伍济钢,曹鸿.基于改进 YOLOv7-tiny 的卸船机抓斗检测方法[J].湖南科技大学学报(自然科学版),2024,39(4):77-86.doi: 10.13582/j.cnki.1672-9102.2024.04.010

WU J G, CAO H. A Detection Method for Ship Unloader Grab Based on Improved YOLOv7-tiny[J]. Journal of Hunan University of Science and Technology (Natural Science Edition), 2024, 39(4): 77-86.doi:10.13582/j.cnki.1672-9102.2024.04.010

基于改进 YOLOv7-tiny 的卸船机抓斗检测方法

伍济钢*,曹鸿

(湖南科技大学 机电工程学院,湖南 湘潭 411201)

摘 要:针对港口码头复杂环境下,现有卸船机抓斗检测方法存在精度低、漏检率高和误检率高的问题,提出基于改进 YOLOv7-tiny 的卸船机抓斗检测方法.首先,为克服低质量样本示例影响模型性能的问题,引入 WIoU 损失函数增强训练效 果,加快模型训练的收敛速度;其次,为解决抓斗图像光照差异变化大,模型存在特征提取困难、参数冗余等问题,引入 GSConv 卷积对 ELAN 高效聚合路径模块进行改进,提高特征提取能力,保证模型参数大小和检测性能的平衡;最后,针对抓 斗图像尺度变化大、背景干扰等导致检测精度低的问题,引入 EMA 高效多尺度注意力机制,提高不同尺度的特征学习能 力,突出重要特征信息.所提出的改进算法在卸船机抓斗数据集中的检测准确率为 97.0%,较改进前提高 5.1%,模型大小减 少 16.3%;与 Faster R-CNN, YOLOv5s, YOLOv7 等方法开展对比验证试验,所提方法的 mAP 值分别提高 11.2,9.3 和 3.5 个 百分点.结果表明:改进算法具有精度高、鲁棒性强等优点.

关键词:抓斗检测;YOLOv7-tiny;损失函数;GSConv卷积;注意力机制 中图分类号:TP391.4 文献标志码:A 文章编号:1672-9102(2024)04-0077-10

A Detection Method for Ship Unloader Grab Based on Improved YOLOv7-tiny

WU Jigang, CAO Hong

(School of Mechanical and Electrical Engineering, Hunan University of Science and Technology, Xiangtan 411201, China)

Abstract: In order to address the existing ship unloader grab detection methods in the complex environment of port terminals, which suffer from low accuracy, high leakage rates, and high false detection rates, we propose a detection and recognition method based on an improved YOLOv7-tiny. Firstly, to mitigate the impact of low-quality sample examples on model performance, we introduce the WIoU loss function to enhance training effects and accelerate model training convergence. Secondly, to address issues related to difficult feature extraction and parameter redundancy due to significant variations in lighting conditions in grab images, we incorporate GSConv into the ELAN efficient aggregation path module, so as to improve feature extraction and fusion capabilities. Finally, to tackle the problem of low recognition accuracy caused by large scale changes in grab images and background interference, we introduce the EMA mechanism to enhance feature learning at different scales and highlight important feature information. The recognition accuracy of the improved algorithm on the ship unloader grab dataset is 97.0%, representing a 5.1% improvement over the previous version, with a 16.3% reduction in model size. Comparative validation experiments show that the above method outperforms Faster R-CNN,

收稿日期:2024-04-19

基金项目:湖南省自然科学省市联合基金重点项目资助(2022JJ50129)

^{*} 通信作者, E-mail: jgwu@ hnust.edu.cn

YOLOv5s, and YOLOv7 by 11.2%, 9.3%, and 3.5%, respectively. Experimental results demonstrate that the improved algorithm offers high accuracy and robustness.

