doi:10.6041/j.issn.1000-1298.2026.13.009

基于 YOLO v5n－DRSW 的机采茶青芽叶形态智能检测研究

郭嘉明[1,2]　王建业[1,2]　夏红玲[3]　郭　鹏[1,2]　丁志武[1,2]　刘妍华[1,2]

（1.华南农业大学工程学院，广州 510642；2.广东省农产品冷链物流工程技术研究中心，广州 510642；
3.广东省农业科学院茶叶研究所/广东省茶树资源创新利用重点实验室，广州 510640）

摘要：精准检测机械采收茶青形态，有利于提高机采茶青自动化分级精度及效率。本研究构建了 YOLO v5n－DRSW 机采茶青形态检测模型，以 YOLO v5n 为基线，通过引入分布移位卷积模块（Distribution shifting convolution，DSConv）至头部网络，在减小了模型复杂度的同时提高了推理速度。此外，在颈部网络引入重参数化泛化特征金字塔网络结构（Reparameterized generalized－FPN，RepGFPN），提升了模型的泛化能力和鲁棒性。针对茶青嫩芽小目标特征，将压缩和激励注意力机制（Squeeze and excitation，SE）嵌入骨干网络，借助全局视野增强对特征图的感知能力，进而提高检测精度。最后，将加权交并比损失函数（Wise intersection over union，WIoU）融入网络的损失计算部分，以动态调整梯度增益。试验结果表明，相较于基线模型，YOLO v5n－DRSW 在检测复杂形态机采茶青方面表现出显著优势：模型精确率提升了 2.1 个百分点，达 98.1%；在推理速度方面，YOLO v5n－DRSW 处理单帧图像仅需 2.11 ms，较基线模型有显著提升；YOLO v5n－DRSW 的浮点运算次数相较于基线模型减少了 2.44%，有效保持了轻量化。实际应用结果表明，机采茶青形态在线检测平均准确率达 94.34%，漏检率不超过 1.1%，检测性能较好。

关键词：机采茶青；卷积神经网络；注意力机制；特征金字塔网络；YOLO v5n；形态检测

中图分类号：S24　　**文献标识码**：A　　**文章编号**：1000-1298(2026)13-0127-13　　**OSID**：

Intelligent Morphological Detection Research of Mechanically Picked Tea Green Buds and Leaves Based on YOLO v5n－DRSW

Guo Jiaming[1,2]　Wang Jianye[1,2]　Xia Hongling[3]　Guo Peng[1,2]　Ding Zhiwu[1,2]　Liu Yanhua[1,2]

（1. *College of Engineering*, *South China Agricultural University*, *Guangzhou* 510642, *China*
2. *Guangdong Engineering Research Center of Agricultural Product Cold Chain Logistic*, *Guangzhou* 510642, *China*
3. *Tea Research Institute*, *Guangdong Academy of Agricultural Sciences/Guangdong Key Laboratory of Innovative Utilization of Tea Tree Resources*, *Guangzhou* 510640, *China*）

Abstract: Accurate morphological detection of mechanically harvested fresh tea leaves is beneficial for improving the accuracy and efficiency of automated grading. YOLO v5n－DRSW, an advanced YOLO v5n－based model specifically designed for the precise recognition of machine-harvested tea leaf morphologies was introduced. The model integrated several innovative features: a distribution shifting convolution (DSConv) module in the head network to reduce complexity and enhance efficiency; a reparameterized generalized feature pyramid network (RepGFPN) in the neck to improve generalization and robustness; and the squeeze-and-excitation (SE) attention mechanism embedded in the backbone to strengthen feature perception for small targets like tender buds. By leveraging a global field of view, this mechanism further enhanced the perception capability of the feature maps. Additionally, the wise intersection over union (WIoU) loss function was used to dynamically adjust gradient contributions during training. Compared with the baseline, YOLO v5n－DRSW exhibited significant advantages in detecting machine-harvested tea leaves with complex morphologies. Experimental results demonstrated that YOLO v5n－DRSW achieved 98.1% accuracy, a 2.1 percentage points improvement over the baseline, with an inference time of just 2.11 ms per frame. This rapid processing speed represented a notable improvement

收稿日期：2026－03－10　修回日期：2026－03－20

基金项目：广东省科技计划项目（2023B0202120001）和高水平农科院建设专项（NYQS202612）

作者简介：郭嘉明（1987—），男，教授，博士，主要从事智能农业装备研究，E-mail：jmguo@scau.edu.cn

通信作者：刘妍华（1979—），女，副教授，博士，主要从事智能农业装备研究，E-mail：cynthial@scau.edu.cn

over the baseline model. The model also reduced floating-point operations by 2.44%, confirming its lightweight nature. In practical applications, it attained an average online recognition accuracy of 94.34% with a miss rate below 1.1%, highlighting its strong potential for enhancing automated tea leaf grading systems. Overall, the model demonstrated excellent detection performance, providing reliable assistance for improving the morphological detection outcomes of fresh tea leaves.

Key words: mechanically picked tea; convolutional neural network; attention mechanism; feature pyramid network; YOLO v5n; form realization

0 引言

当前,我国茶叶产需均居全球首位,2023 年生产总量达 354.1 万吨,同比增长 5.95%[1],需求量呈逐年上升趋势。在此背景下,各产茶地开始新兴机械采收模式。传统的人工采收可以保证优质茶青品质,但效率低下且成本高昂[2];机械采收虽能弥补不足,但易产生枝梗和断碎叶,后续不加以分选会影响成茶品质[3]。一芽一叶和一芽二叶茶青是炒制名优茶的核心原料,而一芽三叶、枝梗、断碎叶等则适用于制作速溶茶,或用于提取茶多酚、咖啡碱等功能性成分。华南农业大学联合广东省农业科学院茶叶研究所研发了茶青智能分级装置,通过滚筒筛分实现茶青的智能分级,但在应用过程中发现茶青形态对分级效果有重要影响,通过精准获取茶青形态占比,可为滚筒转速等参数的调控提供依据,提升筛分精度与效率。

目前,常见的茶叶分选机(如美亚光电等)主要采用基于光电色选技术的神经网络算法对各类成品茶进行分选[4],但针对茶鲜叶在形态特征层面的分选研究仍相对匮乏。为提高模型的泛化能力,Xue 等[5]将 ACmix 和 CBAM 集成至 YOLO v5 中,解决了茶芽易受遮挡等因素干扰的问题。Shi 等[6]通过引入 WIoU 和 Swin Transformer 模块,增强了模型的鲁棒性及检测准确率。为保证模型轻量化特性,Lu 等[7]采用 DyHead 动态头,设计轻量级 C3ghost 模块,与 YOLO v5s 相比,尺寸和参数量分别减少了 31.41% 和 32.21%。黄铝文等[8]基于二维离散小波变换 DWT 和 MobileNetV3 进行融合,准确率提升 2.15 个百分点,参数量降低 25.12%。在对茶青检测速度有较高要求的应用场景中,Zhang 等[9]将原模型的特征提取网络替换为 MobileNetV3,模型尺寸降低至 11.78 M,检测速度提升了 11.68 f/s。

尽管上述研究在茶青形态检测领域已取得了一定成果,但仍存在以下问题:①当前,茶青形态检测的研究多集中于采收前阶段,而对于自动化分选过程中的茶青形态检测研究相对较少。② YOLO 系列算法在精度、模型尺寸与推理速度之间难以兼顾,通常高精度会伴随着较大的模型尺寸和较低的推理速度,故需要在三者之间寻求平衡点。③由于茶芽较小且易被遮挡的特性,对算法检测准确性提出了更高要求。

针对以上问题,本研究提出一种基于 YOLO v5n - DRSW 的机器视觉算法模型,以实现机采茶青形态精准检测,并将其应用于自主研发的滚筒-风筛式茶青智能分级装置中,优化机采茶青筛分精度和效率。

1 数据处理

1.1 数据采集

茶青图像数据采自广东省农业科学院茶叶研究所(英德实验基地),采收对象为英红九号茶树品种,采用浙江川崎茶业机械有限公司生产的川崎 NV60H 型采茶机,采收时间为当日 08:00—10:00,采后立即进行试验,避免因陈放时间过久造成萎凋。图 1 展示了采茶机的工作实况。



图 1 采茶机工作实况

Fig. 1 Tea harvester operating conditions

图像采集系统采用海康威视 MV - CE200 - 10UC 型工业相机,分辨率 5 472 像素 ×3 648 像素,能够清晰捕捉茶青样本的形态特征;相机镜头垂直安装于载物台上方,确保拍摄视角稳定;载物台采用推拉式结构设计,便于快速更换不同批次的茶青样本,提高了试验效率和操作便捷性;相机拍摄、参数设置以及存储等功能由计算机控制,实现图像采集过程的自动化与标准化。图 2 为图像采集系统整体结构示意图。

