

DOI:10.19651/j.cnki.emt.2417256

一种融合几何和深度学习的甘蔗高度测量方法^{*}

蒋正中¹ 杨鸿波¹ 杨明浩² 刘安琪¹

(1. 北京信息科技大学自动化学院 北京 100192; 2. 中国科学院自动化研究所 北京 100190)

摘要: 甘蔗是一种具有全球经济意义的重要经济作物,其株高与产量的关系非常密切,而确定甘蔗的高度表征的传统方法依赖于人工检测,这是劳动密集型与耗时的。因此,本文收集了甘蔗作物在不同场景下的图像数据集,提出了轻量化改进的PSPNet和YOLOv5s模型,用于分割甘蔗主体与检测甘蔗顶梢位置。并且开发了一套自动测量机器人,其部署了改进后的模型来实时推理深度摄像头采集的图像,进而,设计云台自动化操作系统。最后,获取甘蔗位置与深度信息来计算甘蔗高度。实验结果表明,甘蔗高度测量方法的平均绝对误差小于2.4 cm、准确率高于97.61%、成功率高于93%、检测时间低于13.2 s,并且提出的甘蔗主干分割与顶梢检测模型参数大幅减小,精度分别保持在87%与97%左右。

关键词: 甘蔗高度测量机器人;甘蔗图像分割;甘蔗检测;PSPNet;YOLOv5

中图分类号: TP391.4; TN41 **文献标识码:** A **国家标准学科分类代码:** 520.6

Research on sugarcane automated characterization measurement robot based on image segmentation and detection

Jiang Zhengzhong¹ Yang Hongbo¹ Yang Minghao² Liu Anqi¹

(1. College of Automation, Beijing Information Science and Technology University, Beijing 100192, China; 2. Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China)

Abstract: Sugarcane is a significant and globally important economic crop, with a close correlation between plant height and yield. Traditional methods for determining sugarcane height characterization rely on labor-intensive and time-consuming manual detection. Therefore, this study collected image datasets of sugarcane crops in various scenarios and proposed lightweight enhanced PSPNet and YOLOv5s models for segmenting sugarcane bodies and detecting the tip positions. Additionally, an automated measurement robot was developed, deploying the improved models for real-time inference on images captured by a depth camera and implementing a gimbal automation system. Finally, sugarcane positions and depth information were utilized to calculate sugarcane height. The experimental results show that the average absolute error of the sugarcane height measurement method is less than 2.4 cm, the accuracy is higher than 97.61%, the success rate is higher than 93%, and the detection time is less than 13.2 s. In addition, the proposed sugarcane trunk segmentation and top detection model parameters are significantly reduced, with accuracies maintained at around 87% and 97%, respectively.

Keywords: sugarcane height measurement robot; sugarcane image segmentation; sugarcane detection; PSPNet; YOLOv5

0 引言

甘蔗是世界上最重要的经济作物之一,在全球农业和糖生产中发挥着举足轻重的作用^[1]。甘蔗收获的最佳时机是影响作物产量和提取糖质量的关键因素。然而,传统上,确定甘蔗的确切成熟阶段依赖于农业专家的人工检查,这是一个耗时且劳动密集型的过程,容易出现主观性和错

误^[2]。长期以来,这种低效率一直要求创新的解决方案来自动评估甘蔗成熟度,从而改善农业实践并提高作物产量。然而,通过表型性状可以直观、便捷地鉴定出育种目标,甘蔗的表型研究表明甘蔗的株高和直径与产量的关系非常密切^[3],所以,甘蔗株高的测量对于挑选优质甘蔗苗,提升甘蔗产糖量具有重大意义。

近年来,深度学习技术的出现,特别是卷积神经网络

收稿日期:2024-11-01

*基金项目:广西省科技计划重点研发项目(AB24010164,2024AB08020)、国家自然科学基金京津冀基础研究合作专项(F2024205028)资助

(convolutional neural network, CNN)以及图像分类技术的进步,为彻底改变表征测量的方式提供了一个令人兴奋的机会。通过利用计算机视觉和深度学习的力量,现在可以开发能够直接检测甘蔗^[4-6],并且高精度测量出其长度的强大系统。该项目探索尖端技术与农业的融合,以创建一个自动化的甘蔗高度表征测量系统。

目前,有学者提出利用遥感和无人机图像结合图像技术,大概判断甘蔗种植地区的成熟程度。如Khuimphukhieo 等^[7]提出利用三维成像技术,对甘蔗茎秆的表型特征进行定量化测量。研究团队使用激光扫描仪获取甘蔗茎秆的三维形状数据,并通过图像处理和分析方法,提取茎秆的直径、长度、曲率等参数。Simões 等^[8]提出利用无人机获取的多光谱图像,通过分析不同波段的反射特征,实现对甘蔗的高通量表型测量,其研究团队开发了一种基于图像处理和机器学习的方法,可以自动提取甘蔗的生长指标,包括高度、叶面积、叶绿素含量等。这种高效的测量方法为大规模甘蔗表型研究和育种提供了支持。尽管,以上技术可以获取总体甘蔗田地的培育程度,但很难精确获取甘蔗的生长指标。

最近,有很多学者研究目标检测和分割技术以获取更精细的甘蔗生长表征,从而实现更加智能化的培育模式。Sindhe 等^[9]提出使用YOLOv5 模型来识别甘蔗芽,从而可以实现更加精准的甘蔗种植过程,不过实验准确率较低。Zhang 等^[10]提出一种基于改进YOLOv8 模型来准确识别甘蔗芽,在实验结果中,该方法实现了 97.42% 的识别准确率。Zhou 等^[11]提出了一种结合YOLOv3 模型与传统计算机视觉的新方法,并搭建了自动化实验平台,从而实现甘蔗茎节点的快速识别。尽管,实验结果表明,可以实现近乎 100% 的识别率,但是并未验证该方法是否适用于甘蔗田地现场。最近,以上研究表明,采用目标检测技术可以实现精确的甘蔗表征识别。然而,在真实的甘蔗培育田中,往往面临更复杂的环境干扰挑战,其中,甘蔗存在生长倾角,这会导致甘蔗间会互相重叠,并且甘蔗枝叶也会部分遮挡甘蔗主体,这将会导致识别混乱情况。

语义分割技术可以通过算法自动识别图像中的目标区域,并将这些区域划分出来。由于传统的分割方法缺乏泛化能力,如阈值分割和基于区域分割等,在实际应用中难以拟合模型。深度学习方法因其卓越的性能而被广泛用于农业视觉领域。Peng 等^[12]提出一种基于全卷积神经网络的语义分割算法,在野外对荔枝进行语义分割。实验结果表明,该方法 0.765 的平均交并比,达到了较好的检测效果。Qi 等^[13]提出结合YOLOv5 检测模型和SPNet 语义分割模型,实现在野外对荔枝主茎的检测与分割,从而计算得到主茎拾取点的准确位置信息,使机器人可以精准采集荔枝。该研究表明,结合目标检测和语义分割技术可以在复杂的野外环境中,减少环境干扰,更好的完成农业视觉任务。

因此,本文的主要目标是利用图像检测能力,设计一个

能够准确地测量甘蔗高度表征的机器人系统。首先,本文设计了基于改进 PSPNet 网络的甘蔗主干分割模型,用来分割采集来的图像中的甘蔗主体,再设计了甘蔗顶梢检测的轻量化 YOLO 网络算法,进而通过使用庞大而多样化的高分辨率图像数据集训练模型,让其捕捉处于不同生长和成熟阶段的甘蔗顶梢位置信息,再通过自主开发的机器人自动化测量系统去测量出甘蔗高度,最后通过实验验证了系统的有效性。这项研究为甘蔗产业带来了巨大的希望,因为它有可能显着改善作物管理和资源分配,降低劳动力成本,并最终提高整体产量和糖质量。

