

基于 GGWL – YOLO 模型的轻量化熟草莓识别方法

李聪聪^{1,2} 李浩冬^{1,2} 赵志超^{1,2} 李亚南¹ 孟超月^{1,2}

(1. 河北农业大学信息科学与技术学院, 保定 071001; 2. 河北省农业大数据重点实验室, 保定 071001)

摘要: 由于草莓果实受枝叶遮挡以及光线变化的复杂背景影响,草莓果实识别精度较低,影响自动化采摘效率。为实现复杂环境下的草莓果实目标快速识别,本文提出基于 YOLO 11 改进的 GGWL – YOLO 成熟草莓识别模型。首先,引入多层次特征提取的 HGNetV2 网络,并加入多尺度池化技术的 SPPF 模块和结合多头自注意力与前馈神经网络的 C2PSA 模块,构建 GhostHGNetV2 网络作为特征提取主干网络,适用于草莓果实检测,更进一步提取草莓果实的未遮挡部分细小特征。其次,在颈部网络引入 GFPN 网络,用 DySample 改进上采样模块,构建 GDFPN 网络实现多层次跨尺度特征融合,与 GhostHGNetV2 网络细小特征提取能力相结合,改善草莓果实识别的遮挡问题。然后,选用 WIoUv3 改进边界框回归损失函数,改善因光照影响草莓果实识别问题,减少草莓图像低质量样本的影响。最后,使用 Lamp 方法对改进的整体模型进行通道剪枝和模型微调,减少冗余通道数保留关键的通道信息。试验结果表明,改进的 GGWL – YOLO 模型的精确率、mAP@ 0.5、参数量、检测速度分别为 92.5%、94.3%、 3.28×10^6 、91.5 f/s,与 YOLO v8s、YOLO 11s、RT – DETR – r18 模型相比精确率分别提升 1.4、1.2、2.4 个百分点,参数量减少 70.4%、65.1%、83.4%,并且在平均精度均值和浮点运算量也存在优势。改进后的模型可以快速、准确地识别复杂环境下的草莓果实,进一步推动草莓自动化采摘。

关键词: 草莓; 图像识别; 目标检测; YOLO 11

中图分类号: TP391.4 文献标识码: A 文章编号: 1000-1298(2026)14-0039-10

OSID:



Recognition Method of Lightweight Ripe Strawberry Based on GGWL – YOLO Model

Li Congcong^{1,2} Li Haodong^{1,2} Zhao Zhichao^{1,2} Li Ya'nan¹ Meng Chaoyue^{1,2}

(1. College of Information Science and Technology, Hebei Agricultural University, Baoding 071001, China

2. Hebei Province Key Laboratory of Agricultural Big Data, Baoding 071001, China)

Abstract: Due to the influence of the complex background of strawberry fruit and branches and leaves as well as light changes, the recognition accuracy of strawberry fruit is low. It affects the efficiency of automated picking. To achieve rapid recognition of strawberry fruit targets in complex environments, an improved GGWL – YOLO mature strawberry recognition model based on YOLO 11 was proposed. Firstly, HGNetV2 network with multi-level feature extraction was introduced, SPPF module with multi-scale pooling technology and C2PSA module with multi-head self-attention and feedforward neural network were added. GhostHGNetV2 network was constructed as the feature extraction backbone network, which was suitable for strawberry fruit detection. Further the fine features of the uncovered part of strawberry fruit was extracted. Secondly, the GFPN network was introduced into the neck network, the upsampling module was improved with DySample, the GDFPN network was constructed to achieve multi-level and cross-scale feature fusion, and it was combined with the fine feature extraction ability of the GhostHGNetV2 network to improve the occlusion problem in strawberry fruit recognition. Then, WIoUv3 was used to improve the bounding box regression loss function, improve the problem of strawberry fruit recognition due to the influence of light, and reduce the influence of low-quality strawberry picture samples. Finally, Lamp method was used for channel pruning and model fine-tuning of the improved overall model to reduce the number of redundant channels and retain key channel information. The test results showed that the precision, mAP@ 0.5, parameter number and detection speed of the improved

GGWL – YOLO model were 92.5% , 94.3% , 3.28×10^6 and 91.5 f/s, respectively. Compared with YOLO v8s, YOLO 11s and RT – DETR – r18 models, the precision was improved by 1.4, 1.2 and 2.4 percentage points, respectively, while the number of parameters was reduced by 70.4% , 65.1% and 83.4% . Moreover, it also had advantages in mean average accuracy and floating point computation. The improved model can quickly and accurately identify strawberry fruits in complex environments, and further promote automatic strawberry picking.

Key words: strawberry; image recognition; target detection; YOLO 11

0 引言

草莓味道甜美,富含多种营养物质,被誉为“水果皇后”,深受消费者喜爱^[1]。中国是世界上最大的草莓种植国家,随着草莓需求不断增长,其种植面积和产量也随之增加,导致草莓采摘工作量愈加繁重^[2]。现今,草莓采摘机器人已投入使用,由于草莓种植环境错综复杂,果实与枝叶遮挡、光线等因素会影响草莓的精确检测。因此,提高在复杂环境下的草莓识别精度,对推动草莓自动化采摘和智慧农业发展具有重要意义^[3]。

随着人工智能技术的发展,深度学习在农业场景得到广泛应用,果实检测精度大幅提升。与传统机器学习算法相比,深度学习目标检测算法具有适应性好、检测速度快、准确率高等优点。孙俊等^[4]提出了一种在复杂果园环境下的葡萄检测模型 MRW – YOLO v5s,研究表明模型进一步改善了葡萄果实识别遮挡与光照问题。王元昊等^[5]针对柑橘果实的识别遮挡与定位问题,提出了一种基于 YOLO v8 – SAM 的改进算法,运用边缘检测法、最小二乘法、双目立体相机三维稠密深度点云得到被遮挡柑橘有效轮廓。上述研究表明深度学习在果实识别方面具有较强的可行性,并为草莓果实识别积累了重要的研究经验和基础。

Du 等^[6]提出基于 YOLO v7 改进的 DSW – YOLO 模型,降低了草莓果实识别的遮挡影响,改进的网络模型识别精确率为 82.8%。Yang 等^[7]提出了一种 LS – YOLO v8s 草莓成熟度检测模型,检测精度达到 94.4%,检测速度为 19.23 f/s。刘莫尘等^[8]改进 YOLO v8pose 模型对红熟期草莓进行识别与果柄关键点检测,在光线、遮挡与拍摄角度对模型预测的影响下 mAP 分别为 94.52%、95.48%、94.63%。

由上述研究可知,复杂背景下草莓检测存在大量错检、漏检现象,检测精度仍有待提高。针对上述问题,本文选用 YOLO 11n 模型进行改进,提出复杂环境下草莓果实识别的轻量化目标检测模型 GGWL – YOLO (GhostHNetV2 – GDFPN – WIoUv3 – Lamp – YOLO 11)。引入 HNetV2^[9] 网络,并将 HGBlock 模块中轻量化卷积替换为 GhostConv^[10],同时增加

SPPF 和 C2PSA 模块,构建 GhostHNetV2 网络替换 YOLO 11 原有主干网络;在颈部网络引入 GDFPN^[11] 网络,用 DySample^[12] 改进上采样模块,构建 GDFPN 网络实现跨尺度多层次特征融合;选用 WIoUv3^[13] 改进边界框回归损失函数;使用 Lamp^[14] 方法对改进的整体模型进行通道剪枝和微调。以期实现模型轻量化,满足移动设备部署的需求,提高复杂背景下草莓果实检测精度与速度。

