

改进 FCOS 网络的海洋鱼类目标检测^①



张琳, 葛艳, 杜军威, 刘玉鹏

(青岛科技大学 信息科学技术学院, 青岛 266061)

通信作者: 葛艳, E-mail: geyan@qust.edu.cn

摘要: 鱼类的探索与保护是保持海洋生态环境平衡的重要一环, 然而水下环境复杂, 受光照、水质以及遮挡物的影响, 造成水下捕捉鱼类图像成像模糊识别困难, 制约水下鱼类目标的检测速度以及检测精度. 针对以上问题, 提出了一种基于改进 FCOS 的海洋鱼类识别模型. 首先, 该模型以一阶段算法 FCOS 为基本架构, 使用轻量级的 MobileNetv2 作为骨干网络, 既保证检测准确度, 还可以提高检测; 其次, 引入自适应空间特征融合 (adaptively spatial feature fusion, ASFF) 模块, 避免尺度特征的不一致性, 提高检测精度; 最后, 将 center-ness 分支引入到回归分支中, 引入联合交并比损失 (GIoU loss, generalized intersection over union) 提高检测的性能. 实验数据集使用公开数据集 Fish4Knowledge (F4K) 中的图片以及视频帧截取图片, 选取训练性能最优模型进行评估. 结果表明, 提出的新模型在以上数据集的平均检测精度分别为 99.79%、99.88%, 相较于原模型以及其他检测模型本文提出模型的检测精度与识别速度更高, 可为海洋鱼类识别提供参考依据.

关键词: 鱼类识别; 目标检测; FCOS 网络; 特征融合; MobileNetv2; 深度学习

引用格式: 张琳, 葛艳, 杜军威, 刘玉鹏. 改进 FCOS 网络的海洋鱼类目标检测. 计算机系统应用, 2023, 32(3): 309-315. <http://www.c-s-a.org.cn/1003-3254/8965.html>

Improved FCOS Network for Marine Fish Target Detection

ZHANG Lin, GE Yan, DU Jun-Wei, LIU Yu-Peng

(School of Information Science and Technology, Qingdao University of Science and Technology, Qingdao 266061, China)

Abstract: Exploring and protecting fish is an important part of maintaining the balance of the marine ecological environment. However, the complex underwater environment affected by light, water quality, and occlusions makes it difficult to identify blurred fish images captured underwater and consequently restricts the speed and accuracy of underwater fish target detection. To solve the above problem, this study proposes a marine fish identification model based on improved fully convolutional one-stage object detection (FCOS). Specifically, the model takes the one-stage FCOS algorithm as the basic structure and uses the lightweight MobileNetv2 as the backbone network, which not only ensures the detection accuracy but also improves the detection; then, an adaptive spatial feature fusion (ASFF) module is introduced to avoid the inconsistency in scale features and improve detection accuracy; finally, the center-ness branch is introduced into the regression branch, and the generalized intersection over union (GIoU) loss is introduced to improve detection performance. Regarding the experimental dataset, the pictures in the public dataset Fish4Knowledge (F4K) and video frame screenshots are utilized, and the model with the optimal training performance is selected for evaluation. The results show that the average detection accuracy of the proposed new model on the above datasets is 99.79% and 99.88%, respectively. Compared with the original model and other detection models, the proposed model provides higher detection accuracy and identification speed. The model in this study can provide a reference for marine fish identification.

Key words: fish identification; target detection; fully convolutional one-stage object detection (FCOS) network; feature fusion; MobileNetv2; deep learning

① 基金项目: 山东省自然科学基金 (ZR2021MF092)

收稿时间: 2022-07-24; 修改时间: 2022-08-26; 采用时间: 2022-09-07; csa 在线出版时间: 2022-11-04

CNKI 网络首发时间: 2022-11-15

对鱼类进行检测识别是对海洋生物进行了解及保护的重要基础。但由于海洋环境的复杂,光照不均匀、水中杂质、遮挡物多等因素,造成采集到的鱼类图像能见度、对比度较低,加大了识别的难度,容易造成误检情况。鱼类的准确定位和高效识别是实现水下智能作业的关键。

传统的鱼类目标检测流程是先采用滑动窗口对图像遍历产生候选区,使用 SIFT、HOG 进行特征提取与相应的分类器 SVM 等结合综合判断进行分类。滑动窗口为区域选择策略的基础是此类方法的关键难点,导致了时间复杂度高和针对性低;SIFT、HOG 等传统算子对多样性的变化适应性不强,导致特征鲁棒性差。

