doi:10.6041/j.issn.1000-1298.2025.11.046

基于 SAM - CenterNet 的马铃薯与杂质分类识别方法

张　杨[1]　刘发英[2]　杨振宇[1,3]　文永双[1]　耿良月[1]　魏忠彩[4]　李学强[3]

(1. 山东理工大学机械工程学院, 淄博 255000; 2. 山东理工大学电气与电子工程学院, 淄博 255000;
3. 山东省马铃薯生产装备智能化工程技术研究中心, 德州 253600;
4. 山东理工大学农业工程与食品科学学院, 淄博 255000)

摘要: 针对马铃薯收获入库过程中杂质(土块、石块)需要识别并剔除的问题,提出基于 CenterNet 改进的薯杂分类识别方法。首先,搭建六通道薯杂识别与杂质剔除装置,采集源图像;其次,在主干网络特征提取阶段,将传统卷积下采样替换为 SPD - Conv 下采样,提高识别网络对薯杂不同尺度的细粒度识别能力,解决土石块与部分马铃薯存在特征相似导致漏识别的问题;然后,为了抑制背景无效特征,加入 ACmix 注意力机制,增强薯杂表面的有效特征,解决马铃薯表面和图像背景粘连泥土导致误识别的问题;最后,构建多层级特征融合结构,获取更加丰富的薯杂表面纹理特征信息,解决模型输出特征单一而导致识别精度低的问题。构建薯杂图像数据集对改进后的薯杂分类识别方法进行试验,试验结果表明,改进后算法的模型参数量为 1.013×10^7,前向推理时间为 27 ms,平均精度均值为 97.5%。与原 CenterNet 模型相比,模型参数量增加 1.66×10^6,前向推理时间增加 2 ms,但平均精度均值提高 8.2 个百分点。该薯杂分类识别方法能够满足马铃薯收获入库过程中杂质识别并剔除的技术要求。

关键词: 马铃薯; 杂质; SAM - CenterNet; ACmix; 深度学习

中图分类号: S226.5　　**文献标识码**: A　　**文章编号**: 1000-1298(2025)11-0480-10　　**OSID**:

Recognition Method for Classification of Potatoes and Impurities Based on SAM - CenterNet

ZHANG Yang[1]　LIU Faying[2]　YANG Zhenyu[1,3]　WEN Yongshuang[1]　GENG Liangyue[1]
WEI Zhongcai[4]　LI Xueqiang[3]

(1. *School of Mechanical Engineering, Shandong University of Technology, Zibo* 255000, *China*
2. *School of Electrical and Electronic Engineering, Shandong University of Technology, Zibo* 255000, *China*
3. *Shandong Intelligent Engineering Technology Research Center of Potato Production Equipment, Dezhou* 253600, *China*
4. *School of Agricultural Engineering and Food Science, Shandong University of Technology, Zibo* 255000, *China*)

Abstract: In order to address the problem that impurities (soil and stones) need to be identified and rejected in the process of potato harvesting and warehousing, an improved potato impurity classification and identification method based on CenterNet was proposed. Firstly, a six-channel potato impurity recognition and impurity removal device was built to capture the source images. Secondly, in the feature extraction stage of the backbone network, the traditional convolutional downsampling was replaced with SPD - Conv downsampling to improve the recognition network's ability to recognize the fine-grained recognition of potatoes and impurities at different scales, and solve the problem of missed recognition due to the existence of feature similarity between soil and stone blocks and some potatoes. Then, in order to suppress the background invalid features, the ACmix attention mechanism was added to enhance the effective features on the potato and impurity surfaces, and to solve the problem of mud adhering to the potato surface and the image background leading to misrecognition. Finally, a multilevel feature fusion structure was constructed to obtain richer texture feature information of the potato and impurity surfaces, and solve the problem of low recognition accuracy due to the single output feature of the model. The potato and impurity image dataset was constructed to test the improved potato and impurity classification

收稿日期: 2024-07-13　修回日期: 2024-12-02

基金项目: 国家自然科学基金项目(52105266)和山东省科技型中小企业创新能力提升工程项目(2021TSGC1420)

作者简介: 张杨(1999—),男,硕士生,主要从事智能农业装备技术研究,E-mail: 2250771733@qq.com

通信作者: 杨振宇(1973—),男,副教授,博士,主要从事农业机器人与机器人视觉技术研究,E-mail: 05338@163.com

and recognition method, and the test results showed that the number of model parameters of the improved algorithm was 1.013×10^7, the forward inference time was 27 ms, and the average accuracy mean value was 97.5%. Compared with the original CenterNet model, the number of parameters was increased by 1.66×10^6, the forward inference time was increased by 2 ms, but the mean average precision was increased by 8.2 percentage points. In the process of potato harvest and storage, the method of potato and impurity classification can meet the technical requirements of impurity identification and removal.

Key words: potatoes; impurities; SAM－CenterNet; ACmix; deep learning

0　引言

马铃薯作为全球广泛种植和消费的农作物，不仅提供了丰富的营养价值和美味的口感，还具有重要的经济和社会价值[1-2]。中国启动了马铃薯主粮化战略[3]，马铃薯深加工进程正在逐步推进。然而，在马铃薯收获过程中，常常掺杂大量的土块、石块等杂质[4]。掺杂的杂质不仅在输送或存储过程中损伤马铃薯，还会增加存储成本[5]。近年来，国家大力支持农业机械自动化的发展，在播种[6]、收获[7]、分级[8]等方面取得了显著进展。然而，由于杂质形状差异性和自动识别技术的限制，杂质剔除方面的发展受到了一定的影响。目前，马铃薯中的杂质剔除主要依靠人工操作，不仅成本高，而且效率低。

近年来，各国学者对马铃薯与杂质分类检测进行了研究[9-13]。使用传统的机器视觉[14-16]识别马铃薯与杂质，利用颜色纹理特征识别时，马铃薯与土石块形状相似性较高，造成识别精度不高；利用紫外线反射率差异识别，虽然准确率较高，但是马铃薯包裹大量泥土会导致误识别，并且紫外线相机成本较高；利用声信号和光散射差异识别方法，对识别环境要求较高，并且撞击钢板容易损伤马铃薯；利用热红外图像差异识别，输送分选过程需要热风激励，分选效率低。近年来，深度学习在目标检测算法领域展现出卓越的特征提取能力[17-19]，适应环境较强，分选效率和识别准确率高，已逐渐成为农业领域目标检测的主要算法[20-22]。但基于深度学习的马铃薯与杂质快速识别方面研究比较缺乏。大部分研究是基于需要预先设定一组特定宽高比的候选框去做目标检测，但是候选框的设计无法适应马铃薯与杂质所有目标的形状和尺寸，因此容易导致漏检和误检。针对以上问题，无先验框 CenterNet[23-24] 模型在实际应用中越来越受到关注。