为保证图像采集的规范性和可重复性,将茶青分为上下两排,每排 5 株,按 2 cm 间距相隔。这种布置方式旨在尽可能还原机械采摘后茶青的实际分布状态,同时避免因人为过度干预导致的样本偏差。图像采集完成后进行初步筛选,除去因操作不当造

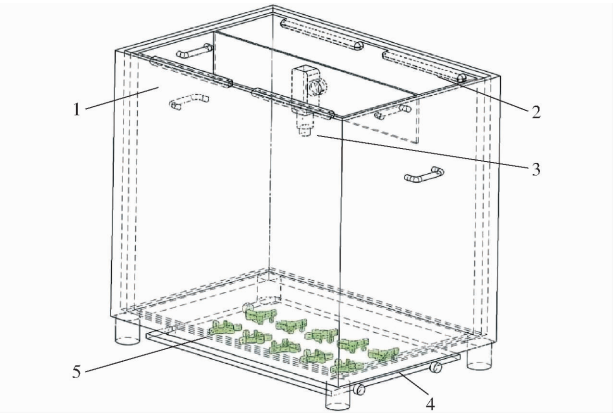

图 2　图像采集系统整体结构示意图

Fig. 2　Overall structure diagram of image acquisition system

1. 箱体　2. 灯带　3. 摄像头　4. 载物台　5. 机采茶青

成的对焦模糊、反光等问题的图像，最终选取 590 幅机采茶青图像，以此构建试验所用数据集，样本示例见图 3。



图 3　机采茶青图像样本示例

Fig. 3　Example of machine-harvested tea shoots images

1.2　茶青形态分类

对于机采茶青，嫩芽与叶片数量的比例是评价茶青形态的关键指标。本研究中，茶青形态特指其外观结构特征。根据相关标准，名优茶原料以芽头、一芽一叶、一芽二叶初展为主[10]，而一芽三叶及以上、单片、破碎芽叶、老叶梗等则适制大宗茶[11]。基于上述标准，将茶青划分为一芽一叶、一芽二叶、一芽三叶、枝梗和断碎叶 5 种形态，并以此作为数据集中茶青图像的标定准则。茶青形态分类如图 4 所示。

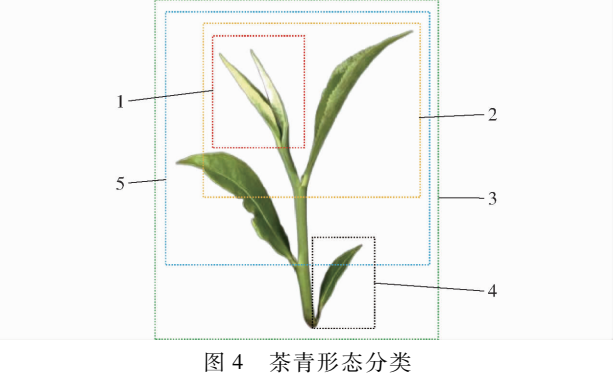

图 4　茶青形态分类

Fig. 4　Classification for form of machine-harvested tea

1. 一芽一叶　2. 一芽二叶　3. 枝梗　4. 断碎叶　5. 一芽三叶

本研究通过对大量机采茶青样本进行实地采样与物理参数测量，发现不同形态茶青在叶长、叶展及单叶质量等关键形态指标上存在显著差异。基于上述差异特征，构建了形态分类依据，具体特征参数见表 1。

表 1　机采茶青特征参数

Tab. 1　Characteristic parameters of machine-harvested tea

茶青种类	特征参数		
	叶长/mm	叶展/mm	质量/g
一芽一叶	30 ~ 60	10 ~ 20	0. 10 ~ 0. 24
一芽二叶	42 ~ 70	35 ~ 66	0. 28 ~ 0. 72
一芽三叶	50 ~ 95	56 ~ 99	0. 45 ~ 0. 96
枝梗	82 ~ 145	66 ~ 119	0. 79 ~ 1. 80
断碎叶	41 ~ 97	25 ~ 45	0. 16 ~ 0. 73

1.3　数据增强

鉴于机采茶青图像样本数量有限且不同形态的茶青数量分布不均衡，为防止模型出现过拟合或欠拟合现象，同时提升模型的泛化能力和鲁棒性，本研究对数据集进行增强处理[12]，具体处理方式包括镜像、增强对比度、高光、添加高斯噪声以及灰度化等，如图 5 所示。通过上述增强方法，最终扩充得到 850 幅茶青样本图像。

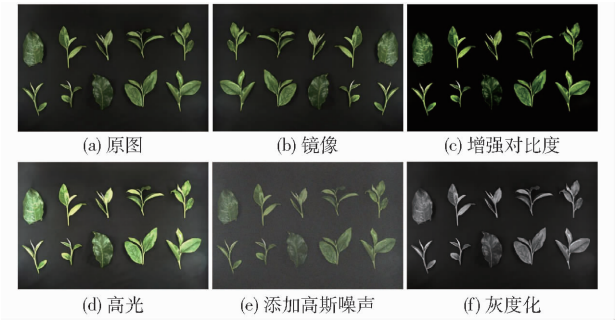

图 5　数据增强示例

Fig. 5　Example of harvested tea date enhancement

1.4　数据集制作

完成数据增强后，通过 Make Sense 网站对样本进行目标区域标注，手动绘制边界框并标注其类别。表 2 为数据集分布情况。

表 2　数据集的标注名称及数量分布

Tab. 2　Labeled names and quantity distribution in data set

茶青种类	茶青数据分布	
	训练集数量	测试集数量
一芽一叶(1Y1Y)	1 100	320
一芽二叶(1Y2Y)	1 200	345
一芽三叶(1Y3Y)	1 250	350
枝梗(ZY)	950	305
断碎叶(DY)	1 450	380

2　试验方法

2.1　YOLO v5 目标检测算法

YOLO v5 是由 Ultralytics 公司推出的一种单阶

段目标检测算法[13]，通过单个神经网络实现端到端的训练，可以自动提取特征，将简单的特征融合为更复杂的特征，通过复杂特征解决检测定位问题[14]，从而完成快速准确的目标检测，其网络结构主要由输入端 Input、主干网络 Backbone、颈部网络 Neck 和头部网络 Head 4 部分组成[15]。鉴于传送带运行速度快且模型需部署于树莓派等低算力设备，本研究以 YOLO v5 为基础，通过结构优化，在保证茶青形态检测精度的同时兼顾推理速度与模型轻量化，实现三者的有效平衡。

2.2 试验环境

本研究使用的计算机 CPU 型号为 Intel Core i7 - 14700HX，GPU 型号为 NVIDIA GeForce RTX 4060，运行内存为 32 GB，操作系统为 Windows 11，编程语言为 Python，深度学习框架为 PyTorch，所有模型的训练及调试均在 PyCharm 集成开发环境中完成。

2.3 评价指标

以精确率(P)、召回率(R)[16]、参数量[17]、浮点运算次数、平均精度均值(mAP)和每秒处理帧数(FPS)[18]等作为评价指标，对各模型进行综合评估。

2.4 算法选择

根据网络深度和宽度的不同，YOLO v5 被划分为 YOLO v5s、YOLO v5n、YOLO v5m、YOLO v5l 和 YOLO v5x 5 种版本[19]，将数据集分别在上述算法中运行 200 轮次，其性能表现如表 3 所示。

表 3 YOLO v5 各个算法的性能

Tab. 3 Performance evaluation of YOLO v5 series models

YOLO v5 算法	P/%	R/%	参数量	浮点运算次数
YOLO v5s	95.6	99.0	7.026×10^6	1.58×10^{10}
YOLO v5n	96.0	98.9	1.766×10^6	4.10×10^9
YOLO v5m	95.9	99.5	2.087×10^7	4.79×10^{10}
YOLO v5l	96.2	99.8	4.613×10^7	1.08×10^{11}
YOLO v5x	95.7	99.7	8.620×10^7	2.04×10^{11}

根据表 3 中的数据可知，不同 YOLO v5 算法在精确率和召回率两项指标上的差异均不超过 1%，表现较为接近；而 YOLO v5n 的模型参数量显著低于其他 4 种算法，模型参数量及计算复杂度较低。

2.5 算法改进

由于算法需在有限的存储资源与实时性约束下高效运行，故本研究以 YOLO v5n 作为基础模型，进行以下改进：①在模型的头部网络引入 DSConv 卷积模块，分别替换第 17、20、23 层的 C3 模块作为 Detect 输入。②在颈部网络引入 RepGFPN 重参数化泛化特征金字塔网络结构，将残差结构应用于 10～13 层的 FPN 网络。③在骨干网络尾部添加 SE 注意力机制。④在损失计算部分引入 WIoU，代替原有的 CIoU 损失函数。YOLO v5n - DRSW 的网络模型结构如图 6 所示。

