DOI:10.20033/j.1003-7241.(2026)07-0031-06

一种面向港口危险区域入侵检测的改进型目标检测方法

刘红林[1], 孙子鉴[2], 史艺康[2]
(1.国家能源集团永州发电有限公司,湖南 永州 425000;2.燕山大学,河北 秦皇岛 066004)

摘　要:针对港口作业现场危险区域范围随工况变化而动态调整、监控画面中目标尺度差异大、复杂背景易造成误检漏检等问题,提出一种面向港口危险区域入侵检测的改进型目标检测方法。该方法以 YOLOv8 为基础,在骨干网络中嵌入通道—空间注意力机制,强化模型对小目标、弱特征目标及关键区域信息的表达能力。在特征融合阶段改进原有路径聚合结构,构建增强特征金字塔模块,以提升多尺度特征的传递效率、融合质量和复杂场景适应能力。结合港口现场实际需求,利用三角锥检测结果对作业危险区域进行动态划定,并通过射线判别策略实现区域入侵识别。为验证方法有效性,构建自制数据集,并在公共数据集上进行对比实验与消融实验。结果表明,改进方法在检测精度、召回能力和实时性之间取得了较好平衡,在自制数据集和公共数据集上的平均精度均优于多种对比模型。现场部署与测试结果进一步表明,该方法具有较好的工程应用价值和推广前景。

关键词:图像处理;深度学习;目标检测;特征融合;危险区域入侵检测;港口监控
中图分类号: TP391.41;U698.5　　**文献标志码**: A　　**文章编号**: 1003-7241(2026)07-0031-06

An improved object detection method for intrusion detection in port hazardous areas

Liu Honglin[1], Sun Zijian[2], Shi Yikang[2]
(1. State Energy Group Yongzhou Power Generation Co. Ltd., Yongzhou 42500, Hunan, China;
2. Yanshan University, Qinhuangdao 066004, Hebei, China)

Abstract: To address the problems that the boundaries of hazardous areas in port operation sites change dynamically with working conditions, target scales in surveillance images vary greatly, and complex backgrounds easily lead to false detections and missed detections, an improved object detection method for intrusion detection in port hazardous areas is proposed. Based on YOLOv8, the proposed method embeds a channel-spatial attention mechanism into the backbone network to enhance the model's ability to represent small targets, weak-feature targets, and key regional information. In the feature fusion stage, the original path aggregation structure is improved and an enhanced feature pyramid module is constructed to improve the transmission efficiency and fusion quality of multi-scale features, as well as the adaptability to complex scenes. In combination with the practical requirements of port sites, the hazardous operation area is dynamically delineated according to the detection results of traffic cones, and regional intrusion is identified through a ray-casting discrimination strategy. To verify the effectiveness of the method, a self-built dataset is constructed, and comparative as well as ablation experiments are conducted on a public dataset. The results show that the improved method achieves a good balance among detection accuracy, recall, and real-time performance, and its mean average precision on both the self-built dataset and the public dataset is superior to that of multiple comparative models. Field deployment and testing further demonstrate that the proposed method has good engineering application value and promising prospects for wider application.

Keywords: picture processing; deep learning; object detection; feature fusion; dangerous area intrusion detection; port control

在港口的现代化建设中,安全生产始终是第一位的,加强对危险区域的入侵检测就是保障手段之一[1]。随着深度学习的出现以及计算机算力的提升,图像检测摆脱了传统的灰度二值化处理方式,使得检测精度大幅提升。目前,基于深度学习的危险区域入侵检测系统[2]已经逐渐开始取代人力监控的方式。但在现场作业中,往往需要根据作业现场动态确定危险区域,此外因为港口区域很大,摄像头较多,而计算设备算力有限,如何在保证模型精度的基础上,尽可能减少计算代价,提高模型的检测速度,使得模型更具实用性是目前最大的问题。

目前主流的目标检测算法可以分为两类:第一类为双阶段算法,算法分为两步分开执行,输入图像先经过候选框生成网络,再经过分类网络,对候选框的内容进行分类。其典型代表有 R-CNN[3]、Faster-RCNN[4]、Mask-RCNN[5]等。双阶段算法在候选区域的提取过程中计算量大、过程复杂、检测速度慢、虽然拥有较高的准确度,但却无法满足

