

基于 SAW - YOLO v8n 的葡萄幼果轻量化检测方法

张传栋¹ 高鹏² 亓璐¹ 丁华立¹

(1. 济宁学院计算机科学与工程学院, 曲阜 273100; 2. 曲阜师范大学网络空间安全学院, 曲阜 273165)

摘要: 葡萄簇幼果实受背景色、遮挡和光照变化的影响,检测难度大。为了实现对背景色、遮挡和光照变化具有鲁棒性的葡萄簇幼果检测,提出了一种融合随机注意力机制(Shuffle attention, SA)的改进 YOLO v8n 模型(SAW - YOLO v8n)。通过在 YOLO v8n 模型的 Neck 结构中融入 SA 机制,增强网络多尺度特征融合能力,提升检测目标的特征信息表示,并抑制其他无关信息,提高检测网络检测精度,在不明显增加网络深度和内存开销的情况下,实现了葡萄簇幼果的高效准确检测;采用基于动态非单调聚焦机制的损失(Wise intersection over union loss, Wise - IoU Loss)作为边界框回归损失函数,加速网络收敛并进一步提高模型的准确率。构建了葡萄簇幼果的数据集 GGrape,该数据集由 3 780 幅复杂场景下的葡萄簇幼果图像及对应标注文件组成。通过该数据集对 SAW - YOLO v8n 模型进行训练和测试。测试结果表明,基于 SAW - YOLO v8n 的葡萄簇幼果检测算法的精度(Precision, P)、召回率(Recall, R)、平均精度均值(Mean average precision, mAP)和 F1 值分别为 92.80%、91.30%、96.10% 和 92.04%,检测速度为 140.85 f/s,模型内存占用量为 6.20 MB。与 SSD、YOLO v5s、YOLO v6n、YOLO v7 - tiny、YOLO v8n 等 5 个轻量化模型相比,其 mAP 值分别提高 16.06%、1.05%、1.48%、0.84%、0.73%,F1 值分别提高 24.85%、1.43%、1.43%、1.09%、1.60%,模型内存占用量分别降低 93.16%、56.94%、37.63%、47.00%、0,是所有模型中最小的,具有明显的轻量化、高精度优势。讨论了不同遮挡程度和光照条件的葡萄幼果检测,结果表明,基于 SAW - YOLO v8n 的葡萄幼果检测方法能适应不同遮挡和光照变化,具有良好的鲁棒性。结果表明,SAW - YOLO v8n 不仅能满足对葡萄簇幼果检测的高精度、高速度、轻量化的要求,且具有较强的鲁棒性和实时性。

关键词: 葡萄幼果; 疏果; 目标检测; shuffle attention; YOLO v8n; Wise - IoU Loss

中图分类号: TP391.4 **文献标识码:** A **文章编号:** 1000-1298(2024)10-0286-09

OSID:



Lightweight Detection Method for Young Grape Cluster Fruits Based on SAW - YOLO v8n

ZHANG Chuandong¹ GAO Peng² QI Lu¹ DING Huali¹

(1. School of Computer Science and Engineering, Jining University, Qufu 273100, China

2. School of Cyber Science and Engineering, Qufu Normal University, Qufu 273165, China)

Abstract: The detection of young grape cluster fruits is challenging due to the influence of background color, occlusion, and lighting variations. To achieve robust detection of young grape cluster fruits for the varying conditions, an improved YOLO v8n model that integrated shuffle attention (SA) mechanism was proposed in the work. By incorporating SA mechanism into the Neck network of the YOLO v8n model, the multi-scale feature fusion ability of the network was enhanced, the feature information representation of the detection target was improved, and other irrelevant information was suppressed, improving the accuracy of the detection network, which achieved efficient and accurate detection of young grape cluster fruits without significantly increasing network depth and memory overhead. Wise intersection over union loss (Wise - IoU Loss) with the dynamic nonmonotonic focusing mechanism was taken as the bounding box regression loss function, to accelerate the network convergence for the better detection accuracy of the model. Herein, a Grape dataset was constructed, which comprised 3 780 images of young grape cluster fruits in complex scenarios along with corresponding annotation files. Training and testing results of the SAW - YOLO v8n model on this dataset showed that the precision (P), recall (R), mean average

收稿日期: 2024 - 04 - 15 修回日期: 2024 - 07 - 11

基金项目: 中国博士后科学基金项目(2023M732022)和济宁市重点研发计划项目(2021ZDZP025)

作者简介: 张传栋(1978—),男,副教授,主要从事图像处理 and 模式识别研究, E-mail: jnzcd@jnxu.edu.cn

precision (mAP), and F1 score of the young grape cluster fruit detection algorithm based on SAW - YOLO v8n were 92.80%, 91.30%, 96.10%, and 92.04%, respectively, where the detection speed was 140.85 f/s, and the model size was only 6.20 MB. Compared with that of SSD, YOLO v5s, YOLO v6n, YOLO v7-tiny, and YOLO v8n, the mAP was increased by 16.06%, 1.05%, 1.48%, 0.84% and 0.73%, respectively, and F1 scores were increased by 24.85%, 1.43%, 1.43%, 1.09% and 1.60%, respectively, and the model weights were reduced by 93.16%, 56.94%, 37.63%, 47.00%, and 0, respectively, which was the smallest among all models and had obvious advantages in lightweight and high accuracy. Moreover, the young grape cluster fruits detection with different degrees of occlusion and lighting conditions were also explored, and the result showed that the young grape cluster fruit detection method based on SAW - YOLO v8n can adapt to different occlusion and lighting changes, and had good robustness. In summary, SAW - YOLO v8n not only met the requirements of high-precision, high-speed, and lightweight detection of young grape cluster fruits, but also had strong robustness and real-time performance.