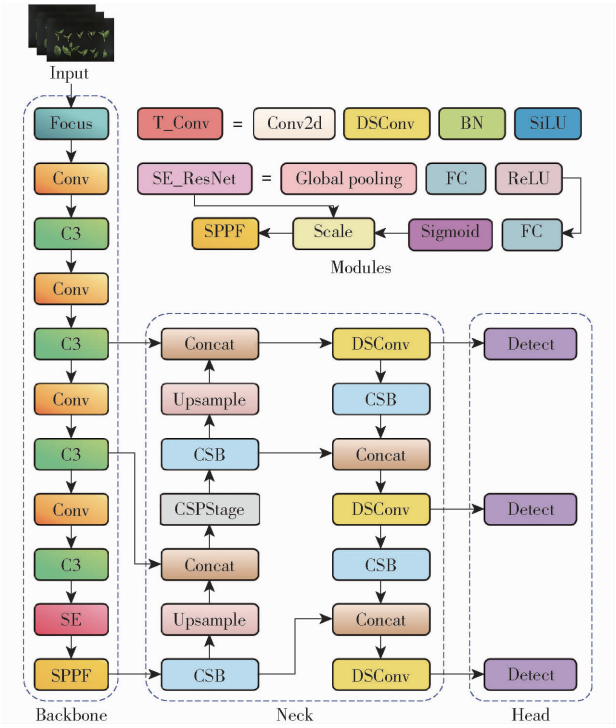

图 6 YOLO v5n - DRSW 网络结构图

Fig. 6 Network structure of YOLO v5n - DRSW

2.5.1 特征提取网络优化

卷积模块主要负责提取图像特征、减少信息丢失、加速模型收敛、引入非线性变换以及适应不同尺度的特征图等，从而增强模型对复杂茶青形态的多尺度检测能力[20]。在 YOLO v5n 的头部网络部分，C3 模块中的 Bottleneck 结构网络层数较多，计算量大且推理速度缓慢[21]，难以满足轻量化及高效实时检测茶青形态的需求，因此引入 DSConv 代替 C3 模块。

DSConv 是传统卷积层的一种变体，通过将卷积核的权重张量分解为 VQK 和 KDS，在 VQK 中仅存储整数值，从而有效降低计算量和内存占用[22]。此外，DSConv 利用低位定点计算显著提升了网络的推理速度。同时，其采用基于内核和基于通道的分布偏移策略，仅导致极小的准确率损失，在资源受限的条件下仍能保持高效运行[23]。DSConv 结构如图 7 所示。其中，KDS 为内核分布偏移，k 为内核的宽度和高度；Expand 表示将原始的 $k\times k$ 核进行扩展，在通道维度上扩展了 BLK 次；VQK 为可变量化内核，它与扩展后的 KDS 在对应位置逐元素相乘，生成重构核 RES；ch_i 是当前层的通道数；Original 为原始卷积核，由 RES 组合出最终的卷积操作逼近原始的卷

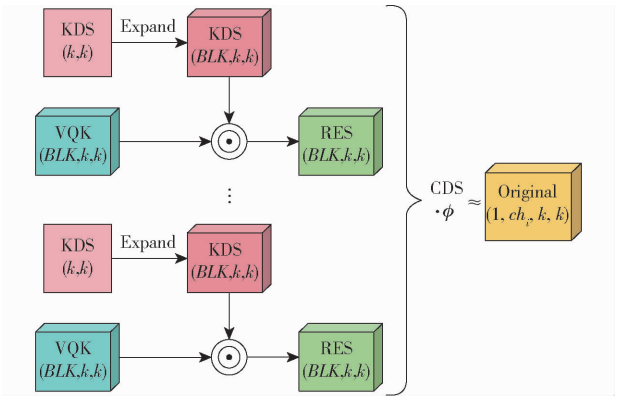

图 7　DSConv 结构图

Fig. 7　Structure of DSConv

积核；ϕ 为通道分布偏移的缩放参数，组合后的重构核通过乘以该参数完成通道级别的分布偏移标准，最终重构出与原始卷积核一致的特征表达。

在 YOLO v5n 中，将 DSConv 卷积模块引入 Head 部分，分别替换第 17、20、23 层的 C3 模块作为 Detect 输入，改进后的模型命名为 YOLO v5n－D。

2.5.2　特征融合网络优化

特征金字塔网络（FPN）通过自上向下的路径和横向连接，能够同时利用高层级的语义信息和低层级的细节信息，实现不同层级特征的融合[24]。然而，在 YOLO v5n 的特征融合过程中，部分特征信息被重复使用，导致参数冗余，增加了非最大值抑制（NMS）的复杂性[25]。为解决以上问题，本研究采用 RepGFPN（重参数化泛化特征金字塔网络）代替 FPN。

RepGFPN 是一种改进的特征融合网络，通过 CSPStage（高效的特征提取结构）和 RepConv（重参数化的卷积结构）模块实现更高效的特征融合与信息传递[26]，该结构能够增强模型的特征提取能力、提高检测准确率，同时降低因上采样导致的特征模糊和冗余。RepGFPN 结构如图 8 所示。其中，不同不同颜色的特征图分别代表浅层（绿色）、中层（蓝色）和深层（红色）语义信息，通过多级 CSPStage 实现跨层特征融合，支持邻近层、斜上层和斜下层的信息传递。其中，浅层语义通道数为 128，空间分辨率为 $H/8 \times W/8$；中层语义通道数为 256，空间分辨率为 $H/16 \times W/16$；深层语义通道数为 512，空间分辨率为 $H/32 \times W/32$。

在 YOLO v5n－D 中，将 RepGFPN 结构引入网络的 Neck 部分，使用残差结构替换 FPN 网络，改进后的模型命名为 YOLO v5n－DR，相关参数计算如下：

CSPStage 通过 1×1 卷积、3×3 卷积和空间金字塔池化（SPP）处理以减少通道数，具体公式为

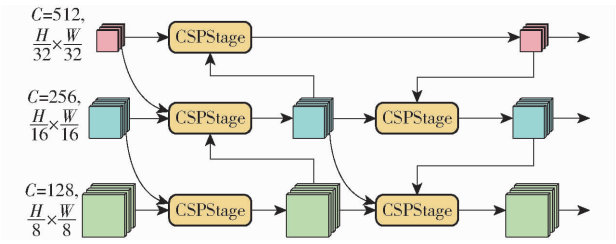

图 8　RepGFPN 结构图

Fig. 8　Structure of RepGFPN

$$y_{csp} = y_1 + y_2 \tag{1}$$

其中　$y_1 = \mathrm{ConvBNAct}(x_1)$

$y_2 = \mathrm{ConvBNAct}(\mathrm{RepConv}(x_2))$

式中　x_1、x_2——输入特征图

y_1——通过 1×1 卷积后输出的特征图

y_2——通过 3×3 卷积和空间金字塔池化（SPP）处理后输出的特征图

y_{csp}——路径特征合并后输出的特征图

ConvBNAct——卷积、批归一化和激活函数

RepConv——重参数化的卷积

RepConv 将 3×3 卷积和 1×1 卷积组合重参数化为一个单独的 3×3 卷积，以减少模型参数和计算量，即

$$\mathrm{RepConv}(x) = \mathrm{Conv}(x) + \mathrm{Conv}(\mathrm{Pad}(\mathrm{Conv}(x))) \tag{2}$$

式中　$\mathrm{Conv}(x)$——对输入特征图 x 应用 3×3 卷积操作

Pad——对 3×3 卷积的输出进行填充

RepGFPN 通过 CSPStage 和 RepConv 模块实现不同尺度特征图的融合。在融合过程中，高层特征图经上采样后与低层特征图合并，然后通过 CSPStage 进行处理，从而实现特征的融合与提升。具体公式为

$$y_{fuse} = \mathrm{CSPStage}(\mathrm{Upsample}(x_{high}) + x_{low}) \tag{3}$$

式中　y_{fuse}——特征融合后输出的特征图

Upsample——上采样操作

x_{high}——高层特征图

x_{low}——低层特征图

嵌入 RepGFPN 特征金字塔网络模块能够提升模型的特征融合和信息传递能力，然而，随着其在颈部网络中嵌入层数的增加，尽管平均准确率有所提升，但模型参数量也随之增长[27]，可能会影响茶青形态检测的推理速度。

为探究 RepGFPN 的最佳嵌入深度，在 YOLO v5n－D 的基础上设计消融试验，分别将 RepGFPN 模块集成至 10～13 层、10～17 层、10～20 层以及 10～23 层，旨在探索模型检测精度与推理速度之间的平衡点。

2.5.3 残差神经网络优化

在 YOLO 目标检测算法中,注意力机制通过模拟人类视觉系统的选择性关注机制,使模型能够聚焦于与当前任务最相关的信息,同时忽略不重要的背景或干扰信息[28]。鉴于本文研究对象为机采茶青,茶芽作为关键特征,具有体积小且易被叶片遮挡的特点,故引入 SE 注意力机制,以增强模型对茶芽的捕捉能力。

