

DOI: 10.13733/j.jcam.issn.2095-5553.2025.08.015

胡永光, 金矿, 刘益航, 等. 改进 YOLOv5s 的自然场景下茶叶嫩芽检测[J]. 中国农机化学报, 2025, 46(8): 103–111

Hu Yongguang, Jin Kuang, Liu Yihang, et al. Improved YOLOv5s algorithm of tea buds detection in natural scenes [J]. Journal of Chinese Agricultural Mechanization, 2025, 46(8): 103–111

改进 YOLOv5s 的自然场景下茶叶嫩芽检测^{*}

胡永光^{1, 2}, 金矿^{1, 2}, 刘益航^{1, 2}, 鹿永宗^{1, 2}, 张志^{1, 2}, 潘庆民^{1, 2}

(1. 江苏大学农业工程学院, 江苏镇江, 212013; 2. 江苏大学现代农业装备与技术教育部重点实验室, 江苏镇江, 212013)

摘要: 茶叶嫩芽采摘是茶叶生产中的重要环节, 智能化采茶需要基于深度学习的茶叶嫩芽检测算法作为技术保障。为提高茶叶嫩芽检测速度和精度, 提出一种改进 YOLOv5s 网络模型的检测算法 TN—YOLOv5s。首先引入 GhostNet 网络结构中的深度可分离卷积 GhostConv, 替代原有 YOLOv5s 模型中特征提取网络与特征融合网络的普通卷积层; 然后在模型特征提取网络的末端加入 CA 空间注意力机制; 其次使用 SIoU_Loss 代替 CIoU_Loss 作为回归损失函数; 最后使用 Soft—NMS 替换 NMS。结果表明, 改进后的模型在自制的茶叶嫩芽数据集上, 相比于原始 YOLOv5s 算法, 模型精确率、召回率和平均精度均值分别高出 7.1%、5.9%、6.4%, 权重大小从 13.7 MB 下降到 7.48 MB。同时与目前主流的检测算法相比, 在检测精度、模型大小和检测速度方面更有优势。改进后的模型降低被遮挡茶叶嫩芽的漏检率, 能够准确快速地实现不同场景下的茶叶嫩芽检测。

关键词: 茶叶嫩芽; 自然场景; YOLOv5s; 目标检测; 遮挡

中图分类号: TP391.4; S57.1 **文献标识码:** A **文章编号:** 2095-5553 (2025) 08-0103-09

Improved YOLOv5s algorithm of tea buds detection in natural scenes

Hu Yongguang^{1, 2}, Jin Kuang^{1, 2}, Liu Yihang^{1, 2}, Lu Yongzong^{1, 2}, Zhang Zhi^{1, 2}, Pan Qingmin^{1, 2}

(1. School of Agriculture Engineering, Jiangsu University, Zhenjiang, 212013, China;

2. Key Laboratory of Modern Agricultural Equipment and Technology, Ministry of Education, Jiangsu University, Zhenjiang, 212013, China)

Abstract: The plucking of tender tea buds is a crucial stage in tea production, and intelligent tea harvesting relies on a deep learning-based tea bud detection algorithm as a technical foundation. To enhance the speed and accuracy of tea bud detection, this paper proposes an improved detection algorithm TN—YOLOv5s based on the YOLOv5s network model. First, GhostConv, a depthwise separable convolution from the GhostNet network structure, is introduced to replace the ordinary convolution layers in the feature extraction and fusion networks of the original YOLOv5s model. Second, the CA spatial attention mechanism is added at the end of the model's feature extraction network. Third, the SIoU_Loss is employed as the regression loss function in place of CIoU_Loss. Finally, Soft—NMS is used to replace NMS. The results of the study show that the improved model achieved 7.1%, 5.9%, and 6.4% higher model precision, recall, and average precision values, respectively, and the weight size decreased from 13.7 MB to 7.48 MB compared to the original YOLOv5s algorithm on the custom tea buds dataset. Furthermore, when compared to current mainstream detection algorithms, the improved model shows better performance in detection accuracy, model size, and detection speed. The improved model reduces the leakage detection rate of obscured tea buds, and enable accurate and rapid detection of tea buds in different scenarios, providing a technical foundation for the development of tea-picking robot technology and equipment.

Keywords: tea buds; natural scenes; YOLOv5s; object detection; occlusion

收稿日期: 2024 年 1 月 18 日 修回日期: 2024 年 3 月 6 日

^{*} 基金项目: 国家自然科学基金地区科学基金项目 (32460782); 江苏高校优势学科建设工程项目 (PAPD—2018—87)

第一作者(通讯作者): 胡永光, 男, 1976 年生, 江苏徐州人, 博士, 教授; 研究方向为茶果园生产技术与装备。E-mail: deerhu@163.com

0 引言

茶叶嫩芽采摘是茶叶生产中的重要环节之一,高质量的采摘对于提升茶叶产量、质量和经济效益具有重要意义^[1]。长期以来,茶叶生产主要利用春季鲜叶,而产量比春茶高数倍的夏秋茶因为采收成本高而被大量弃采。提高夏秋茶的全年利用率,让其产生更高的经济价值具有重要意义^[2]。实现对夏秋茶的机械化采摘是降低采收成本的重要途径,基于深度学习技术实现对茶叶嫩芽的检测是确保机械化采收质量的重要手段。对大多数茶树品种而言,夏秋茶嫩芽颜色与老叶颜色接近^[3],且其在机采蓬面上分布密集,嫩芽间存在相互遮挡,上述原因使得夏秋茶茶叶嫩芽检测难度较大。因此,开展夏秋茶茶叶嫩芽检测方法的研究对于实现夏秋茶的机械化和智能化采摘具有重要的意义。

目前 2 种茶叶采摘方式:人工采摘和机械采摘。名优茶采用人工采摘,工人能够根据茶叶的老嫩程度选择性采摘茶芽,确保鲜叶的外形和品质,但其存在采摘效率低、用工成本高甚至用工荒的问题^[4]。大宗茶的采摘方式多为机械采摘,其采摘效率虽然较高,但对蓬面上的嫩芽和老叶一刀切,易造成嫩芽的损坏,嫩芽完整度低、含杂质高^[5]。目前为兼顾采摘质量和效率,出现了新型的智能采茶机器人概念,由于茶叶采摘的复杂性,其多处于实验室研究阶段。结合图像处理与机器视觉技术,实现对夏茶茶叶嫩芽的检测,可为实现夏茶机械化和智能化采摘奠定基础,是未来智能采茶机器人的核心技术之一。