收稿日期:2024-12-16;**录用日期**:2025-02-24
基金项目:河北省自然科学基金(23280301Z)
作者简介:刘红林(1986—),男,工程师,研究方向为港口自动化,计算机视觉。
引用本文:刘红林, 孙子鉴, 史艺康. 一种面向港口危险区域入侵检测的改进型目标检测方法[J]. 自动化技术与应用, 2026,45(7):31-36. (Liu Honglin, Sun Zijian, Shi Yikang. An improved object detection method for intrusion detection in port hazardous areas[J]. Techniques of Automation and Applications, 2026,45(7):31-36.)

实时性的需求。第二类是单阶段的,主要是利用端到端的思想,采用整张图像来回归预测出目标物体的类别和位置。典型代表有 YOLO[6-8]系列、SSD[9]系列等。随着算法的更新,单阶段算法在精度上逐渐可与双精度算法比拟,但速度比双精度算法更快,更适用于实际应用。

本文提出了一种改进的 YOLOv8[10]算法,该算法在主干网络中集成了上下文感知卷积层和动态上下文融合层,通过自注意力机制强化了对小目标和复杂场景的检测能力。还优化了特征金字塔网络结构,以提高对多尺度特征的整合效率和检测精度。这一系列的改进显著增强了模型的实用性和准确性,尤其适用于港口监控场景[11-12]。本研究的算法在自建危险区域入侵数据集(dangerous zone intrusion, DZI)上进行了广泛的测试,实验结果表明,该方法能够有效地应对港口环境中的实际应用挑战,为港口安全监控提供了一种新的解决方案。

1 本文算法

本文提出的网络结构如图 1 所示,相比于原 YOLOv8 模型,在骨干网络中,集成了 CBAM[13]注意力机制,以增强模型对遮挡或远距离小目标的检测能力;改进原有的 PANet[14]网络结构,增加增强特征金字塔模块整合多尺度的特征信息,并通过增强的特征融合技术优化特征表示,特别适用于复杂的港口监控环境中,其中目标大小多变,背景通常复杂多样。

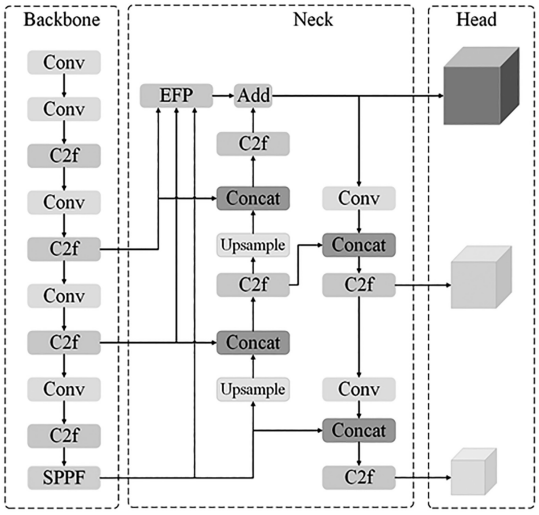


图 1 改进的 YOLOv8 网络结构

Fig. 1 Improved YOLOv8 network architecture

1.1 通道空间注意力模块

为了进一步增强模型对小目标的检测精度,本文引入了通道空间注意力模块(convolutional block attention module, CBAM)对图像特征信息进行更深入的提取。CBAM 是一种简单且有效的注意力机制,其具体结构如图 2 所示。它能够沿着通道和空间两个独立维度依次推断特征图信息,然后将注意力图与输入的特征图相乘,从而实现自适应特征细化。

CBAM 模块主要包括两个子模块:通道注意力模块(channel attention module, LAM)和空间注意力模块(spatial attention module, SAM)。在通道注意力模块中,输入的特征图首先经过全局平均池化和全局最大池化,生成两个不同的特征向量,这些向量通过共享的多层感知机处理后进行相加操作。生成的通道注意力图与原始特征图相乘,从而调整特征图中各个通道的重要性。接着,将调整后的特征图输入到空间注意力模块,空间注意力模块通过平均池化和最大池化在空间维度上进行聚合,并使用卷积操作生成空间注意力图。最终,将该注意力图与输入的特征图相乘,得到经过空间维度调整的特征图。

