

基于深度学习的散货港口接卸运载车辆识别方法

张凤翔¹, 毛锐¹, 周泉², 钱欣漪³, 剪欣², 易双²

(1. 天津港远航国际矿石码头有限公司, 天津 300456; 2. 中交水运规划设计院有限公司(中交集团自动化码头技术研发中心), 北京 100007; 3. 上海海事大学 物流工程学院, 上海 201306)

摘要: 针对散货港口抓斗门机智能化改造过程中, 基于图像特征的陆侧接卸运载车辆识别和定位方法中存在的时间成本高, 效率低下的问题。提出一种基于 LW-YOLOv5 的散货港口接卸运载车辆识别算法。通过色相饱和度明度 (hue saturation value, HSV) 颜色空间实现目标的快速粗定位, 再利用轻量化的算法实现目标的精确定位。最后, 使用该方法应用于实际工程项目。结果显示, 与已有算法相比, 所提方法在不影响识别准确度的同时, 能够降低时间成本, 提高检测速度。

关键词: 散货港口; 接卸车辆; 抓斗门机; YOLOv5; 深度学习; 卷积神经网络; 图像识别

中图分类号: TP391.413

文献标志码: A

文章编号: 1003-7241(2026)02-0058-06

Recognition of bulk cargo port unloading vehicles based on deep learning

ZHANG Fengxiang¹, MAO Rui¹, ZHOU Quan², QIAN Xinyi³, JIAN Xin², YI Shuang²

(1. Tianjin Port Yuanhang International Ore Terminal Co., Ltd., Tianjin 300456, China;

2. CCCC Water Transportation Consultants Co., Ltd. (CCCC Automation Wharf Technology Research and Development Center), Beijing 100007, China; 3. Logistics Engineering College, Shanghai Maritime University, Shanghai 201306, China)

Abstract: In the process of intelligent transformation of grab cranes at bulk cargo ports, issues such as high time costs and low efficiency exist in the recognition and localization of land-side loading and unloading vehicles based on image features. This paper proposes a cargo port loading and unloading vehicle recognition algorithm based on the LW-YOLOv5 model. The algorithm first achieves rapid coarse localization of the target using the hue saturation value (HSV) color space, followed by precise localization through a lightweight model. Finally, the proposed method is applied to an actual engineering project. The results indicate that, compared to existing algorithms, the proposed method reduces time costs and improves detection speed without compromising recognition accuracy.

Keywords: bulk cargo ports; unloading vehicles; grab portal crane; YOLOv5; deep learning; convolutional neural network; image recognition

随着智能技术的不断进步, 人工智能给各行各业带来重大变化和一系列效益。物流业作为受益对象之一, 智能化的物流设备、物流管理在物流业中迅速发展, 促使物流业全面进入智慧物流时代^[1]。散货港口作为我国大宗散货物流领域的重要组成部分, 其智慧化升级进程正随着国家对智慧散货港口的政策性扶持和引导力度的加大而不断加快。抓斗门座式起重机(简称抓斗门机)卸船作业工艺中, 国内很多港口在陆侧采用重型接卸运载车辆进行卸船物料的转运, 如图 1 所示。

散货码头抓斗门机作业过程中, 通常需要对卸车辆的位置和姿态进行识别检测, 其作用主要体现在安全保护和作业识别两个方面。安全保护主要是需要识别出车辆的位置, 避免抓斗从车辆上方飞跃, 避免抓斗中物料砸落

至车辆驾驶室上方, 引起高空坠物伤人事故; 作业识别主要是需要识别出指定接卸区域是否有车辆到位就绪, 以及车辆的朝向、位置和姿态, 从而控制抓斗的运动目标位置控制, 实现抓斗至车辆车厢的物料卸载转运。在传统人工机上手动操作时, 主要依赖操作人员的眼睛进行生物视觉观察和判断。作业安全和作业效率完全取决于操作人员, 自动化和智能化水平亟待提高^[2]。在进行抓斗门机智能化改造后变为无人控制, 机上不再配备操作人员, 全部依赖计算机控制系统进行自动的作业环境感知和对象识别, 并进行相关的控制。为此急需一种行之有效的接卸车辆识别方法, 来保证在抓斗门机作业过程中迅速可靠地识别出接卸车辆的位置, 保证抓斗门机的安全高效作业。