Keywords: grab detection; YOLOv7-tiny; loss function; GSConv; attention mechanism

干散货占全球海运物流的 70%以上,预计到 2027 年,干散货航运将达到 68 亿 t 的市场容量,并在预 测期内以 5.10% 的复合年增长率增长^[1].由于海运物流是干散货最便宜的运输方式,干散货港口的发展势 头迅猛.与成本高昂的港口扩建相比,提高港口运营效率以最大限度地提高港口吞吐量更具成本效益^[2]. 在干散货港口资源有限的情况下,目前码头的生产条件和技术水平仍然相对落后,如何利用先进的技术来 提高作业效率和作业质量是目前港口码头生产发展关注的核心问题^[3].抓斗卸船机是干散货码头卸船作 业中极为重要的一种机械设备,在装卸作业中起着举足轻重的作用.驾驶员可以利用卸船机的抓斗升降机 构,使控制抓斗在船舱中的位置,以便于抓取物料;抓斗关闭和打开机构使得抓斗能够紧握物料并保持稳 定;在物料被成功抓取后,抓斗会被提升到船舱上方的位置,将物料卸载到料斗上方;料斗可以暂存物料, 经由振动给料机后,物料被输送至码头上的皮带输送机;物料通过皮带输送机被运输到料堆,完成整个卸 船过程^[4].目前,卸船机的操作方式主要分为手动和半自动 2 种.这 2 种操作方式各有特点,但它们都存在 一个共同的限制:对驾驶员的依赖,这意味着完全自动化的卸料方式尚未得到广泛应用^[5].在卸船过程中, 抓斗与船体的碰撞、抓斗错误地落在无货区域以及抓斗与货舱内壁碰撞等都是最常见的安全事故^[6].这些 碰撞事故不仅可能导致船体凹陷,增加维修成本和时间,而且还可能直接导致抓斗本身的损坏,降低作业 效率和安全性.更为严重的是,如果抓斗操作不当,碰撞可能会导致人员伤亡的严重事故^[7].

随着深度学习的不断成熟,到目前为止,由于卷积神经网络(Convolutional Neural Networks, CNN)具备的优异性能^[8],基于深度学习的目标检测方法在港口码头等复杂环境下的智能监控建设取得了突破性成果.基于深度学习的目标检测算法主要分为单阶段和两阶段算法^[9],YOLO(You Only Look Once)^[10]系列算法凭借其较高的检测精度和检测速度在多种目标检测中得到广泛应用.基于深度学习的目标检测在卸船机抓斗中的应用较少,相关研究如张文明等^[11]通过改进深度卷积神经网络 YOLOv3-tiny 对门机抓斗进行检测;CUI 等^[12]提出一种基于卷积神经网络的用于船舶检测的深层特征感知框架;张绍文等^[13]提出一种加权感受野和跨层融合的遥感船舶小目标检测方法,可以实现小目标的多尺度定位检测;李佳东等^[14]提出一种轻量级船舶目标检测算法 YOLO-Ship,可以实现高速、高精度的船舶检测;ZHOU 等^[15]在 YOLOv5 网络结构中引入协同注意力机制和加权金字塔网络,提高船舶的定位和识别精度;罗芳等^[16]在 YOLOv5 网络结构中引入特征增强模块,利用自适应特征融合方法实现高性能检测.在 YOLO 系列算法中,YOLOv7^[17]在相同体量下比同系列中其他算法的性能更加优越,YOLOv7-tiny 是 YOLOv7 算法的一个轻量级模型,参数量更小,检测速度更快,能很好满足抓斗目标实时检测的要求,它采用独特的 ELAN 高效层聚合路径网络模块,可以保证高效的检测精度和速度.但是,港口码头的环境复杂多样,抓斗运动会导致目标图像的尺度差异较大,同时抓斗作业可能会造成遮挡,这些因素加上复杂环境和目标的自身特性,使得目标图像的特征不明显^[18],抓斗检测过程仍存在漏检率和误检率高等问题,影响抓斗目标的检测精度.

