM是一种3D注意力模块,通过能量函数计算注意力权值,是一个简单但非常有效的卷积神经网络的注意机制模块。通过设计一种能量函数来计算注意力权值,能够灵活、有效地提高神经网络的表示能力。SimAM注意力模块将空间注意力与通道注意力进行并行或串行组合,在不引入额外参数

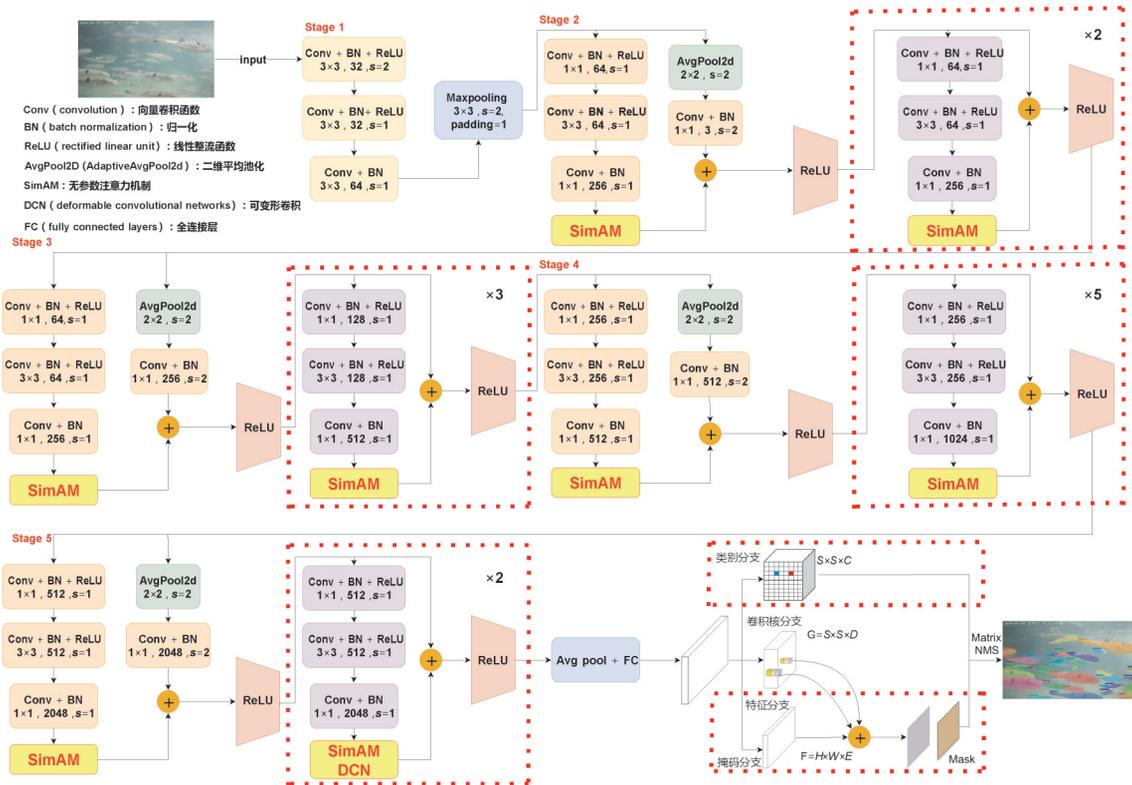


图 2 改进轻量版 SOLOv2 网络结构

Fig.2 Improved light SOLOv2 network architecture

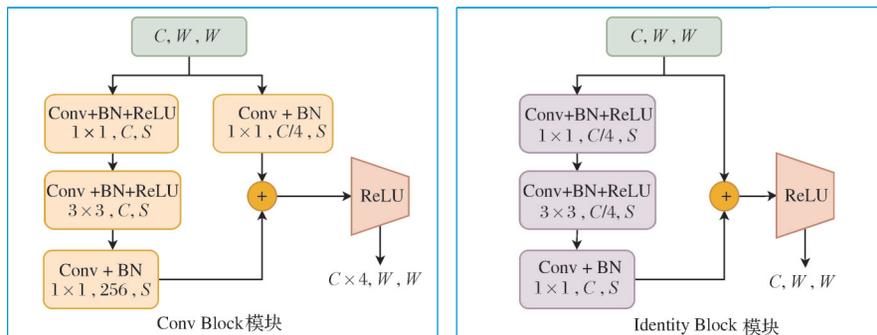


图 3 ResNet50 两个残差模块

Fig.3 Two residual modules of ResNet50

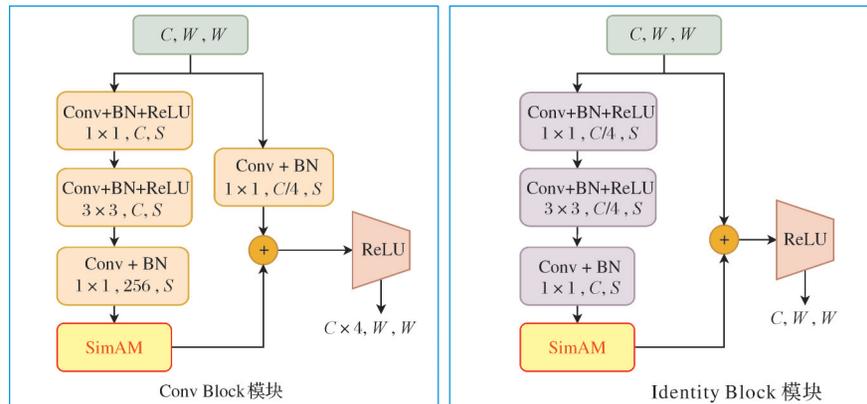


图 4 ResNet50 残差模块结构图

Fig.4 Residual module structure diagram

的情况下为特征图推断3D注意力权值,提出优化能量函数以发掘每个神经元的重要性,利用统一权值的注意力模块,通过优化能量函数以发掘每个神经元的重要性,针对能量函数推导出一种快速解析解,加速注意力权值的计算并得到了一种轻量型注意力模块,结构如图5所示。

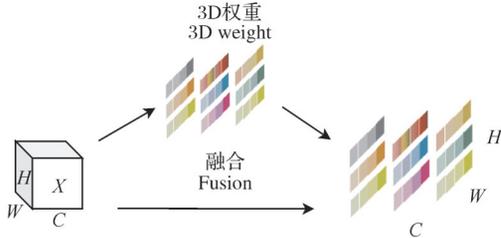


图5 SimAM注意力结构

Fig.5 SimAM attention architecture