1 问题描述

甘蔗高度表征测量,可以通过深度信息直接测量和机械间接式测量,前者仅需要获得完整的甘蔗株体深度图像信息,即可计算出甘蔗高度,但是这种方法需要考虑深度摄像头的感知角度和位置,属于半自动检测方法。后者则是通过云台操作摄像头,结合图像检测技术自动找到甘蔗的顶梢和底端,通过云台角度间接计算甘蔗高度,这种方法不需要特定拍摄场景,自动化程度高,因此,本文着重研究机械间接式测量方法对甘蔗高度表征的自动化测量。

2 基于分割与检测的甘蔗高度表征测量方法

2.1 基于改进 PAPNet 的甘蔗主干分割模型

最初, PSPNet^[14-16] 主要由骨干网络(CNN)和金字塔池化模块组成,金字塔池化模块增强模型对多尺度上下文信息的感知,从而获取图像的全局和局部信息,并将这些信息结合起来,帮助模型在复杂场景中进行更准确的分割。由于本文的研究环境是真实的甘蔗种植棚中,存在杂草和光照不均匀等复杂情况,因此,使用PSPNet 作为甘蔗主干分割实验的基线模型。

随着网络深度的增加,网络能够更好地提取复杂特征,但也带来了一些副作用,如梯度消失和网络退化。通过在ResNet 中加入快速连接分支,这些问题得到了缓解。最初,PSPNet 使用 ResNet50 作为骨干网络。ResNet50 的 Block 包含 3 层卷积层;两层 1×1 卷积层和一层 3×3 卷积层。在该模块中,特征图的维度通过 1×1 卷积层巧妙地减少或扩展,从而使 3×3 卷积层的滤波器数量不受上一层输入的影响,并影响下一层。最初, 1×1 卷积层用于减少中间 3×3 卷积层的维度,从而减少计算量,之后使用另一层 1×1 卷积层恢复维度。这样不仅保持了模型的精度,还减少了网络参数和计算量,节省了计算时间。金字塔池化模块作为PSPNet 的核心组件,其将特征层划分为多个不同大小的区域。每个区域独立生成不同层次的特征图,最终通过拼接这些图像,从而减少不同区域之间信息的损失。该金字塔池化模块是一个四级模块,池化核分别为 1×1 、 2×2 、 3×3 和 6×6 。输入特征图通过 1×1 池化核在金字塔池化模块中进行全局池化,生成一个单一特征图。 2×2

池化核将特征图划分为 2×2 的子区域，并对每个子区域进行池化。类似地， 3×3 和 6×6 池化核将输入特征图分别划分为 3×3 和 6×6 的子区域，并进行池化。最后，通过线性插值扩展输出特征图，使其与最初输入的特征图具有相同的尺寸，并通过拼接获得最终的输出特征图。

根据数据集的特点，PSPNet 进行了改进和优化，构建了基于改进的 PSPNet 模型的甘蔗主干语义分割网络。如图 1 所示，该图展示了在自然环境中用于甘蔗主干的改进

PSPNet。其中 CONV 表示特征图的卷积操作；POOL 表示特征图的池化操作，模型结构主要由骨干特征提取网络和金字塔池化模块组成。前者使用 MobileNetV2 作为模型的骨干网络，在缩小模型规模的同时提高了模型的准确率。在骨干特征提取网络中引入 CBAM 以抑制无用的特征提取，同时增强了对重要特征的关注。改进部分为主干网络，使用改进的主干网络替换原主干网络，在缩小模型规模的同时提高了模型的准确率。

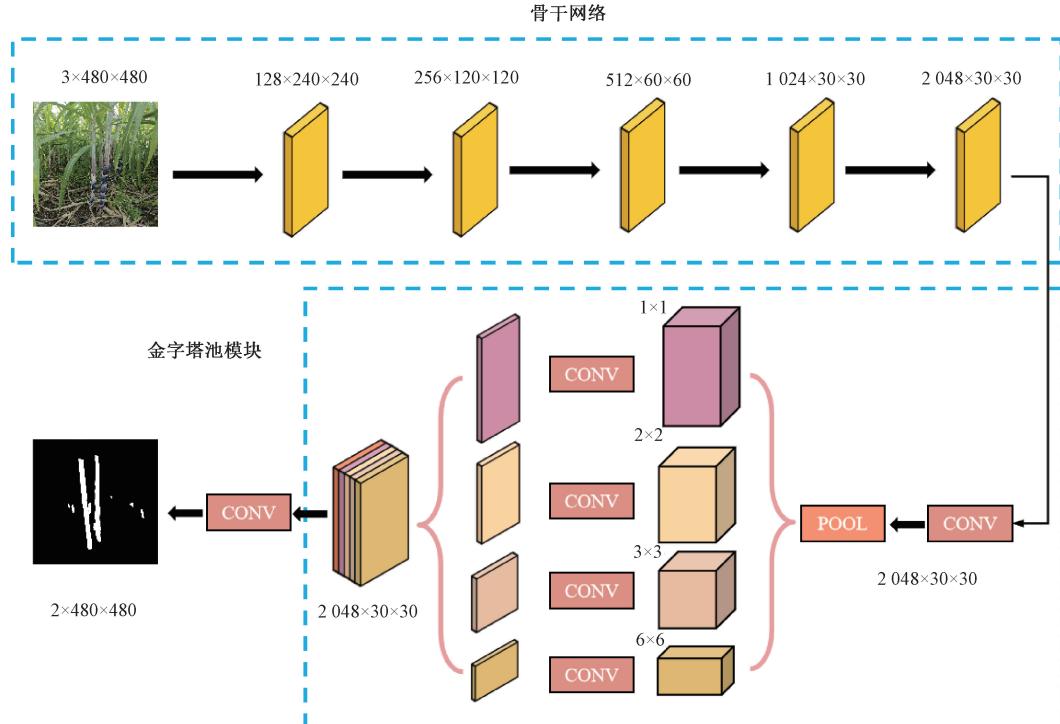


图 1 基于改进的 PSPNet 的甘蔗分割模型结构

Fig. 1 Structure of a sugarcane segmentation model based on an improved PSPNet

1) 基于 MobileNetV2 的骨干网络

原始 PSPNet 模型中常用的主要特征提取网络是 ResNet50。在网络模型的训练过程中，大量的卷积计算减慢了模型参数的更新速度，结果影响了特征提取的速度。在嵌入式设备的计算环境中，Sandler 等^[17-19] 选择了 MobileNetV2 作为骨干网络。MobileNetV1 中使用了深度可分离卷积。与标准卷积相比，当权重参数数量相同时，模型的参数大大减少，网络的计算速度得到提高。MobileNetV2 提出了具有线性瓶颈的倒置残差，进一步降低了模型推理过程中的内存占用，更适合嵌入式硬件设计。

MobileNetV2 使用深度可分离卷积而不是标准卷积从特征图中提取特征。深度可分离卷积包括深度卷积和逐点卷积。深度卷积是一个有通道号的卷积核，只负责一个通道，通过一个通道号和卷积核大小为 3×3 的卷积层，将输入特征图卷积在特征图的单个通道上。逐点卷积由

$1\times 1\times M$ 卷积核组成。对上一步的特征图进行加权，并在深度方向上进行组合，生成新的特征图。与传统的标准卷积相比，深度可分离卷积的参数和操作成本大大降低。在高维空间中，通过在卷积层上执行 ReLU 激活函数的操作可以有效地增加特征的非线性表达式，但在低维卷积层上执行 ReLU 激活函数的操作时，特征图的特征信息会丢失。为了防止信息的丢失，提出了一种线性瓶颈结构。在计算低维卷积层时，ReLU 激活函数不用于非线性变换，从而减少了信息损失。在构建深度网络结构时，提出了瓶颈结构，以缓解梯度消失和网络退化的问题。与标准卷积相比，MobileNetV2 由于深度可分离卷积，通道更少。它是瓶颈通道，然后对特征图进行复杂化以提取特征，导致特征信息的丢失。因此，提出了一种倒置残差结构。特征图的通道由卷积核 1×1 的卷积层扩展，然后由卷积核 3×3 的深度卷积层提取特征，最后通过卷积核 1×1 的逐点卷积对通道进行压缩。

