

基于视觉 Transformer 的马铃薯薯形 与大小自动分级

Automatic grading of potato shape and size based on visual Transformer

樊宏鹏¹ 于鹏飞¹ 杨森²

FAN Hongpeng¹ YU Pengfei¹ YANG Sen²

(1. 兰州资源环境职业技术大学信息工程学院, 甘肃 兰州 730021;

2. 甘肃农业大学机电工程学院, 甘肃 兰州 730070)

(1. Department of College of Information Engineering, Lanzhou Resources & Environment
VOC-teach University, Lanzhou, Gansu 730021, China; 2. College of Mechanical and
Electrical Engineering, Gansu Agricultural University, Lanzhou, Gansu 730070, China)

摘要: [目的] 解决基于人工特征设计分级方法精度低、鲁棒性差的问题。 [方法] 提出了一种强泛化的 Transformer 薯形与大小自动分级方法。基于 Transformer 模型构建 2 个 PotatoViT 模型, 并完成马铃薯薯形分级和大小分级任务; 利用迁移策略和数据增广方法训练出鲁棒性分级模型; 通过测试集定量分析, 验证了研究所提方法在马铃薯分级中的有效性。 [结果] PotatoViT 模型对薯形分级的准确率和模型 F_1 得分分别为 96.36%, 94.75%, 对大小分级的准确率和模型 F_1 得分分别为 89.66%, 85.16%, 分级精度优于 VGG16、ResNet50 和 MobileNetV3 网络模型。 [结论] 研究所提方法对马铃薯薯形与大小的准确、实时检测是可行的。

关键词: 马铃薯; 薯形; 大小; 分级; Transformer

Abstract: [Objective] Aiming at the problems of low accuracy and poor robustness in the previous classification methods based on artificial features. [Methods] A strong generalization automatic classification method of potato shape and size was proposed in this study. First, two potato ViT models were built based on Transformer model to complete potato shape grading and size grading tasks in parallel. Secondly, a robust model was trained by using migration strategy and data augmentation method. Finally, the effectiveness of this method in potato grading was verified by quantitative analysis of test sets.

基金项目: 甘肃省教育厅高校教师创新基金项目 (编号: 2023A-226); 兰州资源环境职业技术大学科研能力提升项目 (编号: X2023A-11)

作者简介: 樊宏鹏, 男, 兰州资源环境职业技术大学讲师, 硕士。

通信作者: 于鹏飞 (1985—), 男, 兰州资源环境职业技术大学副教授, 学士。E-mail: yupfei@lzre.edu.cn

收稿日期: 2023-10-17 **改回日期:** 2024-05-08

[Results] The experimental results show that the potato ViT model achieves 96.36% and 94.75% for potato shape classification, and 89.66% and 85.16% for size grading in terms of accuracy and μ_{F_1} index. The classification accuracy was better than VGG16, ResNet50 and MobileNetV3 network models.

[Conclusion] The results shows that it is feasible to apply the proposed method for the real-time and accurate detection of potato shape and size. The results of this study can provide theoretical and technical support for potato intelligent grading.

Keywords: potatoes; shape; size; grading; Transformer

马铃薯形状和大小的准确分级是马铃薯商品化处理中的关键环节, 对马铃薯的市场占有量起重要作用^[1]。目前, 马铃薯外部检测都是人工进行的, 存在劳动密集、耗时和效率低的缺点, 而且由于工人的主观性分级结果可能存在差异。

因计算机视觉技术对马铃薯的一些外部特征进行检测具有客观性强、成本低、精度高等优点^[2], 机器学习算法逐渐被应用于马铃薯形状、重量、机械损伤和绿皮等检测。孔彦龙等^[3]通过提取马铃薯俯视图像的 6 个不变矩参数, 利用神经网络和回归方法建立马铃薯形状与质量分级模型。Su 等^[4]利用机器视觉技术和体积线性回归模型预测马铃薯样品质量, 并根据深度图像的三维形状信息和图像处理算法自动检测凹凸、弯曲、凹陷等外观缺陷。王红军等^[5]利用数字图像处理技术获得每个马铃薯样本的图像特征参数, 建立马铃薯形状和重量检测模型。邓立苗等^[6]提取马铃薯形状、颜色特征后, 采用近似椭圆法进行形状检测。以上研究利用人工设计特征的方式提取马铃薯的简单几何特征、形状特征, 结合分类器实现无