1 数据集与方法

1.1 数据集建立

数据采自河北省保定市水润嘉禾农业园区的高架温室草莓,拍摄场景如图 1 所示。于 2024 年 4 月 20 日 09:00—18:00,分别在左 45°、正面、右 45° (3 个水平方向),草莓高架 0.3、0.6 m 距离 (2 处位置),顺光、逆光 (2 种光线) 条件下进行拍摄,拍摄设备为红米 K30pro 型手机,共拍摄 4 892 幅图像,分辨率为 3 120 像素 × 4 208 像素,如图 2a ~ 2g 所示。用 LabelImg 软件对草莓图像进行标注,标注类型分为成熟草莓 (ripe) 和未成熟草莓 (unripe),图像标签以 txt 格式保存,如图 2h 所示。将采集到的图像数据划分为 4 个数据集,分别是综合数据集、简单数据集、遮挡数据集、光照数据集,图像数量分别为 4 492、200、200、200 幅 (后 3 个数据集实际共包含 400 幅图像),数据集划分如表 1 所示。将综合数据集按比例 8:1:1 以随机方式划分为训练集、验证集和测试集,其中图像数量分别为 3 594、451、447 幅,标签数量分别为 20 083、2 646、2 466 个。



图 1 草莓种植场景

Fig. 1 Strawberry planting scene

1.2 草莓果实目标检测模型

提出基于 YOLO 11n 改进的 GGWL – YOLO 轻

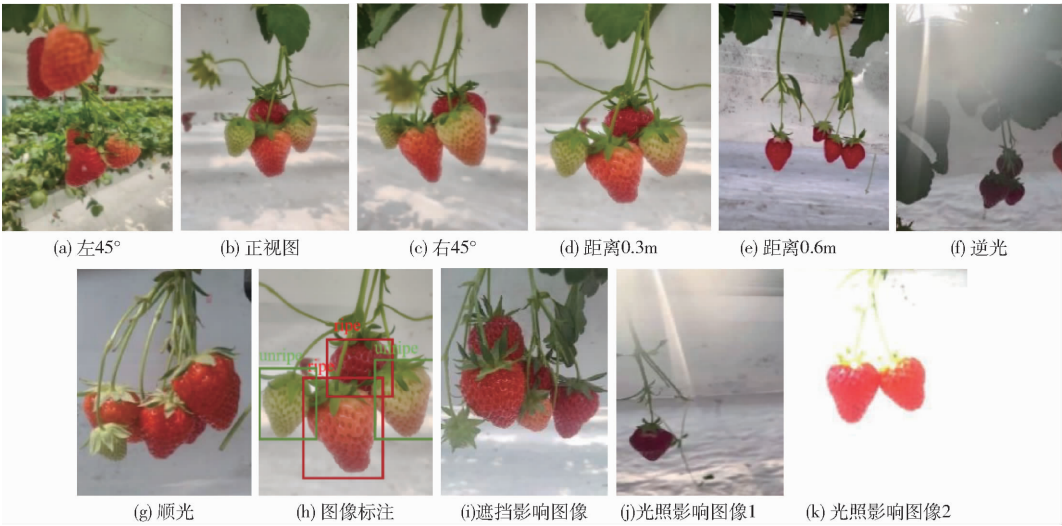


图 2 草莓果实图像样例

Fig. 2 Strawberry fruit image examples

表 1 数据集划分详情

Tab. 1 Datasets partition

数据集	作用	构成
综合数据集	仅用于 GGWL - YOLO 模型训练、验证、测试	4 492 幅复杂背景下受遮挡、光照等因素影响的图像(与以下 3 个数据集独立)
简单数据集	仅用于模块改善遮挡、光照效果试验,(不参与模型的训练、验证、测试)	200 幅既不受遮挡也不受光照影响的图像
遮挡数据集	仅用于模块改善遮挡效果试验(不参与模型的训练、验证、测试)	含 200 幅图像,即 100 幅只受遮挡、不受光照影响的独立图像,100 幅源自简单数据集
光照数据集	仅用于模块改善光照效果试验(不参与模型的训练、验证、测试)	含 200 幅图像,即 100 幅不受遮挡、只受光照影响的独立图像,100 幅源自简单数据集(同遮挡数据集)

量化草莓果实检测模型,改进的模型整体结构如图 3 所示,其中 Bbox Loss 表示回归损失,Cls Loss 表示分类损失,Add 表示特征相加操作,Concat 表示通道拼接操作。主要从 4 方面进行改进:引入 HGNetV2 网络、增加 SPPF 和 C2PSA 模块,构建 GhostHGNetV2 网络替换 YOLO 11 主干网络;引入 GFPN 颈部网络并用 DySample 改进上采样模块,构建 GDFPN 网络;选用 WIoUv3 改进损失函数;采用 Lamp 方法对改进的整体模型进行通道剪枝。

1. 2. 1 GhostHGNetV2 主干网络

GhostHGNetV2 基于 HGNetV2 网络进行改进,由 1 个 HGStem 模块和 4 个 HG_Stage 模组构成,HG_Stage 模组由 GhostHGBlock 模块和 DWConv^[15] 模块构成,

并在最后 1 个模组中引入 SPPF 模块和 C2PSA 模块,GhostHGNetV2 整体结构如图 4 所示,其中Conv - sc 表示通道压缩卷积,Conv - ex 表示通道激励卷积,恒等映射表示将每个元素映射到其自身。HGStem 是 GhostHGNetV2 的预处理层,起对原始草莓图像进行初步的特征提取作用,将特征图尺寸变为原图的 1/4,降低后续计算过程的复杂程度。GhostHGBlock 模块是 GhostHGNetV2 网络结构的核心,旨在用多层次化方式提取草莓果实特征,学习更加丰富和深层特征,更好地刻画草莓果实未遮挡区域的细微特征。GhostHGBlock 模块借鉴了 OSA(One-shot aggregation)^[16] 模块特征聚合思想,每层卷积采用双向连接方式。GhostHGBlock 将输入原始特征与每层 GhostConv 轻量化卷积输出特征与进行通道拼接,然后通过挤压和激励^[17] 操作识别并放大关键特征。

GhostConv 卷积用于替换 HGBlock 模块中由 1 × 1 卷积和 DWConv 组成的轻量化卷积,目的是在增强草莓特征提取效果的前提下减少计算量,不仅保持模型的轻量化,还保留了原始特征信息。DWConv 计算式为

$$y_{ij} = \Phi_{i,j}(y_i')$$

$(i = 1, 2, \cdots, m; j = 1, 2, \cdots, s)$

(1)

式中 y_i' ——第 i 个内在特征图
 $\Phi_{i,j}$ ——负责生成第 j 个 Ghost map 的线性操作

SPPF 模块与 C2PSA 模块结构如图 3 所示。SPPF 通过采用多尺度池化技术在不同尺度下对图像进行特征提取与编码,能够将任意尺寸的输入图像缩放至固定尺寸,并生成具有固定长度的特征向量。C2PSA 模块结合了多头自注意力机制、多尺度