2) 骨干特征提取网络嵌入 CBAM

为了在复杂环境下提取更多的甘蔗主干特征,在分段网络中增加注意力机制是提高网络提取特征能力的方法之一^[20]。本研究的研究对象是甘蔗的主干,它与周围环境具有很高的相似性。因此,在改进的主干提取网络中引入卷积块注意力模块(CBAM)^[21-22],提高了红甘蔗特征的提取能力。

在本研究中,MobileNetV2 作为模型的骨干网络,它包含不同的瓶颈结构模块。它们由相同深度可分离的卷积层组成,层间的跳跃结构克服了深度学习性能中的泛化和梯度爆炸问题,从而更好地学习了深度特征信息。为了提高其特殊的提取能力,在瓶颈结构的第 5 层和第 7 层之后引入了 CBAM。

CBAM 于 2018 年提出。它使用全局平均池化和最大池化来增强目标区域的特征信息。该模块可以增强通道

上的图像特征,提高重要特征的显著性。该模块由两个独立的模块组成:关注通道的通道注意力模块(CAM)和关注特征图空间的空间注意力模块(SAM)。当特征图输入到模块中时,通道注意力模块通过全局最大池化和全局平均池化,从输入特征矩阵中获取两个一维向量,并根据通道维度合并向量,生成通道压缩权重矩阵。将通道压缩权重矩阵与输入特征矩阵相乘,得到通道注意力模块和输出特征图。输出特征映射进入空间注意力模块后,分别通过全局最大池化和全局平均池化得到两个二维向量。在卷积层中通过拼接将两个二维向量卷积,生成空间压缩的权重矩阵。将空间压缩的权重矩阵与输入特征映射相乘,得到空间注意力模块的输出特征映射。CBAM 结构如图 2 所示,其中 Conv 表示卷积层,MLP 表示多层感知器,AvgPool 表示特征图的平均池化操作,MaxPool 表示特征图的最大池化操作。

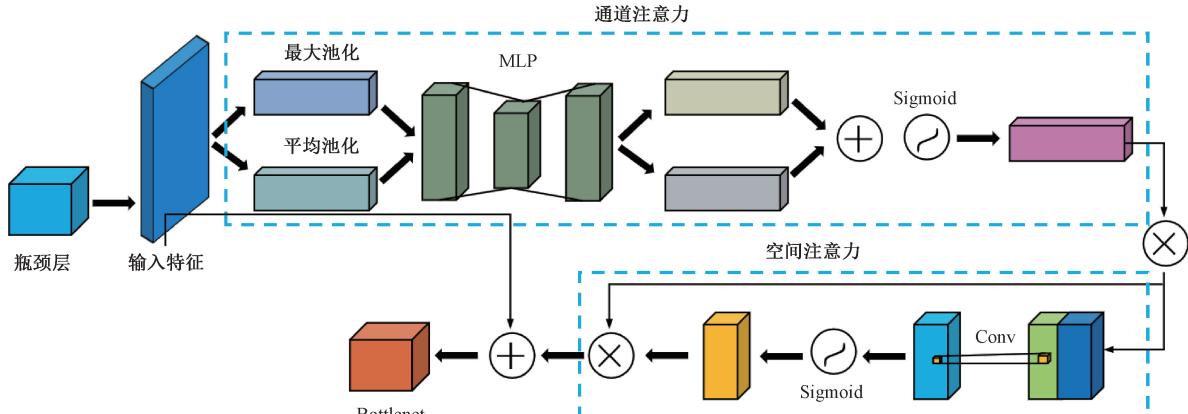


图 2 CBAM 结构

Fig. 2 CBAM structure

2.2 轻量化改进的甘蔗顶梢检测模型

YOLOv5 是一种可以识别数千个对象的算法,当它只用于识别甘蔗顶梢时,它的性能是溢出的,并且有减少参数数量和模型复杂度的空间。

与其他视觉识别算法相比,Ultralytics^[23]发布的 YOLOv5 算法是一种应用更广泛的方法。根据其深度和宽度,YOLOv5 有 4 个版本,分别称为 YOLOv5x、YOLOv5l、YOLOv5m 和 YOLOv5s。YOLOv5s 在 4 个版本中具有最浅的深度和最小的宽度,因此计算和参数的数量最少。其他 3 个版本增加了卷积核的数量和更高的特征宽度,这主要体现在残差块的数量上。针对嵌入式器件的需求,Zhang 等^[24]选择 YOLOv5s 作为甘蔗顶梢检测的基线模型,模型结构如图 3 所示。

1) YOLOv5s 轻量化设计

尽管张福豹等^[25]提出利用 YOLO 系列模型可实现快速有效的目标检测部署,在特定场景下仍需针对性改进。由于甘蔗顶梢在图像中仅占据较小的像素面积。然而,YOLOv5s 网络结构考虑了所有小、中、大目标的检测。在

特征提取过程中,小目标的特征信息主要存在于浅层网络中。为了获得更丰富的语义信息并提高大目标的检测能力,YOLOv5s 的骨干网络需要一个更深的网络结构来提取深层特征。随着网络的加深,输出特征通道的数量呈指数增长。因此,整个网络结构的大部分参数集中在骨干网络的最后几层。此外,在 Neck 结构中,YOLOv5s 采用了 32、16 和 8 倍的下采样,分别生成 20×20 、 40×40 和 80×80 的特征图,以检测大、中、小目标。换句话说,YOLOv5s 在检测甘蔗顶梢时的设计是冗余的。因此,提出了以下针对 YOLOv5s 轻量化改进的策略。首先,移除骨干网络中的 P7 和 P8 层,将下采样次数,采用了较小的感受野。这可以减少模型参数和计算量,同时保留更多细粒度的特征信息,从而提高甘蔗茎顶梢的检测效果。模型轻量化结构如图 4 所示。

2) Ghost-C3 模块

为了进一步减少参数和计算的数量,本研究将 Ghost 模块引入网络,将 Ghost 卷积^[26]替代常规卷积以形成 Ghost-BottleNeck 结构。然后,使用 Ghost-BottleNeck 替

