

# 基于 RetinaNet 的海洋鱼类检测算法

周迎峰, 张荣芬, 刘宇红\*, 李宽

贵州大学大数据与信息工程学院, 贵州 贵阳 550025

**摘要** 为了更好地保护和利用海洋鱼类资源,需要对海洋鱼类进行有效监测,但海洋环境复杂,导致海洋鱼类的识别检测普遍存在检测精度不佳等问题。针对上述问题,本文提出一种基于 RetinaNet 改进的海洋鱼类检测算法。首先,用 DenseNet-121 替换 RetinaNet 原有的主干网络,减少参数数量的同时保留了更多的鱼类图像特征。然后,在主干网络中引入卷积注意力模块,引导神经网络更有针对性地提取图像特征。其次,在原有的 FPN 网络中引入新的卷积层,使得改进后的 PFPN 网络能够融合更多尺度的图像特征。最后,在分类和回归网络中引入 soft-NMS,有效改善了相同类别的鱼距离过近和相互遮挡造成的漏检问题。实验表明,提出算法的平均精度(mAP)达到 92.12%,相比 SSD 等算法的检测效果有明显提高,相比原算法的 mAP 提升了 4.71%,对于海洋鱼类具有较好的检测效果。

**关键词** 图像处理; 鱼类检测; RetinaNet; DenseNet-121; 卷积注意力模块; soft-NMS

中图分类号 TP391.4

文献标志码 A

DOI: 10.3788/LOP213356

## Marine Fish Detection Algorithm Based on RetinaNet

Zhou Yingfeng, Zhang Rongfen, Liu Yuhong\*, Li Kuan

College of Big Data and Information Engineering, Guizhou University, Guiyang 550025, Guizhou, China

**Abstract** Effectively monitoring marine fish is necessary to protect and utilize marine fish resources. However, the complexity of marine environment leads to low accuracy in marine fish identification and detection. Therefore, this study proposes an improved algorithm for detecting marine fish based on RetinaNet. First, DenseNet-121 was used to replace the original backbone network of RetinaNet, thereby reducing the number of parameters and retaining more fish image features. To guide the neural network to extract image features more pertinently, the convolution attention module was introduced into the backbone network. Second, a new convolution layer was introduced in the original FPN such that the improved PFPN network can fuse more image features with more scales. Finally, soft-NMS was introduced in the classification and regression network to effectively address the detection-missing problem owing to close proximity and mutual occlusion of the same fish species. The experimental results indicate that the average accuracy of the proposed algorithm is 92.12%. This value is significantly improved compared with SSD and other existing algorithms and is 4.71% higher than that of the original algorithm. Thus, the proposed algorithm efficiently identifies and detects marine fish.

**Key words** image processing; fish detection; RetinaNet; DenseNet-121; convolution attention module; soft-NMS

## 1 引言

目标检测是计算机视觉领域中的一个重要的研究方向,其主要用于图像和视频中对目标物体进行定位和识别<sup>[1]</sup>。我国海岸线漫长,海洋资源丰富,近年来对于海洋渔业的过度捕捞让海洋生态遭受到很大程度的破坏<sup>[2]</sup>。将深度学习目标检测技术手段用于观察鱼类栖息地中鱼类种类和数量的变化<sup>[3]</sup>,对于保护和合理利用海洋鱼类资源意义重大<sup>[4]</sup>。

Francisco 等<sup>[5]</sup>提出了一种基于背景减法的计算机视觉方法用于估计玻璃水箱内斑马鱼的数量,这种方法主要侧重于在清澈的鱼缸环境检测特定的鱼类。Christensen 等<sup>[6]</sup>提出了一种卷积神经网络 OFDNet,利用水下摄像机获取的鱼类图像进行鱼类检测,检测精度为 66.7%。Wang 等<sup>[7]</sup>基于 YOLOV2 设计了一个自动分类检测系统,用于对渔船上捕捞的死鱼鱼类图像的分类和检测,这个系统的平均检测精度达到了 91.2%。Xu 等<sup>[8]</sup>基于卷积神经网络 RCNN 提出了一

收稿日期: 2021-12-27; 修回日期: 2022-01-23; 录用日期: 2022-03-02; 网络首发日期: 2022-03-12

基金项目: 贵州省科学技术基金(黔科合基础-ZK[2021]重点 001)

通信作者: \*1459539967@qq.com

种鱼类检测算法,用于网箱中蓝鳍金枪鱼的计数,他们提出的算法对于蓝鳍金枪鱼的识别率达到了 91.5%,但其检测目标仅为蓝鳍金枪鱼,应用范围较窄。以上研究成果对于海洋鱼类检测起了很大的推动作用<sup>[9]</sup>,但其应用环境往往都选择在非自然的海洋环境<sup>[10]</sup>,且他们所用海洋鱼类数据集的种类和样本数量都较少。

基于上述研究,本文对 RetinaNet<sup>[11]</sup>算法进行改进,针对自然环境下拍摄的 16 个类别共 22034 张鱼类图像的数据集,提出了一种精度高、适用范围广的海洋鱼类检测算法。本文算法主要从以下 4 个方面对 RetinaNet 进行改进:首先,为了减少参数量以及提高算法提取图像特征能力,用 DenseNet-121<sup>[12]</sup> 替换 RetinaNet 原有的 ResNet 网络;然后,本文在 DenseNet-121 中引入卷积注意力模块(CBAM)<sup>[13]</sup>,引

导神经网络更有针对性地提取特征;其次,在 FPN 网络中引入新的卷积层重构 FPN 网络,使得改进后的 PFPN 网络能够融合更多尺度的图像特征;最后,在分类和回归网络中用 soft-NMS<sup>[14]</sup> 替换原有的 NMS,以此降低漏检率。