由于薯杂形状相近，且马铃薯易粘连大量泥土，使得薯杂特征趋同，加之背景中的输送装置也会粘连泥土，易造成误判。针对薯杂分类检测的需求，CenterNet 模型在目标形状特征检测和检测精度方面存在局限性。为实现薯杂分类的精准、高效，本文提出改进 CenterNet 薯杂分类识别方法。针对部分马铃薯与土石块形状特征相似的情况，在主干网络特征提取阶段，将传统卷积下采样替换成 SPD－Conv 下采样；鉴于薯杂检测的背景存在粘连泥土，引入 ACmix 注意力机制去除无效背景，提高识别精度；由于原始主干网络缺乏薯杂表面纹理特征信息的多层级特征融合，导致检测精度下降，构建多层级特征融合结构，以提升模型的检测精度。

1　薯杂识别与杂质剔除装置结构与原理

薯杂识别与杂质剔除装置主要包括识别控制装置、滚筒输送装置和剔除装置，如图 1 所示。识别控制装置包括上支撑架、工控机、工业相机（4 台）、光电开关、条形光源（6 件）、触摸屏和 PLC 控制系统；滚筒输送装置包括下支撑架、滚筒、链传动装置和电动驱动滚筒，相邻 2 个滚筒支撑马铃薯并带动马铃薯旋转；剔除装置包括侧支架、剔除通道（6 个）、剔除板和气缸。薯杂识别与杂质剔除装置总共有 2 个通道组，每 3 个剔除通道组成一个通道组。每个通道组在薯杂输送方向上依次安装 2 台工业相机。4 台工业相机由光电开关触发同时采集图像。

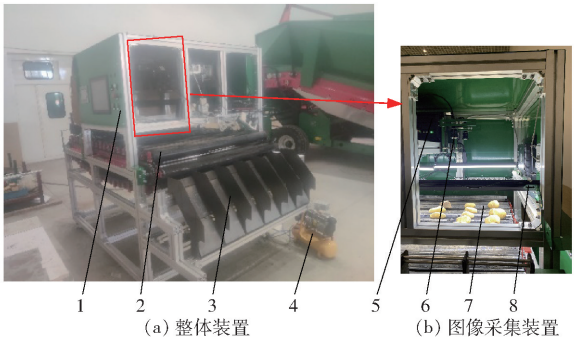


图 1　薯杂识别与杂质剔除装置实物图

Fig. 1　Physical picture of potato impurity identification and impurity removal device

1. 识别控制装置　2. 滚筒输送装置　3. 剔除装置　4. 空气压缩机　5. 工控机　6. 工业相机　7. 薯杂　8. 光源

薯杂识别剔除工作原理如下：薯杂由上料输送带传送至滚筒输送装置实现等间距有序输送。当马铃薯进入图像采集区域时，光电开关触发全部工业相机采集图像，每台工业相机视野范围内可采集 9 个位置的薯杂图像。光电开关的触发响应节拍是每

个视野范围触发一次。在相邻 2 个光电开关触发的过程中,薯杂在滚筒的带动下产生翻转,从而使得同一通道组内的前后 2 台工业相机分别采集到薯杂的正面、背面的图像,且正面图像和背面图像的薯杂位置一一对应。实时识别处理薯杂正面图像和背面图像,获取杂质的位置信息发送给 PLC,PLC 控制剔除装置中的气缸实现对杂质剔除。

2 图像采集

以山东省滕州地区种植的马铃薯“希森 3 号”为研究对象。设置滚筒输送装置的速度为 0.15 ~ 0.25 m/s,使用 40 kg 马铃薯和 40 kg 杂质进行数据采集。图像采集时,马铃薯和杂质混合后依次放置在滚筒输送装置上。通过如图 2 所示的机器视觉系统界面进行多次循环图像采集,共获得了 800 幅薯杂图像,每幅图像的分辨率为 4 096 像素 ×1 980 像素,并以 JPEG 格式保存。使用 LabelImg 软件对构建好的薯杂图像数据集进行打标签。将打好标签图像数据集按照比例 7∶2∶1随机划分为训练集、测试集和验证集[25]。



图 2 机器视觉系统界面

Fig. 2 Machine vision system interface

3 SAM - CenterNet 算法

3.1 SAM - CenterNet 网络结构

CenterNet 模型作为一种经典的无锚点目标检测算法,提供了精确的目标定位信息,能够准确地识别目标的类别。CenterNet 模型主要包括主干网络、解码网络、检测网络。主干网络采用传统卷积神经网络为核心的 ResNet18 模型[26]。解码网络通过 3 次反卷积进行上采样,获得高分辨率特征图。检测网络的分支有 3 个,分别为 Heatmap、Offset 和 Size。Heatmap 为热力图,代表每一个热力点所属的物体类别;Offset 代表该中心点在原始图像中的横纵坐标位置;Size 代表该中心点对应的目标框的宽度和高度。使用 CenterNet 模型对薯杂图像识别时,由于土石块与部分马铃薯存在一定的相似度、马铃薯表面和薯杂图像背景中粘连的泥土以及薯杂表面纹理特征提取不足,导致影响识别精度。主干网络主要负责获取薯杂图像的特征信息,因此在 ResNet18 主干网络中,将传统卷积下采样替换为 SPD - Conv(Space-to-depth-convolution)下采样,提升网络对薯杂图像中的薯杂细粒度识别能力,解决土石块与部分马铃薯相似导致漏识别的问题;在主干网络加入 ACmix(A mixed model of self-attention and convolution)注意力机制,增强薯杂表面的有效特征,抑制背景无效特征,解决马铃薯表面和图像背景粘连泥土导致误识别的问题;在主干网络中设计一个多层级特征融合结构(Multi-level feature fusion structure,MFS),获取薯杂图像中的浅层语义细节信息,增强薯杂表面纹理特征信息提取能力,解决模型输出特征单一而导致识别精度低的问题。通过上述改进,SAM - CenterNet(SPD - ACmix - MFS CenterNet)模型可解决原 CenterNet 模型识别薯杂图像中的问题,提高识别精度。SAM - CenterNet 模型结构如图 3 所示,图中主干网络中主要包括 2 个基本的块,分别为 SPDConv Block 和 Rank Block,SPDConv Block 输入输出特征图的尺寸不一样,作用是改变主干网络特征图尺寸,有利于细粒度信息提取;Rank Block 输入输出特征图的尺寸一样,目的是加深主干网络深度。

3.2 引入 SPD - Conv 模块

使用 CenterNet 模型对薯杂图像中的薯杂形状特征进行检测,由于部分薯杂形状相似度较高,其下采样通过 stride 大于 1 的卷积层或者池化层来实现,导致通道维度中细粒度信息丢失,限制了模型学习目标特征的能力。为了解决上述问题,将 SPD - Conv 下采样替换传统卷积下采样。SPD - Conv 模型结构由 1 个 SPD 层和 1 个 Conv 层组成,如图 4 所示。

