

DOI: 10.13382/j.jemi.B2205610

# 融合边缘监督的改进 Deeplabv3+水下鱼类分割方法<sup>\*</sup>

田志新<sup>1</sup>  廖  薇<sup>2</sup>  茅  健<sup>1</sup>  吴建民<sup>1</sup>  袁  泉<sup>3</sup>  徐  震<sup>1</sup>

(1. 上海工程技术大学机械与汽车工程学院  上海  201620; 2. 上海工程技术大学电子电气学院  上海  201620;  
3. 上海市农业科学院  上海  201400)

**摘  要:**水下环境鱼类分割是实现体长测量、体重估算和群体计数等智能化测量的关键技术,为了提高分割的准确性,提出一种融合边缘监督的改进 Deeplabv3+鱼类分割方法。编码部分采用更少的下采样次数,浅层增加卷积块注意力机制(convolutional block attention module, CBAM),以减少信息丢失并增强浅层语义信息;通过设计混合膨胀卷积(hybrid dilated convolution, HDC)改进空洞空间卷积池化金字塔(atrous spatial pyramid pooling, ASPP)模块,提取深层特征;在解码输出部分结合 Canny 边缘检测算子引入边缘监督,通过边缘损失函数来获得边缘预测和边缘标签的误差以更好地学习边缘特征;最后根据不同类像素比率引入优化的损失函数,进一步提升模型的语义分割性能。该方法在 VOC2012 数据集上 mIoU 达到 84.56%,较 Deeplabv3+方法提升了 3.27%,验证了其泛化能力。在 DeepFish 数据集上做消融实验,mIoU 高达 93.66%,均高于 Deeplabv3+、Unet 和 PSPNet 等常见方法。该研究提升了水下环境鱼类分割的精度,可为水产养殖智能化提供支持。

**关键词:** 鱼类分割;边缘监督;Deeplabv3+;CBAM 注意力机制;混合膨胀卷积

**中图分类号:** TP391.41      **文献标识码:** A      **国家标准学科分类代码:** 520.0

## Improved Deeplabv3+ underwater fish segmentation method combining with edge supervision

Tian Zhixin<sup>1</sup>  Liao Wei<sup>2</sup>  Mao Jian<sup>1</sup>  Wu Jianmin<sup>1</sup>  Yuan Quan<sup>3</sup>  Xu Zhen<sup>1</sup>

(1. School of Mechanical and Automotive Engineering, Shanghai University of Engineering Science, Shanghai 201620, China; 2. School of Electronic and Electrical Engineering, Shanghai University of Engineering Science, Shanghai 201620, China; 3. Shanghai Academy of Agricultural Sciences, Shanghai 201400, China)

**Abstract:** Fish segmentation in underwater environment is the key technology to realize intelligent measurement such as body length measurement, weight estimation and population counting. In order to improve the accuracy of fish segmentation, an improved Deeplabv3+ fish segmentation method combined with edge supervision is proposed. In the encoder part, fewer down sampling times are used, and convolutional block attention module (CBAM) is added in the shallow layer to reduce information loss and enhance the shallow semantic information; By designing hybrid dilated convolution (HDC) to improve atrous spatial pyramid pooling(ASPP) module, deep features are extracted. In the decoder output part, Canny edge detection operator is combined to introduce edge supervision, and the edge prediction and edge label errors are obtained through the edge loss function to better learn edge features. Finally, the optimized loss function is introduced according to different pixel ratios to further improve the semantic segmentation performance of the model. This method achieves 84.56% mIoU on VOC2012 dataset, which is 3.27% higher than Deeplabv3+ method, and verifies its generalization ability. In the ablation experiment on DeepFish dataset, mIoU is as high as 93.66%, which is higher than common methods such as Deeplabv3+, Unet and PSPNet. This research improves the accuracy of fish segmentation in underwater environment and can provide support for intelligent aquaculture.

**Keywords:** fish segmentation; edge supervision; Deeplabv3+; CBAM attention mechanism; hybrid dilated convolution

## 0 引言

水下鱼类分割对于智能化捕捞识别、智能监测、鱼类体长测量、体重估算以及生长健康状态的判别十分重要。语义分割方法分为传统图像处理方式和深度学习方式<sup>[1]</sup>。传统分割方法如聚类法、阈值法、边缘检测等分割方法<sup>[2-3]</sup>,但对于噪点比较敏感,对于环境要求较为苛刻。近年来,随着深度学习和计算机视觉的发展,基于深度卷积网络的深度学习方法由于其高效的信息提取能力,被广泛地应用于智能化水产养殖行业。利用视频以及图像分析鱼的状态,有助于对其进行形体和运动状态识别和分析,监测鱼类生长,从而使得养殖过程更加智能高效。刘斌等<sup>[4]</sup>通过建立深度编解码语义分割模型,对鱼类各个部位进行分割,利用分割结果对鱼的各部位建立质心随体坐标,从而判明鱼体动作姿态。语义分割对于鱼类轮廓和形状的细化效果较好,因此也经常作为体长测量的预处理方法,Fernandes 等<sup>[5]</sup>设计了基于深度学习的语义分割模型,能够正确区分鱼体和背景以及鳍,并且给予分割提取的鱼体面积成功地用于估算鱼体长和鱼体重量等重要信息。通过鱼体分割,对于鱼的生物量估计也提供了一个十分重要的依据,Abinaya 等<sup>[6]</sup>基于深度网络和语义分割模型并结合 YOLOv4 确定头部、躯干、尾部,建立各部位关键点,使用长度-质量关系确定鱼类的生物量,这对于深入了解鱼类群体等信息提供了十分可靠的参考。鱼类各部位的分割,对于智能捕捞机器人的识别和定位也是一种十分重要的图像数据,王红君等<sup>[7]</sup>提出一种 SENet 优化的 DeepLabv3+鱼类分割方法,分割性能有明显的提升。

