



张续博,张家瑜,贲宗友,等. 基于 YOLOv5-RF 的糙米内部裂纹识别方法研究[J]. 南京农业大学学报,2024,47(4):792-802.
ZHANG Xubo,ZHANG Jiayu,BEN Zongyou,et al. Research on internal crack recognition method of brown rice based on YOLOv5-RF[J]. Journal of Nanjing Agricultural University,2024,47(4):792-802.

基于 YOLOv5-RF 的糙米内部裂纹识别方法研究

张续博,张家瑜,贲宗友,王逸凡,曾雨可,陈坤杰*

(南京农业大学工学院,江苏 南京 210031)

摘要:[目的]为准确识别糙米内部裂纹,提出了一种基于改进 YOLOv5l 的糙米裂纹识别算法——YOLOv5-RF。[方法]使用深度卷积代替 CBS(convolution+batch normalization+sigmoid linear unit)模块中的普通卷积构建出 DWBS(depthwise convolution+batch normalization+sigmoid linear unit)模块,然后将 DWBS 模块堆叠成辅助骨干网络,再与原始的 Darknet-53 主干网络相结合形成基于 DWBS 模块的同级复合骨干网络,以提高糙米内部裂纹特征提取能力;在颈部网络使用逆向连接层来增强解码端特征的丰富性,用于提高解码端对糙米内部裂纹的识别能力;利用 CBAM(convolutional block attention module)注意力模块调整 SPPF(spatial pyramid pooling-fast)模块的通道注意力表达,通过构建一种 SPPF 模块与 CBAM 模块结合的网络结构来提高编码阶段的语义信息表达质量。[结果]所提出的 YOLOv5-RF 算法的平均准确率、召回率和准确率分别为 94.01%、86.92%、90.85%,相比 YOLOv5l 算法分别提升了 4.72%、6.23%、1.23%,但添加模块对模型检测速度的影响却很小。[结论]所提方法能够准确进行糙米内部裂纹的识别,可为稻谷爆腰率检测提供技术参考。

关键词:糙米裂纹;目标识别;特征信号增强;逆向连接层

中图分类号:TP391.4;S226.6

文献标志码:A

文章编号:1000-2030(2024)04-0792-11

Research on internal crack recognition method of brown rice based on YOLOv5-RF

ZHANG Xubo,ZHANG Jiayu,BEN Zongyou,WANG Yifan,ZENG Yuke,CHEN Kunjie*

(College of Engineering,Nanjing Agricultural University,Nanjing 210031,China)

Abstract:[Objectives]In order to accurately identify internal cracks in brown rice,a brown rice crack detection algorithm based on the enhanced YOLOv5l,YOLOv5-RF,was proposed. [Methods]A DWBS(depthwise convolution+batch normalization+sigmoid linear unit)module was constructed using deep convolution instead of ordinary convolution in the CBS(convolution+batch normalization+sigmoid linear unit)module,and then the DWBS module was stacked into an auxiliary backbone network,which was then combined with the original Darknet-53 backbone network to form a composite backbone network structure of the same level based on the DWBS module to improve the ability and extract internal crack features in brown rice. The feature richness at the decoding end was improved by using an reverse cross-layer connection in the neck network,which was used to improve the detection of internal cracks in brown rice at the decoding end. The CBAM attention module was used to adjust the channel attention expression of the SPPF(spatial pyramid pooling-fast)module to improve the quality of the semantic information expression in the encoding phase by constructing a network structure that combined the SPPF module with the CBAM module. [Results]The experimental results showed that the proposed YOLOv5-RF algorithm achieved 94.01%,86.92% and 90.85% average precision,recall and precision, respectively,which was an improvement of 4.72%,6.23% and 1.23% compared with the YOLOv5l algorithm. But adding modules had little effect on the speed of model detection. [Conclusions]The proposed method can accurately identify internal cracks in brown rice and can provide a technical reference for detecting the cracking rate of rice grains.

Keywords:brown rice cracks;target recognition;characteristic signal enhancement;reverse cross-layer connection

稻谷是世界四大主粮之一,在全球范围内有广泛消费市场^[1]。稻谷在收获和干燥过程,由于各种因素的作用内部会产生裂纹和爆腰,导致稻谷在后续的脱壳和碾磨等加工中产生碎米,从而降低大米品质^[2-3]。因此,爆腰率是评价稻谷烘干质量和衡量干燥工艺是否合理的重要依据^[4]。尽管稻谷爆腰率特别重要,但检测方法十分落后。根据《粮食、油料检验黄粒米及裂纹粒检验法:GB/T 5496—1985》的规定,目前,稻谷爆腰率的检测,主要是借助放大镜、背光灯等简单的装置,人工识别出糙米内部裂纹,然后统计

收稿日期:2023-10-30

基金项目:江苏省科技计划专项资金(重点研发计划现代农业)项目(BE2021305)

*通信作者:陈坤杰,教授,博导,主要从事农产品加工、无损检测及智能化装备研究,E-mail:kunjiechen@njau.edu.cn。

裂纹数和长度来判断稻谷是否爆腰,存在主观性强、效率低及易出错等问题^[5],已不能满足市场需求。

近些年来,随着计算机技术的发展,许多研究人员开始尝试利用计算机视觉、深度学习和图像处理等技术,进行稻谷爆腰检测方法的研究。陈树人等^[6]使用 Micro-CT 设备扫描稻谷,结合图像处理技术和三维重构技术来获取稻谷内部裂纹高清图像,并建立了稻谷所受载荷与稻谷裂纹程度的对照关系。Shi 等^[7]利用 X 光设备获取稻谷内部裂纹高清图像,结合图像处理技术完成稻谷内部糙米裂纹的识别,识别准确率达到 97%。吴杰等^[8]尝试用电子扫描显微镜获取清晰的稻谷断裂面裂纹图像,并利用图像处理技术获取裂纹长度作为 BP 神经网络的输入向量,以此判断稻谷是否爆腰,判别准确率达到 96.4%。尽管采用上述方法能容易获得稻谷内部裂纹图像进行爆腰的判定,但上述所采用的 Micro-CT 扫描设备、X 光设备与电子扫描显微镜设备获取稻谷内部裂纹图像的成本均过于高昂,不利于设备普及。目前,工业相机因其具备高分辨率、快速采集、耐用稳定、价格便宜且能够满足工业复杂应用场景中对高质量图像获取和分析的需求等特点,已经应用于工业领域中。为此一些研究人员也提出采用工业相机获取稻谷图像,再用图像处理技术计算稻谷爆腰率。许振伟^[9]提出了一种基于双正交 B 样条小波变换的大米识别方法,其通过图像处理技术提取单粒大米并对其进行小波变换以此识别大米裂纹,识别准确率达到 98%,但存在裂纹边缘漏检的情况。邢键等^[10]提出一种使用行灰度均值方法判别大米内部是否存在裂纹,判别准确率达到 98.2%,但根据《粮食、油料检验黄粒米及裂纹粒检验法:GB/T 5496—1985》对稻谷爆腰率的定义,需首先识别裂纹区域,然后进行裂纹数和长度的定量分析来判定稻谷是否爆腰,因此,此方法不能用来解决稻谷爆腰率检测的问题。2022 年,王修善等^[11]设计开发了一种应用深度学习的方法进行稻谷内部裂纹识别技术,为后续裂纹定量分析做准备,但由于其采用的 YOLOv4 目标检测模型对裂纹特征提取能力较弱,无法精准识别小裂纹,导致该方法的准确率仅为 81.53%,远远满足不了实际的检测精度要求。尽管如此,该方法却为稻谷爆腰率检测技术的研究提供了一个有益的参考。