## 2 改进 RetinaNet 算法

### 2.1 本文算法的主要结构

本文算法基于单阶段检测算法 RetinaNet 进行改进<sup>[15]</sup>,本文算法的整体结构如图 1 所示,它由 DenseNet 特征提取模块、PFPN 特征融合模块以及引入 soft-NMS 的分类和回归模块三部分组成。其中,DenseNet 特征提取模块对输入图片进行初步的特征提取<sup>[16]</sup>,PFPN 特征融合模块用于图像特征融合,分类和回归网络用于目标分类和位置回归计算,得到最后的预测结果。

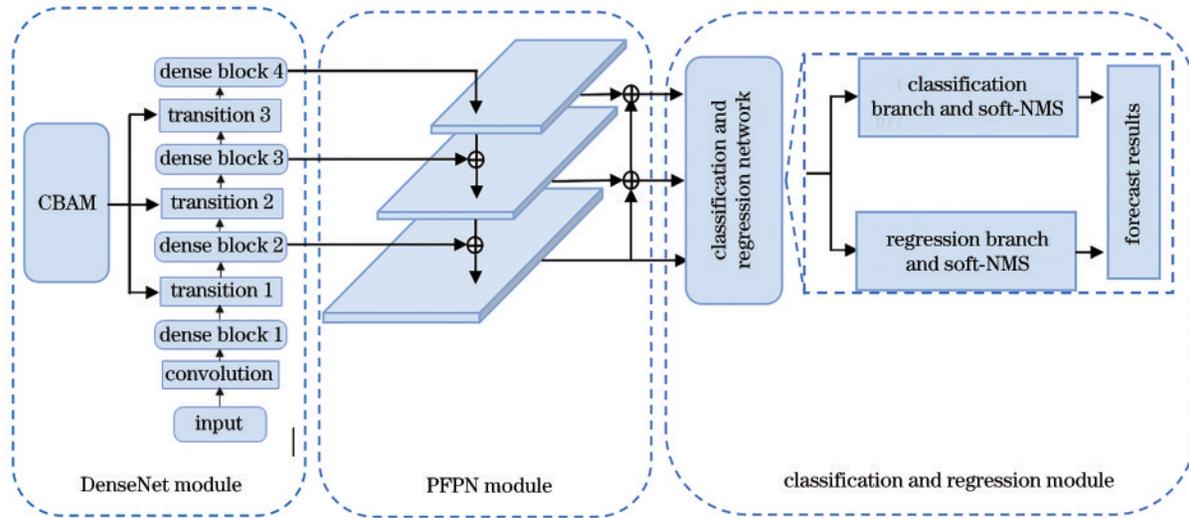


图 1 改进的 RetinaNet 结构

Fig. 1 Structure of improved RetinaNet

### 2.2 主干网络改进

DenseNet 是一种稠密卷积网络,相较于 ResNet, DenseNet 建立了不同层之间的连接,通过这种稠密连接方式减轻了梯度消失问题,极大地减少了网络的参数量,节省了网络计算成本。如图 2 所示,DenseNet 由 dense block 和 transition layer 串联组成。dense block 的结构如图 3 所示,它由多个非线性转换结构组成。卷积层输入  $X_0$ ,通过  $H_1$  卷积层得到输出  $X_1$ ,将  $X_1$  与初始的  $X_0$  进行堆叠,得到一个新的特征层,然后将这个

新的特征层送入  $H_2$  卷积层,得到输出  $X_2$ 。同样, $X_2$  与前面的  $X_1$  和  $X_0$  进行聚合,再送入下一层。重复这个过程, $X_i$  与它前面所有层进行聚合得到模块的最终输出。这种稠密连接方式让 DenseNet 网络更加容易训练,同时加强了传递特征的方式,能够更好地利用每一层的特征。

