

基于图像增强与改进 YOLOv3 的水下生物检测算法

郝 琨¹, 王 阔¹, 赵 璐¹, 王贝贝², 王传启³

(1. 天津城建大学 计算机与信息工程学院, 天津 300384; 2. 天津城建大学 控制与机械工程学院, 天津 300384; 3. 天津凯发电气股份有限公司, 天津 300384)

摘 要: 为准确检测水下生物, 本文采用优化的 MSRCR 算法对水下图像进行增强, 并基于 DenseNet 思想提出一种改进的 YOLOv3 目标检测算法 Den-YOLOv3。针对水下生物图像中存在的图像模糊、色偏严重的问题, 将 ACE 算法作用于 MSRCR 算法增强后的图像, 以优化图像颜色及亮度。同时, 将 YOLOv3 特征提取网络中的残差模块替换成密集块, 得到 Den-YOLOv3 模型, 该模型有效加强了特征传播, 提高了检测精度。实验结果表明: 优化后的 ACE-MSRCR 算法能有效增强水下图像, 进而提高检测精度。Den-YOLOv3 对四类水下生物的检测精度均高于其他模型, 在检测速度上达到 25 ms/帧, 可以满足对水下生物实时检测的要求。

关键词: 计算机应用; 水下生物检测; 水下图像; 图像增强; 深度学习

中图分类号: TP391.4 **文献标志码:** A **文章编号:** 1671-5497(2022)05-1088-10

DOI: 10.13229/j.cnki.jdxbgxb20200970

Underwater biological detection algorithm based on image enhancement and improved YOLOv3

HAO Kun¹, WANG Kuo¹, ZHAO Lu¹, WANG Bei-bei², WANG Chuan-qi³

(1. School of Computer and Information Engineering, Tianjin Chengjian University, Tianjin 300384, China; 2. School of Control and Mechanical Engineering, Tianjin Chengjian University, Tianjin 300384, China; 3. Tianjin Keyvia Electric Co., Ltd., Tianjin 300384, China)

Abstract: In order to accurately detect underwater organisms, this paper adopts the optimized MSRCR algorithm to enhance underwater images, and proposes an improved YOLOv3 target detection algorithm Den-YOLOv3 based on the DenseNet idea. The ACE algorithm is applied to the enhanced MSRCR image to optimize the color and brightness of the image, aiming at the serious problems of image blurriness and color deviation. Meanwhile, the residual module in the YOLOv3 feature extraction network was replaced with dense blocks to obtain Den-YOLOv3 model, which effectively enhanced feature propagation and

收稿日期: 2020-12-19.

基金项目: 国家自然科学基金项目(61902273).

作者简介: 郝琨(1979-), 女, 副教授, 博士. 研究方向: 水下传感器网络, 计算机视觉. E-mail: kunhao@tcu.edu.cn

通信作者: 赵璐(1983-), 女, 讲师, 博士. 研究方向: 机器学习, 深度学习. E-mail: zhaolu@tcu.edu.cn

improved detection accuracy. The experimental results show that the optimized ACE_MSRCR algorithm can effectively enhance the underwater image and improve the detection accuracy. Den-YOLOv3 has a higher detection accuracy than other models for the four types of underwater creatures, and the detection speed reaches 25 ms/frame, which can meet the requirements for real-time detection of underwater creatures.

Key words: computer application; underwater biological detection; underwater image; image enhancement; deep learning

0 引言

当前,水下生物检测技术已广泛应用于海洋生物多样性监测、生态系统健康评估以及智能水产捕捞等领域。在实际应用中,水下环境的复杂性增加了水下生物的检测难度^[1]。随着计算机视觉技术的快速发展,各国学者逐渐将深度学习应用到水下目标检测和分类中,这为水下生物检测提供了新的方向。Banan等^[2]提出一种基于VGG16(Visual geometry group 16)的快速非破坏性的方法确定鱼类种类,该方法在鱼类物种识别方面具有良好的性能,但只适用于图像背景固定且每张图像中只有一个样本的情况;Zhou等^[3]提出了一种基于卷积神经网络(Convolutional neural networks, CNN)和机器视觉的鱼的摄食强度自动分级方法,该方法能有效检测和评估鱼类食欲,但不适用于小样本数据集;Han等^[4]基于Faster R-CNN(Faster regions with convolutional neural network features)和Hypernet的方法进行改进,提出了一种深度CNN网络实现海洋生物的检测和分类,该方法虽然在一定程度上减少了部分检测时间,但整体检测时间依然较高;赵德安等^[5]提出一种基于YOLOv3(You only look once v3)的水下活蟹检测方法,该方法实现了对不同尺寸河蟹的精准检测;Hu等^[6]提出一种基于CNN的ShrimpNet来实现对对虾的识别,但ShrimpNet中CNN层较少,对深层特征提取能力不足,难以实现不同尺寸的对虾检测;Jiang等^[7]基于SSD(Single shot multibox detector)进行改进,通过增加浅层目标检测层和提高类别损失权重提高了水下目标检测精度;Song等^[8]通过将图像增强算法与Mask R-CNN(Mask regions with convolutional neural network features)结合,实现了对小样本数据集的高精度检测,但该方法检测速度低、实用性

