

陈仁凡, 谢知, 林晨. 基于YOLO-ODM的温室草莓成熟度的快速检测[J]. 华中农业大学学报, 2023, 42(4): 262-269.  
DOI: 10.13300/j.cnki.hnlkxb.2023.04.030

## 基于YOLO-ODM的温室草莓成熟度的快速检测

陈仁凡<sup>1</sup>, 谢知<sup>1,2</sup>, 林晨<sup>1</sup>

1. 福建农林大学机电工程学院, 福州 350002; 2. 福建省农业信息感知技术重点实验室, 福州 350002

**摘要** 为解决温室环境下草莓果实快速准确识别问题, 提出一种基于改进YOLOv5s的草莓成熟度检测方法。在主干中引入Shuffle\_Block作为特征提取网络, 从而实现模型轻量化。同时, 在颈部结构中使用全维度动态卷积模块(omni-dimensional dynamic convolution, ODConv), 以提高网络对草莓目标的信息挖掘能力, 降低计算量, 并进一步实现轻量化。结果显示, 改进后的YOLO-ODM(YOLO with ODConv module)模型的平均精度均值达97.4%, 模型体积是7.79 Mb, 在GPU上的单张平均检测时间仅11 ms, 浮点运算量为 $6.9 \times 10^9$ 。与原网络相比, 轻量化的YOLO-ODM方法在提高检测精度的同时, 模型大小缩减43%, 浮点运算量降低52%。以上结果表明, 该轻量化方法可快速准确地对温室环境下草莓果实的成熟度进行检测, 实现草莓的生长状态监测。

**关键词** 草莓; 成熟度; 目标检测; YOLOv5s; Shuffle\_Block; 精准识别

**中图分类号** TP391.4 **文献标识码** A **文章编号** 1000-2421(2023)04-0262-08

草莓(*Fragaria ananassa* Duch.)是一种经济价值和营养价值较高的小浆果<sup>[1]</sup>。据联合国粮食及农业组织的统计数据库(<https://www.fao.org/faostat/zh/#data/QCL>), 中国是全世界草莓种植面积最大的国家, 其2021年种植面积高达129 046 hm<sup>2</sup>。由于草莓果实一般较小且密集, 常有枝叶遮挡, 传统人工对草莓进行生长期观察或采摘都费时费力。目前, 随着我国草莓智能化产业逐渐壮大, 发展具有高准确率的果实成熟度检测系统对于确定成熟度不同的果实分布区域, 草莓生长的自动化监测以及果实采摘都具有重要意义。

传统的草莓识别方法是先获取草莓样本的物理特征, 或者结合光谱成像、电子鼻、激光诱导荧光和生物阻抗数据等, 再使用机器学习的方法进行模式识别<sup>[2-4]</sup>。Raj等<sup>[5]</sup>使用支持向量机结合全光谱图像作为输入研究草莓成熟度分类, 准确率超过98%, 但没有全光谱数据的情况下, 准确率较低。Ibba等<sup>[6]</sup>使用草莓的生物阻抗数据训练机器学习分类器并取得较好成果, 但生物阻抗数据获取过程复杂。Shao等<sup>[7]</sup>采集草莓3个成熟阶段的高光谱图, 结合最小二乘支持向量机和竞争性自适应重加权采样进行成熟度评估, 缺点是室外评估准确率较低。上述方法受

制于传统机器视觉的限制, 且草莓的特征提取繁琐, 易受环境变化的影响, 导致其检测效率低下, 鲁棒性和泛化能力较差。

近年来, 随着深度学习在农作物检测中的兴起, 传统基于特征提取的识别方法正逐步淘汰。常见的深度学习目标检测方法有R-CNN<sup>[8]</sup>、SSD<sup>[9]</sup>、RetinaNet<sup>[10]</sup>和YOLO<sup>[11]</sup>系列网络, 目前大量的研究者将深度学习应用于水果识别和定位中。Fan等<sup>[12]</sup>针对草莓全天候采摘和夜间采集图像光照度低的问题, 提出了一种暗通道增强算法来实现草莓成熟度检测。Su等<sup>[13]</sup>使用MobileNetv1作为YOLOv3的主干网络并结合SE模块, 对番茄成熟度进行分类, 其性能优于其他网络。黄彤鏊等<sup>[14]</sup>在YOLOv5中引入注意力机制, 采用 $\alpha$ -IoU损失函数, 实现在自然环境下柑橘果实的实时监测。但深度学习的模型较大, 计算成本高昂。

为解决上述问题, 本研究提出一种轻量化的基于YOLO-ODM的草莓成熟度检测方法, 在主干中引入Shuffle\_Block作为特征提取网络, 并在颈部结构中使用全维度动态卷积模块ODConv, 以提高网络对草莓目标的信息挖掘能力, 降低计算量, 充分实现模型轻量化, 旨在为温室大棚中草莓生长监测、果实

收稿日期: 2023-02-17

基金项目: 福建省自然科学基金项目(2019J01403)