如图 1 左侧所示,本文采用 DenseNet-121 替换 RetinaNet 原有的 ResNet 网络。DenseNet-121 的结构如表 1 所示。

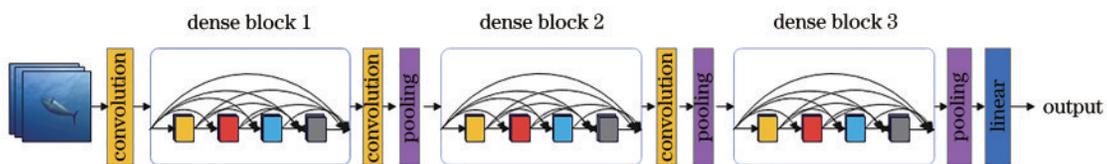


图 2 DenseNet 的结构

Fig. 2 Structure of DenseNet

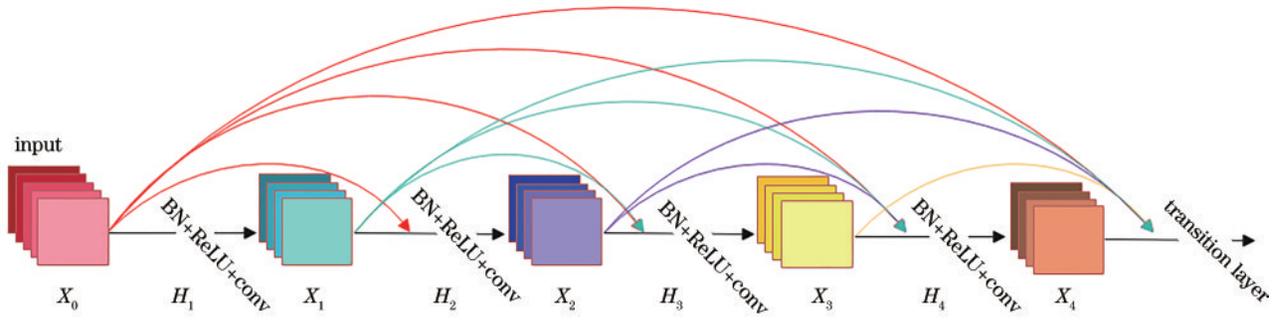


图 3 Dense block 的网络结构  
Fig. 3 Network structure of dense block

表 1 DenseNet-121 的结构  
Table 1 Structure of DenseNet-121

Layer	Setting	Output size
Convolution	7*7 conv, stride 2	112*112
Pooling	3*3 MaxPool, stride 2	56*56
Dense block 1	[1*1 conv, 3*3 conv]*6	56*56
Transition layer 1	1*1 conv, 2*2 AvgPool, stride 2	28*28
Dense block 2	[1*1 conv, 3*3 conv]*12	28*28
Transition layer 2	1*1 conv, 2*2 AvgPool, stride 2	14*14
Dense block 3	[1*1 conv, 3*3 conv]*24	14*14
Transition layer 3	1*1 conv, 2*2 AvgPool, stride 2	7*7
Dense block 4	[1*1 conv, 3*3 conv]*16	7*7
Classification layer	7*7 global AvgPool, 1000D fully-connected, softmax	1*1

### 2.3 引入卷积注意力模块

在卷积神经网络中,每张特征图的侧重点是不同的<sup>[17]</sup>,注意力机制的核心思想是通过一定方法获取到每张特征图重要性的差异,将神经网络的计算资源更多地投入到更重要的任务当中,并利用任务结果反向指导特征图的权重更新,从而高效快速地完成相应任务<sup>[18]</sup>。简而言之,注意力机制的本质就是定位到感兴趣的信息,抑制无用信息。与其他注意力机制不同,如图 4 所示,卷积注意力模块认为,对于卷积网络中的特征图来说,通道和特征图像素点间(通道内部)都蕴含着丰富的注意力信息,而以往的注意力机制只关注了通道上的注意力信息,这对于空间上的注意力信息是

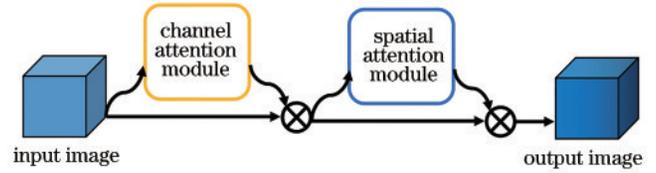


图 4 CBAM 的结构  
Fig. 4 Structure of CBAM

一种浪费<sup>[19]</sup>。CBAM 分为空间注意力模块 (SAM) 和通道注意力模块 (CAM),它将空间和通道两方面的注意力信息通过两个子模块进行汇总,得到更多注意力信息,从而对计算资源进行更合理的分配。

为了进一步选择小目标的和多目标鱼类的重要特征,如图 1 左侧所示,本文在主干网络的每个 transition layer 中加入了卷积注意力模块。CAM 的结构如图 5 所示,本文将输入图像大小设置为 224\*224\*3,经由前面几层神经网络的处理,将经过 1\*1 卷积处理后的特征图 (56\*56\*128) 作为输入特征  $F$ ,首先经过最大池化和平均池化处理得到 1\*1\*128 的两个特征图,然后将它们送入 MLP 层共享参数,接着将 MLP 输出的特征图进行相加和 Sigmoid 激活函数归一化,得到最后的通道注意力特征 (56\*56\*128)。CAM 的运算过程为

$$M_c(F) = \sigma \{ \text{MLP} [ \text{AvgPool}(F) ] + \text{MLP} [ \text{MaxPool}(F) ] \}, \quad (1)$$

式中:  $F$  表示输入的特征图;  $\sigma$  表示 Sigmoid 激活函数; MLP 表示参数共享; AvgPool 表示平均池化; MaxPool 表示最大池化;  $M_c(F)$  表示最后的通道注意力结果。

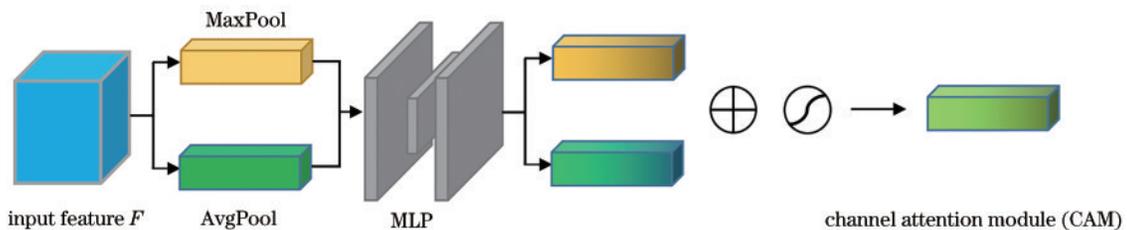


图 5 CAM 的结构  
Fig. 5 Structure of CAM

SAM 的结构如图 6 所示,将通道注意力特征(56\*56\*128)作为输入特征  $F$ ,先经过最大池化和平均池化压缩特征参数量,然后经由卷积压缩特征  $F$  的通道数,最后经过 Sigmoid 激活函数归一化空间权重信息,得到最后的空间注意力特征(56\*56\*128)。

SAM 的运算过程为  $M_s(F) = \sigma \{ f^{7 \times 7} \{ [ \text{AvgPool}(F); \text{MaxPool}(F) ] \} \}$ , (2) 式中:  $F$  表示输入的特征图;  $\sigma$  表示 Sigmoid 激活函数; AvgPool 表示平均池化; MaxPool 表示最大池化;  $M_s(F)$  表示最后的空间注意力结果。