差并且不适合远距离水下目标检测。

尽管有些目标检测算法在某些特定水下环境取得了比较好的检测效果,但在水下生物检测中依然面临一些挑战。一是水下图像能见度低,对比度低,并且存在蓝色和绿色色偏,这些特点不利于目标检测。二是对水下生物的检测既需要较高的实时性,也需要较高的精度,目前常见的检测算法难以平衡精度与速度的关系。

因此,本文将传统无监督的图像增强算法应用到水下,并进行优化,以达到提升水下图像对比度和纠正蓝绿色偏的效果。同时,本文对YOLOv3网络结构进行改进,通过采用密集块替换Darknet53中的残差模块提出一种新的网络结构Den-YOLOv3。将Den-YOLOv3在水下生物数据集上训练,以期提高水下生物检测精度,并在水下生物图像上进行检测以验证能否达到实时的要求。

1 图像增强算法

水下目标检测依赖大量的水下图像,清晰的水下图像能有效提高检测精度,对水下生物检测具有重要的影响。众所周知,大部分水下图像会存在模糊和颜色失真问题,这是由光的散射和吸收导致的。光与水中一些介质的悬浮粒子发生碰撞,传播方向发生改变,导致图像模糊,造成低对比度和雾状效果^[9]。此外,由于水对光的选择性吸收,波长较长的红光衰减最快,波长短的蓝光传播最远,所以水下图像往往呈蓝绿色色调^[10]。水下图像的这两种特点会对目标检测造成很大影响,因此需要对原始水下图像进行增强和复原。

传统无监督的图像处理可分为非物理模型增强方法和基于物理模型的复原方法^[11]。非物理模型方法主要通过调整图像的像素值改善视觉质量,如直方图均衡化^[12](Histogram equaliza-

tion, HE)、限制对比度自适应直方图均衡化^[13] (Contrast limited adaptive histogram equalization, CLAHE)、带色彩恢复的多尺度视网膜增强^[14] (Multi-scale retinex with color restoration, MSR-CR)、自动色彩均衡^[15] (Automatic color equalization, ACE)等方法。其中, CLAHE通过对图像分块进行直方图变换可有效提升图像对比度, 并利用线性插值算法提高速度, 但 CLAHE 会造成图像颜色失真, 且不能纠正水下图像色偏。MSR-CR 是基于 Retinex 理论提出的图像增强算法, Retinex 理论认为一幅图像可表示为光照量和反射量。最早的 Retinex 方法通过去除光照量得到增强图像^[16], MSRCR 通过高斯尺度算子估计图像的光照量并消除其对图像颜色影响, 并通过色彩增益与补偿对图像进行颜色恢复。但 MSRCR 在处理水下图像时, 会出现颜色失真和亮度不高的问题。ACE 通过计算目标点与其他像素点的差值, 得到像素点之间的亮度差异信息校正像素值^[15]。首先, 通过对图像的色彩域/空域调整, 完成色差校正, 接着对 RGB 三通道分别进行线性拓展与动态拉伸得到最终图像。ACE 能较好地处理图像细节, 实现色彩校正, 提升图像亮度和对比度, 具有很好的图像增强效果。而基于物理模型的方法主要根据图像退化过程进行数学建模, 并进行参数估计, 反演得到清晰的图像, 典型的有暗通道先验^[17] (Dark channel prior, DCP) 算法。

2 网络模型

水下目标检测算法可分为两种, 一种是两阶段目标检测, 标志性的算法有 R-CNN (Regions with convolutional neural network features)^[18], Fast R-CNN (Fast regions with convolutional neural network features)^[19] 和 Faster R-CNN^[20]。另一种是单级目标检测, 标志性的算法有 YOLO (You only look once)^[21-23] 和 SSD^[24]。其中, 两阶段目标检测算法在检测准确率和定位精度上有明显优势, 而单级目标检测算法在检测速度上更具

优势, 因此在实际应用场景中, 单级目标检测算法经常被用于实时检测。

2.1 YOLOv3 检测算法

YOLO 的核心思想是利用整张图作为网络的输入, 直接在输出层回归边界框的位置和边界框所属的类别。在特征提取层, 与 YOLOv1、YOLOv2 相比, YOLOv3 采用了更深的网络结构 Darknet53, 使用了大量的 3×3 和 1×1 卷积, 并借鉴残差神经网络的思想, 添加了残差模块。在检测层, YOLOv3 将输入图片划分为 $S \times S$ 个网格, 如果标注物体的中心落入网格, 那么该网格就负责检测物体。每个网格检测 n 个边界框及其置信度以及 m 个类别概率。边界框的宽度和高度是以被检测物体的中心位置相对于网格位置的偏移, 网格的坐标值均被归一化, 其中, 置信度反映是否包含物体以及包含物体情况下位置的准确性, 定义为:

$$C_i^j = P_r(Object) \times IoU_{pred}^{truth} \quad (1)$$

$$P_r(Object) \in [0, 1]$$

式中: C_i^j 为第 i 个网格中第 j 个边界框的置信度; $P_r(Object)$ 为检测边界框包含目标物体的概率; IoU_{pred}^{truth} 为两个边界框的交并比 (Intersection over union, IoU)。

2.2 密集连接网络

密集连接网络 (Dense convolutional network, DenseNet) 是由 Huang 等^[25] 提出的深层卷积神经网络, 其核心思想是以前馈的方式将每一层连接到其他每一层。传统的 L 层卷积神经网络具有 L 个连接, 而 DenseNet 具有 $L(L+1)/2$ 个直接连接。对于每一层, 它之前所有层的特征图被用作输入, 它本身的特征图被用作它所有后续层的输入。DenseNet 神经网络不仅能增强特征传播与特征重用, 同时还能有效减轻由于网络过深导致的梯度消失问题。