陈仁凡, E-mail: 332241094@qq.com

通信作者: 谢知, E-mail: xz@fafu.edu.cn

成熟度的快速准确识别提供技术支持。

## 1 材料与方法

### 1.1 数据及制作

草莓数据集获取于百度飞浆 AI Studio 的公共数据集, 下载网址为 <https://aistudio.baidu.com/aistudio/datasetdetail/147119>。根据试验需求将公共数据集的 json 格式转化为 txt 格式。再使用 labelImg 工具对草莓果实的标注信息进行检查和校正。该草莓数据集拍摄于温室大棚内, 包含 3 100 张图片, 按照 8:1:1 的比例将数据集划分为训练集 (2 480 张)、验证集 (310 张) 和测试集 (310 张)。草莓数据集包含未成熟、半成熟和成熟草莓 3 种样本, 分别为 10 880、2 408 和 2 835 个。数据集的基本信息如表 1 所示。

表 1 草莓成熟度数据集基本信息

Table 1 Strawberry ripeness data set basic information

数据集 Data set	图像 Images	未成熟 Immaturity	半成熟 Medium	成熟 Maturity
训练集 Training set	2 480	8 806	1 916	2 246
验证集 Validation set	310	1 017	241	294
测试集 Test set	310	1 057	251	295
合计 Total	3 100	10 880	2 408	2 835

### 1.2 YOLOv5s 网络模型

为了实现草莓的快速准确定位和识别, 必须同时考虑实时性和准确性, 因此, 本研究选用权重文件较小、推理速度较快、适合部署到边缘设备的 YOLOv5s 作为基础网络。YOLOv5s 由 Backbone、Neck 和 Head 组成, 其网络结构如图 1 所示。其中 Backbone 由一系列标准卷积 CBS 模块和 C3 模块以及 SPPF 模块组成; Neck 为路径聚合网络 PANet, 并融合了主干中的 3 个不同尺度特征图 (P3、P4、P5); Head 为检测网络 Detection。

### 1.3 改进的草莓成熟度检测模型

1) 模型轻量化。本研究使用 ShuffleNetv2 中的模块 (Shuffle\_Block) 替换 YOLOv5s 中的主干特征提取网络。ShuffleNet<sup>[15-16]</sup> 引入了深度可分离卷积, 深度可分离卷积可以分解为逐深度卷积 (Depthwise convolution, DW) 和逐点卷积 (Pointwise convolution, PW)。深度可分离卷积与普通卷积计算量的比值如下:

$$\frac{D_f^2 \cdot D_k^2 \cdot M + D_f^2 \cdot M \cdot N}{D_f^2 \cdot D_k^2 \cdot M \cdot N} = \frac{1}{N} + \frac{1}{D_k^2} \quad (1)$$

其中,  $D_f$  为输入特征图尺寸,  $D_k$  为卷积核尺寸,

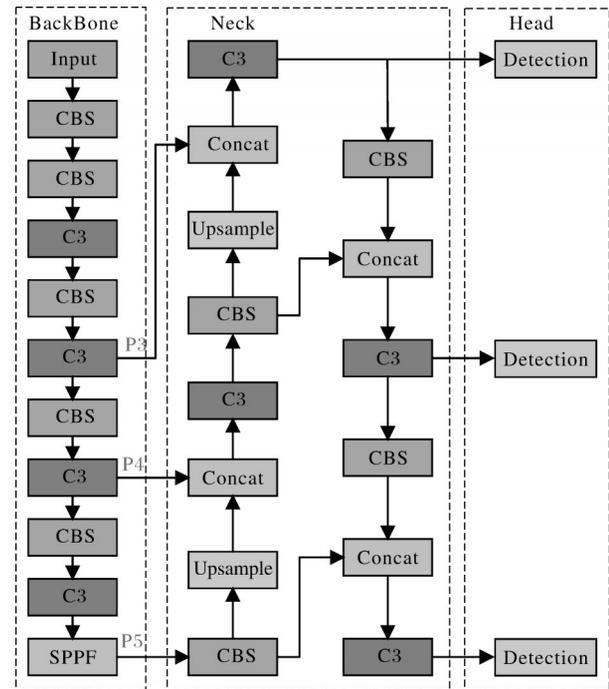


图 1 YOLOv5s 网络结构

Fig.1 YOLOv5s network structure

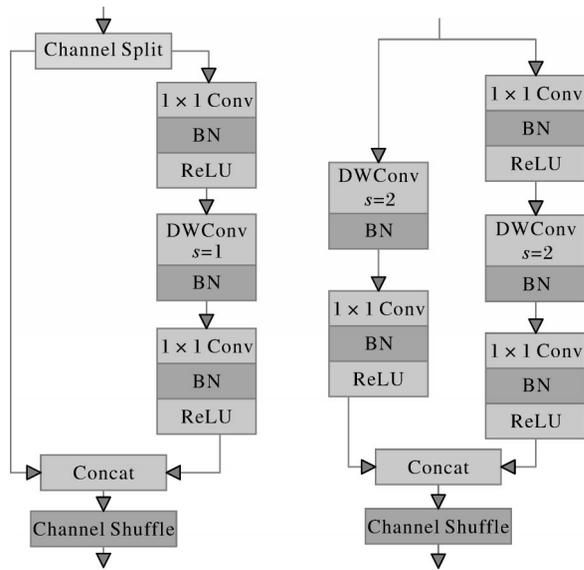
$M$  为输入通道数,  $N$  为卷积核个数。由此可见, 深度可分离卷积可以大幅降低计算量。

