

DOI: 10.13718/j.cnki.xdzk.2025.09.002

王微子, 马伟. 基于改进 YOLOv8n 的套袋柑橘识别 [J]. 西南大学学报(自然科学版), 2025, 47(9): 11-23.

# 基于改进 YOLOv8n 的套袋柑橘识别

王微子, 马伟

中国农业科学院都市农业研究所/国家成都农业科技中心, 成都 610213

**摘要:** 有效识别套袋柑橘目标是实现数字农业产量估测的重要一步。针对在果实重叠和枝叶遮挡环境下识别套袋柑橘较困难的情况, 基于 YOLOv8n 目标检测模型提出了 DH-YOLO(YOLOv8n-DySample-HWD)套袋柑橘目标检测模型。主要对 YOLOv8n 模型的网络结构做了以下改进: 首先, 使用 DySample 模块替换原模型中的上采样模块, 提升模型的准确率; 其次, 在保证模型的准确率和召回率的同时, 引入 HWD 轻量级下采样模块来减少模型内存, 与原始模型相比, 更方便部署。所提出的套袋柑橘目标检测模型在准确率和平均精度方面均提高了 0.5 个百分点, 模型内存减少了 0.2 MB, 具有更高的识别精度和更强的部署能力, 该模型可用于自然环境中套袋柑橘采收和产量预测工作。

**关键词:** 深度学习; 柑橘采收; 目标检测

**中图分类号:** TP391.4; S666

**文献标识码:** A

**文章编号:** 1673-9868(2025)09-0011-13

开放科学(资源服务)标识码(OSID):



## Recognition of Bagged Citrus Fruits Based on the Improved YOLOv8n

WANG Weizi, MA Wei

*Institute of Urban Agriculture, Chinese Academy of Agricultural Sciences/Chengdu National Agricultural Science and Technology Center, Chengdu 610213, China*

**Abstract:** Effective recognition of bagged citrus fruits is a vital step to achieve yield estimation of digital agriculture. In this paper, a bagged citrus fruit recognition model based on YOLOv8n framework named DH-YOLO (YOLOv8n-DySample-HWD) is proposed, in order to solve the problem of poor recognition of bagged citrus in citrus orchards with the disturbance of fruit overlapping, as well as branch and leaf blocking. The following improvements were made mainly to the YOLOv8n structure. Firstly, the DySample module was used to substitute the upsampling module in the original model to improve the precision of the model. Secondly, while ensuring the precision and recall rate of the model, the size of model memory was

收稿日期: 2024-12-12

基金项目: 国家成都农业科技中心地方财政专项资金项目(NASC2024TD04, NASC2022KR08); 四川省自然科学基金青年项目(2025ZNSFSC1123)。

作者简介: 王微子, 硕士研究生, 主要从事目标检测算法研究。

通信作者: 马伟, 博士, 研究员。

reduced by introducing a lightweight downsampling module named HWD, which made model easier to deploy compared to the original one. The proposed DH-YOLO model improved precision and  $mAP_{0.5:0.95}$  by 0.5%, and reduced the size of model memory by 0.2 MB. It has higher recognition precision and is easier to deploy. This model can be used for bagged citrus harvesting and yield estimation in natural environment.

**Key words:** deep learning; citrus harvesting; target detection

柑橘是世界上最受欢迎的水果之一,具有极高的经济价值。中国是世界上最大的柑橘生产国,主要种植区域在长江以南。受气候变化的影响,我国柑橘种植区域逐渐由长江中下游地区向西南地区转移,呈现出“西移内扩”的种植格局<sup>[1-4]</sup>。四川省是我国西南地区柑橘的重要产区,气候类型为亚热带季风气候,地形以丘陵为主,形成了独特的小气候条件,使得冬季阴湿寒冷。为避免损伤柑橘果实,降低生产成本,在冬季果实成熟前会使用白色单层纸袋对其进行套袋处理<sup>[5]</sup>。此外,为防止柑橘果实被鸟类啄食而感染虫害进而影响到柑橘产量,也需在果实成熟前进行套袋处理<sup>[6]</sup>。

传统的柑橘产量估测主要采用人工计数的方法,此方法费时费力,存在漏数或重复计数的情况,误差较为严重。近些年,机器视觉技术在农业生产领域得到广泛应用,从田间管理到采收销售等环节都进行了全面研究,极大促进了我国农业的发展<sup>[7-8]</sup>。机器视觉具有精度高、处理速度快等优势,因此利用机器视觉准确识别到套袋柑橘目标成为实现柑橘产量准确估测的关键<sup>[9-10]</sup>。虽然在绿色背景下,白色的套袋柑橘目标比未成熟的绿色柑橘果实更易识别,但受西南地区小气候影响,阴天较多,柑橘园内采光较差,果实重叠和枝叶遮挡的复杂环境仍然是影响套袋柑橘目标识别的主要问题。

基于机器视觉的图像处理技术和目标检测技术是目前在自然环境下进行目标识别研究的主流方法,目标的颜色和形状是成功检测的重要识别特征。利用图像分割方法可以有效地将目标与图片背景进行分离,达到识别目标的效果。程洪等<sup>[11]</sup>提出了一种基于图像处理和支持向量机的方法,并将此方法应用到青苹果产量估计的模型中;熊俊涛等<sup>[12]</sup>使用 K-means 聚类分割和优化的 Hough 变换圆拟合算法从背景中有效提取柑橘果实目标并应用到柑橘采摘领域中;庄家俊等<sup>[13]</sup>提出了一种改进的 GB 色差图方法并结合支持向量机算法来解决在果实粘连或重叠情况下识别柑橘果实的问题,为未成熟柑橘果实产量的估算提供了依据;白雪冰等<sup>[14]</sup>改进了用于柑橘缺陷样本图像分割的传统 GAC 模型算法,可有效识别柑橘的表面缺陷;邹小林<sup>[15]</sup>使用基于局部邻域二阶差分法的过渡阈值图像分割算法可以更快更好地提取皇帝柑的图像目标;Gong 等<sup>[16]</sup>开发了一种使用图像分割技术在自然环境中识别黄色柑橘的软件系统,该系统可用于柑橘果实产量的估算;Dorj 等<sup>[17]</sup>设计了一种利用图像处理技术识别成熟柑橘的计算机视觉算法,可应用到柑橘产量估算研究中。