近几年, 在目标检测识别领域主要有两种研究方向。

收稿日期: 2024-07-05; 录用日期: 2024-11-22

基金项目: 上海市自然科学基金项目(22ZR1427700); 上海市教育科学基金项目(B2023003)

作者简介: 张凤翔(1982—), 男, 本科, 高级工程师, 研究方向: 散货港口重型起重设备自动控制技术。

通信作者: 周泉(1979—), 男, 博士研究生, 正高级工程师, 研究方向: 港口装备智能化感知与控制技术。

引用本文: 张凤翔, 毛锐, 周泉, 等. 基于深度学习的散货港口接卸运载车辆识别方法[J]. 自动化技术与应用, 2026, 45(2): 58-62, 160. (ZHANG Fengxiang, MAO Rui, ZHOU Quan, et al. Recognition of bulk cargo port unloading vehicles based on deep learning[J]. Techniques of Automation and Applications, 2026, 45(2): 58-62, 160.)

一种是利用传统机器学习算法进行目标检测,使计算机在没有明确编程的条件下拥有学习的能力,并通过对大量数据的学习找出完成任务的方法^[3];另一种是利用深度卷积神经网络算法进行目标检测。它具有局部连接和参数共享的特点,在处理数据时具有时移不变性的优异特性^[4]。



图1 散货港口抓斗门机卸船和接卸车辆

Fig. 1 Grab portal crane and hopper truck of bulk material harbors

机器学习方面,文献[5]将小波变换与特征提取融合,针对复杂遥感图像中丰富的信息,提出了一种新的边缘检测算法提高了图像边缘检测的精度和可靠性。Madey等^[6]开发了一种能够对视频中的车辆进行分类的方法,该方法涉及HOG特征、线性SVM分类器和YOLOv3算法,提高了训练精度。其中特征提取的优劣将直接影响传统机器学习方法的最终识别性能^[7]。而实际图像的特征提取受环境、位置等因素影响,导致传统机器学习方法效果不佳。

现如今,利用深度学习进行目标检测的研究百花齐放,卷积神经网络是所有其他复杂网络的基础,基于深度学习下的目标检测识别领域多以卷积神经网络(convolutional neural networks, CNN)结构为主^[8],文献[9]针对普通神经网络特征提取较为单一和网络退化的问题,提出一种基于卷积神经网络的Inc-Resnet细粒度车型识别方法,有效提取车型特征,提升识别准确率。文献[10]针对大规模网络结构存在其深层网络损耗值大于其浅层网络损耗值的缺点,通过残差映射对网络层的学习函数进行再造,解决这一问题。文献[11]将残差思想,注意力机制和多尺度特征提取三者结合起来识别车辆颜色。文献[12]针对恶劣天气对车牌识别的影响,使用U-net分割模型将车牌细致分割,结合自适应权重和注意力机制提升模型性能与准确率。Zaafouri等^[13]将局部功率谱(local power spectrum, LPS)特征图和卷积神经网络相结合。Pustokhina等^[14]将最优K-mean(OKM)聚类 and CNN模型结合。仅通过车辆的某一个属性来定位一个特定车辆往往很难实现,而且车辆的外在属性可以确定的信息越多,对定位特定车辆帮助则越大^[15-16]。De Oliveira等^[17]同时以车辆的外观和车牌为特征作为输入,提高了识别准确率。相较于准确度,YOLO系列算法因其速度快,模型体积小,结构简洁的优点被广泛应用于工业生产和生活中^[18-20]。

本文提出了一种基于LW-YOLOv5的接卸车辆识别算法,综合考虑了准确率和识别速率的问题,使用颜色特征对目标粗定位,并采用轻量化的算法模型,加快目标的

检出速度。此外,通过工业现场测试结合对比试验证明本方案的优越性。

1 车辆识别与定位

1.1 粗定位算法

实际采集的港口散货接卸车辆的图像,受到恶劣天气,光照不均等因素的影响,不利于对接卸车辆的识别。故首先对目标进行粗定位。

车辆的显著特征去除车牌、车型等就是车辆颜色。针对本文的研究对象散货装载车,选取车辆中颜色特征集中的区域,使用颜色特征用来对目标进行粗定位,帮助模型更有效地训练,以达到更好的性能。