SE 注意力机制是一种在通道维度上引入注意力机制的方法,其核心在于 Squeeze(压缩)和 Excitation(激励),通过全局平均池化聚合各通道的空间特征,形成全局特征向量[29]。该向量经过两个全连接层:第 1 层进行降维并激活,第 2 层恢复维度并使用 Sigmoid 函数输出通道权重。这些权重用于指示各通道的重要性,并据此调整原始特征图,实现特征重标定。标准残差结构与 SE 增强残差结构对比如图 9 所示。

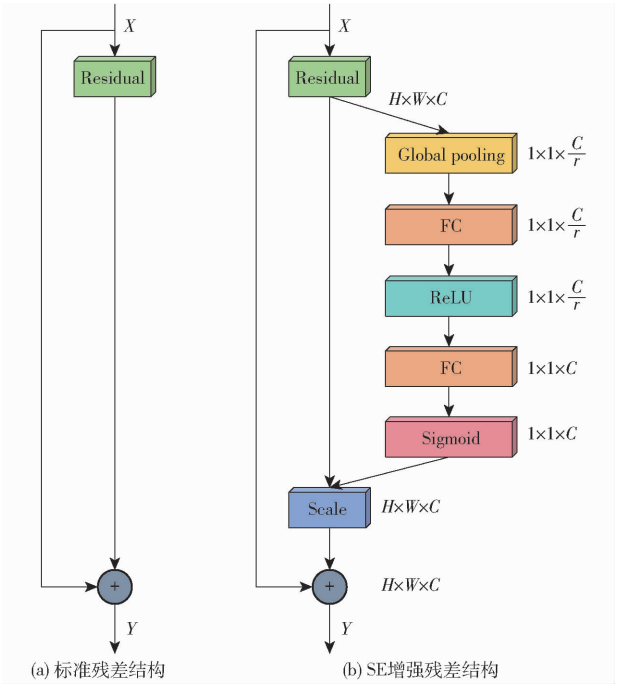

图 9 标准残差结构与 SE 增强残差结构对比

Fig. 9 Comparison between standard residual structure and SE enhanced residual structure

其中,X 为输入特征图,其高度、宽度和通道数分别为 H、W、C;Residual 为残差块;Global pooling 为全局平均池化;FC 为全连接层;ReLU 为矩形激活函数;Sigmoid 输出[0,1]范围的激活函数,得到 $1\times1\times C$ 的通道权重;Scale 使用 SE 权重对特征图进行通道级加权;r 为缩减比,控制 SE 模块的复杂度;+ 为残差连接,将各元素相加;Y 为输出带残差连接后的特征图。

在 YOLO v5n - DR 中,将 SE 引入骨干网络尾部,改进模型命名为 YOLO v5n - DRS,相关参数计算为:

使用全局平均池化(Global average pooling)将输入特征图 x 压缩成一个通道向量的形式 $\boldsymbol{x}_s$,计算公式为

$$\boldsymbol{x}_s = \mathrm{ReLU}(\boldsymbol{W}_s\boldsymbol{x} + \boldsymbol{b}_s) \tag{4}$$

式中 $\boldsymbol{W}_s$——权重 $\boldsymbol{b}_s$——偏置项

接收 Squeeze 层输出的压缩特征 $\boldsymbol{x}_s$,并通过非线性函数 g(此处为 ReLU)计算出权重矩阵,称为兴奋度。然后将兴奋度与压缩特征相乘,得到最终的输出特征 $\boldsymbol{x}_e$。具体公式为

$$\boldsymbol{x}_e = \sigma(g(\boldsymbol{x}_s, \boldsymbol{W})) = \sigma(\boldsymbol{W}_2\delta(\boldsymbol{W}_1\boldsymbol{x}_s)) \tag{5}$$

式中 δ——ReLU 激活函数

σ——Sigmoid 激活函数

$\boldsymbol{W}$——权重参数集合

$\boldsymbol{W}_1$、$\boldsymbol{W}_2$——全连接层的权重

最终,SE 的输出作用于卷积层的输出,得到通道注意力加权的输出特征图。使用得到的 channel-wise 向量,对每个通道特征图的每个元素做加权,即

$$\mathrm{Output} = \boldsymbol{x}\boldsymbol{x}_e \tag{6}$$

2.5.4 训练策略优化

损失函数(Loss function)是深度学习中用来衡量模型预测值与实际值之间差异的函数[30]。其核心作用是为模型训练提供一个可量化的误差指标,指导优化算法调整模型参数,以最小化预测误差。通过不断降低损失值,模型能够逐步提升预测准确性,并在训练数据之外具备更好的泛化能力。基线模型采用 CIoU(Complete intersection over union)损失函数,但处理类别不平衡的目标检测任务时效果欠佳[31]。因此,考虑引入其他类型的 IoU 代替 CIoU,以提高模型对于茶青形态检测的准确性以及泛化能力。

WIoU 是一种基于动态非单调聚焦机制的边界框回归损失函数,与传统 IoU 类损失不同,WIoU 能够根据预测框与真实框之间的重叠质量,自适应地调整对不同样本的关注程度,从而增强模型对不同质量数据集的泛化能力[32]。此外,WIoU 通过计算锚框的离群度动态调整梯度增益[33],进而提高整体检测性能。图 10 为 IoU 预测效果示意图。

本研究将 WIoU 引入损失计算部分,替换原 CIoU 损失函数,改进交并比(IoU)的计算方法,以更准确地评估预测边界框与真实边界框之间的相似度。

WIoU 共有 3 个版本,分别为 v1、v2 和 v3。本研究在 YOLO v5n - DRS 的基础上设计消融试验,分别

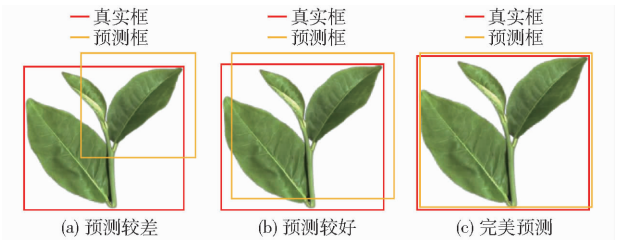

图 10　IoU 预测效果对比图

Fig. 10　Schematics of IoU prediction effect

将这 3 个版本的模块引入损失计算部分，以系统地评估每个版本对模型性能的影响。

3　结果与分析

3.1　模型整体性能分析

本文提出的 YOLO v5n－DRSW 模型在茶青形态检测任务中展现出优异的综合性能，各模型的性能指标变化趋势如图 11 所示。在测试集上，YOLO v5n－DRSW 的分类损失与定位损失分别收敛于 0.011 和 0.019，且在整个训练过程中均呈现稳定下降趋势，最终收敛值显著低于基线模型 YOLO v5n，表明所提出的改进策略有效提升了模型在类别判别与边界框定位两个关键任务上的学习能力与优化效率，从而在整体检测性能上展现出显著优势。就迭代收敛速度而言，YOLO v8n 表现出最快的收敛速度，而 YOLO v3－Tiny 收敛最为缓慢，需更多轮次才能达到相对平稳状态。其余模型收敛速度相近，7 种模型均在第 170 轮次左右训练后趋于稳定。此外，在检测精度方面，YOLO v5n－DRSW 的表现最为突出，其在测试集上的 mAP@0.5 达到 98.3%，mAP@0.5:0.95 达到 75.5%，两项指标均优于其他轻量化对比模型，充分验证了该改进方法在保持模型轻量化的同时有效提升了目标检测精度和鲁棒性。