DenseNet 主要通过内部密集块 (Dense Block) 来实现神经网络的密集连接, 如图 1 所示。其中每个 Dense Block 内部都包含一定数量的神

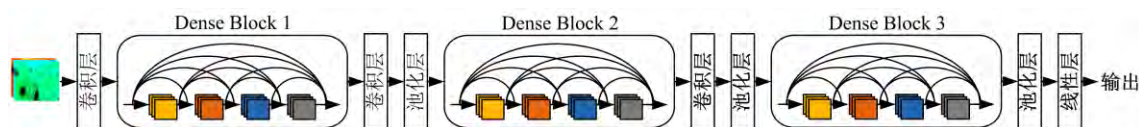


图1 DenseNet 模型结构

Fig. 1 DenseNet model structure

神经元,每个神经元主要由批归一化、激活函数和卷积层组成,每个神经元都与其之前的神经元连接。为了提高网络效率,每个神经元都会增加一个卷积层作为Dense Block的瓶颈层以减少特征图数量。

2.3 Den-YOLOv3模型

本文提出的Den-YOLOv3是以YOLOv3网络结构为基础,在特征提取层借鉴了DenseNet的

思想,采用Dense Block替换Darknet53中的残差模块,加深了网络深度,增强了特征传播,更好地实现了特征重用和表示。在检测层采用了多尺度输出方法,通过下采样输出3个不同尺度的特征图,将输出的小特征图进行上采样并与大特征图融合,使模型既能表示深层语义信息,又能表示浅层的位置信息,提高检测性能,整体网络结构如图2所示。

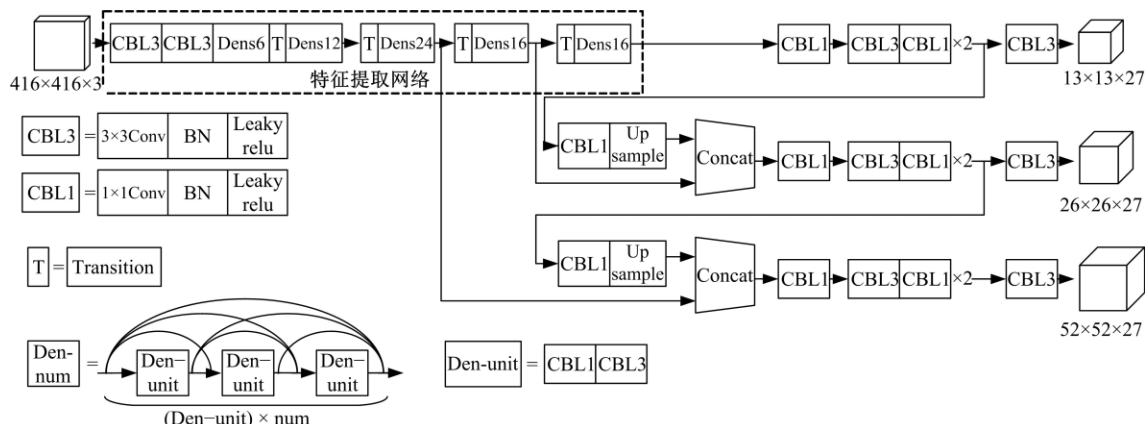


图2 Den-YOLOv3网络模型

Fig. 2 Den-YOLOv3 network model

在特征提取层,原来的Darknet53主要由重复的残差模块构成,每个残差模块由不同个数的残差单元构成,每个残差单元由一个 1×1 卷积层与 3×3 的卷积层和一个跳跃连接组成,输入 1×1 卷积层前的特征与经过 3×3 卷积后的特征通过跳跃连接相加。原来的特征提取网络包括23个残差单元,改进后的主干网络将残差模块替换为Dense Block,整个特征提取层采用了5个Dense Block,每个Dense Block分别包括6、12、24、16、16个密集层(Dense layer),Dense Block之间采用步幅为2的卷积进行下采样,图3展示了本文采用的特征提取网络。公式(2)(3)所示为残差模块与Dense Block的特征传递方式,相比于残差模块的恒等映射后相加的特征传递方式,Dense Block的传递方式融合了当前层之前的所有特征,特征信息更丰富,鲁棒性能更好。

$$X_l = H_l(X_{l-1}) + X_{l-1} \quad (2)$$

$$X_l = H_l[X_0, X_1, \dots, X_{l-1}] \quad (3)$$

式中: l 为网络层数; X_l 为第 l 层输出; $H(\cdot)$ 为每层的非线性函数; $[\cdot]$ 为各层的密集连接。

在检测层中,Den-YOLOv3采用了多尺度预测方法,图片以 416×416 的尺寸输入模型,经过

	类型	卷积核	尺寸	输出
6×	Convolutional	32	3×3	416×416
	Convolutional	64	$3 \times 3/2$	208×208
	Dense Layer	64 32	1×1 3×3	208×208
	Convolutional	128	$3 \times 3/2$	104×104
	Dense Layer	64 32	1×1 3×3	104×104
	Convolutional	256	$3 \times 3/2$	52×52
12×	Dense Layer	64 32	1×1 3×3	52×52
	Convolutional	512	$3 \times 3/2$	26×26
	Dense Layer	64 32	1×1 3×3	26×26
	Convolutional	512	$3 \times 3/2$	13×13
	Dense Layer	64 32	1×1 3×3	13×13
	Convolutional	512	$3 \times 3/2$	13×13

