

基于改进 YOLOv8n 的轻量化茶叶嫩芽检测方法^{*}潘海鸿¹ 陈希良¹ 钱广坤¹ 申毅莉² 陈琳¹

(1. 广西大学机械工程学院 南宁 530004; 2. 梧州学院机械与资源工程学院 梧州 543002)

摘 要: 为解决自然环境下茶叶嫩芽检测场景复杂,模型参数量大无法在嵌入式设备部署等问题,提出一种基于改进 YOLOv8n 的轻量化茶叶嫩芽检测方法。构建一种 MFBNet 轻量化骨干网络,引入 MBCnv 模块后大大减少了模型计算量。同时在骨干网中加入 CBAM 注意力模块,抑制无效信息,提高了模型检测精度;其次引入 AKCnv 模块对 VoVGSCSPC 结构进行改进,提出全新的 AVCStem 模块,并将其替换颈部网络的 C2f 模块,进一步减少模型参数,提升嵌入式设备部署效率;最后采用 GSCnv 模块替换颈部网络结构中的全部 Conv 模块,帮助模型进行快速计算,提高茶叶嫩芽的检测速率。结果表明,本文提出的模型比 YOLOv8n 原模型的 mAP50 和 FPS 分别提升了 3.5%、55.6%,参数量减少了 14.3%,且模型鲁棒性强,满足复杂场景下茶叶嫩芽的轻量化快速检测。

关键词: 茶叶嫩芽检测;轻量化;注意力机制;深度学习;YOLOv8n

中图分类号: TN919.8;TP391.46 **文献标识码:** A **国家标准学科分类代码:** 520.6040

Research on lightweight tea sprout detection method based on improved YOLOv8n

Pan Haihong¹ Chen Xiliang¹ Qian Guangkun¹ Shen Yili² Chen Lin¹

(1. School of Mechanical Engineering, Guangxi University, Nanning 530004, China;

2. School of Mechanical and Resource Engineering, Wuzhou University, Wuzhou 543002, China)

Abstract: To solve the problems of complex tea bud detection scenarios in natural environments and the large number of model parameters that cannot be deployed on embedded devices, a lightweight tea bud detection method based on YOLOv8n is proposed. We construct a lightweight backbone network, MFBNet, introducing the MBCnv module to significantly reduce model computation. Simultaneously, we incorporate the CBAM attention module into the backbone network to suppress irrelevant information, thereby enhancing the model's detection accuracy. Furthermore, the introduction of the AKCnv module improves the VoVGSCSPC structure, proposing the innovative AVCStem module, which replaces the C2f module in the neck network, further reducing model parameters and enhancing the efficiency of embedded device deployment. Finally, we employ the GSCnv module to replace all Conv modules in the neck network structure, facilitating fast model computation and increasing the detection speed of tender tea buds. The results indicate that the proposed model in this paper outperforms the original YOLOv8n model with a 3.5% improvement in mAP50, 55.6% increase in FPS, and a 14.3% reduction in parameters. The model demonstrates strong robustness, meeting the requirements for lightweight and rapid detection of tender tea buds in complex scenarios.

Keywords: tea bud detection; lightweight; attention mechanism; deep learning; YOLOv8n

0 引 言

茶叶采摘是茶叶生产过程中至关重要的一步,不仅影响到茶叶的质量、产量,还影响到茶树的生长和经济效益的增长^[1]。目前我国常见的茶叶采摘方式主要分为人工采茶和机械采茶^[2]。人工采茶速度慢,效率低,且劳动力成本

高;机械采茶不能保证茶叶的完整性,且容易造成挫伤茶树,破坏茶叶结构等问题。随着目标检测技术的快速发展,实现茶叶智能化采摘是未来的发展趋势,其前提和关键是实现茶叶嫩芽的快速识别。因此,茶叶嫩芽检测的研究对于茶叶行业的发展具有重要价值和意义。

传统的茶叶嫩芽检测方法主要是根据茶叶中的颜色差

异以及纹理特征将嫩芽与老叶区分开来^[3-7],通过获取茶叶图像的 RGB 色彩空间中的各个分量,接着使用传统的方法对图像进行分割处理,最后根据嫩芽的形状纹理特征,从而有效地识别出茶叶嫩芽。然而传统的图像识别算法在复杂的环境和场景下的鲁棒性较差,特别是对于茶芽和老叶背景下颜色和纹理特征区别不明显的情况下,识别效果并不理想。

