

基于改进 YOLO v5s 的马铃薯种薯芽眼检测方法

张万枝^{1,2} 曾祥^{1,2} 刘树峰^{1,2} 穆桂脂^{1,2} 张弘毅^{1,2} 郭壮壮^{1,2}

(1. 山东农业大学机械与电子工程学院, 泰安 271018; 2. 山东省农业装备智能化工程实验室, 泰安 271018)

摘要: 芽眼检测是马铃薯种薯智能切块首先要解决的问题,为实现种薯芽眼精准高效检测,提出了一种基于改进 YOLO v5s 的马铃薯种薯芽眼检测方法。首先通过加入 CBAM 注意力机制,加强对马铃薯种薯芽眼图像的特征学习和特征提取,同时弱化与芽眼相似的马铃薯种薯表面背景对检测结果的影响。其次引入加权双向特征金字塔 BiFPN 增加经骨干网络提取的种薯芽眼原始信息,为不同尺度特征图赋予不同权重,使得多尺度特征融合更加合理。最后替换为改进的高效解耦头 Decoupled Head 区分回归和分类,加快模型收敛速度,进一步提升马铃薯种薯芽眼检测性能。试验结果表明,改进 YOLO v5s 模型准确率、召回率和平均精度均值分别为 93.3%、93.4% 和 95.2%;相比原始 YOLO v5s 模型,平均精度均值提高 3.2 个百分点,准确率、召回率分别提高 0.9、1.7 个百分点;不同模型对比分析表明,改进 YOLO v5s 模型与 Faster R-CNN、YOLO v3、YOLO v6、YOLOX 和 YOLO v7 等模型相比有着较大优势,平均精度均值分别提高 8.4、3.1、9.0、12.9、4.4 个百分点。在种薯自动切块芽眼检测试验中,改进 YOLO v5s 模型平均召回率为 91.5%,相比原始 YOLO v5s 模型提高 17.5 个百分点。本文方法可为研制马铃薯种薯智能切块芽眼识别装置提供技术支持。

关键词: 马铃薯芽眼; 目标检测; YOLO v5; 深度学习; 改进解耦头

中图分类号: S223.1; TP391.4 文献标识码: A 文章编号: 1000-1298(2023)09-0260-10

OSID:



Detection Method of Potato Seed Bud Eye Based on Improved YOLO v5s

ZHANG Wanzhi^{1,2} ZENG Xiang^{1,2} LIU Shufeng^{1,2} MU Guizhi^{1,2} ZHANG Hongyi^{1,2} GUO Zhuangzhuang^{1,2}

(1. College of Mechanical and Electronic Engineering, Shandong Agricultural University, Taian 271018, China
2. Shandong Provincial Engineering Laboratory of Agricultural Equipment Intelligence, Taian 271018, China)

Abstract: The first problem to be solved in potato cutting fast is the detection of potato seed bud eyes, an improved YOLO v5s-based potato seed bud eye detection method was proposed to improve seed potato eye detection performance. Firstly, by adding the CBAM attention mechanism, the feature learning and feature extraction of the potato bud eye images were strengthened. The influence of the potato surface background similar to the bud eyes on the detection results was weakened. Secondly, the weighted bidirectional feature pyramid BiFPN was introduced to increase the original information of bud eyes extracted by the backbone network and assign weights to feature maps of different scales, making multi-scale feature fusion more reasonable. Finally, it was replaced with an improved and efficient Decoupled Head to distinguish between regression and classification, speed up the convergence speed of the model, and further improve the performance of potato bud eye detection. The test results showed that the precision, recall rate, and average precision of the improved algorithm were 93.3%, 93.4% and 95.2%, respectively, which was 3.2 percentage points higher than that of the original algorithm in the mean average precision, and the precision and recall rate were improved by 0.9 and 1.7 percentage points. The comparative analysis of different algorithms showed that this algorithm had absolute advantages compared with Faster R-CNN, YOLO v3, YOLO v6, YOLOX and YOLO v7 algorithms. The mAP was increased by 8.4 percentage points, 3.1 percentage points, 9.0 percentage points, 12.9 percentage points and 4.4 percentage points. In the actual detection application, the average recall rate of the improved algorithm was 91.5%, which was 17.5 percentage points higher than that of the original algorithm, and the missed detection rate was reduced. The method can provide technical support for the next step in the development of a sprout-eye identification device for the intelligent cutting of potato seed potatoes.

Key words: potato bud eye; object detection; YOLO v5; deep learning; improved decoupling header

0 引言

我国年均马铃薯种植面积为 $6.0 \times 10^6 \text{ hm}^2$ 以上,总产量位居世界第一,是第一生产大国,但并不是生产强国^[1-3]。我国与美国、荷兰等发达国家相比,马铃薯种植机械化水平偏低。国内传统的切块作业方式仍然处于人工切块,存在耗时耗力、效率低下和出苗率难以保证等问题,为提高种植机械化水平,智能切块取代人工切块将显得尤为重要,而马铃薯种薯芽眼检测是实现智能切块的前提。

为实现马铃薯种薯芽眼检测,国内部分学者通过传统视觉技术进行了相关研究。田海韬等^[4]在彩色和灰度空间中分割出芽眼,再将二者结合进行数学形态学处理得到芽眼标记结果。李玉华等^[5]基于色饱和度三维几何特征进行马铃薯芽眼检测。吕钊钦等^[6]基于 Gabor 特征进行马铃薯图像滤波处理,剔除马铃薯边界连通区域进行芽眼区域提取,完成芽眼的检测。张金敏等^[7]提出使用局部二值模式(LBP)提取特征结合支持向量机(SVM)进行分类,完成对马铃薯芽眼检测。YANG 等^[8]在多光谱图像中结合监督多阈值分割模型和 Canny 边缘检测器,完成马铃薯芽眼检测。以上传统视觉检测是基于颜色、纹理、形状和其他手动标注的特征进行训练,虽然训练模型简单,但模型的泛化性和普适性较差,不利于在不同环境的实际应用。

近年来随着深度学习快速发展,基于深度学习的目标检测技术在农业工程领域也得到了广泛的应用^[9-12]。该技术主要分为基于候选框的双阶段目标检测算法和更快检测速度的单阶段目标检测算法。双阶段目标算法最具有代表的是 R-CNN 系列,其中包括 R-CNN、Fast R-CNN、Faster R-CNN 和 R-FCN^[13-16]。XI 等^[17]提出基于改进 Faster R-CNN 双阶段的目标检测算法,对发芽马铃薯芽眼检测精度可达 96.32%。杨森等^[18]利用迁移学习和复合特征字典方法对 Faster R-CNN 训练,检测马铃薯病害叶片的斑块区域。双阶段目标检测算法先生成一系列样本候选框,再根据生成的候选框区域进行目标分类和回归,检测精度较优,但这也带来检测时间长的弊端,不适用于农业工程实时检测。单阶段目标检测算法与双阶段检测算法相比,候选框和检测同时进行,且兼顾检测速度和精度。此类最具有代表的是 YOLO 系列算法^[19-22]。陈志伟等^[23]采用 YOLO v3 单阶段目标检测算法对马铃薯种薯芽眼检测,并使用图像数据增强和迁移学习思想的方法。王相友等^[24]针对马铃薯中土块石块检测问题对 YOLO v4 模型进行通道剪枝达到轻量化