目前,典型的目标检测模型主要包括两阶段目标检测器 R-CNN 系列和一阶段目标检测器 YOLO 系列。相较于两阶段目标检测器,一阶段目标检测器用于检测裂纹这类小目标速度更快,所需计算资源更小,因此,多数研究人员选择改进 YOLO 模型来解决裂纹识别问题^[12]。改进的途径有 2 种:一种途径是基于 CNN(卷积神经网络)改进 YOLO 模型,另一种途径是基于 Transformer(注意力机制)^[13]。基于 CNN 的改进,主要包括网络结构的调整、损失函数的修改和反向传播算法的优化,其中网络结构作为编码和解码的主要工作组件,在特征提取、特征传输和神经网络中心表达上起着至关重要的作用。基于 Transformer 的改进,主要利用其注意力机制来改进目标检测的精度和上下文理解能力,虽然可以取得较好的效果,但是训练模型所需计算资源成本高,模型不易部署^[14-15]。

从提高糙米的裂纹识别精度和设备成本及普及性考虑,本文以 YOLOv5I 模型为基础,通过引入同级复合骨干网络(composite backbone network structure of the same level)增强小目标裂纹特征信息提取能力,通过添加逆向连接层(reverse cross-layer connection)结构以提高裂纹位置信息的识别能力,以及通过融合 CBAM 注意力模块提高神经网络中特征信息的高质量表达这 3 种方式,基于工业相机,研究开发糙米内部裂纹识别模型——YOLOv5-RF,旨在解决稻谷爆腰率检测的裂纹识别精度问题。

1 材料与方法

1.1 糙米图像采集及数据集建立

1.1.1 样本制作与图像采集系统搭建 选择长粒(‘美香占 2 号’,广东省份汕尾市)、中粒(‘稻花香 2 号’,黑龙江省哈尔滨市)与圆粒(‘南粳 46 号’,江苏省盐城市)3 种稻谷为试验材料。按照图 1 所示流程

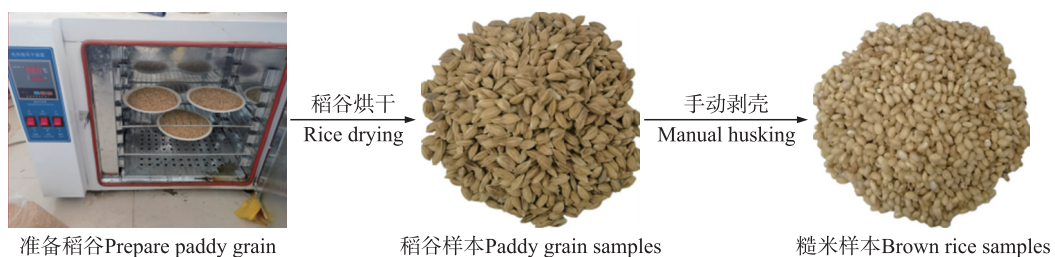


图 1 糙米样本制作流程图

Fig. 1 The flow chart for making brown rice samples

制作样本,然后根据《粮食、油料检验黄粒米及裂纹粒检验法:GB/T 5496—1985》标准用放大镜观察糙米裂纹情况,进行标记。

糙米裂纹图像采集系统如图 2-A 所示,主要包括:糙米托盘、计算机、光源及控制器、灯光盒、工业相机与固定支架等。在图像采集前,固定相机镜头与样本距离为 25 cm,在固定支架上覆盖蓝色吸光布以降低外部光源干扰。

由于糙米凹凸不平,易导致光的散射影响裂纹的可见性,采用蓝色平行背光源和白色环形光源搭建了如图 2-B 所示的光照系统。蓝光波长短,能够更好穿过糙米表皮,使裂纹暴露;白色侧光在裂纹处产生显著的反射和散射效应,使裂纹呈现亮度差异,凸显裂纹的边缘。

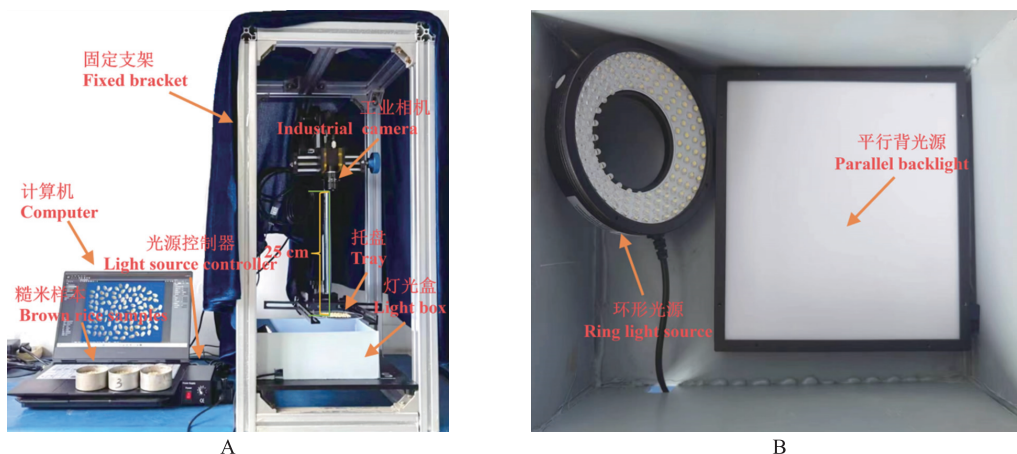


图 2 糙米裂纹图像采集系统

Fig. 2 The image acquisition system for brown rice cracking

A.图像采集系统 Image acquisition system;B.光照系统 Lighting system.

按照《粮食、油料检验黄粒米及裂纹粒检验法:GB/T 5496—1985》与《谷物烘干机:DG/T 017—2021》标准关于爆腰率检测规定,每次取约 100 粒整粒糙米放入托盘采集图像,每张图像只包括 1 种类型的糙米,每种类型糙米采集 175 张图像,共采集 525 张图像,每张图像分辨率大小均为 4 608×3 456。所采集的部分图像如图 3 所示。

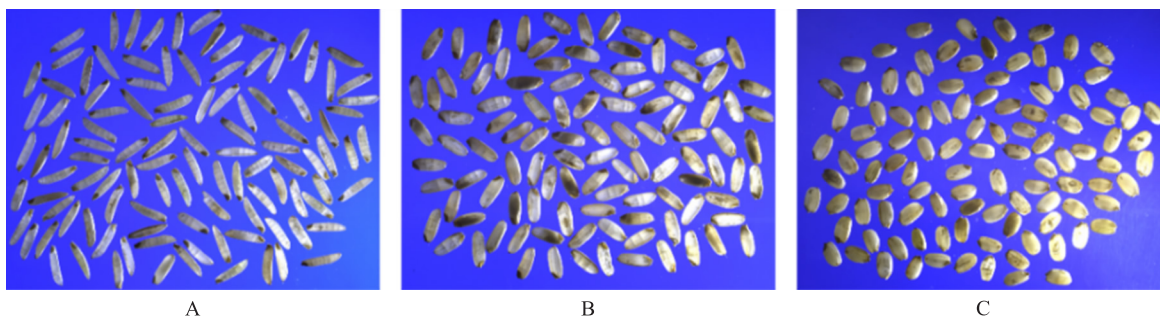


图 3 糙米样本原始图像

Fig. 3 The raw images of brown rice samples

A.长粒糙米 Long brown rice;B.中粒糙米 Medium brown rice;C.圆粒糙米 Round brown rice.