神经科学中,信息神经元相较于其他神经元一般有着不同的激活模式,并且抑制其他神经元,应该更加重视具有空间抑制的神经元。最简单的寻找重要神经元的方法:度量神经元之间的线性可分性。因此,能量函数如式(1):

$$e_i(w_i, b_i, y, x_i) = (y_i - \hat{t})^2 + \frac{1}{M-1} \sum_{i=1}^{M-1} (y_0 - \hat{x}_i)^2 \quad (1)$$

$$\hat{t} = w_i t + b_i \quad (2)$$

$$\hat{x}_i = w_i x_i + b_i \quad (3)$$

其中, t 和 x_i 是输入特征 $X \in R^{C \times H \times W}$ 的单一通道中的目标神经元和其他神经元,公式(2)和(3)分别为 t, x_i 的线性变化, $M = H \times W$ 是该通道上的神经元数量, w_i 和 b_i 是权重和偏差的线性变换, y 为变量,通过最小化公式(1),找到目标神经元 t 与同一通道中所有其他神经元之间的线性可分性,最后通过正则项和二值标签简化公式,最终如式(4):

$$e_i(w_i, b_i, y, x_i) = \frac{1}{M-1} \sum_{i=1}^{M-1} (-1 - (w_i x_i + b_i))^2 + (1 - (w_i t + b_i))^2 + \lambda w_i^2 \quad (4)$$