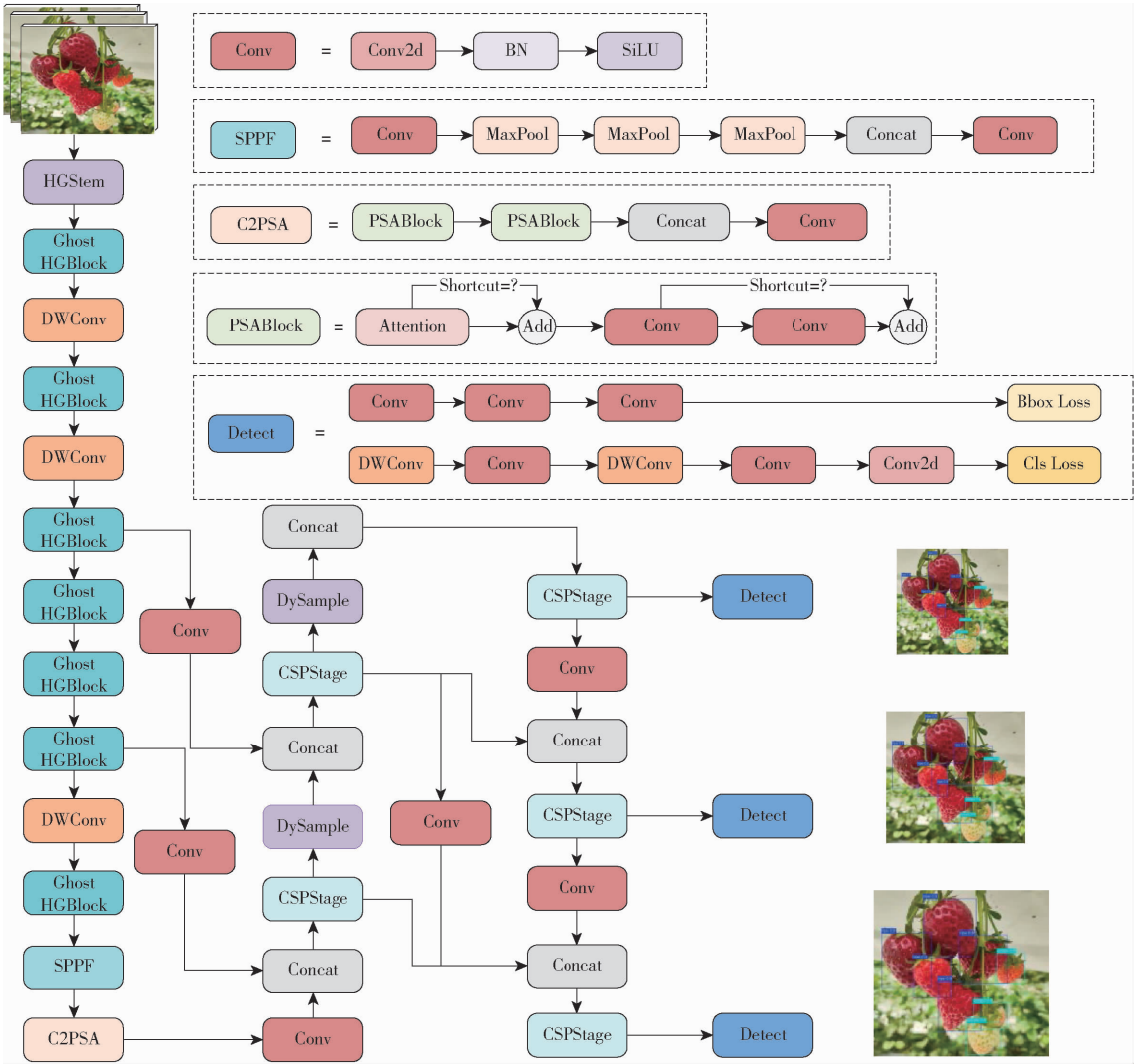


图 3 GGWL-YOLO 网络结构

Fig. 3 GGWL-YOLO network structure

卷积以及通道加权操作,极大地增强了模型对复杂遮挡物体和重要区域的关注能力。为了优化梯度传播和提升网络训练效率,C2PSA 选择性地融入了残差结构(Shortcut),并利用多尺度卷积技术,精准捕捉输入特征的复杂非线性关系,增强模型的特征表示能力。

1.2.2 GDFPN 颈部网络

GDFPN 网络基于 GFPN 网络进行改进,将上采样模块改进为 DySample 模块,包含跳跃层和跨尺度连接的路径融合,能够将深层特征信息和浅层特征信息进行充分交换,防止在特征融合中丢失上下文关系。不同尺度特征使用不同的通道数,在轻量级计算量约束下,灵活控制深层特征和浅层特征的表达能力。CSPStage 连接是基于卷积的特征融合,同时引入重参数化思想和 ELAN 连接,在不增加更多计算量的同时,提升模型精度,整体结构如图 5 所示。

DySample 模块是一个超轻量级的动态上采样

器,相较传统的 CARAFE^[18]上采样方法,能够在不增加额外负担的情况下实现图像分辨率的提升。上采样特征生成过程:采样点生成器基于输入特征 X 生成采样集 S ,GridSample 方法利用采样集 S 中的位置对假设的双线性插值进行重新采样,生成上采样特征 X' ,过程如图 6 所示。上述过程可表示为

$$X' = \text{GridSample}(X, S) \tag{2}$$

式中 S ——大小为 $2 \times H_2 \times W_2$ 的采样集,其中第一维的数字 2 表示 x 轴和 y 轴坐标

采样集 S 生成过程:输入特征 X 通过一个线性层和像素洗牌(Pixel Shuffling)^[19]操作生成偏移量 O ,采样集 S 是偏移量 O 和原始采样网格 G 之和,即