目前,随着基于卷积神经网络的目标检测算法不断发展,先后涌现出区域卷积神经网络(region-based convolutional neural network,R-CNN)^[8]、快速区域卷积神经网络(fast region-based convolutional neural network, Fast R-CNN)^[9]、更快的区域卷积神经网络(faster region-based convolutional neural network, Faster R-CNN)^[10]等 two-stage 网络模型,以及你只需看一次(you only look once, YOLO)^[11-13]、单次多框选择器(single shot multiBox detector, SSD)^[14]等 one stage 网络模型。Chen 等^[15]利用卷积神经网络 Faster R-CNN 来检测茶园复杂环境下的茶叶嫩芽,该方法对于识别检测自然环境下的茶叶嫩芽具有较好鲁棒性;王子钰等^[16]利用改进后的 SSD 算法对茶叶嫩芽进行自动检测识别,实验结果符合预期;Yang 等^[17]提出了一种基于 YOLOv3 算法的茶叶嫩芽采摘点识别方法,通过利用特征金字塔结构获取茶叶嫩芽的特征图,随后采用 K-means 方法对图像目标框的维度进行聚类,该方法在茶叶嫩芽数据集上的准确率达到 90% 以上;黄家才等^[18]对 YOLOv4 算法的主干网络和颈部网络进行改进,分别用 GhostNet 网络替换 Backbone 网络,用 Ghost 卷积替换 Neck 网络中的传统卷积。改进后的模型参数量更小,检测速度更快,为茶叶嫩芽采摘机器人的研究提供了理论支持;王梦妮等^[19]提出一种基于改进 YOLOv5s 网络模型的检测算法,通过修改现有的网络结构来提高模型的检测精度,改进后的网络模型在多个场景下对于茶叶嫩芽的检测具有较高置信度;俞龙等^[20]提出一种改进 YOLOX 的茶叶嫩芽识别算法 SS-YOLOX,通过加入 SE 模块来提高模型的特征提取能力,引入 Soft NMS 算法改善模型检测框重叠较高的问题。消融实验表明,两种改进策略均能有效地提高模型的检测精度。

基于手工特征的传统目标检测算法针对复杂环境下的目标识别泛化能力较差,需要花费足够的时间资源和人力成本进行特征设计和调优,难以满足实际需求。基于卷积神经网络的目标检测算法自适应能力强,可满足不同场景的物体的识别检测,且相比传统的目标检测算法精度更高速度更快。

综上,尽管在复杂环境下茶叶嫩芽图像检测方面取得了一些研究成果,但依然存在一些挑战:复杂背景中茶叶嫩芽密集排列和遮挡造成茶叶嫩芽特征显著性低,算法快速难以从复杂背景中提取有效茶叶嫩芽特征,严重影响嫩芽检测的精度和速度。为此,本文提出一种轻量化茶叶嫩芽

图像检测方法,通过改进网络结构,以加强茶叶嫩芽检测的性能,为茶叶嫩芽智能采摘提供技术支持。

1 网络设计与改进

1.1 改进的 YOLOv8n 网络模型

本文改进的 YOLOv8n 网络模型由 MFBNet 主干网络结构、AVCPFPN 颈部网络结构和头部预测网络结构组成,增强了在目标特征提取、多特征融合和预测输出方面的性能。为了平衡检测精度和速度,本文采用 YOLOv8n 模型作为茶叶嫩芽检测的基础模型,在 YOLOv8n 中引入移动反向瓶颈卷积(mobile inverted bottleneck convolution, MBConv)^[21]模块,并且在 MBConv 结构中融入卷积注意力模块(convolutional block attention module, CBAM)^[22],其次引入由 FasterNet 网络中提出的 BSBLOCK 模块,构建得到 MFBNet 主干网络以提高网络特征提取的能力。此外,提出由 GSConv 模块和 AVCStem 模块组成的 AVCPFPN 颈部网络结构与路径聚合网络(path aggregation network, PANet)^[23],特征金字塔结构结合最终得到 YOLOv8n-Tea 网络模型。YOLOv8n-Tea 网络模型如图 1 所示。

1.2 MFBNet 轻量化主干网络

EfficientNet^[24]是由 Google 提出的一种神经网络架构,其设计旨在保持高度准确性的同时,减少模型的参数量和计算复杂度。通过扩大网络输出的通道数量和模型的网络层数量,不仅能够增强对茶叶嫩芽图像中茶叶检测的关注度,同时能够有效地防止计算量过大以及梯度消失的问题,从而取得最佳检测性能^[25]。EfficientNet 在相对较小的模型尺寸下能够达到与更大的传统网络相媲美的性能,这使得在资源受限的环境下更容易部署和使用,满足在复杂茶叶园林中对茶叶嫩芽检测的实时性。因此,本文采用了 EfficientNet 轻量化网络的混合缩放方法,以获取最优的网络宽度、深度和输入图像分辨率缩放系数,通过应用这些系数对网络的宽度、深度和输入图像分辨率进行缩放。优化后的 YOLOv8n-Tea 模型的主干网络结构 MFBNet 由多个 MBConv 模块和 BSBLOCK 模块组成。本文删除了 MBConv 模块中原有的压缩激励网络(squeeze-and-excitation networks, SENet)^[26],再引入 CBAM 注意力机制,以增加模型对茶叶嫩芽检测区域的关注度。而 BSBLOCK 模块是 FasterNet 网络提出的一种新型特征提取模块,在不大量增加参数的情况下,进一步提升模型精度。将 MFBNet 主干网络替换原模型 YOLOv8n 的主干网络,如表 1 所示。与 YOLOv8n 原主干网络相比,本文提出的主干网络能够使模型的检测性能得到较大提升。在 MFBNet 主干网络结构中,阶段 1 和 2 的 CBS 模块由卷积层、BN 层、SiLU 激活函数层组成,阶段 3~6 是构建的 MBConv 模块以及一个普通卷积模块,通过扩大模型中网络的宽度、深度来提高茶叶嫩芽的检测关注能力;而阶段 7~9 是由一个卷积模块和两个 BSBLOCK 模块组成;阶段 5~7