最近几年,人们在计算机视觉领域取得了累累硕果,深度学习方面技术提高颇多,鱼类目标检测也开始广泛应用卷积神经网络(CNN),这些都是来源于 CNN 出色的特征提取能力。当前基于 CNN 的目标检测算法主要分为两大类,一类是 RCNN^[1]、Faster RCNN^[2] 为代表的双阶段(two-stage)目标检测算法,一类是 YOLO^[3]、SSD^[4]、Retina-Net^[5] 为代表的单阶段(one-stage)目标检测算法。Two-stage 算法首先用到了启发式方法(selective search)或者使用 RPN 结构产生一系列的候选框,下一步通过使用 CNN 对样本做回归和分类,因此这种方法检测的准确率与其定位精度较高,但需要牺牲检测速度为代价;而 one-stage 算法则能够将目标边框定位问题直接转变为回归问题并进行处理,由于略去了候选框,因此这种方法检测速度较快,但检测准确度会比 two-stage 方法略逊一筹。

Alsmadi 等^[6]使用反向传播神经网络、决策树算法以及灰度共现矩阵对 20 类鱼识别分类,在其特定的测试集上取得了 84% 的准确率。李庆忠等^[7]将改进的 YOLO 算法和迁移学习结合出一种 Underwater-YOLO 网络模型,针对小目标和位置重叠鱼类进行有效检测,准确率达 93%。王亚茹等^[8]采用改进的 FML-Centernet 网络模型进行鱼类识别检测,识别平均精度 AP50 可达到 85.99%,平均检测时间在 100 ms 以内。覃学标等^[9]提出了增强混合的鱼类检测方法,有效地改善了模糊水下图像造成的精度退化问题。

通常我们用到的双阶段目标检测算法以及一些单阶段目标检测算法,例如 YOLO 算法,其工作原理是先于原图片中预定义锚框(Anchor),再通过神经网络进行一系列调整,最后预测的目标结果源于以上操作得

到的边界框。此类方法参数多计算繁琐,因此识别效率低。而基于无锚框(Anchor-free)的算法如 CenterNet^[10]、FCOS^[11]、CornerNet^[12],通过 CNN 提取特征之后,取消预先定义锚框,直接在经过卷积得到的特征图(feature map)上进行预测,降低检测的复杂度,既提高检测速度又能保证精度。

FCOS (fully convolutional one-stage object detection) 是一种基于无锚框的全卷积、单阶段逐像素目标检测算法,总的来说有以下几点优势:超参数减少,使得训练更容易,速度有所提高;无锚框可以不用计算 GT boxes 与 anchors 之间的 *IoU*,内存和计算量减少;FCOS 还可以作为 two-stage 检测器的 RPN。FCOS 不仅保证了算法的检测精度且减少了算法复杂度,可以很方便地扩展去解决其他任务,比如鱼类检测识别。本文针对水下鱼类图像检测的难点问题,提出通过 FCOS 网络对鱼类进行分类检测,通过使用 MobileNetv2 轻量级模型作为骨干网络减少计算量,提升模型推理速度满足实时性;同时在 FPN 上引入 ASFF 避免尺度特征的不一致性,提高检测精度;对损失函数进行改进,针对梯度消失问题,使用 *GIOU* 损失来解决。

1 改进 FCOS 的鱼类目标检测模型

FCOS 网络模型主要分为 3 个部分:第 1 部分骨干网络(Backbone)、第 2 部分特征金字塔(feature pyramid networks, FPN)和第 3 部分检测头(Head),其中回归、中心度和分类 3 个分支构成了检测头部分。

1.1 改进 FCOS 网络的检测模型设计

FCOS 默认使用 ResNet-50^[13] 作为网络的骨干,本文在特征提取方面,引入深度可分离卷积,骨干网络使用 MobileNetv2^[14],并且也保证了网络的检测精度,相较于 ResNet-50,使用 MobileNetv2 可以压缩模型,同时降低了参数量和计算量,模型推理速度会有所提升,从而提高对鱼类实时检测性;通过骨干网络输出的特征图经过特征金字塔处理得到 5 层新的特征图,对下层的 3 层使用 ASFF 共享语义特征;将 5 层特征输入到共享的检测头中,将 center-ness 分支引入到回归分支,在检测头部分进行回归分类。改进后的 FCOS 网络模型如图 1 所示。

