

DOI:10.19651/j.cnki.emt.2313205

无人机图像数据驱动的莴苣属株高检测*

贺星耀¹ 冯涛² 梁虹¹ 吴凯香¹ 袁嘉辉¹

(1. 云南大学信息学院 昆明 650504; 2. 云南财经大学信息学院 昆明 650221)

摘要: 莴苣属作物高通量株高数据采集技术的研究和应用较少,对此提出一种基于深度学习和无人机倾斜摄影的莴苣属株高检测方法。首先针对高通量植株株高获取,采用无人机倾斜摄影,生成区域内植株三维模型,提取高程信息;然后利用改进注意力机制的YOLOv5算法,在主干网络C3模块中嵌入CBAM注意力机制,减少浅层噪声信息,提高对小目标及密集目标的检测能力,以达到对区域内植株的目标检测,对应得到每株植株的估算株高。实验结果表明,CBAM-YOLOv5模型识别效果提升明显,对莴苣属作物识别的AP值提升到了96.19%,相较于原始的YOLOv5模型,本文模型的AP值提升了1.5%,植株目标检测对应三维模型计算出的估算值与实测值具有较高的相关性,直线斜率为0.9911, R^2 为0.9311,实现了对莴苣属作物高通量株高数据的检测。

关键词: 莴苣属;株高检测;深度学习;YOLOv5;无人机倾斜摄影;三维建模

中图分类号: TP391.41 **文献标识码:** A **国家标准学科分类代码:** 520.6

Unmanned aerial vehicle image data-driven detection of lettuce plant height

He Xingyao¹ Feng Tao² Liang Hong¹ Wu Kaixiang¹ Yuan Jiahui¹

(1. School of Information Science and Technology, Yunnan University, Kunming 650504, China; 2. School of Information Science and Technology, Yunnan University of Finance and Economics, Kunming 650221, China)

Abstract: The study and application of high-throughput plant height data acquisition technology for lettuce crops are limited. A lettuce plant height detection method based on deep learning and drone oblique photography is proposed to address this. Firstly, oblique photography by drone is used to obtain high-throughput plant height data, and a 3D model of plants within the region is generated to extract elevation information. Then, an improved YOLOv5 algorithm with a CBAM attention mechanism embedded in the C3 module of the backbone network is proposed. This algorithm is designed to reduce shallow noise information, enhance the detection capability of small and dense targets, and achieve target detection of plants in the region. This will result in estimated plant heights for each plant. The experimental results show that the CBAM-YOLOv5 model significantly improves the recognition effect, increasing the AP value for lettuce crop recognition to 96.19%. Compared with the original YOLOv5 model, the AP value of our model has increased by 1.5%. The plant target detection has a high correlation between the estimated values calculated from the 3D model and the measured values, with a linear slope of 0.9911 and R^2 -value of 0.9311, achieving the detection of high-throughput plant height data for lettuce crops.

Keywords: lettuce; plant height detection; deep learning; YOLOv5; drone oblique photography; 3D modeling

0 引言

莴苣属(Lactuca)是一类广泛分布于亚欧两洲的温带和亚热带地区的菊科植物,包括了约75个物种。在中国,约有7种莴苣属植物被广泛栽培和食用。此外,莴苣属植物具有很高的营养价值,尤其是叶子富含多种矿物质和维生素,如铁盐、钙盐和磷盐,以及维生素A、B1、B2、C和P

等。因此,莴苣属植物被广泛用于制作沙拉、汤和炒菜等菜肴,并且在一些地区还被用于药用,例如用于治疗失眠和咳嗽等症状。但目前,研究人员对获取莴苣属作物高通量表型数据的技术探索有限。因此,莴苣属作物表型缺乏产生大数据的基础,阻碍了通过大数据支持智能育种的能力。

在植株的众多表型性状中,植物株高是植物在受基因、管理和环境因素相互作用后,表现出来的结构及生长的特

收稿日期:2023-03-27

* 基金项目:云南大学信息学院第二届院级实践创新项目(Y200211)资助

征和性状,植物株高参数的测量及分析可以有效调节植物的育种环节,最终影响植物的生长发育和作物产量。目前,有多种方法可用于获取植物高度数据,包括使用多种数据采集设备,如电子计算机断层扫描、光合荧光成像、高光谱成像、磁共振成像扫描技术、可见光(RGB color mode, RGB)成像、多光谱成像和红外成像。这些方法可以从不同尺度收集植物高度数据,从单个细胞和器官到整个植物和种群。利用这些方法进行数据采集的植物主要包括:巨芒草^[1]、玉米^[2]、小麦^[3]、高粱^[4]等。同时,基于机器学习的植物株高检测也逐步开展,现已运用在番茄幼苗^[5]、玉米^[6]、小麦^[7]等作物的株高测量中,但基于深度学习高通量株高数据采集技术的研究和应用仍然较少。近年来,针对莴苣属作物的株高数据获取工作已有零星开展,但与其他作物相比,莴苣属作物高通量株高数据采集技术的研究和应用存在许多难点,如莴苣属作物的植株结构通常比较复杂,肉眼很难定义植株高度的起始点和终点;莴苣属作物的种植一般较为密集,存在遮挡和重叠,干扰对目标的检测。