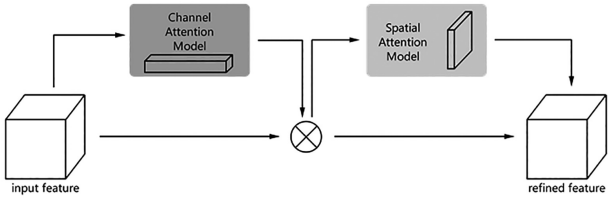


图 2 CBAM 模块

Fig. 2 CBAM module

在本文的网络结构中,CBAM 模块被融合到下采样操作中,使得网络模型能够有效排除环境因素产生的干扰信息,集中关注有用的目标信息。具体而言,在每次下采样操作中,首先对输入特征图应用通道注意力模块,然后对结果特征图应用空间注意力模块,如图 2 所示。通过这种方式,CBAM 模块能够在尽可能少增加参数量的同时显著提升模型对小目标的检测精度。

1.2 特征增强金字塔

为了提高小目标检测的性能并增强对多尺度特征的处理能力,在 YOLOv8 模型中引入了一个改进的结构,即增强特征金字塔(enhanced feature pyramid, EFP)。EFP 模块的设计旨在整合多尺度的特征信息,并通过增强特征融合技术优化特征表示,特别适用于复杂的港口监控环境中,其中目标大小多变,背景通常复杂多样。

增强特征金字塔模块作为颈部网络的一个关键部分,位于主干网络和检测头之间,以优化来自主干网络的 3 个关键特征层级(C3、C4、C5)。C3 层作为最高分辨率层,采用 1×1 和 3×3 的卷积核进行细节提取,同时使用空洞率为 2 的空洞卷积增加感受野,改善对小目标的识别。中等分辨率的 C4 层通过 3×3 的卷积核和同样空洞率为 2 的空洞卷积平衡细节保留与感受野拓展,有效捕捉目标及其上下文信息。对于最低分辨率的 C5 层,选择 5×5 的卷积核配合空洞率为 4 的空洞卷积,专注于捕获广阔的上下文信息,这对识别大型目标尤为重要。

为了整合这些多尺度特征,如图 3 所示 EFP 模块首先将 C5 层的特征图上采样至 C4 层的分辨率,同理,C4 层的特征图上采样至 C3 层的分辨率。通过元素级相加的方式,将这些上采样后的特征图与其相应的上一层进行融合。这种层级融合不仅加强了特征之间的直接联系,还优化了信息的整体流动。每次融合后,采用通道注意力机制(SE 模块[15])来重新校准融合特征图的通道重要性,进一

步提炼特征并增强对关键区域的响应。

最终,经过融合和优化的特征图通过一个 3×3 的卷积层进行最终整合,然后被送往模型的检测头部,用于目标的精确定位和分类。这种详细且针对性的设计确保了 EFP 模块能够有效地处理不同尺度的输入,并通过精心设计的融合策略显著提升整个模型在处理多尺度目标检测任务中的性能,特别是在复杂多变的港口环境中对各种尺寸目标的准确检测。

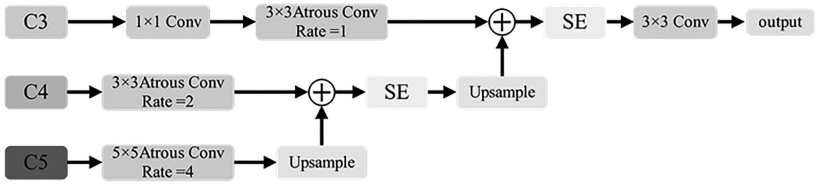


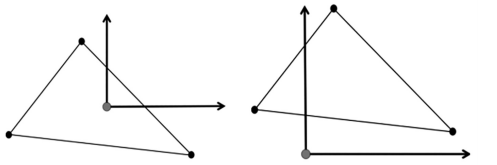
图 3　EFP 模块结构图

Fig. 3　EFP module structure diagram

1.3　危险区域确定和入侵检测

考虑的实际作业中,危险区域不是一个固定的范围,而是随着作业位置流动。因此,本文设计通过现场摆放的三角锥,动态确定危险区域范围。为了划定危险区域,需要检测三角锥并提取坐标,然后根据这些坐标构建连通区域,即为危险区域。

