

赵文博,周德强,邓干然,等.基于改进YOLOv5的甘蔗茎节识别方法[J].华中农业大学学报,2023,42(1):268-276.
DOI:10.13300/j.cnki.hnlkxb.2023.01.032

基于改进YOLOv5的甘蔗茎节识别方法

赵文博¹,周德强^{1,2},邓干然^{3,4},何冯光^{3,4},朱琦¹,韦丽娇^{3,4},牛钊君^{3,4}

1. 江南大学机械工程学院,无锡 214122; 2. 江苏省食品先进制造装备技术重点实验室,无锡 214122;
3. 中国热带农业科学院农业机械研究所,湛江 524088;
4. 农业农村部热带作物农业装备重点实验室,湛江 524091

摘要 针对甘蔗智能切种机作业过程中背景杂乱导致茎节识别精度低等问题,提出了基于改进YOLOv5的一种甘蔗茎节识别方法。采用跨层级连接的方式优化颈部结构,增强不同层级间的信息融合能力;同时改进模型损失函数,一方面引入EIoU损失函数代替原始CIoU损失函数,提高边界框回归精度,另一方面利用Focal loss损失函数替换交叉熵损失函数,解决正负样本比例不均衡问题;最后引入Ghost模块轻量化网络模型。试验结果表明,本研究提出的模型相较于原模型,平均精度值提高了1.4个百分点,达97.80%,单张检测时间为16.9 ms,模型大小仅11.40 Mb,实现了在不同杂乱程度场景下的甘蔗茎节识别,降低了切种时背景杂乱产生的影响。

关键词 甘蔗; 茎节识别; 杂乱背景; YOLOv5; 跨层连接; 损失函数; 精准切割

中图分类号 TP391.4; S238 **文献标识码** A **文章编号** 1000-2421(2023)01-0268-09

目前我国甘蔗产业仅在机耕、培土等环节实现了机械化,但机械种植、机械收获方面仍未普及,导致产业整体机械化程度低。甘蔗的机械种植主要采用预切种式甘蔗种植机,依据农艺要求提前对优质甘蔗进行切种,并完成晒种、浸种和杀菌等农艺流程^[1]。由于人工切种效率低下,目前已开发出防伤芽智能切种方法,能够有效识别茎节位置。但随着切种工作的进行,会产生大量碎屑、杂叶等污垢造成图像采集区域背景污染,导致算法识别精度下降,甚至失效。因此,亟需开发防伤芽的智能切种方法,能够在杂乱背景中精确识别茎节位置,提高茎节识别精度,以实现高效的防伤芽切种。

近年来,针对甘蔗茎节识别定位问题,国内外开展了广泛的研究,通过传统的机器视觉方法实现了甘蔗茎节检测。Zhou等^[2]获取甘蔗图像的感兴趣区域,利用Sobel算子提取边缘特征并采用构造的矩形检测算子积分运算得到特征描述向量,通过检测描述向量的峰值来识别茎节,识别准确率为93%。Chen等^[3]提出一种基于垂直投影函数最小点和局部像素的方法来定位甘蔗茎节,双茎节识别率98.5%。

Yang等^[4]通过Sobel算子计算甘蔗图像在水平方面的梯度信息,利用平均梯度和方差梯度特征识别甘蔗茎节,识别准确率为96.89%。Nare等^[5]利用Sobel算子识别甘蔗边缘,通过调节阈值去除噪音,用识别边缘的附加阈值和边缘总数识别甘蔗芽。张圆圆等^[6]根据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基础上采用中值决策识别甘蔗茎节,识别正确率达到94.7%。以上基于传统机器视觉的检测方法鲁棒性较差,当甘蔗品种更换或背景环境复杂时识别能力会大幅降低,无法按要求完成防伤芽切种工作。随着深度学习越来越广泛地应用于农业领域,目前已取得诸多成果。如李尚平等^[7]通过改变输出特征图尺寸和减少anchors的方法改进YOLOv3,实现甘蔗切种装置对茎节特征的连续、动态的智能识别,识别准确率96.89%。黄彤镔等^[8]通过引入CBAM注意力机制提高特征提取能力,并采用 α -IoU损失函数增强边界框定位精度的方法改进YOLOv5,实现了在自然环境下快速准确的识别柑橘果实。赵德安等^[9]通过替换骨干网络的方法改进YOLOv3,在保证效率和准

收稿日期: 2022-06-20

基金项目: 中央级公益性科研院所基本科研业务费专项(1630132022001)

赵文博, E-mail: 1475817953@qq.com

通信作者: 周德强, E-mail: zhoudeqiang@jiangnan.edu.cn

准确率的前提下实现了复杂环境下苹果的检测。顾伟等^[10]利用ResNet 50网络改进SSD算法的主干结构,提高群体棉籽图像特征提取效率,实现了对群体棉籽的破损检测。