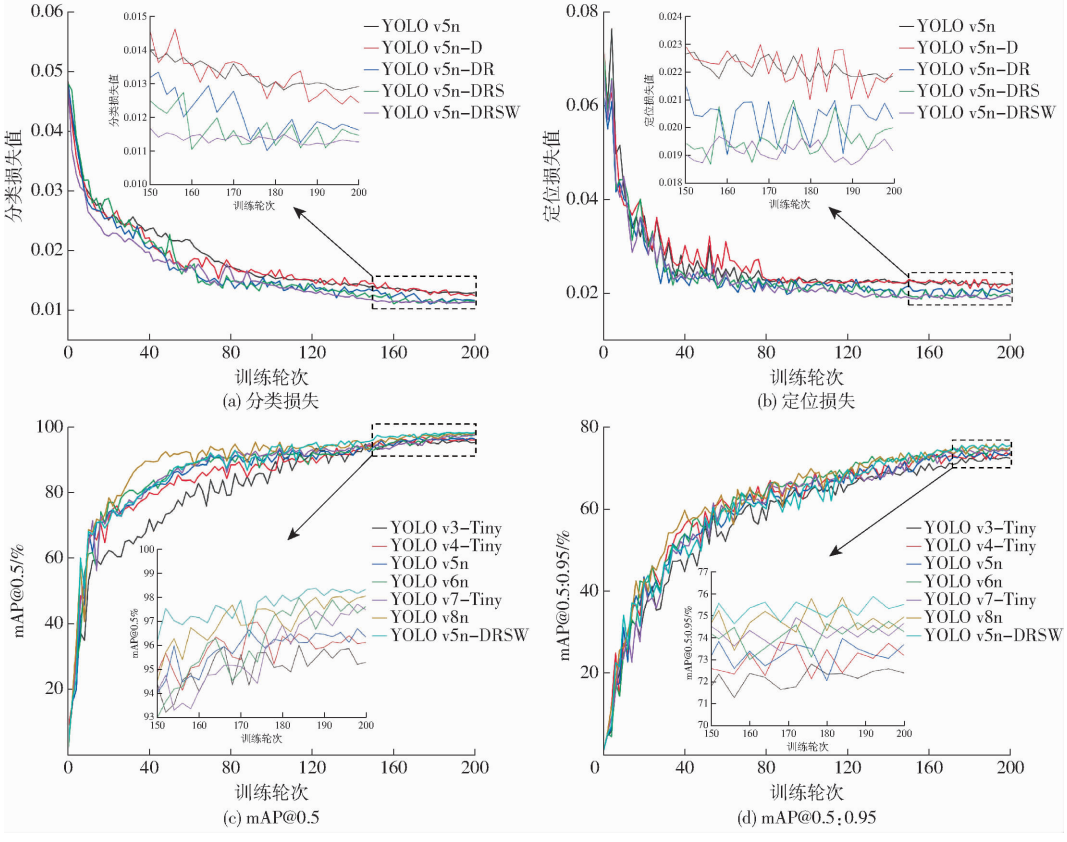

图 11　性能指标变化曲线

Fig. 11　Performance index change curves

3.2　消融试验

为深入探究在基线模型不同位置部署各类模块的性能差异，以设计理念的差异为分类标准，系统地引入了多种类型的轻量化模块，并开展消融试验以评估其对整体性能的影响，如表 4 所示。

试验结果表明，将 DSConv 引入头部网络并替换 C3 模块后，相较于基线模型 YOLO v5n，模型参数量减少了 14.55%，浮点运算次数降低了 7.32%。在显著压缩模型尺寸的同时，检测精确率 P 仅下降了 0.31%，召回率 R、mAP@0.5 与 mAP@0.5:0.95 基本和基线模型持平，而推理速度提升了 50.16%。该改进在几乎不牺牲检测精度的前提下，有效兼顾了模型的轻量化与实时性需求。

为探究 RepGFPN 在颈部网络中的最佳嵌入深

表 4 消融试验结果

Tab. 4 Ablation test results

模型	P/%	R/%	mAP@0.5/%	mAP@0.5:0.95/%	参数量	浮点运算次数	推理速度/(f·s^{-1})
YOLO v5n	96.0	98.9	96.3	73.7	1.766×10^{6}	4.1×10^{9}	317
YOLO v5n-D	95.7	99.0	96.2	73.5	1.509×10^{6}	3.8×10^{9}	476
YOLO v5n-D+RepGFPN 10~13 层	96.6	99.2	97.8	74.9	1.632×10^{6}	4.0×10^{9}	436
YOLO v5n-D+RepGFPN 10~17 层	96.3	99.6	97.4	74.6	1.726×10^{6}	5.3×10^{9}	370
YOLO v5n-D+RepGFPN 10~20 层	96.6	99.3	97.9	75.0	2.054×10^{6}	5.7×10^{9}	318
YOLO v5n-D+RepGFPN 10~23 层	96.9	99.2	98.1	75.2	2.107×10^{6}	5.9×10^{9}	307
YOLO v5n-DR+SE_end	97.8	99.2	97.9	75.1	1.720×10^{6}	4.0×10^{9}	399
YOLO v5n-DR+SE_layer	97.6	99.6	97.7	74.9	2.241×10^{6}	4.4×10^{9}	387
YOLO v5n-DRS+WIoU-v1	98.1	99.7	98.3	75.5	1.759×10^{6}	4.0×10^{9}	475
YOLO v5n-DRS+WIoU-v2	98.0	99.3	98.1	75.2	1.759×10^{6}	4.0×10^{9}	463
YOLO v5n-DRS+WIoU-v3	98.1	99.5	98.1	75.3	1.759×10^{6}	4.0×10^{9}	471

度,在 YOLO v5-D 的基础上,分别将 RepGFPN 模块集成至 10~13 层、10~17 层、10~20 层以及 10~23 层进行试验。结果表明,随着嵌入深度的增加,P、R、mAP@0.5 以及 mAP@0.5:0.95 整体呈上升趋势,但不同试验组间的提升效果差异不大,相较于基线模型 YOLO v5n-D,P 提升差异不超过 1.2%,R 不超过 0.6%,mAP@0.5 不超过 1.9%,mAP@0.5:0.95 不超过 1.7%,且模型尺寸不断增大、推理速度持续下降。当嵌入深度为 10~13 层时,推理速度下降了 8.4%;而当嵌入深度为 10~23 层时,推理速度下降高达 35.5%。综合考虑检测精度与推理效率的平衡,将 RepGFPN 模块集成至颈部网络的 10~13 层。

引入通道注意力模块 SE 主要有两种方式:在骨干网络尾部插入;在每层 C3 模块之后嵌入。因此,通过设计消融试验,在 YOLO v5n-DR 不同位置分别引入 SE 模块进行对照。结果表明,将 SE 部署在 Backbone 尾部时,在 P 和推理速度指标上表现最优:P 提升 1.2%,R 与 YOLO v5n-DR 基本持平,mAP@0.5 和 mAP@0.5:0.95 略有提升,同时参数量仅增加 5.39%,有效兼顾了性能提升与轻量化设计。

WIoU 共有 3 个版本,分别为 v1、v2 和 v3。为系统评估各版本对模型性能的影响,在 YOLO v5n-DRS 的基础上设计了消融试验,分别将这 3 种 WIoU 版本引入损失函数中进行对比。试验结果表明,WIoU-v1 在提升模型精度方面表现最优,且模型参数量与 v2、v3 版本相当。相较于 YOLO v5n-DRS,引入 WIoU-v1 后推理速度提升了 19.05%。综合考虑模型大小、检测精度与推理效率等因素,选择将 WIoU-v1 集成至模型中。通过上述模型改进方法,本文构建了机采茶青形态检测模型 YOLO v5n-DRSW,添加不同改进模块的检测效果对比如表 5 所示。

表 5 添加不同改进模块的检测效果对比

Tab. 5 Comparison of detection effects of different improved modules

模块种类	P/%	R/%	mAP@0.5/%	mAP@0.5:0.95/%	参数量	浮点运算次数	推理速度/(f·s^{-1})
—	96.0	98.9	96.3	73.7	1.766×10^{6}	4.1×10^{9}	317
DSConv	95.7	99.0	96.2	73.5	1.509×10^{6}	3.8×10^{9}	476
DSConv+RepGFPN	96.6	99.2	97.8	74.9	1.632×10^{6}	4.0×10^{9}	436
DSConv+RepGFPN+SE	97.8	99.2	97.9	75.1	1.720×10^{6}	4.0×10^{9}	399
DSConv+RepGFPN+SE+WIoU	98.1	99.7	98.3	75.5	1.759×10^{6}	4.0×10^{9}	475

注:“—”表示未添加改进模块。

相较于基线模型,本文提出的 YOLO v5n-DRSW 在精确率和召回率上分别提升 2.1 个百分点和 0.8 个百分点,同时浮点运算次数减少 2.44%,参数量降低至 1.759×10^{6}。在推理效率方面,其推理速度由 317 f/s 提升至 475 f/s,单帧图像处理时间由 3.15 ms 缩短至 2.11 ms,推理速度显著提高。

上述结果表明,YOLO v5n-DRSW 在降低模型计算复杂度和参数规模的同时,实现了检测精度与推理效率的双重提升,验证了该模型在机采茶青实时、高精度检测任务中的可行性与应用潜力。

为强化模块有效性论证，本研究采用梯度加权类激活映射（Grad－CAM）方法，生成集成不同模块的模型在实际应用场景下的检测热力图。Grad－CAM 通过利用目标类别对卷积层特征图的梯度信息，生成对应的类激活热力图，从而可视化模型在做出预测时所关注的图像区域。由图 12a 可知，尽管集成 DSConv 能够提取茶青芽叶的特征，但表现出了明显的疏漏，其特征提取能力较弱；图 12b 显示，集成了 RepGFPN 后，热力图表现出了更强的多尺度响应，但易受背景噪声干扰；图 12c 为集成了 SE 模块的热力图，热力区域变得更加聚焦和精准；图 12d 表明，添加了 WIoU 模块后，特征提取能力显著增强，且轮廓边界更加清晰，有效提高了茶青芽叶检测的准确性。