$$S = G + O \tag{3}$$

其中

$$O = \text{Linear}(X) \tag{4}$$

最后,通过 GridSample 和采样集 S ,可以生成上采样特征 X' (式(2))。

由于归一化层的存在,局部采样位置的移动范围可能会显著重叠,影响边界附近的预测,会导致逐

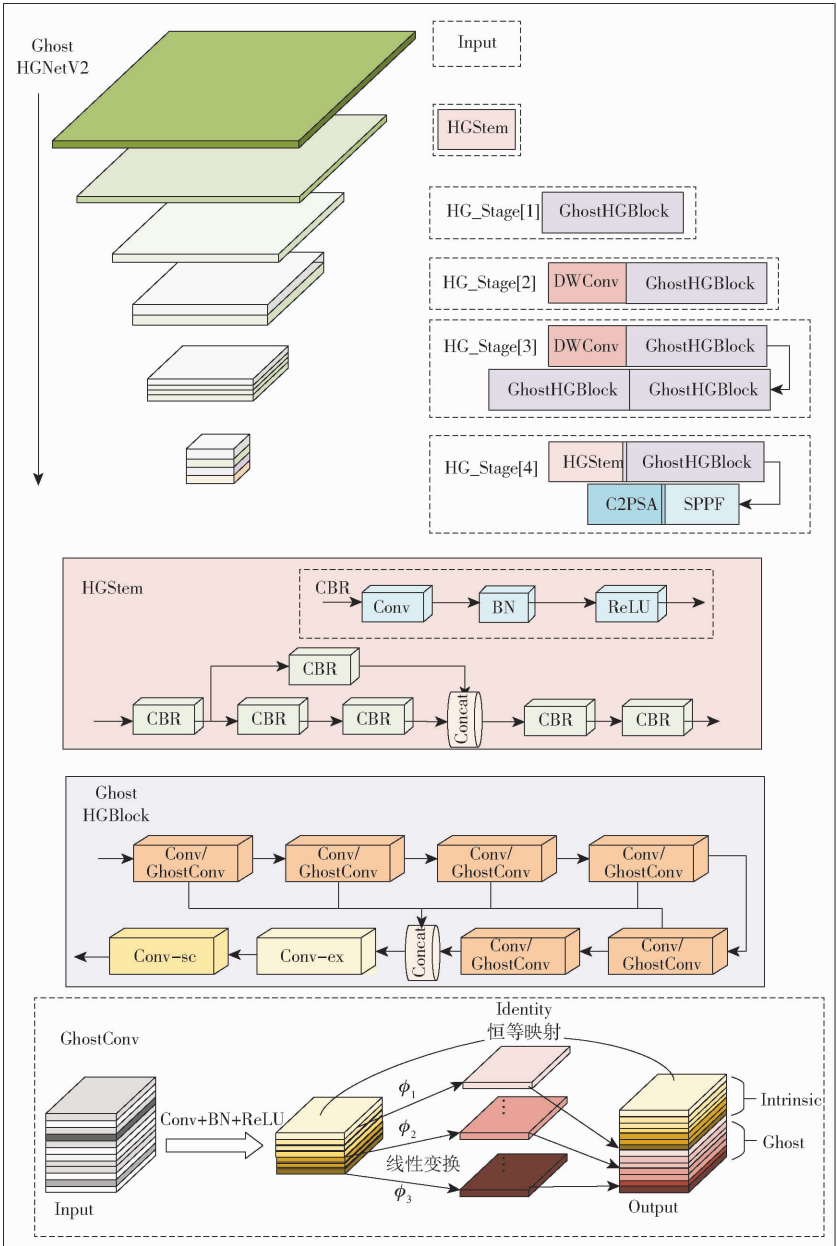


图 4 GhostHGNetV2 网络结构

Fig. 4 GhostHGNetV2 network structure

阶段传播与实际输出产生偏移。为解决这个问题，采用线性投影输入特征生成逐点的“动态范围因子”，使用 Sigmoid 函数和 0.5 的静态因子，动态范围值为 $[0, 0.5]$ ，以 0.25 为中心，满足重叠与非重叠之间的理论边界条件，使采样位置的移动范围在局部受到约束，过程如图 7 所示，因此将式 (4) 重写为

$$O = 0.5 \text{Sigmoid}(\text{Linear}_1(X)) \text{Linear}_2(X) \quad (5)$$

1.2.3 WIoUv3 损失函数

YOLO 11n 采用 CIoU^[20] 损失函数，其对正样本和负样本赋予相同权重。在实际数据集中，正负样本数量往往是不平衡的，大量的负样本会对模型的识别性能产生显著不利影响，CIoU 表现效果不佳，因光照影响存在的低质量样本导致识别效果下降。

因此针对光照影响导致草莓果实识别精度低问题，使用 WIoUv3 损失函数可得到有效改善。

CIoU 只考虑锚框和真实边界框之间的重叠面积、质心距离和纵横比，在此基础上 WIoUv3 引入了动态非单调机制。锚框和真实边界框示意图如图 8 所示，其中紫色矩形表示真实边界框，橙色矩形表示锚框，使用动态非单调频率调制，增强对高质量锚框的关注，减少低质量样本的影响。锚框与真实边界框重叠程度越低， R_{wIoU} 越大，WIoUv1 计算式为

$$L_{wIoUv1} = R_{wIoU} L_{IoU} \quad (6)$$

其中
$$L_{IoU} = 1 - IoU \quad (7)$$

$$R_{wIoU} = \exp\left(\frac{(x - x_{gt})^2 + (y - y_{gt})^2}{(W_g^2 + H_g^2)^*}\right) \quad (8)$$

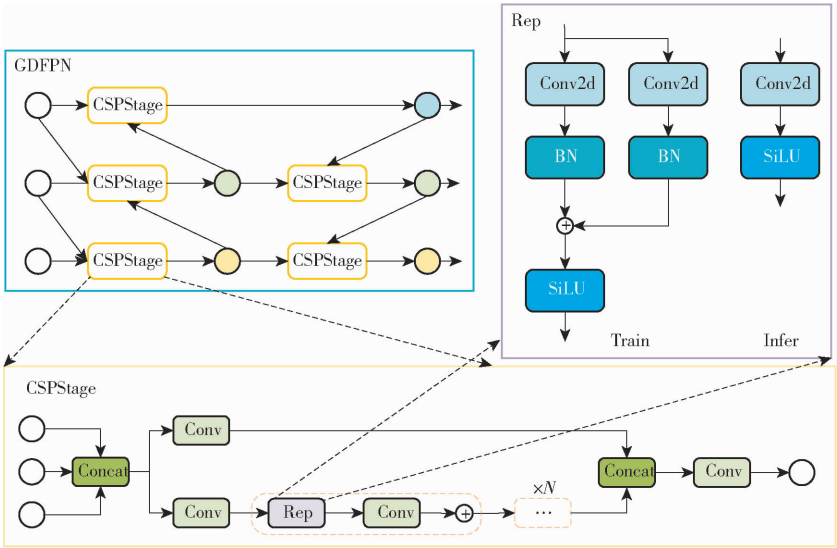


图 5 GDFPN 网络结构

Fig. 5 GDFPN network structure

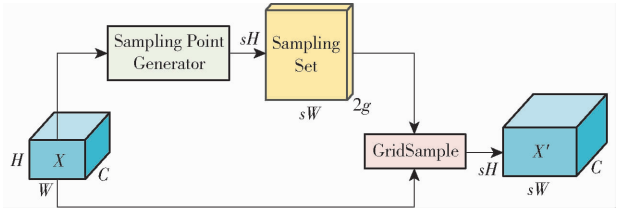


图 6 上采样特征生成

Fig. 6 Upsampling feature generation

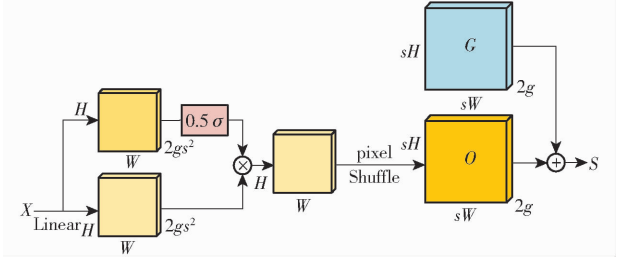


图 7 基于动态范围因子的采样集

Fig. 7 Sampling set based on dynamic range factors

式中 W_g 、 H_g ——锚框与真实边界框最小外接矩形的宽度和高度

x 、 y ——锚框中心点横、纵坐标

x_{gt} 、 y_{gt} ——真实边界框中心点横、纵坐标

R_{WIoU} ——锚框与真实边界框之间中心点距离为更好地减少低质量样本造成的损失,在 $WIoUv1$ 基础上引入动态非单调机制构造 $WIoUv3$, 公式为