ShuffleNetv1 提出了通道重排技术, 通道重排技术将特征图中的通道按照一定规则分成若干组, 然后将每组中的元素重新排列, 从而增加通道之间的信息交互和融合能力, 增加模型的非线性表达能力。ShuffleNetv2 增加了通道划分 (channel split) 操作, 将输入通道划分为 2 组, 然后对其中 1 组进行卷积操作, 从而减少了计算量和模型参数。

使用的 Shuffle\_Block 如图 2 所示。左侧图为 Stride=1 的特征提取操作, 输出通道数和特征图尺寸不变; 右侧图为 Stride=2 的下采样操作, 输出通道数翻倍, 特征图尺寸减半。通过 Shuffle\_Block 搭建的 YOLO 主干网络结构如表 2 所示。

2) 全维度动态卷积 (ODConv)。使用 ODConv 替换 PANet 中的常规卷积以获取充分的上下文信息, 从而使网络更加专注于草莓果实信息。

常规的卷积核不具备动态属性, 其卷积核参数固定不变。动态卷积可以根据输入的图片动态调整卷积核参数, 使得卷积核更加专注于目标信息。为了充分利用动态卷积挖掘上下文线索的能力, 引入了 1 种全维度动态卷积<sup>[17]</sup>, 该卷积不仅赋予卷积核数动态属性, 还关注了卷积的空间大小、输入通道数和输出通道数。



左侧图为  $s=1$  的特征提取操作; 右侧图为  $s=2$  的下采样操作。  
The left figure is the feature extraction operation with  $s=1$ . The right figure is the subsampling operation with  $s=2$ .

图 2 Shuffle\_Block 结构

Fig.2 Shuffle\_Block structure diagram

常规卷积只有 1 个静态卷积核且与输入样本无关, 而传统的动态卷积核使用注意力机制将多个卷积核动态加权再线性组合, 普通动态卷积的计算如式(2)所示。而全维度动态卷积关注于核空间的所有维度, 其计算如式(3)所示。

$$y = (\alpha_{w1}W_1 + \dots + \alpha_{wi}W_i) * x \quad (2)$$

$$y = (\alpha_{wi} \odot \alpha_{\beta} \odot \alpha_{ci} \odot \alpha_{si} \odot W_1 + \dots + \alpha_{wi} \odot \alpha_{\beta} \odot \alpha_{ci} \odot \alpha_{si} \odot W_i) * x \quad (3)$$

其中,  $x$  代表输入特征图,  $y$  代表输出特征图,  $W_i$  代表第  $i$  个卷积核,  $\alpha_{wi}$  为第  $i$  个卷积核的注意力标量,  $\alpha_{si}$ 、 $\alpha_{ci}$  和  $\alpha_{\beta}$  分别表示沿空间、输入通道和输出通道的

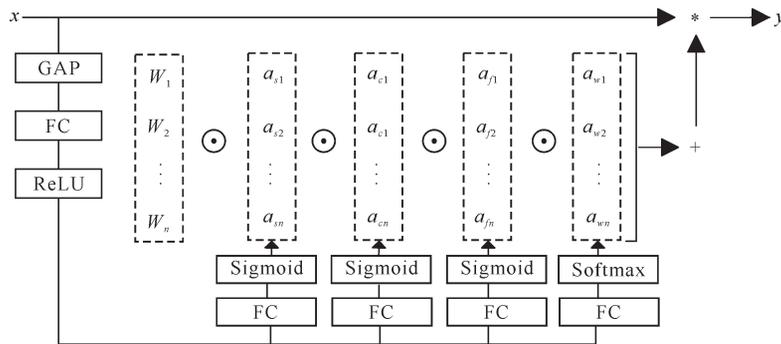


图 3 全维度动态卷积

Fig.3 Omni-dimensional dynamic convolution

3) 损失函数。YOLO 传统的边界框回归损失函数是交并比损失 (intersection over union, IoU), IoU 计算的是预测框 ( $A$ ) 和真实框 ( $B$ ) 的交叠率。IoU 的

表 2 主干网络

Table 2 Backbone network

层 Layer	模块 Block	输出通道数 Output channel	特征图尺寸 Feature map size	步长 Stride	输入到 Neck Input into Neck
0	CBR	32	$320 \times 320$	2	
1	MaxPool	32	$160 \times 160$	2	
2	Shuffle_Block	128	$80 \times 80$	2	
3	Shuffle_Block	128	$80 \times 80$	1	✓
4	Shuffle_Block	256	$40 \times 40$	2	
5	Shuffle_Block	256	$40 \times 40$	1	✓
6	Shuffle_Block	512	$20 \times 20$	2	
7	Shuffle_Block	512	$20 \times 20$	1	✓

注: ✓ 表示此处的输出特征图输入 Neck。Note: "✓" indicates that the output feature map here is input into Neck.