的目的。史方青等^[25]使用 YOLO v3 网络模型将多个遮挡、机械损伤、虫眼的芽眼及杂质的样本加入数据集中,实现对种薯芽眼的检测。

目前基于深度学习的芽眼目标检测虽然取得了一定进展,但针对马铃薯种薯芽眼特征一般仅使用 YOLO 系列原始模型,未区别于一般检测物体特征展开研究。由于马铃薯种薯芽眼随机分布在表面,与整体背景较为相似,不易进行检测,同时种薯芽眼位于表面凹陷处,对于凹陷区域较浅的芽眼经过多次下采样后会造成特征信息丢失,在特征融合处效果较差。YOLO 系列原始模型并未针对以上种薯芽眼特征存在的问题进行具体分析,为实现种薯芽眼精准高效检测,本文提出一种基于改进 YOLO v5s 的马铃薯种薯芽眼检测方法。马铃薯种薯芽眼与马铃薯整体背景较为相似,为抑制除芽眼外的马铃薯背景,使用 CBAM 卷积注意力机制去除冗余特征,增强有用信息,使其对芽眼特征信息更加关注;马铃薯种薯芽眼特征融合效果差,使用加权双向特征金字塔网络 BiFPN,引入骨干原始特征信息,同时为不同尺度特征图赋予不同权重,使得马铃薯种薯芽眼多尺度特征融合更加合理;引入解耦头,将检测头中分类问题和回归任务分开计算,加快模型的收敛速度,以进一步提升马铃薯芽眼检测性能。

1 马铃薯种薯芽眼检测模型

1.1 改进的 YOLO v5s 模型

YOLO v5 网络根据模型深度倍数 (Depth_Multiple) 和层通道倍数 (Width_Multiple) 分为 n、s、m、l 和 x 5 种版本,同时模型的复杂度也依次增加,在牺牲检测速度的情况下提高检测精度。考虑到农业工程应用以实时检测为主,本文选择基于改进的 YOLO v5s 进行马铃薯芽眼检测。YOLO v5s 网络模型分为 4 部分:输入端 (Input)、骨干网络 (Backbone)、颈部网络 (Neck)、预测端 (Prediction Head)。输入端通过 Mosaic 数据增加、自适应锚框计算和自适应图像缩放,以此来达到丰富背景信息,计算最佳锚点框以及原图缩放到统一标准尺寸的目的。骨干网络由 CBS、Bottleneck CSP 和 SPPF 组成,在最新的版本中使用 6×6 的 Conv 卷积模块替换了 Focus 切片操作,使用串行的 SPPF 替换了并行的 SPP,减少计算量,在相同的精度条件下提高检测速度。颈部网络采用 FPN + PAN 结构,FPN 传递高层的语义特征到低层特征中,增强语义信息利于分类,添加自底向上的 PAN 弥补并加强了位置信息。预测端将颈部网络输出的 3 个特征图作为网络的输

出,维度分别为 80×80 、 40×40 和 20×20 ,分别用于检测小目标、中目标和大目标。

本文首先针对马铃薯种薯芽眼与背景相似的问题,在检测头前端引入 CBAM 注意力机制,实现抑制背景的干扰,更加突出种薯芽眼的特征。其次将

FPN + PAN 特征融合更换为更加有效的多尺度特征融合方法 BiFPN,解决特征信息融合尺度不一致问题。最后把传统耦合头替换为解耦头,使得模型收敛速度更快,提升模型整体性能。改进后的 YOLO v5s 芽眼检测整体网络架构如图 1 所示。

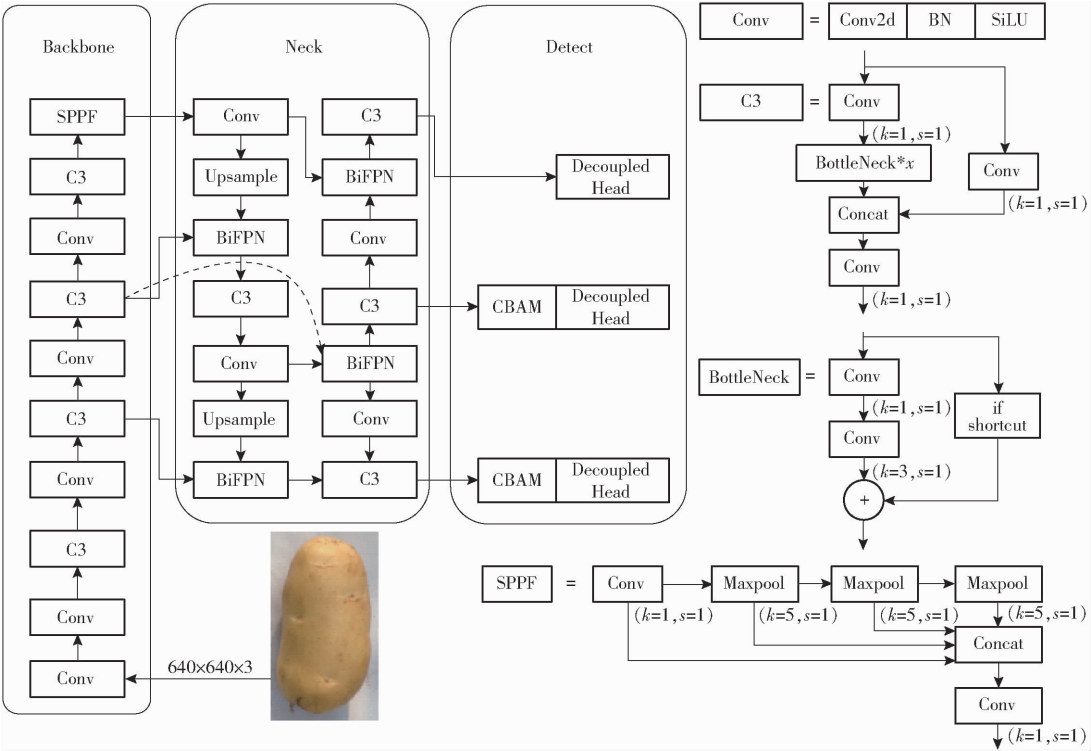


图 1 改进后 YOLO v5s 整体网络架构

Fig.1 Overall network architecture of improved YOLO v5s

1.2 CBAM 注意力机制

为了从复杂的特征信息中获取芽眼关键特征信息,并考虑到芽眼与马铃薯整体较为相似的特性,本文引入了卷积注意力机制 (Convolutional block attention module, CBAM)^[26] 进行强化芽眼目标信息,抑制背景信息。CBAM 结合了特征通道和特征空间 2 个维度的注意力机制,分别负责寻找重要信息和确定重要信息位置,两者结合对特征信息进行权重分配,将有限的计算资源分配给重要的目标,其模块结构如图 2 所示,图中 H 为输入图像高度, W 为输入图像宽度, C 为特征通道数。