SPD - Conv 模型结构对薯杂特征信息提取的步骤如下:

首先,对于输入的 $C_1 \times S \times S$ 特征图 X(其中 C_1、S 分别为输入通道数、特征图的尺寸),根据 scale 比例因子尺寸沿方向 x 和 y 方向同时进行切片操作,切出一系列子特征图。

$$\begin{cases} f_{0,0} = X[0:S:\text{scale},0:S:\text{scale}] \\ f_{\text{scale}-1,0} = X[\text{scale}-1:S:\text{scale},0:S:\text{scale}] \\ f_{0,1} = X[0:S:\text{scale},1:S:\text{scale}] \\ f_{\text{scale}-1,1} = X[\text{scale}-1:S:\text{scale},1:S:\text{scale}] \end{cases} \tag{1}$$

SAM - CenterNet 模型的 stride 值为 2,属于 2 倍下采样,因此选取 scale 值为 2 对特征图 X 进行切片操作,得到 4 个形状为 $C_1 \times S/2 \times S/2$ 的子特征图

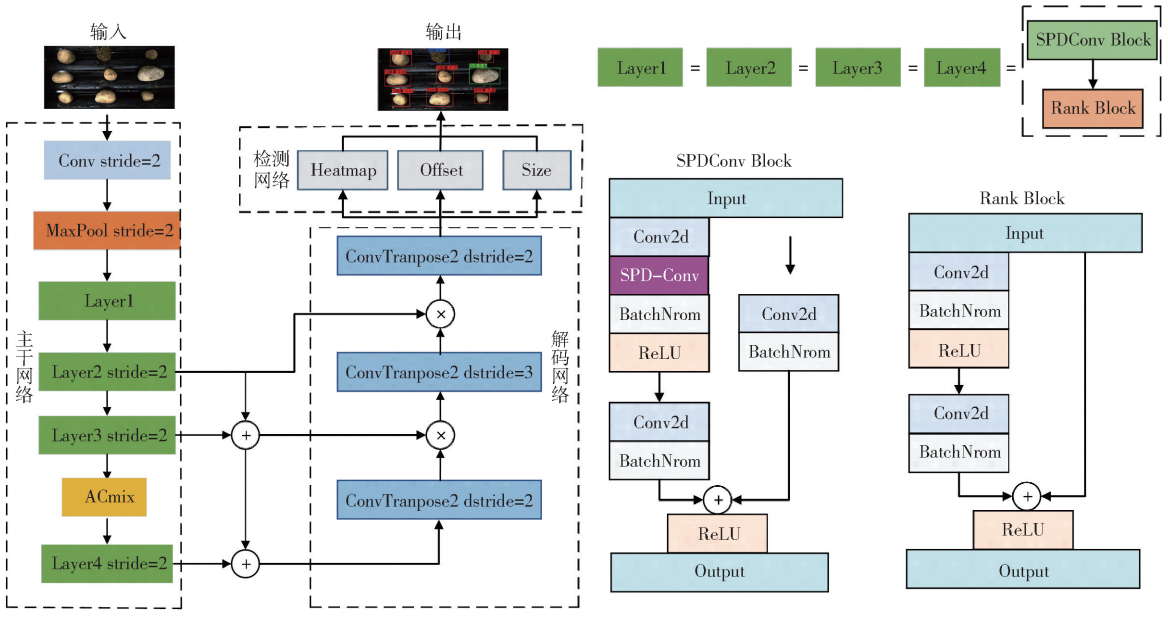


图 3　SAM－CenterNet 模型结构图

Fig. 3　SAM－CenterNet model structure diagram

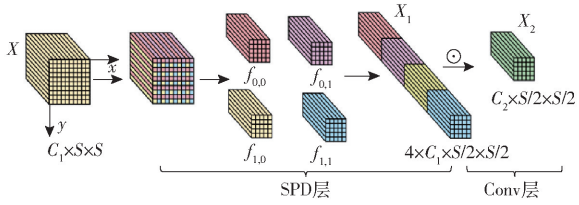


图 4　SPD－Conv 模型结构图

Fig. 4　Structure diagram of SPD－Conv model

$f_{0,0}$、$f_{0,1}$、$f_{1,0}$、$f_{1,1}$，从而保留马铃薯与杂质通道维度的所有信息。

其次，将 4 个子特征图沿着通道进行拼接，得到特征图 $X_1 \in \mathbf{R}^{4\times C_1\times S/2\times S/2}$，计算式为

$$X_1 = \text{Concat}([f_{0,0}, f_{1,0}, f_{0,1}, f_{1,1}]) \tag{2}$$

最后，对 X_1 进行 stride 为 1 的卷积操作，以 1×1 的特征图 F 作为卷积核，对 X_1 进行卷积操作，得到输出特征图 $X_2 \in \mathbf{R}^{C_2\times S/2\times S/2}$，计算式为

$$X_2 = X_1 \odot F \tag{3}$$

SPD－Conv 模型的主要特点在于通过 2 倍下采样切片操作获取的子特征图，能够有效保留特征映射通道维度中的所有信息，解决土石块与部分马铃薯存在特征相似导致漏识别的问题。

3.3　引入 ACmix 注意力机制

在收获过程中，马铃薯和输送滚筒的表面极易附着泥土，导致误识别为土块。为了解决该问题，在主干网络中引入 ACmix 注意力机制，其模型结构如图 5 所示，图中 H、W 和 C 分别为对应特征图的长、宽和通道数。ACmix 的运行过程包括以下 2 个阶段：