$$L_{WIoUv3} = rL_{WIoUv1} \tag{9}$$

其中

$$r = \frac{\beta}{\delta \alpha^{\beta - \alpha}} \tag{10}$$

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

式中 L_{IoU}^* ——单调聚焦因子

β ——锚框离群度

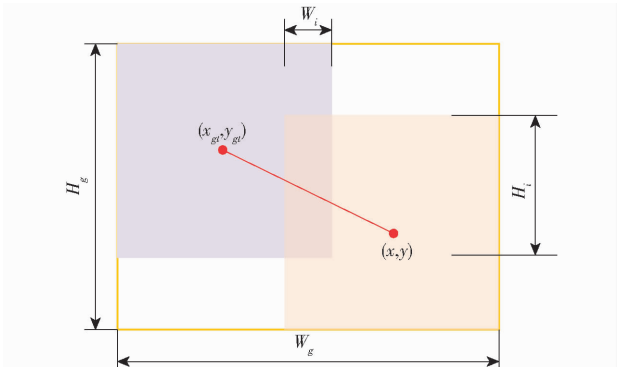


图 8 锚框和真实边界框示意图

Fig. 8 Schematic of anchor box and truth bounding box

r ——非单调聚焦因子

α ——权重函数

β 越高表示锚框质量越差, r 可防止低质量草莓图像样本产生有害梯度, 增加对正确样本的关注, 提高网络模型准确性。

1.2.4 Lamp 通道剪枝算法

模型剪枝算法主要可分为模型剪枝和模型微调。剪枝算法是利用特定的剪裁方法, 删除模型中的冗余信息, 得到一个相对紧凑的网络模型, 在参数量下降的同时精度会有所损失; 模型微调是对剪枝后的模型进行参数微调重新训练, 改善精度下降问题。采用 Lamp 方法优化模型, 将每个权重张量全部压缩为一维向量 W , 计算每个向量 W 的幅度 $|W|$, 将 $|W|$ 按照顺序由小到大排列, u 、 v 表示权重大小 $|W|$ 的索引, $W(u)$ 、 $W(v)$ 表示索引 u 、 v 映射的权重项。因为 $|W|$ 为按照大小顺序排列, 故 $u < v$, 则 $W(u) < W(v)$, 此时 Lamp 评分定义为 $score(u; W)$, 计算式为

$$\text{score}(u; \mathbf{W}) = \frac{(\mathbf{W}(u))^2}{\sum_{v \geq u} (\mathbf{W}(v))^2} \tag{11}$$

其中,分子部分表示目标连接权重幅度平方;分母表示在同一层中剩余连接的权重幅度平方之和。由式(11)可知,权重项越大,其对应的分数越高,

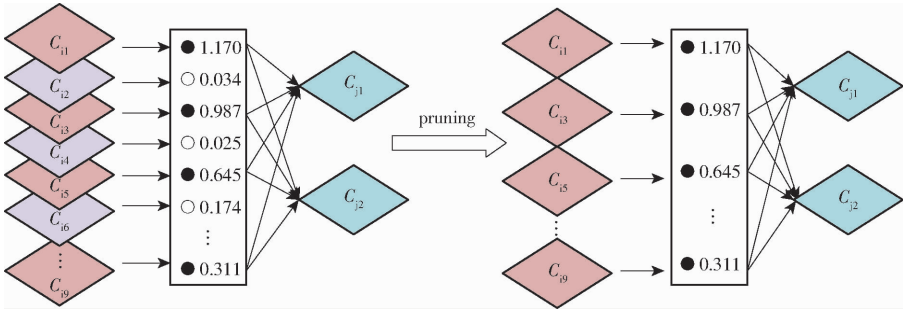


图 9 剪枝示意图
Fig. 9 Pruning diagram

1.3 试验

1.3.1 试验环境

本模型基于 Pytorch 深度学习框架搭建,操作系统为 Windows 10,硬件配置为 Intel(R)Core(TM) i910920X CPU @ 3.5 GHz 处理器,显卡配置为 NVIDIA GeForce RTX 3090 24 GB,内存为 64 GB,软件为 CUDA 11.6、PyTorch 2.5.0、Python 3.9.1,部分训练参数如表 2 所示。

表 2 训练参数
Tab. 2 Training parameters

参数	参数
图像尺寸(像素×像素)	640×640
训练批次数	24
迭代周期数	200
初始学习率	0.01
优化器	SGD
优化器动量	0.937
优化器权重衰减系数	0.000 5

1.3.2 评价指标

为衡量网络结构对草莓目标检测的准确性,采用图像检测的精确率(Precision, P)、召回率(Recall, R)、平均精度均值(mean Average precision, mAP)、参数量、浮点运算量、检测速度和模型占用量评估模型效果,mAP@0.5 为当 IoU 阈值为 0.5 时每个类别平均精度和的均值,而 mAP@0.5:0.95 为当 IoU 阈值从 0.5 到 0.95 时每个类别平均精度和的均值。

2 结果与分析

2.1 消融试验

以 YOLO 11n 目标检测模型为基线模型进行消

对应的连接越重要。分数保证了每层至少有 1 个评分为 1 的最优通道,避免层失效问题。计算出分数后,就会对全局中最小的 score 分数权重进行修剪,不断重复操作,直至满足全局稀疏性约束,如图 9 所示。

融试验,结果如表 3 所示。由表 3 可知,单独改进 GhostHGNetV2、GDFPN 和 WIoUv3 时,效果没有提升甚至稍差。采用 GhostConv 轻量化卷积和采用 DySample 轻量级上采样模块,使得模型参数量仅为 4.37×10^6 。SPPF 与 C2PSA 增强了主干网络对遮挡草莓和细小特征的关注能力,DySample 模块使得颈部网络更充分融合深、浅层特征,与 YOLO 11n 对比精确率提升 0.5 个百分点,表明 GhostHGNetV2 与 GDFPN 同时改进的有效性。在此基础上优化 WIoUv3 损失函数,WIoUv3 利用合理的梯度增益分配策略,加重对低质量草莓图像惩罚,防止其产较大的有害梯度,与 YOLO 11n 对比精确率提高 1.4 个百分点,证明了优化 WIoUv3 损失函数可行性。再加入 Lamp,模型精确率、检测速度分别增加 2.2 个百分点、10.9 f/s,归因于 Lamp 让模型更加关注重要的通道特征,减轻过度拟合使模型重新训练最优参数,加快了模型的推理速度。同时 Lamp 通过删除不重要权重的通道特征,降低了模型参数量并节省了 GPU 计算资源,使得参数量、浮点运算量、模型内存占用量分别降低 24.9%、24.4%、22.7%。由上述分析可知,在可接受范围内保持了模型轻量化,提高了模型检测精度,加快了模型检测速度,且改进模型在检测效果上达到最优。

2.2 不同主干网络性能试验

以 YOLO 11n 模型为基础模型,在保持除主干网络外其他参数一致的基础上,替换主流的特征提取主干网络,对比 6 个不同主干网络对于草莓果实目标的训练效果,分别为 GhostHGNetV2、Starnet^[21]、Mobilenetv4^[22]、SwinTransformer^[23]、Repvit^[24]、EfficientViT^[25]。如表 4 所示,使用 GhostHGNetV2 改进主干网络,使模型拥有最高的精确率和较低的