Key words: young grape cluster fruits; fruit thinning; object detection; shuffle attention; YOLO v8n; Wise - IoU Loss

0 引言

复杂自然场景下幼果期葡萄簇的自动、快速、准确检测是实现智能疏果的重要前提,是葡萄果园自动化和智能化管理的重要基础,对促进智能疏果机器人发展、提高果品品质有重要意义。幼果期葡萄簇形状不规则、浆果微小、分布密集、与叶片颜色高度一致^[1],受自然光照和遮挡的影响,果实识别难度增大^[2-5],制约了检测精度的提高,是幼果期葡萄簇检测面临的重要挑战。

深度学习技术可以学习多尺度特征信息,具有良好的特征提取能力和泛化能力^[6],可以应对姿态、颜色、光照和遮挡等因素的变化^[7],被广泛应用于苹果^[8-9]、油茶果^[10]、樱桃^[11]、番茄^[12]等果实目标识别,并取得了良好的效果,为葡萄目标识别提供了必要借鉴。葡萄簇不同于上述水果,其由多个微小浆果组成,呈簇状密集分布,果实紧密粘连,其姿态、轮廓、尺寸不一,受光照及遮挡程度也不同,这增加了葡萄簇识别难度。不同葡萄品种及同一品种不同生长阶段的葡萄簇颜色有较大差异,针对果品颜色与叶片颜色一致问题研究人员做了相关研究^[13-15]。为提高田间葡萄图像中小目标葡萄检测的精度,李国进等^[16]基于 YOLO v3 算法提出了 WGDM 模型,在公开的酿酒葡萄图像数据集的测试集上平均精度为 81.20%,F1 值为 85.63%。为解决生长环境复杂、枝叶阴影、葡萄重叠等因素导致的无法识别或识别准确率下降的问题,LI 等^[17]提出了 YOLO grape 模型,对多种颜色的葡萄簇进行检测,绿葡萄簇检测的 F1 值为 90.93%,多种颜色葡萄簇 F1 值平均值为 91.42%。ZHANG 等^[18]用 YOLO v5 GAP 检测青葡萄簇,平均精度为 95.13%,可以准确检测生长密集和遮荫的葡萄簇。为了提高青葡萄簇检测精度,WEI

等^[19]用 S - MGDM 模型对成熟期的青葡萄簇进行检测,AP 值为 96.73%,F1 值为 91.0%,分别比原网络模型提高 3.87% 和 3%。孙俊等^[20]提出了一种改进的葡萄检测模型 MRWYOLO v5s,mAP 为 97.74%,较原模型提升 2.32%。综上,这些研究有效地提升了成熟葡萄簇识别效果。

目前,针对成熟期葡萄簇识别研究较多^[21],对幼果期葡萄识别研究较少,幼果期葡萄簇颜色与背景色相近,在光照和遮挡等干扰情况下无法实现高精度识别,识别效果常低于成熟葡萄簇^[15],制约了葡萄生长管理过程中幼果疏果自动化水平的提高。为了实现复杂场景下葡萄幼果精准、快速检测,本文提出一种融合卷积神经网络和注意力机制的葡萄簇幼果检测方法,在 YOLO v8n^[22]模型 Neck 结构中融合 SA^[23]注意力机制,增强网络多尺度特征融合能力,提升检测目标的特征信息表示,并抑制其他无关信息,以提高检测网络的精度,在不明显增加网络深度和内存开销的情况下,以实现葡萄簇幼果的高效准确检测;在 Head 结构中把 CIoU 替换成 Wise - IoU^[24],避免宽高比对模型的过度惩罚,同时平衡高质量和低质量锚框对于模型回归的影响,以增强模型泛化能力。

1 数据集构建

本研究使用的图像采集自山东省曲阜市,葡萄品种为巨峰。拍摄时间为 2022 年 5 月 22—23 日 07:20—10:30 和 14:30—18:00,天气晴朗。如图 1 所示,采集人员站在距离葡萄藤 60 ~ 80 cm 的位置,采用手持拍摄形式,使用 HUAWEI (YAL - AL00 \ BMH - AN10 \ JEF - AN00 型)、OPPO (Find X2 Pro \ Reno6 Pro 5G 型)、VIVO S7 型、OnePlus KB2000 型、Xiaomi M2012K10C 型等多种手机进行拍摄,共采

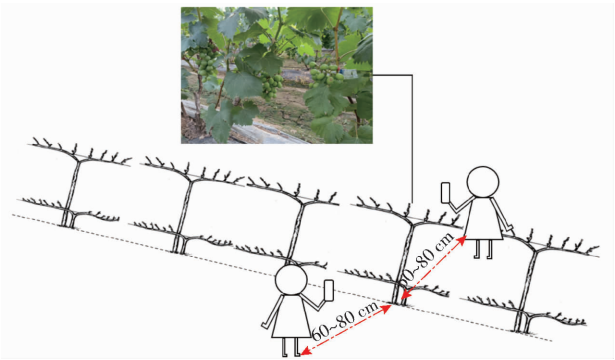


图 1 图像采集过程示意图

Fig. 1 Schematic of image acquisition process

集 4 526 幅葡萄图像,经过数据清洗,最终保留

3 780 幅;获得 2 448 像素×3 264 像素至 4 624 像素×3 472 像素的多种分辨率图像,以提高网络对不同设备采集图像的鲁棒性和普适性。

图像采集时存在拍摄光线差异和不同程度的遮挡与重叠情况。这些光线差异和不同的遮挡与重叠情况使采集的葡萄簇呈现不同的颜色、形状和纹理,导致葡萄簇特征信息更加复杂,增加了检测难度,对检测网络提出了更高的要求。同时,这些自然场景的差异反映了数据集复杂性和多样性,也能更好地验证检测算法鲁棒性。部分葡萄簇幼果图像见图 2,包含不同光照条件和不同遮挡情况。

