

基于改进 YOLO v5s 的甘蔗切种茎节特征识别定位技术

李尚平^{1,2} 郑创锐^{1,2} 文春明^{1,2} 李凯华^{1,2} 甘伟光^{1,2} 李 洋^{1,2}

(1. 广西民族大学电子信息学院, 南宁 530006; 2. 广西高校智慧无人系统与智能装备重点实验室, 南宁 530006)

摘要: 为了实现甘蔗智能横向切种工作站的精准、高效的自动化切种, 针对工厂化切种任务的特点, 提出了一种基于改进 YOLO v5s 的甘蔗茎节特征边缘端识别与定位方法。首先, 利用张正友相机标定法对摄像头进行畸变矫正; 然后对甘蔗茎节数据集进行数据增强, 利用原始的 YOLO v5s 模型进行训练和测试, 结果显示数据增强能一定程度上提高检测精度。针对茎节特征目标小以及模型体积大导致检测精度低、部署难度高等问题, 对 YOLO v5s 的骨干网络进行改进, 在 SPPF 特征融合模块前引入坐标注意力 (Coordinate attention, CA) 模块和 Ghost 轻量化结构, 在 Head 网络中剔除 P5 大目标检测头, 得到了改进后甘蔗茎节检测模型 YOLO v5s - CA - BackboneGhost - p34, 测试结果表明该模型优于其他主流算法和原始模型, 具有高精度、小体积等优势。其中, 平均精度均值 1 和平均精度均值 2 分别提高 5.2、16.5 个百分点, 模型浮点数计算量和内存占用量分别降低 42% 和 51%。最后, 为了提高检测速度和实时性, 将模型部署于边缘端, 利用 TensorRT 技术加快检测速度, 并在传送速度为 0.15 m/s 的甘蔗智能横向切种工作站上完成实际切种实验。实验结果表明, 加速后茎节检测速度达到 95 f/s, 实时检测定位平均误差约为 2.4 mm, 切种合格率为 100%, 漏检率 0.4%, 说明本文提出的模型具有高度可靠性和实用性, 可以为甘蔗横向切种工作站的工厂化、智能化以及标准化应用提供有效的技术支持。

关键词: 甘蔗智能横向切种工作站; 茎节识别与定位; YOLO v5s; 边缘端部署

中图分类号: S24; TP39 **文献标识码:** A **文章编号:** 1000-1298(2023)10-0234-12

OSID:



Stem Node Feature Recognition and Positioning Technology for Transverse Cutting of Sugarcane Based on Improved YOLO v5s

LI Shangping^{1,2} ZHENG Chuangrui^{1,2} WEN Chunming^{1,2} LI Kaihua^{1,2} GAN Weiguang^{1,2} LI Yang^{1,2}

(1. College of Electronic Information, Guangxi Minzu University, Nanning 530006, China

2. Key Laboratory of Intelligent Unmanned System and Intelligent Equipment, Nanning 530006, China)

Abstract: In order to achieve accurate and efficient automated seed cutting in sugarcane intelligent transverse seed cutting workstation, a method based on improved YOLO v5s for identifying and locating the edge end of sugarcane stem node features was proposed for the characteristics of factory seed cutting tasks. Firstly, the camera was corrected for distortion by using the ZHANG Zhengyou camera calibration method, then the sugarcane stem node dataset was enhanced and the original YOLO v5s model was used for training and testing, and the results showed that the data enhancement can improve the detection accuracy to some extent. Then, to address the problems of low accuracy and high model complexity caused by small stem node feature targets, the backbone network of YOLO v5s was improved by introducing the coordinate attention module and Ghost lightweight structure before the SPPF module, and removing the P5 large target detection head in the Head network to obtain the improved sugarcane stem node detection model YOLO v5s - CA - BackboneGhost - p34. The test results showed that the model outperformed other mainstream algorithms and the original model with high accuracy and small size. Among them, mAP@ 0.5 and mAP@ 0.5 : 0.95 were improved by 5.2 and 16.5 percentage points, respectively, and the model computation and size were reduced by 42% and 51%, respectively. Finally, in order to improve the detection speed and real-time performance, the model was deployed at the edge end, and the detection speed was accelerated by using TensorRT technology, and the model was completed on a sugarcane with transmission speed of 0.15 m/s. The actual seed cutting test were

completed on the smart transverse seed cutting workstation with transmission speed of 0.15 m/s. The test results showed that the accelerated stem node detection speed reached 95 f/s, the average error of real-time detection and positioning was about 2.4 mm, the seed cutting qualification rate was 100%, and the leakage rate was 0.4%, which indicated that the model proposed was highly reliable and practical, and can provide effective technical support for the industrialization, intelligence and standardization of sugarcane transverse seed cutting workstation.

Key words: intelligent transverse sugarcane cutting workstation; stem node identification and positioning; YOLO v5s; edge-end deployment

0 引言

甘蔗是我国最重要的糖料作物,在我国国民经济中占有特殊的重要作用^[1]。甘蔗种业是糖料甘蔗生产的基础,是促进甘蔗高产高糖的前提。甘蔗制种的工厂化、标准化以及智能化是促进糖业发展必由之路。

针对甘蔗制种的工厂化、标准化以及智能化生产技术的不足,课题组研发了甘蔗智能横向切种工作站^[2],它具有切种合格率高、伤芽率低等优点,为甘蔗的工厂化制种奠定了良好的基础。本文以提高甘蔗智能横向切种工作站的效率及精度为目标,对其甘蔗茎节识别部分进行优化改进以及部署于边缘设备上,进一步提高甘蔗智能横向切种工作站的实用性与可靠性。

目前甘蔗茎节识别与定位方法相关的研究可以分为传统机器视觉方法、传感器识别分析方法和深度学习方法。在传统机器视觉方法方面,CHEN 等^[3]在甘蔗的感兴趣区域二值化图像内使用垂直投影函数,并对垂直投影函数连续求导后,根据可疑像素值确定茎节,单个茎节点识别准确率为 100%。ZHOU 等^[4]提出了一种基于 Sobel 边缘检测的甘蔗茎节点识别方法,识别准确率为 93%,平均时间为 0.539 s。张圆圆等^[5]利用甘蔗茎节处具有拐点和灰度值不连续的特性,在边缘拟合法和灰度值拟合法的基础上使用中值决策法对甘蔗茎节进行识别,识别率达到 94.7%。在传感器识别分析方法方面,CHEN 等^[6]利用加速度传感器和薄膜压电传感器共同采集位置信息,基于小波分析对甘蔗茎节进行识别,识别精度为 99.63%,单个茎节识别时间为 0.25 s。MENG 等^[7]利用激光传感器获取甘蔗表面轮廓信号,并提出了一种基于多阈值多尺度小波变换的甘蔗茎节识别分析方法,识别率为 100%,识别时间为 0.25 s。在深度学习方法方面,陈延祥^[8]在改进后的 YOLO v3 的基础上使用边缘提取算法和茎节定位算法,识别调和平均值为 97.3%。赵文博等^[9]在 YOLO v5s 的基础上引入 bifpn、ElIoU 损失函数和 Focal loss 损失函数提高识别平均精度,然后在颈部

引入 Ghost 模块使模型轻量化,平均识别精度为 97.8%,模型内存占用量为 11.4 MB。而在甘蔗智能横向切种工作站上使用的识别模型已有两代。第 1 代识别模型为:廖义奎等^[10]利用垂直投影方法确定甘蔗所在位置,然后利用卷积神经网络 AlexNet 搭建甘蔗蔗芽特征识别模型,模型实时动态检测精度为 88%,平均检测时间为 341.09 ms;第 2 代识别模型为:李尚平等^[11]通过改进 YOLO v3 算法实现了甘蔗茎节实时动态识别,识别平均精度为 90.38%,识别平均时间为 28.7 ms,模型内存占用量为 118 MB。

传统机器视觉和传感器分析方法对于甘蔗茎节识别的研究大多数为静态或纵向识别,对场景和目标特征选择的依赖性较强,面对颜色分布差异较大或者外观形状不规律的甘蔗种时,识别能力较差,鲁棒性较低,效率低下,并不能满足工厂化生产的需求。而其它基于深度学习方法的甘蔗茎节识别模型虽然展现出较高的应用前景,但均未部署在嵌入式设备上,缺少实际切种任务的检验,真实效果有待检验,缺乏可靠性和真实性。

因此,本文针对甘蔗智能横向切种工作站应用需求,提出一种基于改进 YOLO v5s 的甘蔗茎节特征识别定位方法以及边缘端部署,优化甘蔗茎节实时检测流程,使其在复杂的光照条件和实时检测的动态环境下,克服特征模糊的影响,提高视觉系统的检测精度及可靠性,并将视觉系统部署于边缘设备上,使其具有更好的实用性与可扩展性,通过在切种工作站上进行实际切种试验,验证本文方法的优越性。