图3 本文提出的密集架构

Fig. 3 Dense architecture proposed in this paper

下采样生成 13×13 的特征图并将此特征图用于小目标检测。Den-YOLOv3将第237层输出的 13×13 特征图进行上采样得到 26×26 的特征图,将其与第178层输出的 26×26 特征图进行融合,并经过 1×1 与 3×3 的卷积进行降维,最终输出 26×26 的特征图用于中型目标检测。将第249层输出的 26×26 特征图上采样得到 52×52 的特征图,将其与第129层输出的特征图融合并经过卷

积降维,最终输出 52×52 的特征图用于大目标检测。

Den-YOLOv3 在多尺度预测方法中使用了先验框机制,每个网格单元使用 9 个先验框。通过采用 K-means 算法对先验框大小进行聚类,其中 K-means 算法采用的距离如下所示:

$$d(box, centroid) = 1 - IoU(box, centroid) \quad (4)$$

式中: d 为距离误差; IoU 为交并比; box 为标注的边界框; $centroid$ 为聚类中心的边界框。

聚类后的先验框大小为 (19.84, 31.36)、(33.28, 52.16)、(49.92, 80)、(75.84, 117.12)、(111.68, 152.32)、(133.76, 224)、(207.68, 256.32)、(248.96, 400.64)、(404.8, 563.84)。

3 实验设置

3.1 数据集

本文采用 2020 年中国水下机器人大赛(湛江)官方数据集,所有图像均拍摄于水下真实环境,每张图片包含不同的目标。整个数据集包括 4 种类型的目标:海参、海胆、扇贝和海星。本文将数据集转化为 VOC 格式进行训练,数据集共有 5543 张 jpg 格式的水下光学图像,其中 4743 张图像用于训练集,800 张图像用于测试集。

3.2 实验平台

实验硬件配置为 Intel(R) Core(TM) i7-7800X CPU@3.50 GHz,模型在 2 个具有 11 GB 显存的 GeForce GTX 1080 Ti 上进行训练。系统采用 Ubuntu 18.04 LTS, CUDA 版本为 10.1, cudnn 版本为 7.1。

3.3 网络训练

模型的输入图像大小为 416×416 ,初始学习率为 0.0001,动量为 0.9,权重衰减为 0.0005,批处理大小为 16。模型总共训练了 300 个 epoch,并分别在 210 个 epoch 与 270 个 epoch 时学习率降为原来的 10%。

3.4 评价指标

为了选取合适的模型,本文主要采用平均精度 AP (Average precision) 和平均精度均值 mAP (mean average precision) 作为评价指标。P-R 曲线是以精准率 P (Precision) 和召回率 R (Recall) 这两个为变量而做出的曲线,而 AP 是由计算 P-R 曲线下覆盖的面积获得, mAP 是对所有类别的

AP 进行平均获得:

$$P = TP / (TP + FP) \quad (5)$$

$$R = TP / (TP + FN) \quad (6)$$

$$AP = \int_0^1 p(r) dr \quad (7)$$

$$mAP = \frac{1}{m} \sum_{i=1}^m AP \quad (8)$$

式中: TP (True positive) 为预测的正样本中包含的正样本数量; FP (False positive) 为预测的正样本中包含的负样本数量; FN (False negative) 为预测的负样本中包含的正样本数量; $p(r)$ 为 P-R 曲线; m 为所检测目标类别的数量。

4 实验结果分析

4.1 水下图像增强

本文分别选取了非物理模型方法中的 CLAHE、MSRCR、ACE 和基于物理模型的 DCP 进行实验,图像处理效果如图 4 所示,可见,经 CLAHE 和 MSRCR 处理过的图像,对比度均有所提升,但经 CLAHE 增强后的图像色偏问题依然严重;经 MSRCR 增强后的图像颜色失真且图像偏暗;经 ACE 增强后图像亮度和颜色均有所改善,但整体上依然呈蓝绿色色调且亮度较低;经 DCP 处理过的图像亮度有所提升,但色偏更严重。

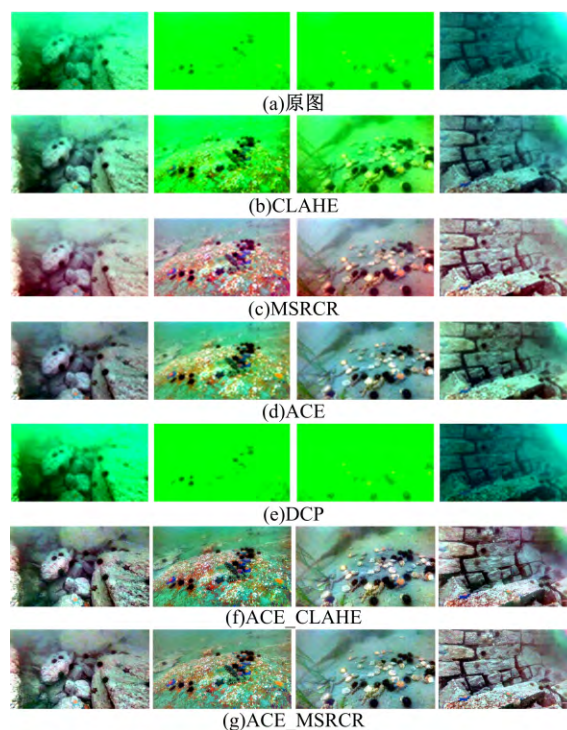


图 4 不同算法的效果对比

Fig. 4 Effect comparison of different algorithms

结合上述结果,ACE直接用于水下图像增强不能取得最佳效果,但能解决CLAHE和MSRCR带来的亮度低以及颜色失真问题,因此将经CLAHE和MSRCR增强后的图像再利用ACE进行处理,以达到对CLAHE和MSRCR优化的目的。ACE不能解决DCP带来的更严重的色偏问题,因此不采用ACE对DCP优化。