传统图像处理技术需要手动进行特征设计,其过程复杂且识别精度较低,难以处理复杂场景和大规模数据。汪建^[6]提出一种基于图像颜色和区域生长相结合的茶叶图像分割算法,其通过在颜色空间上对像素进行划分,并生成种子区域,然后对种子区域基于颜色的相似性和区域的邻接性进行区域生长,最后结合颜色距离和边缘距离进行区域生长和合并来分割出嫩芽。吴雪梅等^[7]根据嫩芽和老叶的 G 和 $G-B$ 分量的颜色信息,利用改进的最大方差自动取阈法计算 G 和 $G-B$ 分量的分割阈值,分割出嫩芽。Zhang 等^[8]提出一种基于改进分水岭算法的茶芽识别分割方法,其通过对蓝分量进行自适应阈值处理,再结合绿分量获得新的组合分量灰度图,最后利用改进的分水岭算法提高嫩芽识别精度。上述方法虽然能够实现对茶叶嫩芽的识别,但在田间环境下,其效果受光照条件和噪声的影响较大,模型的田间鲁棒性和泛化能力较差。

随着深度学习的快速发展和大量应用,其在现代

农业中的研究也在增加。Wang 等^[9]提出了一种基于区域卷积神经网络(R-CNN)的自然场景下茶叶嫩芽识别方法 Mask-RCNN,验证了该检测方法在复杂环境中的通用性和鲁棒性。Li 等^[10]通过对原始嫩芽数据集进行数据增强,提高样本多样性,然后在 YOLOv3 模型中加入空间金字塔池化模块,提出改进 YOLOv3 模型并对茶叶嫩芽进行检测,平均检测精度达到 89.61%。王梦妮等^[11]通过在 YOLOv5s 网络模型主干网络中替换空洞空间卷积池化金字塔结构 ASPP,并在颈部引入双向特征金字塔网络 BiFPN 和卷积注意力机制 CBAM,提出一种改进的 YOLOv5s 模型,相比原模型,准确率、召回率和平均精度值分别提高 4.4%、0.5%、4.0%。

与传统图像处理方法相比,基于深度学习的自然场景茶叶嫩芽检测取得良好效果,但仍然存在以下问题:(1)田间茶园场景下茶叶嫩芽目标小,生长密集,彼此相互遮挡,嫩芽与老叶之间颜色和形状相似,导致识别算法检测精度低,漏检和误检率高;(2)现有的针对 YOLOv5s 算法的改进措施来提高算法的精度,减少漏检和误检的发生,往往会造成改进模型参数量和计算量成倍增加,造成计算资源浪费且不利于模型部署。针对上述问题,本文以 YOLOv5s 模型为基础,根据夏茶茶叶嫩芽检测存在的难点对原网络模型进行改进,将田间拍摄得到的茶叶嫩芽图像进行标注,输入改进后的模型,实现一种轻量化、检测速度快和漏检率低的自然场景下的夏茶茶叶嫩芽检测模型,为智能采茶机器人的视觉系统提供技术支持。

1 数据集构建

1) 图像采集。所用茶叶嫩芽数据集原始图像采集地点为江苏省句容市茅山茶场,采集对象为龙井 43,茶树单拢栽种且长势较好、均一。图像采集设备为红米 K20 pro 手机和 Canon M50 Mark II 相机,图像的分辨率分别是 4 000 像素×3 000 像素和 5 328 像素×4 000 像素,以 JPG 格式保存。数据采集于 2023 年 6 月上旬,拍摄时间为 5:30—18:30。在对茶叶嫩芽进行图像采集时,拍摄设备与茶树距离 10~50 cm,拍摄角度与竖直向上方向呈 30°~60°夹角。共采集原始数据集图像 1 920 幅。采集到不同背景复杂度、拍摄视角高低和嫩芽数目差异如图 1 所示。

2) 数据预处理和增强。使用 LabelImg 图像标注工具,选择一芽二叶的最小外接矩形作为标注框,手动对图像数据集进行标注。自然环境下,光照强度随时间变化而不同,导致同一嫩芽所展现的亮度在不同时间段有所差异,因此光照强度是影响嫩芽识别效果的

重要因素。所以进行嫩芽图像数据集构建时,在一天中不同时间段内进行拍摄,如图 2 所示。

采集数据时沿茶行两侧移动拍摄,拍摄过程中太阳位置和拍摄位置的变化导致茶叶嫩芽图像整体亮度以及图像中部分嫩芽的亮度不同,如图 3(a)和图 3(b)所示。在同一茬口处萌发的嫩芽,经过一段时间的生长,往往分布密集导致相互遮挡,图 3(c)和图 3(d)分别为茶叶嫩芽在无遮挡和相互遮挡条件下的图像。