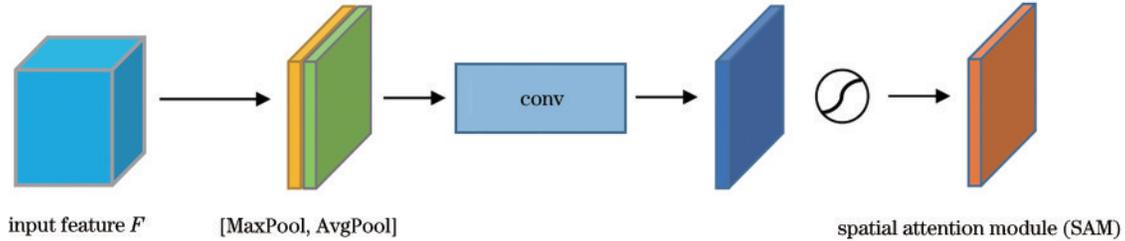


图 6 SAM 的结构  
Fig. 6 Structure of SAM

### 2.4 FPN 网络改进

本文在原有 FPN 网络后添加了自底向上的融合方法对原有的 FPN 网络进行了改进,改进后的 PFPN 网络如图 7 所示。其中, D3、D4、D5 来源于 DenseNet-121 的最后三层输出。D3 首先经过 1\*1 的卷积运算改变了特征图的通道数,再与先经过 1\*1 的卷积运算后

经过上采样的 D4 相加,最后经过 3\*3 的卷积运算生成 P3。D4 首先经过 1\*1 的卷积运算,然后与先经过 1\*1 的卷积运算后经过上采样的 D5 相加,再经过 3\*3 的卷积运算后与经过 3\*3 的卷积运算的 P3 相加生成 P4。重复以上过程得到 P5。D5 经过 3\*3 的卷积运算后得到 P6, P6 经过 3\*3 的卷积运算后得到 P7。

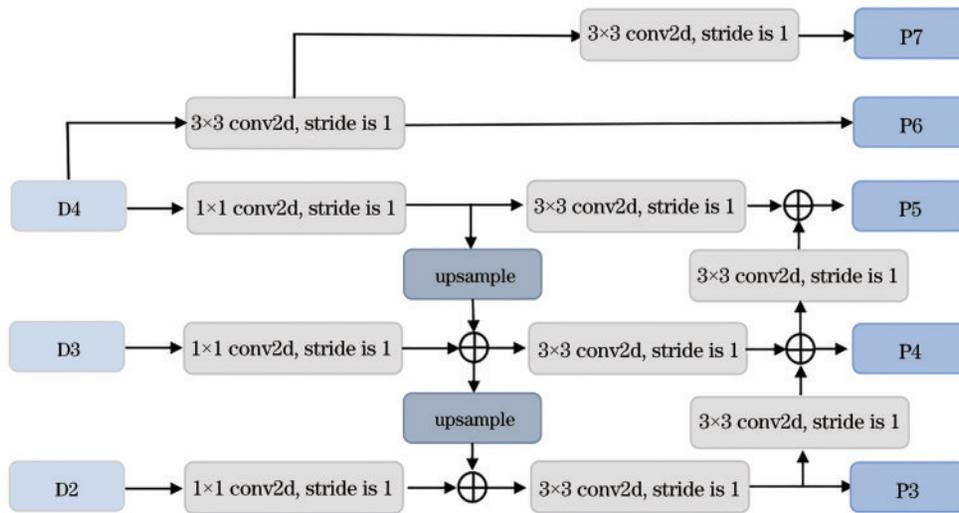


图 7 PFPN 的结构  
Fig. 7 Structure of PFPN

鱼类图像经过 DenseNet-121 特征提取网络后得到不同深度的特征图,深层特征图带有更强的鱼类语义特征信息和较弱的目标位置信息,而浅层的特征图刚好相反。本文改进后的 PFPN 模块首先采用自顶向下的方式将深层特征图与浅层特征图进行叠加融合,达到向浅层特征图传递深层语义特征的目的,然后采用自底向上的方式进一步将深层语义特征和浅层目标位置信息进行融合。这种融合方式充分利用了鱼类图像多尺度特征图的信息,使本文算法能够同时有效地表征各个尺度的目标图像,提高了算法对鱼类图像特征的利用能力。

### 2.5 引入 soft-NMS

本文实验所用的数据集中,有部分图像存在两个及以上同类目标距离过近现象,而按照传统的 NMS 算法对目标进行分类和回归时,这些目标会因为置信度得分低于最大置信度得分而被抑制造成漏检,为了减少这种现象带来的漏检问题,本文算法在分类和回归网络中引入了 soft-NMS,用以提高算法的检测精度和鲁棒性。

非极大值抑制(NMS)是目标检测中非常重要的一部分。针对一个检测对象,NMS 会根据置信度得分对所有检测框进行排序,选择具有最大分数的检测框

作为极大值框,同时根据预定义的置信度阈值抑制具有显著重叠的其余检测框,最后得到预测结果。但是当一张图片中出现两个或多个同类目标重叠时,对于分数低于最高分数的目标会因为抑制而被漏检,为了解决这个问题,在传统的 NMS 中引入连续的置信度抑制函数,根据非极大值框与极大值框重叠 IOU 的大小,对非极大值框的置信度进行不同程度的抑制,这就是 soft-NMS。soft-NMS 有两种表达方式,一种是线性加权的,如式(3)所示,一种是高斯加权的,如式(4)所示。

$$S_i = \begin{cases} S_i, & \text{if } \text{IOU}(A, B_i) < N_i \\ S_i [1 - \text{IOU}(A, B_i)], & \text{if } \text{IOU}(A, B_i) \geq N_i \end{cases}, \quad (3)$$

$$S_i = S_i \exp[-\text{IOU}(A, B_i)^2 / \sigma], \quad (4)$$

式中: $S_i$ 表示保留的预测框结果; $A$ 表示得分最大的预测框; $B_i$ 表示与得分最大预测框相近的预测框; $\text{IOU}$ 表示交并比; $N_i$ 表示预设重叠阈值; $\sigma$ 表示标准差。

