

DOI:10.19651/j.cnki.emt.2312890

基于YOLOv5_PGS的轻量级水下生物识别目标检测*

周新 张春堂 樊春玲

(青岛科技大学自动化与电子工程学院, 青岛 266061)

摘要: 在复杂自然环境下高效探测水下生物资源对中国渔业具有重要意义。针对复杂弱光环境下水下生物资源识别能力低、特征丢失严重等问题,本文提出了一种轻量级的水下生物检测算法。首先,针对水下图像颜色偏差大、清晰度低等问题,提出暗通道一对比限制-光衰减算法用于丰富图像特征信息。其次,引入GhostNet网络和构建C3CA模块提高模型的特征提取和融合能力。最后,对损失函数进行了改进,在降低总损失自由度的同时进一步提升算法的泛化能力。实验结果表明,YOLOv5_PGS算法在水下生物数据集上的检测精度达到了86.22%,较原YOLOv5L算法提高了0.48%。此外,本文算法模型的体积仅为20.4MB,比原模型减少了89.31%,检测速度提高了56.56%。实验结果表明,YOLOv5_PGS算法在水下图像处理中取得了良好的效果,为水下生物资源的实时检测提供了保证。

关键词: 图像增强;水下生物目标;YOLOv5L;轻量级;CAC3;SIU

中图分类号: TP391 **文献标识码:** A **国家标准学科分类代码:** 520.60

Lightweight YOLOv5_PGS based objective detection for underwater biological identification

Zhou Xin Zhang Chuntang Fan Chunling

(College of Automation and Electronic Engineering, Qingdao University of Science and Technology, Qingdao 266061, China)

Abstract: The efficient detection of the underwater biological resources in a complex natural environment is of great significance to China fishery. Aiming at the problems such as low recognition ability and serious feature loss of underwater biological resources in complex low-light environment, a lightweight underwater biological detection algorithm is proposed in this paper. First of all, aiming at the problems of large color deviation and low resolution of underwater images, a Dark channel-contrast limiting-optical attenuation algorithm is proposed to enrich the feature information of underwater images. Thereafter, GhostNet module and C3CA module are used to improve the fusion capability of feature extraction network. Finally, the loss function is improved to reduce the total loss freedom. The experimental results show that the mean average precision of the algorithm reaches 86.22%, which is 0.48% higher than that of the original YOLOv5L. Moreover, the volume of the proposed algorithm model is only 20.4 MB, which is about 89.31% less than that of the original model, and the detection speed of the proposed model is increased by 56.56%. The experimental results show that the improved algorithm achieves good results in underwater images and provides a guarantee for the real-time detection of underwater biological resources.

Keywords: image enhancement; underwater biological targets; YOLOv5L; lightweight; CAC3; SIU

0 引言

随着中国渔业的快速发展,对水下渔业资源的勘探具有重要的研究意义和应用价值。由于水下图像雾化严重导致图像特征模糊,传统的目标检测算法^[1]需要人工识别、提取和分类目标特征,冗余度高,鲁棒性差。自卷积神经网络提出以来,基于深度学习的目标检测算法^[2]可以在同一网

络中对不同尺度特征进行融合,显著提高了目标检测的精度和速度。

目前基于深度学习的目标检测算法主要分为以R-CNN(regions with CNN features)^[3-4]系列算法为代表的双步法与以YOLO(you only look once)^[5]和SSD(single shot multibox detector)^[6]为代表的单步法。两种算法各有优缺点,双步法生成候选区域后再通过卷积神经网络对目标预

收稿日期:2023-02-23

* 基金项目:山东省自然科学基金(ZR2019MEE071)项目资助

测,具有较高的检测精度,单步法提出端到端的概念,将目标框的预测与识别融合,直接回归物体的类别概率和位置坐标值,以牺牲目标检测精度为代价减少了模型的复杂度,更好地均衡了检测精度与速度,适合移动端。

Zhou等^[7]提出了3种专用于水下成像的数据增强方法,并结合Faster R-CNN模型,验证了该方法对海洋生物检测的有效性,但在模型检测精度和速度上仍有提升空间。针对检测精度,Salman等^[8]提出将高斯混合模型与R-CNN算法结合,实现了复杂环境下对鱼类资源的识别,但模型结构过于复杂,难以实现实时检测。针对检测精度与速度,Yang等^[9]提出使用YOLOv3算法将网络结构中的低层次视觉特征与高层次语义特征有机结合,实现了对海胆、海参和扇贝的准确识别,较Faster R-CNN相比平均精度提高6.4%,基本满足实时检测的要求,但其未考虑到水下图像存在失真的问题。Chen等^[10]提出了YOLOv5s-underwater算法来解决图像颜色失真、能见度低等问题,目标的平均检测精度较YOLOv3提升了21.5%。