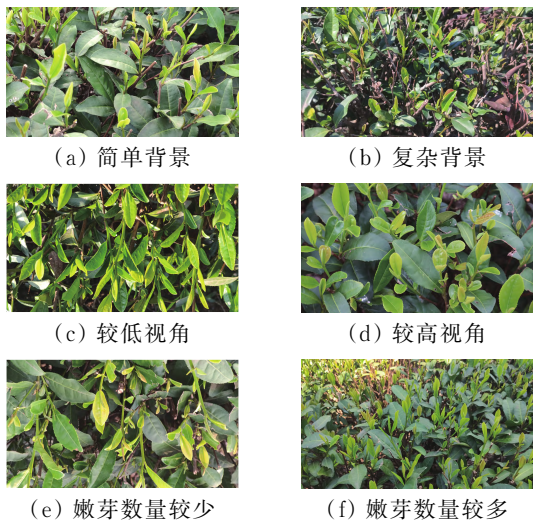


图 1 不同场景下的茶叶嫩芽图像

Fig. 1 Images of tea buds in different scenes

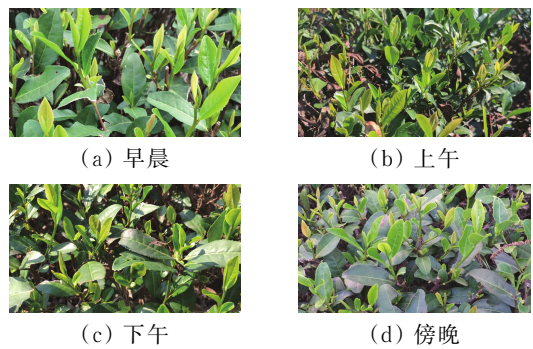


图 2 不同时间段内的茶叶嫩芽

Fig. 2 Tea buds in different time periods



图 3 顺逆光和遮挡下的茶叶嫩芽

Fig. 3 Tea buds under smooth backlighting and shading

对采集到的数据集进行筛选,去除背景虚化严重和模糊的图像,剩余原始图像的采集信息如表 1 所示。

表 1 原始图像的采集信息

Tab. 1 Acquisition information of the original image 幅

序号	影响因素	拍摄条件	图片数量
1	不同时段	早晨	513
		上午	464
		下午	455
		傍晚	489
2	光照情况	顺光	955
		逆光	966
3	遮挡情况	有遮挡	1 420
		无遮挡	500

在早晨、上午、下午、傍晚采集的图像数量分别为 513 幅、464 幅、455 幅、489 幅;顺光和逆光情况下的图像数量分别为 955 幅和 966 幅;有遮挡和无遮挡情况下的图像数量分别为 1 420 幅和 500 幅。为增强试验数据的丰富性,提高模型的泛化能力,对原始图像数据进行翻转、镜像、添加噪声、放大和缩小以扩充数据集,增强效果如图 4 所示。首先,将原始数据集图像进行混合汇总,按 6:2:2 的比例随机划分为训练集、验证集和测试集,其数量分别为 1 152 幅、384 幅、384 幅。然后对训练集中的每幅图像采用上述方法随机组合进行 2 倍数据增强,同时对图像对应的标注文件进行变换。将增强后的图像与原始训练集中的图像汇总得到 3 456 幅图像数据。最终,训练集、测试集和验证集的图片数量分别为 3 456 幅、384 幅、384 幅,使用增强后的数据集进行模型的训练。

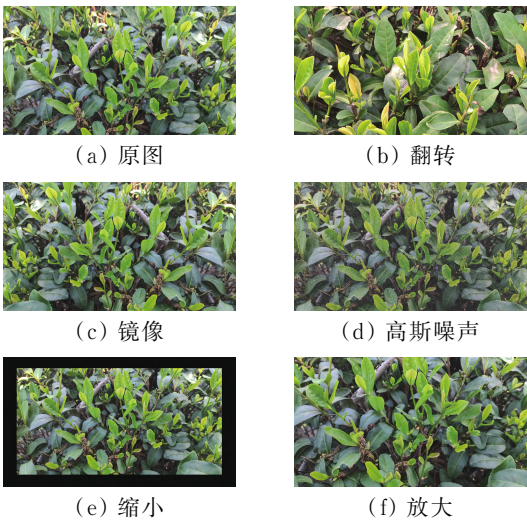


图 4 数据集增强效果

Fig. 4 Dataset enhancement effects

2 嫩芽识别算法及改进

YOLOv5 是一种基于深度学习的目标检测算法,是 YOLO(YOU Only Look Once)系列算法的第五代版本,在更加轻量化的同时提供更高的检测精度,自面世以来,凭借着高性能在学术界和工业界都取得不错

的成果,且其在茶叶嫩芽识别任务中表现出色^[12, 13]。YOLOv5 根据模型大小可分为 5 个不同的版本,分别为 YOLOv5n、YOLOv5s、YOLOv5m、YOLOv5l、YOLOv5x,其中 YOLOv5s 通过采用较小的网络深度和特征图维度来实现较小的权重文件和更快的检测速度^[14]。因此,选择在该模型的基础上进行优化,YOLOv5s 的网络结构如图 5 所示。

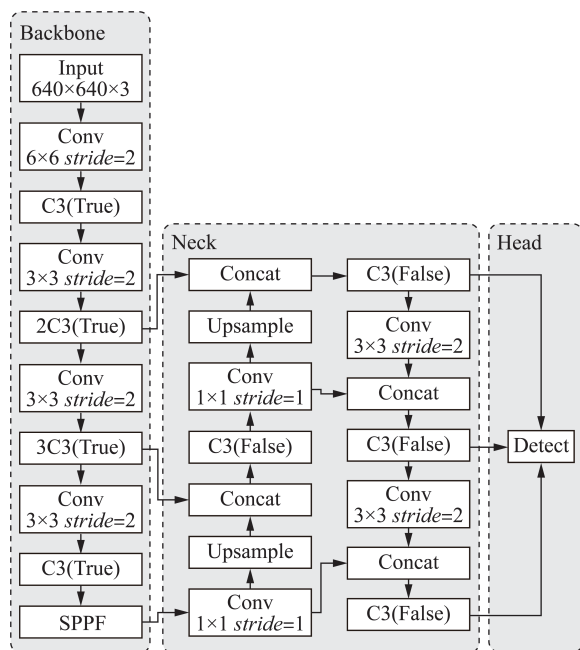


图 5 原始 YOLOv5s 网络结构

Fig. 5 Original YOLOv5s network structure

最终改进后模型的网络结构如图 6 所示。

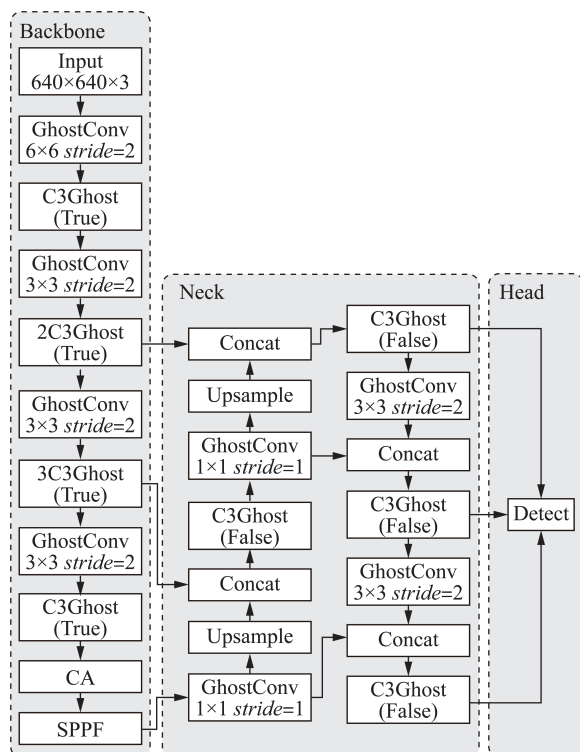


图 6 改进 YOLOv5 网络结构

Fig. 6 Improved YOLOv5s network structure

自制的夏茶茶叶嫩芽数据集场景丰富,检测目标茶叶嫩芽存在相互遮挡、外观颜色接近和目标较小等特点,数据集较为复杂。原始的 YOLOv5s 算法识别精度低,漏检和误检率高,难以满足要求。故提出一种改进 YOLOv5s 网络模型,主要从 4 个方面进行改进。