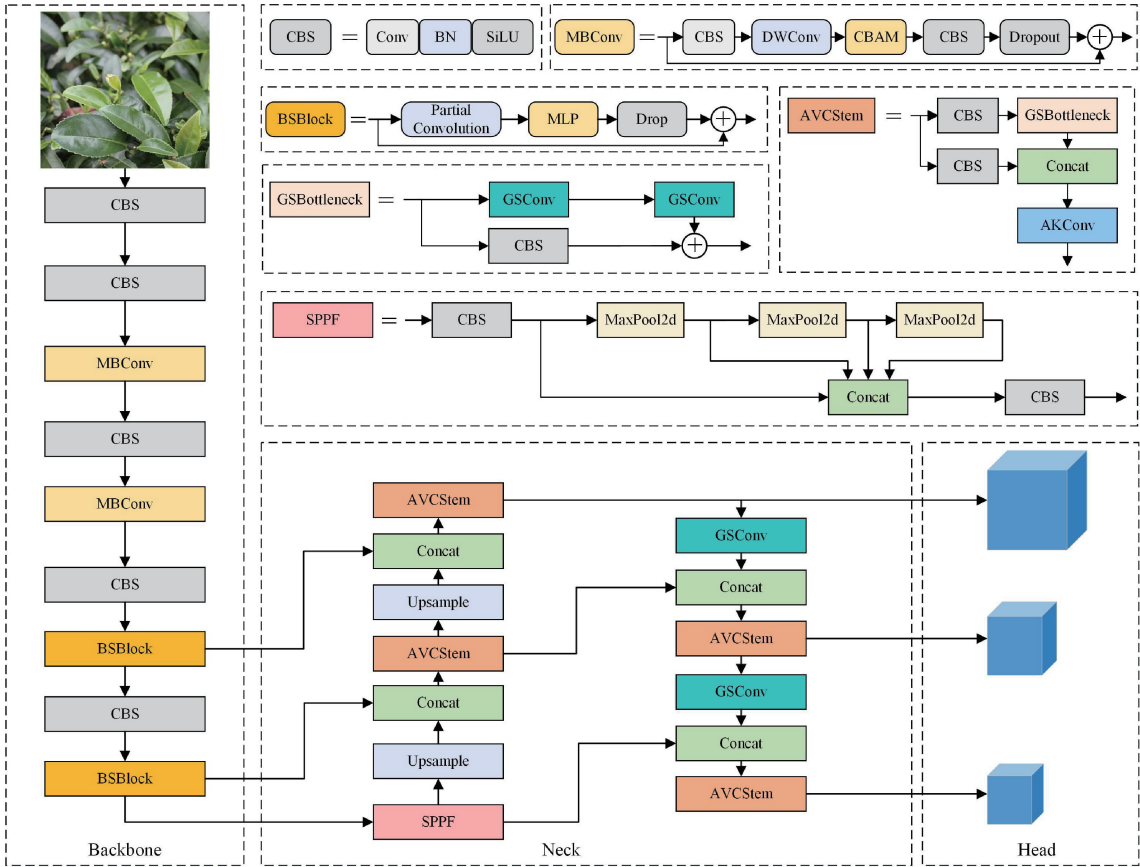


图 1 YOLOv8n-Tea 网络模型

分别提取不同尺度大小的特征图并进一步输入至模型的颈部网络;表 1 中阶段 3 位置的 MBConv 模块通道数较少特征提取速度快但检测精度低,而在阶段 5 位置的 MBConv 模块的通道数扩充为原来的 2 倍,其对茶叶嫩芽特征捕捉性能更加鲁棒。

表 1 MFBNet 轻量化主干网络结构

层数	模块	核大小	输出通道	步长
1	CBS	3×3	64	2
2	CBS	3×3	128	2
3	MBConv	3×3	128	1
4	CBS	3×3	256	2
5	MBConv	3×3	256	1
6	CBS	3×3	512	2
7	BSBlock	1×1	512	1
8	CBS	3×3	1 024	2
9	BSBlock	1×1	1 024	1

因此,本文在骨干网络中使用了 MFBNet 来替换 YOLOv8 原始的主干网络结构。通过 MBConv 和 BSBBlock 的卷积运算可以减少计算和内存消耗并保持高精度。它采用多个跨层连接和注意力机制来提高模型的鲁棒性和泛化能力,使其具有高效性和可扩展性,适用于快