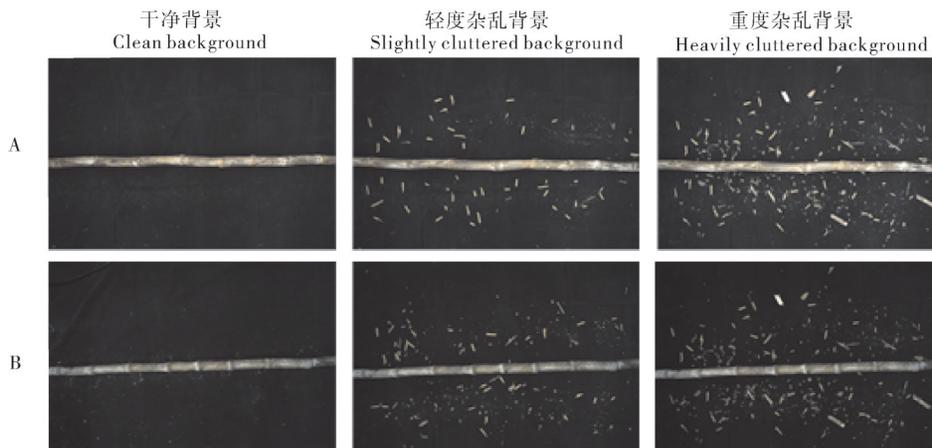
基于传统的机器视觉算法虽然准确率较高,但鲁棒性差,而基于深度学习的算法能增强鲁棒性,但由于模型尺寸较大,需要高昂的计算成本才能完成。为解决上述问题,本研究提出了一种基于改进YOLOv5的甘蔗茎节识别定位方法,以期甘蔗预切种机在杂乱背景下工作提供算法支持。

1 材料与方法

1.1 数据采集与数据集制作

甘蔗样本采集于广东省湛江市中国热带科学院农业机械研究所试验田,采集设备型号为工业相机MER-2000-5GM/C、镜头HN-0826-20M-C1/1X。考虑到不同品种甘蔗直径、茎节形状、颜色以及灰度值变化都会有所不同,选取贵糖49号、新台糖22号2种甘蔗作为样本,人工砍收后统一在研究所试验基

地拍摄。由于农业机械工况恶劣,随着切种工作不断进行,切种机内部图像采集区域会因各种污染造成背景杂乱,如甘蔗切割碎屑、泥土碎屑、切割时飞溅出的蔗糖汁黏液造成的枝叶黏附等。在样本采集时为模拟切种机内部背景杂乱的场景,根据真实内部环境产生的背景污染,分析其干扰原因后在拍摄背景中随机加入碎屑模拟,其中包含因锯片造成的细小碎屑、甘蔗根部黏附的泥土颗粒以及枯叶。模拟环境中加入的干扰元素包含了真实环境中的各类污染,故二者较为相似。拍摄完成后按照背景杂乱程度依次分为干净背景样本、轻度杂乱背景样本、重度杂乱背景样本。最终采集贵糖49号510张、新台糖22号500张,茎节数量共10 094个。采集的单张图片大小为57.7 Mb,分辨率为 $5\,496 \times 3\,672$,存储格式为bmp格式。采用图片标注工具LabelImg将图片标注成VOC格式的xml文件,之后将其转换为yolo格式的txt文件并按照7:2:1的比例划分训练集,测试集和验证集。数据集部分样本如图1所示。



A:新台糖22号 New Taitang 22; B:贵糖49号 Guitang 49.

图1 部分样本图片

Fig.1 Some sample pictures

1.2 YOLOv5架构

随着目标检测算法的不断改进与迭代,YOLOv5较YOLOv4在精度、速度和模型大小上均有较大程度的优化,更适合部署在嵌入式设备上用于工程实践。YOLOv5s是YOLOv5模型系列中尺寸最小的版本,其大小仅14.50 Mb,但在复杂背景环境下容易受到干扰,造成错检、精度不能满足要求等问题。YOLOv5s的主要结构由Input输入端、Backbone骨干网络、Neck网络层和Head检测端4个部分组成,如图2所示。Input输入端通过Mosaic数

据增强丰富数据集,采用K-means自适应锚框算法计算获取最佳锚框尺寸并自适应缩放图片大小以减少计算量。Mosaic数据增强算法^[11]从同一批次的BatchSize个样本中随机抽取4张图片并进行裁剪、缩放,合并为1张图片后输入网络,不仅可以丰富数据集,而且能够提升网络的训练速度。Backbone骨干网络包含C3模块以及SPPF模块,经改进后的C3模块由于使用了残差结构,改善了反向传播过程中梯度爆炸和梯度消息问题。Neck网络主要对骨干网络提取到的特征信息进行融合,共包括FPN特征金