传统的图像分割算法存在算法复杂、效率低等问题。近年来,深度学习算法发展迅速,在农业生产领域研究中也逐渐得到应用。现有的深度学习目标检测算法主要分为 2 种类型,分别是双通道检测和单通道检测。双通道检测应用最广泛的算法为 R-CNN 系列算法,如 Huang 等<sup>[18]</sup>利用改进的 R-CNN 算法提高了分割葡萄簇的能力,有利于果园管理者对葡萄生产进行管理。单通道检测主要有 YOLO 系列算法和 SSD 算法<sup>[19]</sup>,YOLO 系列算法相较于 SSD 算法具有更快的检测速度和更高的识别精度,被诸多学者应用到农业生产中。杨军奇等<sup>[20]</sup>基于 YOLOv4 目标检测算法提出了改进的目标检测模型,提高了在田间检测葡萄叶片目标的平均精度;涂智荣等<sup>[21]</sup>提出了一种基于 YOLOv7-Tiny 模型改进的目标检测模型,该模型更加轻便,提高了识别果园中百香果目标的能力;杨健等<sup>[22]</sup>修改了 YOLOv5 算法的骨干模型结构和损失函数,有效地提高了目标检测模型在温室环境中识别番茄的准确率;黄小玉<sup>[23]</sup>采用 YOLOv3 算法对绿色柑橘进行了识别研究,有助于构建早期柑橘估产模型;伍锡如等<sup>[24]</sup>采用深度学习算法为水果采摘机器人设计了一个视觉识别系统,其识别准确率超过 97%;Wang 等<sup>[25]</sup>提出了 YOLO-CIT 模型,用于识别柑橘果实成熟度;Lin 等<sup>[26]</sup>以 Next-ViT 为基础结构,通过添加注意力创建了 AG-YOLO 网络模型,有效解决了识别成

熟黄色柑橘目标时漏检和多检的问题; Gu 等<sup>[27]</sup>提出了一种基于 YOLOv7-Tiny 的改进算法模型, 提高了在复杂环境下识别柑橘果实目标的能力。

柑橘生长会经历萌芽、开花、结果和成熟 4 个阶段。在开花和结果阶段以及果实成熟前的时间段内, 会对柑橘植株进行疏花疏果工作, 防止因过多的果实消耗植株内养分造成植株过劳乃至死亡。柑橘成熟期主要集中在 12 月至次年 3 月, 并以春节为界分为早熟柑橘品种和晚熟柑橘品种。绿色柑橘果实不易在绿色背景下识别并且无法判断是否对柑橘植株进行过疏果工作。而在套袋期间, 大部分柑橘植株已经完成疏果工作, 更有利于开展柑橘产量估测。因此, 构建柑橘估产模型的关键是如何提高在自然环境下识别套袋柑橘目标的能力, 并为未来轻量化部署创造有利条件。目前有关套袋柑橘目标识别检测的研究较少, 吕佳等<sup>[28]</sup>提出了基于师生模型的 SPM-YOLOv5 算法用于识别套袋柑橘, 其平均精度均值为 77.4%, 为套袋柑橘识别研究提供了参考依据; Sapkota 等<sup>[29]</sup>将 YOLOv8 系列算法与 CNN 算法在不同果园条件下对 2 个不同数据集的实例进行分割比较, 结果表明 YOLOv8 算法具有较高的准确性和效率, 更有利于为模型轻量化部署创造条件。

通过上述理论研究和实际情况分析, 本研究在 YOLOv8n 目标检测模型基础上提出 DH-YOLO (YOLOv8n-DySample-HWD) 套袋柑橘目标检测模型, 用于识别自然环境中的套袋柑橘, 以期对柑橘产量估测研究提供新的参考依据。

## 1 材料与方法

### 1.1 数据来源

本研究的套袋柑橘数据集来源于四川省成都市蒲江县堰儿子村的一处柑橘园, 柑橘品种为“春见”。使用智能手机对套袋柑橘进行了不同角度的拍摄, 包括目标的正面、侧面、反面等, 所采集到的套袋柑橘目标如图 1 所示, 图片像素为  $3\,000 \times 4\,000$ , 总共采集到 500 张图片。采集时间为 2023 年 11 月, 天气以多云、阴天为主, 为西南地区秋冬季节正常天气条件。



图 1 套袋柑橘目标的数据组成

## 1.2 构建数据集

使用 X-AnyLabeling 标注工具对所采集的图片进行手动标注。为提高模型的训练效率,所采用的标注方式为沿套袋柑橘目标轮廓进行标注。为了更加精准地识别单棵果树上的套袋柑橘目标,减少背景中其他果树目标的干扰,主要标注对象以图片中距离摄像头最近的套袋柑橘为主,标注方式如图 2 所示。

随后,将标记好的套袋柑橘图像导出为 YOLO 训练所需的 txt 文件并进行分类,按 4:1 的比例分为训练集(train)和验证集(val)。为了进一步提高模型的识别精度以及验证数据集中图片数量对训练结果的影响,通过提高亮度、加噪声和增加对比度等方式,将训练集扩展到 2 000 张。原始图像对应的 txt 文件也进行了修改,以满足实验要求。图像增强效果如图 3 所示。



图 2 X-AnyLabeling 标注方式



图 3 图像增强效果

## 1.3 套袋柑橘目标检测模型

### 1.3.1 DH-YOLO 模型结构

本研究基于 YOLOv8 系列目标检测算法提出了一种专门识别套袋柑橘目标的检测模型。YOLOv8 系列算法是一种单通道目标检测算法,具有精度高、检测速度快和可扩展性强等优点,根据模型内存大小划分为 5 个等级,由小到大依次为 n、s、m、l、x<sup>[30]</sup>。为了使所修改的模型能投入到实际农业生产中,便于轻量化部署,本研究选择以 YOLOv8n 目标检测模型为基础,进行了相应的改进<sup>[31]</sup>。

