

自动色阶与双向特征融合的水下目标检测算法

杨婷¹, 高武奇², 王鹏^{3*}, 李晓艳⁴, 吕志刚⁴, 邸若海⁴

¹西安工业大学兵器科学与技术学院, 陕西 西安 710021;

²西安工业大学计算机科学与工程学院, 陕西 西安 710021;

³西安工业大学发展规划处, 陕西 西安 710021;

⁴西安工业大学电子信息工程学院, 陕西 西安 710021

摘要 水下环境存在光线差、噪声大等复杂情况, 导致传统水下目标检测方法检测精度较低、漏检率较高。针对上述问题, 在现阶段通用的 Faster R-CNN 的基础上, 提出一种自动色阶与双向特征融合的水下目标检测算法。首先, 采用自动色阶对水下模糊图像进行增强处理; 其次, 采用 PAFPN 进行双向特征融合, 以增强对浅层信息的表达能力; 然后, 在训练前后均引入柔性非极大值抑制 (Soft-NMS) 算法, 来修正并生成候选目标区域; 最后, 采用 FocalLoss 函数, 解决正负样本分配不均衡的问题。实验结果表明, 所提算法在 URPC2020 数据集上的检测准确率可达 59.7%, 召回率可达 70.5%, 相比现阶段通用的 Faster R-CNN 算法, 分别提高了 5.5 个百分点和 8.4 个百分点, 有效提高了水下目标检测的准确率。

关键词 目标检测; 图像增强; 特征金字塔; 柔性非极大值抑制; FocalLoss 函数

中图分类号 TP391.9

文献标志码 A

DOI: 10.3788/LOP213139

Underwater Target Detection Algorithm Based on Automatic Color Level and Bidirectional Feature Fusion

Yang Ting¹, Gao Wuqi², Wang Peng^{3*}, Li Xiaoyan⁴, Lü Zhigang⁴, Di Ruohai⁴

¹School of Ordnance Science and Technology, Xi'an Technological University, Xi'an 710021, Shaanxi, China;

²School of Computer Science and Technology, Xi'an Technological University, Xi'an 710021, Shaanxi, China;

³Development Planning Office, Xi'an Technological University, Xi'an 710021, Shaanxi, China;

⁴School of Electronic Information Engineering, Xi'an Technological University, Xi'an 710021, Shaanxi, China

Abstract Many complex elements such as poor light and high noise in the underwater environment result in low detection accuracy and high missed detection rate in traditional underwater target detection methods. To address these issues, based on the current general Faster R-CNN algorithm, this study proposes an underwater target detection algorithm based on automatic color level and bidirectional feature fusion. First, the automatic color level was used to enhance a blurred underwater image. Second, the path aggregation feature pyramid network (PAFPN) was introduced for feature fusion to enhance the expression for shallow information. Third, the soft non-maximum suppression (Soft-NMS) algorithm was introduced to modify and generate the candidate target regions before and after training. Finally, the FocalLoss function was used to rectify the issue of an unbalanced distribution of positive and negative samples. The experimental results show that the proposed algorithm can reach a detection accuracy of 59.7% on the URPC2020 dataset and a recall rate of 70.5%, which are 5.5 percentage points and 8.4 percentage points respectively higher than the current general Faster R-CNN algorithm, effectively improving the average accuracy of underwater target detection.

Key words target detection; image enhancement; feature pyramid; soft non-maximum suppression; FocalLoss function

1 引言

水下图像作为海洋信息传递的主要载体, 在海洋

的勘探和开发中发挥着重要作用, 水下目标检测即对水下有用目标的定位与识别。然而, 由于水下环境构成复杂、水质浑浊, 存在光的折射与吸收等问题, 水下

收稿日期: 2021-12-02; 修回日期: 2022-01-02; 录用日期: 2022-01-21; 网络首发日期: 2022-01-30

基金项目: 国家自然科学基金(62171360); 陕西省科技厅重点研发计划(2022GY-110)

通信作者: *wp_xatu@163.com

图像存在较为明显的模糊和失真^[1],这给水下目标检测和识别带来巨大的困难与挑战。

传统水下目标检测方法在对目标特征进行有效提取之后将其输入分类器进行检测识别^[2-3]。张志强等^[4]首先提取水下鱼类图像的颜色、长度等特征值,然后对种类进行检测和识别。姚润璐等^[5]在对图像进行背景剔除得到彩色的鱼体图像之后,对其进行分割,提取特征,最后使用 back propagation(BP)神经网络对目标种类进行识别。Rova等^[6]提出了一种用于水下鱼类分类的可变形模板目标识别方法。与原始基于支持向量机纹理的分类方法相比,这些方法大大提高了识别率,然而均只适用于水下环境简单、光照充足、图像清晰、背景对比度高的情况,当受水下生物影响、水下环境异常复杂时,会出现大量误检与漏检,因此传统方法识别效果较差。

近年来,计算机技术和人工智能神经网络发展迅速,水下图像处理逐渐应用到机器学习与深度学习的领域^[7]。王伟等^[8]基于 SSD 算法,将 VGG 网络替换为 ResNet,并利用深度分离可变形卷积模块进行特征提取,提高了整体检测的速度与准确率。刘萍等^[9]将训练好的 UGAN-P 生成网络加入 YOLOv3 中,并通过

K-means 聚类算法来修正 Anchor Box,提高了检测精度。Chen 等^[10]提出了一种新的样本加权超网络,用高分辨率和语义丰富的超特征映射来提高小目标检测的准确性,但同时增加了时间复杂度。宋绍剑等^[11]首先利用带色彩恢复的多尺度颜色增强(MSRCR)算法^[12]增强水下模糊图像,然后用 Mask R-CNN 迁移学习实现水下生物实例分割。徐凤强等^[13]设计了一种基于反馈机制的多信号分析方法,引导机器人自动检测目标并完成自主抓取操作。但是在复杂的海洋环境中,水下暗流、礁石等因素会影响机器人的作业效率。

综上所述,传统目标检测方法和深度学习检测方法均会在背景复杂、图像模糊、目标分布密集时出现检测正确率低、漏检率高等问题。因此,本文针对以上问题以普遍改进的 Faster R-CNN 为基准,提出一种自动色阶与双向特征融合的水下目标检测算法。

2 算法原理

Faster R-CNN^[14]的基本结构如图 1 所示。其主要由卷积神经网络^[15]、区域建议网络(RPN)、感兴趣区域池化(ROI pooling)和回归分类 4 部分组成。

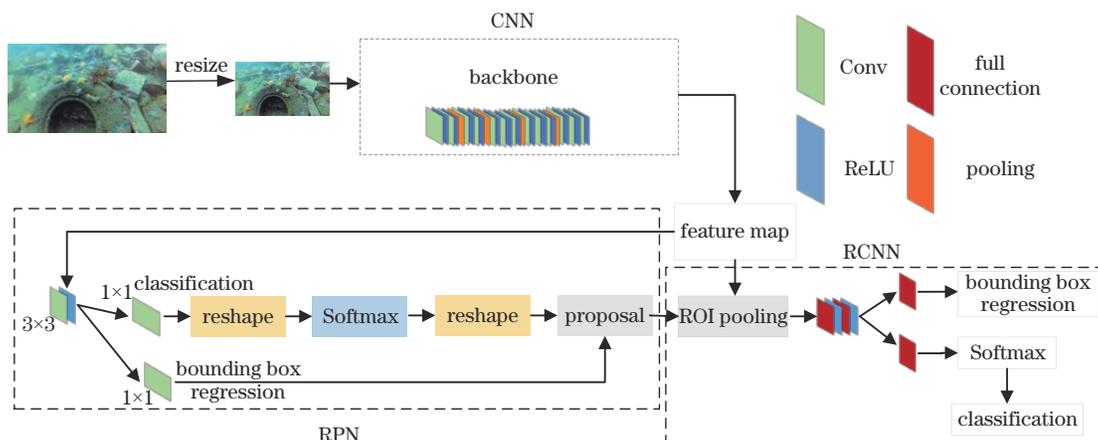


图 1 Faster R-CNN 的基本结构

Fig. 1 Basic structure of Faster R-CNN

1) 特征提取网络中,输入图片通过具有卷积、池化等操作的特征提取网络(如 ZFNET、VGG16^[16]、ResNet^[17]等)后,对图像特征进行提取,提取的特征被用于后续 RPN 和全连接层。