首先对输入的图片进行预处理操作,输入到 Backbone 网络中获得特征图(feature map) C3、C4 和 C5;在 FPN 部分,特征图 P5 直接由 C5 经过 1×1 卷积生

成, P4 和 P3 分别由 C4 和 C3 经 1×1 卷积上采样与 P5 和 P4 相加融合得到, P6 和 P7 是特征图 P5 进行 3×3 卷积的 2 倍下采样得到, 该方法既减少了参数数量, 又增强了特征图语义, 因此最终效果也有所提高. 最终得到了 5 个级别的特征映射 {P3, P4, P5, P6, P7},

下采样步长 stride 分别为 8、16、32、64、128, 即通过 FPN, 将大小不一的对象被不同级别的特征图检测出来. 对 P3、P4、P5 融合 ASFF, 得到 ASFF-1, ASFF-2, ASFF-3, 将 {ASFF-1, ASFF-2, ASFF-3, P6, P7} 这 5 层特征输入到共享检测头中.

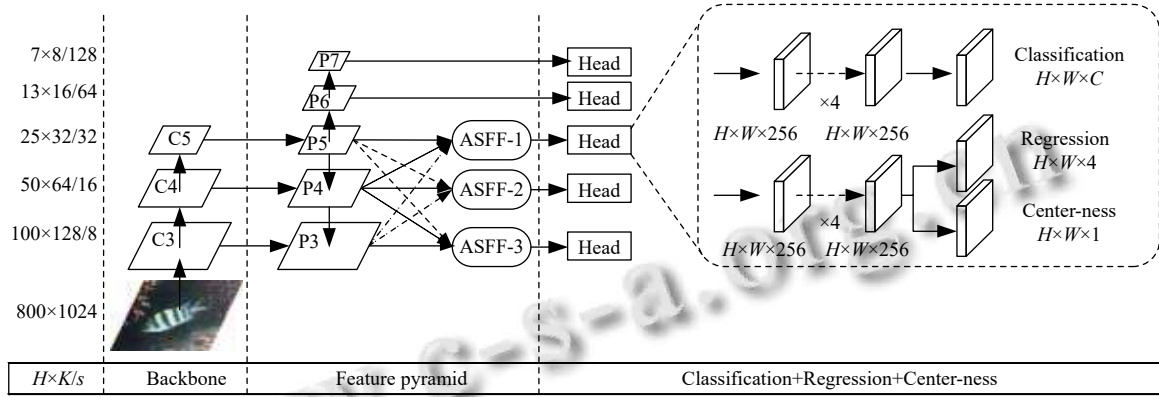


图 1 改进后 FCOS 网络

1.2 深度可分离卷积

轻量级网络 MobileNet 系列最为重要的特点就是深度可分离卷积, 主要分为两部分操作: 逐深度卷积 (depthwise convolution, DW) 和逐点卷积 (pointwise convolution, PW). DW 卷积将卷积核进行拆分, 形成单个 channel, 并且在对每一个 channel 进行卷积操作之后, 得到的特征矩阵深度与输入特征矩阵的深度相同. PW 卷积和标准卷积相同, 但卷积核大小为 1. DW 卷积和 PW 卷积如图 2 所示.

由此可见, 标准卷积计算量是深度可分离卷积的 8-9 倍, 深度可分离卷积在极大程度减少模型参数的同时, 又显著加快了计算速度.

1.3 自适应空间特征融合

金字塔特征表示常用来解决目标检测当中目标尺度变化, 特征金字塔主要是引入了一条自上而下的路径, 对低分辨率和语义丰富的特征进行上采样, 并与网络骨干中输出的高分辨率和语义稀少的特征进行横向结合, 最终形成了一个在各层次上语义共享的特征金字塔. 但对 one-stage 检测器而言, 在内部不同特征尺度之间不一致性是其主要弊端. 为了解决这一问题, 本文融入自适应空间特征融合 (ASFF), ASFF 学习了空间滤波冲突信息以抑制不一致性的方法, 针对有利用价值的信息保留后再加以组合, 提高特征尺度的不变性, 并且引入的计算开销很小, 实现简单. ASFF^[15] 网络结构如图 3 所示.