随着高通量植物表型技术的广泛应用和植物高度数据的快速积累,我们正迈向更深入、全面的植物研究,而利用大数据开展植物育苗育种已逐渐成为研究热点。为此,本文基于深度学习和无人机倾斜摄影技术^[8],建立高精度、多尺度的莴苣属作物株高数据获取模型^[9-10],构建可用于无人机和地表图像处理设备的空地融合莴苣属作物大田植株识别、株高测量的智能算法模型。在此基础上,实现符合莴苣属作物植株特点的低成本、高通量株高数据获取,从而为莴苣属作物农业的数字化、智能化的育种育苗提供数据基础和技术支撑。

1 材料与方 法

1.1 实验地点

实验于 2022 年 9~11 月在云南省昆明市呈贡区云南大学附近一块生菜种植基地(24.50°N,102.50°E)进行,属于亚热带季风性湿润气候,光照良好,降水充沛,作物生长态势良好。同时,该区域的地形相对开阔平坦,有利于无人机在该区域进行大面积的数据采集与分析。如图 1 所示,为选取的实验地点、现场拍摄的照片及使用的设备。

1.2 数据获取与处理

莴苣属包含多种作物,取其中代表性作物生菜作为主要实验对象,研究的模型可广泛应用于同属其他作物。本研究的数据采集主要分为两个阶段,均采用大疆 PHANTOM 4 PRO 无人机采集可见光影像数据,配备有效像素为 2 000 万的相机,焦距为 9 mm,最大分辨率为 5 472×3 956。第一阶段为生菜检测数据集采集阶段,通过设置无人机的图像重叠度、采样距离和飞行高度等参数,可以获得高质量的 RGB 相机数据。常见的设置包括主航线图像重叠度为 80%,主航线间图像重叠度为 85%,飞行高度为 5 m,采样距离为 0.68 cm/pixels;第二阶段为生菜建

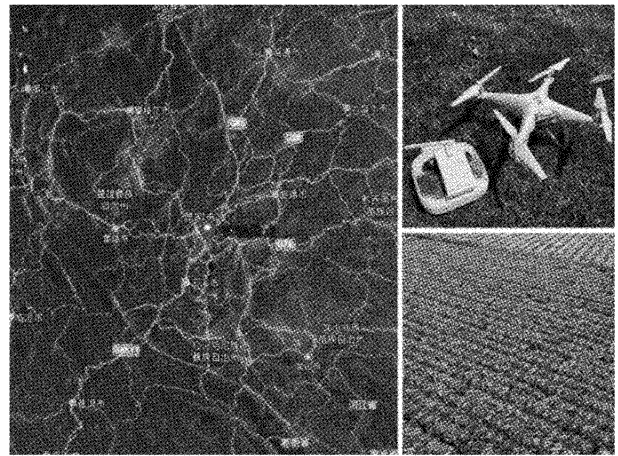


图 1 实验地点、现场及设备照片

模数据采集阶段,无人机环绕飞行采集数据,无人机搭载 RGB 相机系统飞行高度设置为 5、6、7 m,飞行模式设置为兴趣点环绕,以试验田特定位置的圆心为兴趣点圆心,半径为 5 m 进行环绕拍摄,在不同的环绕高度设置不同俯仰角度,分别为 -73° 、 -78° 、 -80° ,环绕飞行速度为 0.4 m/s,RGB 相机设置为等间隔 2 s 拍照,环绕飞行结束后,在兴趣点正上方 8 m,俯仰角为 -90° 拍摄正射影像。

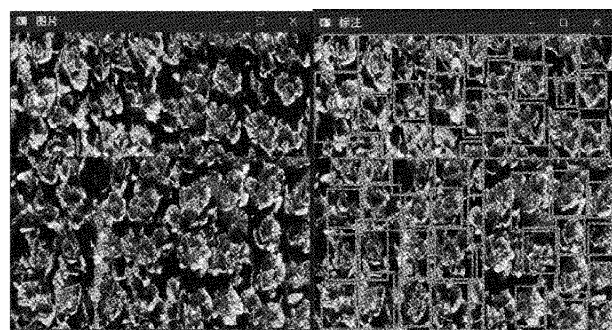
为了确保高质量的图像采集,避免过厚的云层阻挡阳光,飞行是在中午晴朗无风的情况下进行的。这个时间段的照明条件对飞行后的图像分析和处理是有利的。在飞行任务之前,完成飞行前的准备工作,如设置拍摄参数和飞行路线。RGB 相机被设置为自动曝光模式,以避免任何参数对图像的干扰,每次飞行共拍摄 200 张图片。除此以外,记录栽培天数,对该实验地块生菜植株的生菜数量、生菜高度等各项长势指标进行实测和均值的计算。