1.1.2 糙米图像预处理 将所采集的糙米样本原始图像进行预处理,以获取单粒糙米图像来构建糙米裂纹数据集。图像预处理流程如图 4 所示:1)分析输入图像特征,由于图像背景为蓝色,糙米前景为棕黄色,因此在 RGB(红色、绿色和蓝色)三通道中,背景的背景 B 通道值大于糙米前景,R 与 G 通道值小于糙米前景;2)对 525 张图像进行背景去除试验,依据 RGB 通道值得出,当设定背景分割指标值 $S=|(R+G)/2-B|$,分割阈值为 170, S 大于 170 的像素点设为 (0,0,0)、小于 170 的设为 (255,255,255) 时实现了对糙米前景与背景的有效分离;3)对去除图像的背景进行灰度化处理得到二值图像,并进行图像开运算操作填充糙米内部小孔洞;4)获取二值图中每粒糙米轮廓最小外接矩形的属性参数(矩形长、宽、中心点位置),根据所得参数裁剪输入图像的每粒糙米并将结果放置于所建立的黑色掩码下,进行旋转和仿射变换得到单粒糙米图像;5)为避免单粒糙米图像大小变化对模型性能的影响,对图像左右两边同时填充 $(H-W)/2 \times H$ 的

黑色掩码区域使图像长宽相同(H 为单粒糙米图像高度, W 为单粒糙米图像宽度),并改变图像分辨率为 640×640 以建立糙米裂纹数据集。

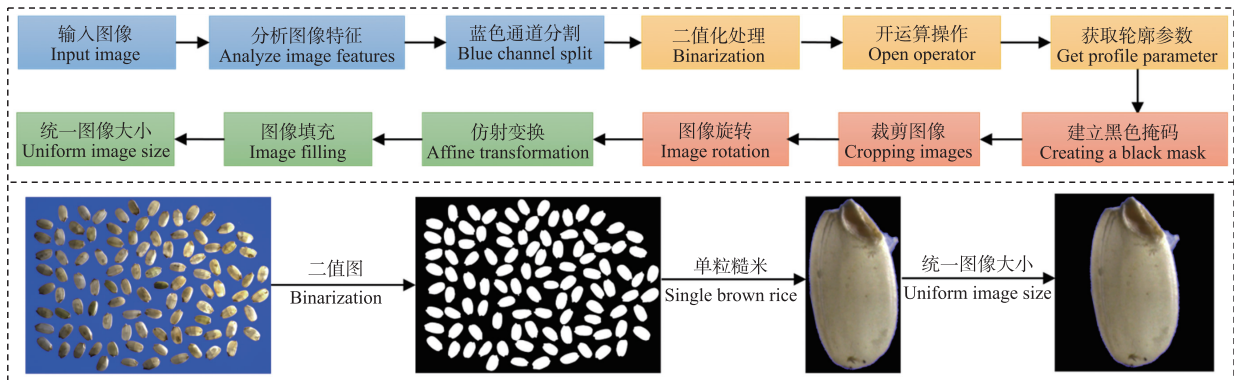


Fig. 4 The schematic diagram of brown rice image pre-processing process

1.1.3 糙米裂纹数据集建立及数据标注 经过图像预处理后生成的单粒糙米图像中,存在 2 676 张糙米裂纹图像。糙米裂纹数据集涵盖长粒、中粒和圆粒 3 种类型糙米;裂纹情况包括单条横向裂纹、多条横向裂纹和龟裂(同时存在横向和纵向裂纹)3 种情况,原始数据集如表 1 所示。

表 1 糙米裂纹数据集分类

Table 1 The classification of the brown rice cracking dataset

裂纹情况 Crack conditions	图像数 Number of images		
	长粒糙米 Long brown rice	中粒糙米 Medium brown rice	圆粒糙米 Round brown rice
单条横向裂纹 Single transverse crack	324	296	307
多条横向裂纹 Multiple transverse cracks	305	316	315
龟裂 Presence of transverse and longitudinal cracks	287	272	254

本文选取了原始数据集中不同长度糙米的 3 种不同裂纹情况的图像进行展示,如图 5 所示。

人工使用 LabelImg 软件进行糙米裂纹标签制作,标注结果如图 6 所示。当标注的矩形框长度 \geq 宽度,裂纹为横向裂纹;矩形框的长度 $<$ 宽度,裂纹为纵向裂纹。在图 6 的标注结果中,裂纹 1 和 3 为纵向裂

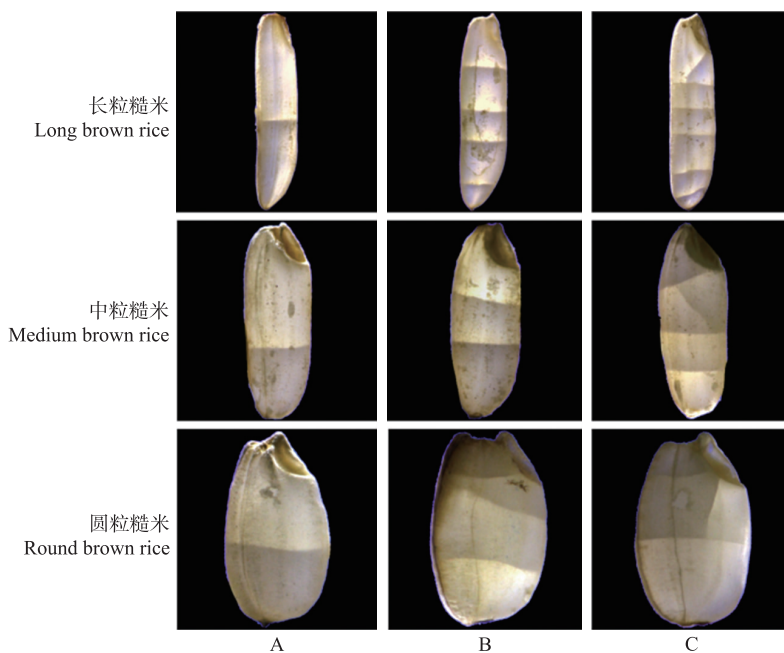


图 5 糙米裂纹图像示意图

Fig. 5 The schematic image of cracked brown rice

A.单条横向裂纹 Single transverse crack; B.多条横向裂纹 Multiple transverse cracks; C.龟裂 Presence of transverse and longitudinal cracks.