Deeplabv3+网络具有多尺度特征,因而对于图像的细节处理,以及小目标的分割会有较好的效果,同时引入深度可分离卷积大大减少了参数量,使得网络模型更加轻量化。基于 Deeplabv3+多尺度特征,各研究者提出的改进模型也层出不穷,邱云飞等<sup>[8]</sup>在网络的浅层特征引入注意力机制,在 VOC2012 数据集上 mIoU 达到 84.44%。任鸿杰等<sup>[9]</sup>提出的改进模型在 ASPP 模块中引入 DAM 注意力模块,并将改进的方法应用于遥感农作物图像分割,分割效果优于传统方法。马冬梅等<sup>[10]</sup>提出在 ASPP 模块中增加带状卷积,增加远程上下文特征信息的提取,在保证准确率的基础上是计算量大大降低。边缘作为一种十分重要的视觉信息在语义分割任务中往往被忽略,提高边缘细节的分割精度成为各研究者们研究热门。Ali 等<sup>[11]</sup>、Cheng 等<sup>[12]</sup>分别深入研究了深度语义分割任务与边缘监督的联系,建立目标与边缘之间的关系,获得更好的边缘感知效果,并分别应用于海陆分割船舶检测等领域。

本文针对 Deeplabv3+的多尺度特征提取特点并结合边缘监督以及 ASPP 模块提出更加细化的改进策略:1)减少下采样次数,从编码层浅层提取出 4、7 两层并增加 CBAM 注意力机制;2) ASPP 层引入更加高效的 HDC 模块,有效地解决了“gridding effect”问题;3) 编码输出层融合边缘监督模块,与分割任务共享神经网络参数,通过边缘损失函数建立关联,实现边缘监督;4) 训练阶段,引入更加优化的损失函数,通过计算不同类的像素数确定损失函数权重,提高模型分割的准确性。将改进的模型在 VOC2012 数据集上进行测试以验证本方法的泛化能力。然后应用在 DeepFish 数据集上,取得了良好的分割结果,对于智能化养殖具有十分重要的研究意义。

## 1 方法原理

### 1.1 Deeplabv3+模型

Deeplabv3+是 Chen 等<sup>[13]</sup>于 2018 年提出的基于编解码结构的分割方法,是在 Deeplabv3 的基础上进行改进的方法。沿用 Deeplabv3 的 ASPP 模块,聚合多尺度特征信息,对 backbone 提取的特征层进行特征信息强化处理。

Deeplabv3+模型如图 1 所示,分为编码部分和解码部分,编码部分 backbone 采用 Mobilenetv2 网络,这是一种基于深度可分离卷积的轻量化网络,从 Mobilenetv2 网络中取出浅层特征和深层特征。深层特征经过 ASPP 模块增强后通过双线性插值上采样,与浅层特征进行拼接。

### 1.2 CBAM 注意力机制

CBAM 注意力机制模块结构如图 2 所示,结合了通道注意力和空间注意力。可以保证在深度网络前向传递的过程中,浅层特征的语义信息和空间位置信息被增强。最大池化倾向于特征比较突出,如轮廓等信息,而平均池化则比较偏向背景信息的保留,CBAM 同时采用两种池化方式,充分的保留了特征信息。

### 1.3 ASPP 模块

Deeplabv3+对于深层特征使用 ASPP 模块,通过使用膨胀率  $r = 6, 12, 18$  的卷积核对特征层进行多尺度特征提取,从而增强输出特征层的信息。通过使用不同膨胀率的卷积核实现不同的感受野。

式(1)、(2)为感受野计算公式:

$$RF_n = RF_{n-1} + (k' - 1) \prod s_i \quad (1)$$

$$k' = r(k - 1) + 1 \quad (2)$$

式中: $RF_n$  为第  $n$  层的感受野, $k'$  为带膨胀率的卷积核的等效卷积核大小, $k$  为当前层的实际卷积核大小, $r$  为膨胀率, $s_i$  为各层卷积的步距。通过计算得到  $r = 6, 12, 18$  的时候对应的感受野大小为 13, 25, 37。这种大膨胀率的卷积核实现了感受野扩大的同时会造成特征提取过程