2.1 轻量化网络 GhostNet

在实际应用场景中,目前茶叶嫩芽检测存在模型参数量大的问题,虽然能够达到检测要求,但是耗费的计算资源较大,不适合在嵌入式平台上部署。为降低模型参数量,提升检测速度,引入 GhostNet 网络结构^[15],该网络在不降低检测精度的同时,减少模型冗余的参数量和计算量,从而提高检测速度。其核心思想是通过引入“Ghost Module”来减小参数量和计算量,其结构如图 7 所示。Ghost Module 可以分为 2 个主要部分,第一部分:使用 1×1 的卷积获得输入特征的必要特征浓缩;第二部分:使用逐层卷积获取额外的特征图,然后将获得的特征图和 1×1 卷积后的结果进行堆叠,以获得最终的结果输出。

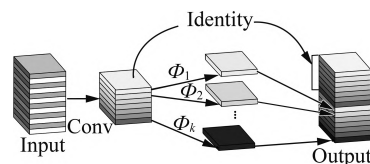


图 7 Ghost 卷积模块

Fig. 7 GhostConv module

Ghost Bottlenecks 是由 Ghost Module 组成的瓶颈结构。Ghost Bottlenecks 的不同步长如图 8 所示, Ghost Bottlenecks 主要由多个 Ghost Module 组成,第 1 个 Ghost Module 作为扩展层扩大感受野,通过卷积的方式增加通道数,第 2 个 Ghost Module 调整合适的输出通道数,便于后续的特征合并。通过短切的方式连接 Ghost Bottleneck 的输入和输出。步长为 2 的 Ghost Bottleneck 模块在步长为 1 的 Ghost Bottleneck 中间插入一个深度卷积,并在每一层后面都应用批量归一化。

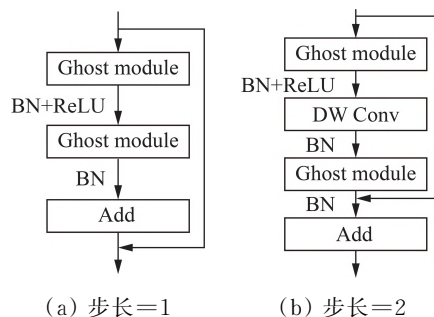


图 8 Ghost Bottleneck 两种不同结构

Fig. 8 Ghost Bottleneck in two different configurations

基于上述,设计 C3Ghost 模块,在相同数量的输

入和输出通道下, C3Ghost 模块是将 C3 模块中 Bottleneck 的 Conv 模块替换为 Ghost 模块, 即使用步长为 1 的 Ghost Bottleneck 结构替代 C3 模块中的 Bottleneck 结构。因此, C3Ghost 模块的参数量和计算量更少。将 YOLOv5 算法主干和颈部网络中原本的 Conv 模块和 C3 模块替换为 Ghost 模块和 C3Ghost 模块, 主干网络和颈部网络的修改实现在提取图片的有效特征和进行特征融合的同时, 进一步降低改进后的基准主干网络模型的大小, 从而提高茶叶嫩芽检测网络的检测速度。由 Ghost 模块构成的 C3Ghost 模块结构如图 9 所示。

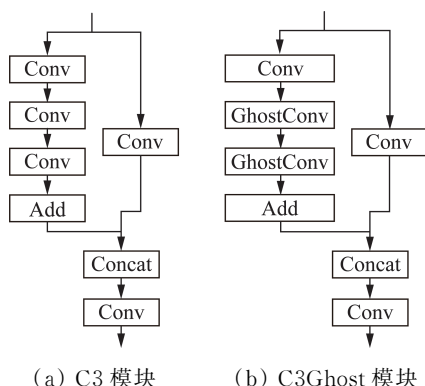


图 9 C3 模块和 C3Ghost 模块

Fig. 9 C3 module and C3Ghost module

2.2 CA 注意力机制

引入注意力机制 CA, 其核心作用在于使网络关注到它所更需要关注的地方, 目前已被广泛使用来加强模型的计算能力。当前有几种优秀的注意力机制, 如 SE9 通道注意力机制^[16]、CBAM9 卷积注意力模块^[17]和 CA9 空间注意力机制^[18]等。在改进模型时, CA 模块被添加到骨干网络中, 以提高捕获位置信息的能力。

CA 将通道注意力分解为 2 个并行的一维(2D)特征编码。也就是说, 输入特征被聚合成垂直和水平 2 个独立的方向特征图。然后这 2 个特征图被编码为 2 个注意力图, 一个用于长期依赖关系, 另一个用于位置信息。因此 CA 模块的操作分为 2 个步骤: 坐标信息嵌入与坐标注意力生成。如图 10 所示, 坐标信息嵌入对应 X 平均池化和 Y 平均池化处理。在此步骤中, 输入每个通道的特征图 X 沿水平和垂直方向编码, 分别使用大小和池化内核 $H \times 1$ 和 $1 \times W$ 。相应地, 得到感知特征图, 表示为 z^h 和 z^w 。为生成坐标注意力, 模型先将 z^h 和 z^w 进行级联, 使用核大小为 1×1 的 F_1 共享卷积提取联合特征。然后, 利用 δ 非线性激活函数生成中间特征图 f 水平和垂直方向的空间信息。 f 沿空间维度被切割成两个独立的张量 f^h 与 f^w , 利用两个 1×1 卷积变换 F_h 和 F_w 将空间维数转换为与输入特