字塔^[12]和PAN路径聚合网络结构^[13]中的2部分信息。Head检测头将Neck网络输出的不同尺度的特

征图进行解析,并通过损失函数和NMS非极大值抑制进行训练或预测。

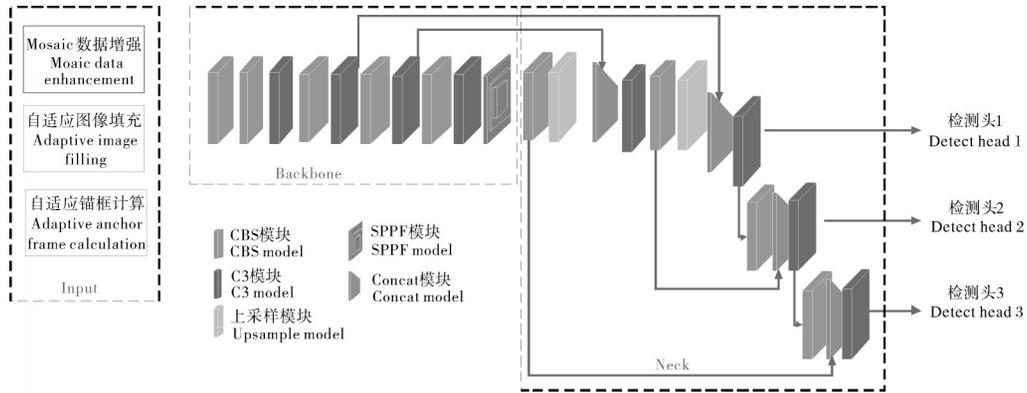


图2 YOLOv5s网络结构

Fig.2 YOLOv5s network structure

1.3 模型改进

1) 改进Neck网络层。YOLOv5s网络的颈部结构设计采用FPN+PAN的结构,FPN结构利用骨干网络最后1层输出特征图进行2次上采样,在3个尺度上构建出高级语义特征图。由于连续上采样会造成浅层信息大量丢失,影响甘蔗茎节识别定位。为实现茎节位置的精确定位,减少杂乱背景带来的干扰,需充分利用主干网络提取到的浅层信息,增强网络对定位框的回归能力。因此,基于BiFPN^[14]提出了跨层加权连接的Neck结构,将颈部和骨干网络同一尺度的特征信息相融合,在不大幅增加计算成本的情况下,将浅层语义信息如边缘、轮廓等融入到PAN结构中,使模型对甘蔗茎节定位的边界框回归更加精确。在信息融合过程中,一般采用将不同层级的特征图直接相加的方法,但实际上不同特征层对最终结果产生的影响不同,因此,引入可学习的方式,在迭代过程中不断调整和学习不同的输入特征层权重。

卷积核。

由式(1)可知,在进行concat拼接时对于不同的特征层采用不同的卷积核运算,模型经过concat融合操作后提取到不同的语义信息,并增加了通道数。由式(2)可知,add操作只是简单地进行像素叠加,会导致信息损失,但共享卷积核参数使参数量较concat操作减少一半,二者均有不同的优势。改进后的Neck结构如图3所示,其中虚线条代表添加可学习权重的特征融合。

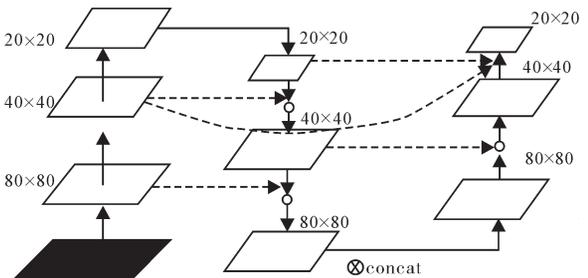


图3 改进后Neck结构

Fig.3 Improved Neck structure

特征融合有多种方式,如采用concat操作将通道叠加进行特征融合、在通道数不变的情况下直接将2张特征图数值相加的add操作以及采用全连接操作进行融合等。因全连接融合方式导致参数量大幅增加,故这里不进行考虑。对add和concat操作分析如下:

2) 损失函数改进。损失函数主要用来评估真实值与预测值之间的差异程度,YOLOv5s的损失函数分别由目标置信度损失、边框定位损失和分类损失3部分构成。改进后的损失函数采用Focal loss计算目标置信度损失,并采用EIoU loss^[15]计算边界框定位损失。

$$out_{concat} = \sum_{i=1}^c x_i \otimes k_i + \sum_{i=1}^c y_i \otimes k_{i+c} \quad (1)$$

$$out_{add} = \sum_{i=1}^c (x_i \oplus y_i) * k_i \quad (2)$$

其中, out_{concat} 和 out_{add} 分别为不同方法融合后的特征图; x_i 和 y_i 分别为不同通道的输入; k_i 表示第 i 个

①采用焦点损失函数 Focal loss^[16] 替换交叉熵损失函数计算目标置信度损失。YOLOv5s的目标置信度损失采用交叉熵损失,如公式(3)。对于二分类问题,总的损失函数如公式(4)。当样本分布不均匀时,损失函数会更加偏重样本数多的类别,即当 $m \gg n$ 时,模型会朝着与预期不同的方向优化。

在检测甘蔗图片时,模型会生成大量候选框,虽然 YOLOv5s 利用跨邻域网格匹配策略增大了正样本的数量,但是负样本仍占较大比例,使损失函数更关注负样本,因此,本研究引入焦点损失函数 Focal loss,在一定程度上降低正负样本比例失衡所带来的影响,有效减少背景带来的干扰。

$$L(y, \hat{p}) = -y \log(\hat{p}) - (1 - y) \log(1 - \hat{p}) \quad (3)$$

$$L = \frac{1}{N} \left(\sum_{y_i=1}^m -\log(\hat{p}) + \sum_{y_i=0}^n -\log(1 - \hat{p}) \right) \quad (4)$$

其中, \hat{p} 为预测概率大小,样本总数 $N = m + n$ 。

Focal loss 的具体形式如公式(5)所示,将公式(6)代入公式(5)可得到公式(7)。

$$L_{\text{focal}} = \begin{cases} -\alpha_t (1 - \hat{p})^\gamma \log(\hat{p}) & \text{if } y = 1 \\ -(1 - \alpha_t) \hat{p}^\gamma \log(1 - \hat{p}) & \text{if } y = 0 \end{cases} \quad (5)$$

$$p_t = \begin{cases} \hat{p} & \text{if } y = 1 \\ 1 - \hat{p} & \text{otherwise} \end{cases} \quad (6)$$

$$L_{\text{focal}} = -\alpha_t (1 - p_t)^\gamma \log(p_t) \quad (7)$$

由式(3)和式(7)可知,Focal loss 其本质在于加入了调制因子 $(1 - p_t)^\gamma$ 和平衡因子 α_t ,其中平衡因子 α_t 纠正正负样本对总损失的影响,反映了分类的难易程度, p_t 越大说明目标置信度越高,样本易于分类; p_t 越小则相反。引入 Focal loss 从特征识别难易角度和样本分布角度出发,解决了样本不均衡带来的模型优化问题。