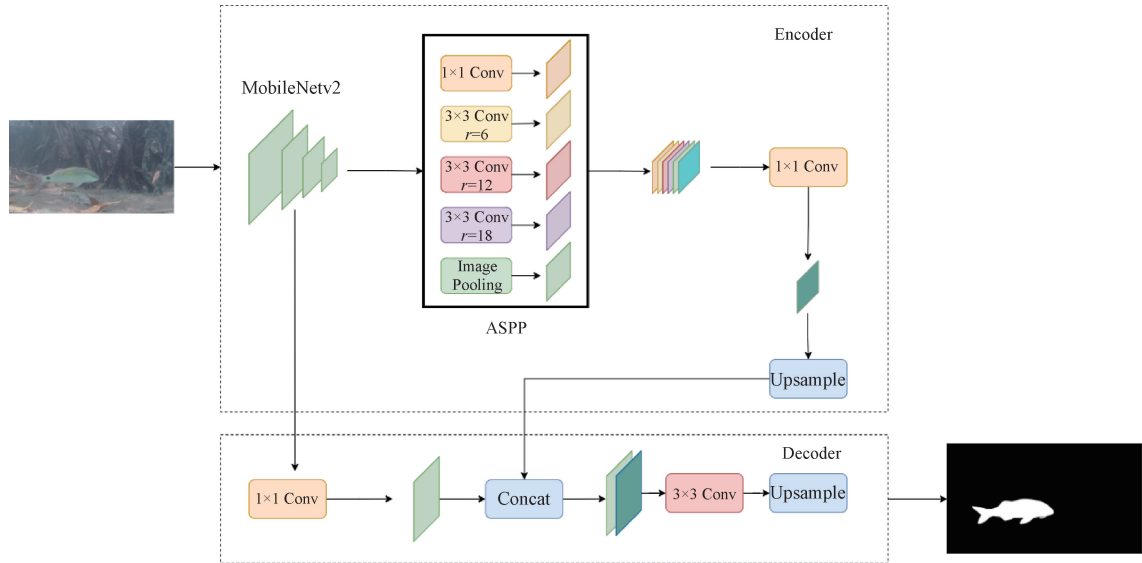


图 1 传统 Deeplabv3+模型  
Fig. 1 Traditional Deeplabv3+ model

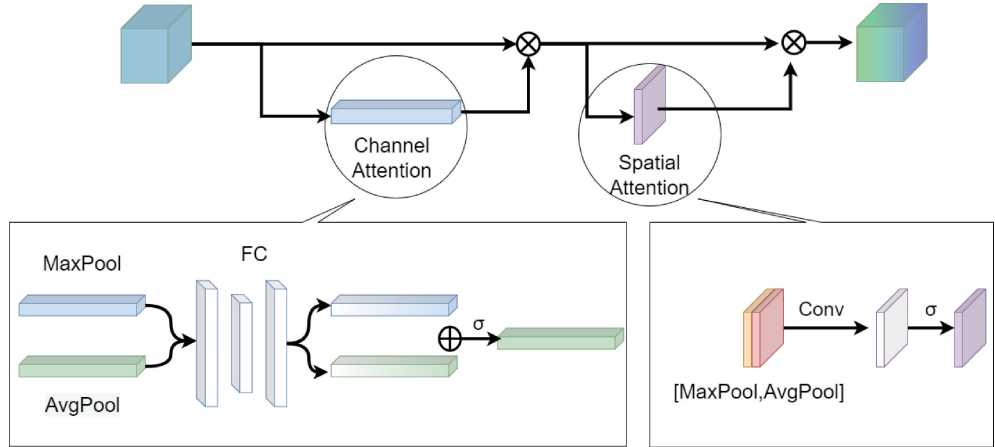


图 2 CBAM 注意力机制模块  
Fig. 2 CBAM attention mechanism module

中的“孔洞”的产生,造成“gridding effect”问题,因此在改进的模型中引入 HDC 模块。

1.4 边缘监督

网络层数的加深和池化操作往往会忽视图像中的细节如边和角等,导致边缘附近的分割效果较差,因此,本文设计了边缘监督模块。

语义分割标签经过 Canny 算子进行边缘提取,生成边缘标签。边缘监督模块与语义分割共享神经网络参数,同时,引入边缘损失函数  $edge\_loss$ ,并通过边缘标签对网络参数进行微调,实现边缘监督功能。

2 改进的 Deeplabv3+模型

2.1 网络结构设计

传统 Deeplabv3+模型对输入图片经过 4 次下采样和

4 次上采样,恢复到原图大小,在这个过程中浅层信息在深度网络前向传递的过程中丢失了很多,上采样对这点的弥补能力略显匮乏。因此在改进的模型中考虑减少下采样次数,并在浅层特征中提取出第 4,7 两层,并且加入 CBAM 注意力机制模块,增强浅层的语义信息。

深层特征层的输出经过改进的 ASPP 模块进行特征增强,使用 1 个 HDC 模块替代膨胀率  $r=6$  的卷积模块,2 个 HDC 模块串联替代  $r=12$  的卷积模块,使用 3 个 HDC 模块串联替代  $r=18$  的卷积模块。根据感受野计算公式计算得引入的 HDC 结构改进后的 3 个膨胀卷积模块的感受野同样也为 13,25,37。实现了在同样的感受野下,避免了“gridding effect”问题,减少了孔洞的产生。

式(3)为 HDC 膨胀率设计时遵循的规则。

$$M_i = \max[M_i + 1 - 2r_i, M_i + 1 - 2(M_i + 1 - r_i), r_i] \tag{3}$$

式中:  $M_i$  为第  $i$  层卷积层的非 0 值之间最大间距, 设计出来的膨胀率  $r$  应当使得  $M_2 \leq \text{卷积核 } K^{[14]}$ 。本文设计 HDC 模块的膨胀率为  $[1, 2, 3]$ , 如图 3 所示。

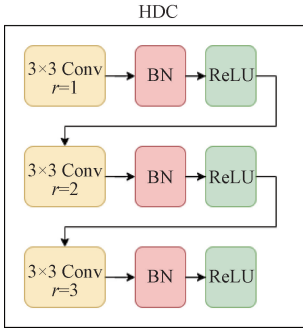


图 3 HDC 模块

Fig. 3 HDC module

在解码输出层设置了边缘标签对结果进行边缘监督, 标签由数据集语义标签经过 Canny 边缘检测得到, 并且根据像素数配置类权重, 设置 edge\_loss 参与总的 loss 损失计算, 共享的神经网络参数以及关联的损失函数使得边缘标签对网络参数产生针对性的微调, 进而得到更加精细的分割结果。改进后的模型如图 4 所示。

## 2.2 优化的损失函数

训练过程中, 由于各样本的像素区域大小不一, 尤其是背景像素在一幅图上占据比例最大, 在计算损失的时候通常会导致学习过程陷入损失函数的局部最小值, 从而导致网络的预测严重偏向于相对较大的类别<sup>[15]</sup>, 从而使得模型对于小物体的分割精度较差。所以在设计损失函数的时候应对于不同的类别给与不同的权重, 基于此,