2) RPN 可以在特征图上生成候选区域。使用分类器判别目标前景和背景,利用非极大值抑制(NMS)^[18]等策略对候选区域位置进行矫正,生成最终的候选目标区域。

3) ROI pooling 用于池化固定尺寸的特征图,以便于后续的目标定位与分类。

4) 回归分类中,采用 L1Loss 函数,通过回归算法对目标进行分类,得到最终目标检测框的边界坐标。

在深度卷积神经网络中,一般通过引入深度残差

网络来解决深层网络结构^[19]带来的梯度消失或者爆炸现象。因此本文根据水下目标尺度较小且图像分辨率较低的特性,把原始 Faster R-CNN 的 VGG16 网络替换为 ResNet50 网络。ResNet50 具有更多的网络层数和更少的运算量,从而能够获取到更为丰富的目标特征。

对于原始 Faster R-CNN 中 ROI pooling 经过两次浮点数取整所带来的量化误差,引入感兴趣区域校准(ROI align)方法。ROI align 使用双线性插值法进行插值,能避免区域不匹配问题,不去掉浮点数,从而减小了浮点数取整所造成的精度损失,使得模型能获得更加准确的候选特征区域。

3 所提算法内容

为了提高对水下目标的检测准确率和召回率,进行如下改进:1)采用自动色阶图像增强算法对水下模糊图像进行增强处理,便于后续目标检测;2)引入 path aggregation feature pyramid network (PAFPN) 进行深层语义特征与浅层边缘特征的融合,增强检测网络对小目标的检测能力;3)针对传统 NMS 算法会

直接过滤置信度较低的候选框的问题,在训练和验证时均采用柔性非极大值抑制(Soft-NMS)算法,用较低的置信度代替原来的置信度,从而解决 NMS 直接过滤的问题,修正候选目标的位置,最终生成候选目标区域;4)采用 FocalLoss 函数来解决数据集中正负样本分布不均衡的问题。所提算法的整体流程如图 2 所示,其中加粗边框部分为所提改进模块。

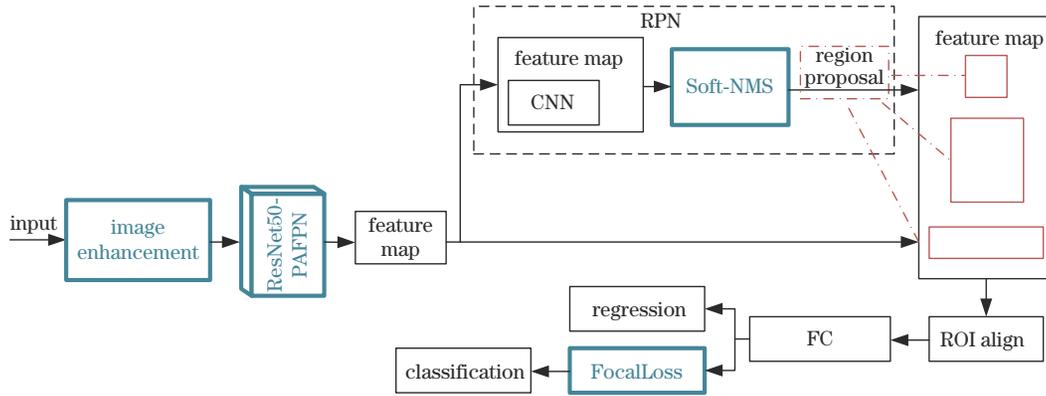


图 2 所提算法的整体网络结构

Fig. 2 Overall network structure of the proposed algorithm

3.1 自动色阶图像增强

选用水下机器人比赛数据集作为水下检测目标,整个数据集场景复杂、色调变化大、目标大小不同、丢失细节较多。在水流、杂质和不均匀光照的影响下,与常见图像相比,水下图像亮度较暗,直方图分布集中,因此检测和识别较为困难。为了解决这个问题,引入

了自动色阶增强算法对模糊图像进行增强处理,可得到更加清晰的水下目标数据集。

自动色阶法通过去除过高和过低像素值,用线性映射或伽马矫正把其余像素值矫正至 $[0, 255]$,来分散图像直方图,增强亮度。其中,线性映射的关系为

$$\begin{cases} f(u) = 0, & u \leq V_{\min} \\ f(u) = 255, & u \geq V_{\max} \\ f(u) = [(u - V_{\min}) / (V_{\max} - V_{\min})] \times 255, & V_{\min} < u < V_{\max} \end{cases} \quad (1)$$

如式(1)所示,首先统计单通道直方图,求出 0.5% 首尾的阈值,记为 V_{\min} 和 V_{\max} ,将低于 V_{\min} 的像素值赋为 0,高于 V_{\max} 的像素值赋为 255,然后对图像其余像素值按照式(1)进行校正^[20]。对红、绿、蓝三个

通道都进行上述的操作后,占比较大的灰度值的动态范围将会扩宽,得到一个对比度增强的彩色图像,其更加清晰,细节更加明显,方便后续的目标识别工作。

图 3 给出了原数据集图像的直方图和自动色阶增

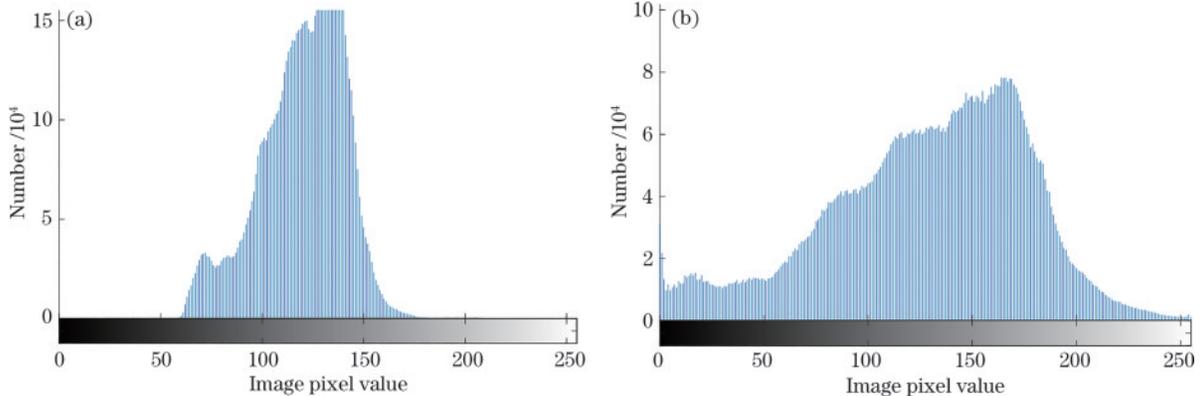


图 3 原图与增强后图像的直方图。(a)原图;(b)增强后图像

Fig. 3 Histograms of original image and enhanced image. (a) Original image; (b) enhanced image