纹,裂纹 2 和 4 为横向裂纹。所生成的标签文件以 PASCAL VOC 数据集格式进行储存。

由于数据集只有裂纹一个标签,相对于背景而言,裂纹像素所占数量较少,导致出现类别不平衡问题。为缓解该问题,采用 Mosaic 和 Mixup 两种数字图像增强算法对原始数据集进行数据增强^[16]。通过增加裂纹目标的数量,减轻类别不平衡问题,帮助模型更好地学习和识别裂纹目标。数据增强后的数据集包括 Mosaic 数据增强的 2 676 张图像、Mixup 数据增强的 2 676 张图像及原始数据集的 2 676 张图像。数据集按照 8:1:1 的比例随机分为训练集 6 422 张图像、测试集 803 张图像、验证集 803 张图像。

1.2 YOLOv5-RF 模型的建立

YOLOv5 模型因其简单的网络结构和极快的推力速度得到了广泛应用。目前,YOLOv5 系列提供了 4 种型号:YOLOv5s、YOLOv5m、YOLOv5l 和 YOLOv5x。这些模型的网络结构类似,主要区别在于对网络特定位置的特征提取模块和卷积核的数量进行了缩放,模型的复杂度和大小不同^[17]。

YOLOv5 模型在大、中目标的检测场景中取得了显著的成绩,但对于小目标检测性能表现不理想^[17]。因此,需要对 YOLOv5 模型的小目标检测性能进行优化和改进,使其适用于糙米裂纹中存在的小目标裂纹检测。但是参数量较大的神经网络模型对训练所需硬件成本及部署设备的计算性能也有相应的较高要求。为了在模型识别精度和设备成本之间维持平衡,选择对 YOLOv5l 模型进行优化和改进。采取 3 个改进措施,分别是引入了辅助骨干网络增强主干网络在粗粒度下的小裂纹特征信息提取能力;在颈部网络添加逆向连接层增强相邻层信息交互,提高裂纹识别能力;在 SPPF 模块后添加 CBAM 模块用于提高语义特征质量。构建 YOLOv5-RF 模型(图 7),模型改进部分用红色表示。



图 6 糙米裂纹图像标注示意图

Fig. 6 The schematic labelling of brown rice cracking images

裂纹 1 和 3 为纵向裂纹,裂纹 2 和 4 为横向裂纹。Cracks 1 and 3 are longitudinal cracks, while cracks 2 and 4 are transverse cracks.

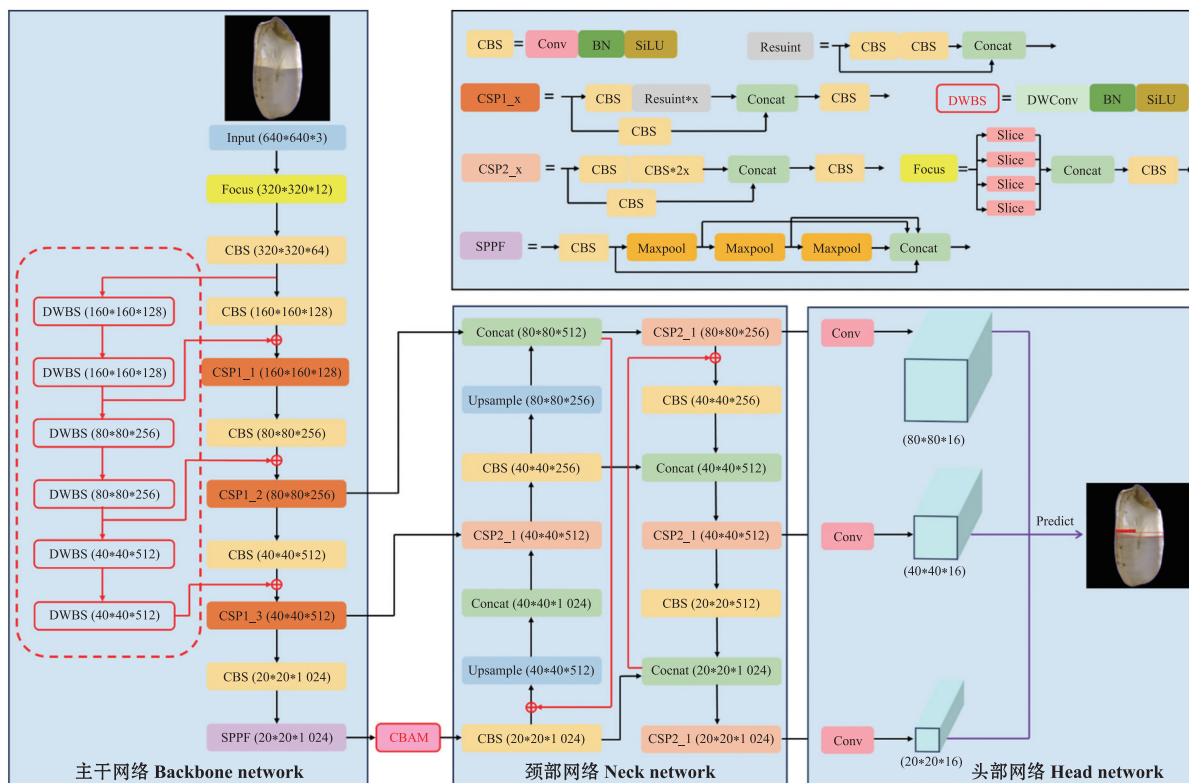


图 7 YOLOv5-RF 网络结构

Fig. 7 YOLOv5-RF network structure diagram

1.2.1 基于深度卷积的同级复合骨干网络 同级复合骨干网络结构^[18]是将 2 个不同的骨干网络结合起来增强网络的表达能力,并充分利用多尺度特征来提高小目标检测性能。本文在 YOLOv5l 主骨干网络中引入同级复合骨干网络结构进行特征提取,提高糙米图像中小裂纹信息提取能力。

由 Zhang 等^[19]的研究可知,深度卷积^[20]可以有效抑制卷积神经网络特征通道中的噪声,继而进一步提高卷积神经网络在粗粒度标签下的特征提取能力和语义信息表达能力。为此,通过使用深度卷积代替 CBS 模块中的普通卷积构建出 DWBS 模块,并将 DWBS 模块不断堆叠形成辅助骨干网络。辅助骨干网络与原始的 Darknet-53 主干网络相结合形成基于 DWBS 模块的同级复合骨干网络,如图 8 所示。

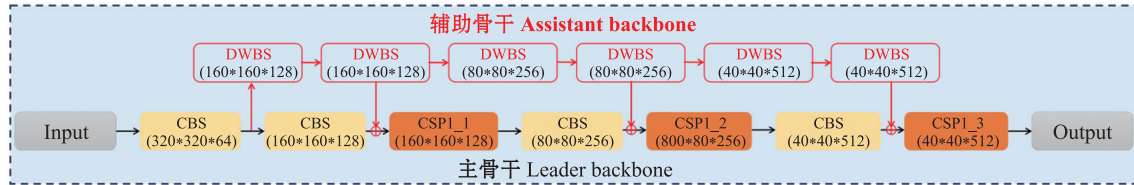


图 8 基于 DWBS 模块的同级复合骨干网络结构

Fig. 8 The composite backbone network structure of the same level based on the DWBS module