表 3 消融试验结果

Tab.3 Ablation experiment results

Ghost HGNetV2	GDFPN	WIoUv3	Lamp	精确 率/%	召回 率/%	mAP@ 0. 5/ %	mAP@ 0. 5;0. 95/ %	参数量	浮点运 算量	检测速度/ (f·s ⁻¹)	模型内存占 用量/MB
×	×	×	×	90. 3	90. 2	94. 1	68. 6	2. 62 × 10 ⁶	6. 6 × 10 ⁹	80. 6	5. 2
√	×	×	×	90. 1	89. 1	93. 9	68. 3	3. 14 × 10 ⁶	7. 7 × 10 ⁹	69. 9	6. 3
×	√	×	×	90. 3	89. 3	93. 6	68. 4	3. 66 × 10 ⁶	8. 2 × 10 ⁹	62. 5	7. 4
×	×	√	×	89. 6	90. 1	93. 9	68. 1	2. 62 × 10 ⁶	6. 6 × 10 ⁹	82. 6	5. 2
√	√	×	×	90. 8	90. 3	94. 1	69. 1	4. 37 × 10 ⁶	9. 8 × 10 ⁹	58. 4	8. 8
√	×	√	×	90. 1	89. 2	93. 9	68. 3	3. 14 × 10 ⁶	7. 7 × 10 ⁹	75. 1	6. 8
×	√	√	×	90. 5	89. 8	93. 8	68. 1	3. 66 × 10 ⁶	8. 2 × 10 ⁹	66. 2	7. 4
√	√	√	×	91. 7	90. 2	94. 2	70. 3	4. 37 × 10 ⁶	9. 8 × 10 ⁹	66. 6	8. 8
√	√	√	√	92. 5	90. 4	94. 3	70. 9	3. 28 × 10 ⁶	7. 4 × 10 ⁹	91. 5	6. 8

注:√表示使用该模块,×表示未使用该模块。

表 4 不同主干网络性能试验结果

Tab.4 Test results of different backbone networks

主干网络	精确 率/%	召回 率/%	mAP@ 0. 5/ %	mAP@ 0. 5;0. 95/ %	浮点运 算量
Starnet	89. 5	88. 2	93. 3	66. 6	5. 00 × 10 ⁹
Mobilenetv4	89. 0	87. 7	93. 4	66. 9	2. 10 × 10 ¹⁰
Repvit	89. 7	89. 7	93. 5	67. 4	1. 70 × 10 ¹⁰
EfficientViT	88. 8	89. 3	93. 1	67. 5	7. 90 × 10 ⁹
SwinTransformer	89. 1	90. 2	93. 5	67. 9	7. 76 × 10 ¹⁰
GhostHGNetV2	90. 1	89. 1	93. 9	68. 3	7. 70 × 10 ⁹

浮点运算量。网络通过使用 2 × 2 卷积学习更丰富的局部特征,并使用更小的通道数以提升分辨率加快草莓果实识别速度,采用多头注意力机制和通道加权操作提高对密集目标特征的关注能力,因此 GhostHGNetV2 取得较高的精确率 90.1%。使用卷积核 5 × 5 的 DWConv,获得更大感受野的同时降低网络的复杂度,使网络浮点运算量仅为 7.7 × 10⁹。由上述分析可知,GhostHGNetV2 主干网络检测效果最好。

2.3 不同损失函数测试试验

改进前的 YOLO 11 目标检测模型使用 CIoU 作为边界框回归损失函数,存在易过拟合的缺点。本试验同时对比损失函数 WIoUv3、PIoU^[26]、ShapeIoU^[27]、SIoU、EIoU、CIoU、DIoU 时的收敛情况,试验结果如图 10 所示。

由图 10 可知,采用 EIoU 时,收敛速度最慢,且收敛后的损失值最高;采用 PIoU、ShapeIoU、SIoU、CIoU 和 DIoU 收敛效果相近,收敛速度高于 EIoU,且收敛后损失值低于 EIoU;采用 WIoUv3 后收敛效果最好,收敛速度最快,收敛后损失值小于其他损失函数。WIoUv3 使用合理的梯度增益分配,保持梯度更新加快参数优化速度,因此使模型拥有更快的收敛速度。通过引入非单调频率调制,加大对低质量

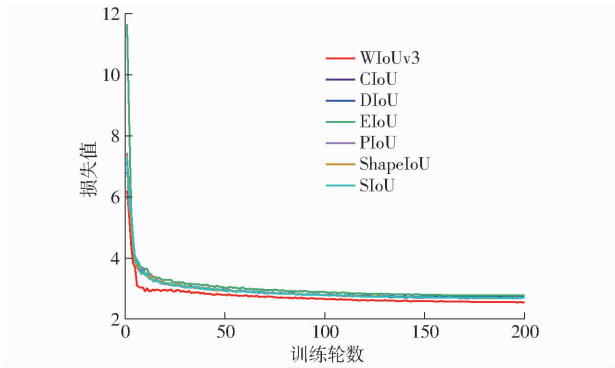


图 10 不同损失函数测试结果

Fig. 10 Test results of different loss functions

样本的惩罚因子,减少其阻碍梯度成长的影响,提高对高质量草莓图像的关注,因此降低了模型训练的损失。

2.4 模型效果分析

2.4.1 遮挡问题

为验证 YOLO 11n + GhostHGNetV2 + GDFPN 模型能否准确识别遮挡草莓果实,使用简单数据集、遮挡数据集与 YOLO 11n、DSW - YOLO 模型进行试验对比,其中 DSW - YOLO 模型对草莓遮挡程度进行分级并提高了对遮挡草莓的识别精度,不同模型在不同数据集上测试结果如表 5 所示。

由表 5 可知,在简单数据集上 YOLO 11n + GhostHGNetV2 + GDFPN 模型比 YOLO 11n、DSW - YOLO 模型识别效果稍好。在遮挡数据集中由于遮挡干扰因素的影响,3 个模型精确率都有所下降,YOLO 11n + GhostHGNetV2 + GDFPN 模型精确率、mAP@ 0. 5 均高于其他模型,证明该模型结构中 GhostHGNetV2 特征提取网络与 GDFPN 特征融合网络的改进效果在一定程度上改善了草莓果实识别的遮挡问题。

2.4.2 光照问题

YOLO 11n + GhostHGNetV2 + GDFPN + WIoUv3

表 5 不同模型在不同数据集上测试结果

Tab.5 Test results of different models on different datasets

模型	数据集	精确率	mAP@ 0.5
YOLO 11n	简单数据集	95.4	97.5
DSW-YOLO		95.1	98.8
LS-YOLO v8s		94.3	96.9
YOLO 11n + GhostHGNetV2 + GDFPN		96.1	98.7
YOLO 11n + GhostHGNetV2 + GDFPN + WIoUv3		96.5	98.4
YOLO 11n	遮挡数据集	91.3	93.6
DSW-YOLO		93.9	95.3
YOLO 11n + GhostHGNetV2 + GDFPN		95.7	98.1
YOLO 11n	光照数据集	91.1	92.3
LS-YOLO v8s		93.9	94.5
YOLO 11n + GhostHGNetV2 + GDFPN		92.5	93.8
YOLO 11n + GhostHGNetV2 + GDFPN + WIoUv3		96.1	97.7