在确定危险区域的基础上,可以进一步判断是否有入侵情况发生。为判断是否有入侵情况发生,提出射线法判断。首先,确定入侵物体的坐标,以入侵物体的坐标为中心点,分别射出 2 条射线,判断射线与连通区域的边界有多少交点。如果中心点在连通区域中,那么每条引出的射线与边界有且仅有一个交点,如果中心点在连通区域外,那么引出的射线则会与边界有 2 个以上的交点。根据交点数判断是否有入侵情况发生。示意图如图 4 所示,图中 4(a)为有入侵情况发生,4(b)为无入侵情况发生。



(a) 为有入侵情况发生　　(b) 为无入侵情况发生

图 4　入侵检测示意图

Fig. 4　Intrusion detection diagram

2　实验环境和数据集介绍

2.1　实验环境

实验硬件环境,i7-6700 3.4 GHz CPU、NVIDIA CTX 2080Ti GPU,操作系统为 Ubuntu 18.04。软件环境,编程语言为 Python3.7,C++。

2.2　实验数据集介绍

本文实验共采用 2 种数据集,第 1 种为(safety-helmet-wearing-dataset,SHWD)数据集,该数据集为公共数据集;第 2 种为 DZI 数据集,该数据集为自己建立的数据集。

SHWD 数据集是一个基于安全帽检测的公共数据集,该数据集分为人员和安全帽 2 个检测目标,并且是一个标准的工地场景数据集,可以很好地适应于港口场景中。SHWD 数据集共 7 581 张图片,以 8∶2 的比例划分训练集和测试集。

DZI 数据集是自己建立的数据集。数据集格式按照 PASCAL VOC 数据集格式建立,视频来源于港口工程器械大臂上的摄像头,共包含不同的工程器械,不同的时间,不同的天气条件等视频 10 条。每个视频按照一定的间隔时间截取图片,然后对截取的图片进行人工筛选,剔除相似、相近、无检测目标的图片,总共制作 3 800 张数据集。数据集以 2∶8 的比例划分测试集和训练集。

标注数据集使用的软件为 LabelImg,数据集共标注有 8 类。由于视频源为摄像头,而摄像头基本的安装位置为大型工程器械的大臂,这就导致了标注目标过于细小并且受环境影响大,如图 5 所示。为解决这个问题,对数据集进行手动标注,然后人工复检,以保证数据集中的目标标注准确。



图 5　数据集标注示意图

Fig. 5　Schematic diagram of the annotation of the dataset

3 实验结果与分析

3.1 评价指标

常用的目标检测模型性能评估通常用准确率(Precision)和召回率(Recall)来评价目标检测模型的综合性指标,mAP 值被用来描述多标签图像分类的性能,AP 是各类别平均精度(average precision,AP)的均值。这些参数用来描述检测性能。相应的计算过程如下。

1) 准确率 P (Precision)。准确率可以衡量模型的精确性,可以定义为模型预测的正类别中有多少是真正的正类别,如公式(5)所示:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{5}$$

其中,TP 表示模型正确地将正例判定为正例的数量,FP 表示模型错误地将负例判定为正例的数量。

2)召回率 R (Recall)。召回率可以衡量模型的全面性,可以定义为所有真正的正类别中有多少被模型成功检测到,如公式(6)所示:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + FN} \tag{6}$$

其中, FN 表示模型为漏报的正样本数量。

3) 平均准确率均值(mean average precision,mAP)。平均准确率均值是一种综合评价指标,它综合考虑了模型在不同类别上的精准率,是所有类别 AP 的均值,公式(7)所示:

$$\text{mAP} = \frac{\sum_{i=1}^{d} \text{AP}_i}{d} \tag{7}$$

其中, AP_i表示单个类别的平均精度,d 表示所有类别的数量。

3.2 DZI 数据集实验结果对比分析

在对 DZI 数据集进行的实验中,对比了本文提出的改进 YOLOv8 模型与其他几种目标检测模型。实验结果如表 1 所示,改进后的 YOLOv8 模型在检测精度方面表现优异,达到了 95.1%,这比 YOLOv4、YOLOv5s、SSD 以及 YOLOv7 模型的表现都要好。同时,本文模型的参数量保持在一个较低水平,实现了对模型轻量化的需求。