输入特征图 F 首先经过通道注意力模块,分别进行 2 个并行的全局最大池化 (Global max pooling, GMP) 和全局平均池化 (Global average pooling, GAP),得到 2 个 $C \times 1 \times 1$ 的特征图,然后将其分别送入共享的 2 层全连接层 MLP 中,激活函数为 ReLU,得到卷积输出的 2 个一维矢量进行 element-wise 的加和操作,最后使用 Sigmoid 激活,计算式为

$$M_c(F) = \sigma(W_1(W_0(F_{avg}^C)) + W_1(W_0(F_{max}^C)))$$

(1)

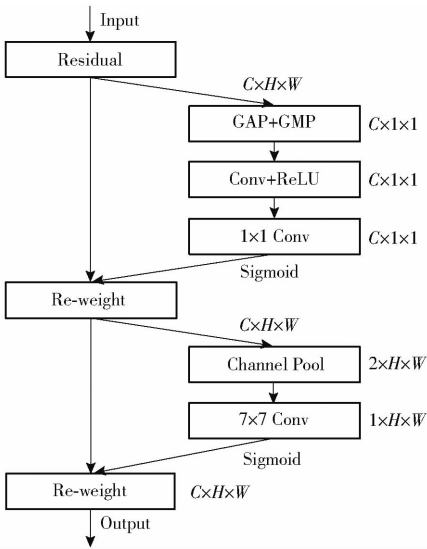


图 2 CBAM 注意力模块结构

Fig.2 CBAM attention module architecture

式中 M_c ——通道注意力模块特征图
 σ ——Sigmoid 激活函数
 F_{avg} ——平均池化特征图
 F_{max} ——最大池化特征图
 W_i ——输入的共享权重

经通道注意力模块处理后的特征图与输入特征图 F 相乘,得到特征图 F' 作为空间注意力模块的输入,首先通过基于通道注意力模块的池化处理,得到 2 个 $1 \times H \times W$ 的特征图再进行相加,然后将经过 7×7 Conv 卷积以及 Sigmoid 激活函数生成的特征图 M_s 与特征图 F' 相乘,计算式为

$$M_s(F) = \sigma(f^{7 \times 7}(F_{avg}^S, F_{max}^S)) \quad (2)$$

1.3 加权双向特征金字塔 BiFPN

为了提升经下采样得到不同种薯芽眼特征图的融合能力,本文将 YOLO v5 的 Neck 网络中的 FPN + PAN 特征融合更换为加权双向特征金字塔网络 (Bidirectional feature pyramid network, BiFPN), BiFPN 的主要贡献为:高效的双向跨尺度连接和带权重的特征融合机制。

PANet 虽然提供自底向上的特征融合层,加强了位置信息,但输入是经 FPN 结构处理的特征信息,缺乏骨干网络提取到的特征中的原始特征信息。为解决该问题, BiFPN 在同一层中的原始输入节点和输出节点之间添加了一条额外的特征融合边,旨在不增加太多成本的情况下融合更多的原始特征,同时删除了只有一条输入边且没有特征融合的节点,从而简化特征融合网络,其结构如图 3 所示。

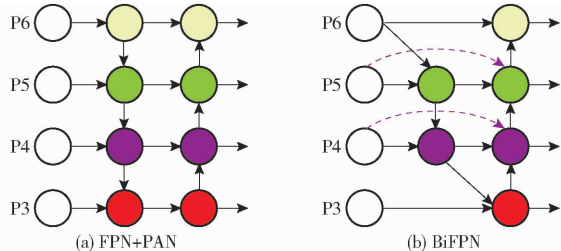


图 3 不同特征金字塔结构

Fig. 3 Pyramid structure with different features

传统的特征融合只是简单地将不同的特征图相加,并未考虑到不同分辨率的特征图对融合输入时贡献度也有所不同,因此简单地对其相加并不是最佳操作。为了解决这一问题, BiFPN 引入快速归一化特征融合机制,对不同尺度的芽眼特征赋予不同的权重,其计算式为

$$O = \sum_i \frac{w_i}{\varepsilon + \sum_j w_j} I_i \quad (3)$$

其中, w_i 表示输入特征 I_i 对应一个可学习的权重,为了保证 $w_i \geq 0$, 在每个 w_i 前采用 ReLU 激活函数,并将初始学习率 ε 设置为 0.001 以此来避免数值不稳定。P4 层的 2 个特征融合过程计算式为

$$P_4^{id} = \text{Conv} \left(\frac{w_1 P_4^{in} + w_2 \text{Resize}(P_5^{in})}{w_1 + w_2 + \varepsilon} \right) \quad (4)$$

$$P_4^{out} = \text{Conv} \left(\frac{w'_1 P_4^{in} + w'_2 P_4^{id} + w'_3 \text{Resize}(P_3^{out})}{w'_1 + w'_2 + w'_3 + \varepsilon} \right) \quad (5)$$

式中 P_i^{in} ——第 i 级的输入特征
 P_i^{id} ——第 i 级的中间特征
 P_i^{out} ——第 i 级的输出特征
 Resize——上采样或下采样操作
 Conv()——卷积

1.4 Decoupled Head 解耦检测头

为了解决种薯芽眼分类问题和回归任务的冲突,加快模型的收敛速度,提高模型的整体性能,本文将 YOLO v5 中 Head 耦合头替换为 Decoupled Head 解耦头结构。SONG 等^[27]提出目标检测算法的分类问题和回归任务是互斥的关系,即分类和回归的关注点不同,分类更加关注目标的纹理内容,回归更加关注目标的边缘信息。原始 YOLO v5 的检测头为耦合头,将分类问题和回归任务合并在一起计算,分类和回归分支共享参数。但这会存在一个问题,若网络模型过分关注分类的信息,则更加明显的马铃薯种薯芽眼特征响应值会很高,只框选出芽眼明显部分,对定位造成很大影响,导致芽眼检测位置不准确,同时也会引起训练过程中收敛速度较慢的问题。为解决该问题,引入 Decoupled Head 解耦头结构可以将分类问题和回归任务分开计算,使二者关注目标更加明确,加快收敛速度。

