

一种基于级联边框精确的 YOLOv8 外集卡放箱方法

王兴峰¹ 王君雄² 王艳宾² 丁睿晟¹ 安津晖¹

1. 天津港太平洋国际集装箱码头有限公司 天津 300463

2. 三一海洋重工有限公司 广东 珠海 519050

摘要: 为了提高场桥自动化作业性能, 针对目前近距外集卡放箱技术在锁头检测时召回率低和检测边框不精确的问题, 提出了一种基于改进YOLOv8的外集卡放箱方法。首先, 在检测头网络中, 设计一种类别无关的感兴趣区域检测分支替换原有的类别检测分支, 使模型可以尽可能检测感兴趣目标而无需区分目标类别, 降低检测难度, 提高模型召回率; 其次, 引入级联结构网络处理感兴趣目标, 在为目标分类的同时逐级精确检测边框; 最后, 构造对应的类别损失函数完善训练步骤。应用结果表明, 改进YOLOv8能更好地赋能自动化外集卡放箱, 提高放箱效率。

关键词: 场桥; 自动化作业; 外集卡放箱; 改进YOLOv8; 级联结构网络

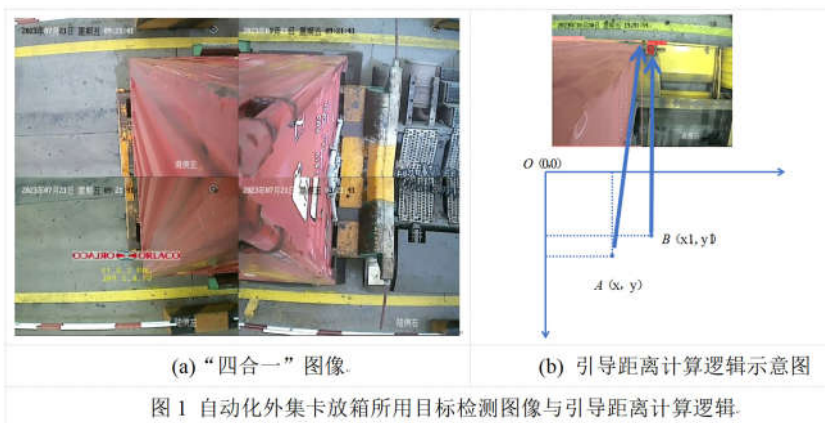
0 引言

在港口和码头作业中, 场桥是关键装卸设备, 特别是在外集卡放置箱子的过程中, 如何准确高效地完成操作一直是技术难点, 传统的集装箱吊放作业依赖大量人为操作, 这不仅增加了劳动成本, 且由于对环境的视觉信息提取有限, 缺乏多角度、多视点的全面信息, 导致作业效率低下和误操作。

近年来, 随着计算机视觉技术在各个领域取得的飞速发展, 将目标检测算法与场桥自动化改造相结合的应用方向也随之日益成熟, 在集卡引导、箱门定位、抓放箱以及场桥大车防撞等功能得到了广泛应用。在自动化外集卡放箱系统中, 目标检测结果的准确性直接影响最终的放箱性能, 在该场景下, 目标检测结果的准确性主要体现在目标检测精度和边框回归精度上, 由于放箱过程中图像运动模糊和复杂的图像背景, 导致现有目标检测模型存在准确性不足的问题。针对该问题, 本文采用了YOLOv8模型^[1]作为基准模型进行放箱目标检测, 并对其进行改进, 从而实现更加稳定高效的自动化外集卡放箱。

1 自动化外集卡放箱技术分析

基于计算机视觉的自动化外集卡放箱技术主要包括目标检测模块、引导距离计算模块和像素映射模块三个计算模块。目标检测模块通过吊具中锁部位安装的4个视觉相机采集图像并按一定尺寸裁剪拼接成“四合一”图像, 如图1(a)所示, 图像利用目标检测模型获得目标类型和目标检测框; 引导距离计算模块根据检测得到的目标类型和目标检测框位置制定对应的引导距离计算规则, 如图1(b)所示, 利用箱体目标检测框右上角坐标与锁头目标检测框的右上角坐标计算引导距离; 最终, 像素映射模块对图像中计算得到的引导距离通过像素映射成实际大车小车需要移动的距离。在自动化放箱过程中, 计算模块实时采集图像进行推理, 最终将推理结果发送给PLC执行操作, 可见, 自动化外集卡放箱技术一方面需要检测模型具有较快的检测速度以保证实时性, 另一方面, 自动化外集卡放箱技术必须有高放箱成功率以确保安全性和可靠性, 而这十分依赖目标检测模块输出结果具有较高的召回率和检测框精度。