而本文所使用的图像背景在RGB颜色空间中有较多元素与车辆的颜色类似,因此选择不同的颜色空间用于提取车辆的颜色特征。颜色空间又称为彩色模型,其中最常见的是RGB模型。然而,在进行颜色识别时,通常不采用RGB模型,原因在于其主要描述的是红、绿和蓝三种原色的颜色通道。这种模型在处理颜色识别时将色调、亮度和饱和度三个参数交织在一起,导致难以有效区分和单独分析这些关键特征。相比之下,在HSV颜色空间通过颜色范围进行滤波,使得目标检测能够更准确、更高效地进行^[18],因此本文将图像转为HSV颜色空间,HSV与RGB的变换公式为

$$V = \max(R, G, B) \quad (1)$$

$$S = \begin{cases} \frac{V - \min(R, G, B)}{V}, & V \neq 0 \\ 0, & V = 0 \end{cases} \quad (2)$$

$$H = \begin{cases} \frac{V - \min(R, G, B)}{V}, & V = R \\ 120 = 60(B - R) / (V - \min(R, G, B)), & V = G \\ 240 + 60(R - G) / (V - \min(R, G, B)), & V = B \end{cases} \quad (3)$$

车辆的色调相对固定;由于现场光照因素的影响,车辆的颜色饱和度与亮度变化较大,而光照变化会影响整张照片的颜色饱和度,可以建立车辆的颜色饱和度与整张照片的颜色饱和度均值之间的关系;车辆的亮度对提取图像的影响不大,为了较好的泛化能力可以选取较宽的阈值范围。对所用图像进行抽样统计散货装载车的颜色特征信息,使用均值法对结果进行均值处理得到如下的关系式和阈值范围(\bar{S} 表示整张照片颜色饱和度的均值)。

$$\begin{cases} H = 42, [-10, 10] \\ S = 1.474 \times \bar{S}, [-25, 25] \\ V = 150, [-50, 50] \end{cases} \quad (4)$$

下一步使用选定的阈值范围对原始图像进行二值化。由于车辆在图像中并非单一的纯色,并且受到车辆姿态与现场光照条件等因素的影响,二值化处理往往会导致车辆被分割为多个不连续的轮廓,从而使得部分车辆特征信息丧失。为有效区分不同的车辆与噪声信息,本文采用形态

学闭运算来整合相邻的轮廓。考虑到车辆在图像中的尺度相对固定,利用轮廓面积参数对提取的目标进行筛选,以保留感兴趣的目标区域。最终,以该目标区域为中心,构建一个边长为车辆平均像素长度两倍的方形,从而实现目标的粗定位。当同一目标因多个轮廓导致多个粗定位框时,此时根据多个定位框的交并比(IoU)选出最合适的粗定位框。目标粗定位的流程如图2所示,粗定位的示例如图3所示。