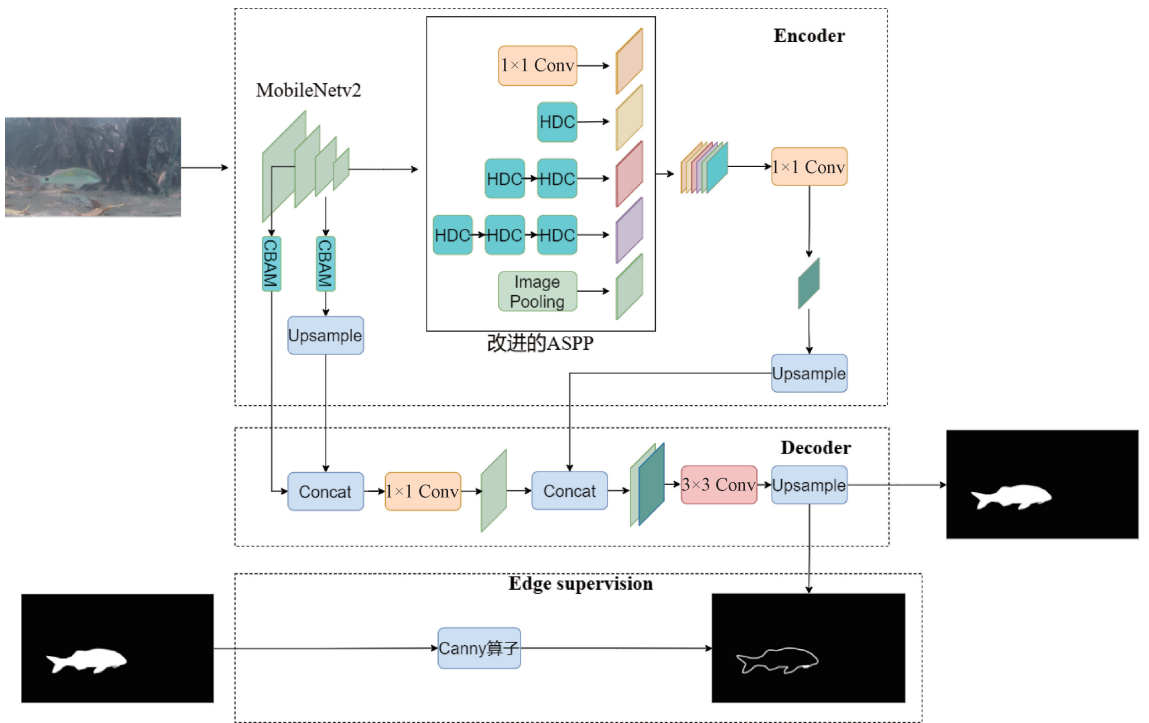


图 4 改进的 Deeplabv3+ 模型

Fig. 4 Improved Deeplabv3+ model

引入 focal loss 和 dice loss, 并在 focal loss 上赋予不同的权重。Dice loss 是一种基于区域的损失, 从全局出发计算损失, 而 focal loss 是基于分布的损失<sup>[16]</sup>, 是交叉熵损失的一种变体, 通过降低简单示例的权重来解决难易样本的不平衡问题<sup>[17]</sup>。侧重于像素级计算损失, 二者结合, 融合各自的优势, 有助于模型的快速收敛并且对于不同权重的目标的分割更加精确。

Dice loss 定义为:

$$L_s = 1 - \frac{2|X \cap Y| + 1}{|X| + |Y| + 1} \quad (4)$$

式中:  $X$  为预测结果的分割区域,  $Y$  为实际标签区域。  $X \cap Y$  为预测结果与真实标签区域的重合部分。

Focal loss 定义为:

$$L_f = -\alpha(1 - p_i)^\gamma \log(p_i) \quad (5)$$

式中:  $L_f$  为 focal loss 损失函数,  $p_i$  分类结果的置信度, 越接近 1, 则认为越容易区分, 其权重则越小。  $\alpha$  和  $\gamma$  为权重的控制程度参数,  $\gamma = 0$  时, 即为交叉熵损失。本文取  $\alpha = 0.5, \gamma = 2$ 。

边缘损失 Edge loss 定义为:

$$L_e = -\lambda_e p_e \log(\hat{p}_e)$$

(6)

式中: $L_e$  为边缘损失函数, $\lambda_e$  为边缘和背景像素权重, $p_e$  为边缘标签值, $\hat{p}_e$  为边缘预测值。

总的 Loss 定义为 3 种损失相加:

$$L = \lambda_i L_f + L_s + L_e$$

(7)

$$\lambda_i = 1 - r_i$$

(8)

式中: $r_i$  为类别像素占比, $\lambda_i$  为各个类别权重, $L_f$  为 focal loss, $L_s$  为 Dice loss, $L_e$  为边缘损失。

### 3 实验与分析

#### 3.1 参数配置与数据集准备

实验设备预装 Python3.7, CUDA10.0, GPU 选择 NVIDIA Tesla P100, 显存 16 GB。系统环境为环境 Ubuntu18.04 系统,深度学习框架使用 Pytorch。

训练集合测试集划分为 7 : 3,每个 batch 的大小设置为 8、16、24 分别进行实验测试。初始学习率为 0.007,采用余弦退火衰减策略。使用 Adam 优化器,动量参数设置为 0.9,为避免网络初始权值的随机性,预先下载与训练权重进行迁移学习。