1 甘蔗智能横向切种工作站构建

1.1 甘蔗智能横向切种工作站组成与工作原理

甘蔗智能横向切种工作站由二级耙结构、图像采集黑箱、切种平台、液压站、甘蔗分拣电机、甘蔗传送电机、摄像头、调节灯、光电传感器、图像识别及控制系统等零部件组成,如图 1 所示。切种过程中,随着二级耙的转动,整根蔗种被横向有序地送入黑箱中,黑箱顶部的摄像头实时采集甘蔗图像序列,接着

由图像识别系统同时检测各茎节位置和计算各切口坐标信息,并将其发送给控制器,由控制器调控多把切刀,同时完成一根甘蔗的双芽段切种工作。

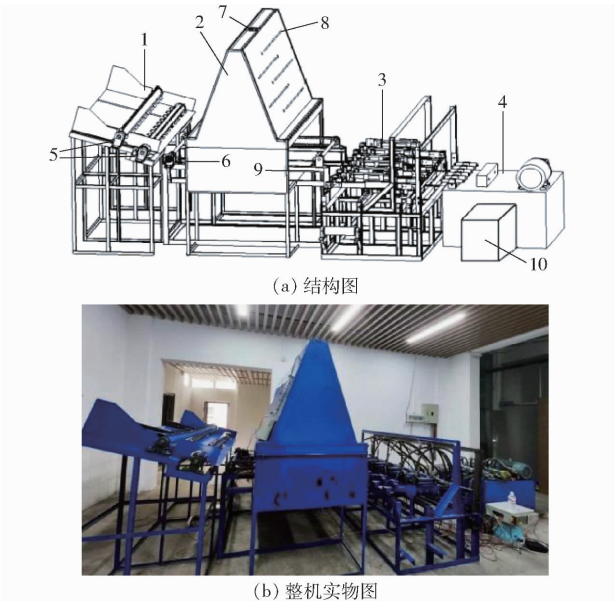


图 1 甘蔗智能横向切种工作站

Fig. 1 Intelligent transverse sugarcane cutting workstation
1. 二级耙结构 2. 图像采集黑箱 3. 切种平台 4. 液压站 5. 甘蔗分拣电机 6. 甘蔗传送电机 7. 摄像头 8. 调节灯 9. 光电传感器 10. 图像识别及控制系统

1.2 图像采集黑箱结构

图像的采集工作主要在黑箱中完成,黑箱中包括顶部摄像头和侧边的矩阵灯带,如图 2 所示。其中,摄像头与传送带平面垂直距离为 1.2 m,该距离的设置可以保证摄像头能够拍摄到传送带上的完整物体,避免甘蔗茎节信息缺失等问题。

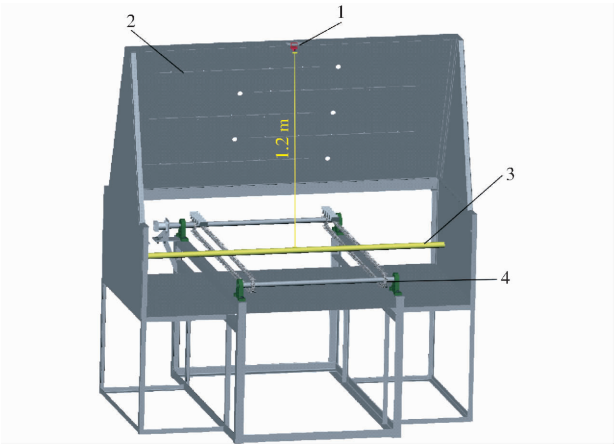


图 2 甘蔗智能横向切种工作站黑箱内部结构图

Fig. 2 Internal structure of black box of intelligent transverse sugarcane cutting workstation
1. 摄像头 2. 调节式矩阵灯带 3. 整根甘蔗 4. 传送链

为了得到最佳的图像茎节特征,需要调节 LED 矩阵灯带的光照分布。经过多次试验,目前平均光照强度为 430.7 lx 时效果最佳,可充分展示甘蔗茎

节的特征。
在选用摄像头时,考虑到实际应用需求和成本限制,使用了 RMONCAM 的 G200 型高清摄像头。该工业相机采集速度为 30 f/s,分辨率为 1 920 像素 × 1 080 像素,曝光方式为卷帘曝光,曝光时间约为 33 ms,图像处理方式为自动曝光。这些参数的设置可以保证图像的清晰度和稳定性,同时满足目标检测的实时性要求。

1.3 相机标定

为了消除图像识别定位的畸变误差影响,需对工业相机进行精准的标定,实现相机坐标和实际坐标之间的转换。在计算机上生成 7 × 9 的棋盘方格,每个方格的尺寸为 2.5 cm × 2.5 cm,如图 3a 所示。利用打印机进行 1:1 打印,再从不同角度对棋盘方格进行拍照,采集 50 幅不同角度的图像,如图 3b、3c 所示。

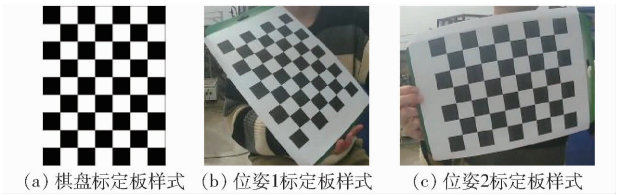


图 3 工业相机标定流程

Fig. 3 Industrial camera calibration process
采用图像处理算法对相机进行标定处理,通过张正友相机标定算法进行计算^[12],可以得到相机内部参数 R_1 和畸变参数 R_2 表达式为

$$R_1 = \begin{bmatrix} 518.306\ 5 & 0 & 309.727\ 6 \\ 0 & 517.915\ 3 & 249.605\ 6 \\ 0 & 0 & 1 \end{bmatrix} \quad (1)$$

$$R_2 = \begin{bmatrix} -0.396\ 5 & 0.158\ 3 & 0 & 0 & 0 \end{bmatrix} \quad (2)$$

2 甘蔗茎节图像数据集采集与预处理

2.1 甘蔗茎节的数据采集

本研究以广西甘蔗为研究对象,在崇左市扶绥县的广西大学广西亚热带农科新城进行甘蔗的采集^[13-14]。为了保证数据的多样性和鲁棒性,从 2022 年 10 月到 2023 年 3 月,通过人工砍收的方式分 3 次收集 404 根甘蔗,甘蔗品种包括中蔗 9 号、桂糖 42 号和桂辐 98-296。平均长度约为 1.8 m,平均直径达到 30 mm,其中保留 20 根作为实际切种的试验样本,其余甘蔗置于切种工作站的传送链上,传送速度分别从 0.1 m/s 逐渐升高至 0.15 m/s,并使用摄像头对整杆甘蔗的茎节特征进行视频采集。通过视频帧分割方式,最终采集到共计 2 336 幅整根甘蔗图像,格式为 JPG。采集到的图像包含不同摆放密度、不同光照、多品种的蔗种图像,以保证目标

检测的准确性与鲁棒性。如图 4 所示。

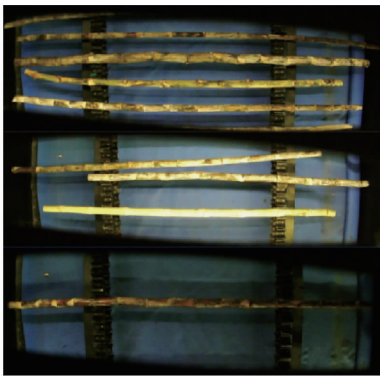


图 4 整杆甘蔗的茎节图像

Fig.4 Image of stem nodes of whole sugarcane

2.2 数据筛选以及数据标注

为提高实时动态环境下的甘蔗茎节识别的精度和速度,因此在筛选过程中,首先从传送速度和曝光度等因素出发,删除一些不符合要求的图像。其次,为了提高模型鲁棒性,着重选择一些模糊但仍具有茎节特征的图像数据。最终保留 449 幅不同疏密、不同光照的甘蔗茎节图像,便于 YOLO 目标检测算法更好地去学习标注好的甘蔗茎节目标细节特征,提高整体模型识别精度。