表 6 不同模型性能试验结果

Tab.6 Results of different models test

模型	精确率/%	召回率/%	mAP@ 0.5/%	mAP@ 0.5:0.95 /%	参数量	浮点运算量	检测速度/(f·s ⁻¹)
YOLO v7tiny	88.1	90.2	93.5	64.4	6.01 × 10 ⁶	1.32 × 10 ¹⁰	53.1
YOLO v8n	88.5	90.5	93.2	67.7	3.01 × 10 ⁶	8.13 × 10 ⁹	86.3
YOLO v8s	89.8	90.9	93.9	70.9	1.11 × 10 ⁷	2.84 × 10 ¹⁰	71.1
YOLO 11n	90.3	90.2	94.1	68.6	2.62 × 10 ⁶	6.61 × 10 ⁹	95.6
YOLO 11s	91.1	90.5	94.2	70.8	9.41 × 10 ⁶	2.13 × 10 ¹⁰	83.5
YOLO 12n	88.9	89.7	93.4	67.7	2.56 × 10 ⁶	6.30 × 10 ⁹	89.4
RT-DETR-r18	90.1	90.6	92.8		1.98 × 10 ⁷	5.69 × 10 ¹⁰	66.3
SSD	90.1	55.6	75.9		2.39 × 10 ⁷	6.12 × 10 ¹⁰	42.5
Faster R-CNN	55.9	85.6	81.2		3.78 × 10 ⁷	1.40 × 10 ¹¹	12.3
LS-YOLO v8s	91.8	90.7	94.1	70.5	1.09 × 10 ⁷	3.75 × 10 ¹⁰	68.3
DSW-YOLO	91.7	90.2	94.3	67.7	2.97 × 10 ⁷	7.76 × 10 ¹⁰	17.7
YOLO v8pose	88.1	86.2	90.7	65.7	5.66 × 10 ⁶	1.16 × 10 ¹⁰	69.7
GGWL-YOLO	92.5	90.4	94.3	70.9	3.28 × 10 ⁶	7.39 × 10 ⁹	91.5

由表 6 可知, GGWL-YOLO 模型精确率、mAP@0.5、mAP@0.5:0.95 均高于其他模型,表明 GGWL-YOLO 可提高成熟草莓识别精度。与 YOLO 11n、RT-DETR-r18、LS-YOLO v8s、DSW-YOLO、YOLO v8pose 相比,浮点运算量分别降低-12.1%(增加)、86.9%、80.2%、36.3%,表明 GGWL-YOLO 仅需较小的计算力。相较于 YOLO 11n,牺牲了小幅检测速度,但换来了性能上的提升是可观的,拥有更高的检测精度,更节约内存占用量和计算资源,更适合部署在移动设备上实现实时性检测。

对模型检测效果进行分析,如图 11 所示,其中深蓝方框表示成熟草莓,青蓝方框表示未成熟草莓,红色×表示类别检测错误,绿色圆圈表示漏检。由

模型将损失函数改进为 WIoUv3,该损失函数平衡高低样本质量,加重对受光照影响的草莓图像的惩罚。为验证该模型可以改善草莓果实识别的光照影响问题进行试验,结果如表 5 所示。

由表 5 可知,在简单数据集中 4 个模型精度相近。在光照数据集中由于光照干扰因素的影响,4 个模型精度都有所下降。同样观察在光照数据集上 YOLO 11n + GhostHGNetV2 + GDFPN + WIoUv3 模型的精确率、mAP@0.5 均高于其他模型,说明该模型抗光照干扰能力更好,证明在模型中 WIoUv3 损失函数能有效降低光照对草莓果实识别的影响。

2.5 不同目标检测模型性能测试

将基于 YOLO 11n 改进的 GGWL-YOLO 模型与 YOLO v7tiny、YOLO v8n、YOLO v8s、YOLO 11n、YOLO 11s、RT-DETR-r18、SSD、Faster R-CNN 等主流目标检测模型以及文献[6-8]提出的草莓果实识别模型 LS-YOLO v8s、DSW-YOLO、YOLO v8pose 进行测试,结果如表 6 所示。

图 11 可知,本文 GGWL-YOLO 模型可成功检测出图像中所有草莓果实目标,不存在误检与漏检现象,同时拥有较高的检测精度。YOLO 11n 存在 1 处误检和 1 处漏检且精确率较低,说明 GGWL-YOLO 模型在一定程度上改善了复杂背景下成熟草莓果实误检、漏检、检测精低情况。RT-DETR-r18 存在 1 处误检且精度较低,LS-YOLO v8s 存在 1 处误检,DSW-YOLO 存在 1 处漏检,YOLO v8pose 存在 1 处误检和 1 处漏检且精度较低,通过与以上模型对比,表明 GGWL-YOLO 模型检测效果最优。

3 结论

(1) 采用 HGNetV2 构建了主干网络

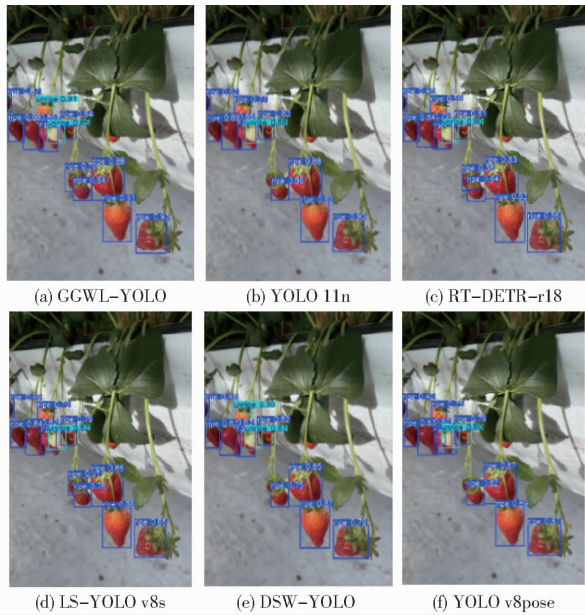


图 11 不同模型检测结果

Fig. 11 Graphs of different model detection results

GhostHNetV2,提取不同尺寸特征图,适应于多尺

度目标检测。与 Mobilenetv4、Repvit、EfficientViT、SwinTransformer 主干网络相比,精确率、召回率、平均精度均值等指标效果均有提升,有效改善了草莓果实识别遮挡问题。

(2)颈部网络引入 GDFPN 网络,并用 DySample 改进上采样模块,实现跨尺度多层次特征融合,进一步提高草莓果实检测精度,精确率提升 0.7 个百分点,mAP@ 0.5 提升 0.2 个百分点,mAP@ 0.5:0.95 提升 0.8 个百分点。

(3)使用 WIoUv3 损失函数,去除 CIoU 中的纵横比惩罚项,同时平衡高质量和低质量锚框对模型回归的影响,增强模型泛化能力,改进后精确率提升 0.9 个百分点,mAP@ 0.5:0.95 提升 0.8 个百分点,进一步缓解光线造成草莓果实识别精度低的影响。

(4)使用 Lamp 方法对改进的整体模型进行通道剪枝,减少冗余通道数保留关键的通道信息,精确率提升 0.9 个百分点,mAP@ 0.5:0.95 提升 0.5 个百分点。

参 考 文 献

[1] 李亮杰,楚宗丽,张惠妹,等. 河南省草莓生产现状调查、经济效益分析及发展建议[J]. 中国果树,2022(7):91-96.
Li Liangjie, Chu Zongli, Zhang Huimei, et al. The investigation and economic benefit analysis of strawberry production in Henan Province and suggestion for the development[J]. China Fruits, 2022(7):91-96. (in Chinese)