1.3 数据标注与增强

本研究采用 LabelImg 工具进行标注,按照 Pascal VOC^[11]数据集格式,对检测任务的 3 600 幅无人机 RGB 图像中的生菜植株进行标注,标注标签为“lettuce”,每个标注好的矩形框会生成对应的 xml 文件,该文件包含每幅图像目标生菜的像素坐标及标签,并划分成训练集、验证集和测试集,其中训练集占 80%,验证集和测试集均为 10%,将其用于目标检测模型的训练、验证及测试任务。

目前,虽然运用深度学习的目标检测算法性能已经有了明显的提升,就其精准度及学习能力也有了长足的进步,但其优异的性能需要基于大量的数据训练及计算。而数据增强是解决数据不足问题的一个好办法。基于 CutMix 数据增强的 Mosaic 是一项强大的技术,它将 4 张图像拼接在一起,创建一个新的图像。这极大地丰富了被检测物体的背景,并能在批量规范化过程中同时计算所有 4 张图像的数据。Mosaic 数据增强的具体流程为:1)读取 4 幅图像 2)分别对 4 张图片进行翻转、缩放、色域变化等,并按照 4 个方向进行定位 3)对图像和方框进行组合。如图 2 所示,可

以模拟密集场景中莴苣属作物的状态,丰富检测物体目标及其背景。



(a) 图片 (b) 标注

图 2 Mosaic 数据增强结果

2 研究方法

本研究主要分为两步,第一步利用无人机拍摄的数据,生成三维模型^[12],进而得到包含经纬度及高程信息的数字表面模型以及数字正射影像,第二步用改进的 YOLOv5 (you only look once) 目标检测模型检测数字正射影像中的生菜目标,对生菜目标中心位置进行定位,并输出中心点位置,将中心点位置与生成的数字表面模型进行对应,得到不同生菜植株其经纬度及高程信息,即株高信息。

2.1 三维数据的生成

用 ContextCapture 软件来处理无人机航拍的生菜 RGB 图像,生成三维点云、数字表面模型(Digital surface model, DSM)和数字正射影像(Digital orthomosaic map, DOM),数据处理流程如图 3 所示,将拍摄的一块试验田内的生菜图片导入 ContextCapture 软件,可以自动读取照片的经纬度、高程、旋转角度、空间姿态等信息,检查照片文件,设置照片位置和空间参考,并将预先处理好的像控点信息导入,提交空三运算,运算结束后开始构建三维模型,进行分块设置,生成三维点云或三维网格,最后得到对应的 DOM 与 DSM 图像,并输出为 tif 格式,而 tif 格式文件我们无法直接使用,还需要用 gdal 中的 GetGeoTransform 对数据进行读取,导出需要的关键信息:经纬度及高程。

2.2 YOLOv5 简介

莴苣属作物植株检测可以运用深层语义特征的深度学习算法来实现,而目前基于深度学习的目标检测算法主要可以分为两类,其中一种是以 R-CNN^[13]、Faster R-CNN^[14]、Mask R-CNN^[15]为代表的基于候选区的两阶段的算法,这种算法有着更高的准确性,但因为需要使用候选区域来生成网络,所以速度较慢,而另一种是基于回归的单阶段算法,其中代表有 SSD (single shot multibox detector)^[16]、Center Net^[17]、YOLO^[18]等算法,不同于第一种算法,其不使用候选区域生成网络,而是采用 regression 的方法来实现物体位置与分类的预测,因此有着更快的速

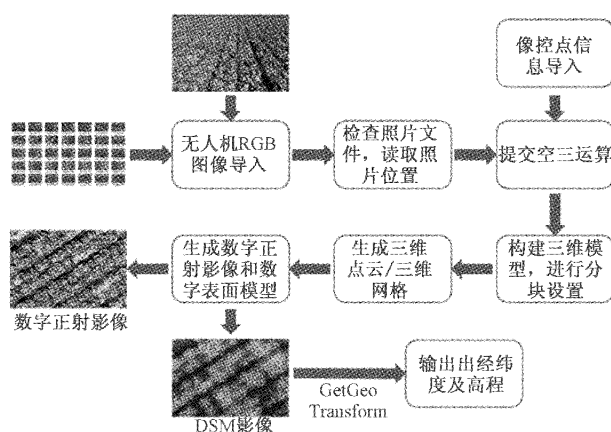


图 3 三维数据生成流程

度,实时性也更好^[19]。考虑到我们需要大规模高速率地识别莴苣属作物,所以选择 YOLOv5 作为莴苣属作物植株检测器。

YOLOv5 的网络结构是在 YOLOv4^[20] 的框架上进行优化的,其融合了数据增强、感受野模块、多正样本匹配和特征金字塔等方法。目前,YOLOv5 的版本已经更新到了 v7.0,同时,设置不同的网络深度和宽度,又将 YOLOv5 分为了 YOLOv5n、YOLOv5s、YOLOv5m、YOLOv5l、YOLOv5x 5 个模型,它们的网络深度和宽度依次递增,相对应的它们的检测速度依次递减。考虑我们对检测速度的需求,YOLOv5s 模型相对更加适合,其网络深度和宽度较小,检测速度更快。YOLOv5s 的网络结构由四部分组成:输入(inputs)、主要特征提取骨干(Backbone)、用于特征融合的颈部(Neck)和检测头(Head)。