YOLOv8 系列算法由主干网络(Backbone)、颈部网络(Neck)和头部网络(Head)3 部分组成。与结构相似的 YOLOv5 目标检测算法相比,YOLOv8 算法的结构用 C2f 模块代替了 YOLOv5 算法结构中的 C3 模块,这种改动使模型具有更丰富的梯度流,可以有效调整不同模型中的通道数量,提高了模型的检测速度;在头部网络中,YOLOv8 算法使用解耦的头部结构,即无锚算法,将 YOLOv5 中的 Anchor-Based 替换为 Anchor-Free,可灵活适应不同目标的形状和大小<sup>[32]</sup>。因此,基于 YOLOv8 算法处理速度快、准确度高优点以及该系列算法中 YOLOv8n 模型的轻量化结构,利于开展套袋柑橘目标识别的研究。

为了提高模型在果实重叠、枝叶覆盖等复杂条件下识别套袋柑橘目标的能力,本研究提出了一种目标定位更准确、识别精度更高的检测模型 DH-YOLO,且与原模型相比更有利于轻量化部署。该模型主要对 YOLOv8n 中的颈部网络结构进行了上下采样模块的替换,所替换的模块在训练过程中可有效学习到复杂

背景中的目标, 从而增强模型识别套袋柑橘的能力, 模型结构如图 4 所示。首先, 特征图在通过模型的 C2f 模块和 SPPF 层后, 会经过侧重于点采样的 DySample 模块, 在样本充足的情况下, 模型的准确率得到了有效提高; 其次, 在执行上采样任务时, 一些特征图被推进到下采样模块中, 将下采样模块替换为 HWD 模块后, 可以更加高效地学习关于特征图的大部分信息, 提高了计算效率, 并减少了模型内存。2 个模块协同工作, 在提高复杂背景下套袋柑橘目标识别精度的同时降低了模型内存大小, 使改进后的模型比原始模型更有利于投入到实际农业生产中。

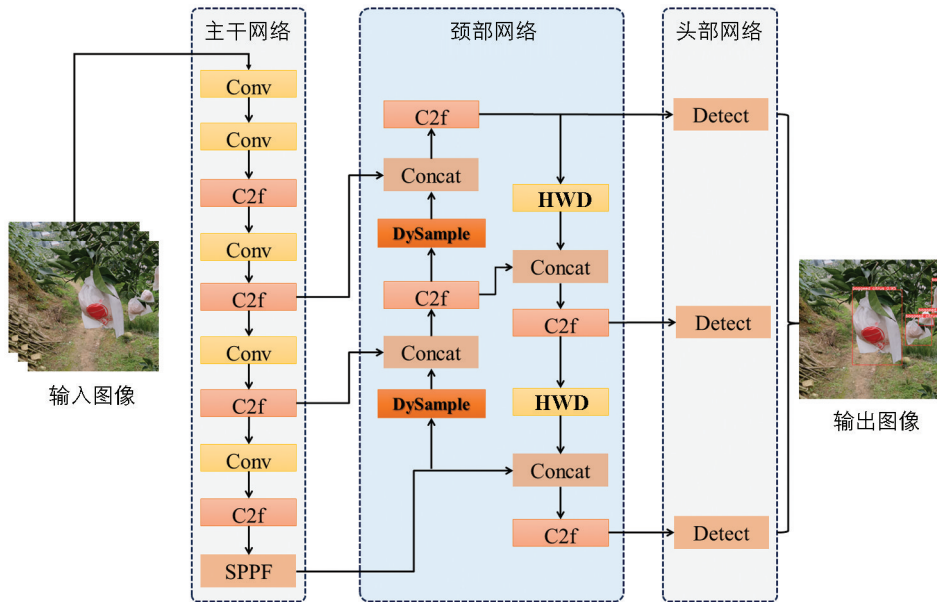


图 4 DH-YOLO 模型结构

### 1.3.2 DySample 动态采样模块

网络模型在训练时, 输入图像通过卷积神经网络后图像变大, 从小分辨率变为大分辨率的过程称为上采样。传统的上采样模块是基于核处理, 在训练过程中需要消耗大量内存并且增加训练时间。为了提高上采样的工作效率, 出现了动态采样模块。本研究使用的 DySample 模块与其他动态采样模块相比, 具有重量较轻、工作效率高、使用范围广等优点<sup>[33]</sup>。该模块的重点是点采样发生器, 此发生器有 2 种采样点生成方法, 分别是静态范围因子(Static Scope Factor)和动态范围因子(Dynamic Scope Factor)。

首先, 静态范围因子使用线性处理和原始采样网格(Original Sampling Grid)获得偏移量。图 5 展示了静态范围因子的结构, 式(1)–式(3)展示了当特征图(Feature Map) $X$  通过静态范围后的计算过程。

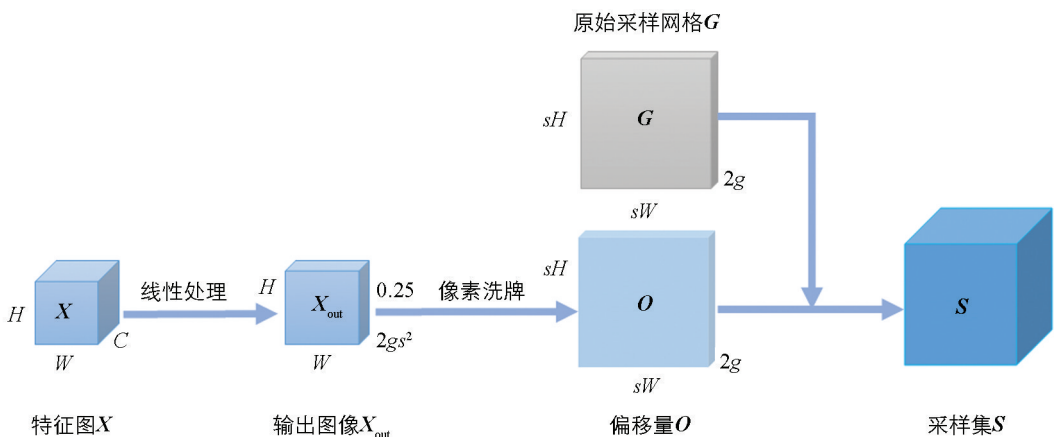


图 5 静态范围因子结构

在式(1)中, 当一个尺寸为  $H \times W \times C$  的特征图  $\mathbf{X}$  输入后, 通过线性处理操作得到  $2s^2$  的偏移, 再通过转换特征图的通道数后得到通道数量为  $2gs^2$  的  $\mathbf{X}_{out}$ 。