② 采用 EIoU^[17] 损失函数计算边框定位损失。YOLOv5s 原采用 CIoU 损失函数计算边框回归定位损失,计算方法如式(8)。

$$L_{\text{CIoU}} = 1 - \text{IoU} + \frac{\rho^2(b, b^{gt})}{c^2} + \alpha v \quad (8)$$

$$\alpha = \frac{v}{(1 - \text{IoU}) + v} \quad (9)$$

$$v = \frac{4}{\pi} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2 \quad (10)$$

其中, $\rho^2(b, b^{gt})$ 代表两框中心点的欧式距离, c 代表能同时覆盖两框的最小闭包区域的对角线距离。

CIoU 损失函数从边界框回归的重叠面积、中心点距离、纵横比三方面进行回归使定位更精确。由公式(10)可得,对预测框形状回归时 CIoU 损失函数对长宽比进行惩罚,当 $w = kw^{gt}$, $h = kh^{gt}$ 时,CIoU 损失函数中对长宽比的惩罚项会退化为 0,导致惩罚失效,最终预测框无法拟合真实框。对公式(10)求偏导可得:

$$\frac{\partial v}{\partial w} = \frac{8}{\pi} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right) * \frac{h}{w^2 + h^2} \quad (11)$$

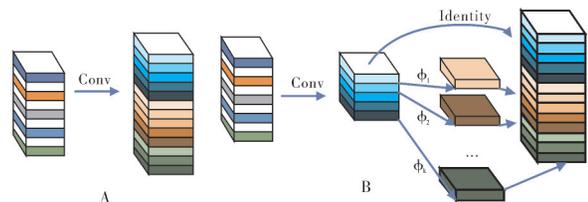
$$\frac{\partial v}{\partial h} = -\frac{8}{\pi} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right) * \frac{w}{w^2 + h^2} \quad (12)$$

由式(11)和(12)可知,在回归过程中, h 和 w 的变化相反,无法使二者同时变小。基于此情况,引入 EIoU 损失函数改进定位回归损失。EIoU 损失函数在 CIoU 损失函数的基础上将对宽高比的惩罚改为采用对目标框和预测框的宽高之差分别惩罚,解决了形状相似时惩罚项退化和无法同时最小化的问题,提高了边界框的回归精度。其定义如公式(13)所示。

$$L_{\text{EIoU}} = L_{\text{IoU}} + L_{\text{dis}} + L_{\text{asp}} \\ = 1 - \text{IoU} + \frac{\rho^2(b, b^{gt})}{c^2} + \frac{\rho^2(w, w^{gt})}{C_w^2} + \frac{\rho^2(h, h^{gt})}{C_h^2} \quad (13)$$

其中, C_w 、 C_h 分别代表能同时覆盖两框的最小闭包区域的高和宽。

3) 引入 Ghost 模块。经上述方法改进后,较原 YOLOv5s 模型识别准确率提高,但同时增加了参数量。在原始模型中使用了大量的卷积核去完成采样、融合等操作,普通卷积输出的特征图中包含了很多冗余特征^[18],为减少模型参数量,使模型能够部署在计算资源有限的硬件设备上,引入 Ghost 模块替换 Neck 中原始普通卷积。Ghost 模块将卷积分为 2 部分,一部分利用常规卷积生成一定数量的特征图,另一部分通过线性运算补全缺失的特征图,从减少冗余特征的角度出发,在不改变输出特征图大小的情况下降低参数量和计算复杂度,结构如图 4 所示。虽然改进后模型实现了轻量化,但由于去除冗余特征图导致精度下降,推理速度变慢等,因此,后续针对模型精度、参数量、推理速度进行消融试验。改进后的网络结构设计如图 5 所示。



A: 常规卷积 Regular convolutional layer; B: Ghost 模块 Ghost module.