表 1 DZI 数据集对比实验

Tab. 1 Comparative experiments on DZI dataset

算法	P/%	R/%	mAP@0.5/%	帧率/(f·s^{-1})
YOLOv4	88.0	85.3	86.5	60
YOLOv4-tiny[2]	67.1	64.6	65.5	90
YOLOv5s	92.3	90.5	91.4	70
YOLO V8	93.6	91.5	92.5	55
SSD	79.3	76.9	77.5	50
YOLO V7	90.1	88.6	89.3	48
本文算法	95.8	93.4	95.1	50

具体来看,YOLOv4 模型的 mAP 为 86.5%,而本文的 YOLOv8 模型则提升了 8.6 个百分点。YOLOv5s 模型虽然也显示了较高的检测精度(91.4%),但本文模型仍然高出 3.7 个百分点。此外,尽管 YOLOv7 提供了相对较高的检测精度(89.3%),本文模型的性能提升依然明显。

在检测速度方面,本文模型以较高的精度和适中的检测速度(50 FPS)保持了良好的平衡。这一点对于需要在资源受限的环境中部署高精度检测模型的应用场景尤为重要。

表 2 DZI 数据集消融实验

Tab. 2 Ablation experiments on DZI dataset

组合	EPF	CBAM	P/%	R/%	mAP@0.5/%
组合 1			87.4	84.6	85.5
组合 2	√		92.7	89.8	90.2
组合 3		√	90.5	87.7	89.5
组合 4	√	√	95.8	93.4	95.1

通过这些比较分析,可以明显看出本文提出的改进 YOLOv8 模型在处理 DZI 数据集时,不仅在减少模型参数量的同时,还显著提高了检测精度,有效地平衡了性能与效率,使其成为复杂港口监控环境中的优选模型。

在 DZI 数据集上进行的消融实验清楚地展示了不同模块对模型性能的具体影响。实验设置包括 4 种不同的配置:未使用任何增强模块,仅使用增强特征金字塔模块,仅使用 CBAM 自注意力模块,以及两者的结合。结果显示,在未应用任何增强模块的基础配置中,模型的 mAP 最低,为 85.5%,这强调了增强模块对提升性能的重要性。引入 EFP 模块后,模型的 mAP 显著提升到 90.2%,表明该模块通过整合多尺度特征信息有效地提高了检测精度。而当使用 CBAM 模块时,mAP 提升到 89.5%,证实了 CBAM 通过加强特征图的通道和空间关注度有效提升了模型性能。最终,当 EFP 和 CBAM 模块同时使用时,模型表现最优,mAP 达到 95.1%。这一结果突出了这两个模块的互补作用,它们的联合应用显著优化了模型的整体表现,特别是在不显著增加参数量的前提下,通过精细化的特征处理显著提高了复杂场景下的检测性能。这种细粒度的特征优化使得改进的模型在实际港口监控应用中表现出较高的实用价值。

3.3 SHWD 数据集实验结果对比与分析

本文选用的是 SHWD 数据集,具体实验结果如表 3 所示,本对比实验的网络输入为 416×416。

在 SHWD 数据集上的实验结果显示,本研究提出的改进型 YOLOv8 模型表现出显著的优越性,尤其是在复杂的工地安全帽检测场景中。具体而言,YOLOv8 模型在此数据集上实现了 90.4% 的平均精确率,明显超越了 YOLOv4、YOLOv7,以及轻量级的 YOLOv4-tiny 和 SSD 模型。这一结果突显了本文模型在精确目标识别和定位能力上的显著提升,特别适合于对高准确度检测有严格要求的实时应用场景。

表 3　DZI 数据集对比实验

Tab. 3　Comparative experiments on DZI dataset

算法	*P*/ %	*R*/ %	mAP@ 0. 5/%	帧率/(f · s^{-1})
YOLOv4	85. 8	83. 0	84. 0	55
YOLOv4-tiny	78. 3	75. 6	76. 5	80
YOLOv5s	88. 5	86. 3	87. 3	70
YOLO V8	90. 4	88. 5	89. 1	60
SSD	80. 1	78. 1	79. 6	65
YOLO V7	86. 6	84. 6	85. 3	58
本文算法	92. 5	90. 0	91. 2	50