优化后的图像增强效果如图4(f)(g)所示,可见,对CLAHE优化后的ACE_CLAHE增强效果显著,图像主体和背景区分明显,色偏基本消失,图像饱和度较高。采用ACE_MSRCR增强后的图像,对比度和暗部区域细节进一步提高,图像整体明亮,取得了很好的视觉效果。本文分别对两种优化后算法进行实验,结果表明用ACE_MSRCR处理过的图像,检测效果更好,因此,最终选择ACE_MSRCR作为水下图像增强算法。

4.2 不同水下图像增强算法检测结果对比

为验证不同的水下图像增强算法对检测结果的影响,采用不同算法增强后的图像分别输入YOLOv3网络中训练并进行测试,实验结果如表1所示,可见,与原图相比,采用经过CLAHE、MSRCR与ACE增强后的图像训练, mAP 均能提高1个百分点左右。其中,对海参的检测精度,

表1 不同图像增强算法对YOLOv3检测结果的影响

Table 1 Influence of different image enhancement algorithms on the detection results of YOLOv3

算 法	AP/%				mAP /%
	海参	海胆	扇贝	海星	
原图	63.5	87.7	66.1	80.4	74.4
CLAHE	66.5	88.2	66.1	81.1	75.5
MSRCR	66.8	88.2	67.8	79.9	75.6
ACE	65.8	87.6	67.8	79.9	75.3
DCP	61.8	87.9	65.7	79.6	73.8
ACE_CLAHE	68.6	87.2	68.5	81.0	76.3
ACE_MSRCR	69.9	87.4	69.1	80.7	76.8

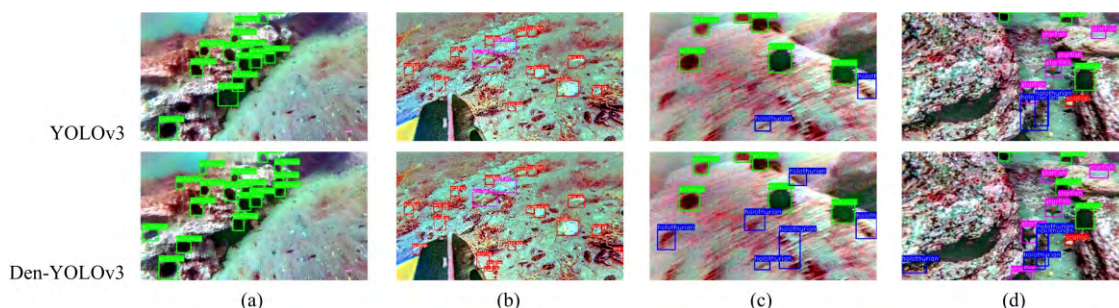


图5 YOLOv3与Den-YOLOv3检测效果对比

Fig. 5 Detection effect of Den-YOLOv3 was compared with YOLOv3

CLAHE提高了3%,ACE提高了2.3%,MSRCR提高了3.3%,这说明增强后的图像更有利于目标检测。同时,经过DCP处理的图像,不同类别的AP均降低, mAP 降低了0.6%,因此DCP并不适用于水下图像增强与复原。

在由4种算法对图像进行预处理的基础上,优化后的ACE_CLAHE与ACE_MSRCR也表现出良好的效果。经ACE_CLAHE增强的图像,与原图相比,海参AP提高了5.1%,扇贝AP提高了2.4%,整体 mAP 提高了1.9%。经过ACE_MSRCR增强的图像,与原图相比, mAP 提高了2.4%,与MSRCR相比, mAP 提高了1.2%。总体上ACE_MSRCR取得最优的效果,因此本文最终选用ACE_MSRCR增强后的水下图像进行后续实验。

4.3 密集块的有效验证

本节验证了密集块对检测结果的影响,如表2所示,与原始YOLOv3相比,Den-YOLOv3对4种水下生物的检测精度均有所提升,其中对海参的检测精度提升了2.9%,对扇贝的检测精度提升了2.5%,总体上 mAP 提高了1.9%。在检测速度上,YOLOv3网络越深检测时间越长,虽然增加了少量检测时间,但Den-YOLOv3仍旧满足实时检测。

图5为YOLOv3与Den-YOLOv3检测效果对比图,可见YOLOv3模型存在漏检情况,图5

表2 YOLOv3与Den-YOLOv3检测结果对比

Table 2 Comparison of detection results between YOLOv3 and Den-YOLOv3

模型	AP/%				mAP /%	时间/ms
	海参	海胆	扇贝	海星		
YOLOv3	69.9	87.4	69.1	80.7	76.8	15
Den-YOLOv3	72.8	88.6	71.6	81.9	78.7	25