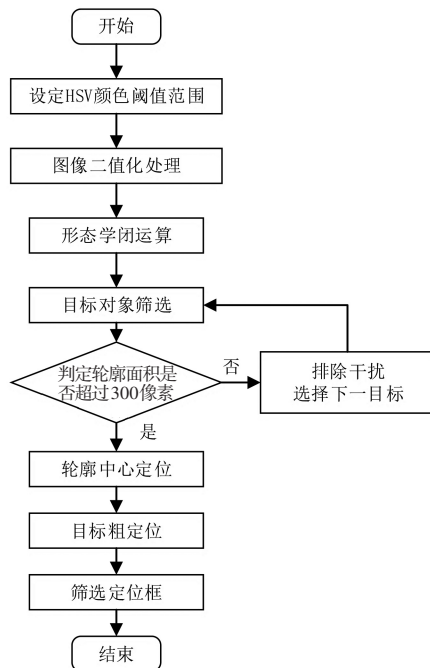


图2 粗定位算法流程图

Fig. 2 Flowchart of the coarse localization algorithm

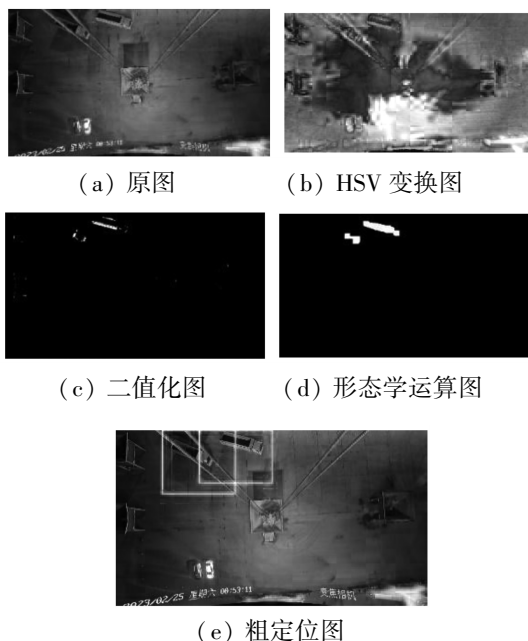


图3 车辆目标粗定位示例

Fig. 3 Example for vehicles of the coarse localization

1.2 精定位算法

LW-YOLOv5 算法是一种轻量化的 YOLOv5 变体,在保持 YOLOv5 整体框架不变的前提下,通过对网络模型进

行轻量化处理,从而提升了检测速度。YOLOv5 的架构主要由3个部分组成。骨干网络(Backbone)、颈部结构(Neck)和检测头(Head)。本文将重点对骨干网络和颈部结构进行改进,以优化整体性能。

骨干结构(Backbone)采用了 CSP 模块和 Focus 模块。跨阶段部分网络(cross stage partial network, CSP)网络结构是一种增强 CNN 学习能力的跨缩阶段局部网络,降低20%计算量的同时提高了 CNN 的能力。在图片进入骨干结构前,对图片进行切片操作,采用隔列采样拼接,把高分辨率特征图拆分成多个低分辨率特征图,减少了模型需要处理的像素数量。由于图片的图像采集位置相对固定,车辆结构在图片中所占比例基本不变。依据所用图片数据中散货接卸车辆的特征,对骨干结构进行优化,删去不必要的卷积层。改进后的骨干结构如图4所示。

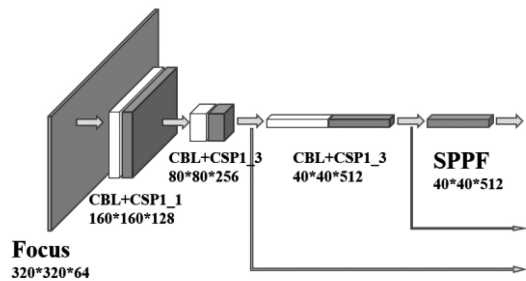


图4 改进的骨干结构

Fig. 4 Enhanced backbone structure

颈部结构(Neck)采用特征金字塔网络(feature pyramid network, FPN)加路径聚合网络(path aggregation network, PAN)双金字塔结构。FPN 通过将低层特征和高层特征融合起来,得到一个识别和定位都准确的目标检测结构。PAN 是一个用于实例分割的路径聚合网络,将不同尺度和不同通道的特征进行聚合,使得模型能够更好地利用全局和局部信息。改进后的颈部结构只保留用来检测中小型物体的锚,删除用来针对大型目标识别的指定卷积层。改进后的颈部结构如图5所示。LW-YOLOv5 模型整体结构如图6所示。

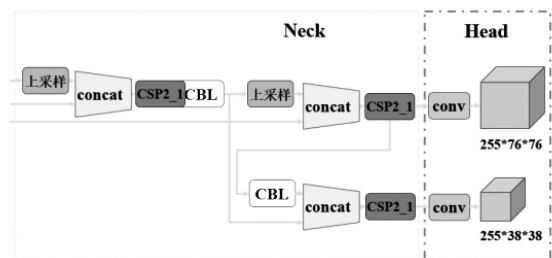


图5 改进的颈部结构

Fig. 5 Enhanced neck structure

具体目标精定位实现过程如下。首先根据粗定位得到的图像框,分别将各个框外图像变成黑色背景,依次将图像输入给训练好的 LW-YOLOv5 模型;其次通过训练好的模型权重对目标进行精确定位;最后将多个预测到的目标整合到一张完整的图像上。精定位示例流程如图7所示。