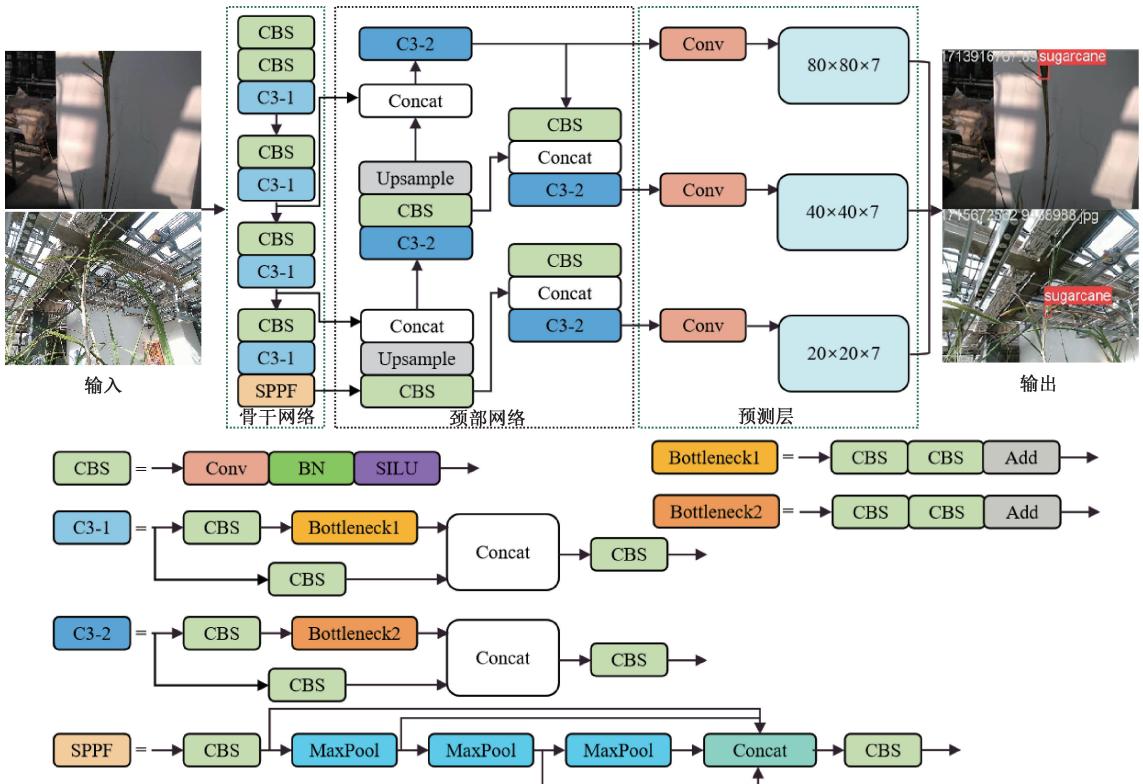


图 3 基于 YOLOv5s 的甘蔗顶梢检测模型结构

Fig. 3 Structure of sugarcane top detection model based on YOLOv5s

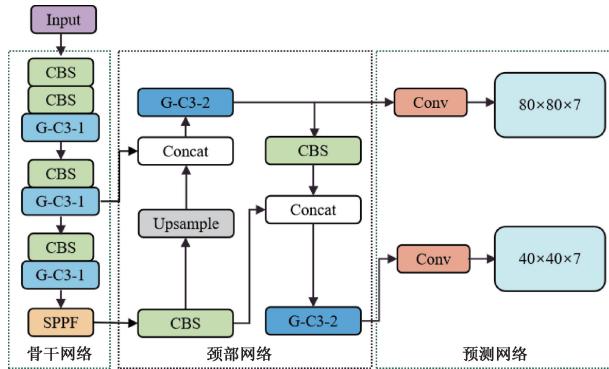


图 4 轻量化设计后的模型结构

Fig. 4 Model structure diagram after lightweight design

换原始 YOLOv5s 网络中的 BottleNeck 模块, 以形成 Ghost-C3 结构, 如图 5 所示。Ghost 卷积使用简单的线性变换来生成幻影特征以替代常规卷积生成的冗余特征。它通过较少的参数和计算获得大量有效的特征图, 使用少量的常规卷积来计算特征图, 并通过廉价的线性变换处理这些特征图以生成大量幻影特征。

最后, 使用身份映射将这两组特征图拼接起来形成新的输出。基于上述分析, 将改进后的 Ghost-C3 模块替换骨干和颈部网络中的所有 C3 模块, 构成 YOLOv5s-L-G 轻量化模型结构。YOLOv5s-L-G 轻量化模型结构, 如图 4 所示。该模型引入 Ghost-C3 模块在不减少影响模型检测准

确性的特征数量的情况下, 进一步减少了参数数量和计算复杂度。

3) SimAM 注意力模块

轻量化模型结构往往会导致检测精度的损失并降低模型的性能。为了提高甘蔗节点的检测精度, 本研究在检测模型中引入了^[27]注意力机制。SimAM 注意力机制基于视觉神经科学理论, 通常拥有丰富神经科学信息的神经元会表现出与周围神经元不同的放电模式, 并抑制周围的神经元。在视觉处理过程中具有空间抑制效应的神经元应被分配更高的重要性。在检测甘蔗茎节点的任务中, 网络中的这些神经元常常需要被增强, 以提高甘蔗茎节点特征的提取效率并提高检测精度。可以通过测量目标神经元与其他神经元之间的线性分离来检测这些神经元。基于这些结果, SimAM 注意力机制提出了以下能量函数:

$$e_t^* = \frac{4(\hat{\sigma}^2 + \lambda)}{(t - \hat{\mu})^2 + 2\hat{\sigma}^2 + 2\lambda} \quad (1)$$

$$\hat{\mu} = \frac{1}{M} \sum_{i=1}^M x_i \quad (2)$$

$$\hat{\sigma}^2 = \frac{1}{M} \sum_{i=1}^M (x_i - \hat{\mu})^2 \quad (3)$$

其中, t 表示目标神经元, x 表示周围神经元。

在式(1)中, 能量 e_t^* 越低, 神经元 t 与周围神经元 x 之间的差异就越大, 该神经元的重要性就越高。因此, 可以

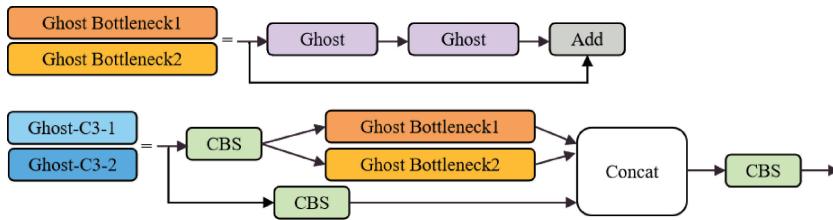


图 5 Ghost-c3 轻量化模型结构

Fig. 5 Ghost-c3 lightweight model structure

通过计算得到每个神经元的重要性 $1/e_i^*$ 。最终的 SimAM 模块基于以下内容进行优化:

$$\tilde{X} = \text{sigmoid}(\frac{1}{E}) \odot X \quad (4)$$

其中, X 是输入特征单通道中的一个神经元, \tilde{X} 是输出特征单通道中的一个神经元, E 用于分类所有通道和空间维度中的能量函数 e_i^* , 并使用 sigmoid 函数限制 E 中可能过大的值。此外, 与现有的 1 D 通道注意力和 2 D 空间注意力机制相比, SimAM 注意力机制可以在不增加额外参数的情况下为特征图生成 D 注意力权重, 使网络能够学习到更多具有判别性的神经元, 提升特征提取能力。这是一个简单但高效的注意力机制模块。

2.3 甘蔗高度测量方法

在对甘蔗进行高度测量之前, 需要先对甘蔗图像分割出甘蔗主干和背景, 进而从左至右对每一个分割出来的甘蔗主体进行标号, 然后对每个甘蔗的中心坐标信息进编码存储。进而, 转动云台对准标号中的第一个甘蔗, 使摄像头视角中心对准甘蔗主体, 实现自动对准甘蔗主体。在自动对准需要测量的甘蔗主体后, 应对不同的环境有以下两个实施方式:

实施方法 1) 本模块主要依靠摄像头自动输出到甘蔗茎底部 d_2 和甘蔗茎顶梢的距离 d_1 以及两个方向的夹角 α , 在自动对准甘蔗主体后, 向下旋转摄像头, 启动甘蔗底部检测器, 检测到甘蔗底部后停止旋转摄像头, 通过深度摄像头获取甘蔗底部到摄像头的距离 d_2 , 再向上旋转摄像头, 同时启动顶梢检测器, 当摄像头视角中出现甘蔗顶梢时停止旋转摄像头, 通过深度摄像头获取甘蔗顶梢到摄像头的距离 d_1 , 云台可以自动获取从底部到顶端抬升的角度 α 。通过公式 $h = \sqrt{d_1^2 + d_2^2 - 2d_1d_2\cos\alpha}$ 来计算甘蔗的高度 h , 机器人在田间作业时, 路面的崎岖不平会导致相机倾斜, 增加误差。由于本测量方案本质上测的是甘蔗生长点到甘蔗顶稍的距离, 不会受到相机倾斜的影响, 利用该方法可以有效减少机器人在作业时因田间道路崎岖不平导致相机倾斜造成的误差, 如图 6 所示。