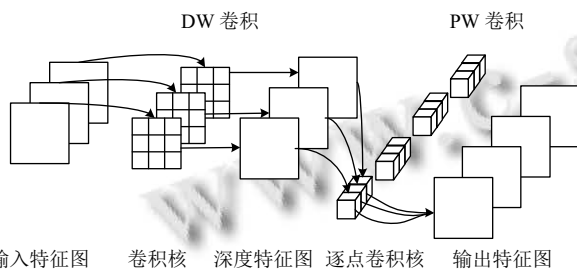


图 2 深度可分离卷积示意图

ASFF 思想的实现可分为两步: 恒等缩放与自适应融合. 以 ASFF-3 为例, 融合后的特征 ASFF-3 为 level1-level3 三个尺度图的特征与融合的权重参数 α , β 和 γ 相乘再相加的结果, 权重自适应调整, 这样对学习不同特征尺度在预测特征图的贡献度方面效果更佳. 最终得到第 1 层特征图的式 (2) 为:

标准卷积的计算量为: $D_K \times D_K \times M \times N \times D_F \times D_F$; D_F 为输入特征矩阵的高和宽, D_K 为卷积核大小, M 为输入特征矩阵的深度, N 为输出特征矩阵的深度.

DW 卷积+PW 卷积的计算量为式 (1):

$$D_K \times D_K \times M \times D_F \times D_F + M \times N \times D_F \times D_F = \frac{1}{N} + \frac{1}{D_K^2} \quad (1)$$

$$y_{ij}^l = \alpha_{ij}^l \cdot \mathbf{x}_{ij}^{1 \rightarrow l} + \beta_{ij}^l \cdot \mathbf{x}_{ij}^{2 \rightarrow l} + \gamma_{ij}^l \cdot \mathbf{x}_{ij}^{3 \rightarrow l} \quad (2)$$

其中, y_{ij}^l 为最终输出特征图对应位置的特征向量; $\mathbf{x}_{ij}^{n \rightarrow l}$

为融合前的第 n 层特征图中位置 (x,y) 的特征向量; α_{ij}^l 、 β_{ij}^l 、 γ_{ij}^l 表示在第 l 层特征图通过学习得到的权重参数, 并且 α_{ij}^l 、 β_{ij}^l 、 γ_{ij}^l 经过拼接之后再通过 Softmax 函数使其范围在 $[0, 1]$ 之间, 规定三者满足以下关系式:

$$\alpha_{ij}^l + \beta_{ij}^l + \gamma_{ij}^l = 1 \quad (3)$$

$$\alpha_{ij}^l = \frac{e^{\lambda_{\alpha_{ij}^l}}}{e^{\lambda_{\alpha_{ij}^l}} + e^{\lambda_{\beta_{ij}^l}} + e^{\lambda_{\gamma_{ij}^l}}} \quad (4)$$

λ^l 为一个 1×1 卷积分别作用在 \mathbf{x}^{n-l} 上得到, 可以通过 BP 算法进行优化学习.

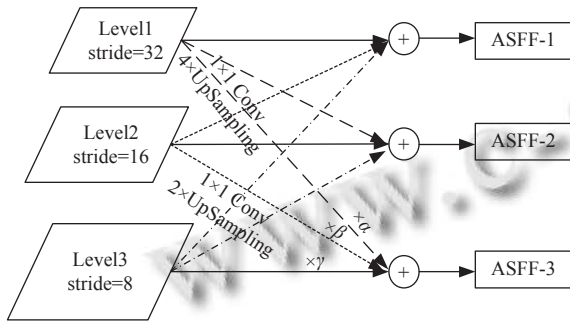


图3 ASFF 网络结构

1.4 损失函数

FCOS 损失函数由 3 部分组成 L_{cls} 、 L_{reg} 、 $L_{center-ness}$. 其中, L_{cls} 采用焦点损失 (focal loss), L_{reg} 采用 IoU loss (L_{IoU}), $L_{center-ness}$ 采用 BCE loss.

1.4.1 中心度损失

低质量预测框通常是由远离物体中心的像素点产生, 中心度 *center-ness* 是为了降低其得分, 使经过 NMS 后更靠近鱼类中心的像素点生成的高得分框排在前面. 中心度表示从该位置到该位置负责的对象中心的标准化距离, t^* 、 r^* 、 b^* 分别表示对于输入图像中的一个像素点位置距离真实框 4 条边的值, 中心度目标被定义为式 (5):

$$center-ness^* = \sqrt{\frac{\min(t^*, r^*)}{\max(t^*, r^*)} \times \frac{\min(t^*, b^*)}{\max(t^*, b^*)}} \quad (5)$$

中心度取值范围从 0 到 1, 位置点越靠近真实框中心, *center-ness* 的预测目标越接近 1, 因此使用二进制交叉熵 (BCE) 损失进行训练. 中心度的作用是降低距离对象中心较远的边界框分数的权重, 所以在高概率情况下, 非最大抑制 (NMS) 会将这些低质量的边界框给过滤掉, 检测性能进而得到显著提高.