第 1 阶段是特征学习模块对输入的特征图实施 3 次 1×1 卷积操作投影，每个卷积输出 N 个片段，因此，最终生成一组 3×N 个特征图的中间特征。

第 2 阶段是特征聚合的过程，包括自注意力和卷积 2 个路径。自注意力路径，将中间特征收集到

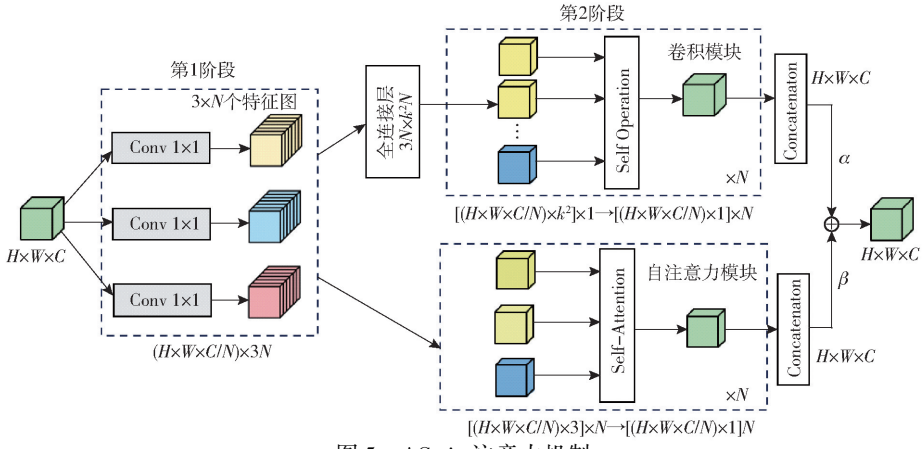


图 5　ACmix 注意力机制

Fig. 5　ACmix attention mechanism

N 组中,每组包含 3 个特征,3 个特征的映射分别为查询、键和值,对特征映射生成的特征图进行平移和聚合操作。通过自注意力模块对输入特征图的全局感知,输出得到 $H \times W \times C$ 的特征图 F_{att}。卷积路径采用全连接层,生成 k^2 个特征映射,对生成的特征进行平移和聚合操作。通过卷积模块处理,收集薯、杂、筒局部颜色、边缘特征,输出得到 $H \times W \times C$ 的特征图 F_{conv}。最后,将自注意力和卷积的输出进行相加,得到输出特征图 F_{out},计算式为

$$F_{out} = \alpha F_{att} + \beta F_{conv} \tag{4}$$

式中 α、β——可学习的参数

F_{att}——自注意力路径输出

F_{conv}——卷积路径的输出

引入 ACmix 注意力机制可有效增强薯杂特征的关注能力和网络的表征能力,实现薯杂特征的提取与整合,区分土块与马铃薯粘连的泥土,抑制背景无效特征,降低误识别。

3.4 多层级特征融合

原 CenterNet 网络主干特征提取输出通道如图 6 所示,只使用了最后一层(Layer4)的通道特征进行输出,忽略了存在浅层信息的其他通道特征,造成马铃薯与杂质的表面纹理特征信息提取不足,导致精度受限。为了提升检测精度,设计一个多层级特征融合结构,多层级特征融合结构如图 7 所示,Layer2、Layer3 和 Layer4 分别对应输出为特征图 F_s、特征图 F_c 和特征图 F_d。多层级特征融合过程如下:首先,将特征图 F_s 与特征图 F_c 进行 1×1 的卷积变换后,作相加融合输出 F_{out1};其次,将 F_{out1} 与特征图 F_d 进行 1×1 的卷积变换后,作相加融合输出 F_{out2},F_{out2} 通过反卷积获得输出 F_{u1};然后,为了增加模型的泛化能力,将 F_{out1} 与 F_{u1} 进行 1×1 的卷积变换后,作相乘融合输出 F_{out3},F_{out3} 通过反卷积获得输出 F_{u2};最后,将特征图 F_s 与 F_{u2} 进行 1×1 的卷积变换后,作相乘融合输出 F_{out4}。多层级特征融合结构能够将浅层特征信息与深层特征信息进行融合,捕捉

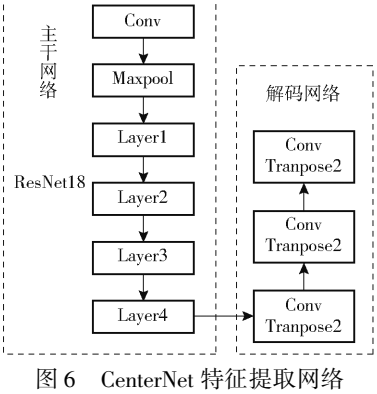


图 6 CenterNet 特征提取网络

Fig. 6 CenterNet feature extraction network

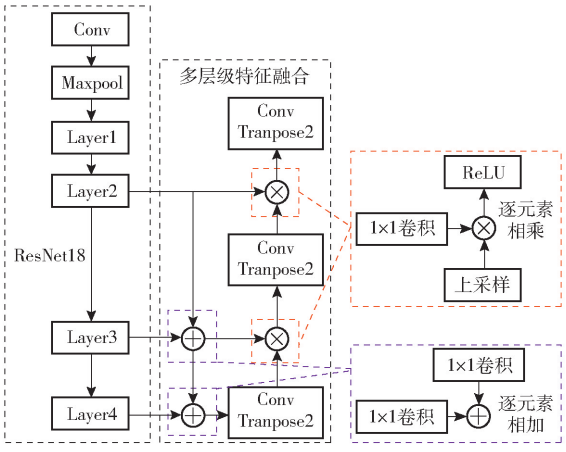


图 7 多层级特征融合

Fig. 7 Multi-level feature fusion

不同层次的特征信息,解决输出特征图单一问题,获取更全面和丰富的语义信息,计算式为

$$\begin{cases} F_{out1} = \varepsilon F_s + \theta F_c \\ F_{out2} = F_{out1} + \lambda F_d \\ F_{out3} = F_{out1} F_{u2} \\ F_{out4} = \varepsilon F_s F_{u1} \end{cases} \tag{5}$$

式中 ε、θ、λ——可学习的参数

4 试验结果及分析

4.1 试验环境与评价指标

试验过程在 Windows 11 操作系统下进行,GPU 型号为 NVIDIA GeForce RTX 3060Ti,内存 24 GB。深度学习框架为 Pytorch,编程平台为 PyCharm,所用编程语言为 Python 3.8。初始学习率为 0.000 1,学习率下降方式使用余弦退火 Cos(Cosine annealing),优化器选用 Adam,优化器内部的 Momentum 参数设置为 0.9,设置输入薯杂图像尺寸为 512 × 512 × 3,训练总轮数设置为 200。本文以精确率(Precision,P)、召回率(Recall,R)、平均精度(Average precision,AP)、平均精度均值(Mean average precision,mAP)、F1 值(F1-value)、模型参数量(Parameters)和前向推理时间(Forward inference time)作为算法性能的评价指标。

4.2 注意力机制引入不同位置的对比试验

注意力机制的引入,在提高识别精度的同时,模型参数量将增加,存在梯度消失和过拟合风险。在主干网络的不同位置引入注意力机制,识别精度提高与参数量增加的关系存在差异。在 ResNet18 作为主干网络和 SPD - Conv 下采样的基础上,按图 8 所示的 4 个引入位置依次将注意力机制引入。4 个不同引入位置的注意力机制模型分别为 ACmix1 模型、ACmix2 模型、ACmix3 模型、ACmix4 模型。不引入注意力机制的模型为 ACmix0 模型。按照不同的