速对茶叶嫩芽进行检测和分类任务。

1.3 GSConv 模块

Slim-neck^[27]最初被引入到无人驾驶车辆的视觉系统中,其核心思想是在轻量化颈部网络的同时保持高效特征提取的骨干,以实现模型尺寸的减小,并确保保持检测精度。在 Slim-neck 中,GSConv 的结构如图 2 所示。

它融合了 GhostNet 和 ShuffleNetv2 的轻量级设计思想。GhostNet 的主要目标是解决标准卷积输出通常包含大量相似特征图的问题,从而导致计算冗余。GhostNet 引入了 Ghost 卷积,这不同于标准卷积,它首先通过标准卷积获取第一部分,接着通过深度卷积对该部分进行处理,生成几个相似的特征图,最终将这两部分连接起来形成输出特征图。ShuffleNetv2 的主要目标是解决深度可分离卷积中信道信息在计算过程中被分离的问题,进而导致信道之间的信息交互相互独立。Shuffle 操作沿着通道方向重新排列特征图,以较低的计算成本保持通道之间的信息相互交互。因此,本文采用 GSConv 模块替代 YOLOv8n 颈部网络中的 Conv 模块,以降低计算复杂性,使其更适合在轻量化设备端进行快速茶叶嫩芽检测。

1.4 AVCSstem 结构

在茶叶嫩芽检测中,YOLOv8n 网络在提取图像特征信息时容易受到周围叶片纹理干扰,使得模型只关注局部

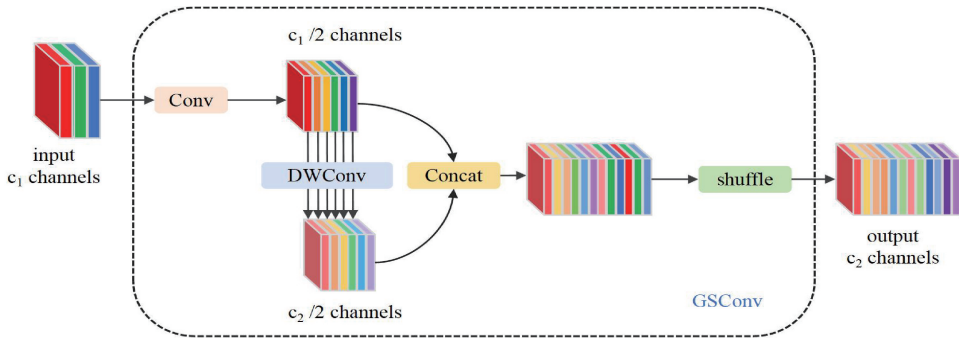


图 2 GSCnv 模块

像素位置。为了有效利用上下文捕获目标信息,有必要重复堆叠增强特征提取的卷积层。但直接重复堆叠这些层会导致计算效率低下和优化模型的困难,本文采用

VoVGSCSPC 结构和 AKConv 模块构建了全新的 AVCStem 结构来避免这一问题。该模块显著提高了模型捕获上下文信息的能力。AVCStem 结构如图 3 所示。

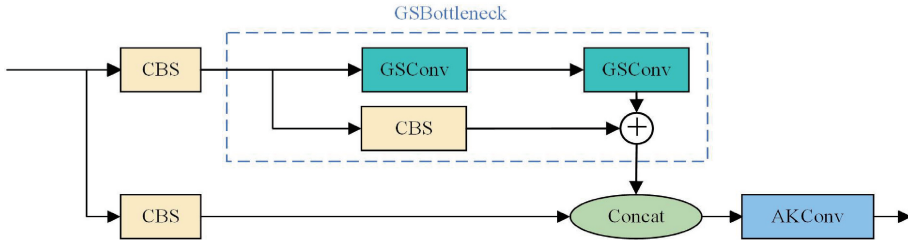


图 3 AVCStem 结构

GSBottleneck 模块包括用于空间特征交互的 CBS 模块和两个 GSCnv 模块,其中两个 GSCnv 模块的核大小分别为 1×1 和 3×3 ,以增强整体 GSBottleneck 模块多尺度特征的融合能力,同时加快特征提取速度。VoVGSCSPC 模块中通过集成 GSBottleneck 模块,创建了一个特征提取能力更强的 AVCStem 模块,具有高效率 and 轻量化的特点。因此,将 AVCStem 模块替换 YOLOv8 中 C2f 模块中,从而使网络对特征提取更加丰富,以满足快速检测和高精度识别目标的需求。

1.5 CBAM 注意力

针对复杂园林下的茶叶嫩芽成像环境特征不明显,引入了 CBAM 注意力来提高整体模型对茶叶嫩芽检测区域的检测关注度。CBAM 模块为了使网络能够专注学习重要特征,忽略不重要的特征,分别考虑了通道和空间特征,形成一种新的卷积块结构,从而更好地进行特征融合,并且有助于模型更加集中地关注输入图像的重要部分,从而提高茶叶嫩芽检测模型的鲁棒性,克服了 YOLOv8n 主干网络在模糊、纹理、强光环境下对局部茶叶嫩芽重要信息特征提取能力不足,泛化能力较差等问题。