征张量 X 通道数相同的张量 g^h 和 g^w , 如式(2)和式(3)所示。其中注意力权重 g^h 和 g^w 在相应的两个方向上, 最终通过特征图获得的 x_c 与注意力权重相乘后输出 y_c , 如式(4)所示。

$$f = \delta(F_1(z^h, z^w)) \quad (1)$$

$$g^h = \sigma(F_h(f^h)) \quad (2)$$

$$g^w = \sigma(F_w(f^w)) \quad (3)$$

$$y_c(i, j) = x_c(i, j) \times g_c^h(i) \times g_c^w(j) \quad (4)$$

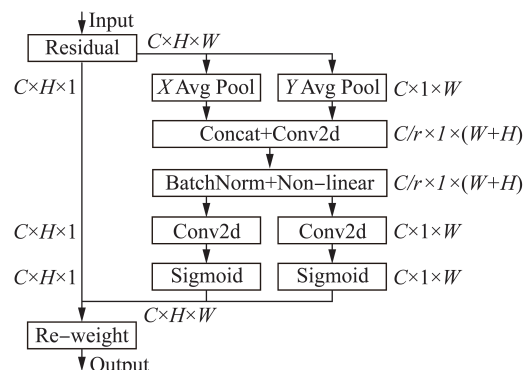


图 10 CA 注意力结构

Fig. 10 CA attention structure

2.3 损失函数的改进

YOLOv5 算法的损失函数由边界框回归损失、类别损失、置信度损失三部分组成, 其中常用交并比 IoU 作为损失函数, 来描述边界框回归损失^[19]。针对损失函数理论的不足, $GIoU_Loss$ 、 $DIoU_Loss$ 、 $CIOU_Loss$ 相继被提出来解决边界框回归问题, YOLOv5 原始模型中选择 $CIOU_Loss$ 作为回归损失函数。

现有方法匹配真实框和预测框之间的交并比 IoU 、中心点距离和宽高比等, 均未考虑真实框和预测框之间不匹配的问题。针对以上问题, 选用 $SIoU$ 损失函数^[20]替换 $CIOU$ 损失函数, $SIoU$ 和 $SIoU$ 损失函数的定义如式(5)和式(6)所示。

$$Loss_{SIoU} = IoU - \frac{\Delta + \Omega}{2} \quad (5)$$

$$\delta_{SIoU} = 1 - Loss_{SIoU} \quad (6)$$

式中: $Loss_{SIoU}$ —— $SIoU$ 损失值;

δ_{SIoU} —— $SIoU$ 损失函数;

Ω ——形状损失;

Δ ——距离损失。

$$\Omega = \sum_{t=\omega, h} (1 - e^{-\omega_t})^\theta \quad (7)$$

$$\Delta = \sum_{t=x, y} (1 - e^{-\gamma \rho_t}) \quad (8)$$

其中, (ω, h) 和 (ω^g, h^g) 分别表示预测框和真实框的宽和高, $\omega_\omega = \frac{|\omega - \omega^g|}{\max(\omega, \omega^g)}$, $\omega_h = \frac{|h - h^g|}{\max(h, h^g)}$, θ 的值通过遗传算法计算取得, 表示对形状损失的关注程度,

$\rho_x = (\frac{b_x^{\text{gt}} - b_x}{c_w})^2$, $\rho_y = (\frac{b_y^{\text{gt}} - b_y}{c_h})^2$, c_w 和 c_h 分别表示最小外接矩形的宽和高, γ 计算如式(9)所示。

$$\gamma = 1 + 2 \sin^2 \left[\arcsin \frac{|b_y^{\text{gt}} - b_y|}{\sqrt{(b_x^{\text{gt}} - b_x)^2 + (b_y^{\text{gt}} - b_y)^2}} - \frac{\pi}{4} \right] \quad (9)$$

式中: (b_x, b_y) 、 $(b_x^{\text{gt}}, b_y^{\text{gt}})$ ——预测框和真实框的中心点坐标。

SIoU 损失函数在 CIoU 损失函数的基础上规定了预测框向真实框靠近的方向,避免可预测框的游荡现象,可以进一步提高模型的回归精度,因此,使用 SIoU 损失函数计算回归损失。

2.4 柔性非极大值抑制 Soft—NMS 算法

使用柔性非极大值抑制算法(Soft—NMS)^[21]替换非极大值抑制算法(NMS)算法作为改进模型的预测框过滤算法。自然环境下茶叶嫩芽间存在遮挡,同一茬口处茶叶嫩芽分布密集,导致被检测茶叶嫩芽重叠比例较高。如果被检测物体密集分布,则边界框之间重叠较大,NMS 算法在选择具有高置信度的边界框时会忽略其他具有较低置信度但仍然有效的边界框,导致一些被遮挡茶叶嫩芽的边界框被错误地抑制或删除,Soft—NMS 算法的替换可有效解决上述错误抑制导致的漏检问题。

与传统 NMS 算法采用类硬阈值来判断相邻检测框是否保留不同,Soft—NMS 首先选定一个分数最高的预测框作为基准框,进而计算基准框与同类别其他预测框的 IoU ,当 IoU 值小于人工设定的阈值时,该预测框不被抑制;反之,则对该预测框的置信度使用得分衰减函数机制进行计算,而非彻底将其得分直接置为 0 分^[22]。

Soft—NMS 的重叠边界框得分衰减机制使得抑制效果在重叠区域内平滑过渡,能够更好地保留被遮挡目标的检测结果,Soft—NMS 计算如式(10)所示。

$$S_i = \begin{cases} S_i & IoU(M, b_i) < N_t \\ S_i e^{-\frac{IoU(M, b_i)^2}{\sigma}} & IoU(M, b_i) \geq N_t \end{cases} \quad (10)$$

式中: S_i ——候选框 b_i 的置信得分;

σ ——控制衰减速度的参数;

M ——最大得分框;

N_t —— IoU 阈值,控制是否对得分进行衰减。

3 模型训练与评价指标

3.1 试验平台

试验环境为 Windows 10 操作系统,处理器型号为 Intel (R) Core (TM) i5—13600KF CPU@

3.5 GHz,系统内存为 32 GB, NVIDIA GeForce RTX 3060 显卡,显存为 12 GB。深度学习框架采用 PyTorch 1.12.1,编程平台为 PyCharm,编程语言为 Python 3.8,所有算法均在相同环境下运行。