YOLOX 在检测头上使用 Anchor-free 的解耦头对分类和回归分支解耦计算,同时新增 2 个额外的 3×3 CBS 卷积模块,虽提升了检测精度,但在一定程度上增加了模型复杂度。为此本文在 YOLOX 基础上进行了精简设计,如图 4 所示,删除了 1×1 Conv 降维卷积,采用检测头的维度由主干和颈部的宽度系数共同缩放,并删除了 2 个分支中额外一个 3×3 CBS 卷积模块。删除这 2 部分的目的是因为过多卷积层会造成计算量增加以及部分信息

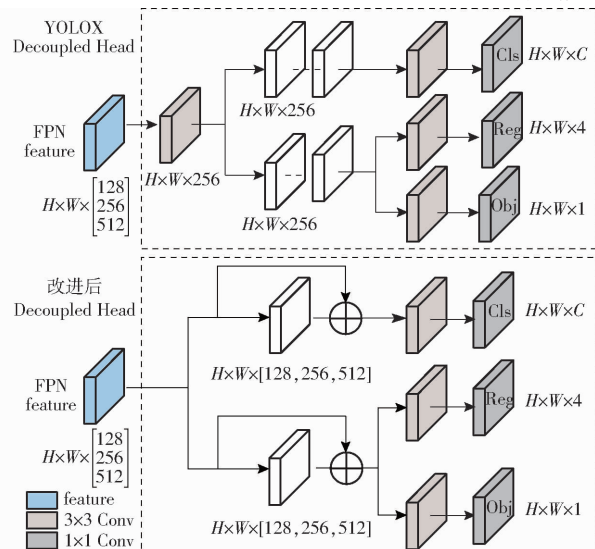


图 4 改进后的 Decoupled Head

Fig. 4 Improved Decoupled Head

丢失,同时为每个 CBS 模块添加残差边,旨在降低网络优化难度,并在不增加成本的情况下提高检测头性能。

2 数据准备与模型训练

2.1 数据集构建

数据集采集地点为山东省滕州市其祥马铃薯种植合作社,样本种薯品种为“荷兰 15 号”和“中薯 3 号”。种薯采集样本的品质会对芽眼的检测造成重要影响,因此选择无病害、干腐、病斑、虫眼等储存状况良好的马铃薯种薯作为采集样本。采集后的数据集经人工剔除冗余、像素质量差的图像,共 931 幅种薯图像。为避免数据集样本过少影响其模型训练效果和泛化性能,本文采用翻转、旋转、改变亮度和添加噪声的方式进行数据扩充,经过扩充及处理后的数据集总共 3 936 幅图像,部分样本图像如图 5 所示。利用 LabelImg 软件对马铃薯种薯图像中芽眼进行标注,标注后的数据集按 9:1 的比例划分为训练集(3 542 幅)和验证集(394 幅)。



图 5 增强后部分样本图像
Fig.5 Part of sample image after enhancement

2.2 试验环境与参数设置

试验过程在 Win 11 操作系统下进行,处理器型号为 12th Gen Intel (R) Core (TM) i7 - 12700K 3.60 GHz,显卡型号为 Nvidia GeForce RTX 3080Ti。深度学习框架为 Pytorch,编程平台为 PyCharm,软件环境为 CUDA 11.6 和 CuDNN 8.0.5 版本,所用编程语言为 Python 3.8。

为提高算法性能,减少过拟合,训练采用的批量大小(batch size)为 16,初始学习率设置为 0.01,动量设置为 0.937,训练总轮数设置为 200,后续不同算法对比试验中也将保持统一变量。

2.3 评价指标

为了通过检测结果来评价改进后 YOLO v5s 模型的性能,选用准确率 (Precision, P)、召回率 (Recall, R)、平均精度均值 (Mean average precision,

mAP) 和帧率 (Frames per second, FPS) 作为评价指标。

3 试验结果与分析

3.1 YOLO v5 版本对比试验

YOLO v5s、YOLO v5m、YOLO v5l 和 YOLO v5x 的 4 个版本在速度和精度之间提供了不同的权衡,以适应不同的计算能力和实时性的需求,为了选择合适的版本作为基线模型,进行了不同版本的 YOLO v5 对比试验,试验结果如表 1 所示。

表 1 YOLO v5 不同版本性能比较
Tab.1 Performance comparison of different versions of YOLO v5

模型	平均精度 均值/%	参数量	浮点运 算数	帧率/ ($\text{f}\cdot\text{s}^{-1}$)	内存占用 量/MB
YOLO v5s	92.0	7.010×10^6	15.8	35.7	13.6
YOLO v5m	92.3	2.103×10^7	50.2	20.0	42.4
YOLO v5l	93.3	4.660×10^7	113.9	15.1	93.7
YOLO v5x	93.3	8.719×10^7	216.9	14.3	175.0

可以看出,YOLO v5 各版本模型的平均精度均值相差并不大的情况下,在参数量 (Parames) 和浮点运算数 (GFLOPs) 方面,其他 3 个版本模型比 YOLO v5s 模型增加 3 ~ 13 倍,同时也引起 FPS 降低和权重文件增大,综合性能不如 YOLO v5s 模型。分析其模型复杂度增加但平均精度均值并未有明显提升的原因可能是芽眼目标较为单一,当使用 m、l 和 x 版本模型时,由于网络所含残差结构的个数和卷积核的个数不断增加,对芽眼特征提取操作次数也越多,造成芽眼特征信息丢失,使得检测精度并未有大幅度提升。

综上所述,本研究所使用的马铃薯种薯芽眼数据集相对较小,目标样本较单调,因此使用较浅层的网络即可充分提取特征,同时考虑后续在实际应用中需要对模型进行边缘端部署等原因,选用复杂度较低且具有较高检测精度的 YOLO v5s 模型作为基线模型进行改进。

3.2 不同解耦头应用于 YOLO v5s 的性能比较

为了评价改进后的解耦头与 YOLOX 解耦头在 YOLO v5s 模型中的性能优劣,分别单独加入 YOLO v5s 模型中进行试验,结果如表 2 所示。可以看出,改进后解耦头的参数量相较于 YOLOX 解耦头参数量 (Parames) 有所下降,并且浮点运算数 (GFLOPs) 减少约 50%,说明删除 1×1 Conv 降维卷积和一个 3×3 CBS 卷积模块后可降低计算量的有效性;同时改进解耦头在计算量下降的情况下,为每个 CBS 模块添加残差边使得平均精度均值仍然有 0.4 个百分点的提