理论上,每个通道有 M 个能量函数, λ 为系数,公式(2)对 w_i 和 b_i 有1个快速的封闭解为:

$$w_i = -\frac{2(t - u_i)}{(t - u_i)^2 + 2\sigma_i^2 + 2\lambda} \quad (5)$$

$$b_i = -\frac{1}{2}(t + u_i)w_i \quad (6)$$

$$u_i = \frac{1}{M-1} \sum_{i=1}^{M-1} x_i \quad (7)$$

$$\sigma_i^2 = \frac{1}{M-1} \sum_{i=1}^{M-1} (x_i - u_i)^2 \quad (8)$$

其中,公式(7)和(8)是该通道中除 t 外的所有神经元的平均值和方差计算。因此,最小能量 e_i^* 公式

如式(9)所示, $1/e_i^*$ 为该神经元的权重。

$$e_i^* = \frac{4(\hat{\sigma}^2 + \lambda)}{(t - \hat{u})^2 + 2\hat{\sigma}^2 + 2\lambda} \quad (9)$$

按照注意力机制的定义,需要对特征进行增强处理,1个神经元对应1个神经元能量 e_i^* ,单个通道的所有神经元的能量组成该通道的能量矩阵 E , X 表示特征图,融合 SimAM 注意力特征图公式如式(10):

$$\tilde{X} = \text{sigmoid}\left(\frac{1}{E}\right) \odot X \quad (10)$$

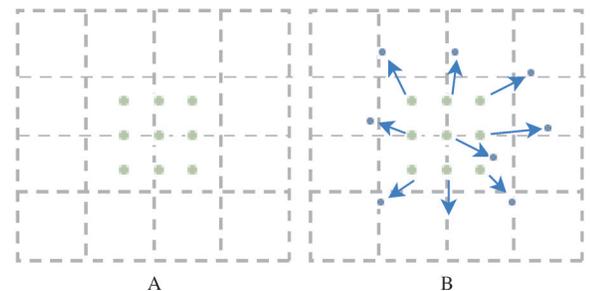
sigmoid 激活函数由公式(11)定义:

$$s(x) = \frac{1}{1 + e^{-x}} \quad (11)$$

改进后的 ResNet 骨干网络,在不添加额外参数的情况下获得不同尺度的特征信息,并且可以提高小目标的分割精度。

3) 可变形卷积 DCN。卷积神经网络有对未知的变化适应性差、泛化能力不强等缺陷,而可变形卷积 DCN 的特点是在感受野中引入偏移量,为卷积核上的每一点学习1个偏移量,使用额外的偏移量来增强模块中的空间采样位置,不需要额外的监督。卷积核可以根据不同的数据学习的卷积结构。通过偏移量的学习,可以使得感受野与物体的实际形状贴近,从而使卷积区域始终覆盖在物体形状周围,无论物体如何形变,都可以通过偏移量进行学习,结构如图6所示,这样的卷积核在学习过程中可以可视化感受野,提高对形变的建模能力,以达到提升实例分割精度的目的,是一种引入极少计算量并提升模型精度的最佳策略。

通过在 Identity Block 模块的最后1层卷积中加入 DCN,利用传统的卷积核提取的特征图作为输入,



A. 绿色是标准的 3×3 卷积 The green is standard 3×3 convolution; B. 给普通卷积加上偏移之后形成的可变形卷积核,蓝色是新的卷积点,箭头是位移方向 The deformable convolution kernel formed after adding the offset to the ordinary convolution, the blue color is the new convolution point, and the arrow is the displacement direction.