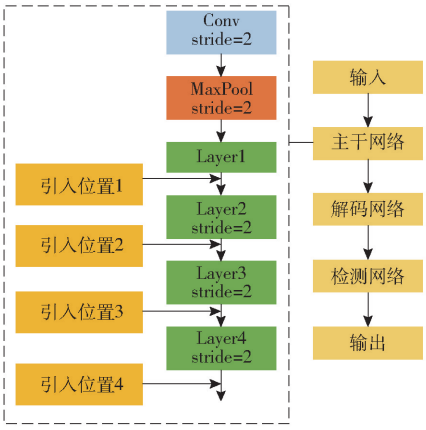


图 8　注意力机制不同位置的引入框图

Fig. 8　Attention mechanisms block diagrams at different positions

注意力机制模型分别对训练集图像进行处理，试验结果如表 1 所示。从表 1 可以看出，4 个不同引入位置的注意力机制模型与 ACmix0 相比，模型参数量依次增加 1×10^4、5×10^4、2×10^5、9×10^5，mAP 依次增加 0.5、1.3、2.7、2.7 个百分点。在 4 个不同引入位置的注意力机制模型中，ACmix3 模型参数量增幅较小，mAP 最大。ACmix4 与 ACmix3 相比，mAP 相同，参数量增加 7×10^5，因此，考察多个引入位置同时引入注意力机制对模型参数量和 mAP 的影响时，不考虑 ACmix4。在主干网络中，同时在 2 个及 3 个引入位置引入注意力机制的模型分别为 ACmix12 模型、ACmix13 模型、ACmix23 模型、ACmix123 模型。ACmix3 与 ACmix12、ACmix13、ACmix23、ACmix123 相比，除了比 ACmix12 模型参数量增加 1.4×10^5，其余分别降低 1×10^4、5×10^4、6×10^4，mAP 分别增加 1.6、0.5、0.3、0.7 个百分点。综合考虑注意力机制不同引入位置的模型参数量和 mAP，ACmix3 模型是最佳的 ACmix 注意力机制引入位置。

表 1　注意力机制不同引入位置的试验结果

Tab. 1　Experimental results of different introduction locations of attention mechanism

参数	注意力机制的模型								
	ACmix0	ACmix1	ACmix2	ACmix3	ACmix4	ACmix12	ACmix13	ACmix23	ACmix123
模型参数量	9.21×10^6	9.22×10^6	9.26×10^6	9.41×10^6	1.011×10^7	9.27×10^6	9.42×10^6	9.46×10^6	9.47×10^6
平均精度均值/%	92.4	92.9	93.7	95.1	95.1	93.5	94.6	94.8	94.4

4.3　训练与测试结果分析

在模型训练过程中，利用了冻结训练与解冻训练相结合的策略。该方法也有助于降低训练过程中的计算量，加快网络的收敛速度。在前 50 次迭代过程中，模型采用冻结训练，主干特征提取网络不发生改变，仅微调模型的其它部分。在后续的训练阶段中，模型则采用解冻训练的方式，对主干特征提取网络的参数进行调整，提高模型的性能。改进前后的 2 个模型对训练集图像进行处理的损失曲线如图 9 所示。

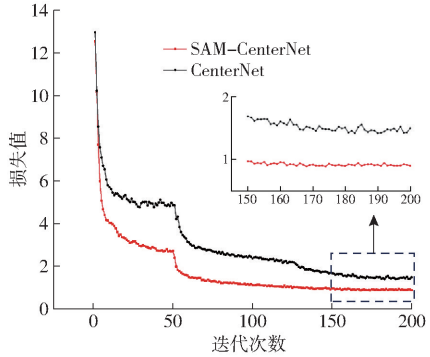


图 9　改进前后模型的训练损失曲线

Fig. 9　Training loss curve of model before and after improvement

从图 9 可以看出，在迭代 50 次内的冻结训练阶段，SAM－CenterNet 模型和 CenterNet 模型收敛速度都有所加快。解冻之后，损失值有一个快速下降的阶段，随后损失值下降逐渐平稳，表明网络仍在进行参数调整。相较于原模型，改进模型迭代 100 次开始收敛，原模型需要迭代到 150 次才达到收敛状态，表明改进后模型对马铃薯与杂质的特征信息识别能力更强。

原模型与改进模型完成训练后，在测试集图像上进行测试，测试识别结果如图 10 和图 11 所示。从图 10 和图 11 可以看出，随着阈值升高，马铃薯和土石块识别精确率呈上升趋势，召回率则呈下降趋势，F1 值先升高后降低。相较于原 CenterNet 模型，SAM－CenterNet 模型精确率具有更快的上升速度和更早趋于稳定的趋势；SAM－CenterNet 模型召回率没有出现过早下降的现象；SAM－CenterNet 模型 F1 值最大值状态保持更长久。以上结果表明传统卷积下采样替换为 SPD－Conv 下采样和构建一个多层级特征融合结构，以及在主干网络中引入 ACmix 注意力机制，能够有效提升算法的识别性能。

精确率高，能够减少误识别问题；召回率高，能够降低漏识别问题。F1 值为精确率和召回率的调和平均数，F1 值越高，模型的整体性能越好。从图 11 中可以看出，马铃薯和土石块识别结果的 F1 值在阈值 0.3 ~ 0.7 之间保持较高位置。使用