强后图像的直方图。由图 3(a)可知,原图像直方图分布集中,像素集中在 100~150,因此图像整体偏暗,难以观察目标。分析图 3(b)可知,自动色阶增强后的直方图较为分散,像素分布均匀,整体亮度得到增强,便于后续视觉观察与分析。

3.2 特征提取网络(PAFPN)

针对水下环境复杂,网络会丢失较多底层特征纹理信息的问题,引入 PAFPN,利用增加的自底向上的特征融合路径来减少底层特征向顶层传播的损失,且通过两次横向连接,可以更好地保留底层特征的信息。

PANet 特征提取网络^[21]属于简单的双向融合特征提取网络。相对于单向融合的 FPN,PAFPN 增加了从底层向上的特征融合途径,同时使用底层的精确定位信息,减少了信号的传递途径,从而降低了底层特征信息向顶层传递的损失;同时每个建议框都使用了金字塔所有层的特性信息,减少了对建议框的随意分配过程,通过全连接层融合,扩大了预测信息的来源。PAFPN 结构如图 4 所示。

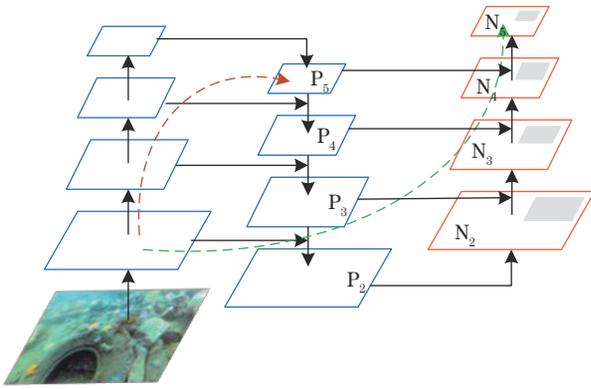


图 4 PAFPN 结构

Fig. 4 Structure of PAFPN

PAFPN 在 FPN 的基础上添加了一个自底向上的传输路径,将底层信息直接向上传输,以增加底层信息的利用率。在图 4 中,左边的虚线箭头说明在 FPN 算法中,将浅层的特征传递到顶层要经历数十或一百多个网络层,经过自底向上的多层传递,会丢失较多的浅层特征信息。右边的虚线箭头即为 PAFPN 新添加的 bottom-up path augmentation 结构,浅层特征横向连接到 P_2 ,再从 P_2 沿着 bottom-up path augmentation 传递到顶层,如此一来,信息传递的网络层数不到 10 层,因此便能较好地保留浅层特征信息。

其中,bottom-up path augmentation 的具体结构如图 5 所示。每一个特征层通过横向连接获得一个更高分辨率的特征图 N_i ,特征图经过步长为 2、尺寸为 3×3 的卷积层下采样为原来的一半,再对其与横向连接传递过来的具有较低分辨率的特征图 P_{i+1} 进行逐元素相加,融合后的特征映射再经过 3×3 的卷积

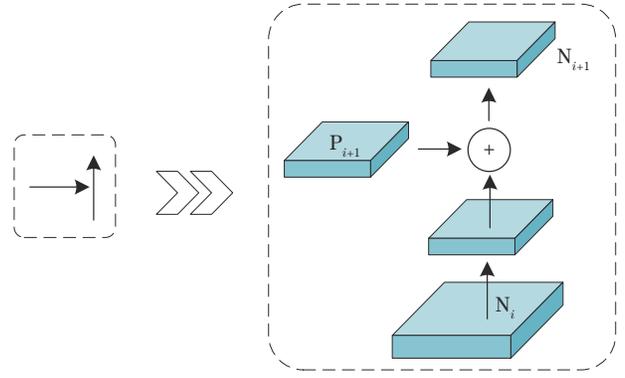


图 5 自底向上模块的结构

Fig. 5 Structure of bottom-up module

处理,生成后续子网络的 N_{i+1} 。这个迭代过程最后融合得到的特征层是 N_2 、 N_3 、 N_4 、 N_5 ,如图 4 所示,并将它们用于后续预测框的分类和回归。其中, \oplus 代表相同尺度的特征图在像素级别上的相加, $N_2 \sim N_5$ 组成特征金字塔。

3.3 Soft-NMS 算法

传统 NMS 算法采用硬阈值判断的方法判断相邻阈值是否保留。其在进行多余候选框筛选时存在的明显缺陷是如果图像上同一个区域目标重叠度较高,则检测框分数被置为 0,这会导致对该目标检测失败,从而降低算法的平均准确率。对此,将传统 NMS 算法替换成 Soft-NMS 算法^[22]用于对候选框的筛选与优化。

Soft-NMS 算法由 NMS 算法改进而来,其分数重置函数为

$$S_i = \begin{cases} S_i, & R_{IoU}(M, b_i) < N_t \\ S_i [1 - R_{IoU}(M, b_i)], & R_{IoU}(M, b_i) \geq N_t \end{cases}, (2)$$

式中: S_i 为第 i 个检测框的得分; M 为得分最大的检测框; b 为检测框集合, b_i 为 b 集合中第 i 个检测框; N_t 为自设定的一个阈值。在式(2)中,Soft-NMS 通过设置衰减函数的方法来避免直接将检测框得分置 0 的问题。当检测框与 M 重叠较大时,则分配较低的分;而重叠区域较小时,则保持原本的分。

传统检测算法只在检测网络训练后进行候选区域筛选。所提改进在网络训练前后均加入 Soft-NMS 算法,经过实验验证,这种方法能够进一步对错误或者重复的候选框进行筛选,同时有效地保留了正确框,进而提高整个算法的检测准确率。

3.4 FocalLoss 函数

使用 FocalLoss 函数作为模型分类损失函数来处理水下数据集正负样本数量不平衡的问题^[23]。FocalLoss 函数利用降低简单负样本权重进行训练的方式来使检测模块侧重于更难分类的样本,这也是一种困难样本挖掘的方式。FocalLoss 函数是基于交叉熵损失函数改进而来的。交叉熵损失函数的表达式为

$$L = \begin{cases} -\log p, & y = 1 \\ -\log(1 - p), & y \neq 1 \end{cases}, \quad (3)$$

式中: p 为预测样本属于 1 的概率; y 为标签, 取 $[-1, 1]$ 。FocalLoss 函数的表达式为

$$L_F = \begin{cases} -\alpha(1 - p)^\gamma \log p, & y = 1 \\ -\alpha p^\gamma \log(1 - p), & y \neq 1 \end{cases}, \quad (4)$$

式中: γ 为调节系数; α 为平衡因子。 $\gamma > 0$ 时, 可以减小易于分类样本的损失, 使模型注重于更加困难和复杂的样本。此外, α 作为平衡因子可以平衡数据集中正负样本分布不均衡的问题。通过实验证明, FocalLoss 函数作为分类损失函数有效提高了分类的准确性。