数据集选择 PASCAL VOC2012 和鱼类分割数据集和鱼类分割数据集 DeepFish<sup>[18]</sup>。DeepFish 数据集来自 20 个不同栖息地的鱼类图像样本,其中,用于语义分割的图像一共 310 张图像,经过图像翻转等操作,将图像扩充至 620 张,以达到模型训练需求。

#### 3.2 结果与分析

##### 1) VOC2012 测试结果

为了验证该方法的普适性,使用 PASCAL VOC2012 公共数据集进行测试。首先计算出各个类标签的像素数比,根据比例分配各类的权重,权重文件保存为 json 格式。训练世代为 100 个 epoch,batch\_size 设置为 24,每个 epoch 记录 loss 值,得到 loss 下降曲线如图 5 所示。

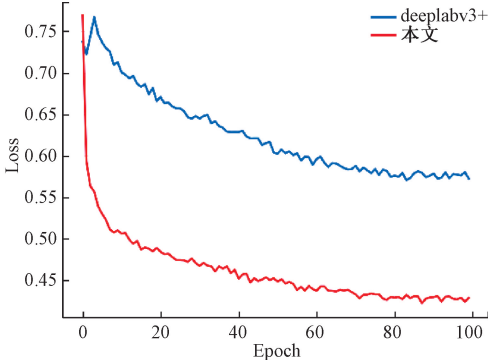


图 5 本方法和 Deeplabv3+训练过程 loss 曲线

Fig. 5 The method and loss curve of Deeplabv3+ training process

图 5 和表 1 表明,改进的模型很快能够收敛,并且最佳 loss 值在 0.42 附近,低于原网络模型,这是因为引入的优化损失函数可以对难识别类、小物体类、像素占比低的类进行着重训练,并且浅层注意力机制会增强浅层语义信息,使得模型对这些类别有较好的识别效果,降低了总体的 loss,提高了总体的 mIoU,mRecall 和 mPrecision。

表 1 不同方法的分割结果对比

| Table 1 Comparison of segmentation results of different methods |        |           |              |
|-----------------------------------------------------------------|--------|-----------|--------------|
| 方法                                                              | mIoU/% | mRecall/% | mPrecision/% |
| Deeplabv3+                                                      | 81.29  | 90.88     | 88.11        |
| 改进的 Deeplabv3+                                                  | 84.56  | 93.85     | 89.31        |

##### 2) DeepFish 数据集测试结果

在模型开始训练前,首先需要对语义分割标签图计算各个类像素个数,确定 loss 函数的权重。对 DeepFish 数据集标签图使用 Canny 算法提取出边缘,作为边缘监督的标签,如图 6 所示,并且将鱼类数据集的所有标签处理成可被网络模型识别的图像。

图 7 为本方法和使用原始 Deeplabv3+对数据集的分割结果。本方法在边缘细节以及小目标的分割效果都较原始 Deeplabv3+方法的效果好,例如图 7(c1)中鱼的尾部分割以及图 7(c2)的小鱼以及边缘细节的分割。浅层特征的注意力机制使得模型对小目标的识别效果更好,并且引入边缘监督使得模型对于边缘细节的处理更好。

训练过程中 batch\_size 的大小会影响到模型的性能,batch\_size 越大,模型性能越好,识别精度就越好<sup>[19]</sup>。本文应用 3 种 batch\_size,分别为 8、16、24,对原 Deeplabv3+、替换改进 ASPP 模块的 Deeplabv3+、浅层引入 CBAM 注意力机制的 Deeplabv3+和引入边缘监督的 Deeplabv3+做了消融实验,得到如表 2 的结果。

表 2 消融实验结果

| Table 2 Results of ablation experiment |          |      |      |              |
|----------------------------------------|----------|------|------|--------------|
| batch_size                             | 改进的 ASPP | CBAM | 边缘监督 | mIoU/%       |
| 8                                      |          |      |      | 87.51        |
| 8                                      | ✓        |      |      | 88.97        |
| 8                                      | ✓        | ✓    |      | 89.74        |
| 8                                      | ✓        | ✓    | ✓    | <b>91.96</b> |
| 16                                     |          |      |      | 88.04        |
| 16                                     | ✓        |      |      | 88.61        |
| 16                                     | ✓        | ✓    |      | 90.52        |
| 16                                     | ✓        | ✓    | ✓    | <b>92.25</b> |
| 24                                     |          |      |      | 88.16        |
| 24                                     | ✓        |      |      | 89.60        |
| 24                                     | ✓        | ✓    |      | 90.31        |
| 24                                     | ✓        | ✓    | ✓    | <b>93.66</b> |