注意力标量 ( $i=1, 2, 3, \dots, n$ ),  $\odot$  代表沿核空间不同维度的乘法运算。ODConv 关注了核空间的 4 个维度 (核尺寸、输入通道数、输出通道数和核个数)。使用了与 SE 注意力机制相同的计算方式, 首先将输入  $x$  通过全局平均池化 (global average pooling, GAP) 压缩到与输入通道同长度的特征向量中, 然后用全连接层 FC (fully connected layers, FC) 将特征向量映射到低维空间, 再经过 ReLU 激活函数, 并分别通过 4 个 FC 层和 Sigmoid 激活函数或 Softmax 函数生成核空间的 4 个注意力标量  $\alpha_{si}$ 、 $\alpha_{ci}$ 、 $\alpha_{\beta}$  和  $\alpha_{wi}$ 。全维度动态卷积的过程如图 3 所示。

计算表达式如下:

$$IoU = \frac{A \cap B}{A \cup B} \quad (4)$$

传统的 IoU 损失有 2 个缺点: 第一, 如果预测框

和真实框不相交,那么无法计算出 2 个框之间的真实距离;第二,预测框和真实框的相交状态无法精确计算。为解决上述问题,一系列改进的 IoU 损失被提出,例如:GIoU、DIoU 和 CIoU 等。YOLOv5s 使用的 IoU 损失函数为 CIoU,CIoU 损失考虑了预测框和真实框的重叠面积、中心点距离和长宽比,但其描述的长宽比仍有一定的模糊性,EIoU 在 CIoU 的基础上拆开长宽比并用长宽的差异值取代长宽比。EIoU 损失由两框之间重叠损失  $L_{IoU}$ 、两框之间中心距离损失  $L_{dis}$  以及两框之间宽和高损失  $L_{asp}$  组成。其公式为:

$$L_{EIoU} = L_{IoU} + L_{dis} + L_{asp} = 1 - IoU + \frac{\rho^2(b, b^{gt})}{c^2} + \frac{\rho^2(w, w^{gt})}{c_w^2} + \frac{\rho^2(h, h^{gt})}{c_h^2} \quad (5)$$

其中,  $\rho$  为 2 个框中心点之间的欧氏距离,  $b$  和  $b^{gt}$  分别代表预测框和真实框,  $c$  代表 2 个框的最小外接矩形的对角线长度,  $w$  和  $h$  分别代表预测框宽和高,  $w^{gt}$  和  $h^{gt}$  为真实框的宽和高,  $c_w$  和  $c_h$  是覆盖预测框和真实框的最小外接矩形的宽和高。

考虑到边界框的回归过程中存在正负样本不平

衡问题,引入 Focal-EIoU<sup>[18]</sup> 损失,在 EIoU 的基础上结合了 Focal 损失,从梯度的角度出发,把高质量和低质量的锚框区分开以解决样本不平衡问题,Focal-EIoU 的计算表达式如下:

$$L_{Focal-EIoU} = IoU^\gamma L_{EIoU} \quad (6)$$

其中,  $\gamma$  为控制异常值抑制程度的参数。

Focal-EIoU 损失不仅考虑到预测框和真实框的中心点距离、重叠面积和长宽的差异值,还考虑到边界框回归中样本不平衡问题。因此,本研究使用 Focal-EIoU 损失。

综合以上三方面的改进,本研究提出的草莓成熟度检测模型 YOLO-ODM 如图 4 所示。该模型以图 1 所示的 YOLOv5s 作为基础网络模型;使用 Shuffle\_Block 替换主干特征提取网络以实现网络结构的轻量化;使用 ODConv 模块替换 PANet 中的普通卷积模块来加强网络的上下文信息挖掘能力,提高模型对草莓果实的检测能力,同时也进一步降低模型运算量;使用 Focal-EIoU 损失以解决样本不平衡问题。