3.5 评价指标

采用平均检测精度 (AP) 和平均召回率 (AR) 作为主要参考指标来评价整个模型的检测性能。在 COCO 数据集给定的目标检测评价指标中, AP 是所有类别平均精度的均值, 即传统意义的平均准确度均值 (mAP)。AR 是所有类别召回率的均值。平均检测精度和平均召回率的计算公式分别为

$$P_A = \frac{N_{TP}}{N_{TP} + N_{FP}}, \quad (5)$$

$$R_A = \frac{N_{TP}}{N_{TP} + N_{FN}}, \quad (6)$$

式中: TP 指正样本被正确识别为正样本; TN 指负样本被正确识别为负样本; FP 指负样本被错误识别为正样本; FN 指正样本被错误识别为负样本。

4 实验分析

4.1 实验环境

实验基于 Intel®Core™ i7-8700 CPU@3.20 GHz × 12 处理器, NVIDIA RTX 2070super 显卡, 12 GB 内存, Ubuntu 16.04 64-bit 操作系统。检测模型在 PyTorch 1.0 框架上实现, 使用 Python 3.7 和 CUDA

10.1, 并利用 CUDNN 7.6.5 进行加速处理。实验训练学习率设置为 0.0025, batch-size 为 2, 共训练 60 个 epoch。

4.2 实验数据集

使用的数据集来自 2020 水下机器人目标抓取大赛 (URPC) 官方提供的水下图像数据集, 由于全部图片均在真实海洋环境下拍摄, 图片质量呈现各种形式、不同程度的衰减, 图片中包含海参、海胆、扇贝和海星 4 种目标, 处理成可用数据集, 有 7593 张图片。由于不同水下生物生活习惯存在较大差异, 大部分图片只存在单个目标, 当所有目标同时出现时, 海参和海胆的数量要远远大于海星和扇贝, 并且所有目标都在图像中呈现非均匀分布, 部分目标的相互遮挡和覆盖给模型的检测带来了更大的难度。数据集部分样本示例如图 6 所示。

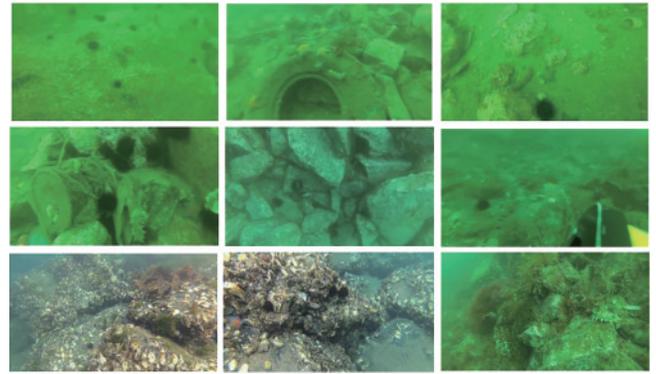


图 6 部分数据样本

Fig. 6 Partial data samples

为了适应所提训练模型, 使用 Python 语言将原始 xml 标注文件转换为 json 文件, 并将全部 7593 张图片以 7:3 的比例划分为训练集和测试集, 即随机选取 5315 张作为训练集样本, 2278 张作为测试集样本, 分别对所提算法进行训练和验证。图 7 给出了数据集的分析结果。图 7(a) 为真实框的宽高比, 可以看到大部分集中在 1.0 左右, 但也有部分在 0.5~1.0, 少部分集

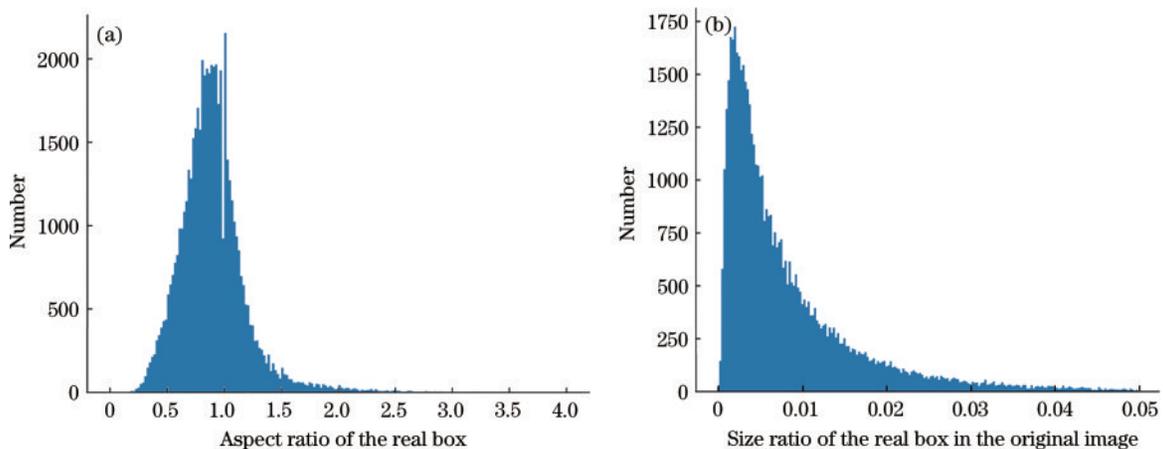


图 7 数据集分析。(a) 真实框的宽高比; (b) 真实框在原图的大小比例

Fig. 7 Dataset analysis. (a) Aspect ratio of the real box; (b) size ratio of the real box in the original image

中在 1.25~2.0。图 7(b) 为真实框在原图的大小比例, 可以看到大部分框只占到了原图的 0.01%, 即数据集中小目标占比较大。

4.3 图像增强效果分析

为了说明所提自动色阶图像增强算法在水下环境中的有效性, 对比了目前较为主流的图像增强算法, 包括多尺度颜色增强 (MSR) 算法^[24] 和具有色彩保护的多尺度颜色增强 (MSRCP) 算法^[25]。

图 8 给出了原始图像和三种增强算法增强图像之

后的结果。分析可知: MSR 算法可以对图像色彩进行整体增强, 但是会造成色调丢失, 最终导致增强后的图像纹理、色彩模糊不清; MSRCP 算法增强后的图像褪色明显, 其中青绿色光投射过强, 红通道退化严重, 造成图像失真; 所提自动色阶增强算法不仅提高了水下目标的可见性而且也使色彩的分布更加均衡, 既保留了水下环境的背景色彩又增加了水下目标的清晰度, 因此达到了较好的视觉感受, 为后续的目标检测提供了良好的基础。