1.2.2 基于逆向连接层的特征融合网络 YOLOv5l 模型的颈部网络(neck)是将骨干网络所获得的特征信息聚合来提高检测头(head)对裂纹位置信息预测的准确率。颈部网络由特征金字塔网络(feature pyramid networks, FPN)结合路径聚合网络(path aggregation network, PANet)结构对特征进行多尺度融合。PANet 结构主要用于连接 2 个 FPN 模块;如图 9 所示的非红色部分,左边的 FPN 模块将深层语义信息向浅层传播;右边的 FPN 模块将浅层语义信息向深层传播,将 M3 深层位置信息和其他阶段的浅层纹理信息进行融合,并传输到 P5 模块中进一步提取特征在 P3 中完成深层实例信息的表达,使模型能够更好地处理不同大小的目标。但是这种方法过于依赖相邻层特征的聚集,导致长行程特征信息无法有效交互,减弱了神经网络对小目标特征信息的特征表达和提取能力。为此,如图 9 所示,在 PANet 的基础上,添加 2 个逆向连接层^[21]路径 ①和②用于聚合非相邻层之间的特征,以此来提高解码端对糙米裂纹识别能力。模型训练过程中,路径①和②均只需执行一次。

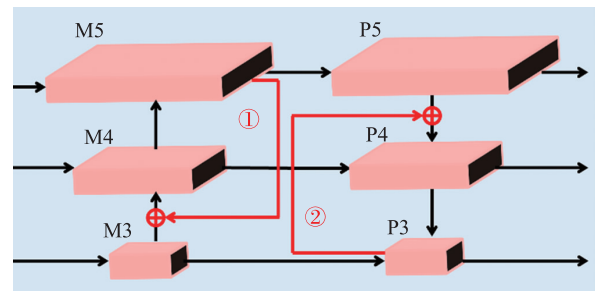


图 9 PANet+逆向连接层结构图

Fig. 9 The structure diagram of PANet+reverse cross-layer connection

红色部分为逆向连接层结构。The red part is the reverse cross-layer connection. M: 自顶向下路径增强 Top-down path augmentation; P: 自底向上路径增强 Bottom-up path augmentation.

路径①: M5 (80×80×512) 经过卷积核尺寸为 3×3、卷积步长为 2 的卷积层后,特征图尺寸变为 40×40×512;对特征图进行步长为 2 的最大池化操作,特征图尺寸变为 20×20×512;将所得特征图与 M3 (40×40×512) 进行相加操作。

路径②: P3 (20×20×1024) 经过卷积核尺寸为 1×1、卷积步长为 1 的卷积层后,特征图尺寸变为 40×40×256;进行上采样操作改变特征图尺寸为 80×80×256;将所得特征图与 P5 进行相加操作。

1.2.3 CBAM 注意力模块 在 YOLOv5l 中同时使用同级复合骨干网络和逆向连接层时,神经网络中出现了耦合表达,网络性能相比于单独添加逆向连接层时出现衰减。其根本原因是编码端的语义信息特征信号无法在解码端被有效解析成实例特征。为此,可使用 CBAM 注意力算法,通过通道注意力算法和空间注意力算法对语义信息进行加权调整,使编码信息可以在解码端进行实例信息解析。

YOLOv5l 神经网络中的 SPPF 模块使用 3 次 Max-pooling 算子对特征通道进行逐元素加权调整,然后通过 Concatenate 操作输出新的特征张量。然而,SPPF 模块中的逐元素通道注意力加权操作,无法有效维持语义信息的高效表达。在 SPPF 模块中使用 CBAM 注意力算法,通过 CBAM 复合注意力算法调整 SPPF 模块的通道注意力表达,是提高编码阶段的语义信息表达质量的有效方法^[22]。因此,构建了一种 SPPF 模块与 CBAM 模块结合的网络结构,如图 10 所示。利用 CBAM 注意力算法中的空间注意力模块,根据通道注意力中空间位置的重要性,自适应调整不同位置的特征响应,有效提高裂纹目标检测的精度。

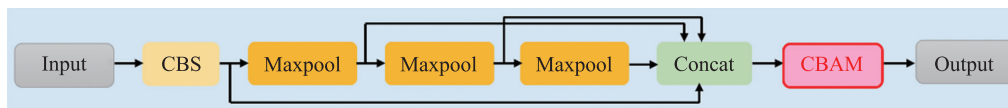


图 10 SPPF 模块结合 CBAM 模块网络结构图

Fig. 10 The structure diagram of SPPF module combined with CBAM module network

1.3 模型评估参数

由于数据集中仅存在稻谷裂纹单个类别,选用平均准确率(average precision, AP)作为模型性能的评价标准之一。其他模型性能评估指标包括准确率(precision)、召回率(recall)、F1 值(F1-score)、每秒传输帧数(frame per second, FPS),模型参数量(parameters)^[23]。准确率、召回率、F1 值和平均准确率的公式如下:

$$Precision = \frac{TP}{TP+FP} \quad (1)$$

$$Recall = \frac{TP}{TP+FN} \quad (2)$$

$$AP = \int_0^1 Precision(Recall) dr \quad (3)$$

$$F1\text{-score} = \frac{2 * Precision * Recall}{Precision + Recall} \quad (4)$$

式中:TP 是真阳性,表示预测为正例,实际为正例的数量;FP 是假阳性,表示预测为正例,实际为负例的数量;FN 为假阴性,表示预测为负例,实际为正例的数量。

1.4 试验条件和模型训练参数设置

试验所用显卡是 RTX3090,操作系统为 Windows 10,使用 Pytorch 深度学习框架构建模型,CUDA 版本为 11.2,CUDNN 版本为 8.1.1。改进的 YOLOv5-RF 选取图像尺寸为 640×640 像素的图像作为输入,图像均为 RGB 格式,冻结训练轮次为 50,非冻结训练轮次为 150,批量大小设为 8,选取模型 SGD 优化器作为训练优化方法,初始学习率为 0.01,循环学习率为 0.1,动量因子为 0.937,权重衰减系数为 0.000 5。模型训练时采用官方提供的 YOLOv5.pth 预训练权重。模型测试阶段,置信度阈值设置为 0.5,非极大值抑制阈值设为 0.3。

2 结果与分析

2.1 改进后模型的性能分析

YOLOv5-RF 模型与 YOLOv5l 模型在 60 轮至 150 轮的平均准确率如图 11 所示。由图 11 可知,2 种模型在迭代至 100 轮之后平均准确率均趋于稳定,未出现明显波动,YOLOv5-RF 模型相较于 YOLOv5l 模型对糙米裂纹的识别具有较高的平均准确率。

采用 YOLOv5-RF 模型对存在龟裂情况的糙米图像样本进行裂纹识别能力验证,结果如图 12 所示。从图 12 可以看出,YOLOv5-RF 模型能够对糙米图像中的裂纹进行准确识别。