通过分析存在的问题,本文提出一种基于 YOLOv7-tiny 的卸船机抓斗检测方法.为了解决数据集普遍存在一定数量的低质量示例影响模型性能的问题,引入加权交并比(Weighted Intersection over Union, WIoU)^[19]损失函数来增强数据集的训练效果,加快模型训练的收敛速度.针对 ELAN 模块堆叠卷积参数冗余,特征融合能力差,在光照变化大等情况下特征提取困难等问题,引入 GSConv 卷积^[20]进行深层特征的提取,尽可能地避免传输过程中的空间信息丢失,加快推理速度并提高模型精度.在特征融合网络中,通过在特征提取网络中引入 EMA^[21]注意力机制,抑制背景干扰,高效融合深浅层特征,提高模型的检测精度.

1 YOLOv7-tiny 算法

YOLOv7-tiny 是 YOLOv7 算法中的一个轻量级模型,其参数量更小,速度更快,它的网络结构由特征提取网络 backbone 以及进行特征融合和实现检测的检测头 Head 组成,特征提取网络由 CBL 卷积块和

ELAN 高效层聚合网络组成,特征融合网络沿用 YOLOv5 系列路径聚合特征金字塔(PAFPN)网络架构,将 特征金字塔网络(FPN)^[22]从高层传递的强语义信息与路径聚合网络(PANet)^[23]自底向上传递的强定位 信息张量拼接,充分融合不同层次特征信息,实现多尺度学习.YOLOv7-tiny 算法的网络结构如图 1 所示.



图 1 YOLOv7-tiny 算法的网络结构

标准的 YOLO 损失函数包括定位损失、置信度损失和分类损失.对于抓斗检测等特定任务,尤其是精细分类任务,定位损失、置信度损失和分类损失的权重设置不合理会导致模型在某一方面的优化不足.针对该问题,可以通过重新调整定位、置信度和分类损失的权重,确保模型在各个方面都能得到充分的优化. YOLOv7-tiny 的 ELAN 结构虽然轻量高效,但其容量可能不足以处理抓斗复杂的姿态特征,导致特征提取 不充分.ELAN 结构的特征融合方式可能无法充分捕捉抓斗姿态的细节,导致检测效果不佳,因此,可以适 当增加 ELAN 结构的层数或对其内部卷积模块进行改进,增强模型的特征提取能力.

本文引入 WIoU 损失函数对原有标准的 YOLO 损失函数 CIoU 进行改进,引入 GSConv 卷积对 ELAN 结构进行改进以提高特征提取能力,引入 EMA 注意力机制提高对目标细节特征的关注,忽略复杂背景的干扰.改进后算法的整体网络结构如图 2 所示.



图2 最终改进算法的整体网络结构

2 改进基于 YOLOv7-tiny 的抓斗检测算法

2.1 引入 WIoU 损失函数

边界框损失函数作为目标检测中的一个至关重要的部分,对于提升目标检测模型的性能起到决定性的作用.现有的研究通常基于一个前提,即训练数据集中的样本质量较高.然而,在数据集的创建过程中,

并没有一个明确的、统一的质量评估标准,这一问题在自制数据集中尤为突出,因为这些数据集可能包含 大量的低质量样本.YOLOv7-tiny的基准模型采用 CloU 作为边界框回归损失函数,而 CloU 针对多尺度目 标对象存在着泛化性较弱和收敛较慢的问题.对于抓斗目标形状相对复杂的情况,CloU 损失函数使用简 单的边界框不能很好地实现匹配,会导致目标检测性能的下降.加权交并比损失函数 WloU 引入动态非单 调聚焦机制得到具有两层注意力机制的 WloU,其构建距离注意机制为

$$L_{\text{WIoU}} = R_{\text{WIoU}} L_{\text{IoU}};$$

$$R_{\text{WIoU}} = \exp\left[\frac{(x - x_{\text{gt}})^2 + (y - y_{\text{gt}})^2}{(W_{\text{g}}^2 + H_{\text{g}}^2)^*}\right].$$

式中: L_{WIoU} 为 WIoU 损失函数; R_{WIoU} 为产生阻碍收敛的梯度; L_{IoU} 为 IoU 损失函数; x, y为预测框的中心 点在 x, y 轴上的坐标值; x_{gt}, y_{gt} 为真实框在 x, y 轴上的坐标值; W_{g}, H_{g} 为被标注目标在图像中占据的宽 度和高度.