在数据标注方面,本文使用 LabelImg 软件对这些图像进行了标注,并将标注信息转换为 YOLO 需要的 txt 格式。在标注过程中,得到包含茎节目标的中心坐标 (x',y') 、宽、高信息的 xml 文件。由于 YOLO 需要的标注文件类型为 txt 文件,因此使用 Python 编程将 xml 文件转化为 txt 格式的标注文件,得到了可以在 YOLO 模型中运行的数据集^[15-18]。YOLO 数据格式如图 5 所示。

class	x	y	w	h
-------	---	---	---	---

图 5 YOLO 数据格式

Fig.5 YOLO data format

图 5 中 class 表示数值化的类别标签,标签从 0 开始索引,就本研究而言,仅有一个类别标签,所以 class 值仅为 0; x,y,w,h 的计算式为

$$\begin{cases} x = \frac{x'}{W} \\ y = \frac{y'}{H} \end{cases} \tag{3}$$

$$\begin{cases} w = \frac{w'}{W} \\ h = \frac{h'}{H} \end{cases} \tag{4}$$

式中 W ——图像宽度 H ——图像高度
 w ——归一化后目标宽度
 h ——归一化后目标高度

x ——归一化后目标中心 x 坐标
 y ——归一化后目标中心 y 坐标
 w' ——目标宽度 h' ——目标高度

图 6 为甘蔗茎节边界框标注示意图。图中的字符与式(3)、(4)对应。图像坐标系的原点位于图像的左上角, x 轴方向沿着图像水平向右, y 轴方向沿着图像竖直向下。

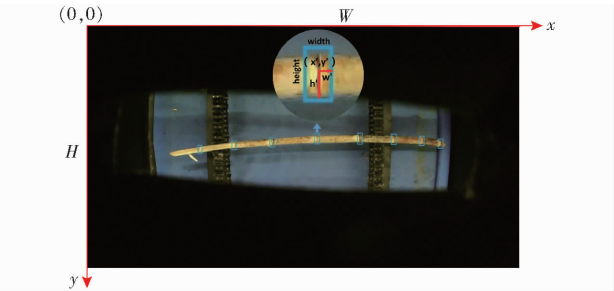


图 6 甘蔗茎节数据标注示意图

Fig.6 Schematic of sugarcane stem node data labeling

通过对标注的甘蔗茎节边界框数据进行统计,449 幅图像经过标注后得到 8 102 个甘蔗茎节实例,因为 1 幅图像可能存在多根甘蔗、多个茎节,所以得到的实例数量较多。接着对图中边界框坐标中心 (x,y) 和边界框的宽、高进行了归一化处理。统计结果如图 7 所示,图中颜色较深的区域表示数据高度重叠。

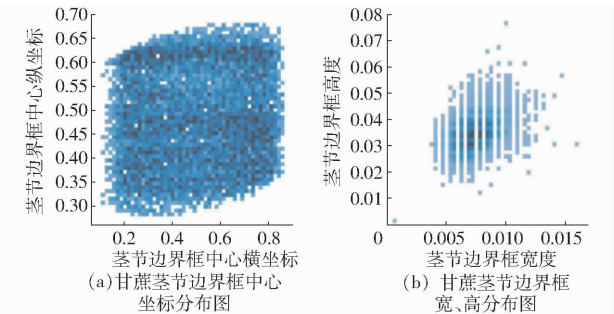


图 7 甘蔗茎节边界框统计结果

Fig.7 Statistical results of box of sugarcane stem node boundary

从图 7a 可知,甘蔗茎节的边界框的中心 (x,y) 主要分布在 y 轴 $0.3 \sim 0.7$ 之间以及 x 轴 $0.2 \sim 0.8$ 之间,这是因为在安装摄像头时,在其外部增加了一个黑色外壳,用于减弱光线的干扰,所以在识别画面出现四周黑边的情况,从而影响了甘蔗茎节边界框的分布。

此外,根据图 7b 可知,每个甘蔗茎节实例的边界框面积都不超过整幅图像区域的 0.12%,总体呈现小目标的特点,这也是在后续算法改进中需要解决的一个重点问题。

2.3 数据增强

采用 Imgaug 图像增强库扩大甘蔗茎节图像数

量,记录数据增强前后的训练、测试效果对比。本实验选择 YOLO v5s 模型作为基准模型,并将原始数据随机抽样出约 10% 数据作为测试集,共 50 幅图像。训练集、验证集和测试集的划分比例为 8:1:1,接着分别对训练集和验证集采用仿射变换、翻转变换、高斯模糊、高斯噪声等方式扩增数据集 7 倍,测试集采用原始数据集分割得到的测试集,数据划分结果如表 1 所示。

表 1 数据集划分
Tab.1 Division of data set

数据集类型	训练集	验证集	测试集	总数
原始数据集	349	50	50	449
Imgaug 数据增强后	2 443	350	50	2 843

分析数据增强前后数据集在每一轮训练后在验证集上的效果,如图 8 所示(图中平均精度均值 1 指 mAP@ 0.5,平均精度均值 2 指 mAP@ 0.5:0.95,YOLO v5s_aug 指在训练 YOLO v5s 模型时使用数据增强后的数据集,YOLO v5s_noaug 指在训练 YOLO v5s 模型时使用原始数据集)。从结果上可以看出,数据增强后的模型在精确率(Precision)、召回率(Recall)和平均精度均值(Mean average precision, mAP)上都有不同程度的提升,并且震荡幅度更小,这是由于数据增强丰富了数据集的多样性,使得模型能够学习到更多的数据分布和场景,从而提高模型的泛化能力和鲁棒性,降低了过拟合风险。

图 9 为 500 轮训练过程中在训练集和验证集下的 box 损失值以及 object 损失值,box 损失值描述预

测框与锚框之间的匹配程度,object 损失值描述预测物体类别与真实类别的差异程度。从图中可以看出,使用数据增强后的 box 损失值越小,object 损失值也相对较低,说明经过数据增强后模型收敛更快。

从图 9 中还可以看出,在前 100 轮的训练以及验证过程中,目标损失出现先上升后下降的趋势,其原因是在训练初期过程中,模型只能学习到一些基本的特征,导致目标损失函数的值较高。随着训练迭代次数的增加,模型会学习到更多的特征,从而降低目标损失函数的值。

数据增强前后的模型在测试集上的检测性能如表 2 所示。从表 2 可以看出,在原始数据集基础上进行数据增强使精确率、召回率、平均精度均值 1、平均精度均值 2 分别提高 3.8、5.2、5.4、8.5 个百分点,其中平均精度均值 2 提升最明显。这是因为数据增强扩展了原始数据集,增加了样本数量和多样性,使得模型能够学习到更多丰富的甘蔗茎节特征,一定程度上提高了模型的检测精度、稳定性和鲁棒性。

YOLO v5s_aug 模型性能参数和损失值如图 10 所示。从图中可以发现,采用数据增强策略能显著且稳定提高模型检测精度和稳定性,加快收敛速度。

3 基于改进 YOLO v5s 的甘蔗茎节识别模型

3.1 YOLO v5s 算法结构优化

YOLO v5s 是 YOLO v5 中模型复杂度较低,且兼具较高检测精度和速度的模型,被广泛应用于各

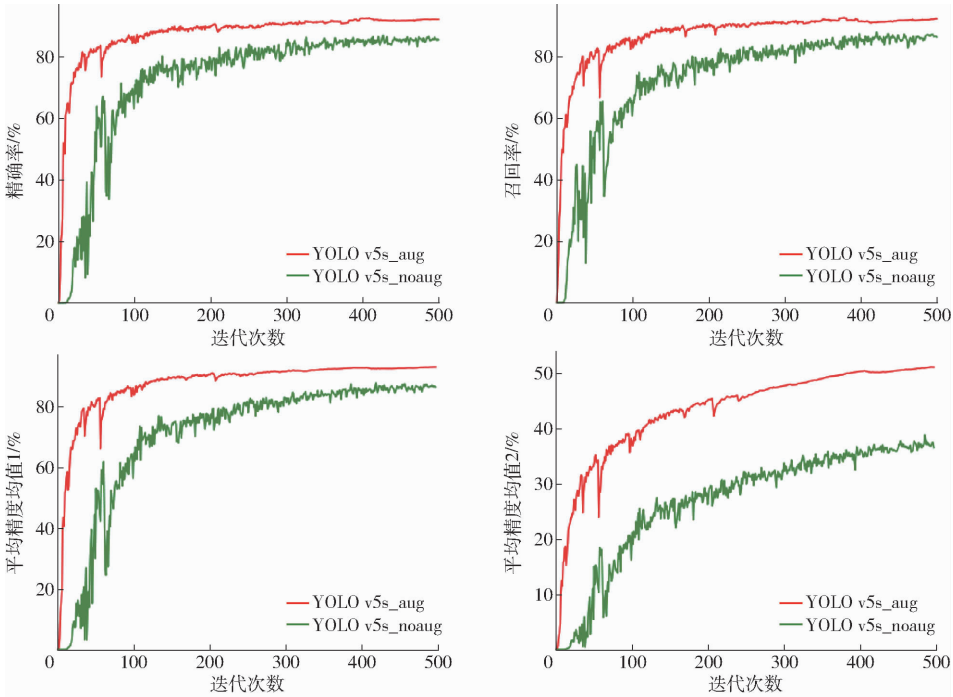


图 8 数据增强前后数据集训练过程验证集性能评估结果

Fig.8 Performance evaluation results of validation set before and after data enhancement for dataset training process