2.2 不同检测算法的性能对比

为客观展示 YOLOv5-RF 模型的优越性,将其与具有代表性的两阶段目标检测 Faster R-CNN、一阶段目标检测算法 SSD、Retina-Net、Efficient-Det、YOLOv5l^[23-25] 进行性能比较,结果如表 2 所示。

从表 2 可见,与 YOLOv5-RF 模型相比,Faster R-CNN 模型的裂纹识别平均准确率达 91.66%,仅次于本文算法,但其参数量增加 77.56 MB,增加了应用端的设备成本和时间成本;SSD 模型参数量减少 25.52 MB,但裂纹识别的平均准确率为 87.62%;Retina-Net 模型裂纹识别的平均准确率为 88.03%,但其召回率为

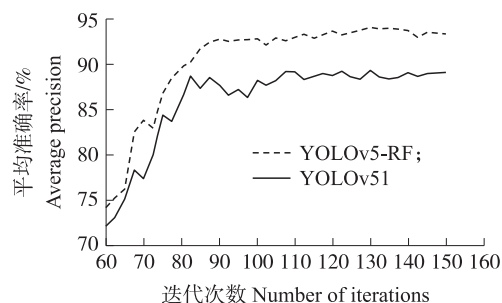


图 11 YOLOv5-RF 模型与 YOLOv5l 模型平均准确率对比

Fig. 11 Comparison of average precision between the YOLOv5-RF and YOLOv5l

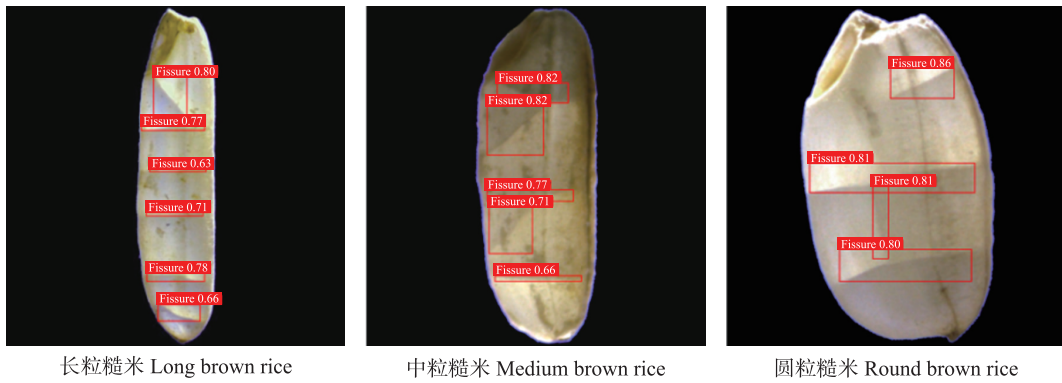


图 12 糙米图像裂纹识别结果

Fig. 12 The crack recognition results of brown rice image

77.65%, 导致其在稻谷爆腰率检测任务中的鲁棒性表现欠佳, 尤其在糙米裂纹识别环节存在裂纹漏检的情况; Efficient-Det 模型虽然参数量减少 23.40 MB, 但因准确率减少 8.48%, 导致其在稻谷爆腰率检测环节中存在裂纹误检情况。

YOLOv5-RF 模型与 YOLOv5l 模型相比, 参数量增加 2.5 MB, YOLOv5-RF 平均准确率较 YOLOv5l 提升 4.72%, 准确率提升 1.23%, 召回率提升 6.23%, F1-score 提升 3.50%。YOLOv5-RF 的 FPS 比 YOLOv5l 模型少 1.8, 对检测速度影响不大。在这几种模型中, 本文算法对裂纹识别的平均准确率最高, 达到 94.01%。

表 2 不同网络模型在糙米裂纹数据集上的测试结果

Table 2 The test results of different network models on the brown rice cracking dataset

模型 Model	参数量/MB Parameters	每秒传输帧数 Frames per second	准确率/% Precision	召回率/% Recall	F1 值 F1-score	平均准确率/% Average precision
Faster R-CNN	126.69	14.58	90.67	84.76	0.87	91.66
SSD	23.61	35.68	90.36	78.82	0.84	87.62
Retina-Net	37.98	23.14	86.42	77.65	0.82	88.03
Efficient-Det	25.73	28.52	82.57	84.81	0.84	85.35
YOLOv5l	46.63	17.56	89.62	80.69	0.85	89.29
YOLOv5-RF	49.13	15.67	90.85	86.92	0.89	94.01

不同模型的准确率-召回率曲线, 如图 13 所示。由图 13 和表 2 可得, YOLOv5-RF 模型在召回率较高的情况下依旧能够保持较高的准确率, 同时平均准确率分别比 Faster R-CNN、SSD、Retina-Net、Efficient-Det 和 YOLOv5l 高 2.35%、6.39%、5.98%、8.66% 和 4.72%。

为进一步验证 YOLOv5-RF 的性能, 选取 YOLOv5l 模型、Faster R-CNN 模型和 SSD 模型, 在图像样本存在龟裂、光照过曝和遮挡等复杂情况下进行检测, 结果如图 14 所示。YOLOv5l、Faster R-CNN 和 SSD 模型出现不同程度漏检和误检, YOLOv5-RF 模型没有发现漏检或误检的情况, 表明 YOLOv5-RF 能够更好地适应不同大小的裂纹目标, 提供更高的检测性能。

2.3 消融试验

为了进一步探究添加不同模块对模型性能的影响, 进行消融试验, 试验结果如表 3 所示。由表 3 可知, 添加任意一种模块, 都有助于提高模型的平均精度。特别是添加一个逆向连接层结构后, 模型的平均精度较添加同级复合骨干网络或 CBAM 注意力模块提高更加明显。

在 YOLOv5l 模型的基础上, 添加同级复合骨干网络或逆向连接层结构后再加入 CBAM 注意力模块, 相比仅添加同级复合骨干网络或逆向连接层结构, 更能提升模型精度。然而, 在同时使用同级复合骨干网络和逆向连接层结构时, 与仅添加逆向连接层结构相比, 模型精度会有所下降, 这可能是因为信号耦合效果导致编码端的语义信息特征信号无法有效解析成实例特征所致。但如果进一步添加 CBAM 注意力模块, 有助于模型更好地聚焦重要特征信息, 通过动态调整通道和空间注意力, 提升编码端的语义信息表达, 从而改善模型性能。