WIoU使用异常值来区分锚框的质量,为异常值较小的高质量锚框和异常值较大的低质量锚框分配小梯度增益,同时削弱低质量示例对锚框回归的不利影响.该方法侧重于中等质量的锚框,并能有效地提高检测器的整体性能.最小的封闭框如图3所示.

2.2 引入 GSConv 卷积改进 ELAN 模块

YOLOv7-tiny 算法主要通过标准卷积的堆叠进行特征 提取,大量的标准卷积导致参数量的激增和特征的冗余, 而且随着层数的不断增加,影响也随之增加,这对于抓斗 的实时检测具有很大的影响.ELAN 高效层聚合网络通过 控制最短和最长梯度路径来优化算法的特征提取能力,当



模型深度达到一定程度时,如果继续堆叠卷积块,将使精度增益越来越小,甚至会导致总体精度比浅层网络差,所以需要在控制最优梯度路径的同时,又要满足抓斗目标对检测速度和精度的要求.因此,本文采用高效轻量级的 GSConv 卷积对 ELAN 模块中的卷积进行改进,GSConv 的结构如图 4 所示.



图 4 GSConv 结构

GSConv 卷积克服了深度卷积计算过程中通道信息分离的缺点,并能有效利用深度卷积计算量低的优势,由深度卷积和标准卷积拼接获得特征图后,采用 Shuffle 混洗对特征图进行完全融合,丰富特征提取信息.GSConv 卷积在保持算法精度没有衰减的前提下,减少计算参数以实现轻量化,加快模型算法的推理速度.

GSConv 对应的时间复杂度为

$$T_{\text{GSConv}} = O\left[WHK_1K_2\frac{C_2}{2}(C_1+1)\right]$$

式中: W 为输出特征图的宽度; H 为输出特征图的高度; K_1 , K_2 为卷积核大小; C_1 为每一个卷积核的通道

数,同时也等于输入特征图的通道数; C₂ 为输出特征图的通道数.

通过引入 GSConv 卷积对 ELAN 模块中的深层 3×3 卷积核的普通卷积进行改进,以此来提高 ELAN 的特征提取能力和通道信息交互能力.ELAN 改进如图 5 所示.



图 5 ELAN 改进

2.3 融合 EMA 注意力机制

当 CNN 网络扩展到多个卷积层时,表现 出显著增强学习特征表示的能力,这样会导 致堆栈更多的卷积,存在消耗大量内存和计 算资源的缺点,通过添加注意力机制来替代 卷积层的堆叠,可以很好地解决这一缺点,同 时有利于学习更具有判别性的特征.

本文通过在特征提取层中引入 EMA 高 效多尺度注意力模块,增强模型对多尺度抓 斗特征的学习能力,更加关注抓斗个体的重 要信息,忽略复杂背景等非重要信息,实现提 高模型性能的目的.常用的通道或空间注意 力机制通过通道降维操作建模跨通道关系来 增强特征信息,这可能会对提取深层特征带 来负面影响,EMA 注意力模块将部分通道重 构为批处理,并将通道维度分组为多个子特 征,使空间语义特征在每个特征组内均匀分 布,实现保留每个通道上的特征信息和减少 计算开销的目的.EMA 模块结构如图 6 所示, 其中 C 为通道数, C 为组数.