表 2 结果表明,随着 batch\_size 的增加,各方法的

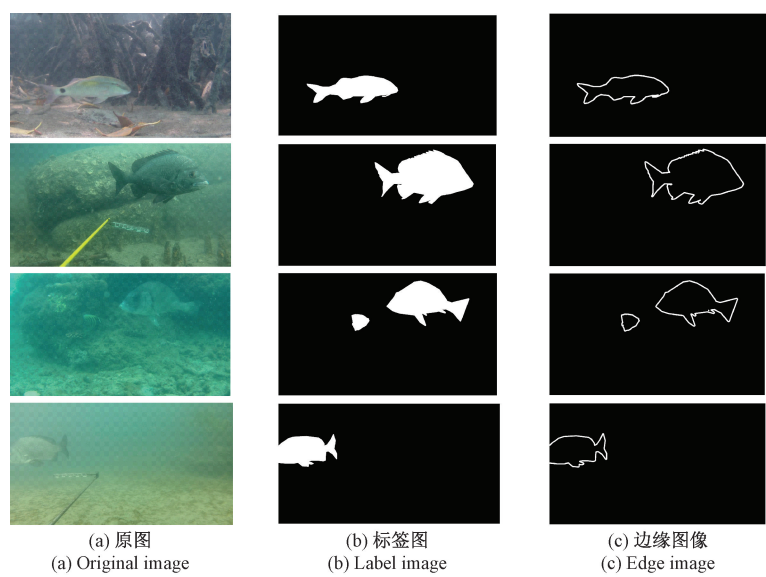


图 6 DeepFish 数据集与标签处理

Fig. 6 DeepFish data set and label processing

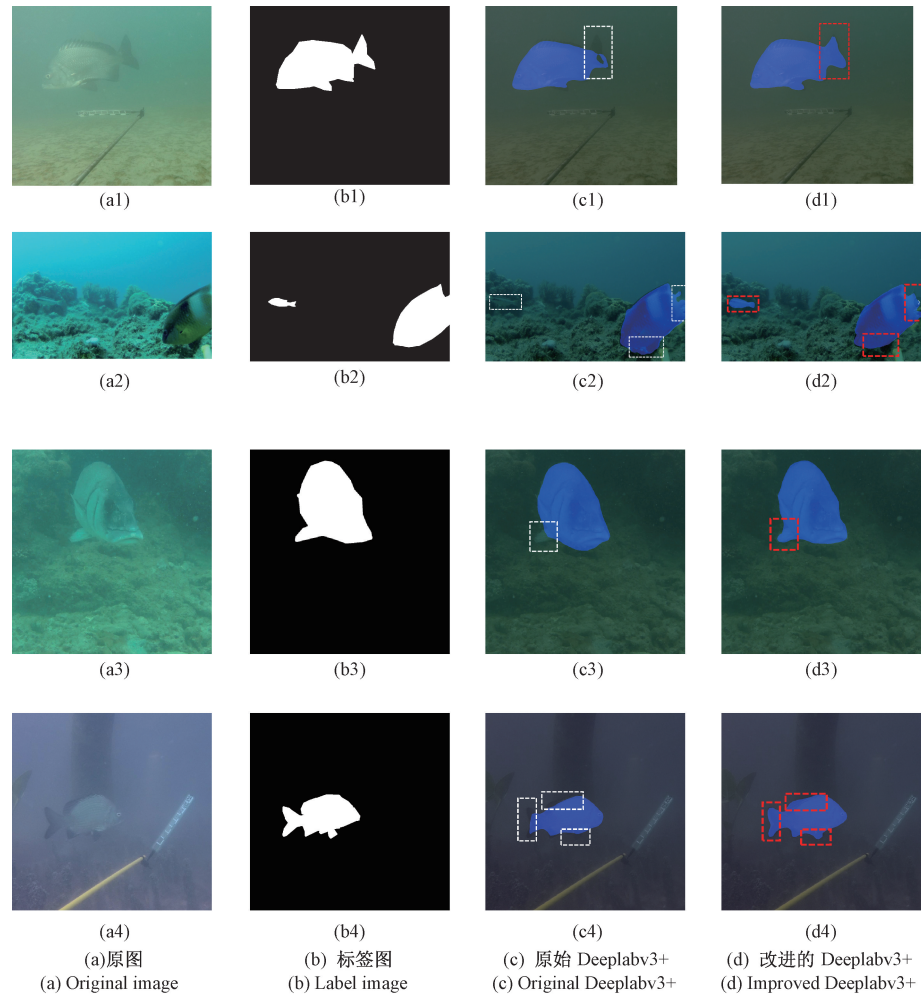


图 7 改进方法和原方法分割结果对比

Fig. 7 Comparison of segmentation results between improved method and original method

mIoU 均有不同程度的增加。在不同的 batch\_size 下,本方法的 mIoU 均表现最佳,在 batch\_size 为 24 时,本方法的 mIoU 可以达到 93.66%,表明使用了更加细密的 ASPP 模块、注意力机制和边缘监督使得分割效果更好。

表 3 本方法和现有方法对比结果

Table 3 Comparison results between new method and existing method

| 方法                   | baseNet     | 背景 IoU/% | 鱼 IoU/% | mIoU/% | 参数量/MB       | 运行时间/s       |
|----------------------|-------------|----------|---------|--------|--------------|--------------|
| PSPNet               | mobilenetv2 | 98       | 46      | 71.66  | 2.03         | 45.03        |
| PSPNet               | resnet50    | 99       | 71      | 84.98  | <b>46.79</b> | <b>72.28</b> |
| Unet                 | vgg         | 99       | 86      | 92.50  | 23.74        | 70.49        |
| Unet                 | resnet50    | 99       | 83      | 91.29  | 41.90        | 52.60        |
| Deeplabv3+           | mobilenetv2 | 99       | 77      | 88.16  | 5.54         | 56.30        |
| Attention_DeepLabv3+ | mobilenetv2 | 99       | 82      | 90.31  | 14.06        | 48.60        |
| ASPP_DeepLabv3+      | mobilenetv2 | 99       | 80      | 89.60  | 13.99        | 63.64        |
| 本方法                  | mobilenetv2 | 99       | 88      | 93.66  | 14.06        | 47.83        |

基于 mobilenetv2 的 PSPNet 使得网络模型更加轻量化,但是分割效果较差;Deeplabv3+采用 encoder-decoder 框架,同样基于 mobilenetv2 的轻量化网络效果提升了 17%;基于 vgg 和 resnet50 的 Unet 网络由于使用了 encoder-decoder 框架,提升了分割效果,但是相较 mobilenetv2,模型更加庞大,参数量是 mobilenetv2 模型的 2~4 倍之多;本方法是基于 Deeplabv3+,引入改进的 ASPP 模块、CBAM 注意力机制和边缘监督,保持模型的相对轻量化的前提下,获得了更好的分割效果,mIoU 达到 93.66%。