$$\mathbf{X}_{out} = \text{linear}(\mathbf{X}) \quad (1)$$

随后, 通过像素洗牌(Pixel Shuffle)技术得到尺寸为  $sH \times sW \times 2g$  的偏移量  $\mathbf{O}$ 。在原始的上采样模块中会出现采样点重叠的情况, 这对预测结果将产生较为负面的影响。为了解决这个问题, 该模块将  $\mathbf{X}_{out}$  乘以 0.25, 对采样位置做了一定约束, 更好地满足重叠和非重叠的理论边界条件, 计算过程如式(2)所示:

$$\mathbf{O} = 0.25 \times \text{linear}(\mathbf{X}) \quad (2)$$

最后, 将偏移量  $\mathbf{O}$  与尺寸为  $sH \times sW \times 2g$  的原始采样网格  $\mathbf{G}$  相结合, 得到采样集  $\mathbf{S}$ , 计算过程如式(3)所示:

$$\mathbf{S} = \mathbf{O} + \mathbf{G} \quad (3)$$

动态范围因子与静态范围因子结构类似, 但在特征图处理和计算上有些许不同。图 6 展示了动态范围因子的结构, 式(4)展示了当特征图  $\mathbf{X}$  输入后得到偏移量  $\mathbf{O}$  的计算过程。

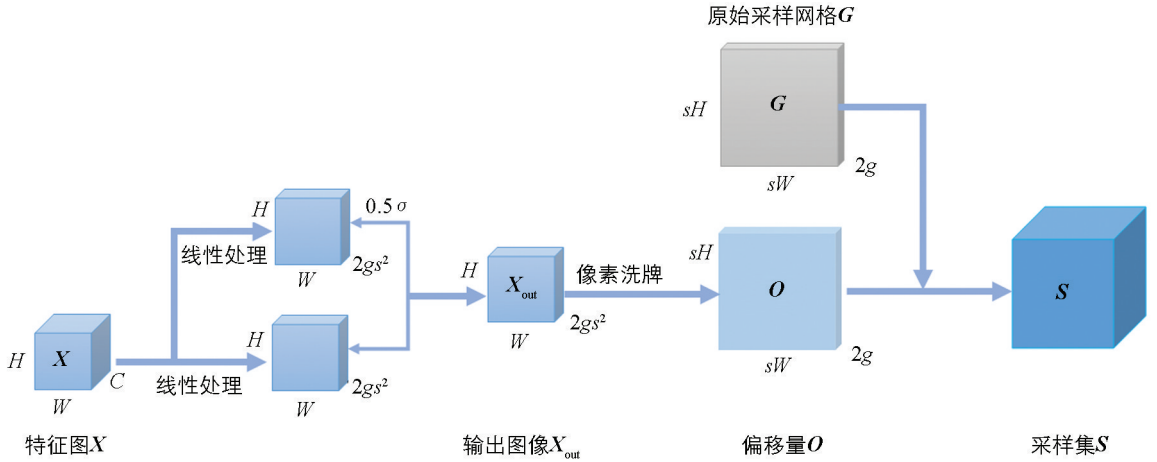


图 6 动态范围因子结构

当特征图  $\mathbf{X}$  输入后被分成 2 部分, 将这 2 部分分别进行线性处理, 其中一部分使用 sigmoid 函数进行线性处理并乘以 0.5, 而另一部分则与静态范围处理相同。将这 2 部分处理结合到一起得到偏移量  $\mathbf{O}$ , 最后与原始采样网格相结合获得最终采样集  $\mathbf{S}$ 。

$$\mathbf{O} = 0.5 \times \text{sigmoid}(\text{linear}_1(\mathbf{X})) \times \text{linear}_2(\mathbf{X}) \quad (4)$$

DySample 上采样模块增强了点采样的抗干扰能力, 与其他动态上采样模块相比, 它减少了参数算法, 提高了识别套袋柑橘目标的准确性<sup>[34]</sup>。

### 1.3.3 基于 Haar 的 HWD 模块

与上采样相反, 下采样是在完成上采样任务后降低图像分辨率。在进行下采样任务时, 图片像素减小容易丢失重要的信息, 为了避免在套袋柑橘识别的训练阶段因像素不一致而导致重要目标丢失的问题, 引入了一个基于 Haar 小波的 HWD 下采样模块<sup>[35]</sup>, 图 7 展示了该模块的结构。HWD 的核心思想是在转换图像分辨率时使用 Haar 小波来保留尽可能多的信息, 该模块由 2 个子模块组成: 无损特征编码模块(Lossless Feature Encoding)和特征学习模块(Feature Learning)。

在无损特征编码模块中, 使用 Haar 小波变换可以有效降低特征图的分辨率, 并保留大量信息。这是因

为 Haar 小波中的滤波器可用于从维度为  $W \times H$ 、通道数为  $C$  的图片中提取近似和低频信息。使用 Haar 小波后, 特征图像被处理成 4 个部分: 近似分量以及在水平、垂直和对角 3 个方向上的细节分量<sup>[35]</sup>。每个部分的尺寸为  $\frac{W}{2} \times \frac{H}{2}$ , 通道数将变为原来的 4 倍。

特征学习模块由 1 个大小为  $1 \times 1$  的标准卷积、1 个批处理归一化层和 1 个 ReLU 激活函数组成。标准卷积层可以调整特征图的通道数量, 允许在下采样过程中将处理后的图像与其他通道进行拼接。此功能能够删除冗余信息, 从而促进对识别目标的学习。最后采用批处理归一化和 ReLU 激活层来提取判别特征。

HWD 模块由 2 个不同功能的因子串联而成, 特征图在通过该模块时, 首先保留所有的图像信息, 同时通过无损特征编码模块降低图像分辨率, 最后通过特征学习模块去除冗余信息。与原下采样模块相比, 降低了目标的漏检率并且减少了一定的计算量, 缩减了模型的大小。

## 2 实验

### 2.1 实验环境和参数设置

该实验基于 Windows 11 系列操作系统, CPU 为 Gen Intel Core i5-13400F 2.50 GHz, GPU 为 NVIDIA GeForce RTX 4060 8 GB, Python 版本为 3.11.7, PyTorch 版本为 2.2.2, CUDA 版本为 12.1。所参与实验的训练模型均达到饱和迭代次数。在训练过程中, 输入图像都被调整为  $640 \times 640$  像素大小。