(a)中漏检了部分海胆,图5(b)(c)漏检了部分扇贝及海参,而Den-YOLOv3检测数量更多,对小目标以及有遮挡的水下生物检测更全。图5(d)中存在多个海参聚集,但YOLOv3只检测出部分海参,相比于YOLOv3,Den-YOLOv3检测效果更好。本文通过使用密集块能有效提升网络提取特征能力,因此Den-YOLOv3检测和分类的精度更高。

为进一步验证密集块的作用,本文分别以不同的迭代次数对YOLOv3与Den-YOLOv3进行训练,实验结果如图6(a)所示,可见随着迭代次数增加,YOLOv3与Den-YOLOv3模型的检测效果均有所提高,但总体上Den-YOLOv3的 mAP 均高于YOLOv3,这体现了密集块的有效性。与残差模块相比,密集块能提取到更丰富的特征,即便在迭代次数很低的情况下,依然能得到比YOLOv3模型更高的精度。同时,由于水下环境多采用小样本训练集,因此将原始数据集划分成不同大小,并验证了YOLOv3和Den-YOLOv3在不同大小训练集下的检测结果。原始训练集中包括4743张水下图像,本文按照1/2、1/4、1/8的比例从原始训练集中随机选取水下生物图像组成新

的训练集,新的训练集分别包括2371、1185、592张水下图像。模型经过新的训练集训练后,统一在原始800张测试集上进行测试。图6(b)所示,在由592张水下图像组成的训练集中,Den-YOLOv3的 mAP 比YOLOv3高2%,在1185张水下图像组成训练集中,Den-YOLOv3的 mAP 比YOLOv3高3.2%,在2371张水下图像组成训练集中,Den-YOLOv3的 mAP 比YOLOv3高3%。在小样本训练集上,和YOLOv3相比,Den-YOLOv3具有更好的检测结果,因此Den-YOLOv3具有更好的泛化能力。

4.4 不同算法检测效果对比

为验证Den-YOLOv3的检测性能,将Den-YOLOv3与其他算法进行对比,实验结果如表3所示,可见在单类水下生物检测中,Den-YOLOv3对海参的检测 AP 为72.8%,低于EfficientDet-D0^[26],但对其他三类水下生物的检测精度均高于EfficientDet-D0,这表明Den-YOLOv3在多分类水下生物检测中更具优势。此外,数据集中的扇贝多为泥土遮挡,是较难检测的一类水下生物,而Faster R-CNN、RetinaNet^[27]和Den-YOLOv3能取得比其他算法更高的精度,这是由于它们具有更深的特征提取网络,能提取到更深层的语义信息,并且Den-YOLOv3的特征提取效果好于Faster R-CNN和RetinaNet,取得了最高的精度。

总体上Den-YOLOv3的 mAP 为78.7%,与SSD、EfficientDet-D0、RetinaNet相比分别提高了3.4%、4.9%、2.9%。其他算法中Faster R-CNN的 mAP 为78.3%,与Den-YOLOv3最接近,但Den-YOLOv3的速度是Faster R-CNN的2.7倍。YOLO-Tiny虽然速度最快,但精度比Den-

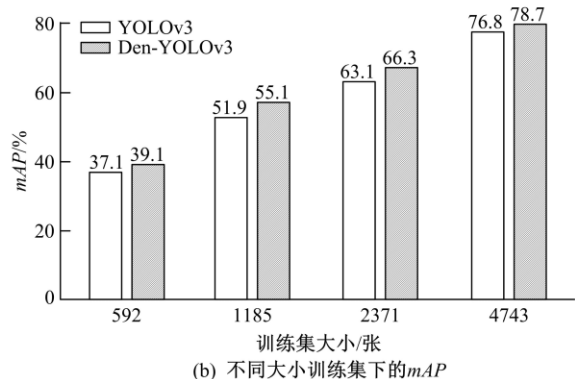
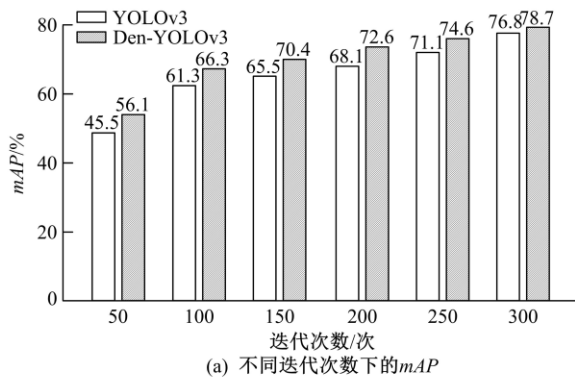


图6 不同情况下的 mAP 对比图

Fig. 6 mAP comparisons in different situations

表3 不同算法检测结果对比

Table 3 Detection results of different algorithms are compared

模 型	AP/%				mAP /%	时间 /ms
	海参	海胆	扇贝	海星		
Faster R-CNN	70.0	88.8	71.3	83.0	78.3	68
SSD	69.3	84.5	66.9	80.5	75.3	43
EfficientDet-D0	76.1	85.7	54.0	79.5	73.8	22
RetinaNet	65.8	85.4	71.0	80.9	75.8	81
YOLOv3-Tiny	53.8	73.7	44.5	71.1	60.8	3
Den-YOLOv3	72.8	88.6	71.6	81.9	78.7	25