2 改进YOLOv8模型

YOLOv8模型是来自Ultralytics的YOLO对象检测模型系列，能提供统一框架，以支持目标检测、实例分割、图像分类的操作，YOLOv8模型具有最快的检测速度，适宜应用在需要实时快速检测的外集卡放箱场景中。常规YOLOv8模型通常先利用主干网络（Backbone）对图像进行特征提取，再通过特征金字塔网络（Neck）融合并输出多尺度特征，最终特征图经过检测头网络内的类别检测分支和边框回归分支进行特征映射得到目标类别和检测框。然而，在放箱场景下，用于检测的图像不仅背景

复杂而且存在运动模糊问题，这导致常规YOLOv8模型在检测锁孔目标时容易漏检，且目标检测框的位置不够精确。

基于此，为了提高放箱场景下的实时检测精度，对YOLOv8模型结构进行改进，在保持主干网络和特征金字塔网络不变的同时，首先设计一种类别无关的感兴趣区域检测分支替换检测头网络中原有的类别检测分支，其次增加级联结构网络处理感兴趣目标，最终进行特征映射输出精确的目标类别和检测框，改进YOLOv8模型整体结构如图2所示。

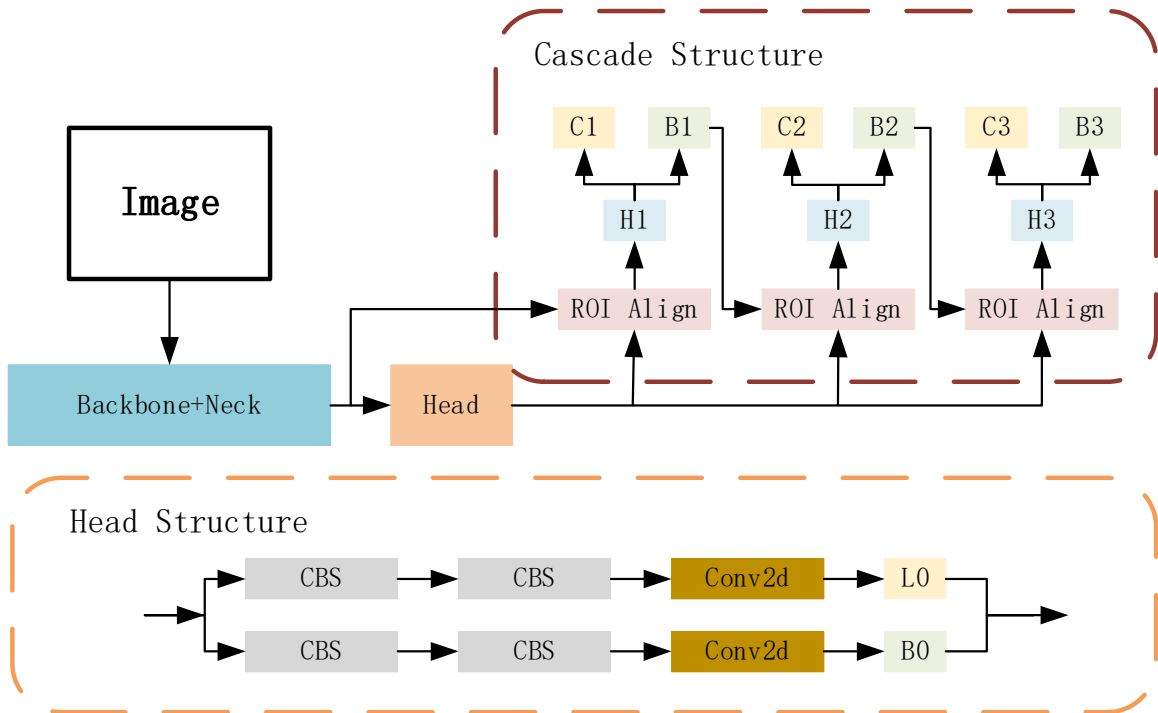


图2改进YOLOv8模型结构

2.1 感兴趣区域检测分支

改进后的检测头网络中，边框回归分支（B0）不变，仍输出检测框数据。而将类别检测分支替换为类别无关的感兴趣区域检测分支（L0），在该分支下，特征图经过特征映射输出目标前景概率，而无需输出具体类别概率，只需区分前景和背景。随后，将前景概率大于阈值的目标作为感兴趣区域输入后续网络中进行具体类别分类。