实施方法 2) 在实际测量中, 由于甘蔗底部光线较暗, 甘蔗底部的检测难度要高于甘蔗顶梢, 在甘蔗底部难以检测成功时, 可使用该实施方式, 在实验前测量出摄像头到地面的距离, 该距离为定值 h_0 , 在自动对准甘蔗主体后此

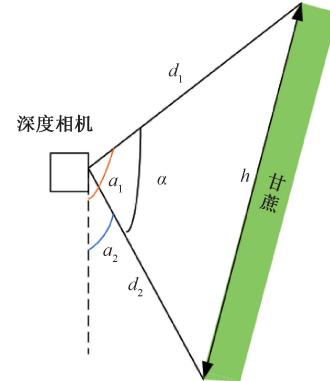


图 6 实施方式 1

Fig. 6 Implementation method 1

时摄像头平视云台, 启动顶梢检测器, 同时向上抬高云台, 当摄像头视角中出现甘蔗顶梢时停止旋转摄像头, 云台可以自动获取从平视到顶端抬升的角度 α , 通过公式 $h = d_1 \sin\alpha + h_0$ 来计算甘蔗的高度 h , 如图 7 所示。

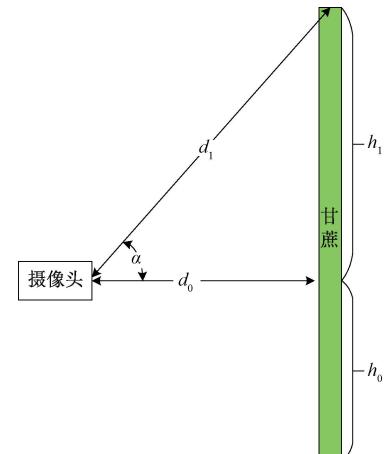


图 7 实施方式 2

Fig. 7 Implementation method 2

3 实验平台与数据集

3.1 移动机器人实验平台

本文使用由移动平台, 舵机云台和深度摄像头组成的机器人平台进行实验, 机器人整体结构如图 8 所示。其中, 移动平台采用的是 MR600 移动机器人, 由于其体积

小、续航长、负载大、精度高、控制简单以及性价比高等优点,选其作为移动平台。并且,在移动平台上搭载了 2 自由度的舵机云台,可以稳定的作为检测所用摄像头的搭载工具。最后,本文选用了 D435i 深度摄像头作为甘蔗高度表征检测实验的感知器件,为后续实现自动甘蔗高度表征测量提供硬件基础。

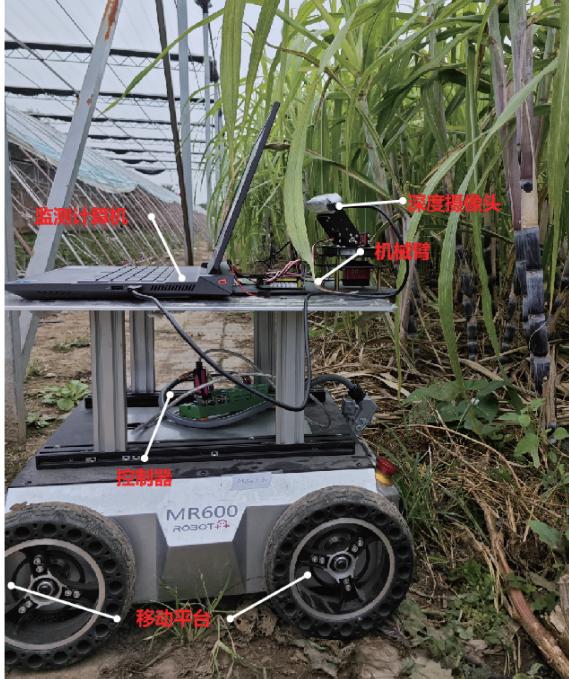


图 8 机器人实验平台结构

Fig. 8 Structure of robot experiment platform

3.2 甘蔗数据集介绍

本文的甘蔗数据采集主要依靠移动平台不断靠近甘蔗,以及云台控制摄像头的移动来检测和测量甘蔗的表型数据。

本研究使用的数据集样本取自北京市遗传与发育生物学研究所以及北京市草莓博士甘蔗种植基地。所选甘蔗的生长期约为 3~10 个月,平均直径为 50 mm,总计 80 根甘蔗。数据集有 3 578 张图像,其中训练数据集有 2 984 张图像,测试集有 394 张图像。为了使训练的模型具有更强的泛化性,该数据集的采集过程,分别针对甘蔗的不同成熟度、光照强度、探测角度和环境复杂度进行收集图像。如图 9 所示,展示了不同场景下采集甘蔗图像数据,以期望模型可以具有更强的实际应用潜力。此外,为了更直观的展示甘蔗高度表征分布,绘制了甘蔗数据集的高度频率直方图,如图 10 所示。

进一步,本文对数据集数据进行图像预处理,包括图像去噪、图像增强以及图像校正。其中,使用了中值滤波器去除图像中的噪点,并且对数据进行了直方图均衡化,最后使用灰度校正方法中的直方图修正对图像进行复原校正,减弱由于运动平台在运动过程中造成的图像失真等情况。

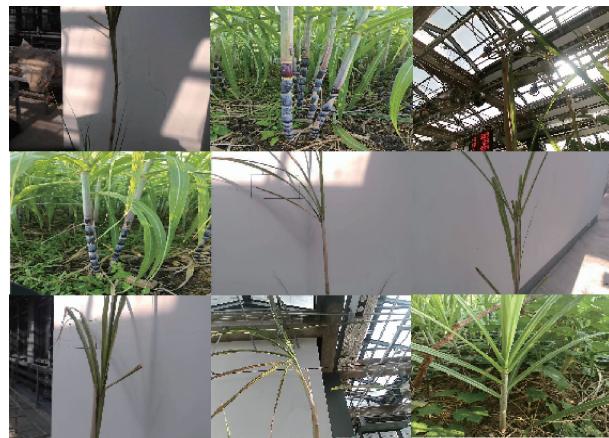


图 9 不同场景下甘蔗数据集

Fig. 9 Sugarcane dataset under different scenarios

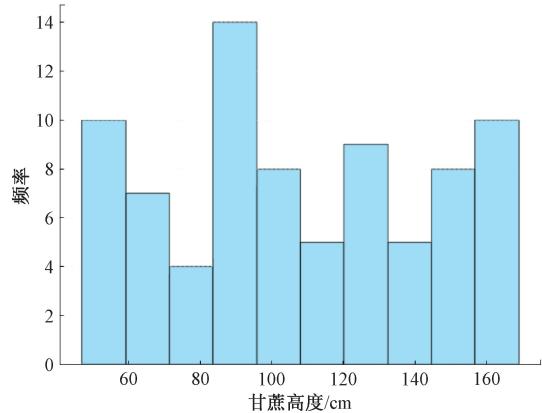


图 10 甘蔗数据集高度频率直方图

Fig. 10 Histogram of the height frequency of the sugarcane dataset

4 实验设计

4.1 实验环境和参数设置

本研究中使用的处理平台是配备操作系统的笔记本电脑 Windows11、3.20 GHz AMD Ryzen7 5800 H CPU、16 GB 内存和 NVIDIA GeForce RTX3060(6 GB)GPU。为了提高网络训练的速度,使用了 GPU 加速。CUDA 版本为 11.6.0, cuDNN 版本为 8.3.2。开发语言为 Python3.8.13 和 PyTorch1.12.1, 初始学习率为 0.01, 最终 OneCycleLR 学习率为 0.01, 动量为 0.937, 权重衰减为 0.0005, 批处理大小为 8, 训练 epoch 设置为 200。每个模型都使用预训练模型方法进行训练。