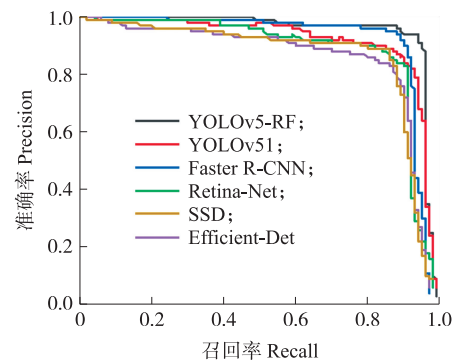


图 13 准确率-召回率曲线

Fig. 13 Precision-recall curve

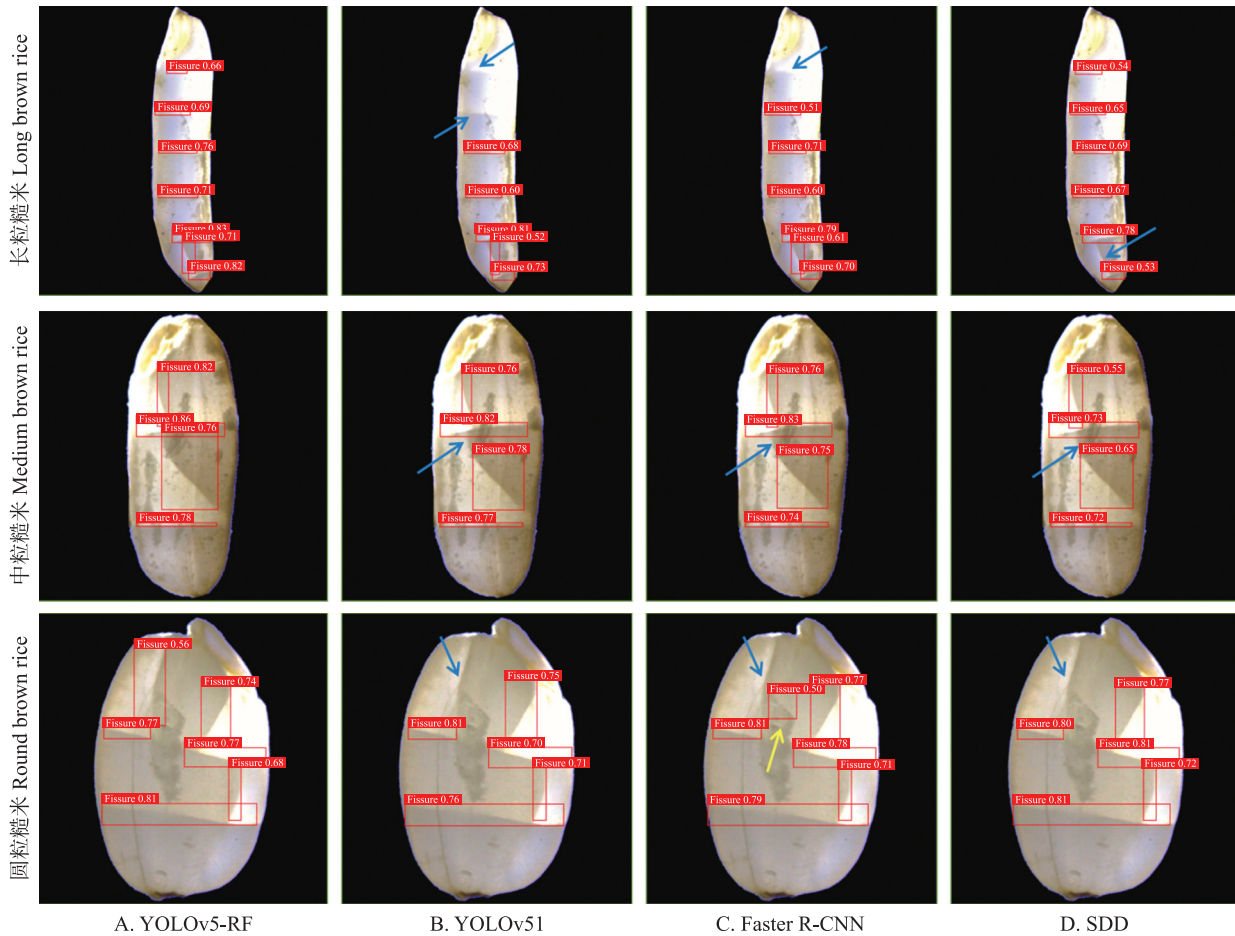


图 14 不同网络模型的糙米裂纹区域识别结果

Fig. 14 The results of brown rice crack region localization with different network models

蓝色箭头指向漏检,黄色箭头指向误检。The blue arrow points to missed detections, and the yellow arrow points to false detections.

表 3 消融试验

Table 3 The ablation experiment

模型 Model	参数量/MB Parameters	每秒传输帧数 Frames per second	平均准确率/% Average precision
YOLOv51	46.63	17.56	89.29
YOLOv51+基于 DWBS 模块的同级复合骨干网络 YOLOv51+the composite backbone network structure of the same level based on DWBS module	47.17	15.98	91.92
YOLOv51+逆向连接层结构 YOLOv51+reverse cross-layer connection	48.34	16.59	93.02
YOLOv51+CBAM	46.89	16.66	91.39
YOLOv51+基于 DWBS 模块的同级复合骨干网络+CBAM YOLOv51+the composite backbone network structure of the same level based on DWBS module+CBAM	47.43	16.26	92.68
YOLOv51+逆向连接层结构+CBAM YOLOv51+reverse cross-layer connection+CBAM	48.60	16.83	93.42
YOLOv51+基于 DWBS 模块的同级复合骨干网络+逆向连接层结构 YOLOv51+the composite backbone network structure of the same level based on DWBS module+reverse cross-layer connection	48.87	15.79	92.86
YOLOv5-RF	49.13	15.67	94.01

在模型参数量和检测速度方面,从表 3 可见,YOLOv5-RF 模型相较于 YOLOv51 模型,参数增加了 2.5 MB,FPS 有略微降低,仅为 2 帧,表明模块的添加对检测速度影响很小。因此,从保持算法性能、平衡硬件设施成本及检测效率方面综合考虑,所提出的 YOLOv5-RF 模型都能较好满足稻谷裂纹检测的实际要求。

3 结论

采用同级复合骨干网络和逆向连接层结构,尽管可以在粗粒度标签下提高特征提取能力和语义信息表达能力,显著提高模型识别精度,但由此产生的耦合表达会导致神经网络性能有所衰减。通过增加注意力算法模块CBAM,可以有效改善耦合表达现象,从而显著提高神经网络的性能。

针对糙米裂纹目标较小、尺度大小不一而不易提取其裂纹信息等问题,提出了一种基于改进YOLOv5网络的YOLOv5-RF模型,对糙米裂纹识别的平均准确率达94.01%,较YOLOv5模型高4.72%,也较主流的Faster R-CNN和SDD模型分别高2.35%和6.39%。从模型的识别精度、硬件成本及检测效率等综合考虑,YOLOv5-RF算法能够精确完成糙米裂纹识别,为稻谷爆腰检测提供技术支持。

参考文献 References:

- [1] 熊书剑,孙卫红,赵鲁宁,等. 稻谷分程干燥工艺优化[J]. 食品科学,2017,38(6):274-281.
Xiong S J, Sun W H, Zhao L N, et al. Optimization of three-stage drying of paddy[J]. Food Science, 2017, 38(6): 274-281 (in Chinese with English abstract).
- [2] 李维强. 减少稻米爆腰与破碎提高稻谷出米率[J]. 粮食加工, 2014, 39(3): 33-36.
Li W Q. Reduce rice bursting and crushing to increase rice yield[J]. Grain Processing, 2014, 39(3): 33-36 (in Chinese).
- [3] 陈坤杰,梁静,蒋维银,等. 储藏温度与初始含水率对稻谷和糙米食用品质的影响[J]. 农业机械学报, 2022, 53(9): 390-398.
Chen K J, Liang J, Jiang W Y, et al. Effect of storage temperature and initial moisture content on eating quality of rice and brown rice[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(9): 390-398 (in Chinese with English abstract).
- [4] 姚英政,董玲,朱宇. 烘干设备干燥效率比较及对稻谷爆腰率的影响[J]. 粮食储藏, 2017, 46(2): 21-24.
Yao Y Z, Dong L, Zhu Y. Comparison of drying efficiency of grain driers and effect on crack percentage of paddy kernel[J]. Grain Storage, 2017, 46(2): 21-24 (in Chinese with English abstract).
- [5] 吴中华,康宁,董晓林,等. 声发射法无损检测稻谷籽粒应力裂纹[J]. 农业工程学报, 2018, 34(16): 274-280.
Wu Z H, Kang N, Dong X L, et al. Non-destructive detection of stress cracking in rice kernel by acoustic emission[J]. Transactions of the Chinese Society of Agricultural Engineering, 2018, 34(16): 274-280 (in Chinese with English abstract).
- [6] 陈树人,徐李,尹建军,等. 基于Micro-CT图像处理的稻谷内部损伤定量表征与三维重构[J]. 农业工程学报, 2017, 33(17): 144-151.
Chen S R, Xu L, Yin J J, et al. Quantitative characterization of grain internal damage and 3D reconstruction based on Micro-CT image processing[J]. Transactions of the Chinese Society of Agricultural Engineering, 2017, 33(17): 144-151 (in Chinese with English abstract).
- [7] Shi H, Siebenmorgen T J, Luo H L, et al. Fissure detection and measurement in rough rice using X-ray imaging[J]. Transactions of the ASABE, 2019, 62(4): 859-866.
- [8] 吴杰,任奕林,王恒志. 稻谷加工断裂面裂纹显微图像识别研究[J]. 广东农业科学, 2014, 41(5): 213-216.
Wu J, Ren Y L, Wang H Z. Research on microscopic images identification of crack on processed grain fracture surface[J]. Guangdong Agricultural Sciences, 2014, 41(5): 213-216 (in Chinese with English abstract).
- [9] 许振伟. 基于多分辨率B样条小波分析的大米裂纹检测研究[J]. 中国粮油学报, 2012, 27(6): 101-104.
Xu Z W. Study about rice fissure detection based on multi-resolution B-spline wavelet analysis[J]. Journal of the Chinese Cereals and Oils Association, 2012, 27(6): 101-104 (in Chinese with English abstract).
- [10] 邢键,罗佳顺. 融合多维特征的稻米外观品质检测技术[J]. 哈尔滨理工大学学报, 2021, 26(5): 76-82.
Xing J, Luo J S. Rice appearance quality detection technology based on multi-dimensional features[J]. Journal of Harbin University of Science and Technology, 2021, 26(5): 76-82 (in Chinese with English abstract).
- [11] 王修善,谢方平,刘大为,等. 基于深度学习的稻谷内部损伤无损检测装置[J]. 实验技术与管理, 2022, 39(12): 112-117.
Wang X S, Xie F P, Liu D W, et al. Nondestructive testing device for internal damage of rice based on deep learning[J]. Experimental Technology and Management, 2022, 39(12): 112-117 (in Chinese with English abstract).
- [12] Du Y C, Pan N, Xu Z H, et al. Pavement distress detection and classification based on YOLO network[J]. International Journal of Pavement Engineering, 2021, 22(13): 1659-1672.
- [13] Amjoud A B, Amrouch M. Object detection using deep learning, CNNs and vision transformers: a review[J]. IEEE Access, 2023, 11: 35479-35516.
- [14] Zhang Z X, Lu X Q, Cao G J, et al. ViT-YOLO: transformer-based YOLO for object detection[C]//2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW). Montreal, BC, Canada. IEEE, 2021: 2799-2808.
- [15] 何家峰,陈宏伟,骆德汉. 深度学习实时语义分割算法研究综述[J]. 计算机工程与应用, 2023, 59(8): 13-27.
He J F, Chen H W, Luo D H. Review of real-time semantic segmentation algorithms for deep learning[J]. Computer Engineering and Applications, 2023, 59(8): 13-27 (in Chinese with English abstract).
- [16] Yan B, Fan P, Lei X Y, et al. A real-time apple targets detection method for picking robot based on improved YOLOv5[J]. Remote Sensing,

- 2021, 13(9): 1619.
- [17] 马瑞峻, 何浣冬, 陈瑜, 等. 基于改进 YOLOv5 的火龙果成熟度识别方法[J]. 沈阳农业大学学报, 2023, 54(2): 196-206.
Ma R J, He H D, Chen Y, et al. Recognition method for maturity of pitaya based on improved YOLOv5[J]. Journal of Shenyang Agricultural University, 2023, 54(2): 196-206 (in Chinese with English abstract).
- [18] Liang T T, Chu X J, Liu Y D, et al. CBNet: a composite backbone network architecture for object detection[J]. IEEE Transactions on Image Processing, 2022, 31: 6893-6906.
- [19] Zhang J Y, Qi C, Mecha P, et al. Pseudo high-frequency boosts the generalization of a convolutional neural network for cassava disease detection[J]. Plant Methods, 2022, 18(1): 136.
- [20] Chollet F. Xception: deep learning with depthwise separable convolutions [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI, USA. IEEE, 2017: 1800-1807.
- [21] Qiao S Y, Chen L C, Yuille A. DetectoRS: detecting objects with recursive feature pyramid and switchable atrous convolution [C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville, TN, USA. IEEE, 2021: 10208-10219.
- [22] Woo S, Park J, Lee J Y, et al. CBAM: convolutional block attention module [C]//European Conference on Computer Vision. Cham: Springer, 2018: 3-19.
- [23] 徐会杰, 黄仪龙, 刘曼. 基于改进 YOLOv3 模型的玉米叶片病虫害检测与识别研究[J]. 南京农业大学学报, 2022, 45(6): 1276-1285.
DOI: 10.7685/jnau.202110039.
Xu H J, Huang Y L, Liu M. Research on pest detection and identification of corn leaf based on improved YOLOv3 model[J]. Journal of Nanjing Agricultural University, 2022, 45(6): 1276-1285 (in Chinese with English abstract).
- [24] 王大庆, 禄琳, 于兴龙, 等. 基于深度迁移学习的 EfficientNet 玉米叶部病害识别[J]. 东北农业大学学报, 2023, 54(5): 66-76.
Wang D Q, Lu L, Yu X L, et al. Maize leaf diseases identification using EfficientNet based on deep-transfer learning[J]. Journal of Northeast Agricultural University, 2023, 54(5): 66-76 (in Chinese with English abstract).
- [25] 朱红春, 李旭, 孟场, 等. 基于 Faster R-CNN 网络的茶叶嫩芽检测[J]. 农业机械学报, 2022, 53(5): 217-224.
Zhu H C, Li X, Meng Y, et al. Tea bud detection based on faster R-CNN network [J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(5): 217-224 (in Chinese with English abstract).

责任编辑: 沈 波