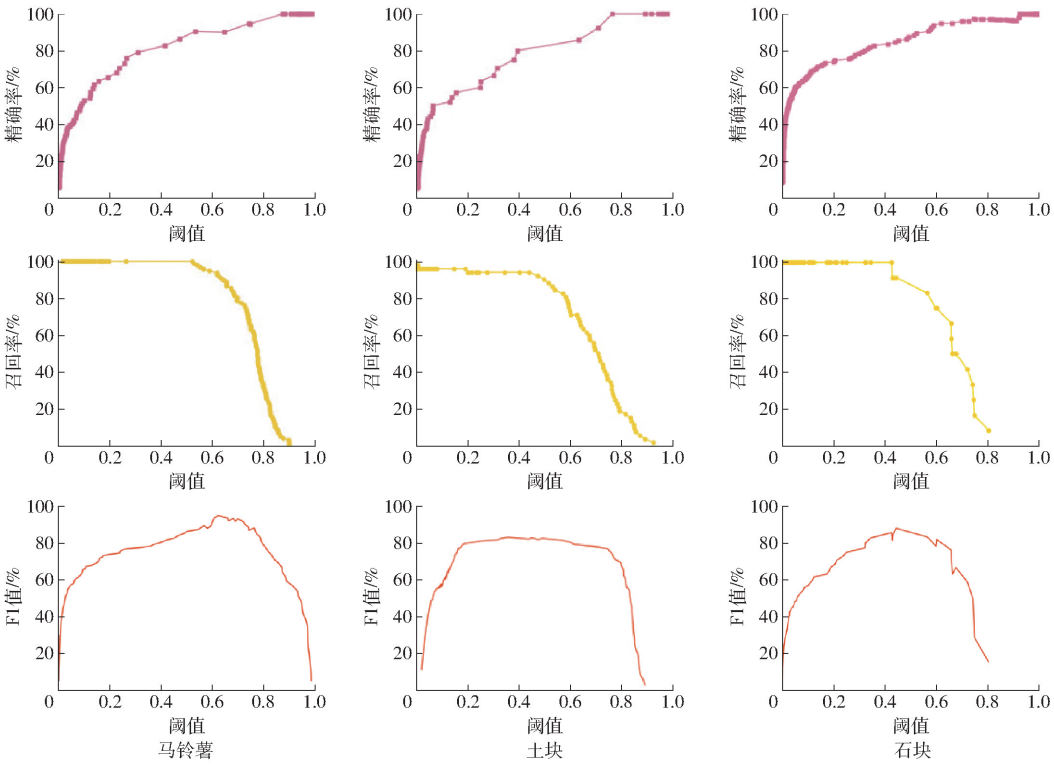


图 10 CenterNet 模型测试识别结果

Fig. 10 Test recognition results of CenterNet model

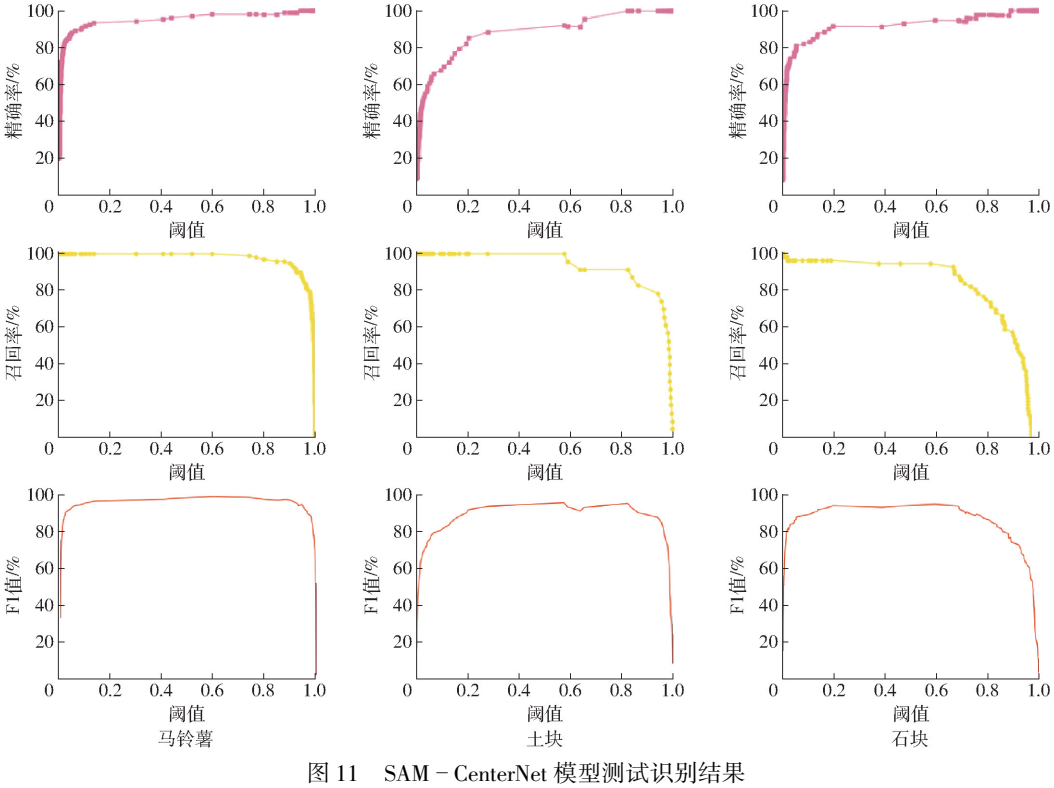


图 11 SAM - CenterNet 模型测试识别结果

Fig. 11 Test recognition results of SAM - CenterNet model

SAM -CenterNet 模型在阈值 0.3 ~0.7 之间对测试集图像进行薯杂识别试验,计算平均精度,试验结果如图 12 所示,阈值从 0.3 ~0.5 的变化过程中,马铃薯和土块识别的平均精度平缓上升,石块识别的平均精度存在微小波动。阈值从 0.5 ~0.7 的变化过程中,马铃薯识别平均精度逐渐下降,土块识别的平均精度急剧下降后趋于平稳,识别石块的平均精度相对平稳后急剧下降。选取阈值 0.5,识别马铃薯、土块和石块的平均精度分别为 98.9%、97.8%、95.6%,阈值 0.5 为综合最佳值,可满足实际生产需

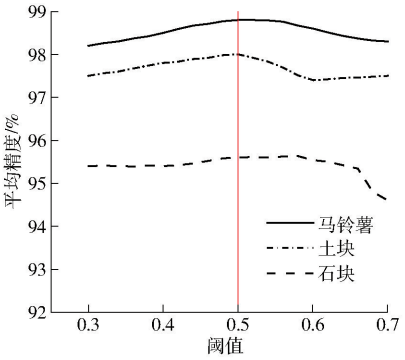


图 12　SAM－CenterNet 模型薯杂识别平均精度

Fig. 12　SAM－CenterNet model potato hybrid recognition AP value

求。因此，选取阈值 0.5 为后续试验参数。

4.4　消融试验

为了验证 SAM－CenterNet 模型的有效性，采用消融试验对比改进前后的效果。方案 1 在原 CenterNet 基础上将传统卷积下采样替换为 SPD－Conv；方案 2 是在方案 1 的基础上加入 ACmix 注意力机制；方案 3 是在方案 1 的基础上加入多层级特征融合结构。试验结果如表 2 所示。

表 2　消融试验结果

Tab. 2　Results of ablation experiments

方法	SPD－Conv	ACmix	MFS	mAP/%	参数量	前向推理时间/ms
CenterNet				89.3	8.47×10^6	20
方案 1	√			92.4	9.21×10^6	22
方案 2		√		93.1	8.67×10^6	21
方案 3			√	91.6	9.19×10^6	24
SAM－CenterNet	√	√	√	97.5	1.013×10^7	27

注：“√”表示该方案含有该模块。

从表 2 可知，方案 1 采用的 SPD－Conv 下采样，检测模型的 mAP 提高 3.1 个百分点，表明 SPD－Conv 下采样提升了模型网络捕捉细粒度特征的能力，解决了因部分薯杂形状相似而导致的漏识别情况；方案 2 加入 ACmix 注意力机制，检测模型的 mAP 提高 3.8 个百分点，表明 ACmix 注意力机制能够获取薯杂的关键特征信息，可以有效抑制背景无效特征。方案 3 采用的多层级特征融合结构，检测模型的 mAP 提高了 2.3 个百分点，表明 MFS 结构能够解决薯杂表面纹理特征信息提取不足导致的精度受限问题。本文提出的 SAM－CenterNet 模型参数量虽然增加 1.66×10^6，前向推理时间增加 7 ms，但 mAP 达到 97.5%，与原 CenterNet 模型相比 mAP 提升 8.2 个百分点，可满足实际生产要求。综上所述，SAM－CenterNet 模型对薯杂识别的平均精度均值得到了有效提升。