YOLOv3低了17.9%,这是由于YOLO-Tiny使用了很多池化层,损失了大量的特征信息,而Den-YOLOv3通过使用密集块丰富了特征,获得了较高的精度。

图7展示了不同算法的检测效果图,为了直观比较检测结果,本文采用原始标注图片作为对比,可见SSD和YOLO-Tiny的检测效果明显低于其他算法,尤其是对扇贝的检测存在明显漏检

情况,同时EfficientDet-D0对扇贝的检测也表现出很差的效果,如图7(d)所示。RetinaNet能检测出大部分水下生物,但定位不准确。而Den-YOLOv3的检测效果与Faster R-CNN基本相同,图7(a)中海胆分布较密集,且目标较大,图7(b)(c)中的目标较小,可以看到Den-YOLOv3均能有效识别,这是由于Den-YOLOv3使用了多尺度预测方法,提高了对不同大小目标检测的检测性能。

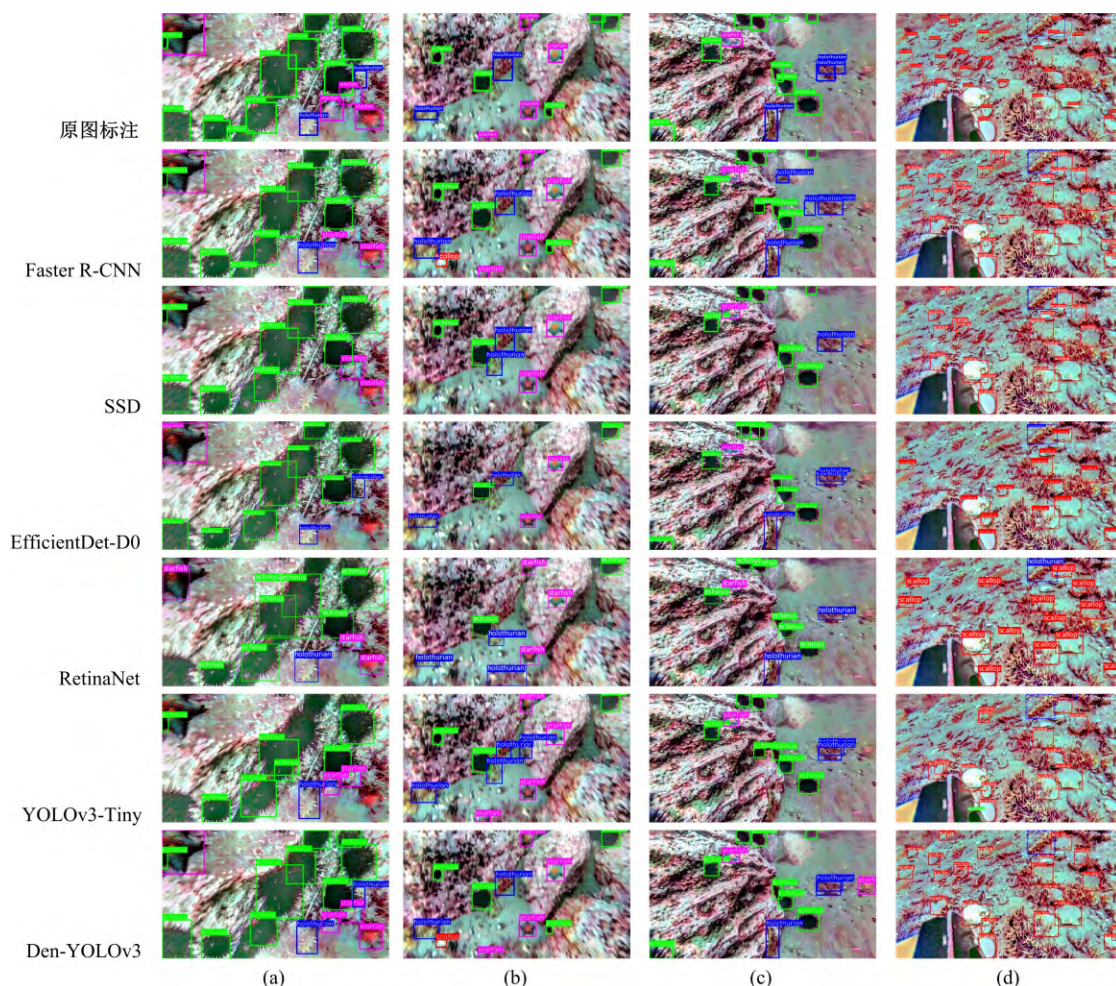


图7 不同算法检测效果对比

Fig. 7 Detection effect of different algorithms is compared

5 结束语

针对水下图像对比度低、色偏严重的情况,提出了有效的水下图像增强方法。通过对比多种传统的图像增强方法,最终选择优化后的ACE-MSRCR算法对水下生物图像进行增强,增强后的水下图像提高了对比度,消除了色偏,更有利于特征学习,并最终提高了2.4%的平均精度。基于YOLOv3目标检测模型,本文提出了一种改进

YOLOv3的水下生物检测模型Den-YOLOv3。通过采用密集块替换YOLOv3中的残差模块,提高了特征提取效率,使更丰富的特征信息在网络中传递,最终提升了1.9%的平均精度。总体上,Den-YOLOv3模型检测精度达到78.7%,速度达到25 ms/帧,能满足水下生物实时检测的条件,具有良好的检测性能。后续工作中,将在满足精度与速度的条件下,压缩网络模型,将其成果应用在嵌入式设备中。