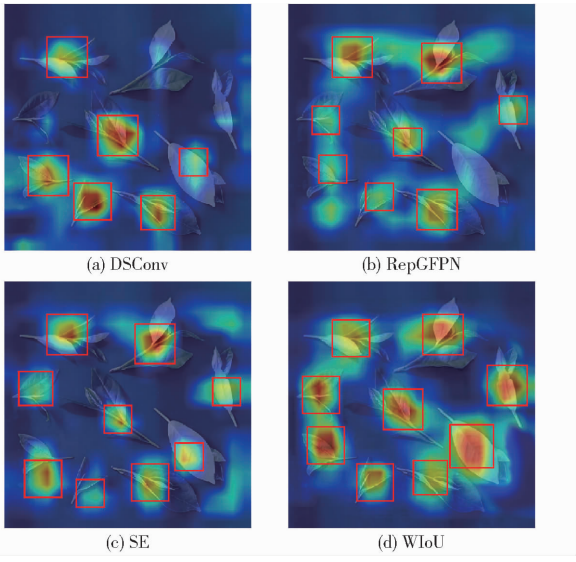

图 12　集成不同模块的检测热力图

Fig. 12　Integrated detection heat maps of different modules

3.3　主流轻量化模型对比试验

为进一步验证 YOLO v5n－DRSW 在茶青实时检测任务中的优越性，本研究选取了模型参数量与计算量相近的 YOLO v3－Tiny、YOLO v4－Tiny、YOLO v5n、YOLO v6n、YOLO v7－Tiny 和 YOLO v8n 共 6 种主流轻量化检测算法开展对比试验，试验结果如表 6 所示。

表 6　YOLO v5n－DRSW 与目前主流轻量化模型的茶青检测效果对比

Tab. 6　Comparison of tea leaf recognition effect between YOLO v5n－DRSW and current mainstream lightweight models

模型	F1 分数/%	mAP@0.5/%	mAP@0.5:0.95/%	参数量	单帧图像处理时间/ms
YOLO v3－Tiny	67.7	95.8	72.4	3.175×10^6	5.35
YOLO v4－Tiny	81.3	96.1	73.2	5.964×10^6	3.24
YOLO v5n	97.4	96.3	73.7	1.766×10^6	3.15
YOLO v6n	86.0	97.6	74.3	4.982×10^6	3.40
YOLO v7－Tiny	95.6	97.4	74.7	8.462×10^6	2.87
YOLO v8n	97.5	98.0	74.9	2.256×10^6	2.49
YOLO v5n－DRSW	98.9	98.3	75.5	1.759×10^6	2.11

由表 6 可知，YOLO v5n－DRSW 在各项关键性能指标上均优于对比模型。在 F1 分数方面，该模型达到 98.9%，较基线模型 YOLO v5n 提升 1.5 个百分点，较 YOLO v8n 高出 1.4 个百分点，表明其在精确率与召回率之间取得了更优的平衡；在目标检测精度方面，模型的 mAP@0.5 为 98.3%，mAP@0.5:0.95 达到 75.5%，获得所有对比模型中最高的平均检测精度，说明该模型在不同 IoU 阈值下均具备更强的鲁棒性和定位能力。在模型尺寸方面，该模型的参数量仅 1.759×10^6，保持了良好的轻量化特性；在推理速度方面，该模型的单帧处理时间仅 2.11 ms，快于 YOLO v5n、YOLO v6n 和 YOLO v3－Tiny，展现出优异的实时检测能力。综上，YOLO v5n－DRSW 在保持轻量化的同时，实现了更高的检测精度与更快的推理速度，优于现有主流轻量化模型，充分验证了其在茶青形态实时检测任务中的实用性。

图 13 为各轻量化模型在测试集上的混淆矩阵。该矩阵以归一化的平均精度 AP 呈现，对角线元素反映各类别的正确检测能力，非对角线元素则表示类别间的误判程度。试验结果表明，早期轻量化模型（YOLO v3－Tiny 和 YOLO v4－Tiny）在部分相似类别之间（如一芽一叶和一芽二叶、一芽三叶和枝梗）存在明显混淆，对角线 AP 值普遍低于 90%；随着 YOLO 架构的演进，YOLO v5n 至 v8n 模型的检测性能显著提升，各类别对角线 AP 值基本超过 95%。其中，YOLO v5n－DRSW 在所有类别上均取得最优表现，各类别对角线 AP 值稳定在 97% 及以上，且非对角线误判值显著低于其他模型，充分验证了其在增强特征判别性与抑制类别混淆方面的有效性。

3.4　机采茶青形态在线检测及分选试验

为更全面地验证 YOLO v5n－DRSW 的实际应用效果，本研究依托智能茶青分级装置开展机采茶

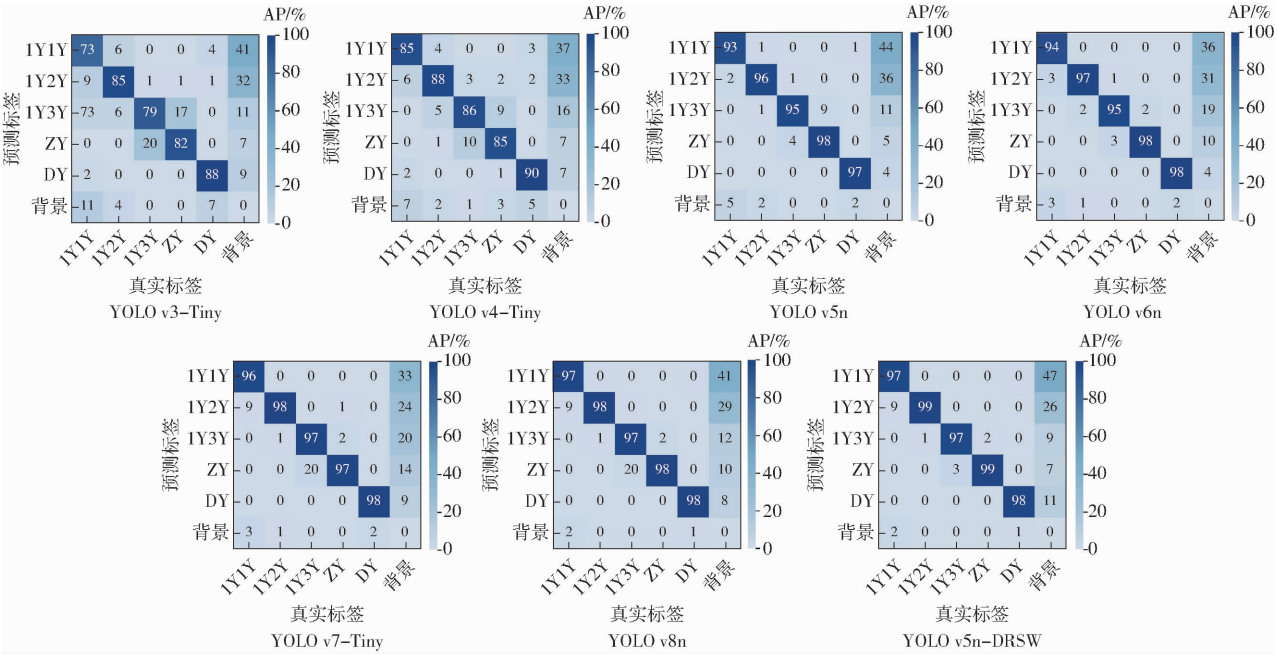

图 13 各轻量化模型在测试集上的混淆矩阵

Fig. 13 Confusion matrix of each lightweight model on test set

青形态实时检测及分选试验,试验场景及茶青智能分级装置分别如图 14、15 所示。



图 14 机采茶青分选场景

Fig. 14 Scene of sorting machine-picked tea leaves



图 15 茶青智能分级装置

Fig. 15 Intelligent grading device for tea leaves

1. 收集框 2. 进料口 3. 摄像头 4. 风机 5. 传送带 6. 称量托盘 7. 分级滚筒

YOLO v5n - DRSW 在实际应用场景中对机采茶青的在线检测表现出优异性能,各类形态的机采茶青在线检测结果如表 7 所示,各主流轻量化模型在 A ~ E 5 个实际应用场景中的部分典型实时检测效果对比见图 16,图中错检案例以红色椭圆形标出,漏检案例以红色矩形框标出。在场景 A 中,

表 7 机采茶青在线检测结果

Tab. 7 Online recognition results of machine-picked tea leaves %

茶青种类	正确检测率	错误检测率	漏检率
一芽一叶	95.3	4.7	1.1
一芽二叶	95.7	4.3	0.7
一芽三叶	93.5	6.5	0.2
枝梗	93.1	6.9	0.1
断碎叶	94.1	5.9	0.4