输入模块包含了自适应锚点计算,这使得网络在训练过程中能够根据初始锚框输出预测框。然后将这一输出与真实框进行比较,以计算出差异,并相应地更新网络参数。此外,还实现了自适应图像缩放,以减少调整图像大小后经常出现的黑边,从而使推理速度得到显著提高。主干网络采用了 CSPDarknet53 和空间金字塔池化模块(spatial pyramid pooling fast, SPPF), CSPDarknet53 模块将 CSPNet^[21]和 Darknet53 网络进行了结合,在保证准确率的同时,降低网络计算量,SPPF 模块的主要目的是通过在融合过程中使用多轮 maxpool 来提取和融合高层次的特征,其目的是捕捉尽可能多的高层次语义特征,SPPF 相对于 SPP^[22],在保证其精准度的同时,使计算速度快了不止两倍。颈部采用了融合特征金字塔网络(feature pyramid networks, FPN)和路径聚合网络(path aggregation network, PANet)^[23],FPN 在所有尺度上采用自上而下的横向联系,来构建高层次的语义特征图,呈现出特征金字塔的经典结构,而 PANet 是一种实例分割算法,具体结构其实是对特征的反复提取,在通过 FPN 的几个层后,低层的目标信息可能会变得模糊。为了弥补这一点,PANet 加入了一个自下而上的途径来加强定位信息。检测头则是沿用