### 2.2 模型评估指标

为了验证模型的有效性, 需要对模型进行评估。深度学习网络训练时常见的评估指标有准确率( $P$ )、召回率( $R$ )和平均精度( $mAP$ )。 $mAP$  分为 2 种类型, 即  $mAP_{0.5}$  和  $mAP_{0.5:0.95}$ 。 $mAP_{0.5}$  表示 IoU 阈值为 0.5 时的平均精度,  $mAP_{0.5:0.95}$  表示从 0.5 到 0.95 范围内的 IoU 阈值的平均精度<sup>[26]</sup>。结合上述 4 个指标可以较好地评估模型的准确性和鲁棒性, 所对应的公式如式(5)~式(8)所示。

$$P = \frac{TP}{TP + FP} \times 100\% \quad (5)$$

$$R = \frac{TP}{TP + FN} \times 100\% \quad (6)$$

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

$$mAP = \frac{\sum_{i=1}^n AP_i}{n} \quad (8)$$

其中:  $TP$  为正确识别的套袋柑橘目标数量;  $FP$  为错误识别的套袋柑橘目标数量;  $FN$  为未识别到的套袋柑橘目标数量;  $n$  为类别总数。

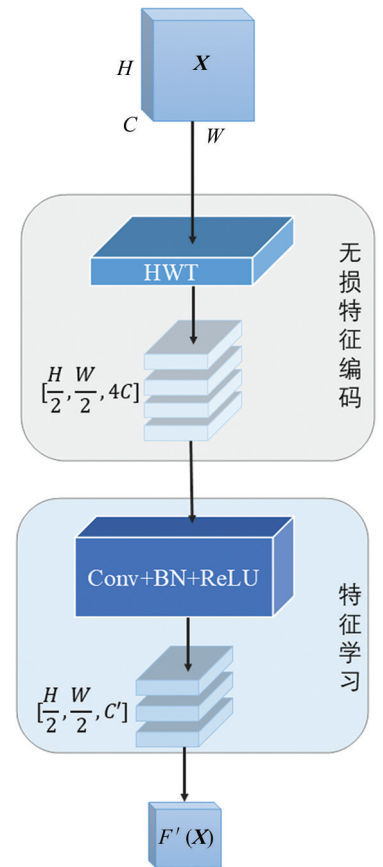


图 7 HWD 模块结构

## 2.3 实验结果

### 2.3.1 不同采样模块的对比实验

为了验证该模型所使用的上采样模块和下采样模块的优越性,参考相关文献后,选取了一些性能较好的采样模块,如 CARAFE<sup>[36]</sup>、ASF<sup>[37]</sup>、HWD<sup>[35]</sup>、ContextGuided<sup>[38]</sup>、ADown 等,并将模块按照类别进行组合。实验结果如表 1 所示,训练中损失值与轮次之间的变化曲线如图 8 所示。

表 1 不同采样模块训练结果对比

%

模型名称	模块类型	$P$	$R$	$mAP_{0.5}$	$mAP_{0.5:0.95}$
YOLOv8n	无	83.6	77.5	85.4	63.6
YOLOv8n-DySample	上采样	79.8	81.0	86.6	62.6
YOLOv8n-CARAFE	上采样	80.1	78.4	85.4	61.9
YOLOv8n-ASF	下采样	84.8	73.5	84.8	60.8
YOLOv8n-HWD	下采样	86.0	74.9	85.2	61.7
YOLOv8n-ADown	下采样	84.6	76.3	86.1	63.6
YOLOv8n-ContextGuided	下采样	83.2	72.6	84.5	59.4
YOLOv8n-DySample-ContextGuided	组合	80.8	76.3	83.5	59.1
DH-YOLO	组合	83.0	79.0	86.6	62.7
YOLOv8n-DySample-ADown	组合	84.8	74.1	85.4	62.9

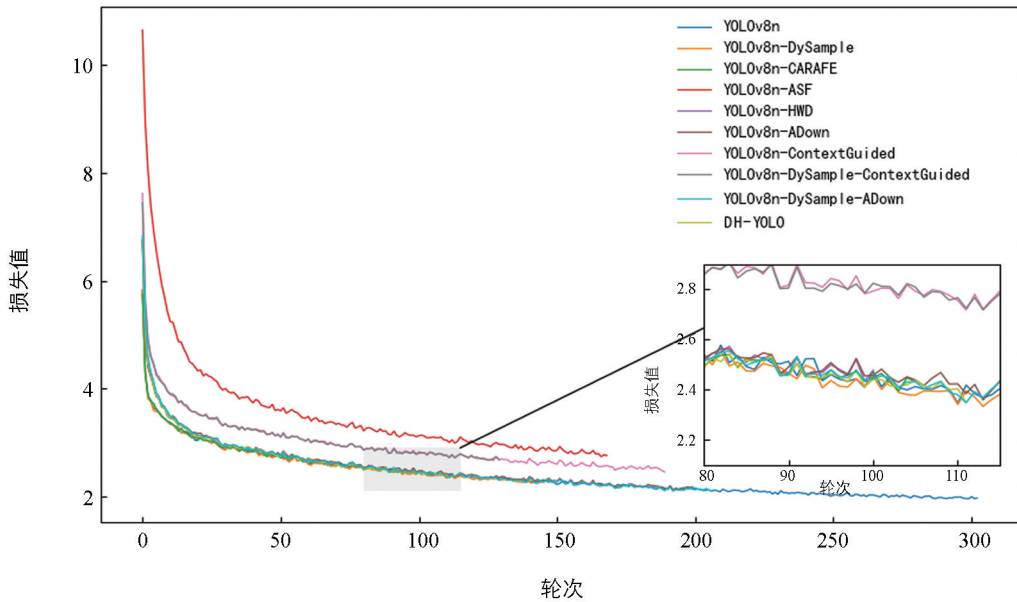


图 8 各模型的损失曲线

由表 1 可知,有 3 种不同模块类型的模型:只替换上采样模块、只替换下采样模块和同时替换上、下采样模块。在只替换上采样模块的模型中,尽管 YOLOv8n-DySample 模型的  $P$  低于 YOLOv8n-CARAFE 模型和 YOLOv8n 模型,但它的  $R$  和  $mAP_{0.5}$  明显高于这 2 种模型,表明引入 DySample 采样器后,降低了目标检测的漏检率,能够更精确地识别到目标。在只替换下采样模块的模型中,YOLOv8n-HWD 模型的  $P$  高于其他模型,但  $R$  相对较低,表明模型识别到的目标有很高的可信度。在同时替换上、下采样模块的模型中,表现最佳的模型是 DH-YOLO,尽管模型的  $P$  相较于 YOLOv8n 模型降低了 0.6 个百分点,