4.2 模型评价指标

本研究针对甘蔗高度表征测量任务,设计了轻量化改进的 PSPNet 与 YOLOv5s 模型,用于甘蔗主体分割与甘蔗顶梢检测,在参数量减低的情况下,同时保持较高的检测精度。因此,有必要评估模型的性能和复杂性。对于图像分割模型,像素精度(PA)、IoU、FPS 和网络参数作为甘蔗分割模型的评价指标。指数值越高,模型越有效,其

计算公式如式(5)~(6)所示。

对于图像检测模型,常用的评估指标包括平均精度(AP)、模型大小、参数数和浮点运算(FLOP)。其计算公式如式(7)~(9)所示。AP 值是由准确率 P 和召回率 R 组成的 P-R 曲线的面积。在这里,TP 表示甘蔗茎节的正确检测数量。FP 表示甘蔗茎节点的检测次数不正确,其中阴性样本表示阳性样本,FN 表示甘蔗茎节点检测次数不正确,其中阳性样本表示为阴性样本。

模型大小对应模型所需的存储空间,主要由模型中的参数数量决定。模型越小,越容易应用于移动终端。FLOP 表示浮点运算的数量,用于衡量模型的计算量和复杂度。通常, FLOPs 值越低,模型所需的推理时间就越少。

对于甘蔗高度测量精度的评估,本文使用平均绝对误差(MAE)衡量测量方法的优劣。其计算公式如式(10)所示。

$$PA = \sum_{i=0}^k \frac{p_{ii}}{\sum_{j=0}^k p_{ji}} \times 100\% \quad (5)$$

$$IoU = \sum_{i=0}^k \frac{p_{ii}}{\sum_{j=0}^k p_{ij} + \sum_{j=0}^k p_{ji} - p_{ii}} \times 100\% \quad (6)$$

$$AP = \int_0^1 P(R) dR \times 100\% \quad (7)$$

$$P = \frac{TP}{TP + FP} \times 100\% \quad (8)$$

$$R = \frac{TP}{TP + FN} \times 100\% \quad (9)$$

$$MAE = \frac{1}{K} \sum_{i=1}^K \|y_i - y_i'\| \quad (10)$$

5 结果与讨论

5.1 甘蔗主干分割实验结果与分析

在这项研究中,甘蔗的主干被视为一类,背景被视为另一类,为了验证改进模型的躯干分割性能,基于 PSPNet,在自然环境下的甘蔗躯干数据集上测试了 PSPNet 模型的不同结构,并验证了不同结构对分割模型性能的有效性。结果如表 1 所示。

作为经典的分割网络模型,PSPNet 可以有效地在复杂环境中分割目标,但在计算资源有限的情况下,很难分割甘蔗主干。在原始 PSPNet 中,使用 ResNet50 作为主要特征提取网络。原始 PSPNet 在甘蔗主干分割模型中的 IoU 和 PA 分别为 78.15% 和 86.26%。在第 13 个卷积残差层加入 CBAM 后,改进的模型增加了参数,并降低了甘蔗分割时的 IoU。以 MobileNetV2 作为特征提取网络的骨干,与原始 PSPNet 相比,甘蔗主干分割模型的 IoU、PA 和 Fps 分别增加了 0.17%、0.53% 和 4.25,模型参数减少了 20 倍。与原始 PSPNet 相比,以 MobileNetV2 为骨干特

表 1 各模块对模型性能的比较结果

Table 1 Comparison results of model performance by various modules

模型(PSPNet)	IoU/%	PA/%	参数量	FPS
ResNet50	78.15	86.26	4.91×10^7	49.77
ResNet50+CBAM	77.71	86.27	5.04×10^7	48.57
MobileNetV2	78.32	86.79	2.45 \times 10^6	54.02
MobileNetV2+CBAM	78.46	86.96	2.48×10^6	51.76

征提取网络,并在第 5 和第 7 个瓶颈结构后引入 CBAM 网络的改进 PSPNet 在 IoU 上有所提升,验证了引入 CBAM 网络的有效性。最后,与原始 PSPNet 的算法相比,以 MobileNetV2 为骨干网络并引入 CBAM 的改进 PSPNet,在 IoU 上提高了 0.31%,在 PA 上提高了 0.70%,在 Fps 上提高了 1.99,并且算法参数仅为原始 PSPNet 的 5.00%,经改进后的 PSPNet 模型处理后的甘蔗分割图像如图 11 所示。



图 11 经改进后的 PSPNet 模型处理后的甘蔗分割图像

Fig. 11 Sugarcane segmentation image processed by the improved PSPNet model

5.2 甘蔗顶梢检测实验结果与分析

根据网络深度和宽度,有 5 种不同的 YOLOv5 网络模型,分别是 YOLOv5n、YOLOv5s、YOLOv5m、YOLOv5l 和 YOLOv5x。从 YOLOv5 官方文档来看,这些模型的性能正在提高,但它们的复杂性也相应增加。由于自建甘蔗顶梢数据集与 COCO 数据集差异较大,YOLOv5 官方文档给出的模型间数据对比不足以在选择基准模型时获得满意的结果。因此,在自建数据集上测试了第六版 YOLOv5 网络模型的性能,如表 2 所示。

YOLOv5s、YOLOv5m、YOLOv5l 和 YOLOv5x 的 AP 值几乎相同,均在 97% 左右。然而,在模型大小、参数数量和 FLOP 方面,后三者是 YOLOv5s 的 2 倍以上,其中 YOLOv5x 的复杂度最高,分别是 YOLOv5s 的 12 倍、12.3 倍和 12.9 倍。可以看出,复杂的网络结构并没有提高甘蔗顶梢的检测精度,反而增加了模型的大小和计算负担。尽管, YOLOv5n 在模型大小、参数和 FLOP 上优于 YOLOv5,但其 AP 值仅为 94%,低于 YOLOv5s 的 96.8%,

表 2 YOLOv5 系列模型在甘蔗数据集上的检测结果
Table 2 Detection results of YOLOv5 series models
on sugarcane dataset

模型(YOLO)	AP/%	模型大小/MB	参数量	FLOPs/G
v5n	94.0	3.9	17 605 180	4.1
v5l	96.9	92.2	46 108 278	107.6
v5m	97.0	42.2	20 852 934	47.9
v5x	97.1	173.1	86 173 414	203.8
v5s	96.9	14.4	7 012 822	15.8
v5s-L	96.5	3.4	1 558 340	10.1
v5s-L-G	96.6	2.6	1 129 340	7.2
v5s-L-G-S	97.2	2.6	1 129 340	7.2