图6 DCN结构

Fig. 6 DCN architecture

在原始卷积中增加额外偏移量的空间采样位置,使卷积核可以根据实际情况调整本身的形状,更好地提取输入的特征,从而适应不同鱼的形状、大小等几何形变^[18]。结构图如图7所示。

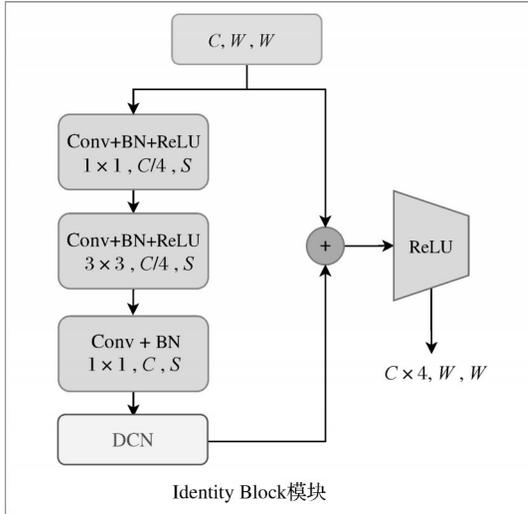


图7 融合DCN残差模块结构图

Fig. 7 Structure diagram of fusion DCN residual module

对于特征图 p 的计算公式为:

$$y(p) = \sum_{p_n \in R} w(p_n) \cdot x(p + p_n + \Delta p_n) \quad (12)$$

式(12)中: $w(p_n)$ 为卷积核的值, n 是在 p 处卷积核的范围内所有采样的位置总数, Δp_n 是偏移矩阵。

4) 损失函数。损失函数包括两部分,分别是分类分支和mask分支,分类loss直接使用FocalLoss,输出采用的是sigmoid激活。 L_{cate} 是用于分类的Focal损失函数, L_{mask} 是用于mask预测的损失函数, λ 设置为3。

$$L = L_{\text{cate}} + \lambda L_{\text{mask}} \quad (13)$$

mask分支的损失函数见式(14):

$$L_{\text{mask}} = \frac{1}{N_{\text{pos}}} \sum_k 1_{\{p_{i,j}^* > 0\}} d_{\text{mask}}(m_k, m_k^*) \quad (14)$$

按从左到右、从上到下索引, $i = [k/s]$, $j = k \bmod s$, N_{pos} 表示阳性样本数, p^* 、 m^* 分别表示类别真值和掩码真值。1是指示函数,如果 $p_{i,j}^* > 1$ 则为1,否则为0。

经过实验,发现Dice loss有很好的稳定性,确定 d_{mask} 使用Dice loss。

$$d_{\text{mask}} = 1 - D(p, q) \quad (15)$$

D 为Dice系数,定义如式(16):

$$D(p, q) = \frac{2 \sum_{x,y} (p_{x,y} \cdot q_{x,y})}{\sum_{x,y} p_{x,y}^2 + \sum_{x,y} q_{x,y}^2} \quad (16)$$

$p_{x,y}$ 和 $q_{x,y}$ 分别为预测掩码和真实掩码在 (x, y) 位置的像素值。

1.3 模型训练

试验运行环境采用百度AI Studio平台和PaddlePaddle,使用平台配置的GPU、Tesla V 100、32G显存以及RAM32G为实验硬件设备。为了提高模型训练速度,减少时间成本,采用基于SOLOv2框架下的轻量版SOLOv2预训练模型进行迁移学习。试验使用红鳍东方鲀数据集,随机划分数据集70%为训练集、20%为测试集、10%为验证集,批量数为16,训练10 000轮。神经网络在刚开始训练的时候非常不稳定,为保证网络能够具有良好的收敛性,通常会采用以较低学习率逐渐增大至较高学习率的方式实现网络训练的“热身”阶段,因此初始学习率设置为0.000 8,权值衰减速率设置为0.000 2。

选用红鳍东方鲀作为数据对象进行模型训练,采用交并比(intersection over union, IoU)、平均精度、平均精度均值、检测速率、参数量、浮点运算次数作为模型的评价指标。本研究选取的判定阈值为0.5和0.75。

2 结果与分析

2.1 模型选取

为了选取基础模型,对标准版SOLOv2模型和轻量版SOLOv2模型进行试验。试验结果如表1所示,可以看出轻量版SOLOv2模型对鱼群的分割精度与标准版SOLOv2模型分割精度相近,有很少量提升,但检测速率较标准版SOLOv2模型提高了34.1个百分点,轻量版SOLOv2模型在分割精度和速度方面都有较好性能,所以,选取轻量版SOLOv2模型作为本研究的基础模型。