3 试验及结果分析

3.1 试验环境及数据集

本文试验平台为戴尔工作站 Precision 5820,操作系统:Ubuntu18.04 LTS;CPU:Intel Xeon(R) W-2255; 内存:32 GB;显卡:Quadro RTX4000×2;显存:16 GB×2;深度学习框架:PyTorch.

试验数据为通过高性能图像采集平台对天津某港口的抓斗式卸船机进行采集制作得到的抓斗数据 集,该数据集用于训练和评估面向抓斗对象的检测算法.数据集由9000张图像组成,包括3种常见的抓斗 姿态类型,分别为抓斗空载闭合、抓斗张开和抓斗满载,抓斗作业姿态如图7所示.通过对抓斗数据集进行 分析可知:抓斗作业过程中,室外作业环境、光照条件等因素的不同会改变抓斗图像的亮度和对比度,影响 深度学习模型提取目标特征:抓斗形状变化和位移范围增大等会导致抓斗图像目标尺度的变化增大:港口 有各种建筑物、设备、交通工具和人员,这些会成为复杂的背景干扰因素.



(a) 抓斗空载闭合

图7 抓斗作业姿态

(c) 抓斗满载

在 YOLOv7-tiny 的训练过程中,设置输入图片的像素为 640×640,动量大小参数为 0.937,初始学习速率 为 0.01,训练批次大小为 16,迭代次数为 100.数据集按照 8:1:1 的比例划分为训练集、验证集和测试集.

3.2 试验评价指标

为充分评估各改进检测算法对抓斗姿态的识别效果,本文采用针对单一类别的平均精度(Average Precision, AP)和针对所有类别的平均精度均值(Mean Average Precision, mAP)2个指标作为主要的评估 标准,分别反映模型对单一类别和整体类别的检测能力和准确性.通过精确率(Precision)和召回率 (Recall)之间的关系来绘制 PR 曲线,在这个曲线上,精确率为纵坐标,召回率为横坐标.平均精度、平均精 度均值、精确率和召回率的计算公式为

$$AP = \int_{0}^{1} P(R) dR \times 100\%;$$

mAP = $\frac{\sum_{i=1}^{n} AP_{i}}{n} \times 100\%;$
Precision = $\frac{TP}{TP + FP} \times 100\%;$
Recall = $\frac{TP}{TP + FN} \times 100\%.$

式中:P为精确率;R为召回率;n为总实例数量;TP为模型正确检测到的正类(目标)实例数量;FP为模 型错误地将负类(非目标)预测为正类的实例数量;FN 为模型未能识别出的正类实例数量,即漏检数量.

3.3 试验结果及分析

本文基于 YOLOv7-tiny 算法网络进行改进 WIoU 损失函数、引用 GSConv 卷积改进 ELAN 模 块以及融合 EMA 注意力的改进,为验证改进算 法对抓斗的检测效果,采用控制变量法进行消 融实验,并对各部分改进进行分析.首先,设置试 验一:验证改进 WIoU 损失函数对抓斗数据集的 检测效果,改进得到 tiny-WIoU 算法,选择 CIoU, EloU, SloU 进行对比试验, 以本文构建的抓斗 数据集为训练对象,训练结果得到的 Loss 曲线 如图 8 所示.改进前后网络模型的性能对比结果 如表1所示.



	•	,		
网络带刑		AD@05/0		
网络侠堂 -	抓斗张开	抓斗空载闭合	抓斗满载	mAP@0.3/%
YOLOv7-tiny	97.0	92.0	86.8	91.9
tiny-WIoU	97.0	92.8	89.8	93.2

表1 tiny-WIoU 网络模型检测性能对比

由表1可知:抓斗满载的检测精度相比于改进前提高3个百分点,整体平均检测精度提高1.3个百分点,这也证明前文构建的图像数据集存在的低质量示例对训练产生的负面影响,而通过引入WloU损失函数可以更加精确地权衡不同大小和形状的抓斗的边界框.这种改进使得网络更倾向于精确地定位和检测抓斗的具体位置,尤其是在抓斗形状多样或部分遮挡的情况下.由于WloU采用动态非单调调频机制,可以有效屏蔽训练过程中的许多负面影响,通过使用WloU损失函数训练的模型大大提高了抓斗满载姿态类别的检测精度.