[2] 刘畅,王晓,李宪松,等. 我国草莓生产态势及国内外比较分析[J]. 中国果树,2023(7):136-140.
Liu Chang, Wang Xiao, Li Xiansong, et al. Strawberry production situation in China and comparative analysis at home and abroad[J]. China Fruits, 2023(7):136-140. (in Chinese)

[3] 董乃深,程泓超,应仇凯,等. 垄作草莓双臂采摘机器人设计与试验[J]. 农业机械学报,2024,55(增刊1):29-40,50.
Dong Naishen, Cheng Hongchao, Ying Qiukai, et al. Design and experiment of picking robot with dual arms for ground planting strawberries[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(S1):29-40,50. (in Chinese)

[4] 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)

[5] 王元昊,姜欢欢,罗红品,等. 基于改进 YOLOv8 算法对被遮挡柑橘的识别与定位优化[J]. 西南大学学报(自然科学版),2025,47(2):171-183.
Wang Yuanhao, Lou Huanhuan, Luo Hongpin, et al. Optimization of identification and localization of occluded citrus based on improved YOLOv8 algorithm[J]. Journal of Southwest University (Natural Science Edition), 2025, 47(2): 171-183. (in Chinese)

[6] Du X, Cheng H, Ma Z, et al. DSW-YOLO: a detection method for ground-planted strawberry fruits under different occlusion levels[J]. Computers and Electronics in Agriculture, 2023, 214: 108304.

[7] Yang S, Wang W, Gao S, et al. Strawberry ripeness detection based on YOLOv8 algorithm fused with LW-Swin Transformer[J]. Computers and Electronics in Agriculture, 2023, 215: 108360.

[8] 刘莫尘,褚镇源,崔明诗,等. 基于改进 YOLO v8-Pose 的成熟期草莓识别和果柄检测[J]. 农业机械学报,2023,54(增刊2):244-251.
Liu Mochen, Chu Zhenyuan, Cui Mingshi, et al. Red ripe strawberry recognition and stem detection based on improved YOLO v8-Pose[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023,54(S2):244-251. (in Chinese)

[9] Zhao Y, Lv W, Xu S, et al. Detrs beat yolos on real-time object detection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 16965-16974.

[10] Han K, Wang Y, Tian Q, et al. GhostNet: more features from cheap operations[C]//Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition,2020:1577-1586.

[11] Zhao G, Ge W, Yu Y. GraphFPN: graph feature pyramid network for object detection[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021: 2763-2772.

- Yu Gaohong, Wang Lei, Sun Liang, et al. Advancement of mechanized transplanting technology and equipments for field crops [J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(9): 1–20. (in Chinese)
- [21] 秦贵, 封俊, 宋卫堂, 等. 移栽机的吊杆运动分析与设计准则[J]. 农业机械学报, 2002, 33(5): 48–50.
Qin Gui, Feng Jun, Song Weitang, et al. The kinematic analysis and design criteria of the dibble-type transplanters [J]. Transactions of the Chinese Society for Agricultural Machinery, 2002, 33(5): 48–50. (in Chinese)
- [22] Kendall S L, Holmes H, White C A, et al. Quantifying lodging-induced yield losses in oilseed rape [J]. Field Crops Research, 2017, 211: 106–113.
- [23] 李鹏斌, 颜华, 王品隆, 等. 移栽机小株距栽植机构杆件优化与试验[J]. 农业机械学报, 2020, 51(增刊 2): 72–78.
Li Pengbin, Yan Hua, Wang Pinlong, et al. Optimization and test of small plant spacing planting mechanism rod for transplanter [J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(S2): 72–78. (in Chinese)
- [24] 尹大庆, 张诺一, 周脉乐, 等. 高速变姿态接苗鸭嘴式栽植机构优化设计与试验[J]. 农业机械学报, 2020, 51(6): 65–72.
Yin Daqing, Zhang Nuoyi, Zhou Maile, et al. Optimal design and experiment of high speed duckbill planting mechanism with variable catch-seedling attitude [J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(6): 65–72. (in Chinese)
- [25] 王永维, 何焯亮, 王俊, 等. 旱地蔬菜钵苗自动移栽机栽植性能试验[J]. 农业工程学报, 2018, 34(3): 19–25.
Wang Yongwei, He Chaoliang, Wang Jun, et al. Experiment on transplanting performance of automatic vegetable potseedling transplanter for dry land [J]. Transactions of the CSAE, 2018, 34(3): 19–25. (in Chinese)
-

(上接第 48 页)

- [12] Liu W, Lu H, Fu H, et al. Learning to upsample by learning to sample [C] // Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023: 6027–6037.
- [13] Tong Z, Chen Y, Xu Z, et al. Wise-IoU: bounding box regression loss with dynamic focusing mechanism [PP/OL]. V3. arXiv(2023–04–08) [2025–05–01]. <https://doi.org/10.48550/arXiv.2301.10051>.
- [14] Lee J, Park S, Mo S, et al. Layer-adaptive sparsity for the magnitude-based pruning [PP/OL]. V2. arXiv (2021–05–09) [2025–05–01]. <https://doi.org/10.48550/arXiv.2010.07611>.
- [15] Han Q, Fan Z, Dai Q, et al. On the connection between local attention and dynamic depth-wise convolution [PP/OL]. V5. arXiv (2022–08–04) [2025–05–01]. <https://doi.org/10.48550/arXiv.2106.04263>.
- [16] Lee Y, Hwang J, Lee S, et al. An energy and GPU-computation efficient backbone network for real-time object detection [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019.
- [17] Hu J, Shen L, Sun G. Squeeze-and-excitation networks [C] // Proceedings of the IEEE Conference on Computer Bision and Pattern Recognition, 2018: 7132–7141.
- [18] Wang J, Chen K, Xu R, et al. Carafe++: unified content-aware reassembly of features [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 44(9): 4674–4687.
- [19] Shi W, Caballero J, Huszar F, et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 1874–1883.
- [20] Zheng Z, Wang P, Liu W, et al. Distance-IoU loss: faster and better learning for bounding box regression [C] // Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 12993–13000.
- [21] Ma X, Dai X, Bai Y, et al. Rewrite the stars [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 5694–5703.
- [22] Qin D, Lechner C, Delakis M, et al. MobileNetV4: universal models for the mobile ecosystem [C] // European Conference on Computer Vision, 2025: 78–96.
- [23] Liu Z, Lin Y, Cao Y, et al. Swin transformer: hierarchical vision transformer using shifted windows [C] // Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021: 10012–10022.
- [24] Wang A, Chen H, Lin Z, et al. Repvit: revisiting mobile cnn from vit perspective [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 15909–15920.
- [25] Liu X, Peng H, Zheng N, et al. Efficientvit: memory efficient vision transformer with cascaded group attention [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 14420–14430.
- [26] Chen Z, Chen K, Lin W, et al. Piou loss: towards accurate oriented object detection in complex environments [C] // European Conference on Computer Vision, 2020: 195–211.
- [27] Zhang H, Zhang S. Shape-iou: more accurate metric considering bounding box shape and scale [PP/OL]. V2. arXiv (2024–01–12) [2025–05–01]. <https://doi.org/10.48550/arXiv.2312.17663>.