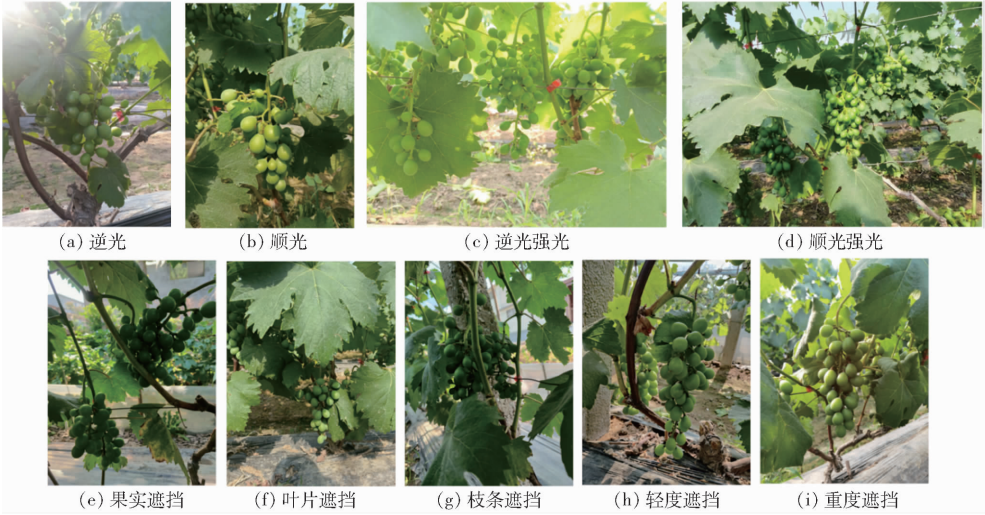


图 2 葡萄簇幼果图像数据集示例

Fig. 2 Examples of young grape cluster fruits image dataset

对采集的图像进行筛选、重写、分辨率缩放(随机按 20%~90% 的比例进行缩放)共获得 3 780 幅图像,使用 LabelImg 对图像中的葡萄簇幼果进行人工标注,最终形成葡萄簇幼果数据集 GGrape。随机将

数据集 GGrape 的 70% (2 646 幅图像)作为训练数据进行训练,数据集 GGrape 的 10% (378 幅图像)作为验证数据集,剩余 20% (756 幅图像)作为测试数据集。葡萄簇幼果数据集 GGrape 的基本信息见表 1。

表 1 GGrap 数据集基本信息

Tab. 1 Basic information of GGrap dataset

数据集类别	各数据集图像数量/幅	遮挡类型	不同遮挡图像数量/幅	光照类别	不同光照图像数量/幅	分辨率/(像素×像素)
训练集	2 646	未遮挡或轻遮挡	1 811	顺光	942	489×652~3 124×4 161
		重度遮挡	835	逆光	1 704	489×652~3 124×4 161
验证集	378	未遮挡或轻遮挡	262	顺光	133	489×652~3 121×4 161
		重度遮挡	116	逆光	245	600×800~3 110×4 147
测试集	756	未遮挡或轻遮挡	461	顺光	282	547×729~3 124×4 161
		重度遮挡	295	逆光	474	600×800~3 124×4 161

注:50% 以上葡萄簇被遮挡或多簇葡萄重叠 30% 以上,属于重度遮挡,其他为轻度遮挡。

2 检测模型建立

2.1 YOLO v8 网络

YOLO v8 作为一阶段目标检测算法,具有速度

和准确性的优势,可以结合不同尺度的不同特征图来学习物体的颜色、形状和纹理特征,有助于在近景色物体检测任务中实现高精度^[25]。YOLO v8 由 Backbone、Neck 和 Head 3 部分组成。Backbone 部

分主要进行特征提取,Neck 部分主要进行特征融合,Head 部分主要预测目标的类别和边界框信息。YOLO v8n 是 YOLO v8 5 个模型中最小的,其网络深度和特征图较小,因此体积更小,检测速度更快。可以保证检测精度的同时,达到更快的检测速度,得到的模型内存占用量更小,满足轻量化、高精度、实时性葡萄簇幼果检测的需要。

2.2 基于 SAW - YOLO v8n 网络的葡萄幼果检测方法

通道注意力机制通过调整特征通道的权重,有选择性地关注对任务更重要的特征,抑制无用的特征表达;空间注意力提升关键区域的特征表达,增强感兴趣的特定目标区域同时弱化不相关的背景区域^[26]。SA 注意力机制是一种结合了空间和通道注意力的轻量级注意力模块,具有较高的计算效率^[23]。在 YOLO v8n 模型 Neck 结构中添加 SA 模块,使模型更加关注与葡萄幼果目标相关的关键特征通道,引导模型关注图像中与葡萄幼果紧密相关的区域,这有助于增强葡萄幼果在特征表示中的显著性,在提高葡萄幼果检测性能的同时,保持较低的模型复杂性;在 Head 结构使用 Wise - IoU 作为损失函数,聚焦于普通质量的锚框,可以进一步提高模型

的准确性。SAW - YOLO v8n 网络模型结构如图 3 所示。

2.2.1 SA 注意力模块

SA 注意力模块将输入的特征图多个子特征沿通道维度分组并并行处理;每个子特征 Split 分成 2 份,分别提取通道注意力和空间注意力,使用“Shuffle Unit”将通道注意力和空间注意力整合到每个组的一个块中,通过 Concat 的方式将组内部的信息进行融合;通过“Channel Shuffle”运算对组进行重排,使不同组之间的信息进行通信,最终输出特征图。SA 模块结构如图 4 所示。SA 注意力模块通过结合空间和通道注意力、分组并行处理、“Channel Shuffle”运算等,为处理复杂场景下的葡萄幼果检测提供了一种高效且有效的方法。

2.2.2 Wise - IoU 损失函数

YOLO v8 模型使用 CIoU 作为边界框回归损失函数,CIoU 考虑了预测框与真实框的重叠面积、中心点距离和宽高比 3 个维度的损失,但由于宽高比描述的是一个相对值,而非真实差异值,当预测框与真实框宽高比相等时,CIoU 宽高比的惩罚项将失效;同时,CIoU 没有考虑高、低质量样本的平衡,当目标检测训练集中含有低质量示例时,会使模型检