YOLO v3 - Tiny、YOLO v4 - Tiny、YOLO v6n 和 YOLO v7 - Tiny 均将一芽二叶错检为一芽三叶,而 YOLO v8n 将枝条错检为一芽三叶;在场景 B 中,YOLO v3 - Tiny、YOLO v4 - Tiny 和 YOLO v7 - Tiny 将一芽一叶错检为一芽二叶;在场景 C 中,YOLO v3 - Tiny、YOLO v4 - Tiny、YOLO v5n 和 YOLO v7 - Tiny 各出现一次漏检;在场景 D 中,YOLO v3 - Tiny 出现一次漏检,而 YOLO v4 - Tiny 和 YOLO v6n 将一芽二叶错检为一芽三叶,YOLO v5n 将一芽一叶错检为一芽二叶。上述漏检和错检主要集中在小尺度茶青目标上,如一芽一叶、一芽二叶以及断碎叶,表明未经改进的轻量化模型普遍在小目标检测方面存在能力不足的问题。

YOLO v5n - DRSW 在各类茶青形态的检测任务中均展现出优异的性能,整体平均检测准确率达 94.34%。其中,一芽三叶与枝梗因在外观特征上比较相似,容易造成混淆,因此二者形态的检测准确率在所有类别中相对最低,但依然保持在 93.1% 以上。此外,该模型的漏检率整体维持在较低水平,即

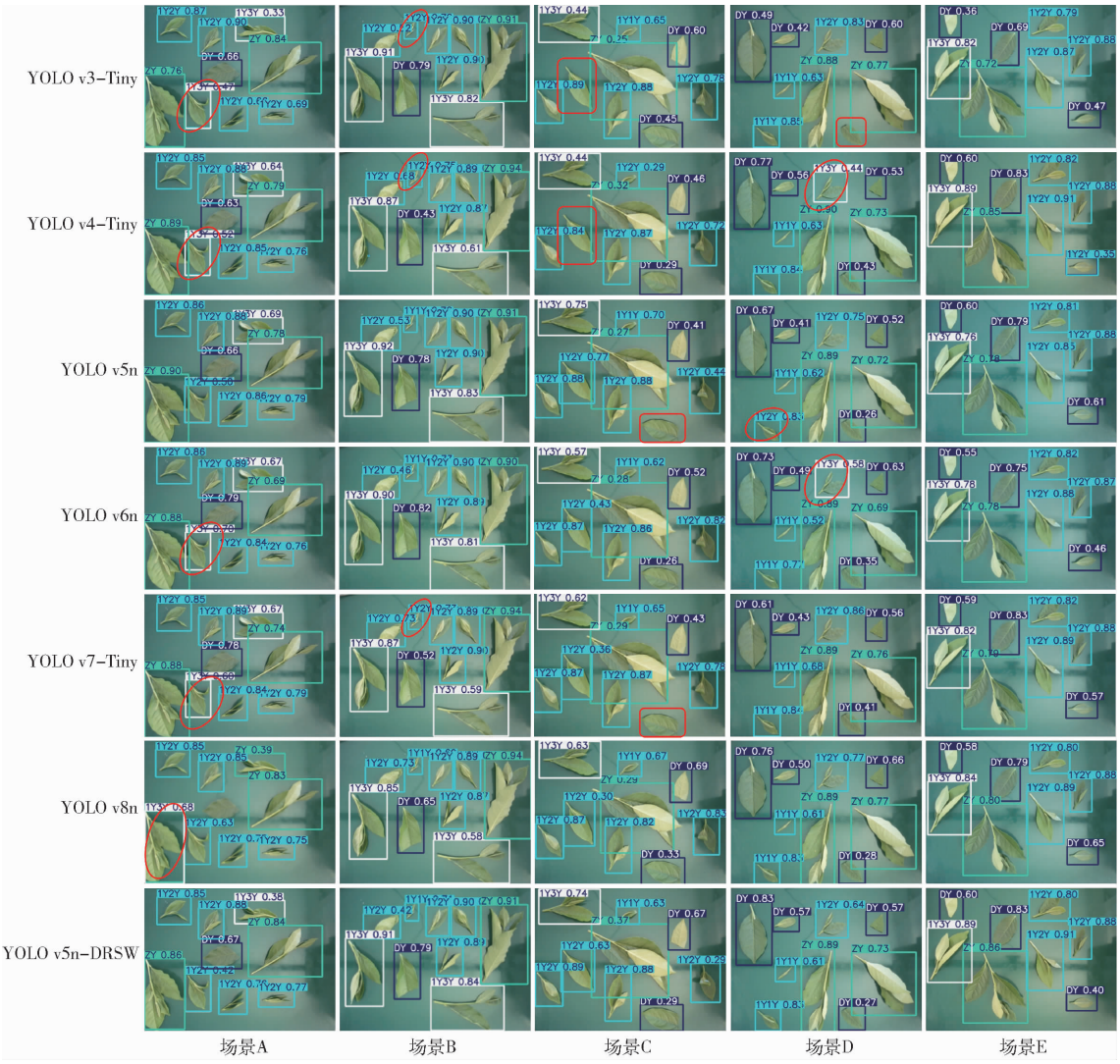

图 16　各主流轻量化模型在实际应用场景中的部分实时检测效果对比

Fig. 16　Comparison of some real-time recognition effects of mainstream lightweight models in practical application scenarios

便是在小目标检测中最具挑战性的一芽一叶形态下，漏检率也未超过 1.1%。上述试验结果充分表明，YOLO v5n－DRSW 具备出色的检测精度、良好的鲁棒性与泛化能力，能够高效、可靠地应对实际生产中茶青在不同形态与姿态下的在线检测需求。

4　结论

（1）为精准检测机械采收茶青形态，提高机采茶青自动化分级精度及效率，本研究提出了 YOLO v5n－DRSW 机采茶青形态检测模型。

（2）YOLO v5n－DRSW 在基线模型 YOLO v5n 的基础上，模型精确率提升了 2.1 个百分点，达 98.1%。此外，该模型显著减少了茶青误检、漏检以及预测框重叠现象，这表明引入 SE 注意力机制有效增强了模型的泛化能力；引入 RepGFPN 金字塔网络显著提升了模型的特征融合能力；将 WIoU 替换原 CIoU 损失函数能够有效缓解样本类别分布不平衡的问题。

（3）YOLO v5n－DRSW 的浮点运算次数相较于基线模型减少了 2.44%，YOLO v5n－DRSW 的单帧图像处理时间仅需 2.11 ms，显著优于 YOLO v5n，这表明将 DSConv 替换原 C3 模块不仅保持了模型的轻量化特性，还大幅提升了推理速度。

（4）依托智能茶青分级装置开展机采茶青形态实时检测及分选试验，结果表明，YOLO v5n－DRSW 在实际应用场景中的在线检测平均准确率达 94.34%，漏检率不超过 1.1%，展现出良好的鲁棒性与泛化能力，能够有效应对不同形态与姿态条件下的茶青在线检测任务。

（5）研究表明，YOLO v5n－DRSW 在机采茶青形态检测精度、模型轻量化与推理速度方面表现出色，适合树莓派等低算力边缘设备部署，可实现复杂形态茶青的高效在线检测，为机采茶青自动化分选提供了可靠的算法支持与技术储备。

参 考 文 献

[1] 中华人民共和国国家统计局. 中国统计年鉴[M]. 北京:中国统计出版社,2023.

[2] 杨肖委,沈强,罗金龙,等. 基于改进 YOLOv8n 的茶树嫩芽识别[J]. 茶叶科学,2024,44(6):949 - 959.
Yang Xiaowei, Shen Qiang, Luo Jinlong, et al. Research on tea bud recognition based on improved YOLOv8n[J]. Journal of Tea Science, 2024, 44(6): 949 - 959. (in Chinese)

[3] 宋彦,杨帅,郑子秋,等. 基于多头自注意力机制的茶叶采摘点语义分割算法[J]. 农业机械学报,2023,54(9):297 - 305.
Song Yan, Yang Shuai, Zheng Ziqiu, et al. Semantic segmentation algorithm based multi-headed self-attention for tea picking points[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(9): 297 - 305. (in Chinese)

[4] 张德高,刘敏基,谢焕雄. 农产品光电色选原理及色选机国内外研究发展现状[J]. 保鲜与加工,2020,20(3):233 - 237.
Zhang Degao, Liu Minji, Xie Huanxiong. Principle of agricultural products electronic color sorter and the domestic and foreign current situation and development[J]. Storage and Process, 2020, 20(3): 233 - 237. (in Chinese)

[5] Xue Z, Xu R, Bai D,et al. YOLO - Tea: a tea disease detection model improved by YOLOv5[J]. Forests, 2023, 14(2): 415.