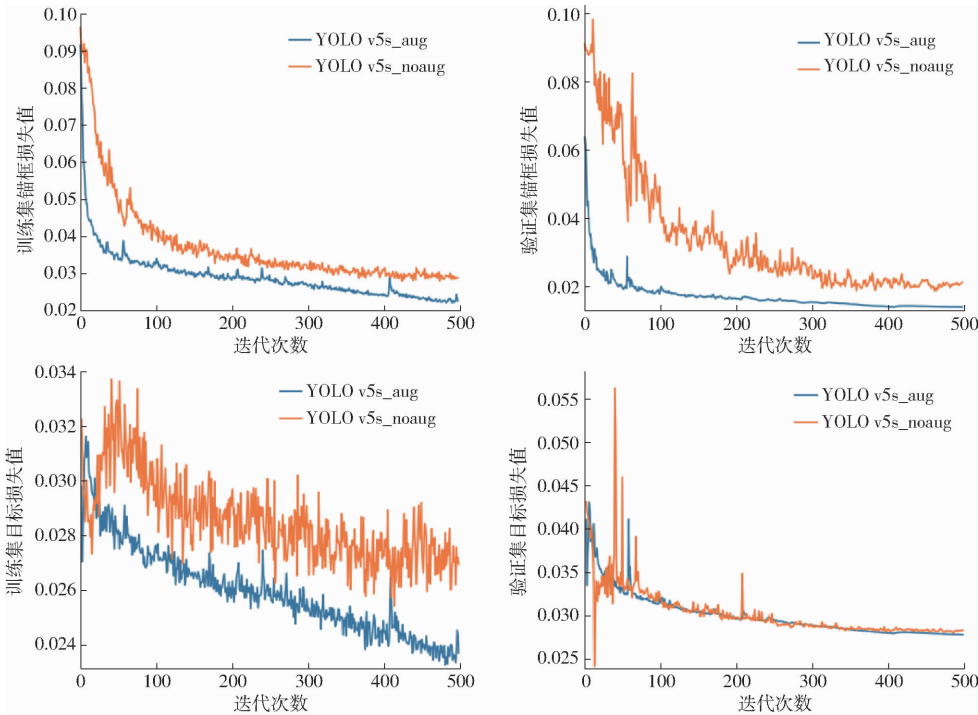


图 9 数据增强前后数据集训练过程损失值变化曲线

Fig. 9 Curves of loss value during training process of dataset before and after data augmentation

表 2 数据增强前后数据集训练所得模型在测试集上性能参数

Tab.2 Performance evaluation results of trained model on test set before and after data augmentation of dataset %

数据集类型	精确率	召回率	平均精度均值 1	平均精度均值 2
	率	率		
原始数据集	88. 7	85. 8	87. 3	44. 7
数据增强后数据集	92. 5	91. 0	92. 7	53. 2

类工业生产活动。将它作为网络优化的基础,可以保证网络结构的可靠性和高效性,避免过多的参数冗余对实际应用造成不良影响。所以,本研究将以 YOLO v5s 算法为基础,对骨干网络和 Head 网络进行优化改进,使其对甘蔗茎节小目标的检测性能提高,模型复杂度降低。

3. 1. 1 骨干网络优化

骨干网络是 YOLO v5s 模型的基础,负责提取

输入图像的特征,它的模块主要包括 Conv 模块、C3 模块以及 SPPF 模块。Conv 模块实现了对输入特征的提取和转化;C3 模块利用残差网络进行残差学习,提高了模型的深度;SPPF 模块是在 SPP 的基础上减少网络层数,并对不同感受野的信息特征进行融合,提高特征图的表达能力和特征融合的速度。但是,原始网络对小目标特征的提取能力有限,提取的特征图中也包含了许多冗余特征信息,增加了计算量和内存占用量。所以,为了进一步提高 YOLO v5s 的特征提取能力以及降低模型复杂度,本文将对它的骨干网络进行优化改进。图 11 为 YOLO v5s 模型的原始骨干网络以及改进后的骨干网络结构示意图。从图 11b 中可以看出,原始骨干网络的 SPPF 层前面引入 CA 模块,并将 Conv 模块和 C3 模块换成轻量级 GhostConv 模块和 C3Ghost 模块,最后将输出结果同 P3、P4 和 P5 特征共同输入至下一级

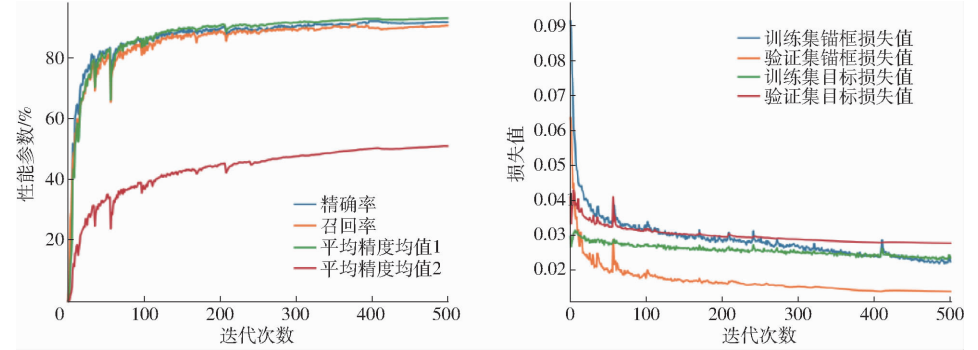


图 10 YOLO v5s_aug 模型性能参数和损失值变化曲线

Fig. 10 Performance parameters and loss curves variations of YOLO v5s_aug model

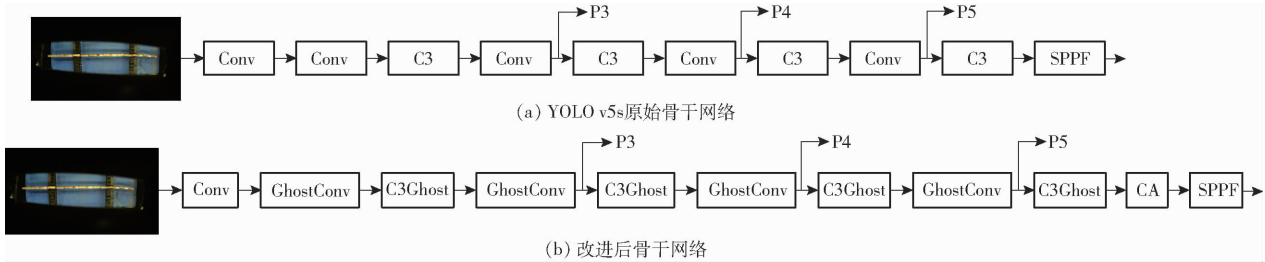


图 11 改进前后的骨干网络对比

Fig. 11 Comparison of backbone networks before and after improvement

Neck 网络。

CA 注意力模块的结构图如图 12 所示,从图中可以发现,CA 注意力模块是通过水平方向和垂直方向上分别进行平均池化,再使用转换器对空间信息进行编码,最后把空间信息通过加权的方式融合进通道中,这有利于 CA 注意力机制全面考虑空间信息和通道信息,增强特征提取能力^[19-20]。

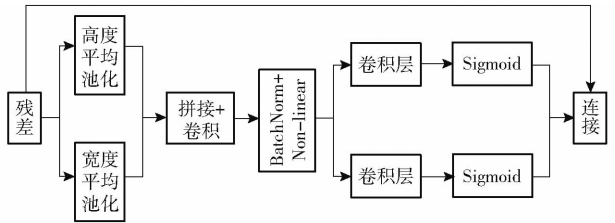


图 12 CA 注意力模块
Fig. 12 CA attention module

Ghost 结构^[21]由 Ghost 卷积模块和 Ghost 瓶颈模块组成。Ghost 卷积模块包括常规卷积和线性变换两部分(图 13a),首先通过有限的常规卷积得到一部分特征图,之后利用廉价的线性变换生成更多 Ghost 特征图,生成的 Ghost 特征图能够极大的表现常规卷积中包含的冗余特征信息,最后通过恒等映射将两组特征图进行组合。Ghost 瓶颈模块(图 13b)是由两个 Ghost 卷积模块构成的,其中一个 Ghost 卷积作为扩展层,用于增加特征维度,扩张通道数;另一个 Ghost 卷积用于减少通道数使其与直连的特征相匹配,最后经过 shortcut 连接后输出特征。