1.4.2 分类损失

在鱼类数据集中, 有的鱼数量较少, focal loss 可以

在不影响检测速度的情况下, 用来解决鱼类目标检测当中样本比例严重不平衡的问题, 降低简单样本的内部加权, 从而更关注难分类的鱼类样本. 分类损失 L_{cls} 公式如下:

$$L_{cls}(p_{x,y}, p_{x,y}^*) = -\alpha_t(1 - p_{x,y}^t)^\gamma \log(p_{x,y}^t) \quad (6)$$

其中, α_t 负责调整正负样本之间比例不平衡问题, γ 用来调节简单样本所占权重降低的速率, $p_{x,y}$ 为置信度.

1.4.3 回归损失

在 FCOS 中使用交并比 (intersection over union, *IoU*) 进行回归损失, *IoU* 表示 BBox (bounding box) 框和 GT (ground truth) 框的交集之比, 如式 (7) 所示:

$$IoU = \frac{|A \cap B|}{|A \cup B|} \quad (7)$$

其中, A 、 B 分别表示预测框与真实框面积, *IoU* 取值范围 $[0, 1]$. 虽然 *IoU* loss (L_{IoU}) 能很好地反应预测框和真实框之间的相关程度, 但是在 BBox 与 GT 不相交的时候 *IoU* 值为 0, 无法计算 L_{IoU} , 梯度也为 0, 不能进行反向传播, 从而也无法回归; 当预测框与真实框有 *IoU* 时, 二者重合率大小是 *IoU* 无法正确预测的. 式 (8) 为 L_{IoU} 计算公式:

$$L_{IoU} = 1 - IoU \quad (8)$$

GIoU 可以解决 BBox 与 GT 因无重叠而造成的梯度消失问题, 聚焦在重叠区及其他的非重叠区域, 而 *IoU* 只聚焦在重叠区域. *GIoU* 的公式如式 (9) 所示:

$$GIoU = IoU - \frac{|C(A \cup B)|}{C} \quad (9)$$

其中, C 表示 BBox 与 GT 最小外接矩阵的面积, *GIoU* 取值范围为 $[-1, 1]$. 当 BBox 与 GT 相重合, 即三者面积相等, $GIoU = IoU = 1$; 当 BBox 框与 GT 框不相交且两者距离无穷远时, $GIoU = -1$. *GIoU* loss (L_{GIoU}) 在 L_{IoU} 之后加了一个惩罚项, 因此可以在不重叠时, 预测框向目标框移动, L_{GIoU} 的计算公式为式 (10):

$$L_{GIoU} = 1 - GIoU \quad (10)$$

L_{GIoU} 的取值范围为 $[0, 2]$.

模型整体损失函数可用式 (11) 表示:

$$L(\{p_{x,y}\}, \{t_{x,y}\}) = \frac{1}{N_{pos}} \sum_{x,y} L_{cls}(p_{x,y}, c_{x,y}^*) + \frac{\lambda_1}{N_{pos}} \sum_{x,y} I\{c_{x,y}^* > 0\} L_{reg}(t_{x,y}, t_{x,y}^*) + \frac{\lambda_2}{N_{pos}} \sum_{x,y} I\{c_{x,y}^* > 0\} L_{center-ness}(center-ness_{x,y}, center-ness_{x,y}^*) \quad (11)$$

其中, x, y 代表特征图上的某一位置 (x, y) ; $p_{x, y}$ 为特征图 (x, y) 点网络预测类别信息, $c_{x, y}^*$ 为真实类别信息; $t_{x, y}$ 为网络预测的目标边界框信息, $t_{x, y}^*$ 为真实目标边界框信息; $center-ness_{x, y}$ 为网络预测的 $center-ness$, $center-ness_{x, y}^*$ 为真实的 $center-ness$; N_{pos} 为正样本数量, λ 为平衡权重系数均取 1, $I\{c_{x, y}^* > 0\}$ 为激活函数, 当 $c_{x, y}^*$ 大于 0 时值为 1, 否则为 0, 即回归分支以及 $center-ness$ 分支计算的是正样本的损失。