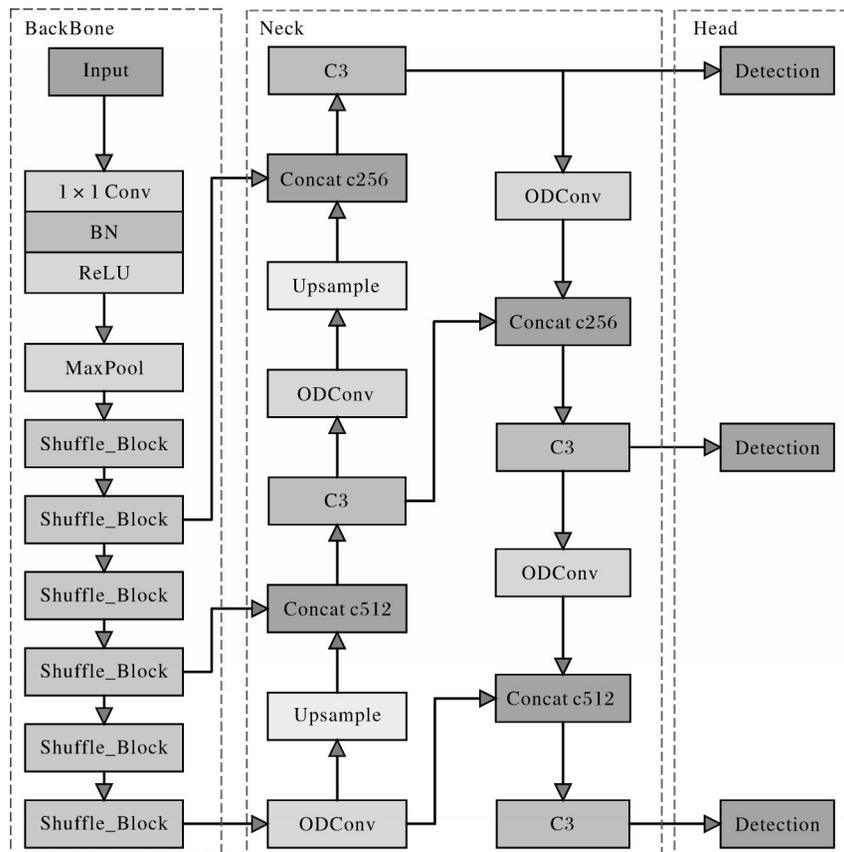


图 4 YOLO-ODM 结构

Fig.4 Structure diagram YOLO-ODM

1.4 评价指标

所用模型的评价指标为精确率(precision,  $P$ )、召

回率(recall,  $R$ )、平均精度均值(mean of average precision,  $P_{mAP}$ )、浮点运算量、模型大小、参数数量和推理

速度。其中  $P_{mA}$  基于 Precision-Recall 曲线得出,能够全面评估算法的准确性和鲁棒性。计算公式如下:

$$P = \frac{T_p}{T_p + F_p} \quad (7)$$

$$R = \frac{T_p}{T_p + F_N} \quad (8)$$

$$P_{mA} = \frac{\sum_{i=1}^c P_{i,A}}{C} \quad (9)$$

其中,  $T_p$  表示真正例,  $F_p$  表示假正例,  $F_N$  表示假负例,  $P_{i,A}$  表示每个类别的平均精度 (average precision),  $C$  表示总类别数。

### 1.5 实验环境及参数设置

本实验的硬件环境为 Intel®core™ i5-12600K 3.7 GHz, NVIDIA GeForce RTX 3060 和 32 GB 内存,软件环境为 Windows10 操作系统,使用 Pytorch1.11 深度学习框架,结合 CUDA11.3 进行训练。

本研究采用 SGD 优化器进行优化,输入图像分辨率为 640 像素 × 640 像素,学习率为 0.01,动量为 0.937,权重衰减为 0.000 5,批尺寸为 16,交并比为 0.5,使用 YOLOv5s 在草莓数据集上训练得到的权重来进行迁移学习并训练 100 个轮次 (Epoch)。

为避免试验中不同 IoU 阈值的干扰,设置 IoU 阈值为 0.5。YOLOv5s 的先验框聚类中心是基于 COCO 数据集并用 K-means 算法聚类生成得到,直接使用原有的先验框存在 2 个问题:第一,COCO 数据集包含的 80 种类别和本研究使用的草莓数据集截然不同;第二,K-means 算法生成的聚类中心有一定的随机性,造成先验框与真实框匹配效果不佳。所以本试验使用改进的 K-means++ 算法重新聚类先验框得到 9 个聚类中心,具体结果见表 3。

表 3 先验框和尺度匹配结果

Table 3 The prior bounding box and scale matching results

尺度 Size	视野 View	先验框 Anchor boxes		
20×20	大 Large	[87, 99]	[93, 164]	[117, 216]
40×40	中 Middle	[58, 98]	[60, 60]	[73, 133]
80×80	小 Small	[22, 31]	[36, 45]	[42, 72]

## 2 结果与分析

### 2.1 模型训练

模型训练过程中的评价指标变化如图 5 所示。随着训练 Epoch 不断进行,召回率上升最快,准确率

和  $P_{mA}$  也不断上升,最终 3 个评价指标在第 20 个 Epoch 左右趋于平缓,但准确率和召回率均有较大幅度波动;当模型训练到 70 个 Epoch 左右,波动开始减弱,最终准确率和召回率分别达到 92.1% 和 93.6%。同时  $P_{mA}$  在第 40 个 Epoch 左右已经达到收敛,最终的  $P_{mA}$  达到 97.4%。由于使用了域内迁移学习,本实验模型可在较短时间内达到收敛状态。

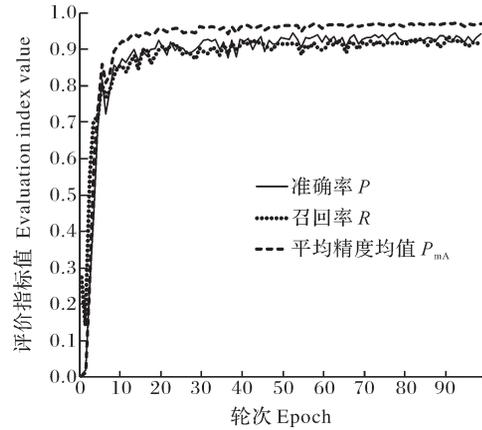


图 5 模型收敛图

Fig.5 Model convergence diagram

### 2.2 消融试验

使用原始网络与改进的 YOLO-ODM 网络在本试验草莓数据集上进行训练,消融试验结果见表 4。其中试验 1 为基线网络 YOLOv5s,试验 2~5 为添加了各种模块的网络。如表 4 所示,试验 2 使用了 Shuffle\_Block 轻量化主干网络使浮点运算量大幅度降低,但也伴随着召回率和  $P_{mA}$  降低。试验 3 使用全维度动态卷积后,模型的信息挖掘能力得到进一步提升,使得模型更加关注草莓果实特征,准确率和  $P_{mA}$  均超过基线网络,同时模型计算量也有所减少。试验 4 在使用 Focal-EIoU 交并比损失函数后使得准确率有所回升。试验 5 在 K-means++ 算法对草莓先验框重新聚类的前提下,3 个评价指标均超过基线网络,其中  $P_{mA}$  提高 0.4 百分点。