表1 模型试验性能

Table 1 Performance of pre-trained model

模型 Models	平均精度 Average precision	平均精度均值 Mean average precision		检测速度/(帧/s) Detection speed
		IoU=0.50	IoU=0.75	
标准版SOLOv2 Original SOLOv2	0.621	0.856	0.724	4.1
轻量版SOLOv2 Light SOLOv2	0.615	0.861	0.727	5.5

2.2 DCN 模块融合位置对比

为了进一步验证可变形卷积DCN对调整模型感受野以及特征的提取能力,对DCN加入模型位置进行对比试验。ResNet50网络有5个阶段,分别在第4、5阶段,第3、4、5阶段和第5阶段融合DCN进行对比试验。结果如表2所示,过多地加入DCN,图像的平均分割精度有很少量的提升,几乎相近,在平均精度增长0.2个百分点的情况下,增加了17%的参数量。

表2 轻量版SOLOv2模型DCN融合位置
试验性能

Table 2 Performance for DCN fusion location test of
lightweight SOLOv2 model

DCN融合位置 DCN fusion position	平均精度 Average precision	主干网络参数量/M Backbone network parameters
Stage(5)	0.644	26.3
Stage(4,5)	0.619	28.1
Stage(3,4,5)	0.646	30.8

从识别效果看,在第3、4、5层加入DCN后模型对图中残缺的小目标鱼类个体的分割效果不好,可能是加入DCN的层数过多,导致特征提取时优化的

不够完善,分割效果如图8所示。对比图8C和图8D可以看出,图8D中部分残缺的目标没有被识别出,小目标的分割精度略减小。所以,最终选取在ResNet50的第5层加入DCN作为本研究试验方案。

2.3 改进SOLOv2模型性能分析

为了评估本研究算法对图像中鱼群的分割效果,将本研究改进的算法与原轻量版SOLOv2模型进行对比,分割结果如图9所示。从图9B中可以看出,本研究算法对鱼群的分割精度明显好于原轻量版SOLOv2模型,在图9C中红圈标记处可以看出,原轻量版SOLOv2模型可能出现目标重复识别和将背景识别为目标等问题,导致图9A中检测出的目标看起来更多,改进后的模型可以很好地提高对特征的提取能力,在图9D红箭头所指区域可以更清晰地看出,改进后的分割精度有明显提升,对小目标鱼群的效果也有显著提升。原主干网络中有许多容易被忽略的小目标特征信息,加入DCN后可以增大模型的感受野,再通过注意力机制模块对这部分特征信息进行注意力重构,突出重要信息,增强模型对细节的感知能力,提高模型对目标的分割精度。