感兴趣区域检测分支将目标分类拆分成两个阶段进行，一方面可以尽可能多的检测到感兴趣的前景区域以减少目标漏检率，另一方面，极大地降低了检测头网络的训练难度，有利于提高整体模型的收敛性能。

2.2 级联结构网络

受启发于级联高精度检测与YOLOv8改进思路^{[2][3][4]}，本文引入级联结构网络处理检测头网络输出的前景目标，

级联结构网络逐层级地对检测目标进行类别再检测和边框精确，一共优化三次。在每一级网络中，首先利用ROIAlign网络层对前景目标所在感兴趣特征图区域进行下采样，其次通过特征映射层($H_{i,i=1,2,3}$)投影得到目标类别($C_{i,i=1,2,3}$)和目标检测框($B_{i,i=1,2,3}$)，若还有下一级，则将目标检测框所在位置得到的感兴趣特征图区域输入到下一层级的ROIAlign网络层，若无下一级，则将当前检测结果作为最终的检测结果。

不同于常规YOLOv8训练时仅采用0.5作为检测框的回归阈值，本文所提YOLOv8模型在级联结构网络不同层级输出的目标检测框分别采用0.5、0.6、0.7作为回归阈值，使得模型在训练时获得逐级精确检测框的能力。

2.3 重构类别损失函数

本文根据所提YOLOv8的模型结构，重新构建了模型

训练的类别损失函数 L_{cls} ，具体公式如下：

$$L_{cls} = -\log[p^*P(C)+(1-p^*)(1-P(C))] \quad (1)$$

其中， \log 表示对数函数， p^* 在检测框内存在目标时为1否则为0， $P(C)$ 表示最终类别预测概率，可通过下式计算：

$$P(C) = P(C|0=1)/P(0=1) \quad (2)$$

其中， $P(0=1)$ 表示感兴趣区域检测分支输出的目标前景概率， $P(C|0=1)$ 表示级联结构网络最后一级输出的目标类别概率。

应用成果

采用本文所提改进YOLOv8模型优化后，模型对于锁头、锁孔等尺寸小易漏检的目标检测能力有显著提升，而对于车板边沿、箱体、邻箱等目标则能输出更精确的检测边框。得益于改进YOLOv8模型较高的目标召回率和检测边框精确度，使后续推理模块可以计算出更精确的放箱引导距离，减少完成放箱过程所需的模型检测轮

次，进而提高场桥外集卡自动化放箱的作业效率和成功率，优化后作业效率高于21自然箱/h，人工介入率降低至3%。

结束语

本文针对现有技术的召回率低和检测边框不精确的问题，提出了一种基于级联边框精确的YOLOv8外集卡放箱方法，主要包括设计一个类别无关的感兴趣区域检测分支，并引入级联结构网络来处理感兴趣目标。同时，还构造了对应的类别损失函数。通过实际效果验证，改进后的模型有效提升了放箱成功率和作业效率。

参考文献

- [1]YaseenM.What isYOLOv8:An in-depth exploration of the internal features of the next-generation object detector. arXiv2024[J].arXivpreprintarXiv:2408.15857.
- [2]程维龙,僧冰枫,刘晓静.基于优化YOLOv8的唐卡图像目标检测方法[J].软件导刊,2025,24(1):186-191.
- [3]张宁,周云翀,徐坤财,等.基于MobileNetV4改进的YOLOv8目标检测算法研究[J].集成技术,2026,15(2):91-103
- [4]吴林,曹雯.YOLO-SCDI:基于改进YOLOv8的车辆检测算法[J].电子与信息学报,2025,47(6):2312-2321.