试验 2 使用轻量化模块替换主干特征提取网络后,  $P_{mA}$  有所降低,原因是使用了轻量化的结构导致模型复杂度降低,使得模型的拟合能力下降。通过采用全维度动态卷积增强了模型的信息挖掘能力和对草莓果实特征的关注力;同时使用 Focal-EIoU 缓解了样本不平衡问题;进而使用 K-means++ 算法对先验框重新聚类以更加贴近草莓真实先验框。结果表明,改进后的轻量化模型不仅提升了准确率、召回率和  $P_{mA}$ ,还明显缩小了模型浮点运算量。

表4 不同网络的消融试验结果  
Table 4 Ablation experiment results

基线网络 Baseline	序号 Order	主干 Shuffle_Block	全维度动态卷积 ODConv	Focal- EIoU	K- means++	准确率/% $P$	召回率/% $R$	平均精度均值/% $P_{mA}$	浮点运算量/ ( $\times 10^9$ ) FLOPs
	1	—	—	—	—	91.2	93.5	97.0	14.4
	2	✓	—	—	—	93.8	90.5	96.8	8.0
YOLOv5s	3	✓	✓	—	—	92.9	93.4	97.2	6.9
	4	✓	✓	✓	—	93.1	92.8	97.2	6.9
	5	✓	✓	✓	✓	92.1	93.6	97.4	6.9

注:使用改进用“✓”表示,未使用改进用“—”表示。Note: Improvements used are represented by "✓", and improvements not used are represented by "—".

### 2.3 不同检测模型的测试对比

选取各种流行的轻量级网络作为YOLOv5s的主干网络进行对比测试,结果如表5所示。同时也将改进模型与SSD、Faster R-CNN和YOLOv4-tiny进行对比测试,结果如表6所示。

表5 不同轻量级主干网络的测试对比

Table 5 Comparison of different lightweight backbone networks

网络 Network	浮点运算量/( $\times 10^9$ ) FLOPs	存储容量/Mb Memory size	平均精度均值/% $P_{mA}$
YOLOv5s	14.4	13.7	97.0
MobileNetv3	7.1	9.2	94.3
GhostNet	6.6	7.8	96.5
ShuffleNetv2	8.0	7.6	96.8

从表5可以看出,不同的轻量级网络在草莓数据集上表现出不同的检测效果。ShuffleNetv2体积最小且 $P_{mA}$ 值较高,因此,本模型选取了ShuffleNetv2

中的模块(Shuffle\_Block)搭建YOLOv5s的主干特征提取网络。表6中一系列模型均固定输入图像尺寸为640像素 $\times$ 640像素,其中SSD和Faster R-CNN使用的主干网络分别为VGG16和ResNet50。测试得到本模型的准确率、召回率和 $P_{mA}$ 分别为92.1%、93.6%和97.4%,其中 $P_{mA}$ 高于其他网络,而模型浮点运算量、体积和参数量均明显低于其他网络。

为验证YOLO-ODM模型在真实场景下草莓果实的检测效果,分别选择高光、遮挡、小目标和密集图像进行测试,如图6A~6D所示。为检验模型在更加复杂的环境背景下的识别效果,使用了与本试验数据集非同源的测试图,即图6E~6H。从图6可以看出,该模型在高光、遮挡、小目标和密集等场景下均有良好的识别效果。此外,在复杂背景下也能有效识别草莓果实,充分证明该模型具有良好的鲁棒性和泛化能力。

表6 不同模型的检测性能对比

Table 6 Test performance comparison table of different models

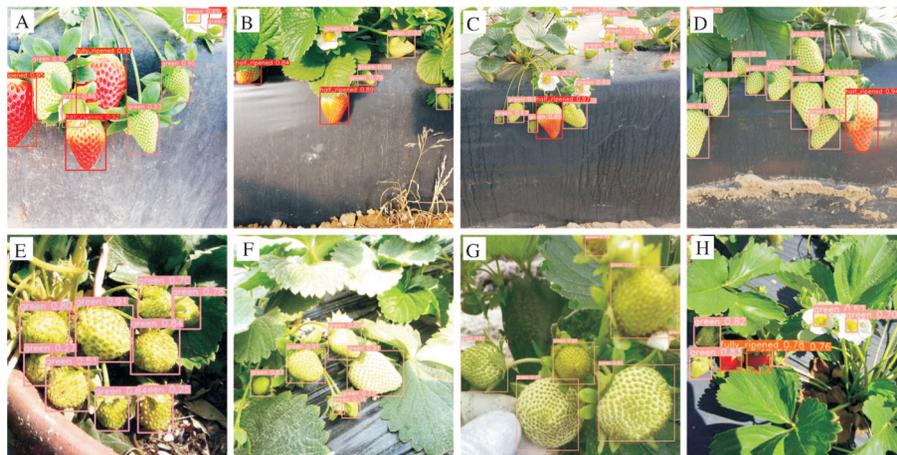
模型 Model	精确率/% $P$	召回率/% $R$	平均精度均值/% $P_{mA}$	浮点运算量/( $\times 10^9$ ) FLOPs	存储容量/Mb Memory size	检测时间/ms Detection time	参数量/( $\times 10^6$ ) Parameter count
SSD(VGG)	89.34	92.47	96.64	137.2	91.60	68.1 681.0	24.0
Faster R-CNN(ResNet50)	73.90	93.18	94.90	20.1	108.00	111.1 1800.0	13.7
YOLOv4-tiny	88.66	90.63	93.89	8.1	22.40	32.3 136.2	5.9
YOLOv5s	91.20	93.50	97.00	14.4	13.70	12.1 101.5	7.0
本研究 This study	92.10	93.60	97.40	6.9	7.79	11.0 99.4	3.9