2 实验数据

2.1 鱼类图像数据集

本文图像数据采用 Fish4Knowledge (F4K) 公开数据集中的 23 类鱼类图像 (数据集图像如图 4 所示), 以及从 F4K 鱼类视频中以帧截取图片组成。图像分辨率缩放至 240×320 像素, 采用 LabelImg 标注软件工具对数据集图像进行手工标注, 并生成 XML 文件供检测使用。



图 4 本文所用数据集图像

2.2 数据增强

由于每种鱼类数量不一, 因此对鱼类图像数据集使用旋转、缩放、水平镜像、明亮度、对比度调整等方式进行数据增强, 同时将图像对应的标注文件进行相应的调整使得标签完整。最终数据增强后的 F4K 数据集共获得 34920 幅图像, 视频帧图像共 1760 幅, 使用 8:2 的合适比例, 将两种数据集各自划分为训练集和验证集。

3 结果与分析

3.1 模型训练平台

本文试验平台主要配置环境为 Windows 11 操作系统, CPU 为 8 核 16 线程 Intel i7-11800H, GPU 为

NVIDIA RTX 3060, CUDA 版本为 11.0, CuDNN 版本为 8.0.3, 基于 PyTorch 1.9 框架下 Python 版本为 3.7 进行训练和测试。

3.2 模型实施

本文试验利用迁移学习, 将在 ImageNet 上进行预训练后得到的 MobileNetV2 模型参数迁移到本实验中, 以更好地获得模型鲁棒性以及提高模型精度。训练过程中, 参数调整如下: 动量 (momentum)、权重 (weight-decay) 和学习率 (learning rate) 分别设置为 0.9、0.0001 和 0.001, 更新模型参数使用的优化器为 SGD (stochastic gradient descent), 权值更新采用 BN (batch normalization) 正则化, batch size 设置为 16, 共迭代 90000 次。

本文在测试集使用各类 AP (average precision), 最终检测结果的评价指标将使用平均值 mAP。根据训练日志中得到的数据, 绘制损失变化曲线, 如图 5 所示, 图 5(a) 表示以 F4K 为数据集得到的总损失变化, 以及分别展示训练过程中分类分支、回归分支、 $center-ness$ 三支的损失变化。从图 5 可知, 改进之后损失函数在前 3000 轮迅速下降。

以视频帧截取图片为例, 展示该模型检测识别结果图, 如图 6 所示。

从图 6 可以看出水下环境模糊、杂质较多、光照不均匀, 但本文模型可以准确定位鱼类目标, 并且可以极大程度上准确识别该目标鱼类。

3.3 不同网络模型的对比试验

将本文研究算法与其他学者在 F4K 数据集上的成果进行比较, 如表 1 所示。董绍江等^[16]提出了一种可以降低计算量且能保证精度的分层精简双线性注意力网络, 精度达到 98.04%; 顾郑平等^[17]通过使用迁移学习得到预训练网络特征, 再结合 SVM 算法进行分类, 取得了 98.6% 的准确率; 焦明连等^[18]把密集卷积网络结合到了胶囊网络中, 为了更新权重, 使用了动态路由算法, 提出了一种改进的胶囊网络 (CapsNet+DenseNet), 在 F4K 数据集上精度达到 98.83%; 赵力等^[19]通过对卷积神经网络改进提出了非对称交互分支网络, 使得模型中的局部特征和全局信息的学习能力有所提高, 在 F4K 数据集上取得了 98.9% 的准确率。贾宇霞等^[20]在结合了迁移学习方法之后, 提出了一种可以实现图像主体自动增强功能的鱼类识别模型, 在 F4K 数据集上取得了较高的检测精度, mAP 达到 99.63%; 本文提出的模型在 F4K 上可以达到 99.79% 的准确率。