图 4 常规卷积层和 Ghost 模块
Fig.4 Regular convolutional layer and Ghost module

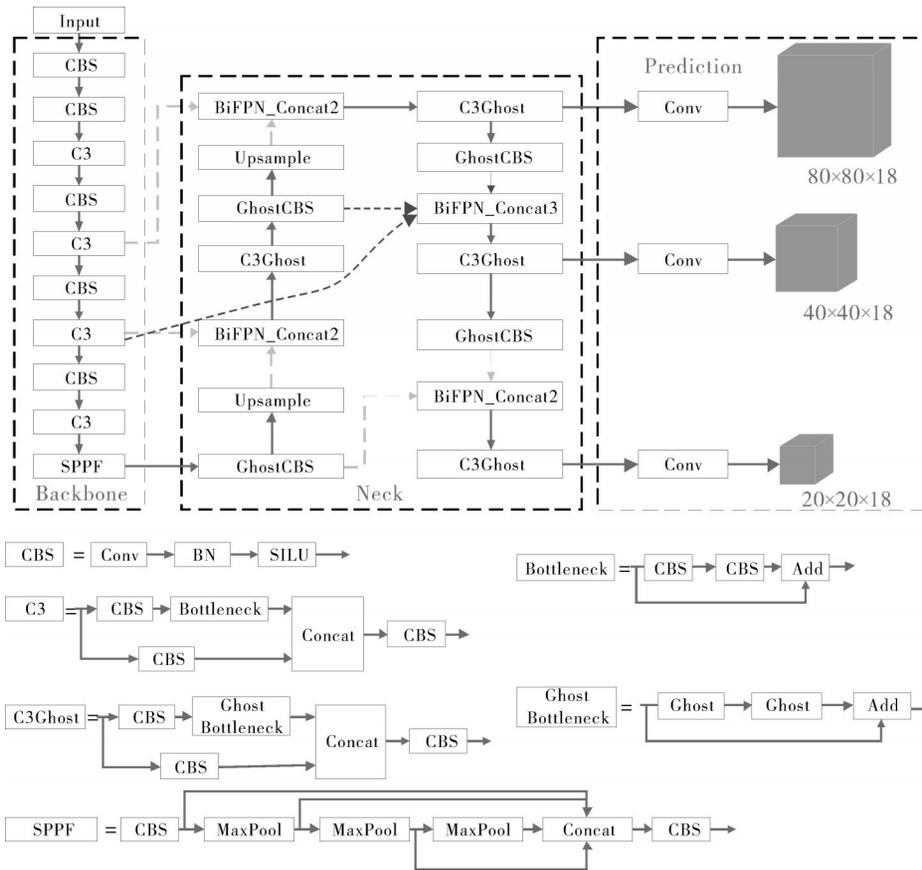


图5 改进后的YOLOv5s结构图

Fig.5 Improved YOLOv5s structure diagram

1.4 试验环境及流程

采用 Windows10 操作系统,计算机内存为 64 G,处理器型号为 Intel(R) Core(TM) i9-10900K CPU @ 3.70GHz 3.70 GHz,显卡型号为 GPU NVIDIA GeForce RTX3060,深度学习框架使用 Pytorch 1.10.0, CUDA 10.2, CUDNN 8.3.3。初始学习率为 0.01,动量参数为 0.937,最终学习率为 0.001,批处理大小 BatchSize 为 32,总迭代次数 epoch 为 300 次,并采用 epoch 为 3、动量参数为 0.8 的 warm-up 方法预热学习率。具体试验流程如图 6 所示。步骤 1,利用工业相机采集甘蔗图像后进行分类、标注处理,并按比例划分好训练集、测试集和验证集;步骤 2,采用 Python 编程语言,利用 Pytorch 深度学习框架构建并改进 YOLOv5s 模型;步骤 3,选取合适的超参数初始化模型并对模型进行训练;步骤 4,利用训练好的权重文件在测试集上测试模型效果;步骤 5,将最终模型与其他模型进行多组横向对比试验。

1.5 评价指标

为系统且客观地评价改进后模型与原模型及其他模型的性能,试验采用准确率(precision, P)、召回

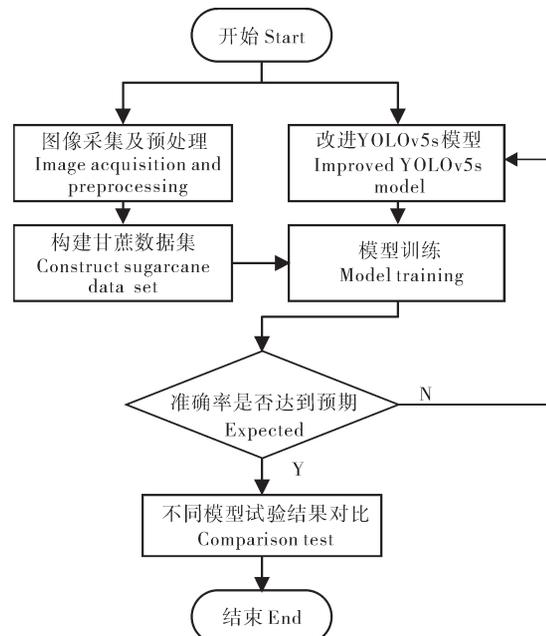


图6 试验流程图

Fig.6 Experiment flow chart

率(recall, R)、平均精度(average precision, P_A)、参数量、计算量(GFLOPs)、模型大小、推理速度

(speed)作为模型的评价标准。其中,推理速度(speed)指对单张甘蔗图像的检测时间。具体计算公式如下:

$$P = \frac{T_p}{T_p + F_p} \times 100\% \quad (14)$$

$$R = \frac{T_p}{T_p + F_N} \times 100\% \quad (15)$$

$$P_A = \int_0^1 PRdR \times 100\% \quad (16)$$

其中, T_p 为真正样本, F_p 为虚假正样本, F_N 为虚假负样本。

2 结果与分析

2.1 采用不同特征融合手段的改进效果

从表1可知,采用add操作实现的特征融合手段改进Neck结构使模型准确率 P 下降了0.2个百分点,

召回率 R 、平均精度 P_A 均提高了1.3个百分点;采用concat操作实现的特征融合改进Neck结构使得模型准确率 P 、召回率 R 和平均精度 P_A 分别提高了0.9、1.3、1.4个百分点。因此,无论用哪种手段实现的跨层加权连接均能使模型平均精度得到提升,但二者使模型的参数量和复杂度有不同程度的增加,add操作实现的特征融合相比于concat操作在模型参数量上高出0.09 Mb,通过分析模型构建过程可以解释这种现象。在模型构建时,需要将第12层、16层、22层进行特征融合,由于add操作需要保证不同层级特征图通道数相同,所以融合之前需加入 1×1 卷积调整通道数,使得本应参数量较小的add操作比concat操作实现的特征融合多出一部分参数量。综合考虑精度和参数量等因素,采用concat操作实现可学习的跨层级连接改进Neck结构。