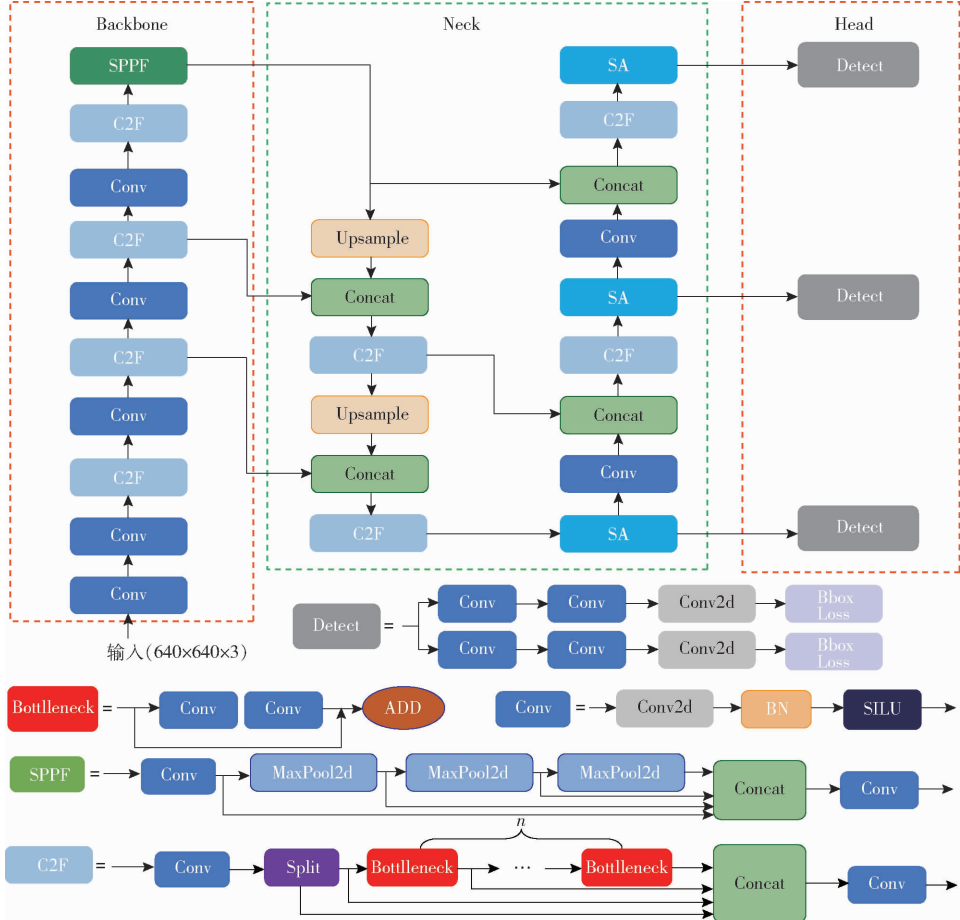


图 3 SAW - YOLO v8n 网络结构图

Fig. 3 SAW - YOLO v8n network structure

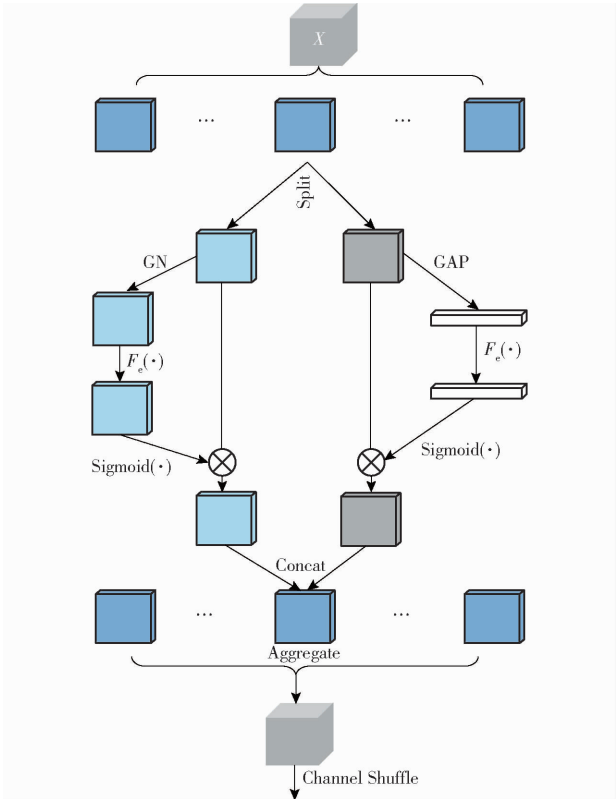


图4 SA 注意力模块结构
Fig.4 Structure of SA module

测性能降低。本研究采用 Wise-IoU v3 损失函数,与 CIoU 相比,Wise-IoU v3 没有引入宽高比的计算,而是先根据距离度量构建距离注意力,得到具有 2 层注意力机制的 Wise-IoU v1 边界框损失,其损失函数公式为

$$L_{\text{WIoUv1}} = R_{\text{WIoU}} L_{\text{IoU}} \tag{1}$$

其中
$$R_{\text{WIoU}} = \exp\left(\frac{(x - x_{gt})^2 + (y - y_{gt})^2}{(W_g^2 + H_g^2)^*}\right) \tag{2}$$

式中 L_{IoU} ——边界框损失
 R_{WIoU} ——距离注意力
 x,y ——预测框中心点横、纵坐标
 x_{gt},y_{gt} ——真实框中心点横、纵坐标
 W_g,H_g ——预测框与真实框最小外接矩形宽、高

*——将 H_g 和 W_g 从计算图分离出来的操作符

然后基于 Wise-IoU v1 的边界框损失 L_{WIoUv1} ,利用离群度构造非单调聚焦系数,得到 Wise-IoU v3 边界框损失,计算公式为

$$L_{\text{WIoUv3}} = r L_{\text{WIoUv1}} \tag{3}$$

其中
$$r = \frac{\beta}{\delta \alpha^{\beta-\delta}} \tag{4}$$

$$\beta = \frac{L_{\text{IoU}}^*}{L_{\text{IoU}}} \in [0, +\infty) \tag{5}$$

式中 r ——梯度增益 β ——离群度

α,δ ——超参数,取 1.9 和 3
 $\overline{L_{\text{IoU}}}$ ——动态滑动平均值

2.3 模型评价指标

为评价葡萄簇幼果检测模型性能,选用 P 、 R 、 mAP 、 F1 值、检测速度和模型内存占用量作为评价指标;其中, mAP 是指 IoU 阈值设为 0.5 时的平均精度,检测速度是指模型每秒检测图像的数量,模型内存占用量是指模型权重文件占内存空间量。