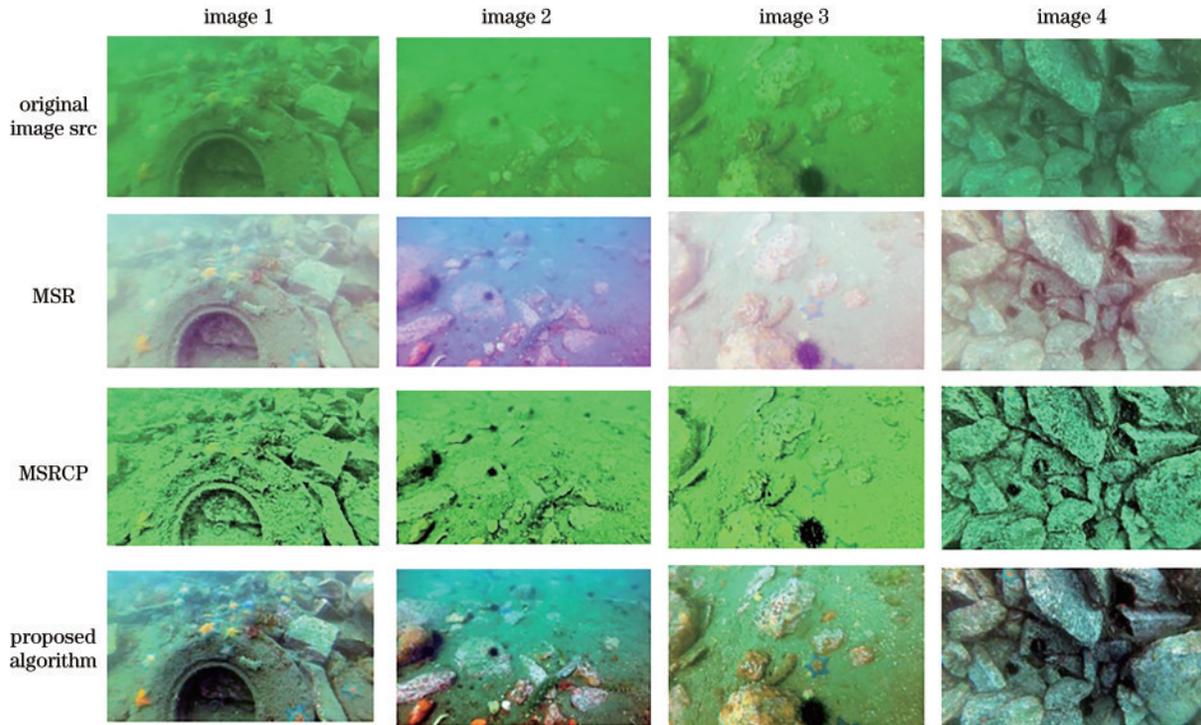


图 8 不同算法增强图像后的结果

Fig. 8 Results of image enhancement obtained by different algorithms

还选用 3 个无参考的图像质量评价指标对 3 种图像增强算法进行定量分析。1) 自然图像质量评价指标 (NIQE), 是一种质量感知统计特征集合。NIQE 不用计算图像的主观评价分数, 而是利用人眼对图像高对比度区域更加敏感的特点, 选取图像中的显著区域建立多元高斯特征模。NIQE 值越低, 越符合人眼的主观评价标准。2) 区域化整合自然图像质量评价指标 (ILNIQE), 首先提取图像的自然场景统计特征, 然后利用 model view controller (MVG) 来计算这种规律被破坏的程度, 从而估计出待测失真图。ILNIQE 不需要失真图像和主观评分就可以对图像进行评价, 泛化能力很好且准确率较高。3) 水下彩色图像质量评价 (UCIQE), 是色彩浓度、饱和度、对比度的线性组合, 用来定量评价水下图像非均匀色偏、模糊和低对比度的情况。UCIQE 值越大, 说明图像色彩越清晰, 图像质量越好。

表 1 给出了图片 1~4 的具体评价结果。从表 1 可以看出: 三种算法的 NIQE 和 UCIQE 均有一定提高;

MSR 算法的 ILNIQE 指标相对于原数据集有所下降, 图像质量降低; 所提自动色阶算法的 ILNIQE 最高; MSRCP 和所提自动色阶算法相差不大, 但是 MSRCP 算法存在色彩失真的问题。因此总体来看, 利用所提自动色阶算法增强后的图像色彩更加平衡, 图像质量更高。

综合考虑主观和客观评价, 所提自动色阶图像增强算法具有更好的增强效果, 无论是从图像质量还是从清晰度的角度来看, 都能够有效地增强水下目标, 提高图像的质量, 得到更加清晰、细节更突出、结构更加明显的图像。所提自动色阶图像增强算法不仅具有较好的评价指标, 而且能够对下一步检测算法提供更好的检测基础。

4.4 目标检测实验分析

为证明所提改进算法的有效性, 对其与改进之前的算法进行一系列对比, 具体方法如表 2 所示。

在表 2 中: Baseline 为基线算法, 在原始 Faster R-CNN 基础上替换骨干网络为 ResNet50, 并加入特征

表 1 图像增强结果评价

Table 1 Evaluation of image enhancement results

Image	Algorithm	NIQE	ILNIQE	UCIQE
1	Src	7.1606	38.2495	0.3061
	MSR	6.8799	39.6676	0.2722
	MSRCP	3.7418	27.1790	0.4390
	Proposed algorithm	6.3834	29.2617	0.2722
2	Src	12.9428	42.7309	0.2785
	MSR	6.8799	39.6676	0.3636
	MSRCP	8.0962	46.2339	0.4708
	Proposed algorithm	10.3499	32.5586	0.4609
3	Src	10.4168	33.5792	0.3585
	MSR	9.1082	46.0194	0.2029
	MSRCP	7.0049	30.098	0.4275
	Proposed algorithm	8.8039	27.8706	0.4276
4	Src	5.7975	31.2044	0.2955
	MSR	5.6905	38.2397	0.2566
	MSRCP	4.1026	34.4918	0.4275
	Proposed algorithm	3.8041	27.3969	0.4668

金字塔网络 (FPN), 引入 ROI align; Method 1、Method 2、Method 3 在 Baseline 基础上分别替换网络损失函数为 IoULoss、FocalLoss、Gaussian FocalLoss, 通过实验选择效果最好的一个函数作为本文损失函数; Our-1 在 Method 2 基础上加入图像增强部分; Our-2 在 Our-1 基础上替换训练前候选框筛选方法为 Soft-NMS; Our-3 在训练前后均采用 Soft-NMS 进行训练; Our-4 表示在 Our-3 基础上引入 PAFPN 进行训练。

为了检验所提改进目标检测模型的性能, 从客观评价结果和主观评价结果两个方面对模型的检测性能

进行评估。对于客观评价结果, 具体实验数据如表 3 所示。其中: AP 采用了 10 个交并比 (IoU) 阈值, 即 IoU 为 0.50:0.05:0.95, 对所有类别求平均值, 即传统意义的平均精确度; $AP^{0.50}$ 表示 IoU 为 0.50 时的检测精度; $AP^{0.75}$ 表示 IoU 为 0.75 时的检测精度; AP^s 表示小目标的检测精度, 即目标面积小于 $32 \times 32 = 1024$ 的目标; AP^m 表示中等目标的检测精度, 即目标面积大于 $32 \times 32 = 1024$ 、小于 $96 \times 96 = 9216$ 的目标; AP^l 表示大目标的检测精度, 即目标面积大于 $96 \times 96 = 9216$ 的目标; AR 是所有类别召回率的均值; AR^s 表示小目标的召回率; AR^m 表示中等目标的召回率; AR^l 表示大目标的召回率。

由表 3 可以看出: Baseline 模型的平均检测精度为 54.2%, 召回率为 62.1%, 随着对网络的改进, 检测精度和召回率相比改进之前都有一定程度的提升; 对比 Method 1、Method 2、Method 3 发现, 当使用 FocalLoss 作为模型损失函数时, 检测效果相对更好, 平均检测精度提高了 1.8 个百分点, 召回率提高了 1.3 个百分点; Our-1 引入了自动色阶增强算法对输入图像进行增强处理, 相比于未增强时, 检测精度提高了 0.7 个百分点, 召回率提高了 0.6 个百分点; 为了验证训练前和训练前后均采用 Soft-NMS 算法对模型准确率提升的有效性, Our-2 和 Our-3 分别只在训练前和在训练前后均采用 Soft-NMS 进行训练, 发现训练前后均采用 Soft-NMS 虽然召回率下降了 0.2 个百分点, 但检测精度提升了 1.2 个百分点; 最后, 为了进一步提高检测模型的准确率, Our-4 在特征加强阶段引入了 PAFPN, 虽然召回率与 Our-2 相同, 但相对于未使用 PAFPN 的算法, 检测精度提高了 0.6 个百分点。综上所述, 可以发现: 所提 Our-4 的准确率达 59.7%, 相比 Baseline 提高了 5.5 个百分点; 召回率达 70.5%, 相比 Baseline 提高了 8.4 个百分点; 其中, 小目标检测的准确率达 33.3%, 相比 Baseline 提高了 6.2 个百分点; 与普遍改进的 Faster R-CNN 相比, 准确率和召回率均有明显提