### 3 分析与讨论

#### 3.1 数据集

本文所用数据集来源于 ImageNet2012 和 Fish4Knowledge。ImageNet2012 数据集源于 ImageNet 官方在 2012 年举办的大规模视觉识别数据竞赛,其中包含多个类别的鱼类图像。Fish4Knowledge 数据集是从水下摄像机实时拍摄的视频中获取的鱼类数据,共包含 23 个类别 27370 张鱼类图像。为了保证数据集的有效性,本文对 ImageNet2012 和 Fish4Knowledge 中单类别样本太少和图像分辨率过低的鱼类图像进行剔除,然后将未标注的图像采用 labelImg 软件进行标注,整理后得到本文所用的海洋鱼类数据集,本文所用的海洋鱼类数据集一共包含 22034 张图像,16 个种类:

dascyllus reticulatus、plectroglyphidodon dickii、chromis chrysurus、amphiprion clarkii、chaetodon lunulatus、myripristis kuntee、white shark、hammerhead shark、crampfish、eel、rock beauty、anemone fish、sturgeon、gar、lionfish、puffer。训练集和验证集的比例按照 9:1 的比例进行随机划分,部分数据集示例如图 8 所示。

#### 3.2 实验环境

如表 2 所示,本文实验采用的深度学习框架为 Pytorch, CPU 型号为 Intel(R) Core(TM) i7-7800X, GPU 型号为 Nvidia GTX 1080Ti, 操作系统为 Ubuntu16.04, 开发软件使用的 PyCharm。

表 2 实验环境

Table 2 Experimental environment

Category	Parameter
CPU	Intel (R) Core (TM) i7-7800X
GPU	Nvidia GTX 1080Ti
Operating system	Ubuntu16.04
Development software	PyCharm
Development framework	Pytorch

#### 3.3 评价指标

本文采用平均精度(AP)来对目标检测的每一类的检测结果进行评估,并利用 mAP,即所有目标类 AP 的平均值来衡量整个算法的性能。AP( $P_{AP}$ )和 mAP ( $P_{mAP}$ )计算公式分别为

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

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

$$P_{AP} = \int_0^1 P(R) dR, \quad (7)$$



图 8 数据集样例示图

Fig. 8 Sample diagram of dataset

$$P_{mAP} = \frac{1}{n} \sum_{i=1}^n AP_i, \quad (8)$$

式中： $n$ 表示检测目标类别数； $R$ 表示召回率，表示样本中的正样本有多少被预测正确； $P$ 为精确度，表示预测为正的样本中有多少是真正的正样本； $S_{TP}$ 为被预测正确的正样本的数量； $S_{FN}$ 为把正样本错误地预测为负的数量； $S_{FP}$ 表示把负样本错误地预测为正的数量； $S_{TP} + S_{FN}$ 为全部正样本数量； $S_{TP} + S_{FP}$ 为全部被分为正样本的数量。

### 3.4 训练过程

本文实验中的输入图像统一设置为 224\*224，同时为了保证训练效果，增强网络算法的鲁棒性，在将图像输入到网络训练之前对图像进行统一的增强处理。为了避免训练过程权重太过随机影响训练效果，本文采用迁移学习的方法训练新模型，首先将 PASCAL VOC 2007 和 PASCAL VOC 2012 数据集在本文算法上训练权重，然后将得到的权重载入本文算法网络进行实验。本文实验的 Batch\_size 设置为 16，初始学习率设置为 0.0001，轮次设置为 150，采用 Adam 优化算法对算法参数进行优化，采用 StepLR 机制动态调整学习率，调整轮次设置为 1，调整参数设置为 0.96。训练效果如图 9 所示，将 RetinaNet 的主干网络替换为 DenseNet-121 后，训练的 Loss 值下降更快，在训练轮次为 130 次之后趋于稳定，充分证明替换主干网络为 DenseNet-121 有效提高了算法训练的收敛速度和收敛效果。

### 3.5 检测效果对比

为了验证本文算法的有效性，将 CenterNet、YOLOV4、Efficientdet、YOLOV3、SSD 和本文提出的算法在同一数据集下进行了对比实验，同时与其他文献对于鱼类检测的精度进行了对比，结果如表 3 所示。

由表 3 可知，改进后的算法精度达到 92.78%，mAP 达到 92.12%。相比于 CenterNet、YOLOV4、Efficientdet、YOLOV3、SSD 等算法，本文算法精度分别提高了 8.13 百分点、10.65 百分点、6.06 百分点、5.19 百分点、8.11 百分点，mAP 分别提高了 7.23 百分

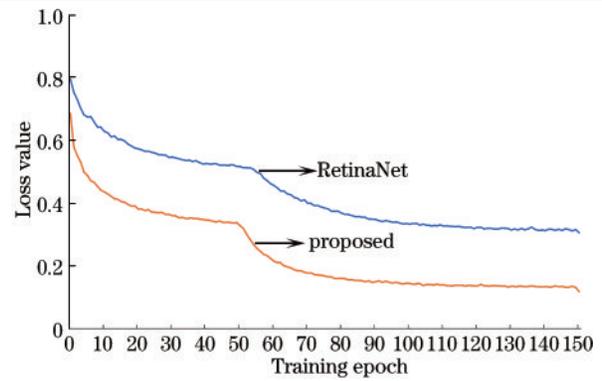


图 9 Loss 值变化曲线

Fig. 9 Loss value change curve

表 3 不同算法精度对比

Table 3 Precision comparison of different algorithms

Algorithm	Precision / %	mAP / %
CenterNet	84.65	84.89
YOLOV4	82.13	85.77
Efficientdet	86.72	87.51
YOLOV3	87.59	85.79
SSD	84.67	86.11
Ref. [20]	-	85.99
Ref. [21]	-	89.00
Proposed	92.78	92.12