3 结果与分析

3.1 试验环境与模型训练

试验平台硬件环境为 AMD Ryzen 7-5800H CPU,32 GB RAM,8 GB Nvidia GeForce RTX 3070 GPU,在 64 位 Windows 10 操作系统下运行,使用 PyTorch 框架、PyCharm 编辑器、Python 编程语言进行网络模型训练。

利用训练集和验证集对 SAW-YOLO v8n 模型进行训练,共训练 300 次,图像尺寸为 640 像素 $\times$ 640 像素。SAW-YOLO v8n 模型损失曲线如图 5 所示,模型训练开始时,损失曲线收敛较快,学习效率很高,当训练迭代次数达到 270 左右时,模型学习效率逐渐达到饱和。最后 10 个迭代次数关闭 Mosaic 增强后,训练损失在 0.57 左右波动。在测试阶段,将 756 幅葡萄簇幼果测试图像输入 SAW-YOLO v8n 检测模型中,得到模型 P 为 92.80%, R 为 91.30%, mAP 为 96.10%, F1 值为 92.04%。

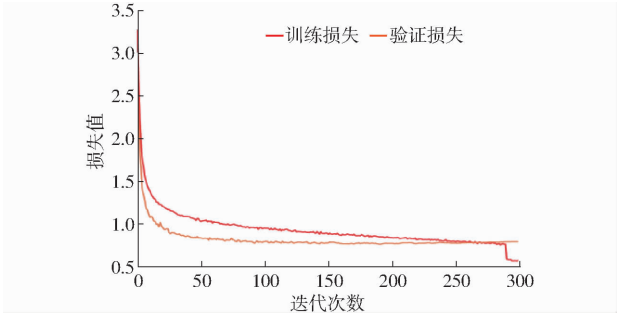


图5 SAW-YOLO v8n 模型损失曲线
Fig.5 Loss curves of SAW-YOLO v8n model

3.2 注意力机制对比试验

为验证不同注意力机制对 YOLO v8n 网络模型的优化效果,分别采用 CBAM、EMA、SimAM、ECA、SA 注意力模块对 YOLO v8n 模型进行改进,并在相同条件下进行检测性能对比试验,试验结果如表 2 所示。可见,SA-YOLO v8n 网络模型相较于其他 5 个网络模型,模型内存占用量无明显增加,检测速度无明显降低; P 虽低于 SimAM-YOLO v8n 网络模型 0.43%,但 F1 值最高,相较于 YOLO v8n、CBAM-YOLO v8n、EMA-YOLO v8n、SimAM-YOLO v8n、

ECA-YOLO v8n 分别提高 0.66%、0.27%、0.44%、0.41%、0.15%。相较于 YOLO v8n 网络模型, P 、 R 分别提高 0.99%、0.33%。综上,5 种注意力机制对 YOLO v8n 网络均有提升效果,其中 SA 注意力机制综合性能优势较为明显,更有助于葡萄幼果的准确检测。

表 2 注意力机制性能对比
Tab.2 Performance comparison of attention mechanisms

目标检测网络	P /%	R /%	mAP/%	F1 值/%	检测速度/($\text{f}\cdot\text{s}^{-1}$)	模型内存占用量/MB
YOLO v8n	91.20	90.00	95.40	90.59	135.14	6.20
CBAM-YOLO v8n	91.60	90.30	95.10	90.94	131.58	6.40
EMA-YOLO v8n	90.40	91.20	95.30	90.79	126.58	6.30
SimAM-YOLO v8n	92.50	89.20	95.30	90.82	135.14	6.20
ECA-YOLO v8n	91.70	90.40	95.50	91.05	142.86	6.20
SA-YOLO v8n	92.10	90.30	95.70	91.19	133.33	6.20

3.3 不同损失函数的性能比较

为验证不同损失函数对模型的优化效果,试验对比 SA-YOLO v8n 模型分别采用 CIoU、DIoU、EIoU、SIoU、MPDIoU、Wise-IoU v3 损失函数的性能表现,试验结果如表 3 所示。可见,相较于采用 CIoU 作为损失函数,采用 Wise-IoU v3 作为损失函数后改进模型的 P 、 R 、mAP、F1 值和检测速度分

别提高 0.76%、1.11%、0.42%、0.93%、5.64%,模型内存占用量没有改变。分别采用 DIoU、EIoU、SIoU、MPDIoU 作为损失函数时,改进模型的 mAP 和 F1 值都有明显的降低,不能满足检测模型的高精度要求。结果表明,Wise-IoU v3 作为损失函数可以聚焦于普通质量的锚框,提高模型检测精度和速度。

表 3 不同损失函数性能对比
Tab.3 Performance comparison of different loss functions

损失函数	P /%	R /%	mAP/%	F1 值/%	检测速度/($\text{f}\cdot\text{s}^{-1}$)	模型内存占用量/MB
CIoU	92.10	90.30	95.70	91.19	133.33	6.20
DIoU	92.20	90.10	95.30	91.14	135.14	6.20
EIoU	92.20	89.40	95.40	90.78	128.21	6.20
SIoU	91.00	90.50	95.40	90.75	135.14	6.20
MPDIoU	92.30	89.60	95.10	90.93	131.58	6.20
Wise-IoU v3	92.80	91.30	96.10	92.04	140.85	6.20

3.4 消融试验

以 YOLO v8n 模型为基础,在 Neck 结构中添加 SA 注意力模块,增强网络多尺度特征融合能力,提升检测目标的特征信息表示;在 Head 结构中使用 Wise-IoU 作为损失函数,以加速收敛的效果,提高网络的准确性。为验证各种改进的有效性,分别在测试集上对 YOLO v8n、SA-YOLO v8n、SAW-YOLO v8n 进行消融试验,表 4 为消融试验结果。从检测结果看,在 YOLO v8n 模型中添加 SA 注意力模块后,所得检测模型 SA-YOLO v8n 的 P 、 R 、mAP、F1 值相对于基线模型 YOLO v8n 分别提高 0.99%、0.33%、0.31%、0.66%,再采用 Wise-IoU 替换