表 2 不同改进检测算法的对比

Table 2 Comparison of different improved detection algorithms

Algorithm	Input image	Backbone	Feature enhancement	ROI	Bbox	Loss function	Epoch
Faster R-CNN	Src	VGG16	None	ROI pooling	NMS(T+V)	L1Loss	None
Baseline	Src	ResNet50	FPN	ROI align	NMS(T+V)	L1Loss	60
Method 1	Src	ResNet50	FPN	ROI align	NMS(T+V)	IoULoss	60
Method 2	Src	ResNet50	FPN	ROI align	NMS(T+V)	FocalLoss	60
Method 3	Src	ResNet50	FPN	ROI align	NMS(T+V)	Gaussian FocalLoss	60
Our-1	IE	ResNet50	FPN	ROI align	NMS(T+V)	FocalLoss	60
Our-2	IE	ResNet50	FPN	ROI align	Soft-NMS(V)	FocalLoss	60
Our-3	IE	ResNet50	FPN	ROI align	Soft-NMS(T+V)	FocalLoss	60
Our-4	IE	ResNet50	PAFPN	ROI align	Soft-NMS(T+V)	FocalLoss	60

表 3 不同改进检测算法的 AP 和 AR
Table 3 AP and AR of different improved detection algorithms

Method	AP / %	AP ^{0.50} / %	AP ^{0.75} / %	AP ^S / %	AP ^M / %	AP ^L / %	AR / %	AR ^S / %	AR ^M / %	AR ^L / %
Baseline	54.2	88.8	59.5	27.1	48.2	60.1	62.1	40.6	57.2	67.0
Method 1	54.9	88.7	61.1	27.4	49.0	60.8	62.9	40.3	58.0	67.8
Method 2	56.0	90.2	62.3	28.5	50.0	61.9	63.4	41.1	58.3	68.3
Method 3	55.5	90.0	61.7	28.9	49.4	61.5	63.0	40.7	58.0	68.1
Our-1	56.7	90.7	63.1	29.0	50.5	62.6	64.0	41.4	58.8	69.1
Our-2	57.9	90.8	66.1	30.6	51.7	63.6	70.5	49.4	65.5	74.8
Our-3	59.1	90.8	67.9	32.4	52.7	65.1	70.3	49.4	65.0	75.0
Our-4	59.7	91.2	68.6	33.3	53.1	67.5	70.5	49.7	65.1	75.3

升。结果说明所提 Our-4 改进算法对水下目标的检测能力更强,效果更佳。

分别以召回率和准确率为横轴和纵轴,拟合得到 PR 曲线,曲线与两坐标轴围成的面积越大,则 AP 值越大,检测精度越高。图 9 给出了不同算法的 PR 曲线。所提算法的 PR 曲线下降速度最慢且位置最高,相比于改进之前的算法,在召回率相同时检测精度最高。相比改进之前的算法,所提算法的 PR 曲线与两坐标轴围成的面积最大,证明其对水下目标的检测性能更好。

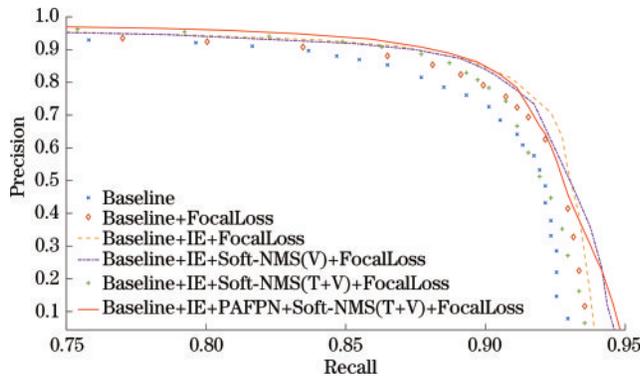


图 9 不同改进算法的 PR 曲线

Fig. 9 PR curves of different improved algorithms

所提网络的训练损失曲线如图 10 所示。由图 10 可以看到,随着检测网络的不断训练,网络总损失函数不断下降,在训练迭代 1500 次之后检测网络趋于稳定,检测模型最终达到收敛。

对于主观评价结果,分别用改进前后的算法对水下目标图像进行检测。图 11 给出了改进前的算法与所提算法的水下目标检测结果对比。

从图 11(a)、(b)可以看出,所提改进算法对小目标的检测精度有很大提升,能够更好地检测出小目标。由图 11(c)、(d)可看出:当目标密度较大、重合度较高时,改进前的算法的检测结果中存在较多无效

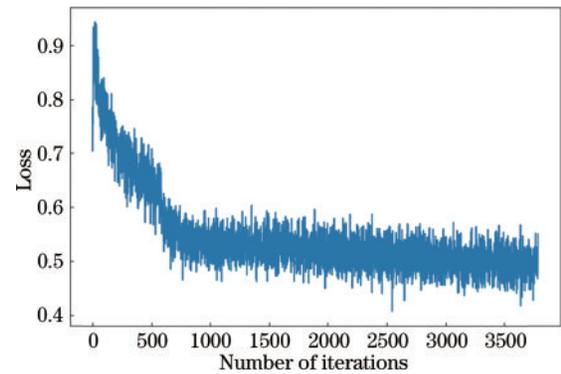


图 10 所提网络的损失函数曲线

Fig. 10 Loss function curve of the proposed network

框,无法正确地检测出全部候选框重叠的目标;而所提改进算法可以更加准确地去掉无效框,检测出检测框重叠较多的目标。由图 11(e)、(f)可以看出,所提改进算法相比于改进之前检测到的目标更多,即降低了整体的漏检率,提高了检测准确率。虽然所提改进算法仍存在少数漏检的情况,例如图 11(b)中的海胆和图 11(d)中的扇贝没有被检测出来,然而综合来看,所提改进算法相对于原算法取得了较优的性能,在环境光线暗、目标密度较高的情况下提高了整体的检测率。

在水下环境复杂、水质浑浊时,所提检测算法能够进一步提升目标的检测精度,图 12 给出了所提算法对模糊图像与清晰图像的检测效果对比。

从图 12 可以看出:在增强之后的图像中,所提改进算法对小目标的漏检率大大降低,对于模糊图像中的小目标,在增强后的图像中所提改进算法都能很好地检测出来;从图 12(e)可以看出,对于存在目标遮挡的图像,进行图像增强之后,所提改进算法的检测效果更好,能够更加准确地识别定位不同的遮挡目标。