了 YOLOv3^[24] 的检测头。YOLOv5 的网络结构如图 4 所示。

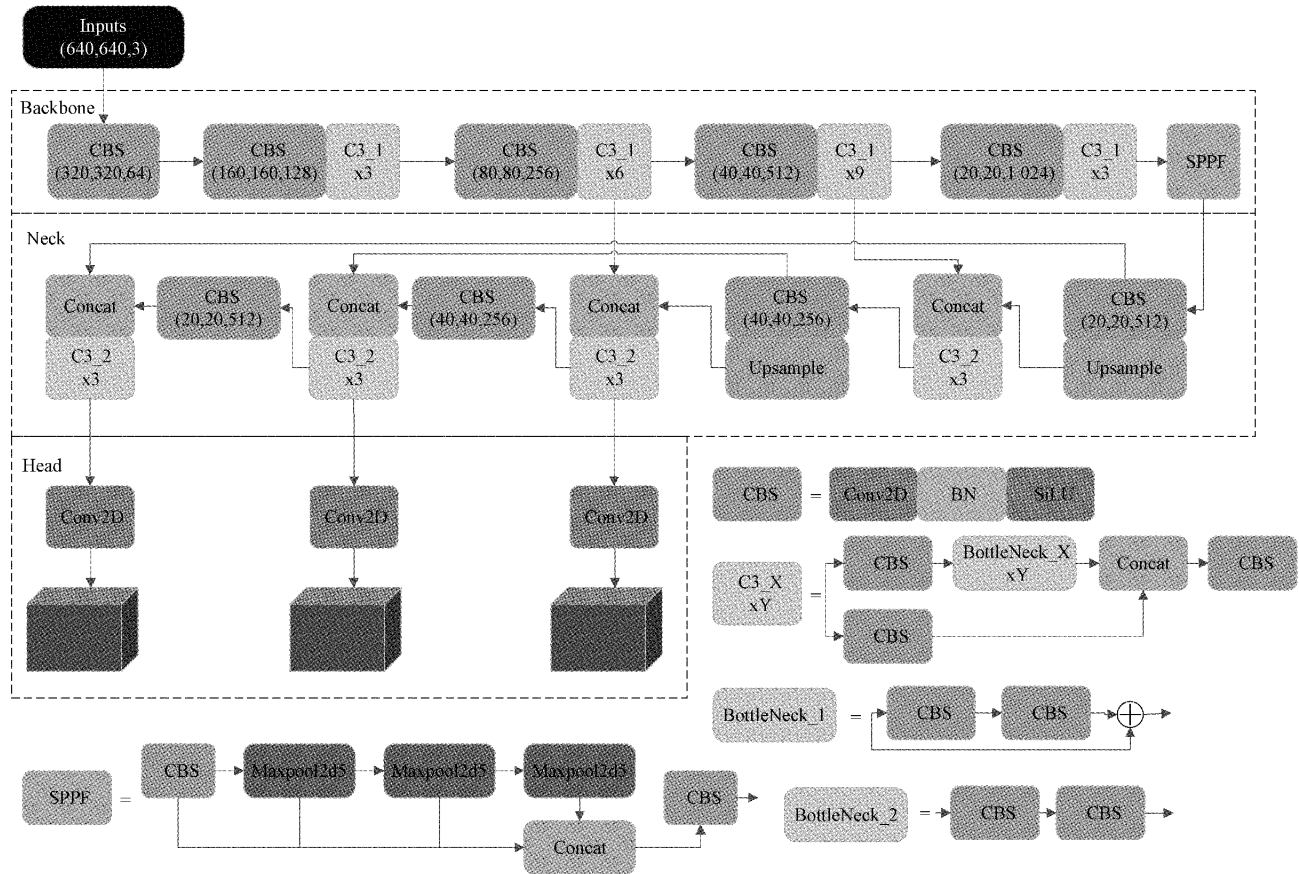


图 4 YOLOv5 的网络结构

2.3 改进的 YOLOv5 模型

为了减少图像背景噪声对检测网络中浅层特征信息的影响,我们在 YOLOv5 物体检测网络引入了注意力机制,以提高对小而密集的目标的检测能力。注意力机制是机器学习常用的一种数据处理技术。当应用于物体检测时,它可以有效地提高检测性能。注意力机制可分为三种类型:通道注意力机制、空间注意力机制和混合域注意力机制,本文应用的是一种基于卷积的混合域注意力机制 CBAM(convolutional block attention module)^[25],CBAM 模块可以增强深度学习中特征的适应性,通过沿着给定特征图的通道和空间维度,来计算注意力图。然后将注意力图与输入特征图相乘,以完善自适应特征。通过训练,CBAM 模块能够区分需要强调的信息和需要抑制的信息,帮助模型在处理数据时做出更好的决策。

CBAM 的网络结构如图 5 所示,给定一张中间特征图 $F \in \mathbb{R}^{C \times H \times W}$ 作为输入,模型可以按顺序得出一维通道注意力图 $M_c \in \mathbb{R}^{C \times 1 \times 1}$ 和二维空间注意力图 $M_s \in \mathbb{R}^{1 \times H \times W}$,输出的整个过程可以概括为:

$$\begin{cases} F' = M_c(F) \otimes F \\ F'' = M_s(F') \otimes F' \end{cases} \quad (1)$$

通道注意力机制模块(channel attention module,

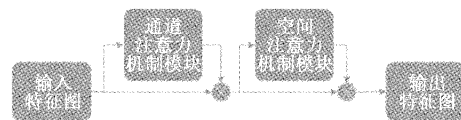


图 5 CBAM 的网络结构

CAM)是一个利用通道间特征关系生成通道注意力图的模块。其结构如图 6 所示。首先,特征图的空间信息用两种操作进行聚合:平均池化和最大池化,以生成两个不同的空间背景数据脚本;平均池化特征 F_{avg}^c 和最大池化特征 F_{max}^c ,然后将这两个数据脚本运行到共享网络,以产生通道注意力图 $M_c \in \mathbb{R}^{C \times 1 \times 1}$,该共享网络由多层感知器(MLP)和一个隐层组成。为了减少参数开销,隐藏层激活大小被设置为 $\mathbb{R}^{C/r \times 1 \times 1}$,其中 r 是减少比率。在对每个数据脚本应用共享网络后,通过对元素进行求和并输出,得到特征向量。通道注意力机制的公式:

$$M_c(F) = \sigma(MLP(AvgPool(F)) + MLP(MaxPool(F))) = \sigma(W_1(W_0(F_{avg}^c)) + (W_1(W_0(F_{max}^c)))) \quad (2)$$

其中, σ 为 sigmoid 函数, $W_0 \in \mathbb{R}^{C/r \times C}$ 和 $W_1 \in \mathbb{R}^{C \times C/r}$ 为 MLP 的权重,他们共享两个输入,RELU 激活函数后面是 W_0 。

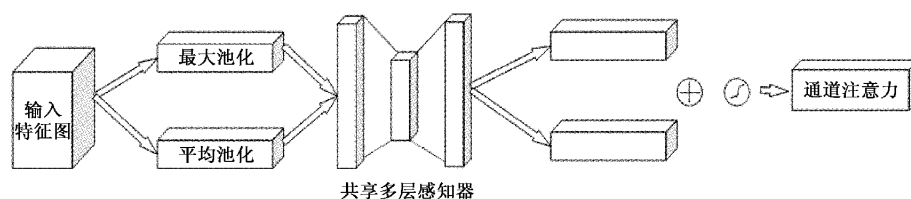


图 6 通道注意力结构

空间注意力机制模块(spatial attention module, SAM)是一个利用特征空间之间的关系来生成空间注意图的模块,如图 7 所示。与通道注意力模块不同,空间注意模块侧重于捕捉特征图中的“位置”信息,是对通道注意力模块的补充。计算空间注意力,首先使用两个池化层来聚合特征映射的通道信息,生成两个二维图:平均池化特征图 $F_{avg}^s \in \mathbb{R}^{1 \times H \times W}$ 和最大池化特征图 $F_{max}^s \in \mathbb{R}^{1 \times H \times W}$,然后通过标准卷积层将它们串联和卷积,生成我们的二维空间注意力图。空间注意力机制的公式为:

$$M_s(F) = \sigma(f^{7 \times 7}([\text{AvgPool}(F); \text{MaxPool}(F)])) = \sigma(f^{7 \times 7}([F_{avg}^s; F_{max}^s])) \quad (3)$$