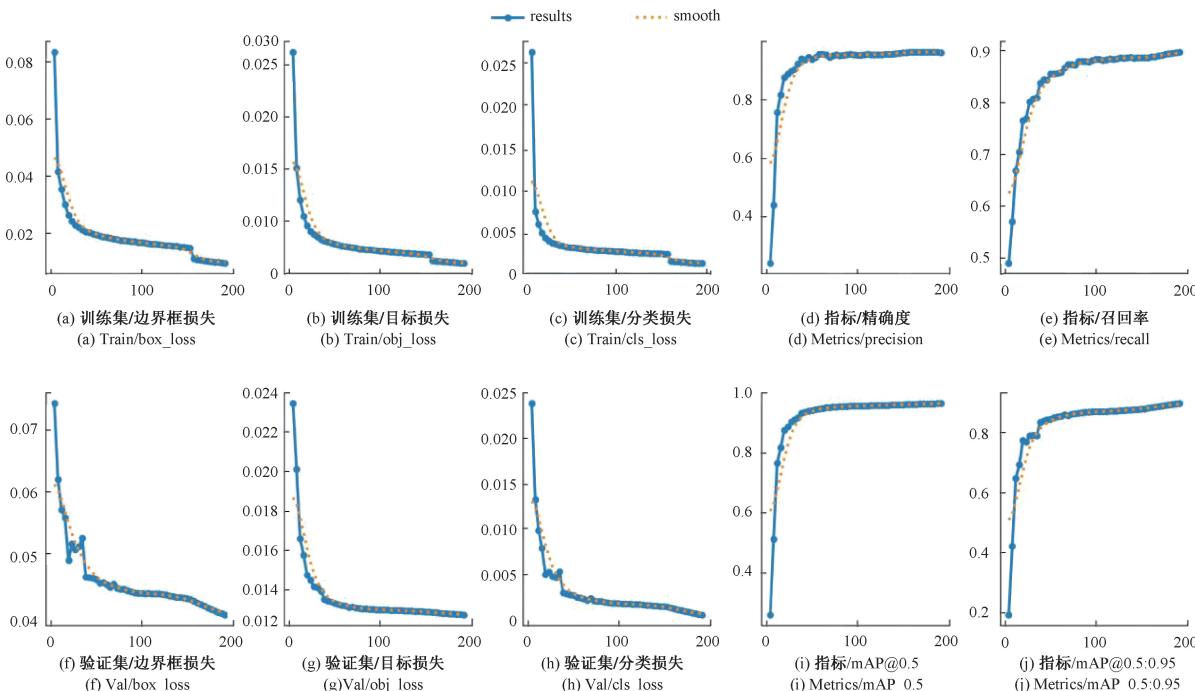


图 12 经 200 轮训练的 YOLOv5s-L-G-S 模型的评价指标图

Fig. 12 Evaluation index diagram of the YOLOv5s-L-G-S model trained in 200 rounds

5.3 甘蔗高度测量实验分析

国内外对于使用机器人自动化测量甘蔗高度的研究非常少, 甘蔗高度测量目前以人工测量为主, 到了 2023 年 11 月底, 才开发出使用双目摄像头的方法进行测量, 现有的使用深度信息的测量方案^[28]主要是通过两个摄像头来进行甘蔗高度测量, 由两相机之间距离 h_0 、底端相机 h_1 与植株底端竖直方向的距离 h_2 三部分组成, 双目摄像头需要将该距离按照转换比例进行转换, 并且需要对两个摄像头进行标定, 而后通过公式 $h_3 = h_0 + h_1 + h_2$ 来计算甘蔗的高度 h 。在现有方案中, 甘蔗的具体高度无法确定, 深度摄像头可以测量的夹角范围仅有 120°, 如果甘蔗高度过高, 则可能无法完成测量; 在现有方案中, 如果甘蔗倾斜, 使用双目摄像头会造成较大误差。并且现有的方案需要进行

因此, 研究选 YOLOv5s 作为研究基准模型, 并对其进行优化设计。

本文所设计的 YOLOv5s-L-G-S 模型进行了消融实验, 如表 2 所示。实验结果表明, 通过轻量化设计和引入 Goast 模型的模型改进方法, 尽管大幅减少了模型参数, 但准确率下降很多, 然而, 通过引入了 SimAM 模块, 加强了模型对目标特征的捕捉能力, 使模型学习到更多有效信息, 最终提升了检测精度。最终, YOLOv5s-L-G-S 的参数量在下降了 83.9% 的基础上, AP 值仍然保持在 97% 左右, 极大提升了模型部署能力。本文提出的 YOLOv5s-L-G-S 经 200 轮训练的评价指标图, 如图 12 所示。甘蔗顶梢与底端检测结果如图 13 所示。

深度相机的标定以及坐标转换, 计算量较大, 如表 3 所示, 测量 MAE 为 4.2。而本文使用单个深度摄像头获取摄像头分别到甘蔗底部与甘蔗顶梢的距离, 云台获取从甘蔗底部到甘蔗顶梢过程中深度摄像头所旋转的角度, 后续直接通过余弦定理求出甘蔗高度, 不涉及多摄像头之间的标定和坐标转换问题, 计算量小, 可以提高甘蔗高度的检测效率, 且大幅度提升了测量精度, 测量 MAE 为 2.0 以及 2.4, 相较于以前的方法误差降低了 50%。

本文选择了验证数据集中的 5 个不同高度的甘蔗, 并对其都进行 20 次实验, 最终取平均值作为实验结果。并部署了相同的测量场景, 使用 3 种实施方法对甘蔗进行高度测量, 其中, 将人工测量结果作为甘蔗真实高度值, 另外 3 种测量方式分别为深度信息测量和实施方式 1、实施方

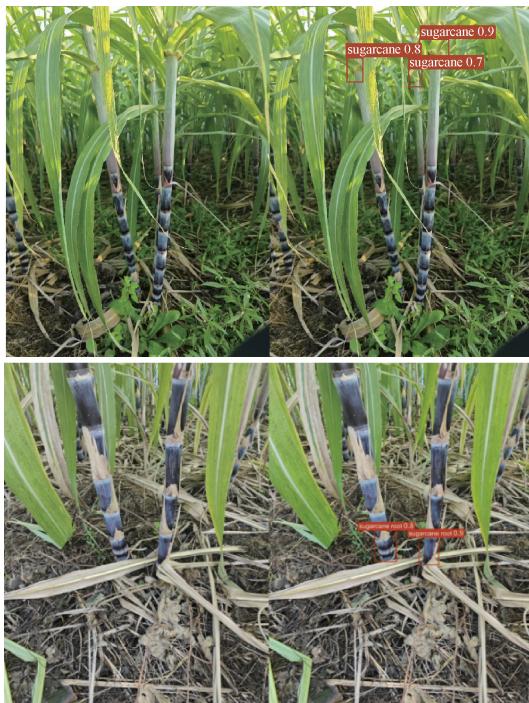


图 13 经 YOLOv5s-L-G-S 模型推理后的甘蔗检测图像

Fig. 13 Sugarcane detection image processed by YOLOv5s-G-S model

表 3 不同测量方法的实验结果

Table 3 Experimental results of different measurement methods

类型	人工	深度信息	实施方法 1	实施方法 2
1	155	161	154	152
2	132	128	135	135
3	104	108	102	102
4	95	91	97	94
5	63	68	61	66
MAE	0	4.2	2.0	2.4
准确率	100%	95.42%	97.98%	97.61%
成功率	100%	78%	93%	95%
检测时间	9.8s	6.1s	13.2s	10.6s

式 2。对于实施方式 1 和实施方式 2 是在 3.1 节描述的移动机器人平台进行测量实验,具体测量原理在 2.3 节中。

测量结果如表 3 所示,由于机器测量需要转动云台,测量时间相比于深度信息测量方式较长,但是本文提出的机器人测量系统不仅可以适应更复杂的测量环境,并且测量精度更高。且测量时间相比于人工测量差距不大,可以满足现场测量的要求,机器人使用两种不同的实施方法测量实验现场如图 14 所示。

除了对各个测量方法的检测精度评价以外,仍需对各个方法的现场自动化测量进行可行性与实时性分析。从实时性的角度来看,尽管受到系统设计和环境干扰的影响,两种机器人测量方法以及通过深度信息测量的检测时



图 14 机器人测量实验现场

Fig. 14 Robot measurement experiment site

间均能满足甘蔗高度测量的要求。从可行性的角度来看,深度信息测量方法必须保证深度摄像机能完整拍摄到整个甘蔗视角,然后使用甘蔗顶梢和底端的深度信息来计算两点的空间距离,由于甘蔗生长参差不齐,这需要在测量过程中对移动机器人增添校对视角步骤,从自动化设计角度来说几乎很难实现。然而,实施方法 1 与实施方法 2 设计了机械臂与视觉算法相结合的测量策略,克服了摄像机与甘蔗的视角问题。在甘蔗底部和顶稍遮挡较少的情况下有较高的测量成功率,可以实现高度自动化的测量流程。