若网络的输入特征图为: $F \in R^{C \times H \times W}$,其中 F 为输入特征图,通道数为 C 、高度为 H 、宽度为 W 的特征图。该模块将根据通道维 $\omega_c \in R^{1 \times 1 \times C}$ 和空间维 $M_s \in R^{1 \times H \times W}$ 来推断注意力图,然后将注意力机制图与输入特征图相乘,自适应细化特征,计算公式如式(1)所示。

$$\begin{aligned} F' &= \omega_c(F) \otimes F \\ F'' &= M_s(F') \otimes F' \end{aligned} \quad (1)$$

式中: \otimes 为元素级乘法, F'' 为最终特征输出。CBAM 各注意力子模块结构图如图 4 所示。

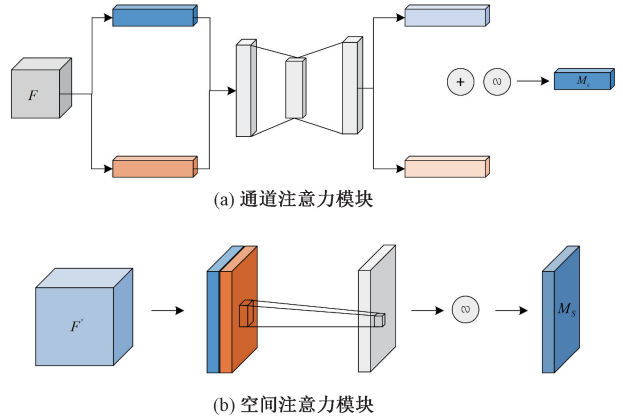


图 4 CBAM 模块结构

如图 4 所示,首先,通过对特征图并行平均池化和最大池化操作,聚合其空间信息,生成两个不同的空间描述符: F_{avg}^c 和 F_{max}^c , 分别表示平均池化特征与最大池化特征。其次,将这两个描述符输入到共享全连接层中,以生成通道注意力特征图,其大小为 $C \times 1 \times 1$,计算公式如式(2)所示。该共享全连接层由多层感知器(MLP)和一个隐藏层组成。为了减少参数,隐藏层的激活大小设置为 $R^{C/r \times 1 \times 1}$,

其中 r 为减少率。

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

式中: σ 为 Sigmoid 激活函数, $W_0 \in R^{\frac{C}{r} \times C}$, $W_1 \in R^{\frac{C}{r} \times C}$ 。

空间注意模块关注的是特征信息在图像中的位置,是对通道注意力模块的补充。首先,空间注意力模块在通道轴上采用平均池化和最大池化操作生成 F_{avg}^S 和 F_{max}^S 两个特征图,并将它们拼接成一个有效的特征描述符。最后由卷积层进行卷积运算,得到相应的空间特征图。具体如式(3)所示。

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

式中: σ 为 Sigmoid 激活函数, $f^{7 \times 7}$ 表示卷积核为 7×7 的卷积。

CBAM 注意力机制联合通道注意力机制与空间注意力机制,在不同层级上提高特征图的表达能力。通过使用平均池化和最大池化操作来聚合特征图的特征信息,从而更好地提取茶叶嫩芽的范围,有选择地关注与所需任务相关的特征^[28]。本文将 CBAM 注意力机制添加到 MFBNet 主干网络 Stage3 和 Stage5 的 MBConv 中,以减少检测茶叶嫩芽与复杂背景之间的互相干扰,从而提高网络模型的鲁棒性和泛化能力。

2 实验数据集

实验所使用的茶叶图像采集于广西三江侗族自治县古宜镇引木屯茶园,图像采集设备为 Canon EOS 550D 数码相机,分辨率为 3456×5184 像素,焦距为 $20 \sim 50$ mm,图像存储格式为 sRGB,总共获取 1 108 张茶叶原始图像,如图 5 所示。使用 Labelimg 图像标记软件对图像进行标记。

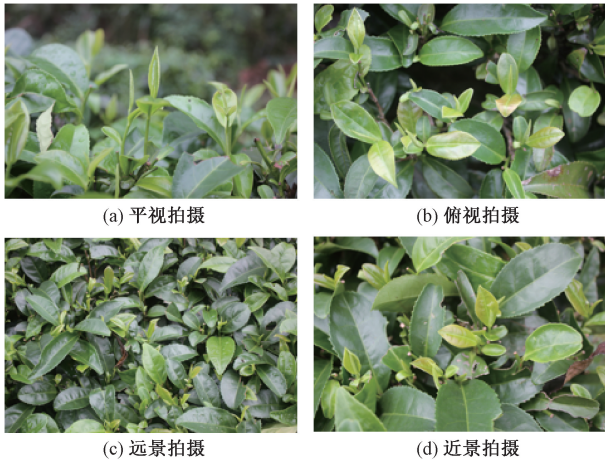


图5 茶叶嫩芽样本图像

为了防止测试集和建模数据集之间的重复,本文采用了先划分后增强的数据处理策略。首先,对训练集、验证

集、测试集按照 7:2:1 的比例随机划分。接着,进行数据增强操作,包括平移、翻转、旋转和亮度调整等,数据扩充后得到训练集 2 326 张、验证集 665 张、测试集 333 张。