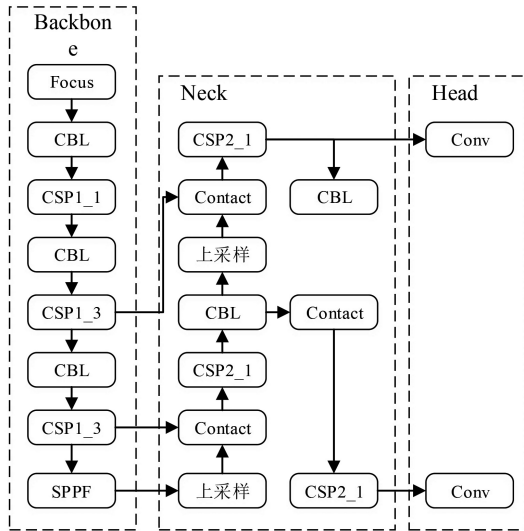


图6 LW-YOLOv5 模型结构

Fig. 6 LW-YOLOv5 model structure

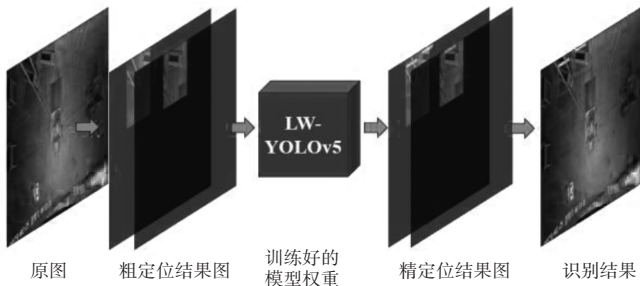


图7 精定位的识别流程

Fig. 7 Flow of fine localization process

2 实验与结果分析

2.1 实验环境配置

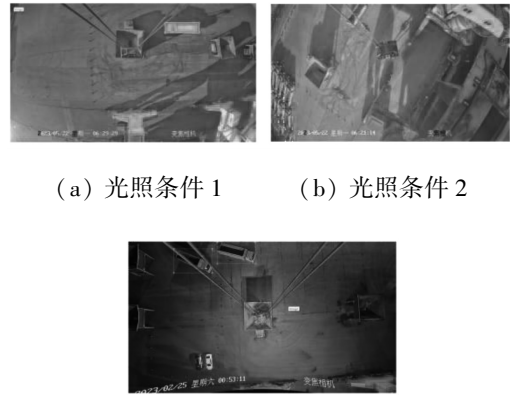
本文采用 i7-9750H+NVIDIA2060 的笔记本电脑,搭配 16 GB 内存,电脑的操作系统为 Windows10。利用 anaconda 构建 python 版本 3.9 的虚拟环境,并使用 CUDA11.8 版本的 pytorch 深度学习框架,使用 PyCharm2023 软件作为环境搭建与显示的界面,同时使用 LW-YOLOv5 模型,以 YOLOv5s 作为原始权重文件。以 LabelImg 作为图像标注工具。

2.2 实验预处理

2.2.1 数据集来源与目标标注

实验中使用到的数据集图片来源于天津某港口实地监控变焦相机拍摄,数据集共 2 000 张图片,图片的分辨率为 2 560×1 440 像素。该数据集由挂在抓斗架上的相机从高空向下俯视拍摄而成,图像包含了照明变化,拍摄角度变化。由于实地拍摄在不影响港口实际工作的前提下与拍摄捕获延迟的影响,部分车辆未全部包含在拍摄图片中,并且图片中存在与散货接卸车辆无关的其他车辆与港口工作设施。由于本实验对象为散货接卸车辆,因此数据集只标注散货接卸车辆(HopperTruck)这一种车辆作为训练识别的对象。将所有的图像数据的 70% 作为训练集,20% 作为验证集,10% 作为测试集,训练集中被遮挡、