在帧率方面,虽然 YOLOv8 的速度不是所有模型中最快的,但它成功地在保持高检测精度的同时,提供了可接受的处理速度,这对于实时监控系统至关重要。这表明本文的模型在确保较高的检测准确性的同时,还能维持必要的操作速度,满足了实际部署中对速度和精度的双重要求。

总结而言,本研究中的数据表明,改进后的 YOLOv8 模型在应对具有复杂背景和多变目标大小的实际场景时,能够显著提升检测精度和速度。其性能表现特别是在工地场景中的应用中显示了模型的优越性,反映了其在实际部署中的高效性和广泛的适用性。

4　算法部署与应用

算法部署以 MFC 为基础,C++为编写语言搭建完成。包括检测系统、入侵系统和报警系统构成,检测系统检测和识别图像中各个物体的位置;经由入侵系统根据三角锥构建危险区域后,判定是否有入侵行为发生;根据入侵系统提供的结果,报警系统给出相应的报警提示,有入侵情况发生时报警灯会变红。

在唐山港对本模型进行部署与测试,港口现场结果如图 6 所示。其中图 6(a)和图 6(b)分别为 2 个不同地点的区域入侵检测。当有无关车辆进入由三角锥划定的危险区域时,就判定为发生区域入侵。根据现场验证,本系统完全满足实践要求。



(a) 一号煤堆



(b) 二号煤堆

图 6　现场部署图

Fig. 6　Field deployment diagram

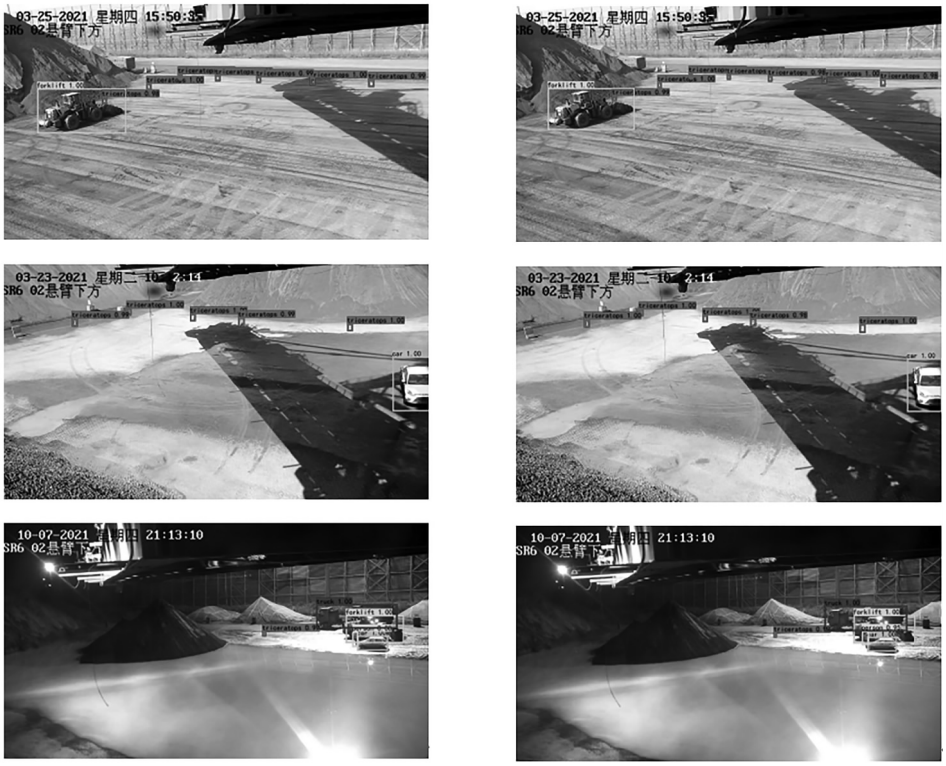
本文对不同场景下改进之后的算法和基础算法测试图片进行对比如图 7 所示。图 7(a)为原 YOLOv8 算法下的测试图片结果,图 7(b)为改进后算法的测试图片结果。