3 实验结果与分析

3.1 实验环境及超参数配置

本文实验环境为 Ubuntu 18.04 版本 64 位操作系统、运行内存 64 GB,显卡为 NVIDIA GeForce RTX 4 090,显存为 24 G,CPU 为 Intel(R) Xeon(R) Platinum i9-13900k。采用 PyTorch 深度学习框架搭建模型,CUDA 版本为 11.1,cudnn 版本为 8.0.4。设置模型训练的输入图像大小为 640×640 ,轮数为 120,批次大小为 32,优化器为 SGD,初始学习率为 0.001,动量为 0.937 以及权重衰减率为 0.000 5。

3.2 评价指标

本文选取的评价指标为召回率(recall, R)、平均均值精度(mean average precision at IoU 0.5, mAP50)、参数量(params)和 FPS。其计算公式如下:

$$\text{Recall} = \frac{N_{TP}}{N_{TP} + N_{FN}} \quad (4)$$

$$\text{AP} = \int_0^1 P(r) dr \quad (5)$$

$$\text{mAP} = \frac{\sum_{i=1}^m \text{AP}_i}{m} \quad (6)$$

式中: $P(r)$ 表示 P-R 曲线, m 表示检测嫩芽类别数量, N_{TP} 表示被正确检测出的嫩芽样本数量, N_{FP} 表示误检的嫩芽样本数量, N_{FN} 表示未被检测到的嫩芽样本数量。

3.3 消融实验

本节验证了 MBConv、BSBlock、AVCStem 和 GSConv 在茶叶嫩芽检测任务中的有效性,实验结果如表 2 所示。在表中,√表示引入了相应的改进策略,而×表示未使用相应的改进策略。

由表 2 可知,本文引入 MBConv 结构在茶叶嫩芽检测方面具有良好效果,相比原模型的 mAP50 值提升了 0.7%,R 提升了 1.3%,这说明 MBConv 结构能够更多提取复杂茶叶嫩芽中的特征信息,扩展特征感受野;引入 BSBlock 模块后,YOLOv8n-v2 模型的 mAP 相比 YOLOv8n 提高了 2.1%,进一步增强了对特征显著性弱的茶叶嫩芽检测性能;在颈部网络引入 AVCStem 模块,虽然 YOLOv8n-v3 模型的 R 值相比 YOLOv8n-v2 下降了,但是参数量整体减少了 0.26 M,并且模型的检测速度有了较大提升,FPS 达到 236,说明本文提出的 AVCStem 模块对网络结构的检测速度和模型轻量化起到了积极的作用;最终 YOLOv8n-Tea 模型在 mAP50 上有较大提高,相比原模型提升了 3.5%,在整体目标检测速度上有较好的表现,FPS 相比原模型提升 55.6%,参数量减少 14.3%,这说明本文提出的 YOLOv8n-Tea 模型减少了参数量,同时也能保留

表 2 消融实验的性能对比

模型	MBConv	BSBlock	AVCStem	GSConv	R/%	mAP50/%	Params/M	FPS
YOLOv8n	×	×	×	×	91.4	94.2	3.00	169
YOLOv8n-v1	✓	×	×	×	92.7	94.9	3.12	219
YOLOv8n-v2	✓	✓	×	×	93.2	96.3	2.92	227
YOLOv8n-v3	✓	✓	✓	×	92.3	96.6	2.66	236
YOLOv8n-Tea	✓	✓	✓	✓	93.0	97.7	2.57	263

更多高层特征信息,能够在复杂背景下大幅度提升对茶叶嫩芽检测精度。YOLOv8n-Tea 的召回率、mAP50 以及参数量分别达到了 93.0%、97.7% 和 2.57 M,具有良好的茶