3) 注意力可视化

将本方法和 PSPNet<sup>[20]</sup>、PSPNet<sub>vgg</sub>、Deeplabv3+、Unet<sub>vgg</sub><sup>[21]</sup>、Unet<sub>resnet50</sub>、基于注意力机制的 Deeplabv3+和改进 ASPP 的 Deeplabv3+在各类 IoU、mIoU、参数量和运行时间上做了对比,表 3 为实验结果。

将本方法和 Deeplabv3+的 ASPP 层和分类层的注意力特征可视化,如图 8 所示。由于改进的 ASPP 模块使用更加细密的 HDC 模块,在不改变感受野的前提下,使得特征图的视野更加细密,相较于原方法,减少了孔洞的产生。如图 8(b)和(d)对比,原方法的 ASPP 可视化结果孔洞较大,特征提取较为粗糙而改进的 ASPP 特征提取较为细密;改进的 Deeplabv3+结构在浅层的两个特征层增加了 CBAM 注意力机制,在分类时,会更多的注意到分类目标,尤其是小目标,图 8(c)和(d)展示了分类层对于目标类的注意力可视化,本方法在对目标类的注意力范围较为准确。

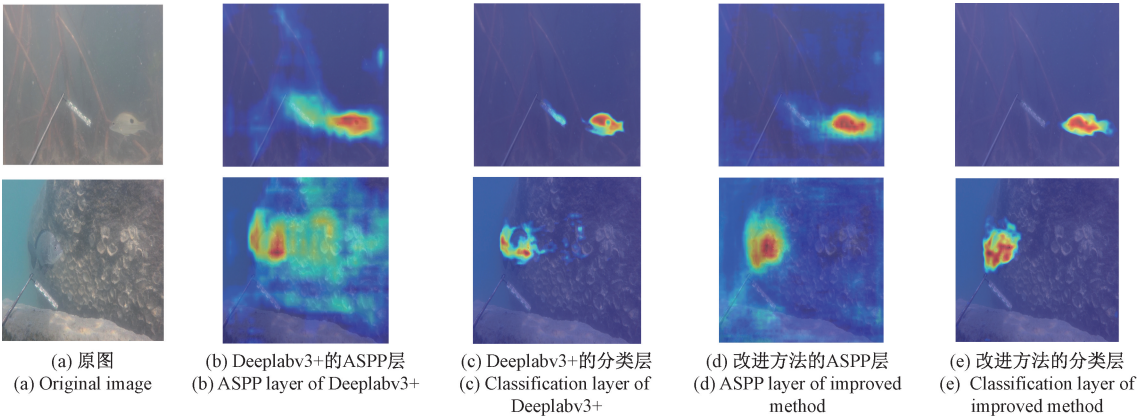


图 8 注意力可视化

Fig. 8 Attention visualization

4 结 论

本文提出一种融合边缘监督的改进 Deeplabv3+水下鱼类分割方法,主要得出如下结论:1) 本方法在 VOC2012 数据集上测试 mIoU 达到 84.56%,比传统

Deeplabv3+方法提升了 3.27%,验证了本方法的泛化能力。2) 通过选择不同的 batch\_size(8、16、24)在 DeepFish 数据集上做消融实验,在各个 batch\_size 下的 mIoU 为 91.96%、92.25%和 93.66%,高出原方法 4.45%、4.21%和 5.50%,验证了本方法的优势。3) 与其他现有方法在 DeepFish 数据集的分割结果作对比,从模型大小、分割结

果准确性以及推断时间上进行分析,本方法在轻量化的同时保证了分割的准确性。4)将本方法和 Deeplabv3+在 ASPP 层和分类层的注意力特征可视化,提出的方法在 DeepFish 数据集上表现出了优异的性能。在后续的研究中,将继续研究兼顾分割精度和更加轻量化的高性能网络,并且将其部署到嵌入式硬件系统中,进一步扩展其在水产养殖工程中的实用价值。