3.2 模型训练参数

模型训练时,使用标准的随机梯度下降法训练模型,动量设置为 0.937,初始学习率为 0.01,权重衰减为 0.000 5。考虑到模型训练时硬件平台的 GPU 内存限制,将输入图片大小调整为 640 像素×640 像素, $batch\ size$ 设置为 16,训练 300 个 $epoch$ 。

3.3 评价指标

采用精确率(P)、召回率(R)、平均精度均值(mAP)、权重大小($Weights$)、检测速度(FPS)、模型参数量($Params$)等指标对模型的性能进行评价。其中精确率 P 、召回率 R 和平均精度均值 mAP 的计算如式(11)~式(12)所示。

$$P = \frac{TP}{TP + FP} \times 100\% \quad (11)$$

$$R = \frac{TP}{TP + FN} \times 100\% \quad (12)$$

式中: TP ——被正确检测出的目标数,即实际为正样本,检测也为正样本;

FP ——被错误检测出的目标数,即实际为负样本,被检测为正样本;

FN ——检测错误的目标数,即实际为正样本,被检测为负样本。

由于本研究仅识别 LJ43 标签嫩芽,即分类的类别数 $C=1$ 。因此, mAP 即为平均精度 AP 。

4 结果与分析

4.1 改进算法的消融实验性能对比

为评估上述改进中每一个关键模块的有效性,进行消融实验,以分析不同模块对检测算法性能的影响。以未做任何改进的原始 YOLOv5s 为基准,将原始模型中特征提取网络和特征融合网络中的普通卷积层替换为深度可分离卷积 GhostConv,添加注意力模块 CA,再替换损失函数 SIoU_Loss 和柔性非极大抑制 Soft—NMS 进行实验,“√”表示引入该方法,使用 4 种改进方法的模型为 TN—YOLOv5s,实验结果如表 2 所示。

为降低模型的参数量,在原始 YOLOv5s 引入 GhostNet 结构,即将普通卷积层替换为 Ghost conv。由表 2 可知,通过引入 GhostConv 模块来构建整个量化模型,模型参数量下降 47.53%,而 mAP 仅下降 0.4%,实验说明 GhostNet 结构的引入能够有效减少模型的参数量。但模型嫩芽精确率略有降低,主要原

因在于引入 GhostNet 结构,简化模型网络结构的同时,损失部分嫩芽特征信息。

为减少部分嫩芽特征损失对模型的影响,在骨干网络末端加入 CA 注意力机制。由表 2 可知,相比未添加注意力机制的 YOLOv5s—GC 模型,模型参数量只提升 0.70%,而 mAP 提升 4.9%。因此,添加注意

力机制能够促进不同通道嫩芽有效信息的交互和传递,抑制无关信息,提升模型的检测性能。

为提升模型对小目标嫩芽检测精度,使用 SIoU 替换 CIoU。由表 2 可知,检测精确率提高 2.7%, mAP 值提升 0.8%。说明 SIoU 的替换考虑嫩芽预测框方向匹配问题,能够更好地优化嫩芽预测框的位置。

表 2 消融实验结果
Tab. 2 Results of ablation test

算法模型	GhostConv	CA	SIoU_Loss	Soft—NMS	$P/\%$	$R/\%$	$mAP/\%$	Params/M
YOLOv5s					74.9	73.5	78.4	7.022
YOLOv5s—G	✓				73.4	73.7	78.0	3.685
YOLOv5s—GC	✓	✓			78.6	76.9	82.9	3.710
YOLOv5s—GCS	✓	✓	✓		81.3	77.4	83.7	3.710
TN—YOLOv5s	✓	✓	✓	✓	82.0	79.4	84.8	3.710

GhostConv 结构的引入、注意力机制 CA 的添加和 SIoU 的替换使模型的检测性能得到提升,但当目标密集分布时,模型存在漏检问题。主要原因在于 YOLOv5s 采用 NMS 作为算法的后处理方法,抑制多余候选框,但对于高密度重叠目标存在一次抑制从而发生漏检现象。由表 2 可知,替换柔性非极大抑制 Soft—NMS 后,模型检测精确率和召回率分别提高 0.7%、2.0%, mAP 提升 1.1%。柔性非极大抑制 Soft—NMS 减少重复嫩芽检测框的影响,有效提高模型对相互遮挡嫩芽的检测性能。

最终,改进后 TN—YOLOv5 模型的精确率、召回率和 mAP 比原 YOLOv5s 模型分别高出 7.1%、5.9%、6.4%。为更直观地展示改进后模型的性能,图 11 为原模型和改进后模型的性能曲线对比。通过消融实验证实改进方法可以改善模型对夏茶茶叶嫩芽的检测性能。

4.2 不同目标检测模型的对比试验

为更全面地验证改进后网络模型的性能,在同一试验环境下与一些经典的目标检测模型进行对比试验。将 TN—YOLOv5s 与 Faster R—CNN^[23]、SSD^[24]、YOLOv3^[25]、YOLOv4^[26]、YOLOv4—Tiny^[27]、YOLOv8s^[28] 算法进行试验对比,结果如表 3 所示。

表 3 不同网络模型的试验结果对比
Tab. 3 Comparison of experimental results of different network models

算法模型	$P/\%$	$R/\%$	$mAP/\%$	权重大小/MB	FPS /(帧·s ⁻¹)
Faster R—CNN	48.69	65.7	53.27	108	12.51
SSD	70.29	44.31	53.68	90.6	47.58
YOLOv3	78.47	54.52	64.25	235.0	61.95
YOLOv4	80.60	64.45	68.39	244.0	51.07
YOLOv4—Tiny	72.49	52.36	60.72	22.4	120.21
YOLOv5s	74.9	73.5	78.4	13.7	107.53
YOLOv8s	70.6	77.9	81.3	21.4	121.95
TN—YOLOv5s	82.0	79.4	84.8	7.48	74.07

由表 3 可知,TN—YOLOv5s 模型相比于其他主流检测模型,不仅有最小的模型体积,同时保持最高的检测精度,推理速度也能够满足对茶叶嫩芽的实时检测。TN—YOLOv5s 模型相比于双阶段检测模型 Faster R—CNN 的平均精度均值高出 31.53%,相比于单阶段目标检测模型 SSD、YOLOv3、YOLOv4、YOLOv4—Tiny、YOLOv5s、YOLOv8s 的平均精度均值分别高出 31.12%、20.55%、16.41%、24.08%、6.4%、3.5%。