叶嫩芽检测性能。

在背景较为密集环境遮挡和较小的茶叶嫩芽目标图像下的检测效果对比,如图 6 所示。

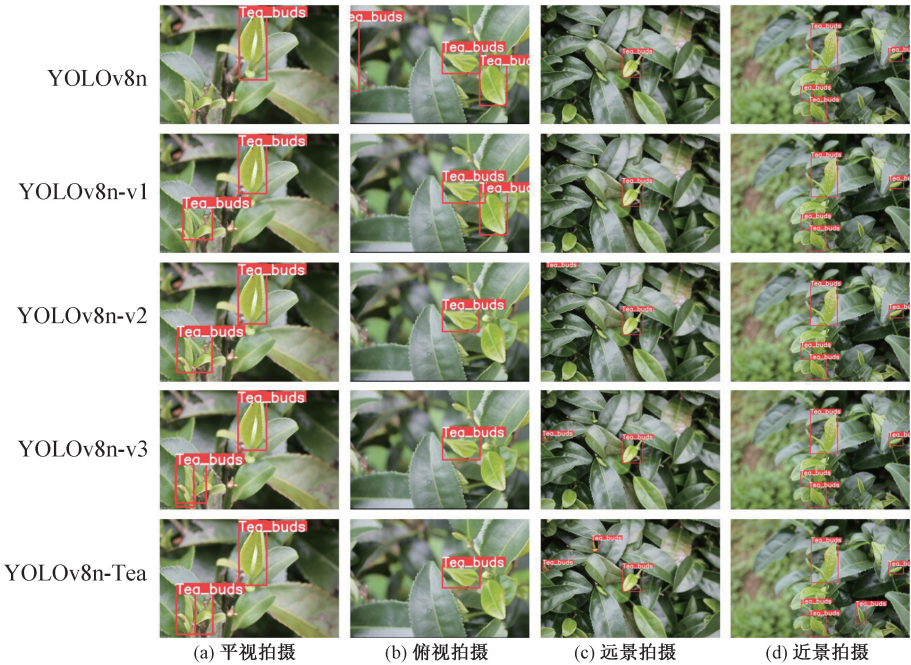


图 6 消融实验茶叶嫩芽检测结果

展示了平视拍摄、俯视拍摄、远景拍摄以及近景拍摄 4 种不同场景下的茶叶嫩芽检测效果。从图 6(a)~(d)中,可以看出 YOLOv8n 对平视拍摄的茶叶嫩芽检测效果较差,不能完全识别出茶叶嫩芽的目标区域。在引入 MBConv 和 BSBlock 模块后,YOLOv8n-v2 模型对嫩芽检测出的数量有所增加、检测位置更准确,说明 MFBNet 主干网络增强了对茶叶嫩芽特征的学习能力,提升了模型实时检测性能;YOLOv8n-v3 模型进一步准确定位图中所检测出的嫩芽区域,说明了本文新提出的 AVStem 模块丰富了模型的多尺度特征融合,加强了对目标区域的关注度;YOLOv8n-Tea 模型可正确识别出所有的茶叶嫩芽区域。从图 6(b)中,可以看出 YOLOv8 存在误检测的情况,图像左侧出现非目标区域检测框。在 YOLOv8n-v1、YOLOv8n-v2、YOLOv8n-v3 中减少了该区域的错误识别情况。这表明本文提出的 YOLOv8n-Tea 模型检测表现定位更加精准,减少了误识别茶叶嫩芽目标的情况。

3.4 对比实验

为了进一步验证 YOLOv8n-Tea 模型的检测性能,本文将 YOLOv8n-Tea 模型与其他深度学习模型在相同数据集上进行对比,实验结果如表 3 所示。

表 3 不同网络模型的性能对比

模型	R/%	mAP50/%	Params/M	FPS
Faster RCNN	86.8	87.8	28.28	32
SSD	58.3	80.9	3.54	68
YOLOv4	23.4	58.4	63.94	41
YOLOv5s	74.0	83.3	7.02	52
YOLOv8n-Tea	93.0	97.7	2.57	263

由表 3 中数据可以看出,本文提出的 YOLOv8n-Tea 模型对茶叶嫩芽检测的 mAP50 指标均高于其他模型,这表明 YOLOv8n-Tea 在进行茶叶嫩芽检测任务时泛化性更

强。在模型参数量方面,YOLOv8n-Tea模型的参数量为2.57 M,同样高于其他模型,表明YOLOv8n-Tea模型在进行茶叶嫩芽检测时速度更快。综上所述,本文所提出的YOLOv8n-Tea模型相较于YOLOv8n原模型以及其他目标检测模型在精度上具有较大的提升,更适合于在复杂的茶叶生长环境下对茶叶嫩芽检测,同时提升了茶叶嫩芽的检测速度。

4 结 论

针对茶叶嫩芽检测图像中模型参数大、检测速度慢以及小目标难以特征捕捉的问题,本文基于YOLOv8n模型,提出了一种茶叶嫩芽轻量化检测模型YOLOv8n-Tea。通过在YOLOv8n的骨干网络中引入BSBlock模块和MBConv并结合CBAM注意力机制,提出了MFBNet网络结构,有效增强了模型对密集环境下目标特征提取能力;其次,在颈部网络结构中构建了AVCStem结构,通过结合VoVGSCSPC结构和AKConv模块,使其能够收集到更丰富的语义信息,实现更高效的特征融合,同时加强了对茶叶嫩芽检测区域更多的关注度,解决了模型对小目标茶叶嫩芽的漏检和误检问题。总体上,YOLOv8n-Tea模型的mAP50达到了97.7%,模型参数量为2.57 M,能够满足复杂茶园环境中的茶叶嫩芽检测要求。在后续的工作中,还应该对模型大小方面进行进一步的优化,期望在确保检测精度的前提下实现更加轻量化的网络模型,并将成果在实际茶叶嫩芽检测中进行部署应用。