点、6.35 百分点、4.61 百分点、6.33 百分点、6.01 百分点。相比于文献[20]和文献[21]，本文算法的 mAP 分别提高了 6.13 百分点和 3.12 百分点。综上所述，在同一数据集下，本文提出的算法对比其他几种算法在精度和 mAP 上提升明显。而与其他文献对于海洋鱼类的检测精度进行对比，结果显示本文算法的效果更佳，表明本文提出的方法在鱼类检测上有一定的优越性。

### 3.6 消融实验

为了充分验证本文算法的检测性能，如表 4 所示，本文以未改进的 RetinaNet 算法为参照，通过多个模块组合的方式进行了 8 组实验来验证各模块的有效性。

表 4 消融实验

Table 4 Ablation experiment

Group	Method	mAP / %	Speed / (ms · image <sup>-1</sup> )	Memory / MB	Model size / MB
0	RetinaNet	87.41	89	4216	242.0
1	DenseNet-121	88.75	46	3247	165.1
2	DenseNet-121+CBAM	91.21	51	3358	168.5
3	DenseNet-121+PFPN	90.53	50	3471	167.7
4	DenseNet-121+soft-NMS	88.95	48	3201	165.5
5	DenseNet-121+CBAM+PFPN	91.75	54	3411	168.7
6	DenseNet-121+CBAM+soft-NMS	91.62	53	3315	168.9
7	DenseNet-121+PFPN+soft-NMS	90.94	52	3398	168.4
8	Proposed	92.12	56	3408	166.6

从表 4 可知,在第一组实验中,将 RetinaNet 的主干网络替换为 DenseNet-121 后的模型 mAP 为 88.75%,相较于替换前 mAP 提升了 1.34 百分点,在检测速度、显存占用以及模型大小上也有大幅改善,这表明 DenseNet-121 不仅能够提取更多的图像特征,对于整个模型的性能也有较大提升。然后在第一组实验的基础上,在主干网络中分别加入 CBAM、PFPN 和 soft-NMS 进行了三组实验,由实验结果可知,分别加入三个模块后的检测精度都有不同幅度的提升,其中加入 CBAM 后的性能提升最明显,其 mAP 达到 91.21%,相较于第一组实验 mAP 提升了 2.46 百分点,充分表明加入 CBAM 之后能够将计算资源更多地用于鱼类图像检测。接着,在第一组实验的基础上,从 CBAM、soft-NMS 和 PFPN 中分三次选取两个不同的模块分别加入网络中进行第五、六、七组实验,由实验结果可知,加入两个不同模块后的检测效果均有提升,其中,加入 CBAM 和 PFPN 后 mAP 达到 91.75%,相较于第一组实验 mAP 提升了 3.00 百分点,加入 CBAM 和 soft-NMS 后 mAP 达到 91.62%,相

较于第一组实验 mAP 提升了 2.87 百分点,加入 PFPN 和 soft-NMS 后 mAP 达到 90.94%,相较于第一组实验 mAP 提升了 2.19 百分点,表明 PFPN 模块能够有效融合多尺度鱼类图像特征图的信息,在分类和回归网络中引入 soft-NMS 能够有效减小漏检率。第八组实验是在第一组实验的基础上加入 CBAM、PFPN 和 soft-NMS,也就是本文所提出的算法,实验结果显示本文算法的 mAP 为 92.12%,对比第一组实验精度提升了 3.37 百分点,相比原算法 mAP 提升了 4.71 百分点(如图 10 所示,本文算法在各个类别上相较于原算法也有明显提升),充分说明这种组合方式的检测效果最好。此外,对比第一组实验和后续七组实验的检测速度、显存占用以及模型大小,都因为参数数量的增加略有下降,但对比原算法都有较大提升。总而言之,在 RetinaNet 中引入 DenseNet-121、CBAM、PFPN 和 soft-NMS 四个模块后,算法整体的检测精度、检测速度、显存占用以及模型大小都有较大改善,充分证明了所加模块对于算法整体的性能有显著提升。