综上,TN—YOLOv5s 算法在保持轻量化的同时有着最高的检测精度,同时能够满足实时性的要求,整体表现较为突出,证明改进模型对茶叶嫩芽识别的可

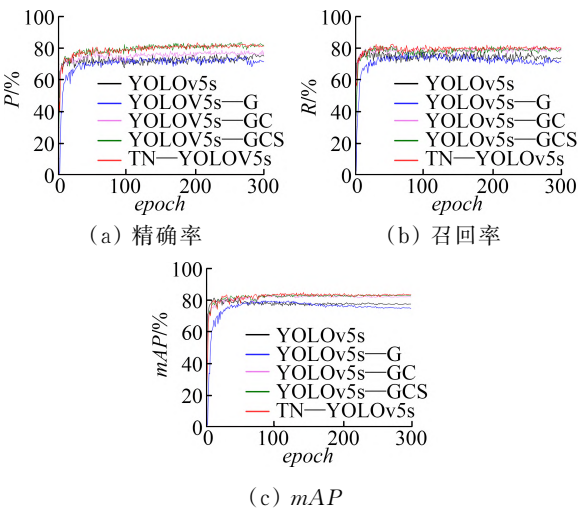


图 11 基于不同改进方法的模型训练过程中性能参数变化对比
Fig. 11 Comparison of performance parameter changes during model training based on different improvement methods

行性和优越性。

4.3 不同场景下的检测效果对比分析

为验证 TN—YOLOv5s 算法在不同场景对茶叶嫩芽检测的泛化性和鲁棒性,更直观地呈现识别效果,将原始 YOLOv5s 模型和 TN—YOLOv5s 模型在 384 张测试集上进行对比试验,随机各选择顺逆光、不同时段和相互遮挡 3 种情况下不同标签对应的 1 张茶叶图像进行展示。原始 YOLOv5s 和 TN—YOLOv5s 模型在顺光和逆光条件的茶叶嫩芽检测效果如图 12 所示。

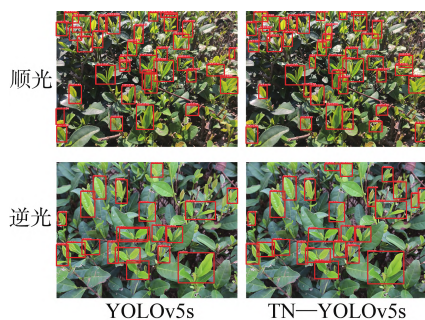


图 12 顺逆光场景 YOLOv5s 改进前后茶叶嫩芽检测效果对比

Fig. 12 Comparison of tea buds detection effect before and after YOLOv5s improvement in smooth and backlight scenes

原始 YOLOv5s 和 TN—YOLOv5s 模型对一天内不同时段的茶叶嫩芽检测效果如图 13 所示,分别为早晨、上午、下午、傍晚时间段下的茶叶嫩芽检测效果。

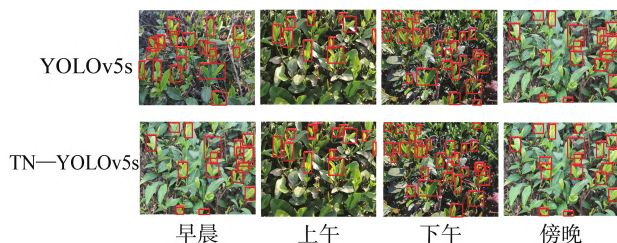


图 13 一天中不同时段场景 YOLOv5s 改进前后茶叶嫩芽检测效果对比

Fig. 13 Comparison of tea buds detection effect of YOLOv5s before and after improvement in different time of day scenarios

TN—YOLOv5s 模型对茶叶嫩芽间相互遮挡导致漏检的改善效果如图 14 所示。

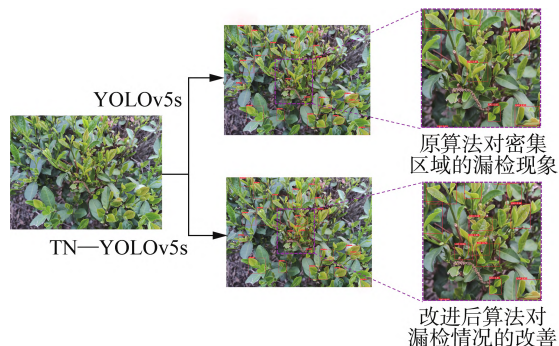


图 14 遮挡情况下茶叶嫩芽漏检情况改善效果图

Fig. 14 Improvement effect of tea buds leakage detection in the case of shading

由图 14 可知,由于茶叶嫩芽目标较小,处于的场景较为复杂,原始 YOLOv5s 模型对于不同场景下的茶叶嫩芽出现漏检现象,且检测出的目标置信度略有降低。TN—YOLOv5s 模型对茶叶图像进行检测时具有更高的置信度分数,且出现的茶叶嫩芽目标全部被识别出。

5 结论

为提高田间自然场景下夏茶茶叶嫩芽检测的实时性和准确性要求,改进 YOLOv5s 检测模型。

1) 在原始 YOLOv5s 基础上引入 GhostNet 网络结构,增加 CA 空间注意力机制,替换 SIOU 损失函数,替换 Soft—NMS 非极大抑制算法,得到优化后的 TN—YOLOv5s 模型。结果表明,改进后 TN—YOLOv5s 模型的精确率、召回率和平均精度均值分别提高 7.1%、5.9%、6.4%,模型参数量降低 47.17%,相比于原模型在保持轻量化的同时提高检测精度。

2) 将 TN—YOLOv5s 分别与 Faster—RCNN、SSD、YOLOv3、YOLOv4、YOLOv4—tiny、YOLOv5s、YOLOv8s 模型进行试验对比。结果表明,改进模型的平均精度比其他 7 种算法分别高出 31.53%、31.12%、20.55%、16.41%、24.08%、6.4%、3.5%,且模型权重大小只有 7.48 MB;虽然检测速度低于 YOLOv4—tiny、YOLOv5s 和 YOLOv8s 模型,但仍能满足对茶叶嫩芽实时检测的要求。