$mAP_{0.5:0.95}$  降低了 0.9 个百分点, 但  $R$  提高了 1.5 个百分点,  $mAP_{0.5}$  提高了 1.2 个百分点。虽然 YOLOv8n-DySample-ADown 模型的  $P$  相较于 YOLOv8n 模型提高了 1.2 个百分点, 但是  $R$  降低了 3.4 个百分点,  $mAP_{0.5:0.95}$  降低了 0.7 个百分点, 表明模型的  $P$  主要依靠降低  $R$  来提升, 在执行检测任务时会产生大量的漏检目标。

从图 8 展示的损失曲线可以看出, 由于所有模型训练均达到饱和, 所有改进模型的训练轮次要远小于 YOLOv8n 模型, 表明改进模型的学习效率均高于 YOLOv8n 模型; 但 YOLOv8n-ASF、YOLOv8n-ContextGuided 和 YOLOv8n-DySample-ContextGuided 3 个模型的损失值下降要远小于其他模型, 表明这 3 个模型虽然学习效率高, 但是模型所学习到的信息要少于其他模型; YOLOv8n-HWD、DH-YOLO 和 YOLOv8n-DySample 最先达到饱和状态, 饱和轮次分别为 101、112 和 145, 并且损失值下降稳定; DH-YOLO 模型的损失曲线波动最为平缓, 可以表明引入 DySample 模块和 HWD 模块后, 提升了模型的训练质量。

基于上述分析, 为了进一步提高模型识别套袋柑橘目标的能力, 将 YOLOv8n-DySample、YOLOv8n-ADown、YOLOv8n-HWD、YOLOv8n-DySample-ADown 和 DH-YOLO 5 种模型在扩充数据集上进行增强实验, 并通过对比分析验证 DySample 模块和 HWD 模块的优越性。

### 2.3.2 扩大数据集对检测性能的影响

深度学习需要大量的图像数据, 因此将数据集扩展到 2 000 张图像, 训练集与验证集的比例依旧保持为 4 : 1, 实验结果如表 2 所示。

表 2 数据集训练结果对比

数据集图片 数量/张	模型名称	模型内存/ MB	$P$ / %	$R$ / %	$mAP_{0.5}$ / %	$mAP_{0.5:0.95}$ / %
500	YOLOv8n	6.3	83.6	77.5	85.4	63.6
	YOLOv8n-DySample	6.3	79.8	81.0	86.6	62.6
	YOLOv8n-ADown	6.0	84.6	76.3	86.1	63.6
	YOLOv8n-HWD	6.0	86.0	74.9	85.2	61.7
	YOLOv8n-DySample-ADown	6.0	84.8	74.1	85.4	62.9
	DH-YOLO	6.1	83.0	79.0	86.6	62.7
2 000	YOLOv8n	6.3	98.8	99.0	99.5	92.9
	YOLOv8n-DySample	6.3	99.3	99.1	99.5	93.8
	YOLOv8n-ADown	6.1	99.1	99.0	99.5	92.8
	YOLOv8n-HWD	6.0	98.6	98.8	99.4	92.9
	YOLOv8n-DySample-ADown	6.1	98.8	99.4	99.4	93.5
	DH-YOLO	6.1	99.3	99.1	99.5	93.4

由表 2 可知, 在数据增强的协助下, 所有模型的  $P$ 、 $R$  和  $mAP$  都有显著提高。虽然 YOLOv8n-DySample 模型具有最高的  $P$  和  $mAP$ , 但模型内存没有明显变化; 而 DH-YOLO 模型的  $P$ 、 $R$ 、 $mAP_{0.5}$  与 YOLOv8n-DySample 模型相同, 仅在  $mAP_{0.5:0.95}$  方面比其少 0.4 个百分点, 但 DH-YOLO 模型的内存小于 YOLOv8n-DySample 模型。尽管 YOLOv8n-DySample-ADown 模型的内存与 DH-YOLO 模型相同, 但

它的  $P$  不及 DH-YOLO 模型,即在识别套袋柑橘目标的精度上 DH-YOLO 模型更具有优势。对评估指标进行综合分析后可知, DH-YOLO 模型在提升  $P$ 、 $R$  和  $mAP_{0.5:0.95}$  的同时有效减少了模型内存,拥有最高的  $P$  和相对较小的模型容量,与原模型相比具有更强的优势。

### 2.3.3 不同 YOLO 模型的对比实验

为了更客观地展示 DH-YOLO 模型在检测套袋柑橘方面的优势,从识别精度和轻量化部署的综合角度出发,选取了 YOLOv3、YOLOv5 以及 YOLOv6 系列算法中最轻量化的模型,这 3 种模型具有较小的内存和较高的识别能力,并且已被广泛应用于各个领域中。将 DH-YOLO 模型与未经改进的 YOLOv3-Tiny、YOLOv5n、YOLOv6n 和 YOLOv8n 进行了比较,主要目的是验证模型在枝叶遮挡以及果实重叠的情况下辨别套袋柑橘目标的能力。上述模型在验证集上的识别能力对比如表 3 所示。

表 3 不同 YOLO 模型的实验结果对比

模型名称	模型内存/MB	$P/\%$	$R/\%$	$mAP_{0.5}/\%$	$mAP_{0.5:0.95}/\%$
YOLOv3-Tiny	24.4	98.8	97.8	98.8	90.1
YOLOv5n	5.3	97.9	97.0	92.2	86.8
YOLOv6n	8.8	98.9	98.9	99.5	91.5
YOLOv8n	6.3	98.8	99.0	99.5	92.9
DH-YOLO	6.1	99.3	99.1	99.5	93.5

与未经改进的 YOLOv3-Tiny、YOLOv5n、YOLOv6n 和 YOLOv8n 模型相比, DH-YOLO 模型在整个数据集上识别套袋柑橘目标方面表现出更优越的综合性能,进一步证实了 DH-YOLO 模型在复杂环境中识别套袋柑橘的可行性。