升,这进一步说明本文解耦头有着更优的性能。

表2 不同解耦头性能比较

Tab.2 Performance comparison of different decoupling heads

模型	参数量	浮点运算数	平均精度均值/%
YOLOX 解耦头	14 325 462	56.2	94.0
改进解耦头	13 209 558	27.1	94.4

改进前后平均精度均值曲线如图6所示。由图6可以看出,使用改进后的解耦头代替原模型耦合头能达到收敛速度更快、性能更好的效果。

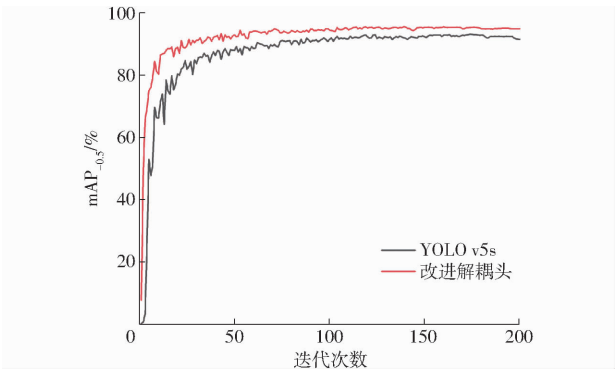


图6 改进前后mAP曲线对比

Fig.6 Comparison of mAP curves before and after improvement

3.3 消融试验

为实现马铃薯种薯芽眼精准高效检测,本文基于原始YOLO v5s模型,提出更有效的网络模型:加入卷积注意力模块CBAM,引入BiFPN特征融合结构,更换解耦头,并分别进行4组消融试验验证,试验结果如表3所示。

表3 消融试验结果

Tab.3 Results of ablation experiments

CBAM	BiFPN	解耦头	P/%	R/%	mAP/%
-	-	-	92.4	91.7	92.0
√	-	-	92.9	93.0	94.4
-	√	-	92.9	92.7	93.6
-	-	√	92.5	92.7	94.4
√	√	√	93.3	93.4	95.2

注:“√”表示添加结构;“-”表示未添加结构。

从表3中可知,加入CBAM注意力机制后,检测模型的mAP提高2.4个百分点,表明加入CBAM后模型能够更加准确地检测种薯芽眼信息;引入BiFPN特征融合结构,能够融合更多主干网络提取的原始特征并且让特征融合更加合理,mAP较原始模型提高1.6个百分点;将YOLO v5s原始模型的耦合头替换为改进后的解耦头,mAP提高2.4个百分点,表明替换解耦头能够解决种薯芽眼特征在分类问题和回归任务的冲突,这不仅使得模型加快收敛

速度,而且还提高了整体检测性能。本文提出的改进模型,准确率达到93.3%,较原始模型增长0.9个百分点,同时召回率增加1.7个百分点,mAP提升3.2个百分点。以上数据说明了与YOLO v5s相比,本文模型有更强的特征提取能力和多尺度融合性能,在马铃薯种薯芽眼检测上有更好的表现。

为了更加直观地展示本文模型与YOLO v5s原始模型对马铃薯种薯芽眼特征的关注程度不同,文中使用Grad-CAM^[28]中类激活热力图分别对2个模型输出层进行可视化分析,即通过类激活热力图中某一区域的亮度来表示该区域在预测输出过程所占的权重,颜色鲜亮区域面积越大,表示预测输出关注度越高,反之亦然。2个模型的类激活热力图如图7所示,可以看出本文模型对马铃薯种薯芽眼的特征预测输出权重分配更高,对于不显眼的马铃薯种薯芽眼也能得到更多关注,避免造成漏检;而且本文模型相较于原始模型在同一处的马铃薯种薯芽眼检测置信度更高,具有更好的鲁棒性。通过类激活热力图可以看出本文模型更适合马铃薯种薯芽眼特征检测。

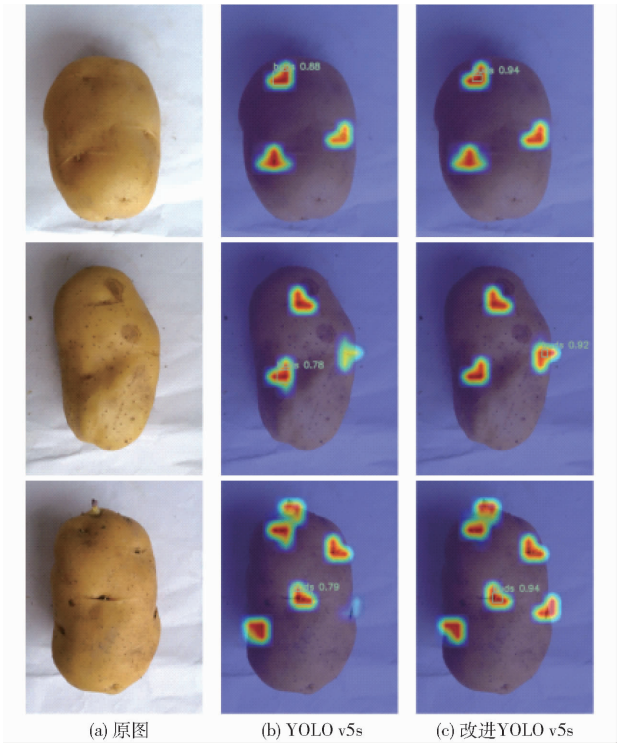


图7 类激活热力图

Fig.7 Class activation heatmaps

3.4 与其他模型对比试验

为了验证本文所提出模型的有效性,将本文改进后的YOLO v5s模型与其他主流检测模型进行对比试验,所有模型均在同一数据集和同一训练设备下采用控制变量原则进行试验,结果如表4所示。

表 4 改进模型与其他检测模型对比

Tab.4 Comparison of improved models with other detection models

模型	P/%	R/%	mAP/ %	内存占用 量/MB	帧率/ (f·s ⁻¹)
Faster R - CNN			86.8	326.0	15.6
YOLO v3	81.9	94.8	92.1	123.4	16.3
YOLO v5s	92.4	91.7	92.0	13.6	35.7
YOLO v6	83.3	84.6	86.2	38.7	30.1
YOLOX	88.3	78.3	82.3	34.3	24.3
YOLO v7	86.2	88.2	90.8	71.3	13.5
改进 YOLO v5s	93.3	93.4	95.2	25.8	32.4