参考文献

- [1] 龙樟,姜倩,王健,等.茶叶嫩芽视觉识别与采摘点定位方法研究[J].传感器与微系统,2022,41(2):39-41,45.
- [2] ZHANG L, LI M, LI X, et al. Summer pruning improves the branch growth and tea quality of tea trees (*Camellia sinensis*)[J]. *Acta Physiologiae Plantarum*, 2021,43(4):1-12.
- [3] SHAO P, WU M, WANG X, et al. Research on the tea bud recognition based on improved K-means algorithm [C]. *MATEC Web of Conferences*. EDPSciences, 2018, 232:03050.
- [4] QI F, XIE Z, TANG Z, et al. Related study based on Otsu watershed algorithm and new squeeze-and-excitation networks for segmentation and level classification of tea buds [J]. *Neural Processing Letters*, 2021, 53(3):2261-2275.
- [5] CHEN Z, HE L, YE Y, et al. Automatic sorting of fresh tea of fresh tea leaves using vision-based recognition method [J]. *Journal of Food Process Engineering*, 2020, 43(9):e13474.
- [6] ZHANG L, ZOU L, WU C, et al. Method of famous tea sprout identification and segmentation based on improved watershed algorithm [J]. *Computers and Electronics in Agriculture*, 2021, 184:106108.
- [7] ISWANTO B H, ALMA A. Texture histogram features for tea leaf identification using visible digital camera[J]. *IOP Conference Series: Materials Science and Engineering*, 2021, 1098(3):1098-1104.
- [8] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]. 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, OH, USA. IEEE, 2014: 580-587.
- [9] GIRSHICK R. Fast R-CNN [C]. 2015 IEEE International Conference on Computer Vision (ICCV), December 7-13, 2015, Santiago, Chile. IEEE, 2016: 1440-1448.
- [10] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks [C]. 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, USA: IEEE, 2015:91-99.
- [11] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]. IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE, 2016:779-788.
- [12] REDMON J, FARHADI A. YOLO9000: Better, Faster, Stronger[C]. IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2017,440(56):6517-6525.
- [13] REDMON J, FARHADI A. YOLOv3: An incremental improvement[J]. *ArXiv Preprint*, 2018, ArXiv:1804.02767.
- [14] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector [C]. European Conference on Computer Vision. Amsterdam, Holland: ECCV, 2016:21-37.
- [15] CHEN Y T, CHEN S F. Localizing plucking points of tea leaves using deep convolutional neural networks-Science Direct [J]. *Computers and Electronics in Agriculture*, 2020,171:105298.
- [16] 王子钰,赵怡巍,刘振宇.基于SSD算法的茶叶嫩芽检测研究[J].微处理机,2020,41(4):42-48.
- [17] YANG H L, CHEN L, CHEN M T, et al. Tender tea shoots recognition and positioning for picking robot using improved YOLO-V3 model[J]. *IEEE Access*, 2019,7:180998-181011.

- [18] 黄家才, 唐安, 陈光明, 等. 基于 Compact-YOLOv4 算法的茶叶嫩芽移动端识别方案[J]. 农业机械学报, 2023;1-11.
- [19] 王梦妮, 顾寄南, 王化佳, 等. 基于改进 YOLOv5s 模型的茶叶嫩芽识别方法[J]. 农业工程学报, 2023, 39(12):150-157.
- [20] 俞龙, 黄楚斌, 唐劲驰, 等. 基于 YOLOX 改进模型的茶叶嫩芽识别方法[J]. 广东农业科学, 2022, 49(7): 49-56.
- [21] SANDLER M, HOWARD A, ZHU M, et al. MobileNetV2: Inverted residuals and linear bottlenecks [C]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2018.
- [22] 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.
- [23] YANG J, FU X, HU Y, et al. PanNet: A deep network architecture for pan-sharpening [C]. 2017 IEEE International Conference on Computer Vision (ICCV). IEEE, 2017, DOI:10.1109/ICCV.2017.193.
- [24] TAN M, LE Q. Efficientnet: Rethinking model scaling for convolutional neural networks [C]. International Conference on Machine Learning, 2019: 6105-6114.
- [25] 范天浩, 顾寄南, 王文波, 等. 基于改进 YOLOv5s 的轻量化金银花识别方法[J]. 农业工程学报, 2023, 39(11):192-200.
- [26] HU J, SHEN L, SUN G, et al. Squeeze-and-excitation networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017 (99): 7132-7141.
- [27] LI H, LI J, WEI H, et al. Slim-neck by GSConv: A better design paradigm of detector architectures for autonomous vehicles [J]. ArXiv Preprint, 2022, ArXiv: abs/2206.02424.
- [28] 郑美琳, 高建瓴. 融合多注意力机制与 PointRCNN 的三维点云目标检测[J]. 电子测量技术, 2022, 45(9):127-132.

作者简介

潘海鸿, 博士, 教授, 主要研究方向为人工智能、机器人深度学习。

陈希良, 硕士研究生, 主要研究方向为图像处理, 计算机视觉。

钱广坤, 硕士研究生, 主要研究方向为机器学习。

申毅莉, 硕士, 副教授, 主要研究方向为机电结合。

陈琳(通信作者), 博士, 教授, 主要研究方向为智能机器人。

E-mail:gxdxcl@163.com