如图 9 所示, DH-YOLO 模型实际检测效果明显优于其他模型。黄色椭圆形区域表示模型在执行识别套袋柑橘任务时出现了多个目标识别错误,这类情况在 YOLOv5n、YOLOv6n 和 YOLOv8n 模型中较为常见;深蓝色矩形区域表示模型在执行检测任务时将背景识别成目标的情况,主要集中在 YOLOv3-Tiny、YOLOv5n、YOLOv6n 和 YOLOv8n 模型中;深红色矩形区域表示在复杂背景下检测套袋柑橘目标时模型未识别目标的情况,这类情况易存在于 YOLOv6n 和 YOLOv8n 模型中。除 DH-YOLO 模型外,所有模型都难以在复杂背景下准确识别套袋柑橘。从轻量化部署的角度来看,虽然 YOLOv5n 模型具有最小的内存,但在识别效果上不如 DH-YOLO 模型。PR 曲线可以更加直观地衡量模型识别目标准确率和召回率之间的关系,在本次实验中,曲线越接近右上角表明模型的性能越高。5 种模型在验证集上的 PR 曲线如图 10 所示,从图中也可以看出 DH-YOLO 模型具有最好的性能。

## 3 结论

本研究旨在解决果实重叠、枝叶遮挡等复杂自然环境下,提高模型识别套袋柑橘目标的能力,并为轻量化部署创造有利条件。主要开展了以下研究工作:1) 柑橘园中拍摄套袋柑橘影像,使用 X-AnyLabeling 标注软件构建了套袋柑橘数据集;2) 以 YOLOv8n 目标检测模型为基础,通过改进 YOLOv8n 模型结构,提出了一种专门识别套袋柑橘的目标检测模型 DH-YOLO,为柑橘采收和产量估算研究提供了有利依据。

YOLOv8 系列目标检测算法与 YOLO 系列的前代算法相比,具有更高的识别精度,但在复杂环境下识别套袋柑橘方面存在较大不足。在本研究中,通过将 DySample 模块代替原始的上采样模块,提高了在



a. YOLOv3-Tiny



b. YOLOv5n



c. YOLOv6n



d. YOLOv8n



e. DH-YOLO

图 9 不同模型的识别效果

复杂环境中识别套袋柑橘目标的准确率；由于上采样模块计算量较高，模块容量较大，而在执行下采样任务时，选择了基于 Haar 小波的下采样模块 HWD，可以在保证模型准确率的同时学习目标特征，减少了目标的漏检率，同时降低了模型容量，使模型能够更好地应用于实际农业生产。

本研究所提出的 DH-YOLO 套袋柑橘目标检测模型与 YOLOv8n 模型相比，虽然整体指标相差不大，但在实际应用过程中，由于背景环境中存在枝叶遮挡的问题造成 YOLOv8n 模型容易漏检，而 DH-YOLO 模型能够较好地解决此问题，并且模型内存减少了 0.2 MB，表明模型具有更强的部署能力且在实际应用中表现良好，为我国数字农业果实产量估测领域的研究提供了新的参考方案。

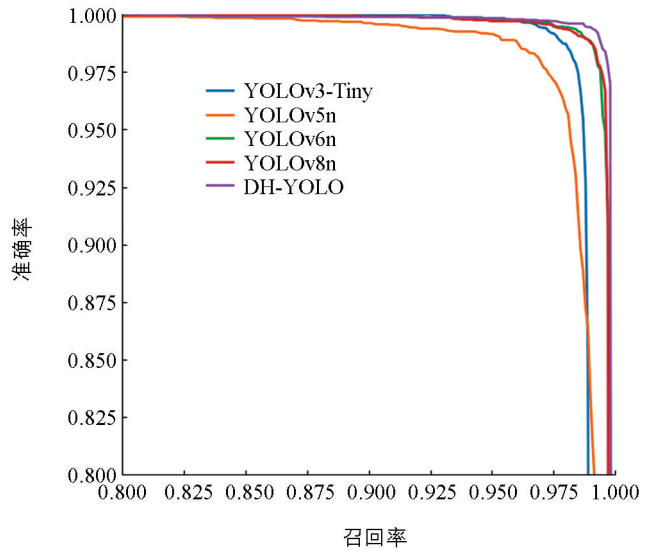


图 10 各模型在验证集上的 PR 曲线

#### 参考文献:

- [1] 黄伟华, 祁春节. 中国柑橘生产空间布局演化及驱动因素研究 [J]. 华中农业大学学报(社会科学版), 2022(4): 90-103.
- [2] 胡友, 陈昕, 祁春节. 中国柑橘生产布局变迁驱动机制研究——基于农户决策视角 [J]. 中国农业资源与区划, 2022, 43(5): 122-131.
- [3] 徐浩, 王玉雯, 罗自威, 等. 养分优化管理提高我国柑橘产量和品质的整合分析 [J]. 应用生态学报, 2024, 35(5): 1301-1311.
- [4] 林春桃, 王明新, 范水生, 等. 基于非期望产出的柑橘生产效率及其时空差异 [J]. 中国农业资源与区划, 2020, 41(12): 76-83.
- [5] 田福敏, 张绍彬, 罗涛, 等. 套袋对柑橘果面绿斑病及果实品质的影响 [J]. 贵州农业科学, 2018, 46(1): 33-35.
- [6] 石莹, 陈思怡, 曾译可, 等. 套袋提升椪柑果实品质的作用机制 [J]. 中国农业科学, 2023, 56(14): 2776-2786.
- [7] 温哈秋子, 郑永强, 刘杨. 柑橘大数据研究与应用 [J]. 农业大数据学报, 2021, 3(1): 33-44.
- [8] PENG K Q, MA W, LU J Z, et al. Application of Machine Vision Technology in Citrus Production [J]. Applied Sciences, 2023, 13(16): 9334.
- [9] HOSAINPOUR A, KHEIRALIPOUR K, NADIMI M, et al. Quality Assessment of Dried White Mulberry (*Morus alba* L.) Using Machine Vision [J]. Horticulturae, 2022, 8(11): 1011.
- [10] KHEIRALIPOUR K, NADIMI M, PALIWAL J. Development of an Intelligent Imaging System for Ripeness Determination of Wild Pistachios [J]. Sensors, 2022, 22(19): 7134.
- [11] 程洪, DAMEROW L, BLANKE M, et al. 基于图像处理与支持向量机的树上苹果早期估产研究 [J]. 农业机械学报, 2015, 46(3): 9-14, 22.
- [12] 熊俊涛, 邹湘军, 彭红星, 等. 扰动柑橘采摘的实时识别与采摘点确定技术 [J]. 农业机械学报, 2014, 45(8): 38-43.
- [13] 庄家俊, 唐宇, 郭琪伟, 等. 基于图像分割和决策层融合的未成熟柑橘检测方法 [J]. 仲恺农业工程学院学报, 2020, 33(4): 46-52.
- [14] 白雪冰, 林鑫, 武云鹏, 等. 基于改进 GAC 模型的柑橘缺陷图像分割方法研究 [J]. 西北林学院学报, 2019, 34(6): 152-159.
- [15] 邹小林. 基于过渡区的皇帝柑图像分割 [J]. 江苏农业科学, 2019, 47(14): 248-250.
- [16] GONG A P, YU J L, HE Y, et al. Citrus Yield Estimation Based on Images Processed by an Android Mobile Phone [J]. Bio-