上述算法虽然在特定的水下环境下目标检测精度有所提高,但仍存在一些挑战:一是水下生物照片存在蓝色或绿色的色偏,二是水下生物检测既需要较高的检测精度,也需要较高的实时性。目前常用的检测算法不能很好地平衡检测精度和速度之间的关系。因此,本文提出了一种基于YOLOv5L的改进算法YOLOv5_PGS(YOLOv5_predict

ghost strength),在纠正水下图片蓝绿色偏的同时,更好地兼顾水下生物资源检测的实时性与准确性。本文主要贡献如下:

1)提出一种用于水下图像增强的DCL(dark channel-contrast limiting-optical attenuation)算法。通过优化暗通道的透射率、子块限制映射和估计场景深度图来增强初始数据集的特征,构建高质量特征的数据集。

2)针对提取特征图存在的冗余问题,使用GhostNet^[11]替换原骨干网络CSPDarknet,利用Ghost模块减少模型的复杂度,提升网络的推理速度

3)构建C3CA模块,在C3模块的基础上引入CA(channel attention)^[12]模块,用于增强网络对小目标区域的特征融合能力,提高模型的检测精度。

4)在损失函数方面,引入SIOU(scylla IoU)^[13]函数减少损失的自由度以准确定位预测框的位置。

1 改进的YOLOv5算法

YOLOv5_PGS的网络结构主要包括3个模块:Ghost模块构成的骨干网络;引入C3CA模块的特征融合模块;预测模块。网络的输入为 $640 \times 640 \times 3$ 的图像,经过骨干网络特征提取后输出3种不同尺度的特征图送入特征融合网络,特征融合后将获得的3个加强特征层送入SIOU损失函数进行模型训练。YOLOv5_PGS结构如图1所示。

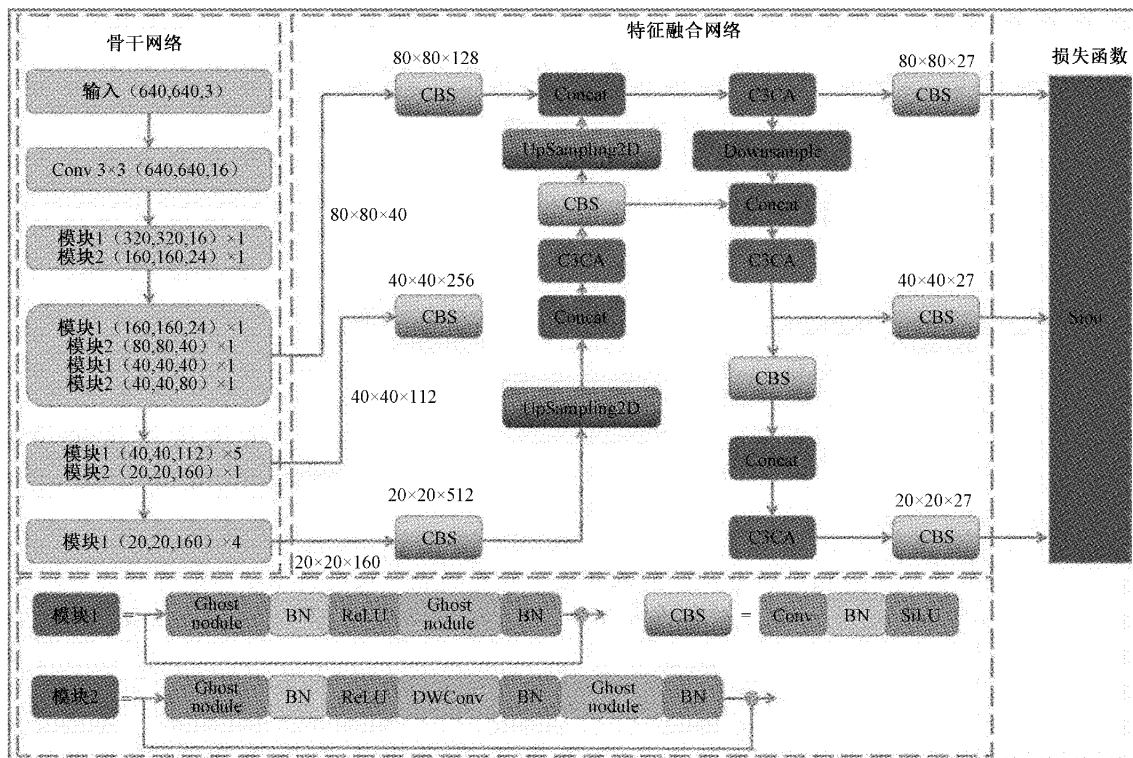


图1 YOLOv5_PGS网络结构

1.1 图像增强

由于水下介质和混合粒子对光线的吸收和散射作用,

水下图像经常出现细节模糊、对比度低、色彩模糊等问题。考虑到暗通道先验算法^[14]在图像去雾中性能优越,本文在

暗通道先验去雾算法的基础上结合水下图像的 RGB 颜色空间和水下光衰减等信息对水下图像进行增强。

如图 2 所示,本文提出的 DCL 图像增强算法主要分为 3 个阶段,第 1 阶段先对 RGB 通道最小值组成的灰度图像进行腐蚀,得到暗通道,随后使用引导滤波对透射率优化得到去雾图像。第 2 阶段考虑到输入图像区域特性问题,