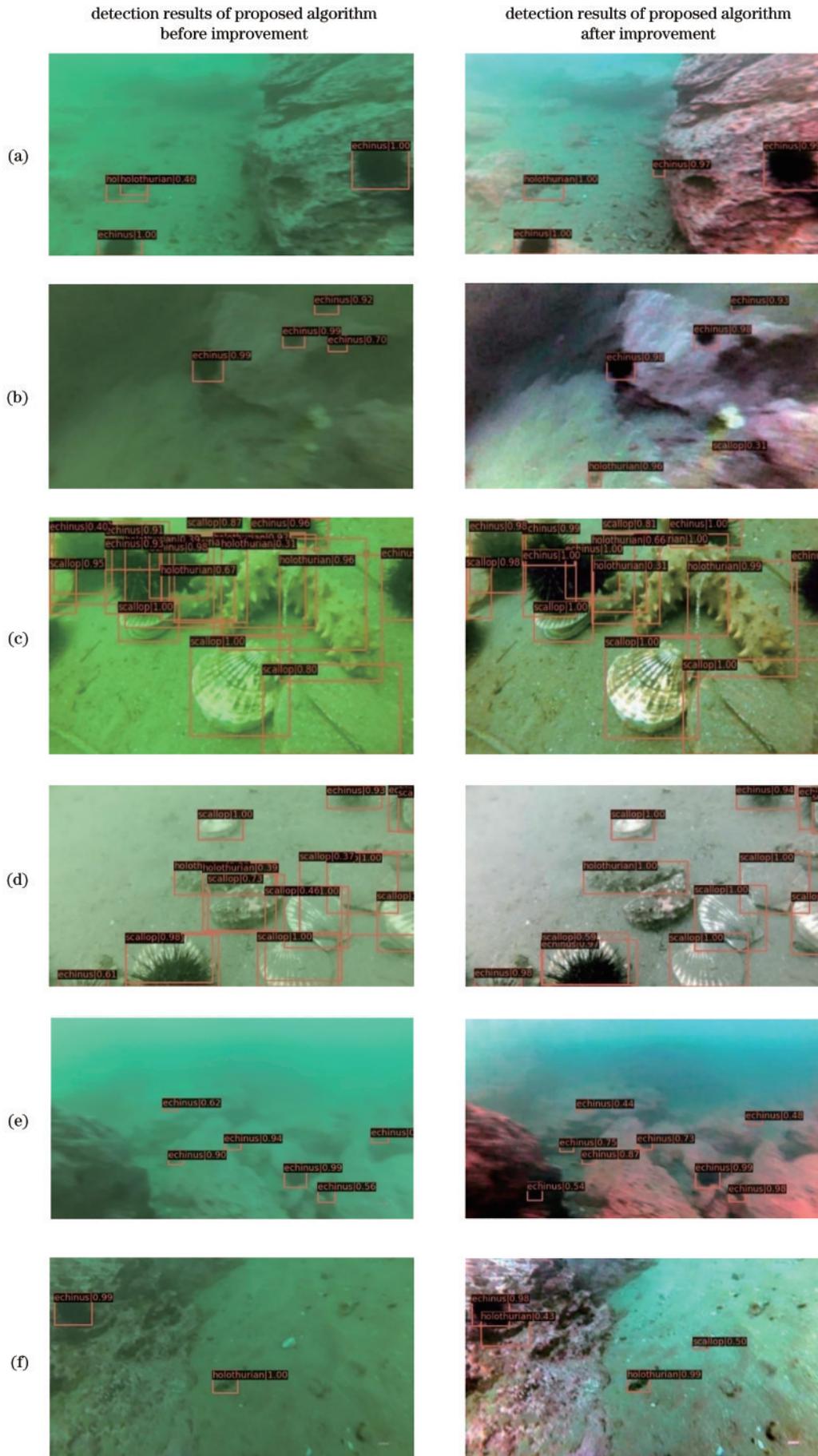


图 11 所提算法改进前后的检测结果对比

Fig. 11 Comparison of detection results of proposed algorithm before and after improvement

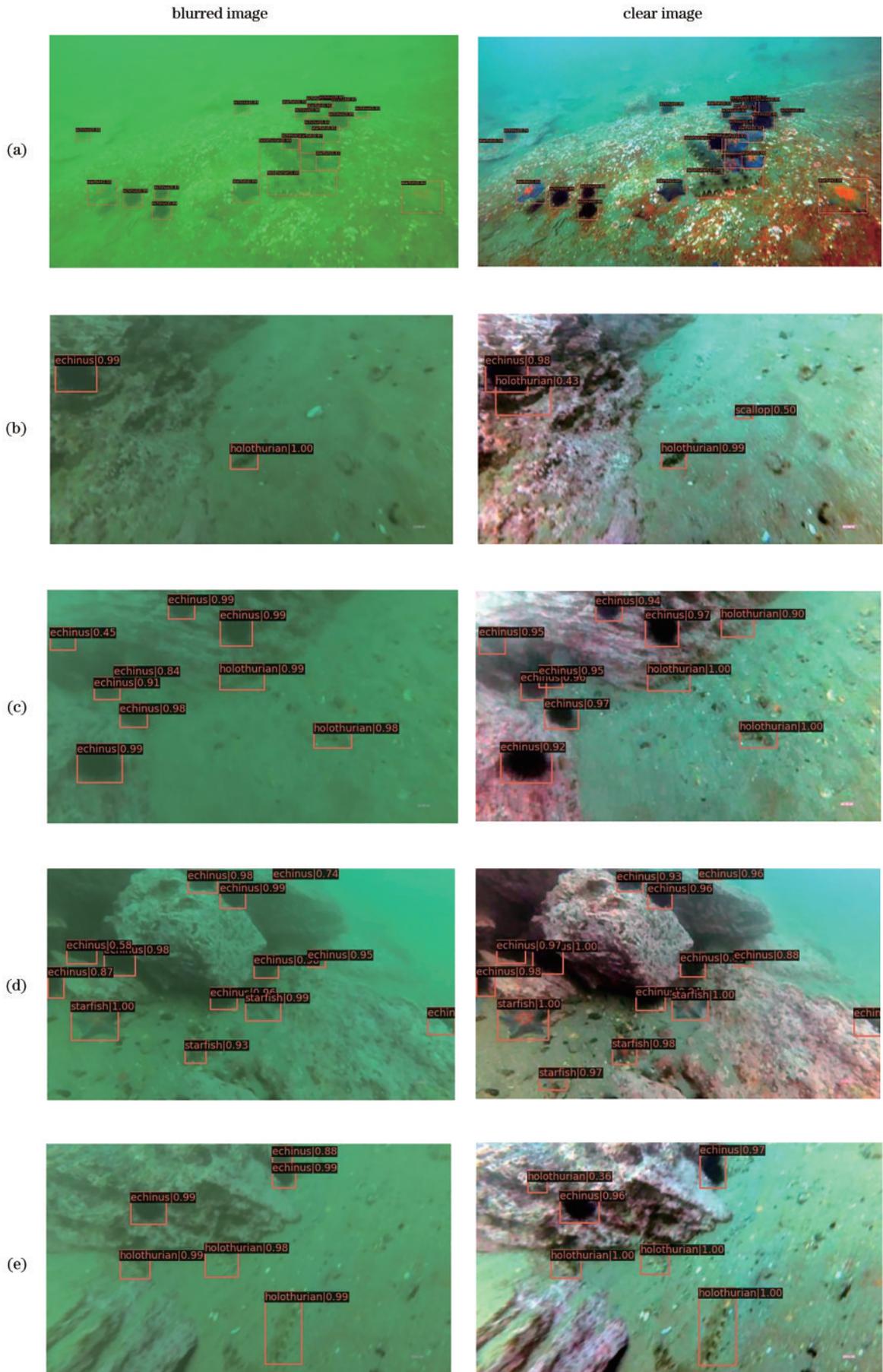


图 12 对模糊图像与清晰图像的检测效果对比

Fig. 12 Comparison of detection effects for blurred image and clear image

5 结 论

针对现有 Faster R-CNN 改进方法的不足,提出了一种自动色阶与双向特征融合的水下目标检测算法,可缓解水下目标模糊难以检测的问题,提高了对水下目标的检测精度,在海底生物探测、海产品捕捞等方面具有非常广阔的应用前景。但由于数据集分布不均匀,所提算法对贝壳类目标检测的精确度仍不是很高,可考虑数据增强方法扩大目标数量,或者进一步改进算法以提高检测精度。