未拍全等目标不予标注,标注示例如图 8 所示,将标注信息保存为 YOLO 格式的 txt 文件。



(a) 光照条件 1 (b) 光照条件 2

(c) 光照条件 3

图8 不同场景图片标注示例

Fig. 8 Example of different scenes picture annotation

2.2.2 图片畸变校正

对散货接卸车辆进行拍摄时是将三维物体成像到二维平面图片之中,受镜头、相机的拍摄角度和距离、环境的光线条件的影响,使得拍摄完成的图片与实际拍摄物体间存在差异,拍摄完成的图片存在畸形形变和失真。本次实验中散货接卸车辆的图片主要受径向畸变影响。径向畸变公式(5)如下,其中, (x, y) 是理想的无畸变的归一化图像坐标, r 为图像像素点到图像中心点的距离, k_1, k_2, k_3 是径向畸变系数。

$$\begin{cases} \hat{x} = x(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) \\ \hat{y} = y(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) \end{cases} \quad (5)$$

为了校正图像的径向畸变,本文采用张正友棋盘格标定法。使用棋盘格作为标定板,通过拍摄不同角度和位置的图像,提取图像中的角点坐标并结合棋盘格的实际物理尺寸,利用数学算法计算出相机的内外参数以及畸变系数,利用这些参数来校正镜头畸变。

2.3 实验结果

根据图 9 可以看出,当经过 329 次的训练迭代之后, box_loss 约为 0.02, obj_loss 趋近于 0.02。 cls_loss 为 0,是因为本文的图像识别对象只涉及散货接卸车辆,只有散货接卸车辆这唯一的图像分类,因此不存在分类损失。精确率和召回率的波动情况都较为平缓, $mAP@0.5$ 为 0.995, $mAP@0.5:0.95$ 为 0.863。通过模型评估可以看出实验所涉及的模型有着较高的检测精度和较强的泛化能力,有着良好的图像识别能力,为之后应用于实际中提供了数据支持。

为直观验证所提方案的可行性,本文选取了 YOLOv5、LW-YOLOv5 以及粗定位结合 LW-YOLOv5 等模型,针对门座式起重机抓斗在不同时间段的卸船作业进行了对照实验,实验结果如上所示。结果表明,车辆的颜色特征可以作为神经网络模型进行定位的辅助方法之一。然而,由于作业场景中存在诸如散货卸料斗、地面环境以及光照条件等干扰因素,粗定位的效果可能会受到一定影响。