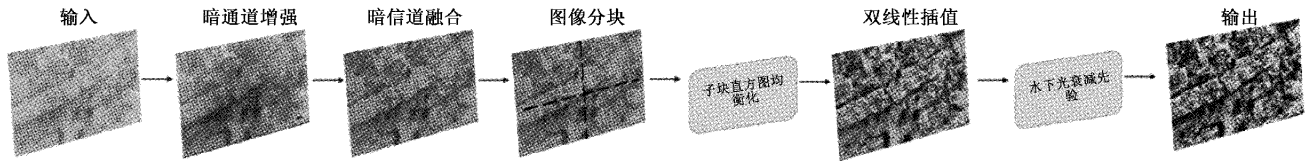


图 2 图像增强总体流程图

1.2 主干特征提取网络

YOLOv5L 网络的骨干网络为 CSPDarknet,具有较强的特征提取能力,但模型计算复杂,导致网络参数量剧增,运算内存较大,不利于实时水下生物检测。针对上述问题,本文引入 GhostNet 网络作为弱光水下生物的骨干网络。

GhostNet 网络的基本组成单元是 Ghost 模块,如图 3 所示。Ghost 卷积先使用 1×1 卷积获得特征图的浓缩特征,再使用廉价的线性运算获得更多的特征图,理论上使用 Ghost 卷积模块可以节约 s 倍的运算时间,缩减 s 倍的参数。

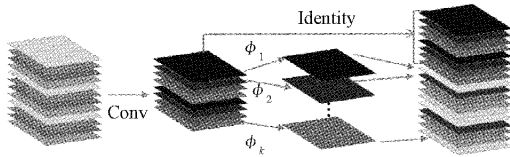


图 3 Ghost 卷积模块

1.3 C3CA 注意力模块

注意力是指聚焦于局部信息的机制,可以帮助计算机注意到更多重要信息。不同于 SE(squeeze and excitation)注意力模块^[15]只考虑通道维度上的注意力,CA 模块将位置信息嵌入到通道注意力中,在捕获跨通道的信息的同时,还能捕获方向感知和位置感知的信息。

本文构建的 C3CA 模块可使得模型具备更高的迁移和特征融合能力,其结构如图 4 所示。C3CA 模块借鉴了残差模块的思想中间直接通过“捷径”连接到输出,这样可以避免因网络加深带来的梯度爆炸等问题,左侧通过两次卷积用于特征提取与融合,右侧通过坐标信息嵌入和坐标注意力生成使得网络更加关注有用的特征信息。

1.4 SIOU 损失函数

损失函数的选择在很大程度上会影响训练的速度和推理的准确性,传统的损失函数并未考虑真实框与预测框之间匹配方向的问题,这可能会使训练收敛延迟以及预测框“四处游荡”并导致最终模型变差。针对此问题,本文提出使用 SIOU 损失函数代替原 CIOU(complete IoU)损失函数。

SIOU 由角度损失、距离损失、形状损失、IOU 损失 4

对图像划块处理,将获得的去雾图像划分为 $M \times N$ 个不相交的矩形子块,分块均衡化水下图像的 RGB 3 个颜色通道,随后使用双线性插值优化块与块之前的过渡,第 3 阶段基于图像形成模型恢复图像,利用水下光衰减的先验信息估计场景深度图,通过深度图回归水下图像的真实亮度。