选用测试集图像验证分析，对改进前后算法的检测结果进行了可视化比较，试验结果如图 13 所示。经可视化对比结果分析，改进后的模型与原模型相比，不仅解决杂质漏识别和误识别问题，能够精确地检测到目标，而且还提高了识别的置信度。为了更直观地展示置信度提升的程度，将测试集图像识别结果的所有数据导入 Origin 软件，对比了马铃薯、石块和土块的 CenterNet 模型改进前后的置信度数据结果，如图 14 所示。图中识别效果判别线是斜率为 1 的直线，其下方的数据点表示改进后识别结果置信度低于改进前识别结果置信度，上方的数据点表示改进后识别结果置信度高于改进前识别结果置信度。采用改进后置信度减去改进前置信度再除以改进前置信度的计算方法计算 CenterNet 模型改

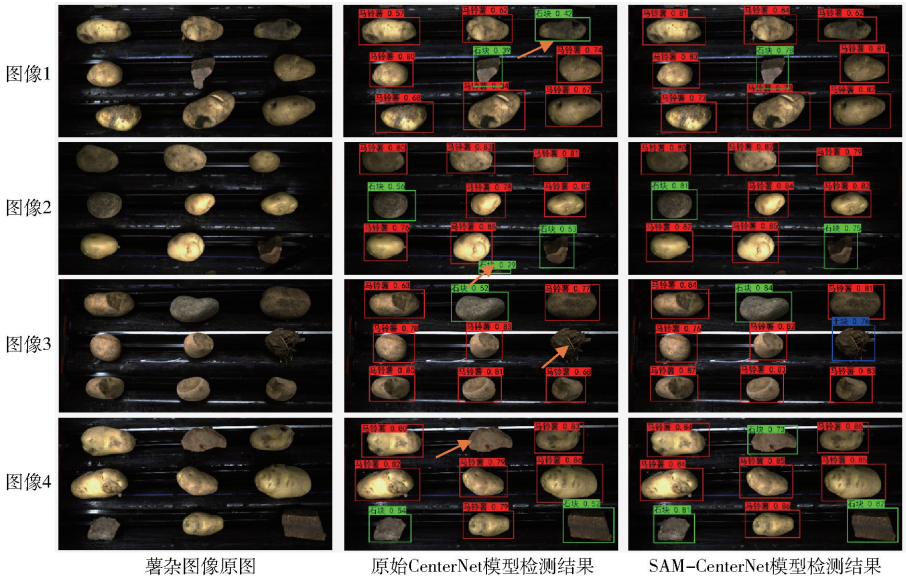


图 13　改进前后 CenterNet 模型检测结果可视化

Fig. 13　Visual inspection results of CenterNet model before and after improvement

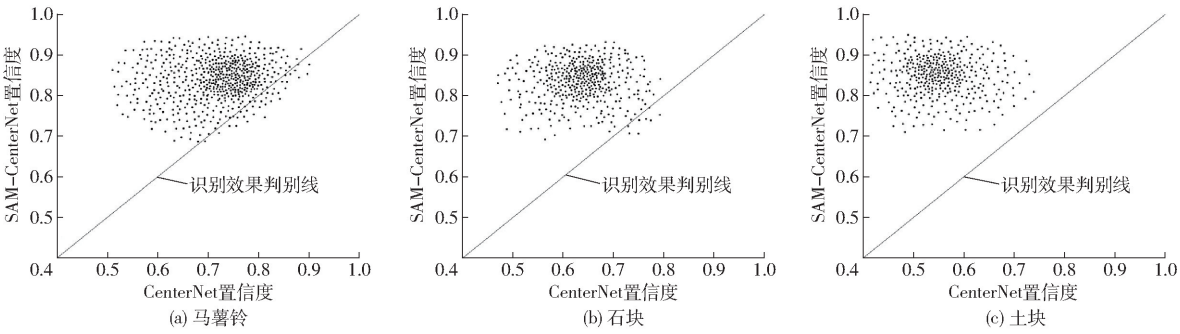


图 14 CenterNet 改进前后的置信度对照图

Fig. 14 Confidence graphs before and after CenterNet improvements

进前后的置信度提升率。由计算结果可知,SAM - CenterNet 模型识别马铃薯、石块和土块的置信度提升率分别为 10%、40%、60%。

5 结论

(1)为提高马铃薯与杂质的识别精度,提出了基于 SAM - CenterNet 模型的薯杂分类识别方法。将传统卷积下采样替换为 SPD - Conv 下采样,提升网络对薯杂图像中的薯杂细粒度识别能力,解决土石块与部分马铃薯相似导致漏识别的问题;在主干网络中引入 ACmix 注意力机制,增强薯杂表面的有效特征,抑制背景无效特征,解决马铃薯表面和薯杂图像背景粘连泥土导致误识别的问题;同时,设计多层级特征融合结构,获取薯杂图像中的浅层语义细节信息,解决薯杂表面纹理特征信息提取不足问题,提高模型的检测精度。

(2)本文提出的 SAM - CenterNet 模型薯杂分类识别方法,可以对马铃薯与杂质进行分类识别,mAP 可达 97.5%,模型参数量为 1.013×10^7,前向推理时间为 27 ms,相比原 CenterNet 模型,mAP 提高 8.2 个百分点。满足薯杂的快速检测要求,可为后续杂质的智能剔除提供技术支撑。

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 Qiyou, LUN Runqi, GAO Mingjie, et al. Strategic path of high-quality development of potato industry in China from 2021 to 2025[J]. Chinese Journal of Agricultural Resources and Regional Planning, 2022, 43(3): 37 - 45. (in Chinese)

[3] 章启祥,刘发英,杨振宇,等. 基于离散元的马铃薯分选装置运行参数优化与试验[J]. 中国农机化学报,2023,44(4):96 - 103.
ZHANG Qixiang, LIU Faying, YANG Zhenyu, et al. Operation parameter optimization and experiment of potato sorting device based on discrete element[J]. Journal of Chinese Agricultural Mechanization, 2023, 44(4): 96 - 103. (in Chinese)

[4] 王虎存,赵武云,孙伟,等. 马铃薯机械化收获技术与装备研究进展[J]. 农业工程学报,2023,39(14):1 - 22.
WANG Hucun, ZHAO Wuyun, SUN Wei, et al. Reserch progress on the technology and equipment for potato mechanized harvesting[J]. Transactions of the CSAE, 2023, 39(14): 1 - 22. (in Chinese)