其中, σ 为 sigmoid 函数, $f^{7 \times 7}$ 表示此处的滤波器为大小为 7×7 的卷积核。

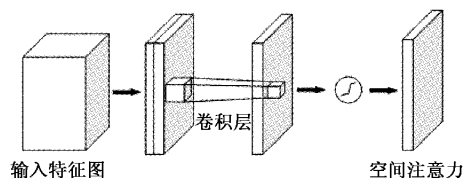


图 7 空间注意力结构

无人机拍摄的莒茛属作物目标大多为小且密集目标,而在深层网络中,这些小而密集的目标在特征图中的特征往往会逐渐减弱,导致漏检和误检。CBAM 模块的集成可以有效增强特征图中的特征信息,使其更容易提取通道和空间特征。因此,这有助于缓解漏检和误检的问题。为了达到最好的效果,通过对比实验,本文尝试了如下几种方案:将 CBAM 模块置于 Backbone 的最后一层、置于 Backbone 的卷积层中和置于 C3 模块中,最后发现将 CBAM 模块加载于 Backbone 中 C3 模块的 BottlenNeck 效果最佳,Backbone 中的 BottlenNeck 原本是由两个 Conv_BN_SiLU 模块组成,本文在第二个 Conv_BN_SiLU 模块后嵌入 CBAM 模块,嵌入结构如图 8 所示,新加入的模块在图中展示为蓝色。嵌入结构后不影响原模型的基础结构,因此训练方式不变。

3 实验结果与分析

3.1 实验环境

本研究使用的无人机为大疆 PHANTOM 4 PRO 型无人机(深圳市大疆创新科技有限公司),目标检测模型训练使用的机器,操作系统为 windows10,处理器 12th Gen

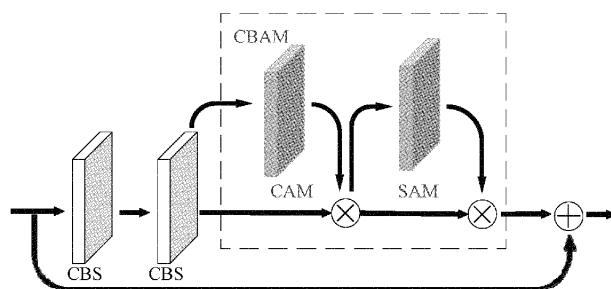


图 8 YOLOv5 改进部分

Intel(R) Core(TM) i9-12900K 5.50 GHz,图形处理器为 NVIDIA GeForce RTX 3070 8 GB,驱动程序版本 31.015.2698,Python 版本 3.9.15,pytorch 版本 1.13.0, cuda 版本 11.7.64,cudnn 版本 8.5.0。

3.2 三维模型计算与网络训练

将无人机环绕拍摄的图像导入 ContextCapture,空间坐标系选择 WGS 84/UTM zone 48 N (EPSG:32648),生成三维网格,再计算其对应的 DOM/DSM 图像,进而得到高程模型。CBAM-YOLOv5 网络的训练,输入网络的图像尺寸被调整为 640×640 ,batch-size 设置为 12,epochs 设置为 150 次,衰减系数设置为 0.000 5,初始学习率为 0.01,因为设置了余弦学习率,当 epoch 达到 150 时,学习率衰减至 0.000 232,网络在迭代次数达到 100 次时,各项指标趋于平稳,模型逐渐收敛。

3.3 性能指标

目标检测模型网络性能的评定指标,一般为被检目标的准确率(precision, P)、召回率(recall, R)以及平均精度(AP),其中平均精度为不同召回率情况下准确率的均值(等同于 P-R 曲线的积分),主要指标的计算公式如下:

$$\begin{cases} P = \frac{TP}{TP + FP} \\ R = \frac{TP}{TP + FN} \\ AP = \sum_{i=1}^{Rank} (R_i - R_{i-1}) \times P_i \end{cases} \quad (4)$$

其中, P 为准确率; R 为召回率; TP 为预测值和真实值相同,预测值为正样本; TN 为预测值和真实值相同,预测值为负样本; FP 为预测值和真实值不同,预测值为正样本; FN 为预测值和真实值不同,预测值为负样本; Rank 为迭代次数。

3.4 实验结果与分析

1) 目标估算的株高检测性能分析

为了评估改进的目标检测算法在所需应用环境下的性能优劣,在相同的实验环境下,使用不同的目标检测算法,进行了多组对比实验,其中 Faster-RCNN 是目前最常用的基于候选区的两阶段的算法,YOLOv3、YOLOv4 则为 YOLO 网络的前几代模型,也为常用的基于回归的单阶段算法,对比实验的评价指标为 AP、Recall 和检测速度,实验结果如表 1 所示。

表 1 不同目标检测算法结果对比

网络结构	AP/%	Recall/%	Time/s
Faster-RCNN	93.58	90.66	0.041
YOLOv3	89.56	85.16	0.027
YOLOv4	93.10	87.85	0.024
YOLOv5	94.72	90.80	0.022
CBAM-YOLOv5	96.19	93.87	0.015
本文模型			

根据表 1 的数据可以得知,尽管两阶段算法 Faster-RCNN 在平均精度方面表现良好,但由于其结构更深,导致检测时间变长。相比于 YOLOv3,CBAM-YOLOv5 引入了 Mosaic 数据增强、自适应图片缩放、自适应锚框计算,改进了激活函数并融合了 CBAM 注意力机制,其 mAP 值提升了 6.63%,召回率提升了 8.71%,同时检测时间也得到缩短;由消融实验可得相比于原始的 YOLOv5,本文模型 AP 值提升了 1.5%,Recall 提升了 3%,检测时间也缩短了 32%。