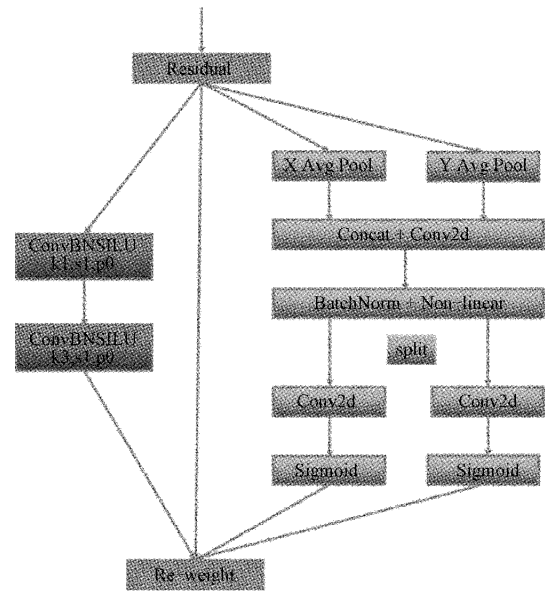


图 4 C3CA 模块

部分组成。SIOU 如式(3)所示。

$$L_{box} = 1 - IOU + \frac{\Delta + \Omega}{2} \quad (1)$$

其中, Δ 指距离损失, Ω 指形状损失。IOU(intersection over union)表示预测框与真实框之间的交并比。

SIOU 将预测框的角度也考虑在内,使得预测框在稳定回归的同时加快了收敛速度。

2 实验环境的搭建与评价指标

2.1 实验环境与数据集

实验环境为 Ubuntu16.04 64 位操作系统, GPU 为 NVIDIA GeForce RTX 2080Ti, 11 GB 显存;模型测试的环境配置为 Windows10 64 位操作系统, GPU 为 NVIDIA GeForce RTX 1050, 4 GB 显存,基于 Pytorch 深度学习框架,使用 Python3.6 语言编写程序。根据硬件条件,模型训练过程分为冻结和解冻两个阶段。在冻结阶段,骨干特征提取网络没有变化,批次大小为 16,训练为 50 轮,在解冻阶段,骨干特征提取网络开始变化,批次大小为 8,训练为

250 轮。两阶段均采用 SGD 优化器,初始学习率为 0.01,动量因子为 0.937,采用余弦衰减法。

本文采用 DCL 算法对公开数据集“underwater detection”进行增强。然后利用增强后的图像验证 YOLOv5_PGS 算法的可靠性。该数据集由 7 543 张图像组成,分为海胆、海参、扇贝和海星 4 类,其中(训练集:验证集):测试集=(9:1):1。为了增强模型的泛化能力,还会对数据进行马赛克和混合等方式进行数据增强。

2.2 评价指标

为了验证 DCL 算法对初始数据集的增强效果,考虑到

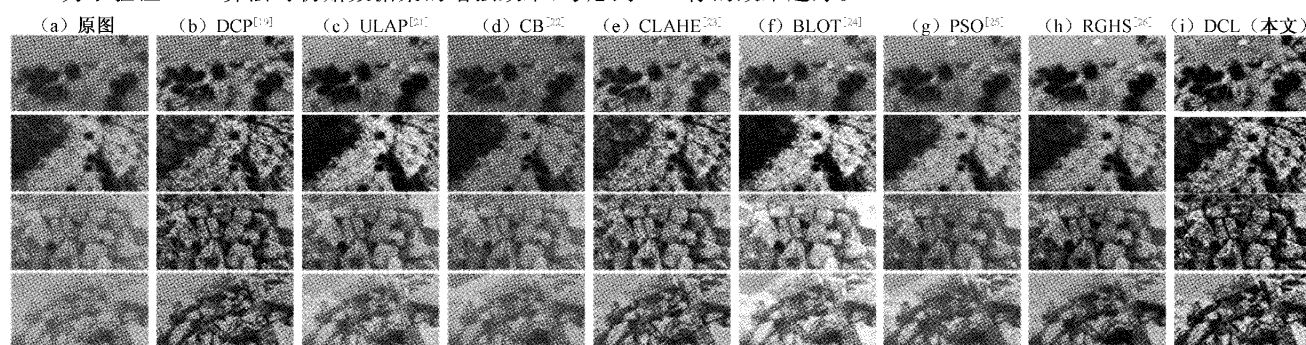


图5 水下图像增强算法的比较

表1 水下图像质量评价

| 方法 | 图5第1行 | | | 图5第2行 | | | 图5第3行 | | | 图5第4行 | | |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| | UCIQE | UIQM | IE | UCIQE | UIQM | IE | UCIQE | UIQM | IE | UCIQE | UIQM | IE |
| CB | 0.856 | 3.009 | 7.543 | 0.653 | 3.336 | 6.984 | 0.484 | 3.050 | 7.136 | 0.477 | 3.195 | 7.153 |
| CLAHE | 0.842 | 2.806 | 7.832 | 0.637 | 3.155 | 7.193 | 0.610 | 2.977 | 7.648 | 0.568 | 3.022 | 7.617 |
| BLOT | 2.299 | 2.039 | 7.617 | 0.590 | 3.037 | 7.389 | 0.482 | 2.462 | 7.310 | 0.458 | 2.699 | 7.383 |
| PSO | 0.695 | 2.949 | 7.679 | 0.596 | 2.538 | 7.117 | 0.542 | 2.209 | 7.237 | 0.525 | 2.585 | 7.310 |
| RGHS | 1.238 | 1.990 | 7.792 | 1.391 | 2.876 | 7.546 | 1.046 | 2.746 | 7.698 | 0.937 | 2.508 | 7.667 |
| DCP | 0.697 | 1.987 | 7.563 | 0.658 | 1.856 | 6.990 | 0.524 | 2.029 | 7.196 | 0.478 | 1.839 | 7.049 |
| ULAP | 2.380 | 2.208 | 7.692 | 1.110 | 2.401 | 7.407 | 0.506 | 2.605 | 7.476 | 0.469 | 2.761 | 7.377 |
| UMC | 2.802 | 1.938 | 7.800 | 3.666 | 2.117 | 7.433 | 2.844 | 2.582 | 7.916 | 1.375 | 2.838 | 7.933 |