为验证引入 GSConv 卷积改进 ELAN 模块的 YOLOV7-tiny 算法的性能,设置试验二:应用 GSConv 卷积 对原有算法中的 ELAN 模块进行改进,得到 tiny-GSConv 算法.将得到的新算法在数据集上进行训练,改进 前后网络模型的性能对比结果如表 2 所示.改进后得到的 mAP 值为 92.4%,相对于原始算法,抓斗张开和 抓斗空载闭合的精度有所下降,初始模型中检测精度最差的抓斗满载闭合姿态种类得到提高,证明 GSConv 模块中的 Shuffle 混洗操作能够通过将通道信息均匀打乱,增强提取到的语义信息,加强特征信息 的融合,提高图像特征的表达能力.结合前文改进 WIoU 算法得到 tiny-WIoU-GSConv 算法,其中对 YOLOv7-tiny 和 tiny-WIoU-GSConv 得到的 Precision 和 Recall 曲线进行对比,结果如图 9 所示.由图 9 可知: 改进后曲线的波动越来越小,进一步说明模型训练的效果得到提升.

网络齿刑		AP/%		描刊十小/MD	
网络侠型 -	抓斗张开	抓斗空载闭合	抓斗满载	mAP@0.3/%	候型入小VMD
YOLOv7-tiny	97.0	92.0	86.8	91.9	12.3
tiny-GSConv	96.6	90.1	90.4	92.4	10.3
Precisio 1.0 0.8 0.6 0.4 0.2 0 0 50	on 1.0 0.8 件 0.6 学 0.4 0.2 0	Recall	Preci 1.0 0.8 0.6 一 0.4 0.2 0 0 50	sion 1.0 0.8 登 0.6 使 0.4 0.2 0 100	Recall

表 2 tiny-GSConv 网络模型检测性能对比



迭代次数

迭代次数

(a) YOLOv7-tiny

迭代次数

迭代次数

(b) tiny-WIoU-GSConv

在试验二得到的 tiny-WIoU-GSConv 算法的基础上,设置试验三:验证融合 EMA 注意力对模型性能的 影响,在 backbone 特征提取网络层中引入 EMA 注意力机制,得到 tiny-EMA 算法,输入数据集训练得到的 结果如表 3 所示.由表 3 可知:tiny-EMA 算法的 mAP 值为 93.9%,其中抓斗张开时的 AP 值为 97.1%,抓斗 空载闭合时的 AP 值为 93.0%,抓斗满载时的 AP 值为 91.7%,整体精度有较大提高.热力图特征对比效果 如图 10 所示.由图 10 可知:原始模型对于抓斗目标特征的提取能力存在一定的局限性,EMA 注意力机制 使得模型能够更加专注于抓斗的细节特征,忽略那些无关紧要的背景信息.这说明在算法中加入 EMA 注 意力机制能增强模型对多尺度目标特征的学习能力.结合前文改进 WIoU 损失函数和 GSConv 卷积最终得 到 tiny-WIoU-GSConv-EMA 算法,最终改进后训练得到的结果如图 11 所示.

表3 tiny-EMA 网络模型检测性能对比

网络模型 —		AP/%			
	抓斗张开	抓斗空载闭合	抓斗满载	mAP@0.3/%	mar @ 0.3 · 0.9/ %
YOLOv7-tiny	97.0	92.0	86.8	91.9	83.9
tiny-EMA	97.1	93.0	91.7	93.9	86.9



图 10 热力图特征对比



图 11 tiny-WIoU-GSConv-EMA 算法网络的 PR 曲线

将试验一~试验三得到的结果与 Faster R-CNN, YOLOv5s, YOLOv7 算法在抓斗数据集中的检测性能进行对比,结果如表 4 所示.