表1 不同融合手段的试验结果

Table 1 Comparison experiments of fusion means

模型 Model	参数量/Mb Parameters	计算量 GFLOPs	准确率/% Precision	召回率/% Recall	平均精度/% Average precision
YOLOv5s	7.01	15.8	95.8	96.2	96.4
BiFPN_add	7.16	16.4	95.6	97.5	97.7
BiFPN_concat	7.07	16.0	96.7	97.5	97.8

2.2 YOLOv5s 消融试验

为验证上述改进后模型在复杂背景下甘蔗茎节识别过程中的优化效果,进行了一系列消融试验。首先通过添加学习权重的跨层连接改进颈部结构,然后将YOLOv5s的目标置信度损失和边框定位损失分别替换为Focal loss和EIoU损失函数。训练过程采用同样的超参数,试验结果见表2。将颈部改进后,使网络将浅层的边缘信息进一步与提取后的深层语义信息相结合,模型对茎节的检测也更加准确,

准确率 P 提高0.9个百分点,平均精度 P_A 提升1.4个百分点,参数量仅增加0.06 Mb。引入EIoU损失函数计算边界回归的损失函数,改变了对预测框长和宽的惩罚方法,提高了模型边界框的定位精度,在不增加参数量的情况下,较原模型准确率 P 提高1.2个百分点,平均精度 P_A 提高1.4个百分点。此外,引入Focal loss焦点损失函数后,解决正负样本分布不均匀问题,使准确率 P 提高了0.8个百分点。由表2可知,由于几种改进措施都均未大幅增加模型参数量和复杂度,故单张图片推理时间几乎没有改变。

表2 YOLOv5s 消融试验

Table 2 YOLOv5s ablation experiment

序列 Order number	模型 Model	参数量/Mb Parameters	模型大小/Mb Size	准确率/% Precision	召回率/% Recall	平均精度/% Average precision	推理时间/ms Reasoning time
0	YOLOv5s	7.01	14.5	95.8	96.2	96.4	14
1	+BiFPN_concat	7.07	14.6	96.7	97.5	97.8	15
2	+EIoU	7.01	14.6	97.0	97.1	97.8	14
3	+Focal loss	7.01	14.5	96.6	97.2	97.3	14
4	+1+2+3	7.07	14.6	97.2	98.0	98.8	15

2.3 引入Ghost的消融试验

将改进后的YOLOv5s模型采用Ghost模块替换不同位置的卷积进行训练并测试,对比不同参数量

下的模型精度,评估模型性能,结果如表3所示,其中Baseline为经过表2中的序列1、2、3分别改进后的YOLOv5s模型。经比较可以看出,不论采用Ghost

改进骨干网络还是改进颈部,或者全部改进,均会使模型参数量和模型大小减小,但会导致模型的平均精度值有不同程度的降低。其中,对颈部替换的参数量较原模型减少1.34 Mb,模型大小减小3.1 Mb,准确率 P 提高1.3个百分点,平均精度 P_A 增长了1.4

个百分点。虽然单张推理时间增加了2.9 ms,但仍满足切种工作对于图像处理时间的要求。因此,采用Ghost模块对颈部Neck进行轻量化改进,能够很好地兼顾检测精度和模型参数量大小。

表3 Ghost模块替换结果比较

Table 3 Comparison of Ghost module replacement results

模型 Model	参数量/Mb Parameters	模型大小/Mb Size	准确率/% Precision	召回率/% Recall	平均精度/% Average precision	推理时间/ms Reasoning time
YOLOv5s	7.01	14.5	95.8	96.2	96.4	14.0
Baseline	7.07	14.6	97.2	98.0	98.8	15.0
替换 Backbone Replace Backbone	4.18	8.9	96.3	97.0	96.6	15.2
替换 Neck Replace Neck	5.67	11.4	97.1	97.9	97.8	16.9
全部替换 All repalce	3.67	7.9	96.3	97.4	97.3	16.0

2.4 与不同主流检测模型的对比

采用同一甘蔗数据集在同等硬件资源下对Faster R-CNN、YOLOv4-tiny、YOLOv3进行对比试验。由表4可知,相比于YOLOv3和YOLOv4-tiny,双阶段算法Faster R-CNN在平均精度上性能优于YOLOv3和YOLOv4-tiny,但单张推理时间远远大于单

阶段算法YOLO。改进后的YOLOv5s模型相较于YOLOv3和YOLOv4-tiny,模型尺寸更小,平均精度更高,更适宜部署在计算资源有限的嵌入式硬件上。相比于原模型YOLOv5s,由于提升了网络特征提取能力以及改进了损失函数,使得改进后模型在杂乱背景下表现更好,平均精度更高。