## 参考文献

- [ 1 ] 邵晓琦,徐海燕,侯一兴,等. 基于多通道增强融合的自适应液晶屏图像分割[J]. 电子测量与仪器学报, 2020, 34(12): 76-84.  
SHAO X Q, XU H Y, HOU Y X, et al. Adaptive LCD image segmentation based on multi-channel enhancement fusion [J]. Journal of Electronic Measurement and Instrumentation, 2020, 34 (12): 76-84.
- [ 2 ] 潘峰,孙红霞. 基于蝗虫算法的图像多阈值分割方法[J]. 电子测量与仪器学报, 2019, 33(1): 149-155.  
PAN F, SUN H X. Multi-threshold image segmentation method based on locust algorithm [J]. Journal of Electronic Measurement and Instrumentation, 2019, 33 (1): 149-155.
- [ 3 ] 梁新宇,罗晨,权冀川,等. 基于深度学习的图像语义分割技术研究进展[J]. 计算机工程与应用, 2020, 56(2): 18-28.  
LIANG X Y, LUO CH, QUAN J CH, et al. Research progress of image semantic segmentation technology based on deep learning [J]. Computer Engineering and Application, 2020, 56 (2): 18-28.
- [ 4 ] 刘斌,王凯歌,李晓蒙,等. 基于语义部位分割的条纹斑竹鲨鱼体运动姿态解析[J]. 农业工程学报, 2021, 37(3): 179-187.  
LIU B, WANG K G, LI X M, et al. Analysis of shark body movement posture of striped bamboos based on semantic segmentation [J]. Journal of Agricultural Engineering, 2021, 37 (3): 179-187.
- [ 5 ] FERNANDES A F A, TURRA E M, DE ALVARENGA É R, et al. Deep learning image segmentation for extraction of fish body measurements and prediction of body weight and carcass traits in Nile tilapia [J]. Computers and Electronics in Agriculture, 2020, 170: 105274.
- [ 6 ] ABINAYA N S, SUSAN D, SIDHARTHAN R K. Deep learning-based segmental analysis of fish for biomass estimation in an occulted environment [J]. Computers and Electronics in Agriculture, 2022, 197: 106985.
- [ 7 ] 王红君,季晓宇,赵辉,等. SENet 优化的 Deeplabv3+淡水鱼体语义分割[J]. 中国农机化学报, 2021, 42(2): 158-163.  
WANG H J, JI X Y, ZHAO H, et al. Semantics segmentation of DEPLABV3+freshwater fish optimized by SENet [J]. Chinese Journal of Agricultural Machinery Chemistry, 2021, 42 (2): 158-163.
- [ 8 ] 邱云飞,温金燕. 基于 DeepLabV3+与注意力机制相结合的图像语义分割[J]. 激光与光电子学进展, 2022, 59(4): 130-139.  
QIU Y F, WEN J Y. Image semantic segmentation based on DeepLabV3+ and attention mechanism [J]. Advances in Laser and Optoelectronics, 2022, 59 (4): 130-139.
- [ 9 ] 任鸿杰,刘萍,岱超,等. 改进 DeepLabV3+网络的遥感影像农作物分割方法[J]. 计算机工程与应用, 2022, 58(11): 215-223.  
REN H J, LIU P, DAI CH, et al. Crop segmentation method of remote sensing image based on improved DeepLabV3+ network [J]. Computer Engineering and Application, 2022, 58 (11): 215-223.
- [ 10 ] 马冬梅,李鹏辉,黄欣悦,等. 改进 DeepLabV3+的高效语义分割[J]. 计算机工程与科学, 2022, 44(4): 737-745.  
MA D M, LI P H, HUANG X Y, et al. Improving efficient semantic segmentation of DeepLabV3+ [J]. Computer Engineering and Science, 2022, 44 (4): 737-745.
- [ 11 ] ALI U, BAYRAMLI B, ALSARHAN T, et al. A lightweight network for monocular depth estimation with decoupled body and edge supervision [J]. Image and Vision Computing, 2021, 113: 104261.
- [ 12 ] CHENG D, MENG G, XIANG S, et al. FusionNet: Edge aware deep convolutional networks for semantic segmentation of remote sensing harbor images[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2017, 10(12): 5769-5783.
- [ 13 ] CHEN L C, ZHU Y, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation [C]. Proceedings of the European Conference on Computer Vision (ECCV), 2018: 801-818.
- [ 14 ] WANG P, CHEN P, YUAN Y, et al. Understanding Convolution for Semantic Segmentation[A]. 2018 IEEE Winter Conference on Applications of Computer Vision (WACV) [C]. Lake Tahoe, NV: IEEE, 2018: 1451-1460.
- [ 15 ] WANG Z, WANG J, YANG K, et al. Semantic segmentation of high-resolution remote sensing images based on a class feature attention mechanism fused with

- DeepLabv3+[J]. Computers & Geosciences, 2022, 158: 104969.
- [16] YEUNG M, SALA E, SCHÖNLIEB C-B, et al. Unified focal loss: Generalising dice and cross entropy-based losses to handle class imbalanced medical image segmentation [J]. Computerized Medical Imaging and Graphics, 2022, 95: 102026.
- [17] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection [C]. Proceedings of the IEEE International Conference on Computer Vision, 2017: 2980-2988.
- [18] SALEH A, LARADJI I H, KONOVALOV D A, et al. A realistic fish-habitat dataset to evaluate algorithms for underwater visual analysis[J]. Scientific Reports, 2020, 10(1): 14671.
- [19] RADIUK P M. Impact of training set batch size on the performance of convolutional neural networks for diverse datasets [J]. Information Technology and Management Science, 2017.
- [20] ZHAO H, SHI J, QI X, et al. Pyramid Scene Parsing Network [A]. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) [C]. Honolulu, HI: IEEE, 2017: 6230-6239.
- [21] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation [C]. International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham, 2015: 234-241.

## 作者简介



**田志新**, 2018 年于河北工业大学获得学士学位, 现为上海工程技术大学硕士生, 主要研究方向为图像处理、计算机视觉及其应用。

E-mail: 2212720600@qq.com

**Tian Zhixin**, received his B. Sc. degree from Hebei University of Technology in 2018. Now he is a M. Sc. candidate of Shanghai University of Engineering Science. His main research interests include image processing, computer vision and its application.



**徐震**(通信作者) 2006 年于华东师范大学获得学士学位; 2013 年获得上海大学博士学位; 2014 年 1 月到 2016 年 1 月在中科院上海应用物理研究所从事博士后研究; 现为上海工程技术大学讲师, 主要研究方向为数据驱动的机器学习和计算机视觉算法发展及其应用。

E-mail: lcxuzhen@163.com

**Xu Zhen** (Corresponding author) received his B. Sc. degree from East China Normal University in 2006, Ph. D. degree from Shanghai University in 2013. From January 2014 to January 2016, he worked as a postdoctoral researcher in Shanghai Institute of Applied Physics, Chinese Academy of Sciences. Now he is a lecturer at Shanghai University of Engineering Science. His main research interests include the development and application of data-driven machine learning and computer vision algorithms.