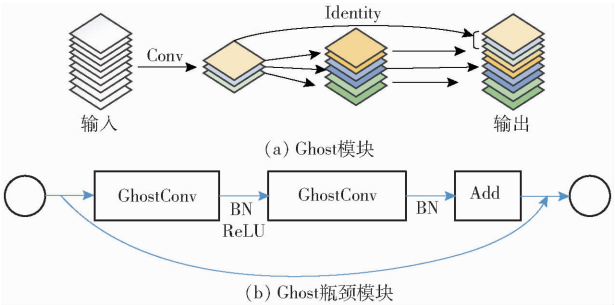


图 13 Ghost 结构图
Fig. 13 Ghost structure diagram

相比于原始骨干网络,在 SPPF 层前加入 CA 注意力模块,不仅可以加强特征的表现力,使网络更加

关注重要的特征,而且可以在多尺度提取不受影响的情况下,保证特征图精细程度,使得模型更加稳定和准确。这一点可以帮助骨干网络有意识地提取出属于甘蔗茎节的特征点,从而实现更好的检测效果以及更高的检测精度。

同时,引入 Ghost 网络结构的骨干网络可以在不影响模型特征提取能力的情况下,降低模型计算量和参数量,减少模型的特征冗余。

3.1.2 Head 网络的优化

Head 网络通常接在 Neck 网络之后,主要负责预测目标检测任务中的类别、位置、置信度等信息。它的输入特征分别选取下采样为 32、16 和 8 的卷积输出结果,即分别对应图中的 P5、P4 和 P3,在输入图像尺寸为 640×640 的情况下,分别用于检测大小在 32×32 以上的大目标、16×16 以上的中目标以及 8×8 以上的小目标。然而在甘蔗茎节数据集的目标面积占比均小于 0.12% 的任务中,Head 网络的大目标检测分支不适用于此目标检测任务。因此本文针对 Head 输出网络进行了如下优化:将输出的大目标检测分支剔除,只保留小目标和中目标检测,这有利于小目标物体的位置信息检测,降低模型后处理阶段的容错率,使模型更加专注于训练有用的特征,减少了网络计算量和参数量,加快模型的推理速度,提高模型精度。其优化结果如图 14。

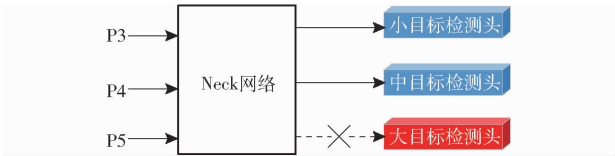


图 14 Head 网络优化示意图
Fig. 14 Illustration of Head network optimization

3.2 算法优化实验对比

训练数据集均统一采用数据增强后的数据,测试集采用原始数据 10% 的数据,表 3 为不同模型及其对应说明。

3.2.1 算法优化改进对比

不同 YOLO v5 改进算法在 500 轮训练过程中的性能指标如图 15 所示。

表 3 模型名称及其说明

Tab.3 Model names and their descriptions

模型名称	说明
YOLO v5s	YOLO v5s 模型
YOLO v5s - CA	骨干网络添加 CA 注意力机制模块
YOLO v5s - p34	Head 网络剔除 P5 检测头
YOLO v5s - CA - p34	骨干网络添加 CA 注意力机制模块，Head 网络剔除 P5 检测头
YOLO v5s - CA - BackboneGhost - p34	骨干网络添加 CA 注意力机制模块，并且引入 Ghost 结构,Head 网络剔除 P5 检测头
YOLO v5s - CA - NeckGhost - p34	骨干网络添加 CA 注意力机制模块，Neck 网络引入 Ghost 结构,Head 网络剔除 P5 检测头
YOLO v5s - CA - Ghost - p34	骨干网络添加 CA 注意力机制模块，骨干网络和 Neck 网络引入 Ghost 结构，Head 网络剔除 P5 检测头

从图 15 可得,单独使用 CA 注意力机制或者单独剔除大目标检测头都能显著提升模型性能,其中

后者的优化策略效果更佳。而同时使用 CA 注意力机制与剔除大目标检测头的优化策略,可获得最佳性能表现。具体来说,单独或者结合使用这两种优化策略对模型的精确率、召回率、平均精度均值 1 和平均精度均值 2 都有提升,在精度、召回率和平均精度均值 1 上差距较小,但在平均精度均值 2 上提升效果最为明显。

在 YOLO v5s - CA - p34 的基础上,对其骨干网络、Neck 网络引入 Ghost 结构,发现在骨干网络或 Neck 网络中单独引入 Ghost 结构的模型在平均精度均值等性能指标上都略微高于同时引入 Ghost 结构的模型,特别是在平均精度均值 2 上的差异最大。而相比 YOLO v5s - CA - p34 模型,3 种轻量化模型检测精度都有细微下降,满足在兼顾检测精度的情况下实现模型轻量化需求。

通过训练得到 7 个模型的优化参数后,使其在统一的测试集上进行测试,结果如表 4 所示。

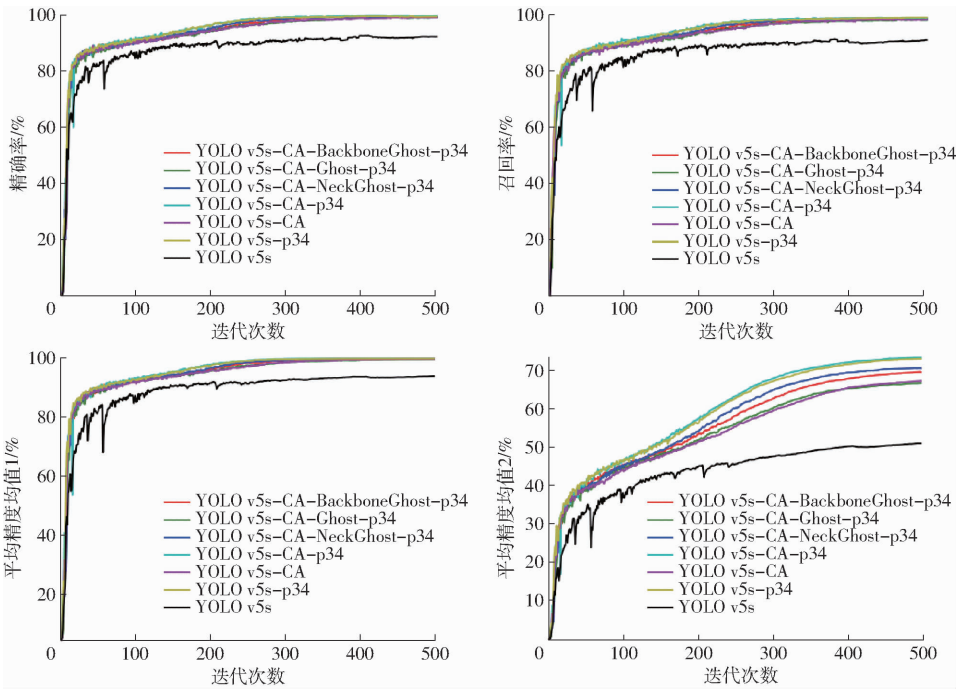


图 15 不同 YOLO 改进算法在训练过程中的性能结果

Fig. 15 Performance results of different YOLO improvement algorithms in training process

表 4 算法优化实验测试集性能评估结果

Tab.4 Performance evaluation results of algorithm optimization experiment test set

模型	精确率/ %	召回率/ %	平均精度 均值 1/%	平均精度 均值 2/%	浮点数 计算量	模型内存占 用量/MB
YOLO v5s	92.5	91.0	92.7	53.2	1.58×10^{10}	13.7
YOLO v5s - CA	97.6	96.6	97.9	67.6	1.58×10^{10}	13.7
YOLO v5s - p34	97.1	97.0	98.0	72.2	1.43×10^{10}	10.2
YOLO v5s - CA - p34	97.3	97.1	98.3	72.6	1.44×10^{10}	10.3
YOLO v5s - CA - Ghost - p34	97.1	96.8	97.8	67.0	7.40×10^9	5.8
YOLO v5s - CA - NeckGhost - p34	97.4	97.2	98.1	70.7	1.26×10^{10}	9.4
YOLO v5s - CA - BackboneGhost - p34	97.1	96.9	97.9	69.7	9.20×10^9	6.7

从表 4 可得,相比于 YOLO v5 原始模型,单独采用 CA 注意力模块、单独采用剔除大目标检测头优化策略的模型在测试集上的平均精度均值 1 分别提升 5.2、5.3 个百分点,在平均精度均值 2 上分别提高了 13.6、19 个百分点,当两种优化方法结合时,其平均精度均值 1 和平均精度均值 2 分别提高 5.6 个百分点和 19.4 个百分点。

通过引入 Ghost 消融实验,发现在 Backbone 网络引入 Ghost 结构的模型相比 YOLO v5s-CA-p34 模型在平均精度均值 1 仅降 0.4 个百分点,但是浮点数计算量和模型内存占用量分别减少 36% 和 35%。同时,与另外两种引入 Ghost 结构的模型相