CIoU 后,所得检测模型 SAW-YOLO v8n 的 P 、 R 、mAP、F1 值进一步得到提高,相对于基线模型 YOLO v8n 分别提高 1.75%、1.44%、0.73%、1.60%。可见,本研究所提的方法可以提高近景色葡萄幼果的检测性能。

3.5 不同检测模型性能比较

为了验证所提方法对葡萄簇幼果检测的有效性,将改进模型与 SSD、YOLO v5s、YOLO v6n 和 YOLO v7-tiny 等轻量级目标检测模型进行测试对比,测试结果如表 5 所示。从检测效果来看,改进模型 SAW-YOLO v8n 的 P 值仅低于 YOLO v6n 模型 0.32%,分别高于 SSD、YOLO v5s、YOLO v7-tiny 模型 8.03%、1.20%、1.75%;相对于 SSD、YOLO v5s、YOLO v6n、YOLO v7-tiny 模型,mAP 值分别提高 16.06%、1.05%、1.48%、0.84%,F1 值分别提高 24.85%、1.43%、1.43%、1.09%,模型内存占用量分别降低 93.16%、56.94%、37.63%、47.00%;SAW-YOLO v8n 网络模型内存占用量是所有模型

表 4 消融试验结果

Tab.4 Results of ablation test %				
模型	P	R	mAP	F1 值
YOLO v8n	91.20	90.00	95.40	90.59
SA-YOLO v8n	92.10	90.30	95.70	91.19
SAW-YOLO v8n	92.80	91.30	96.10	92.04

中最小的,仅为 6.20 MB,具有明显的轻量化优势,易于在移动端设备进行部署。综上,SAW - YOLO v8n 网络模型综合性能最优,在保证识别高准确率、

低误检率的基础上,又体现了轻量化和高速度的优势,能满足近景色葡萄簇幼果的识别高精度、低错误、轻量化的要求。

表 5 不同检测模型性能对比

Tab.5 Performance comparison of different detection models						
检测模型	P/%	R/%	mAP/%	F1 值/%	检测速度/(f·s ⁻¹)	模型内存占用量/MB
SSD	85.90	79.60	82.80	73.72	68.97	90.60
YOLO v5s	91.70	89.80	95.10	90.74	133.33	14.40
YOLO v6n	93.10	88.50	94.70	90.74	75.76	9.94
YOLO v7 - tiny	91.20	90.90	95.30	91.05	135.14	11.70
SAW - YOLO v8n	92.80	91.30	96.10	92.04	140.85	6.20

4 讨论

葡萄种植在野外环境中,环境复杂,干扰较多。为了更准确地验证算法的鲁棒性,并进一步评估算法的有效性,对不同遮挡情况、光照条件对葡萄簇检测影响进行测试,并对误检情况进行分析。

4.1 不同遮挡情况检测效果

把测试集不同遮挡情况的葡萄幼果图像分别输入 SAW - YOLO v8n 模型进行遮挡影响测试,测试结果如表 6 所示。由表 6 可知,轻度遮挡情况的 P、R、mAP、F1 值为 92.70%、92.20%、96.40%、92.45%,分别比重度遮挡情况高 2.66%、2.33%、1.26%、2.49%,说明遮挡或重叠情况会影响葡萄幼果的检测效果。图 6 为不同遮挡情况下葡萄幼果检测图像,图中所有葡萄幼果被成功检测出来,特别是图 6b 中 3 处被严重遮挡的葡萄幼果,表明 SAW - YOLO v8n 网络模型在葡萄幼果严重遮挡时依然能实现准确识别,具有较强的鲁棒性。

表 6 不同遮挡情况检测结果

Tab.6 Detection results of different occlusion conditions				
conditions	%			
遮挡情况	P	R	mAP	F1 值
轻度遮挡	92.70	92.20	96.40	92.45
重度遮挡	90.30	90.10	95.20	90.20

4.2 不同光照条件检测效果

为了评估不同光照条件对葡萄簇幼果检测的效果,采用 SAW - YOLO v8n 网络分别对测试集不同光照条件的葡萄幼果图像进行测试,测试结果如表 7 所示。在顺光条件下 P、R、mAP 和 F1 值为 93.90%、91.40%、96.20% 和 92.63%,分别比逆光条件高 3.99%、0.99%、1.05%、2.47%,说明在逆光条件下葡萄幼果颜色发白、轮廓模糊,葡萄幼果会失去原有颜色、形状和纹理等特征。图 7 为不同光照条件下葡萄幼果检测图像,图中所有葡萄幼果被成功检测出来,包括图 7a 中 2 处强光下葡萄幼果和

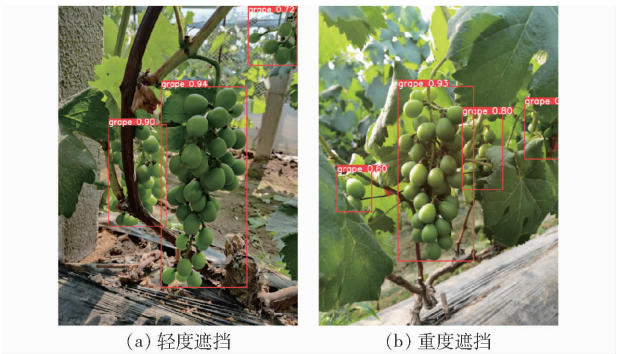


图 6 不同遮挡情况下葡萄幼果检测图像

Fig.6 Detection of young grape cluster fruits under different occlusion conditions