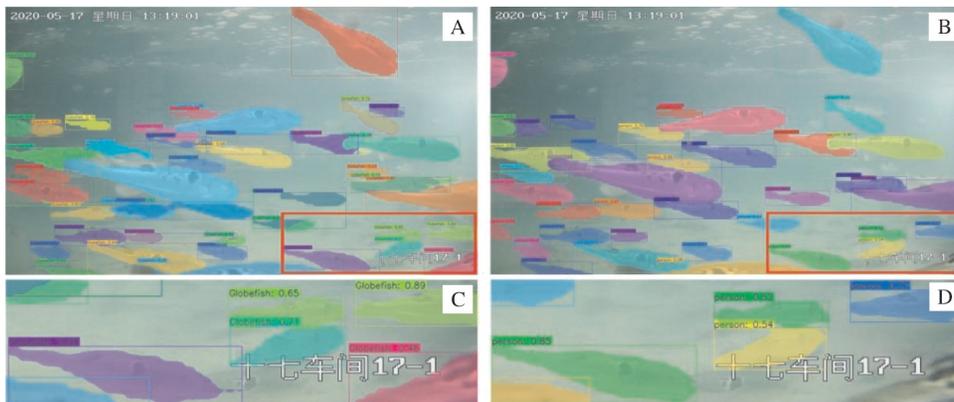


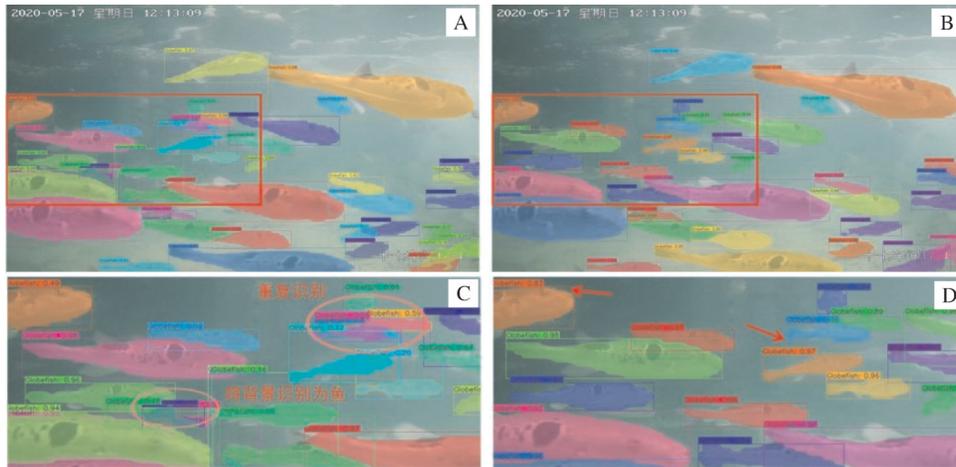
图8 加入DCN后试验分割结果
Fig.8 Experimental segmentation results

由表3可以看出,改进后的SOLOv2模型的各项评价指标均有显著提升,平均分割精度提升了3.7个百分点。原轻量化SOLOv2对小目标的分割效果较弱,加入SiMAM注意力模块后的模型对小目标鱼群的分割精度由0.454提高至0.468,提高了1.4个百分点,说明改进后的模型对小目标也有较好的分割效

果。虽然改进后的模型参数量增加了1.1 M,计算量增加了3.1 G,在几乎不改变模型大小和增加计算量的条件下,使模型保持轻量化,提升模型的分割性能。

2.4 消融实验

为了评估本研究提出的DCN和SimAM两种改



A. 轻量版 SOLOv2 模型分割结果 Model segmentation result diagram of light SOLOv2; B. 改进后模型分割结果图 Model segmentation result diagram of ours; C. 图 A 红框放大视图 Enlarged view of the red frame in figure A; D. 图 B 红框放大视图 Enlarged view of the red frame in figure B.

图 9 SOLOv2 模型试验分割结果

Fig.9 Experimental segmentation results

表 3 模型分割性能指标

Table 3 Model segmentation performance evaluation index

模型 Model	平均精度 Average precision	平均精度均值 Mean average precision			
		IoU=0.50	IoU=0.75	小目标 Small target	大目标 Large target
轻量版 SOLOv2 Light SOLOv2	0.615	0.861	0.727	0.454	0.655
轻量版 SOLOv2+DCN+SimAM Light SOLOv2 +DCN+SimAM	0.652	0.892	0.743	0.468	0.682

进策略对轻量版 SOLOv2 模型分割性能的影响,在数据集上进行消融实验,判断每个改进策略的有效

性,在原有模型基础上依次加上 DCN 和 SimAM,结果如表 4 所示。

表 4 改进模型的性能指标

Table 4 Evaluation index of model improvement performance

模型 Method	DCN	SimAM	平均精度 Average precision	平均精度均值 Mean average precision	
				IoU=0.50	IoU=0.75
轻量版 SOLOv2 Light SOLOv2	—	—	0.615	0.861	0.720
轻量版 SOLOv2 + DCN Light SOLOv2 + DCN	✓	—	0.644	0.884	0.735
轻量版 SOLOv2 + SimAM Light SOLOv2 + SimAM	—	✓	0.647	0.890	0.739
轻量版 SOLOv2 + DCN+ SimAM Light SOLOv2 + DCN+ SimAM	✓	✓	0.652	0.892	0.743

注：“✓”表示采用此方法，“—”表示未采用。Note：“✓” indicates that this method is taken, and “—” indicates that it is not taken.