参 考 文 献

- [1] Zhou Y, Yan K M. Domain adaptive adversarial learning based on physics model feedback for underwater image enhancement[EB/OL]. (2020-02-20)[2021-02-03]. <https://arxiv.org/abs/2002.09315>.
- [2] 刘有用, 张江梅, 王坤朋, 等. 不平衡数据集下的水下目标快速识别方法[J]. 计算机工程与应用, 2020, 56(17): 236-242.
Liu Y Y, Zhang J M, Wang K P, et al. Fast underwater target recognition with unbalanced data set[J]. Computer Engineering and Applications, 2020, 56(17): 236-242.
- [3] 袁红春, 张硕. 基于 Faster R-CNN 和图像增强的水下鱼类目标检测方法[J]. 大连海洋大学学报, 2020, 35(4): 612-619.
Yuan H C, Zhang S. Detection of underwater fish based on Faster R-CNN and image enhancement[J]. Journal of Dalian Ocean University, 2020, 35(4): 612-619.
- [4] 张志强, 牛智有, 赵思明. 基于机器视觉技术的淡水鱼品种识别[J]. 农业工程学报, 2011, 27(11): 388-392.
Zhang Z Q, Niu Z Y, Zhao S M. Identification of freshwater fish species based on computer vision[J]. Transactions of the Chinese Society of Agricultural Engineering, 2011, 27(11): 388-392.
- [5] 姚润璐, 桂詠雯, 黄秋桂. 基于机器视觉的淡水鱼品种识别[J]. 微型机与应用, 2017, 36(24): 37-39.
Yao R L, Gui Y W, Huang Q G. Recognition of freshwater fish species based on machine vision[J]. Microcomputer & Its Applications, 2017, 36(24): 37-39.
- [6] Rova A, Mori G, Dill L M. One fish, two fish, butterfly, trumpeter: recognizing fish in underwater video[C]//Proceedings of IAPR Conference on Machine Vision Applications, MVA 2007, May 16-18, 2007, Tokyo, Japan. [S.l.: s.n.], 2007: 404-407.
- [7] 朱世伟, 杭仁龙, 刘青山. 基于类加权 YOLO 网络的水下目标检测[J]. 南京师大学报(自然科学版), 2020, 43(1): 129-135.
Zhu S W, Hang R L, Liu Q S. Underwater object detection based on the class-weighted YOLO net[J]. Journal of Nanjing Normal University (Natural Science Edition), 2020, 43(1): 129-135.
- [8] 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 W, He Y Y, Guo Y J, et al. Exploring underwater target detection algorithm based on improved SSD[J]. Journal of Northwestern Polytechnical University, 2020, 38(4): 747-754.
- [9] 刘萍, 杨鸿波, 宋阳. 改进 YOLOv3 网络的海洋生物识别算法[J]. 计算机应用研究, 2020, 37(S1): 394-397.
Liu P, Yang H B, Song Y. Marine biometric recognition algorithm based on improved YOLOv3 network[J]. Application Research of Computers, 2020, 37(S1): 394-397.
- [10] Chen L, Liu Z H, Tong L, et al. Underwater object detection using Invert Multi-Class Adaboost with deep learning[C]//2020 International Joint Conference on Neural Networks (IJCNN), July 19-24, 2020, Glasgow, UK. New York: IEEE Press, 2020.
- [11] 宋绍剑, 朱靖旭. 基于 Mask R-CNN 和迁移学习的水下生物目标识别研究[J]. 计算机应用研究, 2020, 37(S2): 386-388, 391.
Song S J, Zhu J X. Object recognition research of underwater creature based on Mask R-CNN and transfer learning[J]. Application Research of Computers, 2020, 37(S2): 386-388, 391.
- [12] Jobson D J, Rahman Z, Woodell G A. 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.
- [13] 徐凤强, 董鹏, 王辉兵, 等. 基于水下机器人的海产品智能检测与自主抓取系统[J]. 北京航空航天大学学报, 2019, 45(12): 2393-2402.
Xu F Q, Dong P, Wang H B, et al. Intelligent detection and autonomous capture system of seafood based on underwater robot[J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(12): 2393-2402.
- [14] Ren S Q, He K M, 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, 2017, 39(6): 1137-1149.
- [15] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 770-778.
- [16] Fan Y C, Zhang L X, Guo H C, et al. Image processing for laser imaging using adaptive homomorphic filtering and total variation[J]. Photonics, 2020, 7(2): 30.
- [17] Theckedath D, Sedamkar R R. Detecting affect states using VGG16, ResNet50 and SE-ResNet50 networks[J]. SN Computer Science, 2020, 1(2): 79.
- [18] Neubeck A, van Gool L. Efficient non-maximum suppression[C]//18th International Conference on Pattern Recognition (ICPR'06), August 20-24, 2006, Hong Kong, China. New York: IEEE Press, 2006: 850-855.
- [19] 刘荻, 张焱, 赵琰, 等. 基于特征重聚焦网络的多尺度近岸舰船检测[J]. 光学学报, 2021, 41(22): 2215001.
Liu D, Zhang Y, Zhao Y, et al. Multi-scale inshore ship detection based on feature re-focusing network[J]. Acta Optica Sinica, 2021, 41(22): 2215001.
- [20] 庄秀玲, 谭福奎, 李震, 等. 基于暗通道先验和优化自动色阶的图像去雾算法[J]. 计算机应用与软件, 2021,

- 38(7): 190-195.
- Zhuang X L, Tan F K, Li Z, et al. Image defogging algorithm based on dark channel prior and optimized auto-color[J]. *Computer Applications and Software*, 2021, 38(7): 190-195.
- [21] Liu S, Qi L, Qin H F, et al. Path aggregation network for instance segmentation[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 8759-8768.
- [22] 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.
- [23] 鞠默然, 罗江宁, 王仲博, 等. 融合注意力机制的多尺度目标检测算法[J]. *光学学报*, 2020, 40(13): 1315002.
- Ju M R, Luo J N, Wang Z B, et al. Multi-scale target detection algorithm based on attention mechanism[J]. *Acta Optica Sinica*, 2020, 40(13): 1315002.
- [24] Xue C H, Lu S J, Zhang W. MSR: multi-scale shape regression for scene text detection[C]//Proceedings of the 28th International Joint Conference on Artificial Intelligence, August 10-16, 2019, Macao, China. New York: ACM, 2019: 989-995.
- [25] 杨文静, 陈明, 冯国富. 基于图像增强的水下视频鱼类识别方法[J]. *激光与光电子学进展*, 2021, 58(22): 2211002.
- Yang W J, Chen M, Feng G F. Fish recognition method for underwater video based on image enhancement[J]. *Laser & Optoelectronics Progress*, 2021, 58(22): 2211002.