- systems Engineering, 2013, 115(2): 162-170.
- [17] DORJ U O, LEE M, YUN S S. An Yield Estimation in Citrus Orchards via Fruit Detection and Counting Using Image Processing [J]. Computers and Electronics in Agriculture, 2017, 140: 103-112.
- [18] HUANG X, PENG D D, QI H N, et al. Detection and Instance Segmentation of Grape Clusters in Orchard Environments Using an Improved Mask R-CNN Model [J]. Agriculture, 2024, 14(6): 918.
- [19] KOIRALA A, WALSH K B, WANG Z, et al. Deep Learning for Real-Time Fruit Detection and Orchard Fruit Load Estimation: Benchmarking of ‘MangoYOLO’ [J]. Precision Agriculture, 2019, 20(6): 1107-1135.
- [20] 杨军奇, 冯全, 王书志, 等. 基于改进YOLOv4的田间密集小目标检测方法 [J]. 东北农业大学学报, 2022, 53(5): 69-79.
- [21] 涂智荣, 凌海英, 李帼, 等. 基于改进YOLOv7-Tiny的轻量化百香果检测方法 [J]. 广西师范大学学报(自然科学版), 2024, 42(5): 79-90.
- [22] 杨健, 杨啸治, 熊串, 等. 基于改进YOLOv5的番茄果实识别估产方法 [J]. 中国瓜菜, 2024, 37(6): 61-68.
- [23] 黄小玉. 基于机器视觉的近色背景绿色果实识别研究 [D]. 重庆: 西南大学, 2020.
- [24] 伍锡如, 雪刚刚, 刘英璇. 基于深度学习的水果采摘机器人视觉识别系统设计 [J]. 农机化研究, 2020, 42(2): 177-182, 188.
- [25] WANG C L, HAN Q Y, LI C J, et al. Fusion of Fruit Image Processing and Deep Learning: A Study on Identification of Citrus Ripeness Based on R-LBP Algorithm and YOLO-CIT Model [J]. Frontiers in Plant Science, 2024, 15: 1397816.
- [26] LIN Y S, HUANG Z F, LIANG Y, et al. AG-YOLO: A Rapid Citrus Fruit Detection Algorithm with Global Context Fusion [J]. Agriculture, 2024, 14(1): 114.
- [27] GU B, WEN C J, LIU X Z, et al. Improved YOLOv7-Tiny Complex Environment Citrus Detection Based on Lightweighting [J]. Agronomy, 2023, 13(11): 2667.
- [28] 吕佳, 李帅军, 曾梦瑶, 等. 基于半监督SPM-YOLOv5的套袋柑橘检测算法 [J]. 农业工程学报, 2022, 38(18): 204-211.
- [29] SAPKOTA R, AHMED D, KARKEE M. Comparing YOLOv8 and Mask R-CNN for Instance Segmentation in Complex Orchard Environments [J]. Artificial Intelligence in Agriculture, 2024, 13: 84-99.
- [30] 唐兴萍, 王白娟, 杨红欣, 等. 基于YOLOv8的水果外观检测与分类方法 [J]. 食品与机械, 2024, 40(7): 103-110.
- [31] 王金鹏, 何萌, 甄乾广, 等. 基于COF-YOLOv8n的油茶果静、动态检测计数 [J/OL]. 农业机械学报, (2024-01-17) [2024-12-12]. <https://link.cnki.net/urlid/11.1964.S.20240117.0910.002>.
- [32] 张新月, 胡广锐, 李浦航, 等. 基于改进YOLOv8n的轻量化红花识别方法 [J]. 农业工程学报, 2024, 40(13): 163-170.
- [33] LIU W Z, LU H, FU H T, et al. Learning to Upsample by Learning to Sample [C] //2023 IEEE/CVF International Conference on Computer Vision (ICCV), October 1-6, 2023, Paris, France. IEEE, 2023: 6004-6014.
- [34] 李洪, 石艳, 范桃. 改进YOLOv8n的O型密封圈表面缺陷检测算法研究 [J]. 计算机工程与应用, 2024, 60(18): 126-135.
- [35] XU G P, LIAO W T, ZHANG X, et al. Haar Wavelet Downsampling: A Simple but Effective Downsampling Module for Semantic Segmentation [J]. Pattern Recognition, 2023, 143: 109819.
- [36] WANG J Q, CHEN K, XU R, et al. CARAFE: Content-Aware ReAssembly of FEatures [C] //2019 IEEE/CVF International Conference on Computer Vision (ICCV), October 27-November 2, 2019, Seoul, Korea. IEEE, 2019: 3007-3016.
- [37] KANG M, TING C M, TING F F, et al. ASF-YOLO: A Novel YOLO Model with Attentional Scale Sequence Fusion for Cell Instance Segmentation [J]. Image and Vision Computing, 2024, 147: 105057.
- [38] WU T Y, TANG S, ZHANG R, et al. CGNet: A Light-Weight Context Guided Network for Semantic Segmentation [J]. IEEE Transactions on Image Processing, 2021, 30: 1169-1179.