比,本模型在平均精度均值等性能上的优势十分有限,但是在浮点数计算量和模型内存占用量的消减效果上优化幅度较大,更符合实际部署应用的需求。

从表 4 还可以发现,无论是进行骨干网络还是 Head 网络优化改进方法,其精确率差距小于 0.5 个百分点,召回率的差距小于 0.6 个百分点。该结果与在训练过程中验证集的性能结果相近,说明模型并没有过拟合。最后通过对骨干网络和 Head 网络共同优化改进,得到 YOLO v5s-CA-BackboneGhost-p34 模型,其性能参数和损失值如图 16 所示。

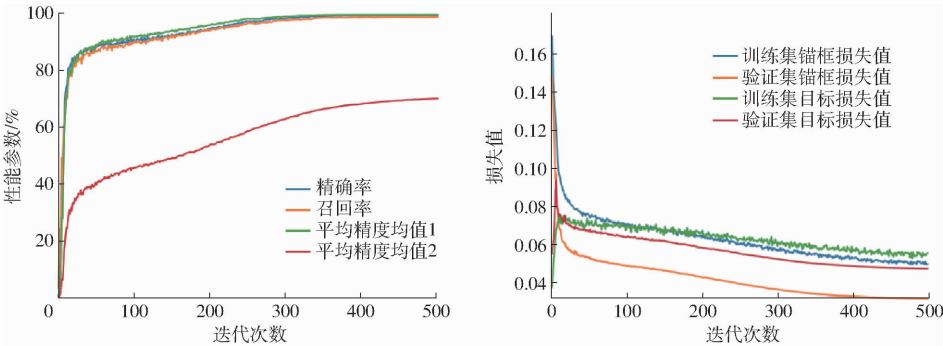


图 16 YOLO v5s-CA-BackboneGhost-p34 模型性能参数与损失值变化曲线

Fig. 16 Performance parameters and loss curves variations of YOLO v5s-CA-BackboneGhost-p34

3.2.2 与其他主流算法对比

为验证本文提出的改进算法相较于其他检测速度较快的主流目标检测算法的优越性,本文将改进算法与 DETR^[22]、CenterNet^[23]、YOLO v3-tiny、YOLO v4-Mish^[24-25]、YOLO v5s、YOLO v5-lite-g、YOLOX-s^[26]、YOLO v7-tiny^[27] 等算法在相同条件下进行比较实验,实验结果如表 5 所示。

表 5 主流算法性能对比实验

Tab. 5 Comparative experiment of mainstream algorithms

算法	准确率/%	召回率/%	平均精度均值 1/%	模型内存占用量/MB	浮点数计算量
DETR	93.0	91.6	93.3	15.9	
CenterNet	90.0	79.0	83.3	124.0	
YOLO v3-tiny	94.3	92.3	94.7	16.6	1.29×10^{10}
YOLO v4s-Mish	90.7	90.8	92.6	20.6	2.06×10^{10}
YOLO v5s	93.0	92.6	94.0	13.7	1.58×10^{10}
YOLO v5-lite-g	96.3	95.7	97.1	14.2	2.53×10^{10}
YOLOX-s	95.8	95.1	96.6	15.5	2.16×10^{10}
YOLO v7-tiny	96.9	94.7	96.2	11.6	1.30×10^{10}
本文方法	97.1	96.9	97.9	6.7	9.20×10^9

通过表 5 可得,相较于其他主流的目标检测算法,本文提出的改进算法具有更高的检测精度和更

小的体积,优势非常明显。

3.3 算法优化后甘蔗茎节检测实例

使用上述优化算法,针对甘蔗茎节图像实例在相同参数条件下进行了测试,如图 17 所示(紫色圈表示漏检,绿色圈表示误检)。在甘蔗茎节实例检测中,检测效果可以较为直观通过误检情况、漏检情况与正例的置信度区间 3 个指标评价。

本文选择了不同光照强度、不同茎节疏密程度的 3 个画面进行甘蔗茎节实例检测的对比分析。从图 17 可以看出,顶部图像为正常光照场景下,YOLO v5s 模型存在漏检甘蔗茎节的情况,而其余改进后的模型均能够识别出全部甘蔗茎节;中间图像为茎节密集且灯光分布不均匀的场景下,大部分模型都出现了漏检和误检情况,漏检位置主要出现在光线较暗的一端;底部图像为光线较暗的场景下,YOLO v5s-CA-BackboneGhost-p34 识别出了所有茎节,而其他两种轻量化模型都出现了误检情况。

综上所述,本文进行算法优化得到的 YOLO v5s-CA-BackboneGhost-p34 模型在复杂光照场景下取得较好的检测效果,也具有较小的体积,更符合作为边缘端部署的模型。

4 模型边缘端部署

将训练好的 YOLO v5s-CA-BackboneGhost-p34

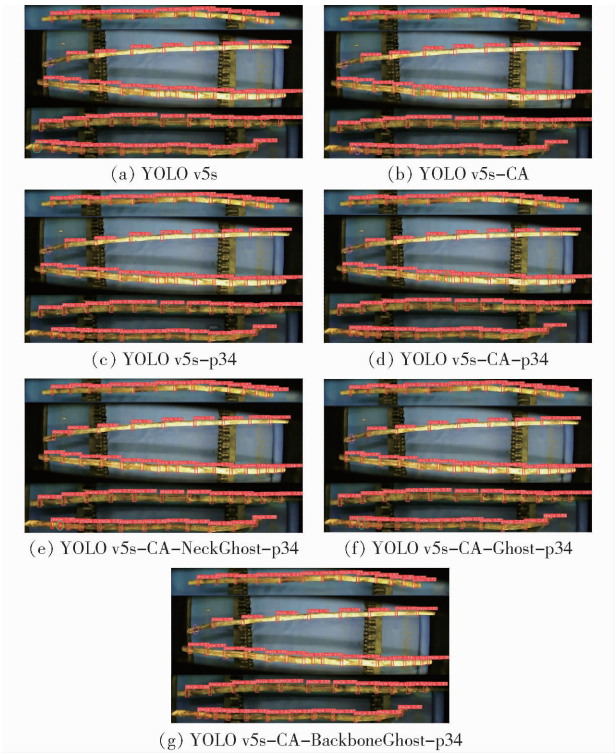


图 17 算法优化后甘蔗茎节检测实例对比

Fig. 17 Example comparison of sugarcane stem node detection after algorithm optimization

模型部署到边缘设备上,这将有利于提高数据传输过程中的响应速度和数据安全性,同时为了提高模型检测速度,利用 TensorRT 进行加速,进一步验证了模型的可靠性,并且提高了检测效率^[28]。TensorRT 加速前后速度对比如图 18 所示。

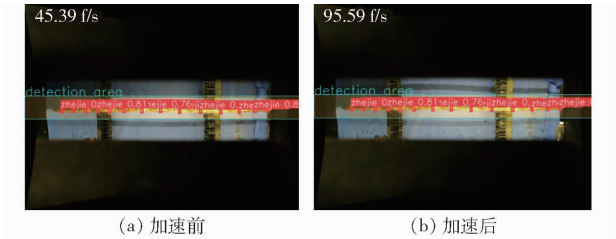


图 18 TensorRT 加速前后对比

Fig. 18 TensorRT acceleration before and after comparison

从图 18 中可知,将模型部署在边缘设备上,并通过 TensorRT 加速,模型检测速度比原来提高 1.1 倍,检测速度仅需 10.5 ms,具有更高的实用性与便捷性,为实际工厂化切种提供了良好的基础。

最后,利用 VNC(Virtual network computing)软件进行桌面共享和远程操作,实现检测画面的实时显示。

本模型的开发环境如下:开发和调试是在 PyCharm 专业版上进行,模型的训练和测试是在 Windows 10 系统上,GPU 为英伟达的 3060 显卡。模型部署采用的嵌入式设备是英伟达的 Jetson Orin NX 16GB 版,它具有强大的计算能力和丰富的输入

输出接口,适合进行边缘计算任务,部署环境是在 Ubuntu 20.04.5 系统上,深度学习框架采用 Pytorch 2.0 版本,采用 TensorRT 8.5.2.2。

5 实际切种试验与结果分析

5.1 切种质量评价准则

根据广西壮族自治区甘蔗良繁基地的用种情况以及农艺的需求,提出以下切种质量评价准则(如图 19 所示):切口与茎节距离必须大于 5 mm,否则属于伤芽情况;蔗种必须包含 2 个或 3 个茎节,否则属于单芽段蔗种。这两种类型的种子都不利于甘蔗种植。因为切口离茎节太近会破坏茎节的结构,影响种子吸收营养和水分,而种植单芽段的蔗种无法保证高出芽率,影响甘蔗产量。所以在进行切种作业时,既要保证切割位置的准确性,还要保证足够的茎节数量。