由表4可知:与两阶段目标检测算法 Faster R-CNN 相比,本文改进算法 tiny-WIoU-GSConv-EMA 的 mAP 值提高 11.2 个百分点;对于 YOLO 系列先进的 v5s 和 v7 版本,本文最终改进得到的 tiny-WIoU-GSConv-EMA 算法的 mAP 值分别提高 9.3 个百分点和 3.5 个百分点;与本文采用的基准算法 YOLOv7-tiny 相比,tiny-WIoU-GSConv-EMA 的 mAP 值提高 5.1 个百分点,模型大小减少 16.3 个百分点.以上结果表明:本文提出的方法可以更好地满足智能抓斗现场监控的应用.从表4 中也可以得到:WIoU 结合 GSConv 或者 EMA 模块后的改进结果证明损失函数的改进能给整体模型带来很好的性能提升.针对原始模型 ELAN 模 块参数冗余、特征融合提取能力差的问题,引入的 GSConv 高效卷积对整体网络的特征提取能力有大幅提

升,不仅提升了网络模型的整体检测性能,而且也在一定程度上减少了参数量和模型大小,方便实际工业 应用的部署.融合 EMA 注意力机制后,网络模型的整体检测性能大幅度提升,证明在抓斗目标检测中,复 杂背景干扰和尺度变化大导致特征不明显等因素严重影响抓斗姿态检测的精度.注意力机制很好地忽略 复杂背景的干扰,可以更好地对抓斗特征进行提取与区分.将本文改进得到的最终算法结果应用于抓斗检 测,检测出3种抓斗姿态:抓斗空载闭合、抓斗张开、抓斗满载,检测效果如图12所示.由图12可知:本文 提出的方法可以准确地检测抓斗,且具有更好的鲁棒性.

算法	mAP@ 0.5/%	mAP@ 0.5 : 0.9/%	Precision/%	Recall/%
Faster R-CNN	85.8	76.6	88.5	78.3
YOLOv5s	87.7	79.2	89.2	80.6
YOLOv7	93.5	86.2	96.8	86.0
YOLOv7-tiny	91.9	83.9	93.7	81.5
tiny-WIoU	93.2	85.4	95.6	85.3
tiny-GSConv	92.4	84.4	95.4	83.4
tiny-EMA	93.9	86.9	92.6	86.9
tiny-WIoU-GSConv	93.8	86.6	94.1	87.7
tiny-GSConv-EMA	94.9	88.0	94.3	88.9
tiny-WIoU-EMA	96.1	90.5	95.2	92.2
tiny-WIoU-GSConv-EMA	97.0	91.1	93.7	90.7

表4 各类改进算法结果对比



(a) 抓斗空载闭合

图 12 tiny-WIoU-GSConv-EMA 算法的检测效果

(c) 抓斗满载

结论 4

1)改进 YOLOv7-tiny 网络结构可以提升抓斗图像在复杂背景、光照变化和尺度变化等情况下的检测 精度和特征提取能力.改进后的模型在检测精度和轻量化方面表现出较好的普适性,满足实际抓斗检测场 景的实时监控需求.

2) 提出的改进算法 tiny-WIoU-GSConv-EMA 在精度和模型轻量化方面优于 Faster R-CNN 和 YOLO 系 列(如 YOLOv5, YOLOv7).相比现有方法,改进后的算法在多尺度目标检测、复杂背景干扰和特征提取能 力方面更具优势.

3)在不同光照条件下或更复杂背景环境中的性能检测方面,所提算法可能存在一些未完全解决的误 检和漏检现象,尤其是在极端环境下,检测的准确性有待进一步研究.

4)本研究为目标检测模型在特征提取和参数优化上提供了新的改进思路.在实际应用中,改进算法具 备高精度和轻量化的特点,满足抓斗实时监控的需求,具有较高的应用价值.