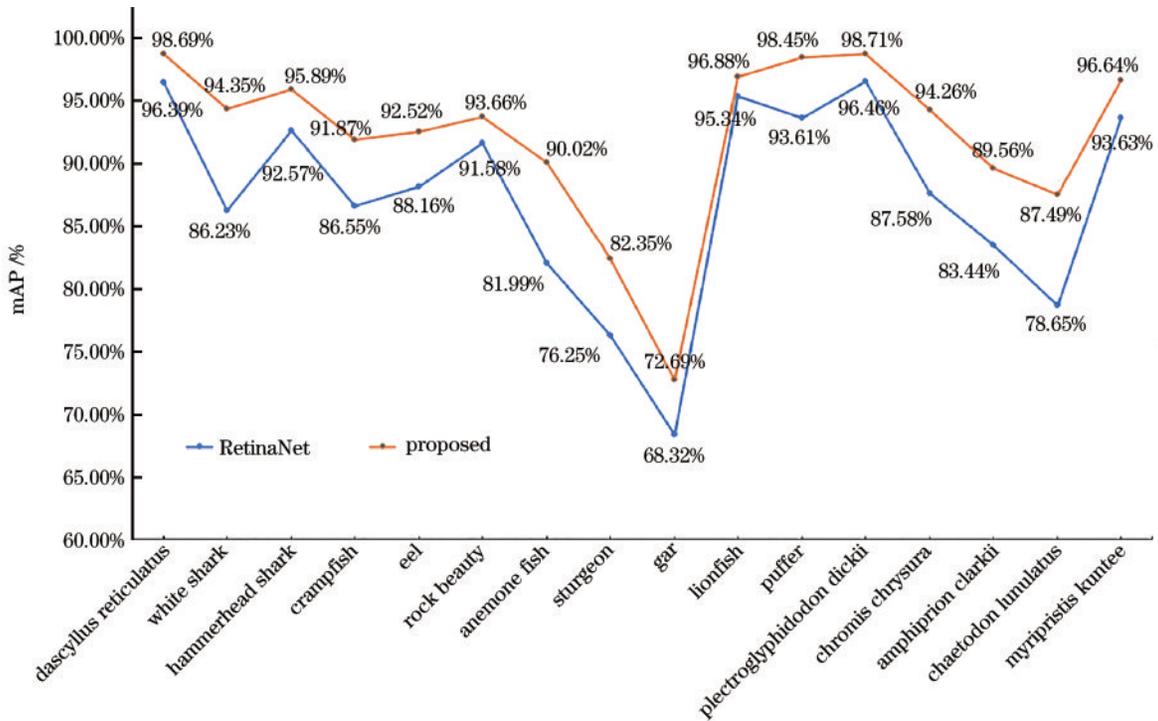


图 10 各类别精度对比

Fig. 10 Comparison of accuracy of each category

### 3.7 算法检测效果展示

本文算法的检测效果如图 11 所示,从图中可以看出,改进后的算法对比原算法有以下明显优势:1)对于小目标鱼类检测效果提升明显;2)对于相互遮挡的同类目标也有较好的检测效果;3)对于鱼类图像整体的检测精度有明显提高。

## 4 结 论

提出了一种基于 RetinaNet 改进的海洋鱼类检测算法。网络改进部分,首先,将 RetinaNet 的主干网络替换为 DenseNet-121 用于提取图像特征,改进后的主干网络参数量更少,能够保留更多的图像特征。然后,

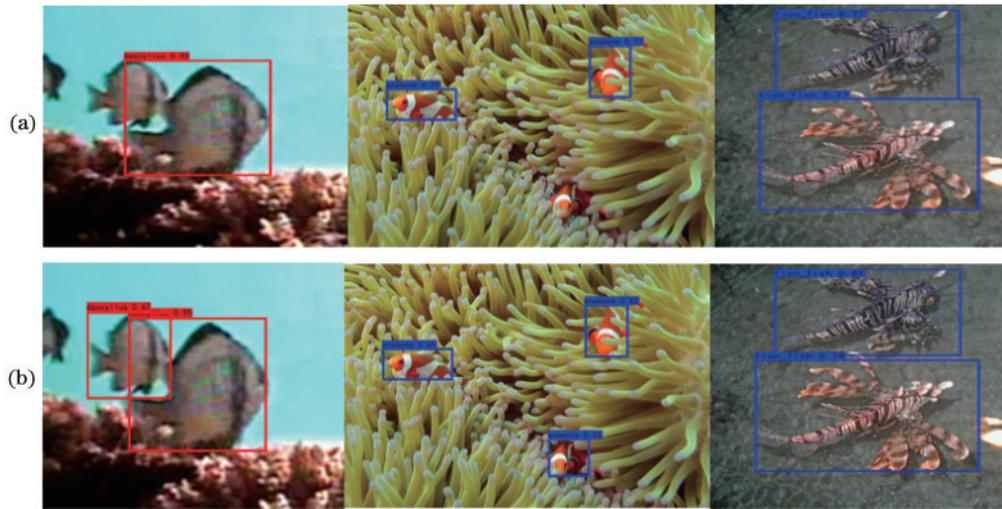


图 11 检测效果对比。(a)原算法检测结果;(b)改进算法检测结果

Fig.11 Comparison of detection results. (a) Detection result of original algorithm; (b) detection result of improved algorithm

在DenseNet-121中引入CBAM,将更多的计算资源用于鱼类图像特征提取。其次,在原有的FPN网络后添加了自底向上的融合方法,进一步将深层语义特征和浅层目标位置信息进行融合。最后,将分类和回归网络中原有的NMS替换为soft-NMS,有效降低了漏检率。实验部分,以mAP为主要衡量指标和其他算法进行对比,实验结果显示提出的算法精度更佳。此外,通过递进式消融实验进一步展示了所加模块对于鱼类检测的有效性,结果表明,本文改进的算法mAP更高,检测效果更好,为海洋鱼类的识别检测和渔业资源保护提供了可靠的技术支撑。