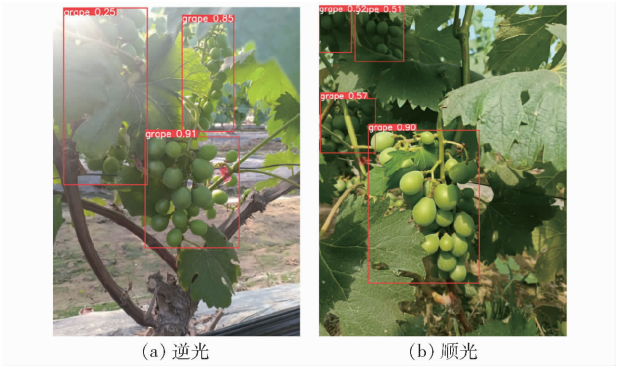


图 7 不同光照条件下葡萄幼果检测图像

Fig.7 Detection of young grape cluster fruits under different lighting conditions

表 7 不同光照条件检测结果

Tab.7 Detection results under different lighting conditions				
conditions	%			
光照条件	P	R	mAP	F1 值
顺光	93.90	91.40	96.20	92.63
逆光	90.30	90.50	95.20	90.40

图 7b 中 3 处阴影中的葡萄幼果。试验结果表明,SAW - YOLO v8n 网络模型在葡萄幼果图像偏亮和偏暗时仍能够实现准确检测,具有较强的鲁棒性。

4.3 葡萄幼果误检分析

图 8 为对葡萄幼果的错误检测示例,图 8a 橙色

椭圆处的葡萄幼果 80% 以上被遮挡导致重复检测，图 8b 橙色椭圆处葡萄幼果图像模糊导致漏检，图 8c 橙色椭圆处的 2 串葡萄簇重叠 90% 以上被检测出一串，导致漏检。错误检测葡萄簇的原因为：①如图 8a、8c 所示，葡萄幼果出现严重遮挡或重叠时，人眼难分，这增加了检测模型提取特征的难度，导致误检。②如图 8b 所示，图像出现模糊现象时，葡萄幼果特征消失，导致漏检。

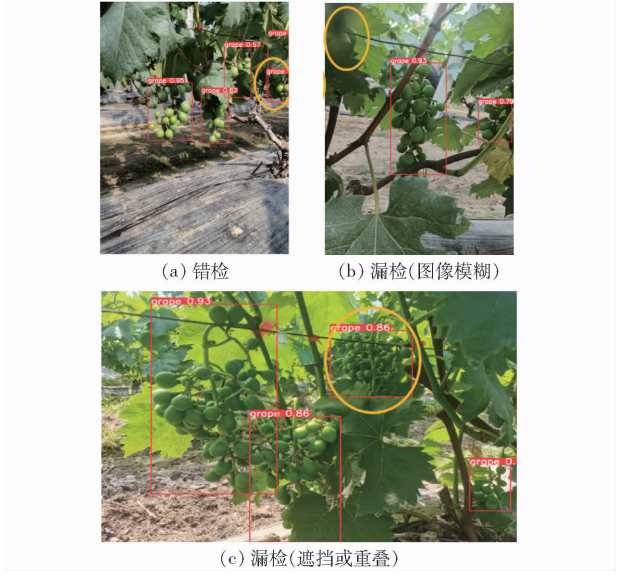


图 8 错误检测葡萄幼果示例

Fig. 8 Example of erroneous detection of young grape cluster fruits

5 结论

(1) 提出了一种基于轻量级 SAW - YOLO v8n

模型的实时检测方法。所提方法的 P 、 R 、 mAP 值和 $F1$ 值分别为 92.80%、91.30%、96.10% 和 92.04%；检测速度可达 140.85 f/s，网络模型内存占用量为 6.20 MB，满足轻量化、实时性、高精度的要求。

(2) 分别采用 CBAM、EMA、SimAM、ECA、SA 注意力模块对 YOLO v8n 模型进行改进测试，测试表明，5 种注意力机制对 YOLO v8n 网络均有提升效果，其中 SA 注意力机制综合性能优势较为明显，更有助于葡萄幼果的准确检测。在 Head 结构使用 Wise-IoU 作为损失函数，可以提高网络模型准确性。

(3) 对 SAW - YOLO v8n 模型与 SSD、YOLO v5s、YOLO v6n、YOLO v7 - tiny、YOLO v8n 等 5 种模型测试结果进行了比较，相对于 SSD、YOLO v5s、YOLO v6n、YOLO v7 - tiny、YOLO v8n 模型，SAW - YOLO v8n 模型的 mAP 分别提高 16.06%、1.05%、1.48%、0.84%、0.73%， $F1$ 值分别提高 24.85%、1.43%、1.43%、1.09%、1.60%，模型内存占用量分别降低 93.16%、56.94%、37.63%、47.00%、0；SAW - YOLO v8n 模型内存占用量是所有模型中最小的，仅为 6.20 MB，具有明显的轻量化优势，能满足近景色葡萄簇幼果的识别高精度、低错误、轻量化的要求。

(4) 对于不同遮挡程度和光照条件的葡萄幼果检测，提出的检测方法可以很好地识别葡萄簇幼果，说明基于 SAW - YOLO v8n 的葡萄簇幼果检测方法能适应不同遮挡和光照变化，具有良好的鲁棒性。

参 考 文 献

[1] LIU S, LI X, WU H, et al. A robust automated flower estimation system for grape vines[J]. Biosystems Engineering, 2018, 172: 110 - 123.

[2] DU W, ZHU Y, LI S, et al. Spikelets detection of table grape before thinning based on improved YOLO V5s and Kmeans under the complex environment[J]. Computers and Electronics in Agriculture, 2022, 203: 107432.

[3] RUDOLPH R, HERZOG K, TÖPFER R, et al. Efficient identification, localization and quantification of grapevine inflorescences in unprepared field images using fully convolutional networks[J]. Vitis Journal of Grapevine Research, 2019, 58(3): 95 - 104.

[4] PALACIOS F, BUENO G, SALIDO J, et al. Automated grapevine flower detection and quantification method based on computer vision and deep learning from on-the-go imaging using a mobile sensing platform under field conditions[J]. Computers and Electronics in Agriculture, 2020, 178: 105796.

[5] 刘平,朱衍俊,张同勋,等. 自然环境下贴叠葡萄串的认可与图像分割算法[J]. 农业工程学报,2020,36(6):161 - 169.