引入 DCN 之后,模型的平均精度提高了 2.9 个百分点,平均精确均值也有明显提升;在引入 SimAM 注意力之后,模型可以关注到许多易被淹没的语义信息,平均精度提高了 3.2 个百分点,对小目标的分割效果也有很好的改善。同时加入 DCN 和 SimAM 注意力模型的分割精度提高到 65.2%。试验

结果表明,2 种改进方案均能有效提高模型的分割精度,验证了本研究所提方法的有效性。

3 讨论

本研究提出了改进的轻量版 SOLOv2 实例分割网络框架,相比于其他深度模型,SOLOv2 模型既有

很高的分割精度又有快速分割的能力,通过结合可变形卷积DCN和引入无参数注意力模块SimAM,扩大模型感受野,提升模型泛化能力,在增加少量参数的前提下提升模型分割的精度,并且对小目标的分割能力也有很大的提高,改善了水下红鳍东方鲀养殖密度大导致多遮挡等复杂情况下漏检以及小目标物体分割效果差的问题。试验结果表明,改进后的轻量版SOLOv2模型的平均精度为65.2%,检测的平均精度比原模型提高了3.7个百分点,单张图像平均分割的时间为0.195 s,性能优于传统的SOLOv2模型,满足实时分割的要求,为实现智能化、精准化鱼类养殖奠定基础。但由于改进后模型的结构较大,需要大量的算力才能满足边缘部署实时运行的需求。为使改进后的模型更好应用于实际,提高部署的运行效率,今后的研究重点将侧重优化网络结构,提高模型的推理速度,并部署到移动设备上,支持在复杂的环境中进行实时精准的鱼群分割工作。