由表 4 可以看出,与 YOLO 其他系列原始模型相比,YOLO v5s 原始模型综合性能要优于其他系列模型,证明了本文选择 YOLO v5s 作为基线模型的可行性。在针对马铃薯种薯芽眼检测,改进 YOLO v5s 模型准确率最优,与 Faster R - CNN、YOLO v3、YOLO v6、YOLOX 和 YOLO v7 等模型相比,mAP 分别提高 8.4、3.1、9.0、12.9、4.4 个百分点,并且在模型内存占用量和帧率上也具有较大优势。召回率对比中,改进 YOLO v5s 模型略低于 YOLO v3 模型,主要是因为准确率和召回率存在一定的此消彼长关系,当模型降低了对芽眼检测要求,芽眼漏检数量将会减少,代表着预测出了更多的种薯芽眼,但这也会增加芽眼检测错误数量。由于本文模型加入高效解耦头,导致模型内存占用量与原始模型相比由13.6 MB 增加到 25.8 MB、检测帧率由 35.7 f/s 降至 32.4 f/s,但由于在种薯芽眼检测中更加注重准确率的提升,且帧率的略微下降并不影响实际应用,因此本文模型在兼顾准确率和召回率情况下,仍能保证最高平均精度均值和较高的帧率,综合性能最优。

3.5 种薯自动切块芽眼检测试验

为了评估本文所提出改进 YOLO v5s 模型在实际种薯自动切块芽眼检测的性能,在自制的马铃薯种薯自动切块装置进行试验,如图 8 所示。将种薯放置在 CCD 相机下,经相机采集到种薯图像后,传输给计算机中改进的 YOLO v5s 模型进行芽眼检测,便于下一步根据检测到的种薯芽眼位置,通过切刀决策方法控制切刀运动实现种薯切块。从种薯表面检测出芽眼是自动切块的重要前提,同时考虑到种薯图像受拍摄设备及环境影响,且实际应用中希望检测得到更多的种薯芽眼,相较于准确率,召回率能更好地反映本文模型的性能。因此在试验结果中主要对召回率进行计算,并将置信度阈值设置为 0.35。

试验结果如表5所示。可以看出,3种不同品

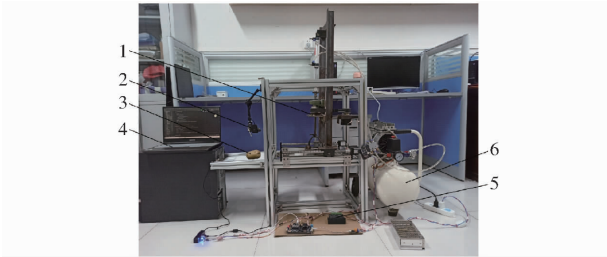


图 8 马铃薯种薯自动切块装置
Fig.8 Potato seed potato automatic block cutting machine structure device

1. 切块装置 2. CCD 相机 3. 马铃薯种薯 4. 计算机 5. 步进电机驱动器 6. 供气装置

表 5 马铃薯种薯芽眼检测结果统计

Tab.5 Statistics of potato bud eye detection results

种薯品种	模型	马铃薯种薯芽眼数量			召回率/
		总数	检出数	漏检数	%
荷兰 15 号	YOLO v5s	87	67	20	77.0
	改进 YOLO v5s	87	79	8	90.8
中薯 3 号	YOLO v5s	149	103	46	69.1
	改进 YOLO v5s	149	136	13	91.3
尤金 885	YOLO v5s	157	119	38	75.8
	改进 YOLO v5s	157	145	12	92.4

种的种薯进行实际检测应用中,改进 YOLO v5s 模型漏检数明显下降,不同种薯品种的平均召回率达到 91.5%,相较于原始模型提高 17.5 个百分点,说明了本文模型在实际种薯自动切块芽眼检测中拥有更强的多尺度性能以及对芽眼具有更多的关注度,能够降低漏检率,从而保证后续切刀根据芽眼位置做出姿态调整的精准性,避免切伤芽眼造成种薯浪费。

图 9 分别为荷兰 15 号、尤金 885 和中薯 3 号 3 种品种种薯的漏检与误检情况。可以看出,YOLO v5s 原始模型存在不同程度的漏检现象(图中椭圆形所示),即使较为明显的芽眼特征也未被检测出,其原因是因为种薯芽眼特征在种薯表面不明显,在通过主干网络下采样以及 FPN 结构上采样过程后会造特征信息丢失,导致在检测头处因特征信息不足引起漏检。而本文所提出的改进 YOLO v5s 模型能够有效避免该问题,并且能够检测到种薯边缘部分较小区域芽眼,漏检率较低。同时不同品种检测过程中 YOLO v5s 原始模型也存在主要以下 3 种误检情况:图 9a 原始模型将种薯表面的破损表皮误检为芽眼;图 9b 原始模型将种薯表面的裂痕误检为芽眼;图 9c 原始模型将种薯表面存在的泥土误检为芽眼。使用改进后的 YOLO v5s 模型进行检测可有效改善上述误检情况,降低误检率。

通过将本文所提模型应用到实际种薯自动切块

装置中,与原始模型相比能够检测到更多芽眼,并且能够有效避免误检情况的发生,这说明位于检测头前端的 CBAM 注意力机制增加了对芽眼特征的权重分配,能够从相似背景下检测出芽眼;BiFPN 中增添额外的特征融合边,引入了更多下采样过程中的原始特征信息,丰富了芽眼特征信息,同时带权重的

特征融合机制也为芽眼特征分配到更多的权重,避免了特征信息的丢失;改进的解耦头能够解决分类和回归的冲突,分别更好地关注了芽眼目标的纹理内容和边缘信息。综上可知,本文提出的改进 YOLO v5s 模型能够更好地应用于实际种薯自动切块芽眼检测。