3) 验证结果表明,TN—YOLOv5s 模型在顺逆光、一天中不同时间段和遮挡情况下均表现出较好的识别效果,并明显改善原始算法因茶叶嫩芽间相互遮挡产生的漏检情况。说明基于 TN—YOLOv5s 的夏茶茶叶嫩芽检测算法具有较好的田间鲁棒性。

参 考 文 献

- [1] 宋志禹,韩余,丁文芹,等. 茶园机械研究“十三五”进展及“十四五”发展方向[J]. 中国茶叶, 2021, 43(10): 26—33.
Song Zhiyu, Han Yu, Ding Wenqing, et al. Tea garden machinery research progress during the 13th Five-Year plan period and development direction in the 14th Five-Year plan period [J]. China Tea, 2021, 43(10): 26—33.
- [2] 王宏亮. 夏秋茶多酚提取分离、纯化及淀粉样纤维—EGCG 水凝胶安全性评价[D]. 南京: 南京农业大学, 2020.
- [3] 李亚涛. 茶叶采摘机器人的视觉检测与定位技术研究[D]. 杭州: 浙江理工大学, 2022.
- [4] 郑航,傅童,薛向磊,等. 茶叶机械化采摘技术研究现状与展望[J]. 中国农机化学报, 2023, 44(9): 28—35.
Zheng Hang, Fu Tong, Xue Xianglei, et al. Research

- status and prospect of tea mechanized picking technology [J]. Journal of Chinese Agricultural Mechanization, 2023, 44(9): 28–35.
- [5] 罗泽涌, 陈建, 方晶晶, 等. 我国丘陵山区茶园种植机械化现状与发展研究[J]. 农机化研究, 2020, 42(2): 1–7.
- [6] 汪建. 结合颜色和区域生长的茶叶图像分割算法研究[J]. 茶叶科学, 2011, 31(1): 72–77.
- Wang Jian. Segmentation algorithm of tea combined with the color and region growing [J]. Journal of Tea Science, 2011, 31(1): 72–77.
- [7] 吴雪梅, 张富贵, 吕敬堂. 基于图像颜色信息的茶叶嫩叶识别方法研究[J]. 茶叶科学, 2013, 33(6): 584–589.
- [8] 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.
- [9] Wang T, Zhang K, Zhang W, et al. Tea picking point detection and location based on Mask-RCNN [J]. Information Processing in Agriculture, 2023: 10(2): 267–275.
- [10] Li Y, He L, Jia J, et al. High-efficiency tea shoot detection method via a compressed deep learning model [J]. International Journal of Agricultural and Biological Engineering, 2022, 15(3): 159–166.
- [11] 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.
- Wang Mengni, Gu Jinan, Wang Huajia, et al. Method for identifying tea buds based on improved YOLOv5s model [J]. Transactions of the Chinese Society of Agricultural Engineering, 2023, 39(12): 150–157.
- [12] 寇雷雷, 张红娜. 基于改进YOLOv5的苹果采摘机器人多目标识别技术研究[J]. 中国农机化学报, 2023, 44(11): 162–168.
- Kou Leilei, Zhang Hongna. Research on multi-target recognition technology of apple picking robot based on improved YOLOv5 [J]. Journal of Chinese Agricultural Mechanization, 2023, 44(11): 162–168.
- [13] Gui Z, Chen J, Li Y, et al. A lightweight tea bud detection model based on YOLOv5 [J]. Computers and Electronics in Agriculture, 2023, 205: 107636.
- [14] 尚钰莹, 张倩如, 宋怀波. 基于YOLOv5s的深度学习在自然场景苹果花朵检测中的应用[J]. 农业工程学报, 2022, 38(9): 222–229.
- Shang Yuying, Zhang Qianru, Song Huaibo. Application of deep learning using YOLOv5s to apple flower detection in natural scenes [J]. Transactions of the Chinese Society of Agricultural Engineering, 2022, 38(9): 222–229.
- [15] Han K, Wang Y, Tian Q, et al. Ghostnet: More features from cheap operations [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 1580–1589.
- [16] 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.
- [17] 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.
- [18] Hou Q, Zhou D, Feng J, et al. Coordinate attention for efficient mobile network design [C]. 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2021: 13708–13717.
- [19] 李光明, 弓皓斌, 袁凯. 基于轻量化YOLOv5s的花椒簇检测研究[J]. 中国农机化学报, 2023, 44(4): 153–158.
- Li Guangming, Gong Haobin, Yuan Kai. Research on lightweight pepper cluster detection based on YOLOv5s [J]. Journal of Chinese Agricultural Mechanization, 2023, 44(4): 153–158.
- [20] Gevorgyan Z. SIOU loss: More powerful learning for bounding box regression [J]. arXiv preprint arXiv: 2205.12740, 2022.
- [21] Bodla N, Singh B, Chellappa R, et al. Soft-NMS improving object detection with one line of code [C]. Proceedings of the IEEE International Conference on Computer Vision, 2017: 5561–5569.
- [22] 郝鹏飞, 刘立群, 顾任远. YOLO-RD-Apple果园异源图像遮挡果实检测模型[J]. 图学学报, 2023, 44(3): 456–464.
- Hao Pengfei, Liu Liqun, Gu Renyuan. YOLO-RD-Apple orchard heterogenous image obscured fruit detection model [J]. Journal of Graphics, 2023, 44(3): 456–464.
- [23] 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.
- [24] Liu W, Anguelov D, Erhan D, et al. SSD: Single shot multibox detector [C]. Computer Vision—ECCV. Springer International Publishing, 2016: 21–37.
- [25] Redmon J, Farhadi A. YOLOv3: An incremental improvement [J]. arXiv preprint arXiv: 1804.02767, 2018.
- [26] Bochkovskiy A, Wang C Y, Liao H Y M. YOLOv4: Optimal speed and accuracy of object detection [J]. arXiv preprint arXiv: 2004.10934, 2020.
- [27] Jiang Z, Zhao L, Li S, et al. Real-time object detection method based on improved YOLOv4-tiny [J]. arXiv preprint arXiv: 2011.04244, 2020.
- [28] Terven J, Cordova-Esparza D. A comprehensive review of YOLO: From YOLOv1 to YOLOv8 and beyond [J]. arXiv preprint arXiv: 2304.00501, 2023.