为了试验出最适合检测莒荻属作物目标的网络模型改进方式,还对相同网络不同改进策略下的模型效果进行了对比试验,分别将 SENet (squeeze-and-excitation networks)、ECANet (efficient channel attention networks)、CANet (coordinate attention networks)、CBAMNet (convolutional block attention module networks) 嵌入到网络模型中,其中嵌入到 Backbone 最后一层的为后缀为 a 的模型,嵌入到 Backbone 中 C3 模块的 Bottleneck 中的为后缀为 b 的模型,对比实验的评价指标为 AP、Recall、Precision 和检测速度,实验结果如表 2 所示。

由表 2 可知,SENet 为通道注意力机制,加入网络后,

表 2 不同改进 YOLOv5 模型在测试集上的性能评估

网络结构	改进策略	AP/%	Recall/%	Precision/%	Time/s
YOLOv5		94.72	90.80	89.86	0.022
SE-YOLOv5	a	94.27	88.65	93.91	0.017
	b	94.91	91.10	90.46	0.013
ECA-YOLOv5	a	95.25	89.18	89.13	0.011
	b	95.54	89.89	90.79	0.012
CA-YOLOv5	a	95.01	92.02	92.73	0.014
	b	95.20	92.94	90.55	0.016
CBAM-YOLOv5	a	95.61	90.67	91.34	0.022
	b(本文模型)	96.19	93.87	93.67	0.015

各项性能指标都有所提升;ECANet 将 SENet 中的 MLP 替换为了一维卷积,可以有效减少参数计算量,所以相较于 SENet 可以大大缩短网络训练的时间,检测时间也更短^[26];CANet 在 SENet 原始的网络模型基础上再加入了位置信息,使得模型能更准确地定位和识别目标区域,可以看到各个性能指标提升了接近 1%;CBAMNet 则是在 ECANet 基础上再加上了空间注意力模块,实现了通道注意力和空间注意力的双机制,更利于莒荻属作物小、密目标的检测;而模型 b 的各个性能指标皆优于模型 a,因此最后选择了 CBAM-YOLOv5-b 模型,可以看到,相较于引入其他注意力机制模型,本文模型的 AP 值最多高出 1.9%,Recall 最多提升了 5%,Precision 最多提升了 4%。

2) 株高测定性能分析

为评估株高测定的性能优劣,对试验区范围内的 200 棵植株进行人工测高,并与机器估算值进行比较,生成的

拟合曲线、 R^2 和 RESM 如图 9 所示。可以看到,估算的株高与实测值具有很好的相关性,直线斜率为 0.991 1, R^2 达到了 0.931 1。

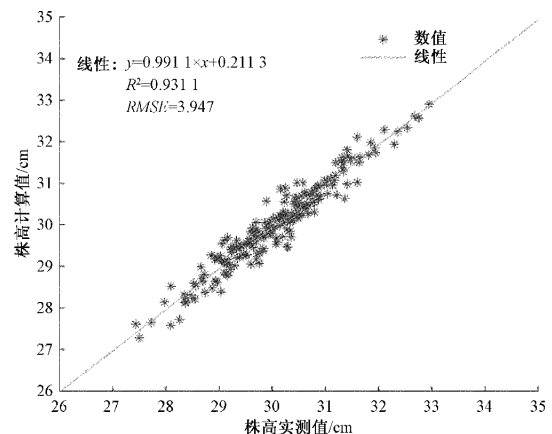


图 9 株高测定结果

4 结 论

为了实现对大面积密集莠苣属作物的株高测量,本文提出了一种结合深度学习目标检测和无人机 RGB 图像建模的株高测量方法。首先针对无人机图像中密集目标,我们提出了改进的目标检测模型 CBAM-YOLOv5,有效提升了模型对密集目标的检测效果,在密集目标的对比实验中,我们模型的 AP 值达到了 96.19%, Precision 为 93.67%,均高于同类型的其他模型。然后使用 ContextCapture 对无人机航拍影像进行三维建模,并生成需要的 DSM 与 DOM 图像,建立使用 GetGeoTransform 的智能算法模型,对 DSM 图像中大量的数据进行读取,导出莠苣属作物对应位置高程数据。由实验结果可知,误差控制在极小的范围内,使用本文株高检测模型,可以满足农业上对大规模高通量莠苣属作物株高的测算。但本文方法还存在一些局限性:植株侧弯、畸变会对株高的计算造成影响;遮挡严重会造成植株目标检测的丢失。对此,已经开展对相关问题的研究,从更多的维度去改进本文的模型,从而为莠苣属作物农业的数字化、智能化的育种育苗提供更坚实的数据基础和技术支撑。