图 19 甘蔗种子切割质量分类

Fig. 19 Classification of sugarcane seed cutting quality

5.2 实际切种试验

为测试本文优化改进模型的效果,在本课题组开发的横向智能切种工作站试验样机上进行双芽蔗切种验证试验。

(1) 样本准备:本文将保留的 20 根长度约为 1.8 m 的甘蔗进行标号,并测量记录每根甘蔗的茎段长度和茎节数量,如图 20a、20b 所示。

(2) 茎节特征检测识别与定位:在传送速度为 0.15 m/s 的情况下,蔗种被有序地送入检测区域(图 20c),对整根甘蔗的茎节特征进行实时检测识别与定位,并将数据保存。接着,系统会根据畸变矫正后的茎节坐标的 x 轴从小到大排序,将相邻的两个茎节的中心坐标计算出来,并将偶数下标的中心坐标作为切割位置发送给控制器。

(3) 切种作业:当甘蔗经过切刀平台前的光电传感器时,6 把切刀将根据切割位置进行自动调刀与切种,如图 20d、20e、20f 所示。

(4) 结果统计分析:通过比较人工测量茎段长度与系统计算茎段长度,评测茎节识别方面的定位误差,如图 21 所示。最后,对切种结果进行人工统计,统计结果见表 6。

经统计,20 根甘蔗的茎段定位精度平均误差约为 2.4 mm,在均值加减 2 个标准差($\pm 2\sigma$)的范围

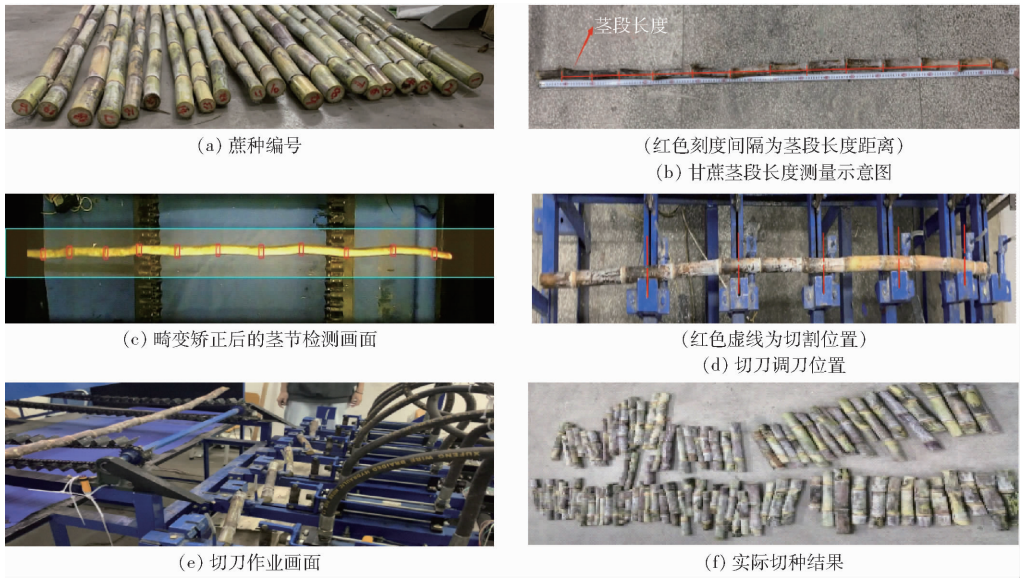


图 20 实际切种试验流程

Fig. 20 Actual seed cutting test flow

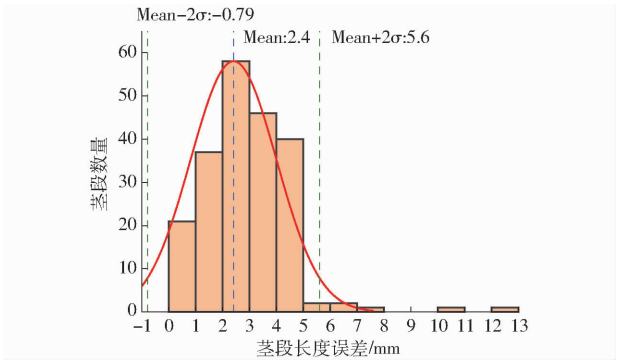


图 21 甘蔗茎段的检测误差分析图

Fig. 21 Histogram of detection error of sugarcane stem segments

内覆盖大约 97% 的数据。甘蔗茎节检测有 1 个漏检,切种合格率为 100%。由于漏检的茎节位置位于甘蔗末端,所以并未对切种任务造成影响。因此,本试验足以验证本模型的优越性、可靠性以及实用性。

6 结论

(1)经实验测试,数据增强可以有效提高模型精度和泛化能力,降低过拟合的风险。其中精度、召回率、平均精度均值 1 以及平均精度均值 2 分别提高 3.8、5.2、5.4、8.5 个百分点。

表 6 切种结果统计

Tab. 6 Statistics of seed cutting results

参数	甘蔗 总数/根	实际茎节 数量/个	检测茎节 数量/个	漏检茎节 数量/个	漏检率/ %	合格蔗段数(包括双 芽段和三芽段)/段	单芽段数/ 段	伤芽段数/ 段	切种合格 率/%
数值	20	230	229	1	0.4	109	0	0	100

(2)在 YOLO v5s 的基础上,在骨干网络中添加 CA 注意力机制以及引入 Ghost 结构,可以稳定提升模型对甘蔗茎节的特征提取能力,帮助模型更好的检测茎节这类小目标,并且可以在确保高精度的情况下降低模型复杂度,为部署在嵌入式设备上奠定基础;在 Head 网络剔除大目标检测头,有利于茎节小目标的位置信息检测,降低模型后处理阶段的容错率,使模型更加专注于训练有用的特征,减少网络浮点数计算量和参数量,加快模型推理速度,提高模型精度。试验表明,同时使用 CA 注意力模块以及剔除大目标检测策略对模型提升效果最好,并且仅在骨干网络引入 Ghost 结构可以兼顾模型精度和大小,更符合边缘端部署

的应用需求。经实验测试, YOLO v5s - CA - BackboneGhost - p34 较原模型平均精度均值 1 提升 5.2 个百分点,平均精度均值 2 提升 16.5 个百分点,浮点数计算量和模型内存占用量分别降低 42% 和 51%,相比其他主流的目标检测方法,其具有更高的检测精度和更小的体积。

(3)将 YOLO v5s - CA - BackboneGhost - p34 模型部署在 Jetson Orin NX16GB 边缘设备上,并通过 TensorRT 加速后,每幅图的检测时间仅需 10.5 ms,检测速度提高 1 倍多。

(4)经实际切种试验验证,在传送速度为 0.15 m/s 的情况下,切割 20 根长度约为 1.8 m 的甘蔗,计算与测量的茎段长度平均误差为 2.4 mm,切

种合格率为 100%。

(5)所提方法可以有效提高甘蔗智能横向预切切种的需求,为甘蔗横向切种工作站的工厂化、智能化以及标准化应用提供有效的技术支持。

种工作站的切种精度和效率,适于现场生产环境下

参 考 文 献

[1] 张木清,姚伟. 现代甘蔗栽培育种学[M]. 北京:北京科学技术出版社,2021.

[2] 李尚平,向锐,廖义奎,等. 智能甘蔗预切种工作站结构与切种精度分析[J]. 中国农机化学报,2020,41(10):74-82.
LI Shangping, XIANG Rui, LIAO Yikui, et al. Structural design and seed cutting accuracy analysis on intelligent sugarcane seed pre-cutting workstation[J]. Journal of Chinese Agricultural Mechanization,2020,41(10):74-82. (in Chinese)

[3] CHEN J, WU J, QIANG H, et al. Sugarcane nodes identification algorithm based on sum of local pixel of minimum points of vertical projection function[J]. Computers and Electronics in Agriculture, 2021, 182: 105994.

[4] ZHOU D, FAN Y, DENG G, et al. A new design of sugarcane seed cutting systems based on machine vision[J]. Computers and Electronics in Agriculture, 2020, 175: 105611.

[5] 张圆圆,何永玲,王跃飞,等. 颜色空间图像处理技术在蔗节识别上的应用[J]. 农机化研究,2020,42(1):231-236.
ZHANG Yuanyuan, HE Yongling, WANG Yuefei, et al. Application of color space image processing technology in sugarcane node recognition[J]. Journal of Agricultural Mechanization Research,2020, 42(1):231-236. (in Chinese)

[6] CHEN M, XU Q, CHENG Q, et al. Sugarcane stem node detection based on wavelet analysis[J]. IEEE Access, 2021, 9: 147933-147946.

[7] MENG Y, YE C, YU S, et al. Sugarcane node recognition technology based on wavelet analysis[J]. Computers and Electronics in Agriculture, 2019, 158: 68-78.