为了客观评价各网络模型的有效性,采用平均精确均值(mean average precision, mAP)、帧数(frame per second, FPS)、模型容量、参数量与浮点运算量对模型进行评价。其中 AP 计算公式如式(14)~(16)所示。

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

$$AP = \int_0^1 P(r) dr \quad (4)$$

式中:Precision 为精确率,Recall 为召回率,TP 为模型正确预测的正样本数,FP 为模型错误预测的负样本数,FN 为模型错误预测的正样本数,对 Precision-Recall 曲线积分可得到平均精度值 AP,平均精度均值 mAP 即为多类别平均精度。

本文提出的 DCL 算法主要用于增加图像的特征,采用以下 3 个评价指标对 DCL 算法处理的结果进行评价:IE (information entropy)、UCIQE (underwater colour image quality evaluation) 和 UIQM (underwater colour quality measure)。IE 反映了水下图像所包含的平均信息量和信息内容的丰富程度。UCIQE 是水下图像中颜色浓度、饱和度和对比度的线性组合。它用于量化水下图像中不均匀的颜色偏差、模糊和对比度。UIQM 从颜色、锐度和对比度 3 个方面评估图像质量。以上 3 个指标的值越高,图像的效果越好。

3 实验结果与分析

3.1 图像增强

本文从数据集中选取了 4 张具有代表性的图像来验证 DCL 算法的有效性和泛化能力,并采用 DCL 算法将所选图像与基于内容的推荐算法(content based, CB)^[16]、限制对比度自适应直方图均衡算法(contrast limited adaptive histogram equalization, CLAHE)^[17]、特征分析算法 BLOT^[18]、粒子群优化算法(particle swarm optimization, PSO)^[19]、相对全局直方图拉伸算法 RGHS(relative global histogram stretching)^[20]、暗通道先验去雾算法(dark channel prior, DCP)和水下光衰减的先验信息 ULAP (underwater light attenuation prior)^[21]算法进行比较。如图 5 所示,CB 和 PSO 算法处理的图像在某些区域存在严

重的色移;CLAHE、RGHS、DCP 和 ULAP 算法不存在局部色偏问题,但处理后的图像整体存在蓝色或绿色色偏;经过 BLOT 算法处理的图像存在一定的雾化现象。DCL 算法可以很好地解决上述问题。

采用 UCIQE、UIQM 和 IE 对实验结果进行了进一步的客观评价。从表 1 可以看出,DCL 处理的图像在 UCIQE 和 IE 中都取得了很好的效果。CB 算法在指标 UIQM 上实现了所有的第 1 名,但 CB 算法处理的照片

存在明显的区域颜色偏差,这可能与 UIQM 计算本身有关。

3.2 骨干网提取结果比较

为实现对水底生物的高效检测,本文以 GhostNet 为主干特征提取网络设计 YOLOv5_PGS 轻量级架构模型。为客观分析算法的有效性,在相同实验环境下,分别以 ConvNeXt^[22]、MobileNetV3^[23]、Swin Transformer^[24] 作为主干特征提取网络做对照实验,实验结果如表 2 所示。

表 2 不同主干特征提取网络结果对比

| Method | Backbone | mAP/% | FPS/(frame/s) | Size/MB | Parameters/M | GFLOPs/G |
|-----------|------------------|-------|---------------|---------|--------------|----------|
| YOLOv5L | CSPDarknet | 85.74 | 22.24 | 187.1 | 46.648 | 114.610 |
| YOLOv5_C | ConvNeXt | 85.35 | 19.47 | 194.0 | 42.793 | 101.861 |
| YOLOv5_ST | Swin Transformer | 86.18 | 18.12 | 193.0 | 41.954 | 100.473 |
| YOLOv5_M | MobileNetV3 | 81.79 | 39.14 | 34.2 | 5.793 | 9.356 |
| YOLOv5_G | GhostNet | 82.59 | 35.17 | 21.3 | 5.492 | 8.066 |

由表 2 可以看出,当使用 ConvNeXt 和 Swin Transformer 作为骨干特征提取网络时,模型 mAP 与原模型相比略有波动,模型复杂度的数量级不变。当 MobileNetV3 与 GhostNet 作为主干特征提取网络时模型容量、参数量、浮点运算量方面均减少 1 个数量级,但 GhostNet 的 mAP 仅略微下降,FPS 增长 58%。因此本文将 GhostNet 作为模型的骨干网络,用于构建