参考文献

- [1] ZHANG L, GRIFT T E. A LiDAR-based crop height measurement system for *Miscanthus giganteus* [J]. *Computers and Electronics in Agriculture*, 2012, 85: 70-76.
- [2] 郭海,樊江川,李英伦,等.基于 RGB-D 点云的田间原位玉米株高测量试验研究[J].*农机化研究*, 2021, 43(10):102-109.
- [3] JIMENEZ-BERNI J A, DEERY D M, ROZAS-LARRAONDO P, et al. High throughput determination of plant height, ground cover, and above-ground biomass in wheat with LiDAR [J]. *Frontiers in Plant Science*, 2018, 9: 237.
- [4] HU P, CHAPMAN S C, WANG X, et al. Estimation of plant height using a high throughput phenotyping platform based on unmanned aerial vehicle and self-calibration: Example for sorghum breeding[J]. *European Journal of Agronomy*, 2018, 95: 24-32.
- [5] YAMAMOTO K, GUO W, NINOMIYA S. Node detection and internode length estimation of tomato seedlings based on image analysis and machine learning[J]. *Sensors*, 2016, 16(7): 1044.
- [6] XIE Q, WANG J, LOPEZ-SANCHEZ J M, et al. Crop height estimation of corn from multi-year RADARSAT-2 polarimetric observables using machine learning[J]. *Remote Sensing*, 2021, 13(3): 392.
- [7] JIN X, MADEC S, DUTARTRE D, et al. High-throughput measurements of stem characteristics to estimate ear density and above-ground biomass[J]. *Plant Phenomics*, 2019.
- [8] 陈梅香,张瑞瑞,陈立平,等.无人机农林业应用全球研究态势分析[J].*智慧农业*, 2021, 3(3):22-37.
- [9] 罗会兰,彭珊,陈鸿坤.目标检测难点问题最新研究进展综述[J].*计算机工程与应用*, 2021, 57(5):36-46.
- [10] 梁鸿,王庆玮,张千,等.小目标检测技术研究综述[J].*计算机工程与应用*, 2021, 57(1):17-28.
- [11] EVERINGHAM M, ESLAMI S M A, VAN GOOL L, et al. The pascal visual object classes challenge: A retrospective[J]. *International Journal of Computer Vision*, 2015, 111: 98-136.
- [12] 郑特,曹明,宋伟.一种利用三维模型自动获取数字正射影像和数字表面模型的方法[J].*测绘与空间地理信息*, 2022, 45(11):211-212, 216.
- [13] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2014: 580-587.
- [14] REN S, HE K, GIRSHICK R, et al. Faster r-cnn: Towards real-time object detection with region proposal networks [J]. *Advances in Neural Information Processing Systems*, 2015, 28.
- [15] HE K, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN [C]. *Proceedings of the IEEE International Conference on Computer Vision*, 2017: 2961-2969.
- [16] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector [C]. *Computer Vision-ECCV 2016: 14th European Conference*, 2016: 21-37.
- [17] ZHOU X, WANG D, KRÄHENBÜHL P. Objects as points[J]. *ArXiv Preprint*, 2019, ArXiv:1904.07850.
- [18] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection [C]. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016: 779-788.
- [19] 单明陶,高玮玮.改进 YOLOv4 的内丝接头密封面缺陷检测算法[J].*电子测量与仪器学报*, 2022, 36(5): 120-127.
- [20] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. Yolov4: Optimal speed and accuracy of object detection[J]. *ArXiv Preprint*, 2020, ArXiv:2004.10934.
- [21] WANG C Y, LIAO H Y M, WU Y H, et al. CSPNet: A new backbone that can enhance learning capability of CNN [C]. *Proceedings of the IEEE/CVF*

- conference on computer vision and pattern recognition workshops, 2020: 390-391.
- [22] HE K, ZHANG X, REN S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1904-1916.
- [23] WANG K, LIEW J H, ZOU Y, et al. Panet: Few-shot image semantic segmentation with prototype alignment [C]. Proceedings of the IEEE/CVF International conference on computer vision, 2019: 9197-9206.
- [24] REDMON J, FARHADI A. YOLOv3: An incremental improvement[J]. ArXiv Preprint,2018, ArXiv:1804.02767.
- [25] WOO S, PARK J, LEE J Y, et al. Cbam: Convolutional block attention module[C]. Proceedings of the European Conference on Computer Vision(ECCV), 2018: 3-19.
- [26] HU J, SHEN L, SUN G. Squeeze-and-excitation

networks[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 7132-7141.

作者简介

贺星耀, 硕士, 主要研究方向为计算机视觉、图像识别及三维建模相关应用。

E-mail: he_xingyao@foxmail.com

冯涛(通信作者), 硕士, 教授, 中国计算机学会会员。主要研究方向为机器学习、智能信息处理和网络安全。

E-mail: vonpower@ynufe.edu.cn

梁虹, 硕士, 教授, 主要研究方向为深度学习和信号处理。

E-mail: liangh@ynu.edu.cn

吴凯香, 硕士, 主要研究方向为 WebGIS 系统开发、三维数字重现。

E-mail: wu1831499391@163.com

袁嘉辉, 硕士, 主要研究方向为基于深度学习的图像分类和识别及其应用。

E-mail: joseph_yung@sina.com