表4 不同模型检测结果对比

Table 4 Comparison of detection results of different models

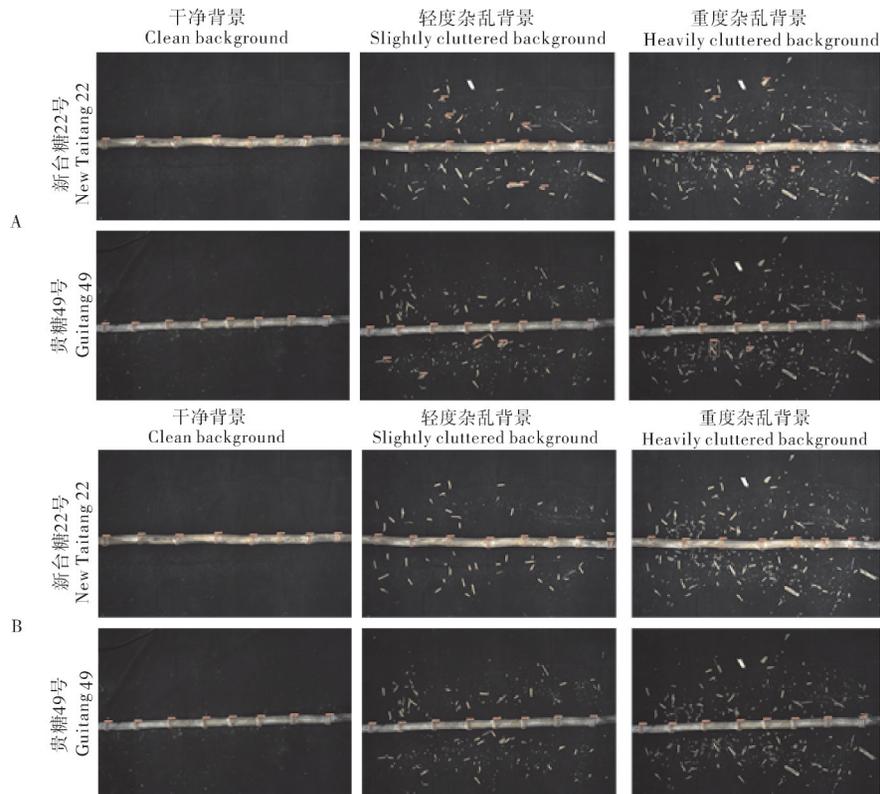
模型 Model	模型大小/Mb Size	平均精度/% Average precision	推理时间/ms Reasoning time
YOLOv3	235.00	93.90	35.0
YOLOv4-tiny	246.17	91.54	29.0
Faster R-CNN	460.12	94.56	68.0
YOLOv5s	14.50	96.40	14.0
改进后 YOLOv5s Improved YOLOv5s	11.40	97.80	16.9

图7为改进前后YOLOv5s在不同杂乱程度背景下茎节识别的效果图,其中图7A为改进前识别效果,图7B为改进后识别效果。从图7A可以看出,当背景逐渐变得杂乱,未改进前模型存在着错检等问题。改进后YOLOv5s模型不但对干净背景下茎节检测效果准确率提高,而且随着背景逐渐杂乱,对茎节的检测效果仍较好。

3 讨论

本研究针对甘蔗智能切种机作业过程中背景杂乱导致茎节识别困难等问题,提出了一种基于改进YOLOv5模型的茎节识别方法,通过将主干网络与颈部同尺寸特征图跨层加权连接提高特征提取能力,减少杂乱背景造成的精度下降;引入Focal loss替

换交叉熵损失函数解决正负样本比例失衡问题,并采用EIoU损失函数计算预测框回归损失提高定位精度,最终结合Ghost模块在保证模型精度的同时对模型进行轻量化改进。改进后的模型在甘蔗数据集上的准确率达97.1%,较原模型提高了1.3个百分点,平均精度为97.80%,模型大小仅11.40 Mb,单张推理时间为16.9 ms。试验结果表明,在杂乱背景下,改进后的算法能够在兼顾模型大小的同时保证较高精度,为复杂工况下的防伤芽切种工作提供精确的定位信息。由于物距较大导致单个相机无法较好地完成整根甘蔗图像的清晰采集,后续考虑减小物距并增加多组相机进行拍摄,通过图像拼接得到整根甘蔗图像,并进一步将模型部署在嵌入式设备上,最终完成真实环境下的切割任务。



A:改进前识别效果图 Recognition effect before improvement; B:改进后识别效果图 Recognition effect after improvement.

图7 不同程度杂乱背景下检测效果图

Fig.7 Detection effect under different degrees of clutter background

参考文献 References

- [1] 赵欢欢,孙东磊,卢颖林,等.我国甘蔗研究近况、热点分析与展望[J].甘蔗糖业,2018(3):65-70.ZHAO H H,SUN D L,LU Y L, et al.Status, hotspot analysis and outlook of sugarcane research in China[J].Sugarcane and canesugar,2018(3):65-70(in Chinese with English abstract).
- [2] ZHOU D Q,FAN Y L,DENG G R, et al.Research on algorithm of sugarcane nodes identification based on machine vision [C]//2019 Nicograph International (NicoInt).Yangling,China.[S.I.]:IEEE,2019:111-116.
- [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/OL].Computers and electronics in agriculture, 2021, 182: 105994 [2022-06-20].https://doi.org/10.1016/j.compag.2021.105994Get rights and content.
- [4] YANG R,LI J,LIU Q, et al.Gradient-based method for the identification of multi-nodes in sugarcane[J].Information processing in agriculture,2020,7(4):491-499.
- [5] NARE B,TEWARI V K,KUMAR CHANDEL A, et al.A mechatronically integrated autonomous seed material generation system for sugarcane: a crop of industrial significance[J].Industrial crops and products,2019,128:1-12.
- [6] 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 Y Y,HE Y L,WANG Y F, 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 with English abstract).
- [7] 李尚平,李向辉,张可,等.改进YOLOv3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 S P,LI X H,ZHANG K, et al.Increasing the real-time dynamic identification efficiency of sugarcane nodes by improved YOLOv3 network[J].Transactions of the CSAE,2019,35(23):185-191(in Chinese with English abstract).
- [8] 黄彤斌,黄河清,李震,等.基于YOLOv5改进模型的柑橘果实识别方法[J].华中农业大学学报,2022,41(4):170-177.HUANG T B,HUANG H Q,LI Z, et al.Citrus fruit recognition method based on the improved model of YOLOv5[J].Journal of Huazhong Agricultural University,2022,41(4):170-177(in Chinese with English abstract).
- [9] 赵德安,吴任迪,刘晓洋,等.基于YOLO深度卷积神经网络的复杂背景下机器人采摘苹果定位[J].农业工程学报,2019,35(3):164-173.ZHAO D A,WU R D,LIU X Y, et al.Apple positioning based on YOLO deep convolutional neural network for picking robot in complex background[J].Transactions of the CSAE,2019,35(3):164-173(in Chinese with English abstract).
- [10] 顾伟,王巧华,李庆旭,等.基于改进SSD的棉种破损检测[J].华中农业大学学报,2021,40(3):278-285.GU W,WANG Q H,LI Q X, et al.Improved SSD based detection of damaged cottonseed[J].Journal of Huazhong Agricultural University,2021,40(3):278-285(in Chinese with English abstract).
- [11] BOCHKOVSKIY A,WANG C Y,LIAO H Y M.YOLO4: op-