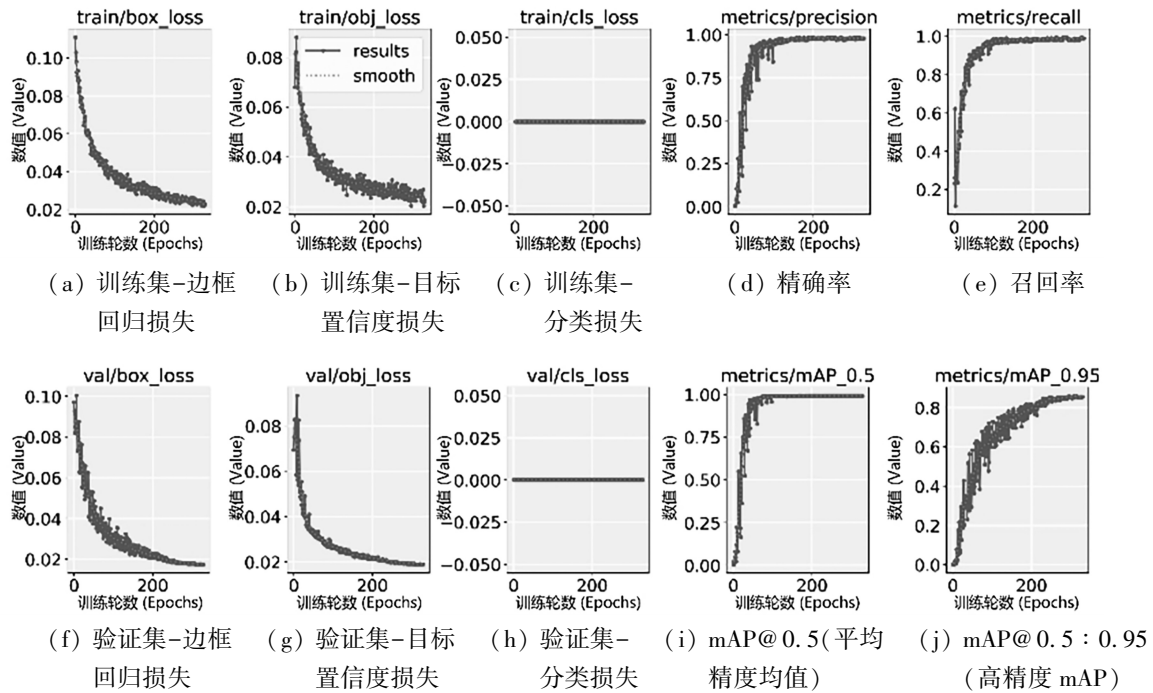


图9 模型评估图

Fig. 9 Model evaluation result

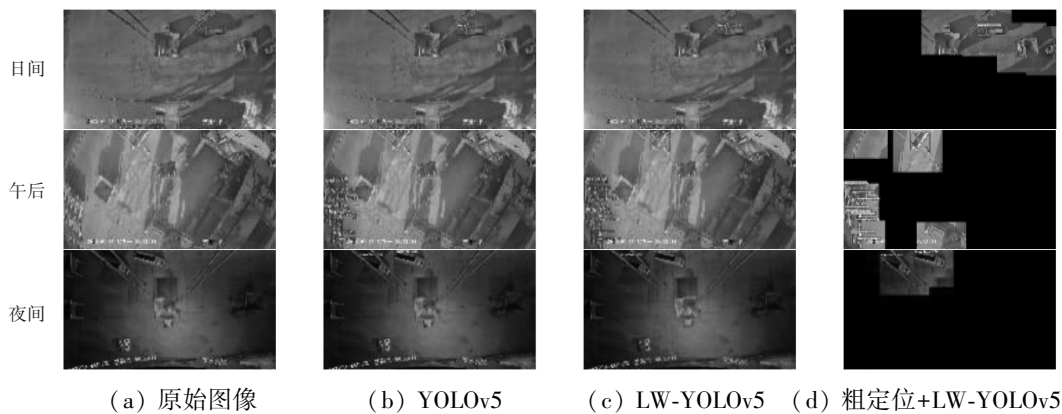


图10 检测效果对比图

Fig. 10 Comparison chart of detection results

表1 不同模型检测速度对比表

Tab. 1 Comparison of detection effectiveness

模型	推理速度/s	准确度/%	召回率/%
YOLOv5 原模型	0.036 9	99.39	99.81
LW-YOLOv5 模型	0.028 9	98.53%	98.91
粗定位+LW-YOLOv5	0.026 3	98.86	99.23

尽管如此,颜色粗定位依然展现出显著的优势。从表1数据可以看出,颜色粗定位在提升神经网络模型预测速度方面具有积极作用。颜色特征能够在复杂场景中实现目标物体的初步分割,从而有效缩小搜索范围,减少精确定位所需的计算资源。此外,颜色信息在光照条件相对稳定时表现出较强的鲁棒性,能够在动态变化的环境中帮助模型维持较高的定位精度。通过将颜色粗定位与深度学习模型的精确定位方法相结合,不仅能够显著提升系统的整体性能,还可以有效减少误差的累积,进而实现更为高

效的作业流程。

3 结论

本文主要针对散货港口抓斗门机陆侧散料接卸过程中存在的车辆识别问题,提出了一种基于 LW-YOLOv5 的散货接卸车辆检测与识别算法。该算法首先利用车辆的颜色特征进行初步定位,随后采用轻量化的算法模型对目标进行精细定位。为验证该方案的可行性,在工业现场进行了实地测试。实验结果表明,与传统算法相比,该方案在检测速率方面有所提升,证明了其有效性。

参考文献

[1]高明. 人工智能在物流行业的应用与发展探讨[J]. 全国流通经济, 2023(17):30-33.
[2]孙菊. 加强干散货码头的智能化应用[J]. 中国水运(下半月), 2019,19(1):113-114.

(下转第160页)