YOLOv5_PGS。

3.3 消融实验

为验证本文改进点的有效性,对通过使用 GhostNet 网络改进的方法 1,使用 GhostNet 与 DCL 算法改进的方法 2,使用 GhostNet、DCL 和 C3CA 模块改进的方法 3 以及本文算法 YOLOv5_PGS 在测试集上进行消融实验,实验结果如表 3 所示。

表 3 消融实验结果

| Model | GhostNet | DCL | C3CA | SIUO | AP | | | | mAP/% | FPS/(frame/s) |
|------------|----------|-----|------|------|--------|-------------|---------|----------|-------|---------------|
| | | | | | urchin | holothurian | scallop | starfish | | |
| YOLOv5L | × | × | × | × | 90.62 | 78.68 | 89.01 | 86.66 | 85.74 | 22.24 |
| Method 1 | √ | × | × | × | 89.36 | 72.40 | 84.18 | 84.42 | 82.59 | 35.17 |
| Method 2 | √ | √ | × | × | 89.57 | 76.40 | 86.17 | 85.21 | 83.89 | 35.08 |
| Method 3 | √ | √ | √ | × | 91.61 | 78.29 | 86.39 | 87.14 | 85.86 | 34.42 |
| YOLOv5_PGS | √ | √ | √ | √ | 91.73 | 79.20 | 86.55 | 87.40 | 86.22 | 34.82 |

从表 3 可以看出,方法 1 的 mAP 比 YOLOv5L 算法低 3.15%,说明引入 GhostNet 轻量级网络后,模型的特征提取能力有所下降,但帧数 FPS 提高了 58.14%。引入 DCL 算法后,方法 2 的 mAP 比方法 1 提高了 1.3%。其中,海参和扇贝的检测效果明显提高。引入 C3CA 模块后,方法 3 的 mAP 比方法 2 的 mAP 提高了 1.97%,说明注意机制的引入增强了模型检测小目标的能力,增强了特征融合网络的能力。本文的 YOLOv5_PGS 算法较方法 3 算法在 4 个种类检测方面均有所细微提升,与 YOLOv5L 算法相比,mAP 提高了 0.48%,帧数 FPS 提高了 56.56%,满足实时检测的要求。

表 4 给出了 YOLOv5_PGS 和 YOLOv5L 算法模型的复杂度比较。由表 4 可以看出,YOLOv5_PGS 算法的模

型容量仅为 YOLOv5L 的 10.9%,参数仅为 YOLOv5L 的 13.6%,GFLOPs (giga floating point operations) 仅为 YOLOv5L 的 10.0%。YOLOv5_PGS 模型的复杂性显著降低有利于提高模型的移植适应能力。

表 4 YOLOv5L 与 YOLOv5_PGS 的模型复杂度

| Model | Size/MB | Parameters/M | GFLOPs/G |
|------------|---------|--------------|----------|
| YOLOv5L | 187.1 | 32.815 | 76.773 |
| YOLOv5_PGS | 20.4 | 5.200 | 7.673 |

3.4 与其他目标检测算法的比较

为了验证 YOLOv5_PGS 算法的优越性,将 YOLOv5_PGS 算法与 Faster R-CNN、SSD、CenterNet、YOLOv5s 和

文献[10]算法在同一数据集上进行实验,结果如表5所示。从表5中可以看出,YOLOv5_PGS算法在模型复杂度上比Faster R-CNN、SSD和CenterNet算法有1个数量级的下降,检测精度也高于这3种算法。在检测精度方面,

YOLOv5_PGS算法比YOLOv5s算法提高了20.12%,比文献[10]算法提高了1.85%,但与YOLOv5s和文献[10]算法相比,其帧数略有下降。综合来看,本文算法在复杂环境下对小目标的检测和识别能力更强,整体性能更好。

表5 YOLOv5_PGS与其他目标检测算法的对比

| Model | AP | | | | mAP/ % | FPS/ (frame/s) | Size/ MB | Parameter/ M | GFLOPs/ G |
|--------------|-------|-------|-------|-------|-----------|-------------------|-------------|-----------------|--------------|
| | 海胆 | 海参 | 扇贝 | 海星 | | | | | |
| Faster R-CNN | 83.60 | 73.56 | 71.42 | 83.12 | 77.92 | 7.46 | 113.5 | 137.099 | 370.210 |
| SSD | 74.33 | 62.63 | 46.56 | 69.75 | 63.32 | 52.99 | 96.6 | 26.285 | 62.747 |
| CenterNet | 90.20 | 77.12 | 84.40 | 85.60 | 84.33 | 38.02 | 131.0 | 32.665 | 70.217 |
| YOLOv5s | 67.70 | 61.50 | 65.30 | 69.90 | 66.10 | 41.32 | 27.1 | 7.277 | 17.156 |
| 文献[10] | 85.60 | 84.80 | 82.30 | 84.80 | 84.37 | 36.71 | 35.25 | 8.385 | 22.384 |
| YOLOv5_PGS | 91.73 | 79.20 | 86.55 | 87.40 | 86.22 | 34.82 | 20.4 | 5.200 | 7.673 |