[6] Shi M, Zheng D, Wu T, et al. Small object detection algorithm incorporating swin transformer for tea buds[J]. PLOS ONE, 2024, 19(3): e0299902.

[7] Lu J, Luo H, Yu C, et al. Tea bud DG: a lightweight tea bud detection model based on dynamic detection head and adaptive loss function[J]. Computers and Electronics in Agriculture, 2024, 227: 109522.

[8] 黄铝文,关非凡,谦博,等. 基于 2D DWT 与 MobileNetV3 融合的轻量级茶叶病害识别[J]. 农业工程学报,2023,39(24):207 - 214.
Huang Lyuwen, Guan Feifan, Qian Bo, et al. Recognizing tea diseases with fusion on 2D DWT and MobileNetV3[J]. Transactions of the CSAE, 2023, 39(24): 207 - 214. (in Chinese)

[9] Zhang Z, Lu Y, Zhao Y, et al. TS - YOLO: an all-day and lightweight tea canopy shoots detection model[J]. Agronomy, 2023, 13(5): 1411.

[10] 周宇杰,吴强,贺磊盈,等. 名优茶采摘机器人的系统设计与试验[J]. 机械工程学报,2022,58(19):12 - 23.
Zhou Yujie, Wu Qiang, He Leiying, et al. Design and experiment of intelligent picking robot for famous tea[J]. Journal of Mechanical Engineering, 2022, 58(19): 12 - 23. (in Chinese)

[11] 吴敏,郇晓龙,陈建能,等. 融合 2D 激光雷达与航向姿态参考系统的采茶机仿形方法研究与试验[J]. 茶叶科学,2023,43(1):135 - 145.
Wu Min, Huan Xiaolong, Chen Jianneng, et al. Research and experiment on profiling method of tea picker based on fusion of 2D - LiDAR and attitude and heading reference system[J]. Journal of Tea Science, 2023, 43(1): 135 - 145. (in Chinese)

[12] 杨信廷,郭向阳,韩佳伟,等. 基于 TiDE - PatchTST 模型的柑橘冷藏效率时序预测模型优化[J]. 农业机械学报,2024,55(7):396 - 404.
Yang Xinting, Guo Xiangyang, Han Jiawei, et al. Optimization of citrus cold storage efficiency time-series prediction model based on TiDE - PatchTST[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(7): 396 - 404. (in Chinese)

[13] 张昌松,李伟. 基于 YOLO v5 的直播油菜激光间苗系统设计与试验[J]. 农业机械学报,2024,55(4):40 - 52.
Zhang Changsong, Li Wei. Research and experiment of direct seeding rape laser seedling system based on YOLO v5[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(4): 40 - 52. (in Chinese)

[14] 罗志聪,李鹏博,宋飞宇,等. 嵌入式设备的轻量化百香果检测模型[J]. 农业机械学报,2022,53(11):262 - 269,322.
Luo Zhicong, Li Pengbo, Song Feiyu, et al. Lightweight passion fruit detection model based on embedded device[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(11): 262 - 269, 322. (in Chinese)

[15] 聂鹏程,钱程,汪清平,等. 赤点石斑鱼氨氮应激行为嵌入式表征研究[J]. 农业机械学报,2025,56(2):503 - 510,522.
Nie Pengcheng, Qian Cheng, Wang Qingping, et al. Behavioral representation for ammonia-nitrogen stress of *Epinephelus akaara* for embedded system[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(2): 503 - 510, 522. (in Chinese)

[16] Zhai X, Zong Z, Xuan K, et al. Detection of maturity and counting of blueberry fruits based on attention mechanism and bi-directional feature pyramid network[J]. Journal of Food Measurement & Characterization, 2024, 18(7): 6193 - 6208.

[17] Bai B, Wang J, Li J, et al. T - YOLO: a lightweight and efficient detection model for nutrient buds in complex tea-plantation environments[J]. Journal of the Science of Food & Agriculture, 2024, 104(10): 5698 - 5711.

[18] Han Y. Comparative analysis of improved YOLOv5 models for corrosion detection in coastal environments[J]. Journal of Marine Science and Engineering, 2024, 12(10): 1754.

[19] Bist R, Yang X, Subedi S, et al. Mislaying behavior detection in cage-free hens with deep learning technologies[J]. Poultry Science, 2023, 102(7): 102729.

[20] Gu Z, Zhu K, You S. YOLO－SSFS: a method combining SPD－Conv/STDL/IM－FPN/SIoU for outdoor small target vehicle detection[J]. Electronics, 2023, 12(18): 3744.

[21] Wang B, Hua J, Xia L, et al. A defect detection method for Akidzuki pears based on computer vision and deep learning[J]. Postharvest Biology and Technology, 2024, 218: 113157.

[22] 王会征,孙良晨,李新龙,等. 基于改进 YOLOv7－tiny 的番茄叶片病虫害检测方法[J]. 农业工程学报,2024,40(10): 194－202.
Wang Huizheng, Sun Liangchen, Li Xinlong, et al. Detecting tomato leaf pests and diseases using improved YOLOv7－tiny [J]. Transactions of the CSAE, 2024, 40(10): 194－202. (in Chinese)

[23] 汤文权,陈灼廷,王东桥,等. 基于改进 YOLO v5 的皮蛋裂纹在线检测方法[J]. 农业机械学报,2024,55(2):384－392.
Tang Wenquan, Chen Zhuoting, Wang Dongqiao, et al. Crack detection method for preserved eggs based on improved YOLO v5 for online inspection[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(2): 384－392. (in Chinese)

[24] 董宝鑫,王江涛. 基于重参数化广义金字塔与扩张残差的遥感图像旋转框算法[J]. 电子测量与仪器学报,2024, 38(12):54－61.
Dong Baoxin, Wang Jiangtao. Remote sensing image rotation box algorithm leveraging reparameterized generalized pyramid and dilated residual[J]. Journal of Electronic Measurement and Instrumentation, 2024, 38(12): 54－61. (in Chinese)

[25] Li K, Jiao P, Ding J, et al. Bearing defect detection based on the improved YOLOv5 algorithm[J]. PlOS ONE, 2024, 19(10): e0310007.

[26] Liu H, Li L, Li Y, et al. A hybrid attention mechanism and RepGFPN method for detecting wall cracks in high-altitude cleaning robots[J]. IEEE Photonics Journal, 2024, 16(5): 1－8.

[27] 徐红梅,杨浩,李亚林,等. 基于改进 YOLO－Pose 的复杂环境下拖拉机驾驶员关键点检测[J]. 农业工程学报,2023, 39(16):139－149.
Xu Hongmei, Yang Hao, Li Yalin, et al. Detecting the key points of tractor drivers under complex environments using improved YOLO－Pose[J]. Transactions of the CSAE, 2023, 39(16): 139－149. (in Chinese)

[28] 郭凯威,杨奎武,张万里,等. 面向文本识别的对抗样本攻击综述[J]. 中国图象图形学报,2024,29(9):2672－2691.
Guo Kaiwei, Yang Kuiwu, Zhang Wanli, et al. A review of adversarial examples for optical character recognition[J]. Journal of Image and Graphics, 2024, 29(9): 2672－2691. (in Chinese)

[29] Liu C, Liu X, Chen C, et al. Soft thresholding squeeze-and-excitation network for pose-invariant facial expression recognition [J]. The Visual Computer, 2023, 39(7): 2637－2652.

[30] 付玲,佘玲娟,颜镀镭,等. 基于内嵌物理信息与注意力机制 BiLSTM 神经网络的臂架系统疲劳损伤预测模型[J]. 机械工程学报,2024,60(13):205－215.
Fu Ling, She Lingjuan, Yan Dulei, et al. Fatigue damage prediction framework of the boom system based on embedded physical information and attention mechanism BiLSTM neural network[J]. Journal of Mechanical Engineering, 2024, 60(13): 205－215. (in Chinese)

[31] 韦柳梅,罗雪梅,康健. 改进 YOLOv8 的航拍图像小目标检测模型[J]. 计算机工程与科学,2025,47(12):2204－2215.
Wei Liumei, Luo Xuemei, Kang Jian. An improved YOLOv8 small object detection model in aerial image[J]. Computer Engineering & Science, 2025, 47(12): 2204－2215. (in Chinese)

[32] Yu C, Zhou L, Liu B, et al. G－YOLO: a YOLOv7-based target detection algorithm for lightweight hazardous chemical vehicles[J]. PLOS ONE, 2024, 19(4): e0299959.

[33] 黄华,张昊,胡晓林,等. 基于改进 YOLO v5 的复杂环境下花椒簇识别与定位方法[J]. 农业机械学报,2024,55(3): 243－251.
Huang Hua, Zhang Hao, Hu Xiaolin, et al. Recognition and localization method for pepper clusters in complex environments based on improved YOLO v5[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(3): 243－251. (in Chinese)