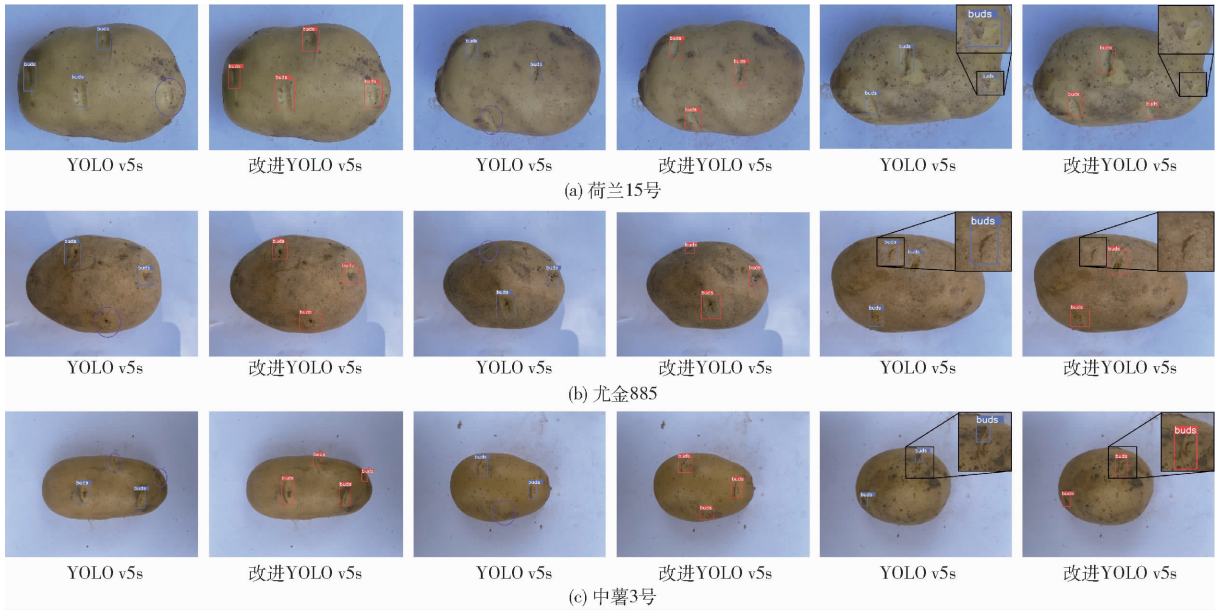


图 9 实际应用检测结果对比
Fig. 9 Comparison of actual application test results

4 结论

(1)为实现种薯芽眼精准高效检测,本文提出一种基于改进 YOLO v5s 的马铃薯种薯芽眼检测方法。首先通过加入 CBAM 注意力机制,加强对马铃薯种薯芽眼图像的特征学习和特征提取,同时弱化与芽眼相似的马铃薯种薯表面背景对检测结果的影响;其次引入加权双向特征金字塔 BiFPN 增加经骨干网络提取的种薯芽眼原始信息,为不同尺度特征图赋予不同权重,使得多尺度特征融合更加合理;最后替换为改进的高效解耦头 Decoupled Head 区分回归和分类,加快模型收敛速度,进一步提升马铃薯种薯芽眼检测性能。

(2)试验结果表明,改进 YOLO v5s 模型准确率、召回率和平均精度均值分别为 93.3%、93.4% 和 95.2%;相比原始 YOLO v5s 模型,平均精度均值提高 3.2 个百分点,准确率和召回率分别提高 0.9、1.7 个百分点;不同模型对比分析表明,改进 YOLO v5s 模型与 Faster R - CNN、YOLO v3、YOLO v6、YOLOX 和 YOLO v7 等模型相比有着较大优势,平均精度均值分别提高 8.4、3.1、9.0、12.9、4.4 个百分点。

(3)在种薯自动切块芽眼检测试验中,改进 YOLO v5s 模型平均召回率为 91.5%,相比原始 YOLO v5s 模型提高 17.5 个百分点,本文提出的改进 YOLO v5s 模型能够更好地应用于实际种薯自动切块芽眼检测。

参 考 文 献

[1] 李紫辉,温信宇,吕金庆,等. 马铃薯种植机械化关键技术与装备研究进展分析与展望[J]. 农业机械学报,2019,50(3): 1 - 16.
LI Zihui, WEN Xinyu, LÜ Jinqing, et al. Analysis and prospect of research progress on key technologies and equipments of mechanization of potato planting[J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(3): 1 - 16. (in Chinese)

[2] 罗其友,伦闰琪,高明杰,等. 2021—2025 年我国马铃薯产业高质量发展战略路径[J]. 中国农业资源与区划, 2022, 43(3): 37 - 45.
LUO Qiyu, LUN Runqi, GAO Mingjie, et al. Strategy path of high quality development of potato industry in China during from 2021 to 2025[J]. Chinese Journal of Agricultural Resources and Regional Planning, 2022, 43(3): 37 - 45. (in Chinese)

[3] 王相友,祝珊,李学强,等. 定向排列纵横切分马铃薯种薯切块机设计与试验[J]. 农业机械学报, 2020, 51(6): 334 - 345.