将本文改进算法训练好的权重对测试集部分数据进行目标检测,检测结果如图6所示。根据图像原始标签及第1张对比图中可以看出,YOLOv5L存在远处扇贝缺失的问题,YOLOv5_PGS对远处贝壳的识别能力优于

YOLOv5L。从第2张对比图可以看出,YOLO_PGS中加入C3CA模块,不仅减少了漏检的情况,而且对远端探测目标的置信度也比YOLOv5L高。

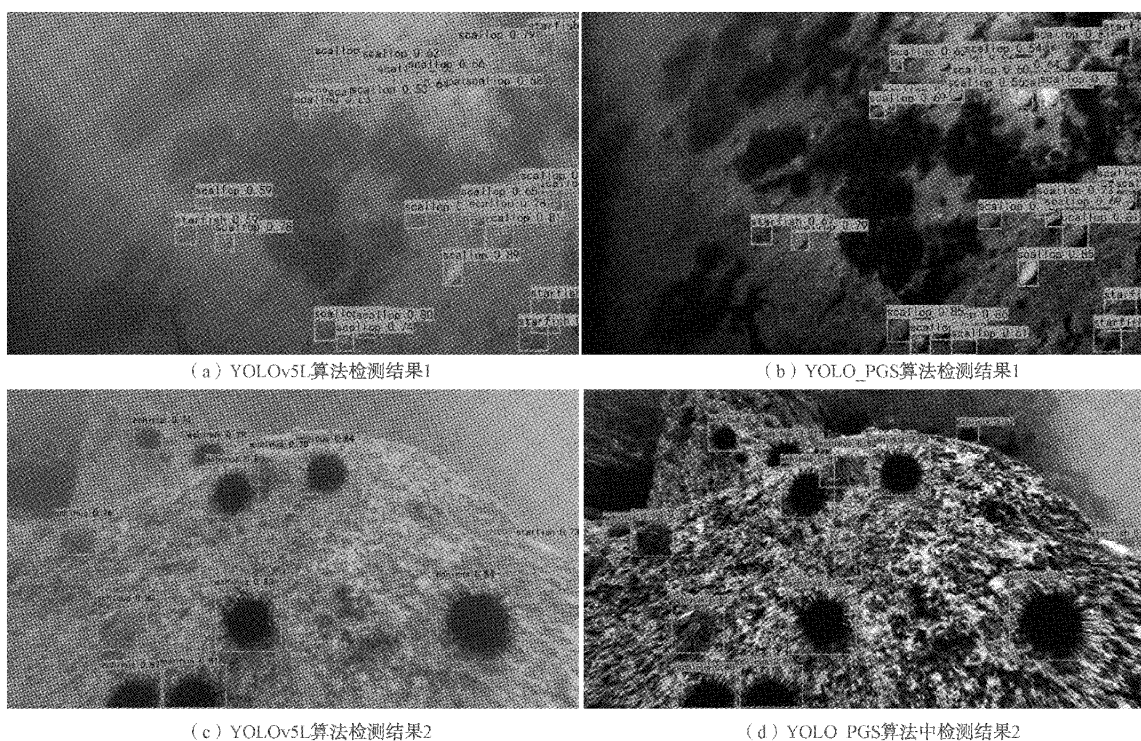


图6 YOLOv5L算法与YOLO_PGS算法检测结果

综合分析表明,采用DCL算法后,数据集的颜色和可见性得到恢复,改善了数据集的特征信息,解决了数据集的颜色偏差问题。在此基础上,改进的YOLO5_PGS算法更容易识别,降低了水下生物的漏检率和误检率。

4 结论

在本研究中,提出了一种基于YOLOv5_PGS的轻量

级水下生物识别目标检测的算法。在图像增强中,DCL算法提高了生物特征的清晰度,显著增加了图片的特征信息。在特征提取网络部分,用GhostNet作为主干网络,将模型复杂度降低了1个数量级,大大提升了模型的检测速度。在特征融合部分提出的C3CA模块增强了网络特征融合能力,采用SIOU作为边界框损失函数,提高了模型训练的速度和推理的准确性。

与其他目标检测方法相比,YOLOv5_PGS 算法解决了图像色偏失真的问题,促使模型更好地提取水下生物特征,具备良好的泛化能力,当其他水下生物数据集样本足够时,针对其他水下生物可做有效性推广。下一步将研究本算法对其他水下生物检测的有效性与机器人携带摄像机进行水下生物探测。