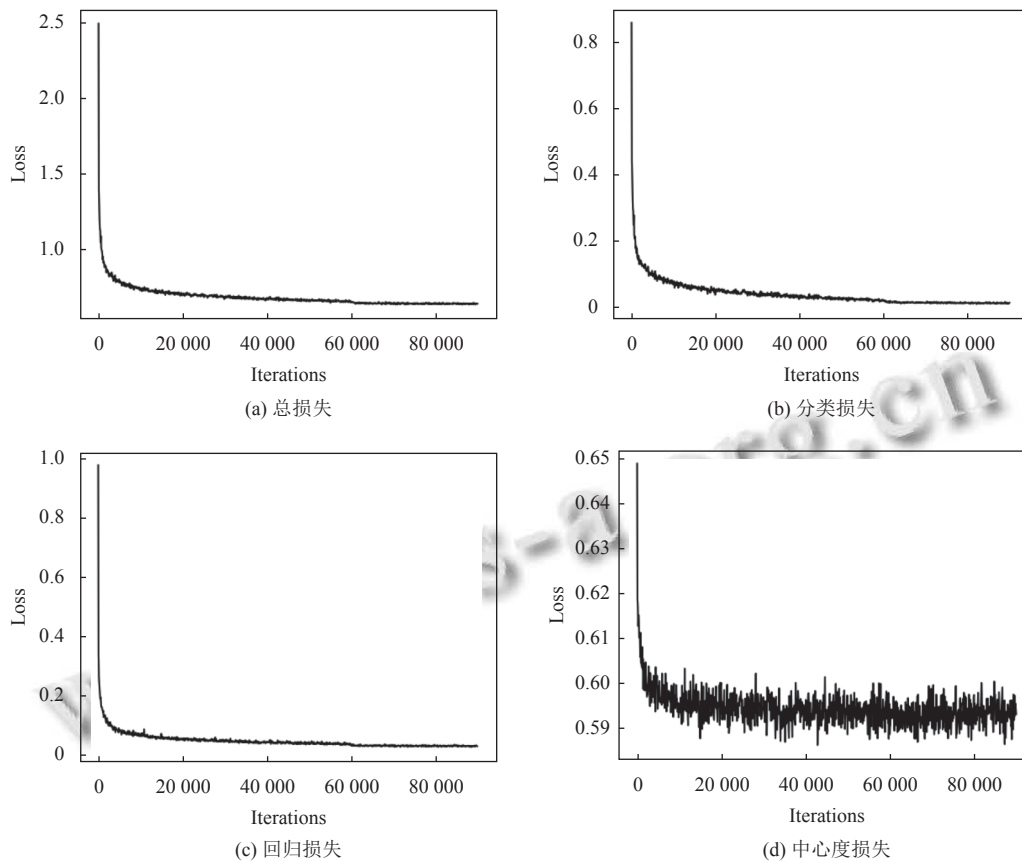


图5 F4K 训练时损失值变化



图6 鱼类检测效果图

表1 在 F4K 数据集上不同网络模型对比 (%)

网络	检测精度mAP
STN-H-CBP ^[16]	98.04
PreCNN+SVM ^[17]	98.60
CapsNet+DenseNet ^[18]	98.83
交互双分支网络 ^[19]	98.90
改进ResNet50 ^[20]	99.63
本文	99.79

3.4 对比试验

将 F4K 数据集与截取的视频帧图片放在模型中分别进行对比试验. 首先是对 F4K 数据集进行测试, 结果如表 2 所示. 可以看到以 ResNet50 为骨干网络时检测精度是 98.67%, 以 MobileNetv2 为骨干网络时检测精

度可以达到 98.37%, 准确率略有降低, 由于 Mobile-Netv2 为轻量级模型, 参数及计算量较少, 因此检测速度有所提升; 将 center-ness 分支从分类分支切换到回归分支, 损失函数使用 *GIoU*, 且特征金字塔融合 ASFF 之后, 其 mAP 比原 FCOS 均有提升, 模型的平均检测时间只牺牲了不到 1 ms.

为进一步检测网络本文对视频帧图片进行测试, 结果如表 3 所示. 本文的改进网络相比于只用 Mobile-Netv2 骨干网络的检测效果提升了 1.62%. 由此说明改进的 FCOS 能够在保证检测速度的前提下提升网络的检测性能.

表2 使用 F4K 改进前后结果对比

网络	On-reg	<i>GIoU</i>	ASFF	检测精度	平均检测时间
				mAP (%)	(ms)
FCOS-ResNet-50	—	—	—	98.67	26.38
FCOS-ResNet-50	√	√	—	99.46	25.75
FCOS-MobileNetv2	—	—	—	98.37	21.29
FCOS-A	√	—	—	98.94	21.54
FCOS-B	√	√	—	99.37	21.71
本文	√	√	√	99.79	22.07

表3 使用数据帧图片改进前后结果对比

网络	On-reg	<i>GloU</i>	ASFF	检测精度 mAP (%)	平均检测时间 (ms)
FCOS-MobileNetv2	—	—	—	98.26	26.10
本文	√	√	√	99.88	28.58