5) 在未来研究中对算法的鲁棒性需要进行进一步验证, 特别是在不同光照和更复杂背景条件下的表 现.同时,可探索更多适合移动端部署的轻量化结构优化方案,以进一步提升模型的适用范围.

参考文献:

- [1] ZHANG X Y, LI J J, YANG Z L, et al. Collaborative optimization for loading operation planning and vessel traffic scheduling in dry bulk ports[J]. Advanced Engineering Informatics, 2022, 51: 101489.
- [2] WANG J Y, ZHANG X Y, GUO W Q, et al. Disruption management-based coordinated scheduling for vessels and ship loaders in bulk ports[J]. Advanced Engineering Informatics, 2023, 56: 101989.
- [3] 王熠杨, 胡志华. 考虑装卸设备阶段性均衡使用的散货码头靠泊计划[J]. 大连海事大学学报, 2023, 49(2): 103-112.
- [4] 毛国明,高峰,王悦州,等.抓斗式卸船机智能电气控制系统设计[J].今日制造与升级,2023(9):137-139.
- [5] 罗佩玉.新型抓斗与翻转集装箱多功能卸船机电气设计[J].起重运输机械,2023(18):54-59.
- [6] 周久华.港口装卸设备安全事故[J].中国安全生产,2022,17(4):52-53.
- [7] 陈建松.自动化控制港口码头散料卸船机的设计创新[J].中国机械,2023(21):22-25.
- [8] XUE B, HE Y, JING F, et al. Robot target recognition using deep federated learning[J]. International Journal of Intelligent Systems, 2021, 36(12): 7754-7769.
- [9] XUE B, YI W J, JING F, et al. Complex ISAR target recognition using deep adaptive learning [J]. Engineering Applications of Artificial Intelligence, 2021, 97: 104025.
- [10] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2016: 779-788.
- [11] 张文明,刘向阳,李海滨,等.基于深度学习的门机抓斗检测方法[J].光电工程,2021,48(1):200062.
- [12] CUI Z Y, LI Q, CAO Z J, et al. Dense attention pyramid networks for multi-scale ship detection in SAR images [J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 57(11): 8983-8997.
- [13] 张绍文,史卫亚,张世强,等.基于加权感受野和跨层融合的遥感小目标检测[J].电子测量技术,2023,46(18): 129-138.
- [14] 李佳东,张丹普,范亚琼,等.基于改进 YOLOv5 的轻量级船舶目标检测算法[J].计算机应用,2023,43(3):923-929.
- [15] ZHOU W N, PENG Y J. Ship detection based on multi-scale weighted fusion [J]. Displays, 2023, 78: 102448.
- [16] 罗芳,刘阳,何道森.复杂场景下自适应特征融合的多尺度船舶检测[J].计算机应用,2023,43(11):3587-3593.
- [17] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors [C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2023: 7464-7475.
- [18] XUE B, ZHENG Q H, LI Z N, et al. ISAR weak feature enhancement with perturbation defense using hybrid clustering oversegmentation [J]. IEEE Transactions on Aerospace and Electronic Systems, 2024, 60(5): 6256-6274.
- [19] TONG Z J, CHEN Y H, XU Z W, et al. Wise-IoU: bounding box regression loss with dynamic focusing mechanism [EB/ OL]. 2023: arXiv: 2301.10051. http://arxiv.org/abs/2301.10051
- [20] LI H L, LI J, WEI H B, et al. Slim-neck by GSConv: a lightweight-design for real-time detector architectures [EB/OL]. 2022: arXiv: 2206.02424. http://arxiv.org/abs/2206.02424.
- [21] OUYANG D L, HE S, ZHANG G Z, et al. Efficient multi-scale attention module with cross-spatial learning[C]//ICASSP 2023 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2023: 1-5.
- [22] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2017: 936-944.
- [23] 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. IEEE, 2018: 8759-8768.