参考文献

- [1] SCHRAMM K D, MARNANE M J, ELSDON T S, et al. A comparison of stereo-BRUVs and stereo-ROV techniques for sampling shallow water fish communities on and off pipelines[J]. *Marine Environmental Research*, 2020, 162: 105198.
- [2] 张培培, 王昭, 王菲. 基于深度学习的图像目标检测算法[J]. *国外电子测量和技术*, 2020, 39(8): 34-39.
- [3] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. *Advances in neural information processing systems*, 2016, 39(6): 1137-1149.
- [4] 郁岩, 齐继阳. 改进 Faster R-CNN 的微型扁平电机 FPC 表面焊点缺陷检测[J]. *电子测量技术*, 2022, 45(7): 146-151.
- [5] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016: 779-788.
- [6] 贾君霞, 史珂鑫. 改进型 SSD 道路行人目标检测算法[J]. *国外电子测量技术*, 2022, 41(12): 26-32.
- [7] ZHOU H, HUANG H, YANG X, et al. Faster R-CNN for marine organism detection and recognition using data augmentation [C]. *Proceedings of the International Conference on Video and Image Processing*, 2017: 56-62.
- [8] SALMAN A, SIDDIQUI S A, SHAFAIT F, et al. Automatic fish detection in underwater videos by a deep neural network-based hybrid motion learning system[J]. *ICES Journal of Marine Science*, 2020, 77(4): 1295-1307.
- [9] YANG H, LIU P, HU Y Z, et al. Research on underwater object recognition based on YOLOv3[J]. *Microsystem Technologies*, 2021, 27: 1837-1844.
- [10] 陈宇梁, 董绍江, 孙世政, 闫凯波. 改进 YOLOv5 的弱光水下生物目标检测算法[J]. *北京航空航天大学学报*, 2022: 1-13.
- [11] HAN K, WANG Y, TIAN Q, et al. Ghostnet: More features from cheap operations[C]. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020: 1580-1589.
- [12] HOU Q, ZHOU D, FENG J. Coordinate attention for efficient mobile network design[C]. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021: 13713-13722.
- [13] GEVORGYAN Z. SIOU loss: More powerful learning for bounding box regression [J]. *ArXiv Preprint*, 2022, ArXiv: 2205. 12740.
- [14] WANG J B, HE N, ZHANG L L, et al. Single image dehazing with a physical model and dark channel prior[J]. *Neurocomputing*, 2015, 149: 718-728.
- [15] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018: 7132-7141.
- [16] ANCUTI C O, ANCUTI C, DE VLEESCHOUWER C, et al. Color balance and fusion for underwater image enhancement[J]. *IEEE Transactions on Image Processing*, 2017, 27(1): 379-393.
- [17] SETIAWAN A W, MENGKO T R, SANTOSO O S, et al. Color retinal image enhancement using CLAHE [C]. *International Conference on ICT for Smart Society*, 2013: 1-3.
- [18] SONG W, WANG Y, HUANG D, et al. Enhancement of underwater images with statistical model of background light and optimization of transmission map [J]. *IEEE Transactions on Broadcasting*, 2020, 66(1): 153-169.
- [19] AZMI K Z M, GHANI A S A, YUSOF Z M, et al. Natural-based underwater image color enhancement through fusion of swarm-intelligence algorithm [J]. *Applied Soft Computing*, 2019, 85: 105810.
- [20] ALMAHAIRI A, RAJESHWAR S, SORDONI A, et al. Augmented CycleGAN: Learning many-to-many mappings from unpaired data [C]. *International Conference on Machine Learning*, 2018: 195-204.
- [21] SONG W, WANG Y, HUANG D, et al. A rapid scene depth estimation model based on underwater light attenuation prior for underwater image restoration[C]. *Advances in Multimedia Information Processing-PCM 2018: 19th Pacific-Rim Conference on Multimedia*, 2018: 678-688.
- [22] LIU Z, MAO H, WU C Y, et al. A convnet for the 2020s[C]. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022: 11976-11986.
- [23] HOWARD A, SANDLER M, CHU G, et al. Searching for mobilenetv3 [C]. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019: 1314-1324.

- [24] LIU Z, LIN Y, CAO Y, et al. Swin Transformer: Hierarchical vision transformer using shifted windows[C]. Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021: 10012-10022.

作者简介

周新,在读硕士研究生,主要研究方向为机器视觉和信息检测与处理。

E-mail: XinZhou1026@163.com

张春堂(通信作者),副教授,主要研究方向为深度学习、图像处理、模式识别、检测技术与智能装置。

E-mail: zct1999@163.com

樊春玲,博士,教授,硕士生导师,主要研究方向为图像信息处理技术和多相流传感系统与流体流动。

E-mail: chunlingfan@qust.edu.cn