- timal speed and accuracy of object detection[J/OL].arXiv, 2020, 2004.10934 [2022-06-20]. <https://doi.org/10.48550/arXiv.2004.10934>.
- [12] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection[J/OL].arXiv, 2017, 1612.03144 [2022-06-20].<https://doi.org/10.48550/arXiv.1612.03144>.
- [13] LI H, XIONG P, AN J, et al. Pyramid attention network for semantic segmentation[J/OL].arXiv, 2018, 1805.10180 [2022-06-20].<https://doi.org/10.48550/arXiv.1805.10180>.
- [14] WANG W H, XIE E Z, SONG X G, et al. Efficient and accurate arbitrary-shaped text detection with pixel aggregation network[J/OL]. arXiv, 2020, 1908.05900 [2022-06-20]. <https://doi.org/10.48550/arXiv.1908.05900>.
- [15] TAN M, PANG R, LE Q V. Efficientdet: scalable and efficient object detection[J/OL].arXiv, 2020, 1911.09070 [2022-06-20]. <https://doi.org/10.48550/arXiv.1911.09070>.
- [16] ZHENG Z, WANG P, LIU W, et al. Distance-IoU loss: faster and better learning for bounding box regression[J/OL]. arXiv, 2020, 1911.08287 [2022-06-20].<https://doi.org/10.48550/arXiv.1911.08287>.
- [17] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[J]. IEEE transactions on pattern analysis and machine intelligence, 2020, 42(2):318-327.
- [18] ZHANG Y F, REN W Q, ZHANG Z, et al. Focal and efficient IOU loss for accurate bounding box regression[J]. Neurocomputing, 2022, 506:146-157.
- [19] HHAN K, WANG Y, TIAN Q, et al. Ghostnet: more features from cheap operations[J/OL].arXiv, 2020, 1911.11907 [2022-06-20].<https://doi.org/10.48550/arXiv.1911.11907>.

Sugarcane stem node recognition method based on improved YOLOv5

ZHAO Wenbo¹, ZHOU Deqiang^{1,2}, DENG Ganran^{3,4}, HE Fenguang^{3,4},
ZHU Qi¹, WEI Lijiao^{3,4}, NIU Zhaojun^{3,4}

1. College of Mechanical Engineering, Jiangnan University, Wuxi 214122, China;

2. Jiangsu Key Laboratory of Advanced Food Manufacturing Equipment & Technology, Wuxi 214122, China;

3. Institute of Agricultural Machinery, Chinese Academy of Tropical Agricultural Sciences, Zhanjiang 524088, China;

4. Ministry of Agriculture and Rural Affairs Key Laboratory of Agricultural Equipment for Tropical Crops, Zhanjiang 524091, China

Abstract As one of the important raw materials for sugar production, sugar cane extracted is not only a necessity but also belongs to national strategic reserve materials. At present, the sugarcane industry as a whole is inefficient, especially in terms of machine seeding, and the efficiency of manual seed cutting is low. The developed methods for sugarcane stem node identification are all based on clean background conditions, but the working environment of agricultural machinery is disgusting. A large amount of debris, miscellaneous leaves and other dirt produced by seed cutting will lead to background pollution in image acquisition area and reduce the recognition ability of algorithm. Therefore, a sugarcane stem node identification method is put forward based on improving YOLOv5 to solve the problems of low recognition accuracy of sugarcane stem node under complex background. To optimize the neck structure and enhance the information fusion capability among different levels by using cross-level connection. At the same time, the model loss function is improved. On the one hand, EIou loss is introduced to replace original CIou loss to improve the precision of boundary box regression; on the other hand, Focal loss function is used to replace the cross-entropic loss function to solve the problem of unbalanced proportion of positive and negative samples. Finally, Ghost module is introduced to lightweight network model. The experimental results show that compared with the original model, the average precision value of the model proposed in this study is increased to 97.80%, the single detection time is 16.9 ms and the memory of the model is only 11.4 Mb, which realizes the identification of sugarcane stem joints in different chaotic scenes and reduces the impact of background chaos when cutting injurious buds.

Keywords sugarcane; stem node identification; cluttered background; YOLOv5; cross-layer weighted connection; loss function; precise cutting

(责任编辑:陆文昌)