4 结论

本文提出了一种基于FCOS的自适应特征融合鱼类检测识别算法。该算法将从MobileNetv2骨干网络中提取的特征输出到特征金字塔中,提高了网络的检测速度;再将自适应空间特征融合(ASFF)与特征金字塔部分融合,避免特征尺度的不一致性;使用*GloU*损失替换*IoU*损失,解决预测框与真实框之间无重叠造成的梯度消失问题,将center-ness分支引入到回归分支,进一步提升模型精度。最终实验结果表明,本文所提出的网络模型在鱼类公共数据集进行检测,其检测精度与识别速度均有所提升,对海洋鱼类检测与识别有很好的适应性。

参考文献

- Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation. 2014 IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2014. 580–587.
- Ren SQ, He KM, Girshick R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137–1149. [doi: 10.1109/TPAMI.2016.2577031]
- Redmon J, Farhadi A. YOLOv3: An incremental improvement. arXiv:1804.02767, 2018.
- Liu W, Anguelov D, Erhan D, et al. SSD: Single shot multibox detector. Proceedings of the 14th European Conference on Computer Vision. Cham: Springer, 2016. 21–37.
- Lin TY, Goyal P, Girshick R, et al. Focal loss for dense object detection. Proceedings of the 2017 IEEE International Conference on Computer Vision. Venice: IEEE, 2017. 2999–3007.
- Alsmadi MKS, Bin Omar K, Noah SA, et al. Fish recognition based on the combination between robust features selection, image segmentation and geometrical parameters techniques using artificial neural network and decision tree. International Journal of Computer Science and

- Information Security, 2016, 6(2): 215–221.
- 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. 模式识别与人工智能, 2019, 32(3): 193–203. [doi: 10.16451/j.cnki.issn1003-6059.201903001]
- 王亚茹, 刘雨青, 黄璐瑶. 基于FML-Centernet算法的鱼类识别检测方法. 激光与光电子学进展, 2021: 1–15. http://kns.cnki.net/kcms/detail/31.1690.tn.20210712.1608.016.html, 2022-06-26.
- 覃学标, 黄冬梅, 宋巍, 等. 模糊水下图像多增强与输出混合的鱼类检测方法. 农业机械学报, 2022, 53(7): 243–249. [doi: 10.6041/j.issn.1000-1298.2022.07.025]
- Zhou XY, Wang DQ, Krähenbühl P. Objects as points. arXiv:1904.07850, 2019.
- Tian Z, Shen CH, Chen H, et al. FCOS: A simple and strong anchor-free object detector. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 44(4): 1922–1933.
- Law H, Deng J. CornerNet: Detecting objects as paired keypoints. International Journal of Computer Vision, 2020, 128(3): 642–656. [doi: 10.1007/s11263-019-01204-1]
- He KM, Zhang XY, Ren SQ, et al. Deep residual learning for image recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas: IEEE, 2016. 770–778.
- Sandler M, Howard A, Zhu ML, et al. MobileNetV2: Inverted residuals and linear bottlenecks. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018. 4510–4520.
- Liu ST, Huang D, Wang YH. Learning spatial fusion for single-shot object detection. arXiv:1911.09516, 2019.
- 董绍江, 刘伟, 蔡巍巍, 等. 基于分层精简双线性注意力网络的鱼类识别. 计算机工程与应用, 2022, 58(5): 186–192. [doi: 10.3778/j.issn.1002-8331.2009-0349]
- 顾郑平, 朱敏. 基于深度学习的鱼类分类算法研究. 计算机应用与软件, 2018, 35(1): 200–205. [doi: 10.3969/j.issn.1000-386x.2018.01.035]
- 焦明连, 朱鹏, 蓝玉珍, 等. 胶囊网络的海洋鱼类图像分类方法. 江苏海洋大学学报(自然科学版), 2020, 29(3): 54–58.
- 赵力, 宋威. 基于非对称双分支交互神经网络的水下生物识别. 计算机应用研究, 2021, 38(4): 1240–1244, 1255. [doi: 10.19734/j.issn.1001-3695.2020.03.0083]
- 贾宇霞, 樊帅昌, 易晓梅. 基于显著性增强和迁移学习的鱼类识别研究. 渔业现代化, 2020, 47(1): 38–46. [doi: 10.3969/j.issn.1007-9580.2020.01.006]

(校对责编: 牛欣悦)