[8] 陈延祥. 基于机器视觉的甘蔗多刀切种装备设计与研究[D]. 无锡:江南大学,2022.
CHEN Yanxiang. Design and research on sugarcane multi-cutter cutting equipment based on machine vision[D]. Wuxi: Jiangnan University, 2022. (in Chinese)

[9] 赵文博,周德强,邓干然,等. 基于改进 YOLO v5 的甘蔗茎节识别方法[J]. 华中农业大学学报,2023,42(1):268-276.
ZHAO Wenbo, ZHOU Deqiang, DENG Ganran, et al. Sugarcane stem node recognition method based on improved YOLO v5[J]. Journal of Huazhong Agricultural University,2023, 42(1):268-276. (in Chinese)

[10] 廖义奎,李尚平,文春明. 一种基于深度学习的甘蔗蔗芽特征识别装置:CN108875789B[P]. 2021-04-27.

[11] 李尚平,李向辉,张可,等. 改进 YOLO v3 网络提高甘蔗茎节实时动态识别效率[J]. 农业工程学报,2019,35(23):185-191.
LI Shangping, LI Xianghui, ZHANG Ke, et al. Increasing the real-time dynamic identification rate of sugarcane nodes by improved YOLO v3 network[J]. Transactions of the CSAE, 2019, 35(23): 185-191. (in Chinese)

[12] ZHANG Z. A flexible new technique for camera calibration[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(11): 1330-1334.

[13] 董舒蕾. 广西甘蔗产业发展对策研究[D]. 南宁:广西大学,2022.
DONG Shulei. Research on the development countermeasures of Guangxi sugarcane industry[D]. Nanning: Guangxi University, 2022. (in Chinese)

[14] 邓宇驰,刘晓婷,黄莹,等. 广西崇左蔗区 2022 年糖料蔗生产调查[J]. 中国种业,2022(10):48-51.

[15] 张志远,罗铭毅,郭树欣,等. 基于改进 YOLO v5 的自然环境下樱桃果实识别方法[J]. 农业机械学报,2022,53(增刊 1):232-240.
ZHANG Zhiyuan, LUO Mingyi, GUO Shuxin, et al. Cherry fruit detection method in natural scene based on improved YOLO v5[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(Supp. 1):232-240. (in Chinese)

[16] 王卓,王健,王泉雄,等. 基于改进 YOLO v4 的自然环境苹果轻量级检测方法[J]. 农业机械学报,2022,53(8):294-302.
WANG Zhuo, WANG Jian, WANG Xiaoxiong, et al. Lightweight real-time apple detection method based on improved YOLO v4[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(8):294-302. (in Chinese)

[17] 修春波,孙乐乐. 基于改进 YOLO v4 网络的马铃薯自动育苗叶芽检测方法[J]. 农业机械学报,2022,53(6):265-273.
XIU Chunbo, SUN Lele. Potato leaf bud detection method based on improved YOLO v4 network[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(6):265-273. (in Chinese)

[18] 宋怀波,王亚男,王云飞,等. 基于 YOLO v5s 的自然场景油茶果识别方法[J]. 农业机械学报,2022,53(7):234-242.
SONG Huaibo, WANG Ya'nan, WANG Yunfei, et al. Camellia oleifera fruit detection in natural scene based on YOLO v5s[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(7):234-242. (in Chinese)

[19] HOU Q, ZHOU D, FENG J. Coordinate attention for efficient mobile network design[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 13713-13722.

[20] 吴珺,董佳明,刘欣,等. 注意力优化的轻量目标检测网络及应用[J]. 智能系统学报,2023,18(3):506-516.
WU Jun, DONG Jiaming, LIU Xin, et al. Lightweight object detection network and its application based on the attention optimization[J]. CAAI Transactions on Intelligent Systems, 2023, 18(3):506-516. (in Chinese)

[21] HAN K, WANG Y, TIAN Q, et al. Ghostnet: more features from cheap operations[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 1580-1589.

- [8] LUCA B, JACK V, JOAO F H, et al. Fully-convolutional siamese networks for object tracking[C] // European Conference on Computer Vision. Amsterdam: Springer, 2016: 850 – 865.
- [9] LUCA B, JACK V, JOAO F H, et al. Learning feed-forward one-shot learners[C] // Advances in Neural Information Processing Systems, 2016.
- [10] LUCA B, JACK V, JOAO F H, et al. Staple: complementary learners for real-time tracking[C] // Conference on Computer Vision and Pattern Recognition (CVPR), 2016.
- [11] LI B, YAN J J, WU W, et al. High performance visual tracking with siamese region proposal network[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 8971 – 8980.
- [12] LI B, WEI W, WANG Q, et al. Siamrpn + : evolution of siamese visual tracking with very deep networks[C] // Conference on Computer Vision and Pattern Recognition (CVPR), 2019: 4282 – 4291.
- [13] ZHANG Z P, PENG H W, FU J L, et al. Ocean: object-aware anchor-free tracking[C] // European Conference on Computer Vision, 2020: 771 – 787.
- [14] GUO D Y, WANG J, CUI Y, et al. Siamcar: siamese fully convolutional classification and regression for visual tracking[C] // Conference on Computer Vision and Pattern Recognition (CVPR), 2020: 6269 – 6277.
- [15] CHEN Z D, ZHONG B N, LI G R, et al. Siamese box adaptive network for visual tracking[C] // Conference on Computer Vision and Pattern Recognition (CVPR), 2020: 6668 – 6677.
- [16] FENG T, QIANG L. Learning to rank proposals for siamese visual tracking[J]. IEEE Transactions on Image Processing: A Publication of the IEEE Signal Processing Society, 2021, 30: 8785 – 8796.
- [17] HELD D, THRUN S, SAVARESE S. Learning to track at 100 fps with deep regression networks[C] // Conference on Computer Vision and Pattern Recognition (ECCV), Springer, 2016: 749 – 765.
- [18] SANDLER M, HOWARD A, ZHU M L, et al. Mobilenetv2: inverted residuals and linear bottlenecks[C] // Conference on Computer Vision and Pattern Recognition (CVPR), 2018: 4510 – 4520.
- [19] MA N N, ZHANG X Y, ZHENG H T, et al. Shufflenet v2: practical guidelines for efficient CNN architecture design[C] // European Conference on Computer Vision, 2018.
- [20] KUMAR P, VASU A, JAMES G, et al. An improved one millisecond mobile backbone[J]. arXiv preprint: 2206.04040, 2022.
- [21] OLGA R, DENG J, SU H, et al. Imagenet large scale visual recognition challenge[J]. International Journal of Computer Vision, 2015, 115(3): 211 – 252.
- [22] LIN T Y, MAIRE M, SERGE B, et al. Microsoft COCO: common objects in context[C] // European Conference on Computer Vision, 2014: 740 – 755.
- [23] HUANG L H, ZHAO X, HUANG K Q. GOT-10k: a large high-diversity benchmark for generic object tracking in the wild[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019.
- [24] REAL E, SHLENS J, MAZZOCCHI S, et al. YouTube-BoundingBoxes: a large high-precision human-annotated data set for object detection in video[C] // Conference on Computer Vision and Pattern Recognition (CVPR), 2017: 5296 – 5305.
- [25] FAN H, LIN L T, YANG F, et al. LaSOT: a high-quality benchmark for large-scale single object tracking[C] // Conference on Computer Vision and Pattern Recognition, 2019: 5374 – 5383.
- [26] WANG Q, ZHANG L, LUCA B, et al. Fast online object tracking and segmentation: a unifying approach[C] // Conference on Computer Vision and Pattern Recognition (CVPR), 2019: 1328 – 1338.
- [27] JACK V, LUCA B, JOAO H, et al. End-to-end representation learning for correlation filter based tracking[C] // Conference on Computer Vision and Pattern Recognition, 2017: 2805 – 2813.
- [28] SANH V, WOLF T, RUSH A M. Movement pruning: adaptive sparsity by fine-tuning[C] // Proceedings of the Advances in Neural Information Processing Systems, 2020: 20378 – 20389.

(上接第 245 页)

- [22] CARION N, MASSA F, SYNNAEVE G, et al. End-to-end object detection with transformers[C] // European Conference on Computer Vision. Cham: Springer International Publishing, 2020: 213 – 229.
- [23] DUAN K, BAI S, XIE L, et al. Centernet: keypoint triplets for object detection[C] // Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 6569 – 6578.
- [24] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLO v4: optimal speed and accuracy of object detection[J]. arXiv Preprint, arXiv:2004.10934, 2020.
- [25] MISRA D. Mish: a self regularized non-monotonic activation function[J]. arXiv Preprint, arXiv:1908.08681, 2019.
- [26] GE Z, LIU S, WANG F, et al. YOLOx: exceeding yolo series in 2021[J]. arXiv Preprint, arXiv:2107.08430, 2021.
- [27] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLO v7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 7464 – 7475.
- [28] JEONG E J, KIM J, HA S. Tensorrt-based framework and optimization methodology for deep learning inference on jetson boards[J]. ACM Transactions on Embedded Computing Systems (TECS), 2022, 21(5): 1 – 26.