注:检测时间列的左右子列分别为在GPU和CPU上的检测情况。Note: The left and right subcolumns of the detection time are the detection on the GPU and CPU, respectively.

## 3 讨论

针对温室环境下草莓果实的准确快速定位问题,本研究提出1种改进的YOLO-ODM草莓成熟度检测模型,可以结合监控摄像头实时监测草莓生长

状态,也可以和嵌入式草莓分拣系统相结合进行实时草莓分类。以Shuffle\_Block作为主干特征提取网络使模型轻量化,用全维度动态卷积充分挖掘草莓数据集的上下文信息,通过Focal-EIoU损失函数缓解了草莓样本的不平衡问题,并使用K-means++算



A:高光目标 Highlight objects; B:遮挡目标 Occluded objects; C:小目标 Small objects; D:密集目标 Dense objects; E、F、G、H:与本试验数据集非同源的目标 Non-homologous targets in the data set of this experiment.

图6 YOLO-ODM模型不同场景下的检测效果

Fig.6 YOLO-ODM model with different scene detection effects

法对草莓数据集的先验框进行重新聚类,提升预测框的准确度。改进后的YOLO-ODM检测模型的 $P_{mA}$ 为97.4%,在GPU上的单张平均检测时间为11 ms,浮点运算量为 $6.9 \times 10^9$ ,模型体积大小为7.79 Mb。相比于原始网络 $P_{mA}$ 提高0.4个百分点,模型大小缩减了43%,浮点运算量减少52%。试验结果表明,在其他杂乱环境下,改进后的检测模型在满足高精度识别草莓果实的同时,计算量大幅减少,提高了检测效率和实时性,并且模型较小,有利于在实际应用中的边缘设备上部署。与张小花等<sup>[19]</sup>的基于EfficientDet-D1的草莓检测模型(1 636张样本图,数据集划分为未成熟和成熟2个类别,CPU平均检测时间0.34 s)相比,本研究数据样本数量(3 100)更大,进一步将数据集划分成未成熟、半成熟和成熟3个类别,并且在CPU上的平均检测时间约为0.1 s。因此,本研究模型的数据集划分更加科学,具有更强的泛化能力和实时性。

由于本研究是在计算机上进行的,检测模型未置于实际嵌入式设备上进行测试,所以真实环境中可能会存在一定误差,后面还需将模型部署到移动端设备,以完成真实场景下的草莓检测。下一步应继续增加更多复杂背景以及其他场景下的草莓果实图片,进一步加强训练和测试,以提升模型的泛用性,更好地满足实际检测的需求。

## 参考文献 References

- [1] 刘小刚,范诚,李加念,等.基于卷积神经网络的草莓识别方法[J].农业机械学报,2020,51(2):237-244.LIU X G, FAN C, LI J N, et al. Identification method of strawberry based on convolutional neural network[J]. Transactions of the CSAM, 2020, 51(2):237-244(in Chinese with English abstract).
- [2] ZHAO Y S, GONG L, HUANG Y X, et al. Robust tomato recognition for robotic harvesting using feature images fusion[J/OL]. Sensors, 2016, 16(2): 173[2023-02-17]. <https://pubmed.ncbi.nlm.nih.gov/26840313/>. DOI: 10.3390/s16020173.
- [3] 宋凯,任晓哲.基于YCbCr色彩空间的玉米叶部病斑的图像分割[J].农业工程学报,2008,24(S2):202-205.SONG K, REN X Z. Image segmentation of disease speckle of corn leaf based on YCbCr color space[J]. Transactions of the CSAE, 2008, 24(S2):202-205(in Chinese with English abstract).
- [4] 赵嘉玮.基于SVM算法的目标分类筛选方法研究[J].无线互联科技,2022,19(23):123-125.ZHAO J W. Research on strawberry classification and screening method based on SVM algorithm[J]. Wireless internet technology, 2022, 19(23): 123-125 (in Chinese with English abstract).
- [5] RAJ R, COSGUN A, KULIĆ D. Strawberry water content estimation and ripeness classification using hyperspectral sensing[J/OL]. Agronomy, 2022, 12(2): 425[2023-02-17]. <https://doi.org/10.3390/agronomy12020425>.
- [6] IBBA P, TRONSTAD C, MOSCETTI R, et al. Supervised binary classification methods for strawberry ripeness discrimination from bioimpedance data[J]. Scientific reports, 2021, 11(1):1-13.
- [7] SHAO Y Y, WANG Y X, XUAN G T, et al. Assessment of strawberry ripeness using hyperspectral imaging[J]. Analytical letters, 2020, 54(10):1547-1560.
- [8] HE K M, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN[C]//2017 IEEE International Conference on Computer Vision Workshops (ICCV Workshops 2017), October 22-29, 2017, Venice, Italy. New York: IEEE, 2017:2961-2969.
- [9] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector[DB/OL]. arXiv, 2016:1512.02325[2023-02-17]. <https://doi.org/10.48550/arXiv.1512.02325>.