参考文献:

- [1] 强伟,贺昱曜,郭玉锦,等. 基于改进SSD的水下目标检测算法研究[J]. 西北工业大学学报, 2020, 38(4): 747-754.
- Qiang Wei, He Yu-yao, Guo Yu-jin, et al. Research on underwater target detection algorithm based on improved SSD[J]. Journal of Northwestern Polytechnical University, 2020, 38(4): 747-754.
- [2] Banan A, Nasiri A, Taheri-Garavand A. Deep learning-based appearance features extraction for automated carp species identification[J]. Aquacultural Engineering, 2020, 89: 102053.
- [3] Zhou C, Xu D, Chen L, et al. Evaluation of fish feeding intensity in aquaculture using a convolutional neural network and machine vision[J]. Aquaculture, 2019, 507: 457-465.
- [4] Han F, Yao J, Zhu H, et al. Marine organism detection and classification from underwater vision based on the deep CNN method[J]. Mathematical Problems in Engineering, 2020: 3937580.
- [5] 赵德安,刘晓洋,孙月平,等. 基于机器视觉的水下河蟹识别方法[J]. 农业机械学报, 2019, 50(3): 151-158.
- Zhao De-an, Liu Xiao-yang, Sun Yue-ping, et al. Detection of underwater crabs based on machine vision[J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(3): 151-158.
- [6] Hu W C, Wu H T, Zhang Y F, et al. Shrimp recognition using ShrimpNet based on convolutional neural network[J]. Journal of Ambient Intelligence and Humanized Computing, 2020(12): 1-8.
- [7] Jiang Z, Wang R. Underwater object detection based on improved single shot multibox detector[C]//Conference Proceeding of 2020 3rd International Conference on Algorithms, Computing and Artificial Intelligence, Sanya City, 2020: 1-7.
- [8] Song S, Zhu J, Li X, et al. Integrate MSRCR and Mask R-CNN to recognize underwater creatures on small sample datasets[J]. IEEE Access, 2020, 8: 172848-172858.
- [9] 王德兴,吴若有,袁宏春,等. 基于多尺度注意力融合和卷积神经网络的水下图像恢复[J]. 吉林大学学报:工学版, 2021, 51(4): 1396-1404.
- Wang De-xing, Wu Ruo-you, Yuan Hong-chun, et al. Underwater image restoration based on gradient channel and optimize[J]. Journal of Jilin University (Engineering and Technology Edition), 2021, 51(4): 1396-1404.
- [10] 郭继昌,茹丽,郭春乐. 基于梯度通道和优化透射率的水下图像复原[J]. 吉林大学学报:工学版, 2020, 50(4): 1435-1442.
- Guo Ji-chang, Ru Li, Guo Chun-le. Underwater image restoration based on gradient channel and optimize[J]. Journal of Jilin University(Engineering and Technology Edition), 2020, 50(4): 1435-1442.
- [11] Han F, Yao J, Zhu H, et al. Underwater image processing and object detection based on deep CNN method[J]. Journal of Sensors, 2020(9): 1-20.
- [12] Hummel R. Image enhancement by histogram transformation[J]. Computer Graphics and Image Processing, 1977, 6(2): 184-195.
- [13] Zuiderveld K. Contrast limited adaptive histogram equalization[J]. Graphics Gems, 1994: 474-485.
- [14] Jobson D, Rahman Z. A multiscale retinex for bridging the gap between color images and the human observation of scenes[J]. IEEE Transactions on Image Processing, 1997, 6(7): 965-976.
- [15] Getreuer P. Automatic color enhancement (ace) and its fast implementation[J]. Image Processing on Line, 2012(2): 266-277.
- [16] Song X, Huang J, Cao J, et al. Multi-scale joint network based on Retinex theory for low-light enhancement[J]. Signal Image and Video Processing, 2021, 15(6): 1257-1264.
- [17] He K, Sun J, Tang X. Single image haze removal using dark channel prior[C]//IEEE Conference on Computer Vision and Pattern Recognition, Miami, 2009: 1956-1963.
- [18] Girshick R, Donahue J, Darrelland T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, 2014: 580-587.
- [19] Girshick R. Fast R-CNN[C]//IEEE International Conference on Computer Vision, Santiago, 2015: 1440-1448.
- [20] Ren S, He K, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 39(6): 1137-1149.
- [21] 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 Vi-

- sion and Pattern Recognition, Las Vegas, 2016: 779-788.
- [22] Redmon J, Farhadi A. YOLO9000: better, faster, stronger[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, 2017: 6517-6525.
- [23] Redmon J, Farhadi A. YOLOv3: an incremental improvement[J]. Computer Vision and Pattern Recognition, arXiv: 1804.02767.
- [24] Liu W, Anguelov D, Erhan D, et al. SSD: single shot multibox detector[C]//European Conference on Computer Vision, Amsterdam, 2016: 21-37.
- [25] Huang G, Liu Z, Maaten L, et al. Densely connected convolutional networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, 2017: 4700-4708.
- [26] Tan M, Pang R, Le V. Efficientdet: scalable and efficient object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Seattle, 2020: 10781-10790.
- [27] Lin T, Goyal P, Girshick R, et al. Focal loss for dense object detection[C]//Proceedings of the IEEE International Conference on Computer Vision, Venice, 2017: 2980-2988.