- WANG Xiangyou, ZHU Shan, LI Xueqiang, et al. Design and experiment of directional arrangement vertical and horizontal cutting of seed potato cutter[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(6): 334 – 345. (in Chinese)
- [4] 田海韬, 赵军, 蒲富鹏. 马铃薯芽眼图像的分割与定位方法[J]. 浙江农业学报, 2016, 28(11): 1947 – 1953.
TIAN Haitao, ZHAO Jun, PU Fupeng. A method for recognizing potato's bud eye[J]. Acta Agricultural Zhejiangensis, 2016, 28(11): 1947 – 1953. (in Chinese)
- [5] 李玉华, 李天华, 牛子孺, 等. 基于色饱和度三维几何特征的马铃薯芽眼检测[J]. 农业工程学报, 2018, 34(24): 158 – 164.
LI Yuhua, LI Tianhua, NIU Ziru, et al. Potato bud eyes recognition based on three-dimensional geometric features of color saturation[J]. Transactions of the CSAE, 2018, 34(24): 158 – 164. (in Chinese)
- [6] 吕钊钦, 亓协腾, 张万枝, 等. 基于 Gabor 特征的马铃薯图像芽眼识别[J]. 农机化研究, 2021, 43(2): 203 – 207.
LÜ Zhaoqin, QI Xieteng, ZHANG Wanzhi, et al. Buds recognition of potato images based on Gabor feature[J]. Journal of Agricultural Mechanization Research, 2021, 43(2): 203 – 207. (in Chinese)
- [7] 张金敏, 杨添玺. 基于 LBP 与 SVM 的马铃薯芽眼检测[J]. 山东农业大学学报(自然科学版), 2020, 51(4): 744 – 748.
ZHANG Jinmin, YANG Tianxi. Potato bud recognition based on LBP and SVM[J]. Journal of Shandong Agricultural University (Natural Science Edition), 2020, 51(4): 744 – 748. (in Chinese)
- [8] YANG Y, ZHAO X, HUANG M, et al. Multispectral image based germination detection of potato by using supervised multiple threshold segmentation model and Canny edge detector[J]. Computers and Electronics in Agriculture, 2021, 182: 106041.
- [9] 张振国, 邢振宇, 赵敏义, 等. 改进 YOLOv3 的复杂环境下红花丝检测方法[J]. 农业工程学报, 2023, 39(3): 162 – 170.
ZHANG Zhenguo, XING Zhenyu, ZHAO Minyi, et al. Detecting safflower filaments using an improved YOLOv3 under complex environments[J]. Transactions of the CSAE, 2023, 39(3): 162 – 170. (in Chinese)
- [10] 黄硕, 周亚男, 王起帆, 等. 改进 YOLOv5 测量田间小麦单位面积穗数[J]. 农业工程学报, 2022, 38(16): 235 – 242.
HUANG Shuo, ZHOU Ya'nan, WANG Qifan, et al. Measuring the number of wheat spikes per unit area in fields using an improved YOLOv5[J]. Transactions of the CSAE, 2022, 38(16): 235 – 242. (in Chinese)
- [11] 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检测[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)
- [12] 张慧春, 张萌, 边黎明, 等. 基于 YOLO v5 的植物叶绿素含量估测与可视化技术[J]. 农业机械学报, 2022, 53(4): 313 – 321.
ZHANG Huichun, ZHANG Meng, BIAN Liming, et al. Estimation and visualization of chlorophyll content in plant based on YOLO v5[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(4): 313 – 321. (in Chinese)
- [13] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014: 580 – 587.
- [14] GIRSHICK R. Fast R – CNN[C] // Proceedings of the IEEE International Conference on Computer Vision, 2015: 1440 – 1448.
- [15] REN S, HE K, GIRSHICK R, et al. Faster R – CNN: towards real-time object detection with region proposal networks[J]. IEEE Trans. Pattern Anal. Mach. Intell., 2017, 39(6): 1137 – 1149.
- [16] DAI J F, LI Y, HE K M, et al. R – FCN: object detection via region-based fully convolutional networks[C] // Proc. of the 30th International Conference on Neural Information Processing Systems, 2016: 379 – 387.
- [17] XI R, HOU J, LOU W. Potato bud detection with improved Faster R – CNN[J]. Transactions of the ASABE, 2020, 63(3): 557 – 569.
- [18] 杨森, 冯全, 张建华, 等. 基于深度学习与复合字典的马铃薯病害检测方法[J]. 农业机械学报, 2020, 51(7): 22 – 29.
YANG Sen, FENG Quan, ZHANG Jianhua, et al. Identification method for potato disease based on deep learning and composite dictionary[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(7): 22 – 29. (in Chinese)
- [19] 武星, 齐泽宇, 王龙军, 等. 基于轻量化 YOLOv3 卷积神经网络的苹果检测方法[J]. 农业机械学报, 2020, 51(8): 17 – 25.
WU Xing, QI Zeyu, WANG Longjun, et al. Apple detection method based on light-YOLOv3 convolutional neural network[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(8): 17 – 25. (in Chinese)
- [20] 李菊霞, 李艳文, 牛帆, 等. 基于 YOLOv4 的猪只饮食行为检测方法[J]. 农业机械学报, 2021, 52(3): 251 – 256.
LI Juxia, LI Yanwen, NIU Fan, et al. Pig diet behavior detection method based on YOLOv4[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(3): 251 – 256. (in Chinese)
- [21] MA J, LU A, CHEN C, et al. YOLOv5-lotus an efficient object detection method for lotus seedpod in a natural environment [J]. Computers and Electronics in Agriculture, 2023, 206: 107635.
- [22] 王小荣, 许燕, 周建平, 等. 基于改进 YOLOv7 的复杂环境下红花采摘识别[J]. 农业工程学报, 2023, 39(6): 169 – 176.
WANG Xiaorong, XU Yan, ZHOU Jianping, et al. Safflower picking recognition in complex environments based on an improved YOLOv7[J]. Transactions of the CSAE, 2023, 39(6): 169 – 176. (in Chinese)
- [23] 陈志伟, 张万枝, 张涛, 等. 基于 YOLOv3 算法的马铃薯种薯芽眼检测[J]. 农机化研究, 2022, 44(11): 19 – 23, 30.
CHEN Zhiwei, ZHANG Wanzhi, ZHANG Tao, et al. Potato seed tuber sprout eye detection based on YOLOv3 algorithm[J]. Journal of Agricultural Mechanization Research, 2022, 44(11): 19 – 23, 30. (in Chinese)

[24] 王相友,李晏兴,杨振宇,等. 基于改进 YOLO v4 模型的马铃薯中土块石块检测方法[J]. 农业机械学报,2021,52(8): 241 – 247,262.
WANG Xiangyou,LI Yanxing,YANG Zhenyu, et al. Detection method of clods and stones from impurified potatoes based on improved YOLO v4 algorithm[J]. Transactions of the Chinese Society for Agricultural Machinery,2021,52(8):241 – 247,262. (in Chinese)

[25] 史方青,王虎林,黄华. 基于卷积神经网络的马铃薯芽眼检测检测研究[J]. 中国农机化学报,2022,43(6):159 – 165.
SHI Fangqing, WANG Hulin, HUANG Hua. Research on potato buds detection and recognition based on convolutional neural network[J]. Journal of Chinese Agricultural Mechanization, 2022, 43(6): 159 – 165. (in Chinese)

[26] 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.

[27] SONG G, LIU Y, WANG X. Revisiting the sibling head in object detector[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020; 11563 – 11572.

[28] SELVARAJU R R, COGSWELL M, DAS A, et al. Grad-cam: visual explanations from deep networks via gradient-based localization[C] // Proceedings of the IEEE International Conference on Computer Vision, 2017; 618 – 626.

(上接第 253 页)

[17] ZHU H C, HUANG W, LIU H Y. Loess terrain segmentation from digital elevation models based on the region growth method [J]. Phys. Geogr. , 2018, 39(1): 51 – 66.

[18] SOFIA G, BAILLY J S, CHEHATA N, et al. Comparison of pleiades and LiDAR digital elevation models for terraces detection in farmlands[J]. IEEE J. Sel. Top Appl. Earth Observ. Remote Sens. , 2016, 9(4): 1567 – 1576.

[19] VON GIOI R G, JAKUBOWICZ J, MOREL J M, et al. LSD: a fast line segment detector with a false detection control[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 32(4): 722 – 732.

[20] 侯建才,李占斌,李勉. 黄土高原丘陵沟壑区一副区小流域淤地坝系效益分析——以王茂沟小流域为例[J]. 水土保持研究, 2007, 61(2): 34 – 36.
HOU Jiancai, LI Zhanbin, LI Mian. Benefit analysis of check dams system in small watershed in first sub-region of hilly-gully loess plateau—a case study in Wangmaogou watershed[J]. Research of Soil and Water Conservation, 2007, 61(2): 34 – 36. (in Chinese)

[21] GARBRECHT J, MARTZ L W. Digital elevation model grid size, landscape representation, and hydrologic simulations-comment[J]. Water Resources Research, 1996, 32(5): 1461 – 1462.

[22] TAROLLI P. High-resolution topography for understanding earth surface processes: opportunities and challenges [J]. Geomorphology, 2014, 216: 295 – 312.

[23] 萨出拉,刘万青. DEM 分辨率对小流域地形水文特征提取的影响[J]. 测绘工程, 2015, 24(5): 72 – 76.
SA Chula, LIU Wanqing. Effects of DEM resolution on extracting topographical and hydrological parameters of small drainage basin[J]. Engineering of Surveying and Mapping, 2015, 24(5): 72 – 76. (in Chinese)

[24] LINDSAY J B. Pit-centric depression removal methods[C] // Proceedings of the GEOMORPHOMETRY 2020 Conference, 2020.