LIU Ping, ZHU Yanjun, ZHANG Tongxun, et al. Algorithm for recognition and image segmentation of overlapping grape cluster in natural environment[J]. Transactions of the CSAE, 2020, 36(6): 161 - 169. (in Chinese)

[6] 宋怀波,尚钰莹,何东健. 果实目标深度学习识别技术研究进展[J]. 农业机械学报,2023,54(1):1 - 19.

SONG Huaibo, SHANG Yuying, HE Dongjian. Review on deep learning technology for fruit target recognition[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(1): 1 - 19. (in Chinese)

[7] SANTOS T T, SOUZA L L D, SANTOS A A D, et al. Grape detection, segmentation, and tracking using deep neural networks and three-dimensional association[J]. Computers and Electronics in Agriculture, 2020, 170: 105247.

[8] 宋怀波,江梅,王云飞,等. 融合卷积神经网络与视觉注意机制的苹果幼果高效检测方法[J]. 农业工程学报,2021,

37(9):297–303.

SONG Huaibo, JIANG Mei, WANG Yunfei, et al. Efficient detection method for young apples based on the fusion of convolutional neural network and visual attention mechanism[J]. Transactions of the CSAE, 2021, 37(9): 297–303. (in Chinese)

[9] 宋怀波,马宝玲,尚钰莹,等. 基于 YOLO v7-ECA 模型的苹果幼果检测[J]. 农业机械学报,2023,54(6):233–242.

SONG Huaibo, MA Baoling, SHANG Yuying, et al. Detection of young apple fruits based on YOLO v7-ECA model[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(6): 233–242. (in Chinese)

[10] 宋怀波,王亚男,王云飞,等. 基于 YOLO v5s 的自然场景油茶果识别方法[J]. 农业机械学报,2022,53(7):234–242.

SONG Huaibo, WANG Ya`nan, WANG Yunfei, et al. *Camellia oleifera* fruit detection in natural scene based on YOLO v5s[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(7): 234–242. (in Chinese)

[11] 张志远,罗铭毅,郭树欣,等. 基于改进 YOLO v5 的自然环境下樱桃果实识别方法[J]. 农业机械学报,2022,53(增刊 1):232–240.

ZHANG Zhiyuan, LUO Mingyi, GUO Shuxin, et al. Cherry fruit detection method in natural scene based on improved YOLO v5[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(Supp. 1): 232–240. (in Chinese)

[12] 何斌,张亦博,龚健林,等. 基于改进 YOLO v5 的夜间温室番茄果实快速识别[J]. 农业机械学报,2022,53(5): 201–208.

HE Bin, ZHANG Yibo, GONG Jianlin, et al. Fast recognition of tomato fruit in greenhouse at night based on improved YOLO v5[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(5): 201–208. (in Chinese)

[13] WANG J, ZHANG Z, LUO L, et al. SwinGD: a robust grape bunch detection model based on Swin Transformer in complex vineyard environment[J]. Horticulturae, 2021, 7(11): 492.

[14] LIU B, ZHANG Y, WANG J, et al. An improved lightweight network based on deep learning for grape recognition in unstructured environments[J]. Information Processing in Agriculture, 2024, 11(2): 202–216.

[15] LU S, LIU X, HE Z, et al. Swin-Transformer-YOLOv5 for real-time wine grape bunch detection[J]. Remote Sensing, 2022, 14(22): 5853.

[16] 李国进,黄晓洁,李修华,等. 采用轻量级网络 MobileNetV2 的酿酒葡萄检测模型[J]. 农业工程学报,2021,37(17):168–176.

LI Guojin, HUANG Xiaojie, LI Xiuhua, et al. Detection model for wine grapes using MobileNetV2 lightweight network[J]. Transactions of the CSAE, 2021, 37(17): 168–176. (in Chinese)

[17] LI H, LI C, LI G, et al. A real-time table grape detection method based on improved YOLOv4-tiny network in complex background[J]. Biosystems Engineering, 2021, 212: 347–359.

[18] ZHANG T, WU F, WANG M, et al. Grape-bunch identification and location of picking points on occluded fruit axis based on YOLOv5-GAP[J]. Horticulturae, 2023, 9(4): 498.

[19] WEI X, XIE F, WANG K, et al. A study on Shine-Muscat grape detection at maturity based on deep learning[J]. Scientific Reports, 2023, 13: 4587.

[20] 孙俊,吴兆祺,贾忆琳,等. 基于改进 YOLOv5s 的果园环境葡萄检测[J]. 农业工程学报,2023,39(18):192–200.

SUN Jun, WU Zhaoqi, JIA Yilin, et al. Detecting grape in an orchard using improved YOLOv5s[J]. Transactions of the CSAE, 2023, 39(18): 192–200. (in Chinese)

[21] 刘云玲,张天雨,姜明,等. 基于机器视觉的葡萄品质无损检测方法研究进展[J]. 农业机械学报,2022,53(增刊 1):299–308.

LIU Yunling, ZHANG Tianyu, JIANG Ming, et al. Review on non-destructive detection methods of grape quality based on machine vision[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(Supp. 1): 299–308. (in Chinese)

[22] TERVEN J R, CORDOVA-ESPARZA D M. A comprehensive review of YOLO: from YOLO v1 to YOLO v8 and beyond[EB/OL]. (2024-02-04)[2024-03-06]. <https://doi.org/10.48550/arXiv.2304.00501>.

[23] ZHANG Q, YANG Y. SA-Net: shuffle attention for deep convolutional neural networks[C]// ICASSP- IEEE International Conference on Acoustics, 2021: 2235–2239.

[24] TONG Z, CHEN Y, XU Z, et al. Wise-IoU: bounding box regression loss with dynamic focusing mechanism[EB/OL]. (2023-04-08)[2024-02-10]. <https://arxiv.org/abs/2301.10051>.

[25] REIS D, KUPEC J, HONG J, et al. Real-time flying object detection with YOLOv8[EB/OL]. (2023-05-17)[2024-02-06]. <https://arxiv.org/abs/2305.09972>.

[26] GUO M, XU T, LIU J, et al. Attention mechanisms in computer vision: a survey[J]. Computational Visual Media, 2022, 8: 331–368.