对改进后的模型在不同场景下的入侵检测准确率进行统计,统计结果如表 4 所示。通过统计结果可以看出,文中的算法检测准确率完全符合实际应用需要。既能在白天视野条件好的时候精准检测,也能在光照不足的夜间完成入侵检测,适用性广泛。

表 4　不同场景入侵检测准确率

Tab. 4　Intrusion detection accuracy in different scenarios

序号	危险区域类别	危险场景	入侵检测准确率/%
1	违章进入	1 号煤堆	95. 19
2	违章进入	2 号煤堆	94. 84
3	违章进入	夜间 3 号煤堆	95. 04



（a）YOLOv4 （b）本文算法

图7 测试图片对比

Fig. 7 Comparison of test pictures

5 结论

本文提出了一种改进的YOLOv8模型,专门针对复杂港口环境的危险区域入侵检测。此模型通过集成通道空间注意力机制,显著提高了对遮挡或远处小目标的检测效率,优化了原始PANet结构,引入了增强特征金字塔模块,以更好地处理多尺度特征和应对目标大小多变及背景复杂的监控场景。经过广泛的测试,改进后的模型在自制的DZI数据集上表现出了优越的检测性能,部署结果显示,本文算法在唐山港作业现场应用良好,有效地保障了港口作业人员的生命财产安全。

参考文献

[1]陈家豪, 刘克中, 陈默子, 等. 基于信道状态信息的船舶敏感区域入侵检测方法[J]. 大连海事大学学报, 2019, 45(1):89-95.

[2]朱鹏浩, 李军, 张世义, 等. 基于深度学习的动态危险区域入侵检测方法[J]. 中国安全生产科学技术, 2024, 20(7):170-178.

[3]Ross G, Jeff D, Trevor D, et al. Region-based convolutional networks for accurate object detection and segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 38 (1): 142-58.

[4]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, 2017, 39(6): 1137-1149.

[5]Min H, Wilson D, Huang Y, et al. Fully automatic computer-aided mass detection and segmentation via pseudo-color mammograms and Mask R-CNN [C]//2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI). Iowa City, IA, USA: IEEE, 2020:1111-1115.

[6]Tian Y, Yang G, Wang Z, et al. Detection of apple lesions in orchards based on deep learning methods of CycleGAN and YOLOV3-Dense [J]. Journal of Sensors, 2019 (1): 1-13.

[7]Mahto P, Garg P, Seth P, et al. Refining Yolov4 for vehicle detection [J]. International Journal of Advanced Research in Engineering and Technology, 2020, 11 (5): 409-419.

[8]朱梦瑞, 牛宏侠. 改进YOLOv3模型的铁路异物入侵识别算法[J]. 北京交通大学学报, 2022, 46(2):37-45.

[9]Liu W, Anguelov D, Erhan D, et al. SSD: Single shot multiBox detector [C]//Computer Vision-ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, 2016, Proceedings, Part I. Cham: Springer, 2016: 21-37.

[10]Fatma M T, Hanaa Z. An improved fire detection approach based on YOLOv8 for smart cities [J]. Neural Computing and Applications, 2023, 35 (28): 20939-20954.

[11]福斌, 殷浩然, 任莉芸, 等. 基于改进的YOLOv8尺度自适应大渣检测算法研究[J]. 自动化技术与应用, 2026, 45(3):128-132, 146.

[12]王宁生, 解志斌, 邵长斌. 融合多尺度信息的水下目标检测算法研究[J]. 自动化技术与应用, 2026, 45(4):1-5.

[13]Pang B. Classification of images using EfficientNet CNN model with convolutional block attention module (CBAM) and spatial group-wise enhance module (SGE) [C]//International Conference on Image, Signal Processing, and Pattern Recognition (ISPP 2022). Bellingham, WA: SPIE, 2022: 1224707.

[14]Wang K, Liew J H, Zou Y, et al. PANet: few-shot image semantic segmentation with prototype alignment[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, Korea (South): IEEE, 2019: 9197-9206.

[15]Jian Z, Siwen W, Ming Q, et al. Improved aircraft flared tube defect detection algorithm of YOLOv4 network structure [J]. Journal of Physics: Conference Series, 2022, 2252(1): 012050.