6 结 论

本文针对甘蔗高度表型的自动化准确测量问题,开发了甘蔗高度测量机器人系统,对于系统的甘蔗顶梢检测问题,提出了轻量化改进的 YOLOv5s 模型,该模型通过删减冗余结构,并引入 Ghost 模块与 SimAM 模块,减少了模型参数和计算量,提升了模型的鲁棒性,结果显示,参数量降低了 83.9%,顶梢检测准确率仍保持在 97% 左右。对于系统的甘蔗主干分割问题,提出了轻量化改进的 PSPNet 网络,更换 MobileNetV2 为骨干网络,并且引入 CBAM 注意力机制,参数量仅为原模型的 5%,并且精度保持在 86% 左右。在实际甘蔗高度测量方法中,由于现有的对于甘蔗测量技术较少,本文设计了强鲁棒性的云台辅助测量方法,相比于深度信息直接测量方法,机器人辅助测量方法将机器人云台角度信息加入高度计算中,测量 MAE 为 2.0 以及 2.4,并且不再受限制于甘蔗拍摄视角问题,可以实现更高度自动化的甘蔗高度表型测量。

参考文献

- [1] Sugarcane biorefinery, Technology and perspectives [M]. Pittsburgh: Academic Press, 2019.
- [2] SOM-ARD J, ATZBERGER C, IZQUIERDO-VERDIGUIER E, et al. Remote sensing applications in sugarcane cultivation: A review [J]. Remote

- sensing, 2021, 13(20): 4040.
- [3] SILVA M A, SILVA J A G, ENCISO J, et al. Yield components as indicators of drought tolerance of sugarcane[J]. *Scientia Agricola*, 2008, 65: 620-627.
- [4] THILAGAVATHI K, KAVITHA K, PRABA R D, et al. Detection of diseases in sugarcane using image processing techniques [J]. *Bioscience Biotechnology Research Communications, Special Issue*, 2020(11): 109-115.
- [5] MILITANTE S V, GERARDO B D. Detecting sugarcane diseases through adaptive deep learning models of convolutional neural network [C]. 2019 IEEE 6th International Conference on Engineering Technologies and Applied Sciences(ICETAS). Kuala Lumpur:IEEE, 2019: 1-5.
- [6] MILITANTE S V, GERARDO B D, MEDINA R P. Sugarcane disease recognition using deep learning[C]. 2019 IEEE Eurasia Conference on IOT, Communication and Engineering (ECICE). Yunlin: IEEE, 2019: 575-578.
- [7] KHUIMPHUKHIEO I, MARCONI T, ENCISO J, et al. The use of UAS-based high throughput phenotyping (HTP) to assess sugarcane yield [J]. *Journal of Agriculture and Food Research*, 2023, 11: 100501.
- [8] SIMôES I O, DO AMARAL L R. UAV-based multispectral data for sugarcane resistance phenotyping of orange and brown rust [J]. *Smart Agricultural Technology*, 2023, 4: 100144.
- [9] SINDHE P S, HEGADI R S. Sugarcane bud detection using YOLOv5 [C]. International Conference on Recent Trends in Image Processing and Pattern Recognition. Cham: Springer Nature Switzerland, 2023: 202-212.
- [10] ZHANG F, DONG D, JIA X, et al. Sugarcane-YOLO: An improved YOLOv8 model for accurate identification of sugarcane seed sprouts [J]. *Agronomy*, 2024, 14(10): 2412.
- [11] ZHOU D, ZHAO W, CHEN Y, et al. Identification and localisation algorithm for sugarcane stem nodes by combining YOLOv3 and traditional methods of computer vision[J]. *Sensors*, 2022, 22(21): 8266.
- [12] PENG H X, XUE CH, SHAO Y Y, et al. Semantic segmentation of litchi branches using DeepLabV3 + model[J]. *IEEE Access*, 2020, 8: 164546-164555.
- [13] QI X, DONG J, LAN Y, et al. Method for identifying litchi picking position based on YOLOv5 and PSPNet[J]. *Remote Sensing*, 2022, 14(9): 2004.
- [14] LI X, DUAN F, HU M, et al. Weed density detection method based on a high weed pressure dataset and improved psp net[J]. *IEEE Access*, 2023, 11: 98244-98255.
- [15] ZHAO J, LI Z, LEI Y, et al. Application of UAV RGB images and improved PSPNet network to the identification of wheat lodging areas[J]. *Agronomy*, 2023, 13(5): 1309.
- [16] LU ZH G, CHEN F M, CHAO Y, et al. Research on extracting special plant planting plots from high-resolution remote sensing images using I-PSPNet semantic segmentation model [J]. *Remote Sensing Technology and Application*, 2024, 39(1): 222-233.
- [17] SANDLER M, HOWARD A, ZHU M, et al. Mobilenetv2: Inverted residuals and linear bottlenecks[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2018: 4510-4520.
- [18] LU J, LIU X, MA X, et al. Improved MobileNetV2 crop disease identification model for intelligent agriculture [J]. *PeerJ Computer Science*, 2023, 9: 1595.
- [19] LIU J, WANG X. Early recognition of tomato gray leaf spot disease based on MobileNetv2-YOLOv3 model[J]. *Plant Methods*, 2020, 16: 1-16.
- [20] LI J, JIN K, ZHOU D L, et al. Attention mechanism-based CNN for facial expression recognition [J]. *Neurocomputing*, 2020, 411: 340-350.
- [21] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module[C]. European Conference on Computer Vision(ECCV), 2018: 3-19.
- [22] WANG Y, ZHANG Z, FENG L, et al. A new attention-based CNN approach for crop mapping using time series sentinel-2 images [J]. *Computers and Electronics in Agriculture*, 2021, 184: 106090.
- [23] JOCHER G, CHAURASIA A, STOKEN A, et al. Ultralytics/yolov5: v6. 1-tensorrt, tensorflow edge tpu and openvino export and inference[J]. Zenodo, 2022, DOI:10.5281/zenodo.6222936.
- [24] ZHANG J, WANG J, ZHAO M. A lightweight crop pest detection algorithm based on improved YOLOv5s [J]. *Agronomy*, 2023, 13(7): 1779.
- [25] 张福豹, 吴婷, 赵春峰, 等. 基于弱光增强与 YOLO 算法的锯链缺陷检测方法[J]. 电子测量技术, 2024, 47(6): 100-108.
- SUN F B, WU T, ZHAO CH F, et al. Saw chain defect detection system based on low-light enhancement and YOLO algorithm [J]. *Electronic Measurement Technology*, 2024, 47(6): 100-108.
- [26] HAN K, WANG Y, TIAN Q, et al. Ghostnet: More

- features from cheap operations [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 1580-1589.
- [27] ZHOU H, SU Y, CHEN J, et al. Maize leaf disease recognition based on improved convolutional neural network ShuffleNetV2 [J]. Plants, 2024, 13 (12): 1621.
- [28] 孙大洋, 叶昕禹, 田瑞琦, 等. 一种基于机器视觉的农作物植株高遥测方法及系统:CN202311200339.1[P]. 2023-11-28.
- SUN D Y, YE X Y, TIAN R Q, et al. A machine vision based remote measurement method and system for crop plant height: CN202311200339.1[P]. 2023-11-28.

作者简介

蒋正中,硕士研究生,主要研究方向为计算机视觉和图像处理。

E-mail:1748187573@qq.com

杨鸿波(通信作者),教授,硕士生导师,主要研究方向为计算机视觉和图像处理。

E-mail:anonbo@bistu.edu.cn

杨明浩,副研究员,硕士生导师,主要研究方向为多通道环境信息感知与交互、机器人灵巧操作交互学习、交互引导的机器人技能优化。

E-mail:mhyang@nlpr.ia.ac.cn

刘安琪,硕士研究生,主要研究方向为计算机视觉和图像处理。

E-mail:2023020375@bistu.edu.cn