### 参 考 文 献

- [1] 王凤随,王启胜,陈金刚,等.基于注意力机制和Soft-NMS的改进Faster R-CNN目标检测算法[J].激光与光电子学进展,2021,58(24):2420001.  
Wang F S, Wang Q S, Chen J G, et al. Improved faster R-CNN target detection algorithm based on attention mechanism and soft-NMS[J]. Laser & Optoelectronics Progress, 2021, 58(24): 2420001.
- [2] 李冲冲.基于YOLOv3的水下鱼类目标的检测与识别[D].杨凌:西北农林科技大学,2020.  
Li C C. Detection and recognition of underwater fish targets based on YOLOv3[D]. Yangling: Northwest A & F University, 2020.
- [3] Li X, Shang M, Hao J, et al. Accelerating fish detection and recognition by sharing CNNs with objectness learning [C]//OCEANS 2016-Shanghai, April 10-13, 2016, Shanghai, China. New York: IEEE Press, 2016.
- [4] Boom B, Huang P, Beyan C, et al. Long-term underwater camera surveillance for monitoring and analysis of fish populations[C]//Workshop on Visual observation and Analysis of Animal and Insect Behavior (VAIB) in Conjunction with ICPR 2012, January 1, 2012, Tsukuba, Japan. Red Hook: Curran Associates, Inc., 2012.
- [5] Silvério F J, Certal A C, de Ferro C M, et al. Automatic system for zebrafish counting in fish facility tanks[M]//Campilho A, Karray F. Image analysis and recognition. Lecture notes in computer science. Cham: Springer, 2016, 9730: 774-782.
- [6] Christensen J H, Mogensen L V, Galeazzi R, et al. Detection, localization and classification of fish and fish species in poor conditions using convolutional neural networks[C]//2018 IEEE/OES Autonomous Underwater Vehicle Workshop (AUV), November 6-9, 2018, Porto, Portugal. New York: IEEE Press, 2018.
- [7] Wang M F, Liu M Y, Zhang F H, et al. Fast classification and detection of fish images with YOLOv2 [C]//2018 OCEANS-MTS/IEEE Kobe Techno-Oceans, May 28-31, 2018, Kobe, Japan. New York: IEEE Press, 2018.
- [8] Xu G, Chen Q C, Yoshida T, et al. Detection of bluefin tuna by cascade classifier and deep learning for monitoring fish resources[C]//Global Oceans 2020: Singapore-U. S. Gulf Coast, October 5-30, 2020, Biloxi, MS, USA. New York: IEEE Press, 2020.
- [9] 邓权.基于深度学习的海洋鱼类检测技术研究[D].济南:山东大学,2021.  
Deng Q. Research on ocean fish detection technology based on deep learning[D]. Jinan: Shandong University, 2021.
- [10] Almero V J, Concepcion R, Rosales M, et al. An aquaculture-based binary classifier for fish detection using multilayer artificial neural network[C]//2019 IEEE 11th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM), November 29-December 1, 2019, Laoag, Philippines. New York: IEEE Press, 2019.
- [11] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection[C]//2017 IEEE International Conference on Computer Vision, October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 2999-3007.
- [12] Huang G, Liu Z, van der Maaten L, et al. Densely connected convolutional networks[C]//2017 IEEE

- Conference on Computer Vision and Pattern Recognition, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 2261-2269.
- [13] Woo S, Park J, Lee J Y, et al. CBAM: convolutional block attention module[EB/OL]. (2018-07-17)[2022-02-04]. <https://arxiv.org/abs/1807.06521>.
- [14] Bodla N, Singh B, Chellappa R, et al. Soft-NMS: improving object detection with one line of code[C]// 2017 IEEE International Conference on Computer Vision, October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 5562-5570.
- [15] 于敏, 屈丹, 司念文. 改进型 RetinaNet 的目标检测算法[J/OL]. 计算机工程: 1-12[2021-12-09]. DOI: [10.19678/j.issn.1000-3428.0062134](https://doi.org/10.19678/j.issn.1000-3428.0062134).  
Yu M, Qu D, Si N W. Target detection algorithm of improved RetinaNet[J/OL]. Computer Engineering: 1-12 [2021-12-09]. DOI: [10.19678/j.issn.1000-3428.0062134](https://doi.org/10.19678/j.issn.1000-3428.0062134).
- [16] 刘广进, 王光辉, 毕卫华, 等. 基于 DenseNet 与注意力机制的遥感影像云检测算法[J]. 自然资源遥感, 2022, 34(2): 88-96.  
Liu G J, Wang G H, Bi W H, et al. Cloud detection algorithm of remote sensing image based on DenseNet and attention mechanism[J]. Remote Sensing for Natural Resources, 2022, 34(2): 88-96.
- [17] Kim J, Yu S C. Convolutional neural network-based real-time ROV detection using forward-looking sonar image [C]//2016 IEEE/OES Autonomous Underwater Vehicles (AUV), November 6-9, 2016, Tokyo, Japan. New York: IEEE Press, 2016: 396-400.
- [18] 张宸嘉, 朱磊, 俞璐. 卷积神经网络中的注意力机制综述[J]. 计算机工程与应用, 2021, 57(20): 64-72.  
Zhang C J, Zhu L, Yu L. Review of attention mechanism in convolutional neural networks[J]. Computer Engineering and Applications, 2021, 57(20): 64-72.
- [19] 刘嘉敏, 谢文杰, 黄鸿, 等. 基于空间和通道注意力机制的目标跟踪方法[J]. 电子与信息学报, 2021, 43(9): 2569-2576.  
Liu J M, Xie W J, Huang H, et al. Spatial and channel attention mechanism method for object tracking[J]. Journal of Electronics & Information Technology, 2021, 43(9): 2569-2576.
- [20] 王亚茹, 刘雨青, 黄璐瑶. 基于 FML-Centernet 算法的鱼类识别检测方法[J]. 激光与光电子学进展, 2022, 59(16): 1615002.  
Wang Y R, Liu Y Q, Huang L Y. Fish identification and detection method based on FML-Centernet algorithm[J]. Laser & Optoelectronics Progress, 2022, 59(16): 1615002.
- [21] 李庆忠, 李宜兵, 牛炯. 基于改进 YOLO 和迁移学习的水下鱼类目标实时检测[J]. 模式识别与人工智能, 2019, 32(3): 193-203.  
Li Q Z, Li Y B, Niu J. Real-time detection of underwater fish based on improved YOLO and transfer learning[J]. Pattern Recognition and Artificial Intelligence, 2019, 32(3): 193-203.