参考文献 References

- [1] 史磊,秦宏,刘龙腾.世界海洋捕捞业发展概况、趋势及对我国的启示[J].海洋科学,2018,42(11):126-134. SHI L, QIN H, LIU L T. Development situation and trend of world marine fishing industry and its enlightenment to China[J]. Marine sciences, 2018, 42(11): 126-134 (in Chinese with English abstract).
- [2] 王振忠,任鹏,赵红光,等.中国水产种业发展现状与对策研究[J].中国农业科技导报,2017,19(1):1-7. WANG Z Z, REN P, ZHAO H G, et al. Development status and countermeasures of the aquatic breeding industry in China[J]. Journal of agricultural science and technology, 2017, 19(1): 1-7 (in Chinese with English abstract).
- [3] 王柯力.深度学习在渔业图像识别中的应用研究[D].上海:上海海洋大学,2018. WANG K L. Research on the application of deep learning in fishery image recognition[D]. Shanghai: Shanghai Ocean University, 2018 (in Chinese with English abstract).
- [4] 朱纯.基于深度学习的目标实时检测模型的研究与实现[D].南京:东南大学,2019. ZHU C. Research and implementation of real-time target detection model based on deep learning[D]. Nanjing: Southeast University, 2019 (in Chinese with English abstract).
- [5] DENG L. Deep learning: methods and applications[J]. Foundations and trends[®] in signal processing, 2014, 7(3/4): 197-387.
- [6] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.
- [7] HINTON G, DENG L, YU D, et al. Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups[J]. IEEE signal processing magazine, 2012, 29(6): 82-97.
- [8] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84-90.
- [9] HE K M, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN[C]//2017 IEEE International Conference on Computer Vision (ICCV). October 22-29, 2017, Venice, Italy. Venice: IEEE, 2017: 2980-2988.
- [10] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). July 21-26, 2017, Honolulu, HI, USA. Honolulu: IEEE, 2017: 936-944.
- [11] BOLYA D, ZHOU C, XIAO F Y, et al. YOLACT++ better real-time instance segmentation[J]. IEEE transactions on pattern analysis and machine intelligence, 2022, 44(2): 1108-1121.
- [12] WANG X L, ZHANG R F, KONG T, et al. SOLOv2: dynamic and fast instance segmentation[DB/OL]. arXiv, 2020: 2003.10152 [2022-09-29]. <https://doi.org/10.48550/arXiv.2003.10152>.
- [13] 吴逢斌,曹国,时昊.基于改进SOLOv2的转炉炼钢钢液检测[J].计算机应用,2022,42(S1):321-326. WU F B, CAO G, SHI H. Steel liquid detection in converter steelmaking by using improved SOLOv2[J]. Journal of computer applications, 2022, 42(S1): 321-326 (in Chinese with English abstract).
- [14] 花玮,顾梅花,李立瑶,等.改进SOLOv2的服装图像分割算法[J].纺织高校基础科学学报,2021,34(4):74-81. HUA W, GU M H, LI L Y, et al. Clothing image segmentation algorithm based on improved SOLOv2[J]. Basic sciences journal of textile universities, 2021, 34(4): 74-81 (in Chinese with English abstract).
- [15] DAI J F, QI H Z, XIONG Y W, et al. Deformable convolutional networks [DB/OL]. arXiv, 2017: 1703.06211 [2022-09-29]. <https://arxiv.org/abs/1703.06211>.
- [16] YANG L X, ZHANG R Y, LI L D, et al. SimAM: a simple, Parameter-free attention module for convolutional neural networks [C]//Proceedings of the 38th International Conference on Machine Learning. New York: PMLR, 2021: 11863.
- [17] SOFIIUK K, BARINOVA O, KONUSHIN A. AdaptIS: adaptive instance selection network [DB/OL]. arXiv, 2019: 1909.07829 [2022-09-29]. <https://doi.org/10.48550/arXiv.1909.07829>.
- [18] 袁利毫,管英飞,钟声华,等.基于YOLOv3的水下小目标自主识别[J].海洋工程装备与技术,2018,5(S1):118-123. YUAN L H, ZAN Y F, ZHONG S H, et al. Small underwater target recognition based on YOLOv3 [J]. Ocean engineering equipment and technology, 2018, 5(S1): 118-123 (in Chinese with English abstract).

Instance segmentation method of *Takifugu rubripes* based on improved light SOLOv2

HE Jiaqi¹, ZHOU Siyi¹, TANG Xiaomeng¹, HU Xianhui¹, WANG Wei¹, CAI Kewei²

1. *College of Information Engineering, Dalian Ocean University, Dalian 116023, China;*

2. *College of Mechanical and Electrical Engineering, Dalian Minzu University, Dalian 116650, China*

Abstract In order to solve the problems of low image segmentation accuracy and poor segmentation results for small targets caused by the uneven density of *Takifugu rubripes*, an instance segmentation method based on improved light SOLOv2 is proposed. Firstly, the structure of deformable convolutional networks (DCN) is optimized by adjusting the receptive field of the convolution using offset parameters. This adjustment enables the receptive field to be closer to the actual shape of the object, leading to better segmentation accuracy. Next, the parameter-free attention mechanism SimAM is fused in the last layer of the residual module to capture more local information in the image, obtain target features at different scales, and optimize the performance of the model for small target segmentation. The experimental results show that the average segmentation accuracy of the improved lightweight SOLOv2 model was improved by 3.7 percentage, and the segmentation accuracy of small targets was improved by 1.4 percentage compared with the original model. After adding both DCN and SimAM attention modules, the segmentation accuracy of the model increased to 65.2%. The results show that the improved SOLOv2 model can improve the detail perception at the boundary, strengthen the model's ability to extract the features of small target fish stocks, and can be used for accurate instance segmentation in high-density scenarios to achieve accurate pixel-level segmentation of *Takifugu rubripes*.

Keywords instance segmentation; segmentation accuracy; attention mechanism; small objects; precision farming

(责任编辑:边书京)