[5] 申阳,胡虎标,许广泰,等. 马铃薯的累积损伤特性[J]. 科学技术与工程,2023,23(2):536 - 541.
SHEN Yang, HU Hubiao, XU Guangtai, et al. Acumulative damage characteristics of potato[J]. Science Technology and Engineering, 2023, 23(2): 536 - 541. (in Chinese)

[6] 杨然兵,王婕,尚书旗,等. 圆台格盘式马铃薯育种试验播种机设计与试验[J]. 农业机械学报,2023,54(1):84 - 93.
YANG Ranbing, WANG Jie, SHANG Shuqi, et al. Design and experiment of potato breeding planter with rotary table grid[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(1): 84 - 93. (in Chinese)

[7] 魏忠彩,王兴欢,李学强,等. 履带自走式分拣型马铃薯收获机设计与试验[J]. 农业机械学报,2023,54(2):95 - 106.
WEI Zhongcai, WANG Xinghuan, LI Xueqiang, et al. Design and experiment of crawler self-carrying sorting type potato harvester[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(2): 95 - 106. (in Chinese)

[8] 姜宏,于永波,章翔峰,等. 马铃薯外部品质分级方法综述[J]. 科学技术与工程,2022,22(14):5519 - 5527.
JIANG Hong, YU Yongbo, ZHANG Xiangfeng, et al. Potato external quality grading methods: a review[J]. Science Technology and Engineering, 2022, 22(14): 5519 - 5527. (in Chinese)

[9] KHAZAEE Y, KHEIRALIPOUR K, HOSAINPOUR A, et al. Development of a novel image analysis and classification algorithms to separate tubers from clods and stones[J]. Potato Research, 2022, 65(3): 707 - 728.

[10] AL-MALLAHI A, KATAOKA T, OKAMOTO H, et al. Detection of potato tubers using an ultraviolet imaging-based machine vision system[J]. Biosystems Engineering, 2010, 105(2): 257－265.
[11] HOSAINPOUR A, KOMARIZADE M H, MAHMOUDI A, et al. High speed detection of potato and clod using an acoustic based intelligent system[J]. Expert Systems with Applications, 2011, 38(10): 12101－12106.
[12] 耿金凤. 含杂马铃薯中土块和石块的激光背向散射成像检测方法研究[D]. 杭州:浙江大学,2019.
GENG Jinfeng. Detection of clods and stones from impurified potatoes using laser backscattering imaging[D]. Hangzhou: Zhejiang University, 2019. (in Chinese)
[13] 孙卫孝,刘发英,杨振宇,等. 基于主动热红外成像的马铃薯与杂质分类方法[J]. 中国农机化学报,2024,45(2):143－150.
SUN Weixiao, LIU Faying, YANG Zhenyu, et al. Classification of potato and impurities based on active thermal infrared imaging[J]. Journal of Chinese Agricultural Mechanization, 2024, 45(2): 143－150. (in Chinese)
[14] 刘平,刘立鹏,王春颖,等. 基于机器视觉的田间小麦开花期判定方法[J]. 农业机械学报,2022,53(3):251－258.
LIU Ping, LIU Lipeng, WANG Chunying, et al. Determination method of field wheat flowering period based on machine vision [J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(3): 251－258. (in Chinese)
[15] REZAEI P, HEMMAT A, SHAHPARI N, et al. Machine vision-based algorithms to detect sunburn pomegranate for use in a sorting machine[J]. Measurement, 2024, 232: 114682.
[16] GAN N, WANG Y, REN G, et al. Design and testing of a machine-vision-based air-blow sorting platform for famous tea fresh leaves production[J]. Computers and Electronics in Agriculture, 2023, 214: 108334.
[17] HOSAIN M T, JIM J R, MRIDHA M F, et al. Explainable AI approaches in deep learning: advancements, applications and challenges[J]. Computers and Electrical Engineering, 2024, 117: 109246.
[18] 张万枝,张弘毅,刘树逢,等. 基于改进 YOLO v5s 的马铃薯种薯芽眼检测方法[J]. 农业机械学报,2023,54(9):260－269.
ZHANG Wanzhi, ZHANG Hongyi, LIU Shufeng, et al. Detection method of potato seed bud eye based on improved YOLO v5s [J]. Transactions of the Chinese Society for Agricultural Machinery 2023,54(9):260－269. (in Chinese)
[19] SALEH M E, WAZERY Y M, ALI A A. A systematic literature review of deep learning-based text summarization: techniques, input representation, training strategies, mechanisms, datasets, evaluation, and challenges[J]. Expert Systems with Applications, 2024, 252: 124153.
[20] 刘毅君,何亚凯,吴晓媚,等. 基于改进 Faster R－CNN 的马铃薯发芽与表面损伤检测方法[J]. 农业机械学报,2024,55(1):371－378.
LIU Yijun, HE Yakai, WU Xiaomei, et al. Potato sprouting and surface damage detection method based on improved Faster R－CNN[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(1):371－378. (in Chinese)
[21] 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 potatoes based on improved YOLO v4 algorithm[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(8): 241－247,262. (in Chinese)
[22] 李栋,仲婷,王笋,等. 基于改进 CenterNet 的柑橘叶片病害检测[J]. 光电子·激光,2024,35(9):934－941.
LI Dong, ZHONG Ting, WANG Sun, et al. Disease detection of citrus leaves based on improved CenterNet[J]. Journal of Optoelectronics· Laser, 2024, 35(9): 934－941. (in Chinese)
[23] ZHOU T, DU Y, MAO J, et al. Parallel attention multi-scale mandibular fracture detection network based on CenterNet[J]. Biomedical Signal Processing and Control, 2024, 95: 106338.
[24] 李利荣,陈鹏,张云良,等. 基于改进 CenterNet 的输电线路电力器件及异常目标检测[J]. 高电压技,2023,49(11): 4757－4765.
LI Lirong, CHEN Peng, ZHANG Yunliang, et al. Detection of power devices and abnormal objects in transmission lines based on improved CenterNet[J]. High Voltage Engineering, 2023, 49(11): 4757－4765. (in Chinese)
[25] 郭辉,陈海洋,高国民,等. 基于 YOLO v5m 的红花花冠目标检测与空间定位方法[J]. 农业机械学报,2023,54(7): 272－281.
GUO Hui, CHEN Haiyang, GAO Guomin, et al. Safflower corolla object detection and spatial positioning methods based on YOLO v5m[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(7): 272－281. (in Chinese)
[26] 牛智有,于重洋,吴志陶,等. 基于改进 ResNet18 模型的饲料原料种类识别方法[J]. 农业机械学报,2023,54(2):378－385,402.
NIU Zhiyou, YU Chongyang, WU Zhitao, et al. Identification of feed raw material type based on improved ResNet18 model [J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(2): 378－385,402. (in Chinese)