- [10] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection [DB/OL]. arXiv, 2018: 1708.02002 [2023-02-17]. <https://doi.org/10.48550/arXiv.1708.02002>.
- [11] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: optimal speed and accuracy of object detection [DB/OL]. arXiv, 2020: 2004.10934 [2023-02-17]. <https://doi.org/10.48550/arXiv.2004.10934>.
- [12] FAN Y C, ZHANG S Y, FENG K, et al. Strawberry maturity recognition algorithm combining dark channel enhancement and YOLOv5 [J/OL]. Sensors, 2022, 22 (2): 419 [2023-02-17]. <https://doi.org/10.3390/s22020419>.
- [13] SU F, ZHAO Y P, WANG G H, et al. Tomato maturity classification based on SE-YOLOv3-MobileNetV1 network under nature greenhouse environment [J/OL]. Agronomy, 2022, 12 (7): 1638 [2023-02-17]. <https://doi.org/10.3390/agronomy12071638>.
- [14] 黄彤颖, 黄河清, 李震, 等. 基于YOLOv5改进模型的柑橘果实识别方法 [J]. 华中农业大学学报, 2022, 41 (4): 170-177. HUANG T B, HUANG H Q, LI Z, et al. Citrus fruit recognition method based on the improved model of YOLOv5 [J]. Journal of Huazhong Agricultural University, 2022, 41 (4): 170-177 (in Chinese with English abstract).
- [15] ZHANG X Y, ZHOU X Y, LIN M X, et al. ShuffleNet: An extremely efficient convolutional neural network for mobile devices [DB/OL]. arXiv, 2017: 1707.01083 [2023-02-17]. <https://doi.org/10.48550/arXiv.1707.01083>.
- [16] MA N N, ZHANG X Y, ZHENG H T, et al. ShuffleNet v2: Practical guidelines for efficient CNN architecture design [DB/OL]. arXiv, 2018: 1807.11164 [2023-02-17]. <https://doi.org/10.48550/arXiv.1807.11164>.
- [17] LI C, ZHOU A J, YAO A B. Omni-dimensional dynamic convolution [C/OL] // ICLR 2022 conference. [S.l.]: ICLR, 2022 [2023-02-17]. <https://openreview.net/forum?id=DmpCfq6Mg39>.
- [18] ZHANG Y F, REN W Q, ZHANG Z, et al. Focal and efficient IoU loss for accurate bounding box regression [J]. Neurocomputing, 2022, 506: 146-157.
- [19] 张小花, 李浩林, 李善军, 等. 基于EfficientDet-D1的草莓快速检测及分类 [J]. 华中农业大学学报, 2022, 41 (6): 262-269. ZHANG X H, LI H L, LI S J, et al. Rapid detection and classification of strawberries based on EfficientDet-D1 [J]. Journal of Huazhong Agricultural University, 2022, 41 (6): 262-269 (in Chinese with English abstract).

## YOLO-ODM based rapid detection of strawberry ripeness in greenhouse

CHEN Renfan<sup>1</sup>, XIE Zhi<sup>1,2</sup>, LIN Chen<sup>1</sup>

1. College of Mechanical and Electronic Engineering, Fujian Agriculture and Forestry University, Fuzhou 350002, China;

2. Fujian Province Key Laboratory of Agricultural Information Perception Technology, Fuzhou 350002, China

**Abstract** An improved YOLOv5s-based method for rapidly detecting strawberry ripeness was proposed to solve the problem of rapid and accurate identification of strawberry fruits in greenhouse. The Shuffle\_Block was introduced as a feature extraction network in the backbone to lightweight the model. Meanwhile, the omni-dimensional dynamic convolution (ODConv) module was used in the neck structure to enhance the information mining ability of model for strawberry targets, reduce computational complexity, and further achieve lightweight. The results showed that the average precision of the improved YOLO-ODM model reached 97.4%. The model size is 7.79 Mb. The average detection time on the GPU is only 11 ms per image, and the floating-point operations are  $6.9 \times 10^9$  FLOPs. Compared with the original network, the lightweighted YOLO-ODM method improved the accuracy of detection while reducing model size by 43% and floating-point operations by 52%. It is indicated that the lightweighted method can rapidly and accurately detect the ripeness of strawberry fruit in greenhouse, monitor the growth status of strawberries.

**Keywords** strawberry; ripeness; object detection; YOLOv5s; Shuffle\_Block; accurate identification

(责任编辑: 赵琳琳)