损检测和分级。但该类分级模型需要人工设计特征,过程非常繁琐、费力,且模型稳定性差。

随着机器学习技术的发展,深度学习技术在农业领域目标识别^[7]、检测^[8]和语义分割^[9]中得到了广泛的应用。利用反向传播技术训练的深度卷积神经网络(DNN)可以自动学习特征提取器,很大程度上提高目标分类和检测精度^[10-11]。深度学习方法也被用于马铃薯和水果分级检测,同时也实现了较高的精度。杨森等^[12]基于Xception网络构建轻量化深度卷积网络,通过迁移学习实现了马铃薯5类外部缺陷的快速识别,平均识别准确率达到96.04%。汪成龙等^[13]利用ResNet、DenseNet和CaffeNet3种深度卷积神经网络模型开展了马铃薯畸形检测,试验结果表明,网络的准确率能够满足实际设备运行要求。张兆国等^[14]针对马铃薯联合收获机在作业过程中分级清选问题,利用改进的YoloV4模型对自然作业环境下的马铃薯数量以及损伤情况进行检测,完成了正常马铃薯和机械损伤马铃薯的精确区分。史方青等^[15]利用YOLOv3网络检测马铃薯芽眼,平均识别精度达到98.44%。以上深度学习方法在图像上使用卷积滤波器来提取重要特征,以理解图像中感兴趣的对象。计算机视觉中注意力机制能够关注图像重要特征,对模型性能改善显著。但由于图像中的像素数量远大于文本中的单词数量,使得注意力机制在视觉应用上受到了很大的限制^[16]。考虑到马铃薯外观形态的多样性和使用CNN网络检测薯形和大小的缺陷,研究拟提出一种基于视觉Transformer的马铃薯薯形与大小自动分级方法,以期实现马铃薯外观品质的快速、无损和准确分级。

1 图像数据的采集与处理

图像收集对象来自马铃薯主产区甘肃定西,品种为新大坪、陇薯7号和青薯9号。根据马铃薯市场品质分级准则,将薯形划为类圆形、椭圆形和畸形。同样,在马铃薯大小分级方面,根据NY/T 1066-2006《马铃薯等级规格》,将马铃薯划分为大(L)、中(M)、小(S)3个等级。其中,单薯质量≥300g为L等级,100g<单薯质量<300g为M等级,单薯质量≤100g为S等级。试验过程中通过三脚架将相机固定高度为25cm,设置背景为白色,保持相同光照条件对每类样本进行图像采样,图像分辨率为2041像素×2721像素。在采集马铃薯大小分级样本时,采用电子秤(惠州市英衡电子科技有限公司)对马铃薯进行称重(精度0.1g),标记出对应样本等级标签,并利用图像采集装置获取样本图像,马铃薯薯形样本示例如图1所示。为了增加样本的多样性和容量,同时消除拍摄角度对分级精度的影响,在图像采集过程中获取马铃薯正视图、左视图和右视图。研究共建立薯形和大小分级2个数据集,薯形数据集样本500幅图像,大小分级数据集



图1 马铃薯薯形样本示例

Figure 1 Example of potato shape sample

300幅图像。每类数据集按照7:3分成训练集和测试集。

2 马铃薯薯形检测与大小分级模型

2.1 整体技术路线

马铃薯薯形与大小自动分级所采用的技术路线如图2所示。先利用数据扩增方法扩充图像数据集,并标记数据标签;再采用迁移学习策略将ViT模型在公共数据集上预训练,形成马铃薯分级预训练模型;最后利用不同数据集评估ViT模型在马铃薯薯形和大小分级中的潜力。

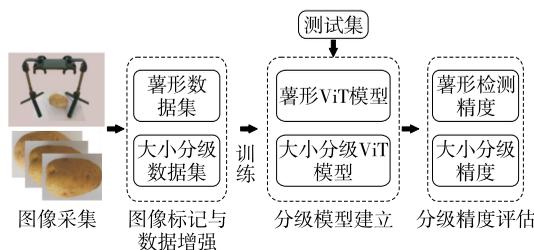


图2 技术路线

Figure 2 Technical route

2.2 薯形与大小分级模型的构建

2.2.1 视觉Transformer模型 目前视觉Transformer(ViT)在图像分类、检测中具有突出优势,该网络在需要较少计算资源的情况下获得了较好分类性能。在图像处理中使用ViT模型,能够显著提高网络训练速度,进一步增强模型在马铃薯分级中的应用性能^[17]。传统的CNN网络结构通常利用带有局部接受域的过滤器提取信息,而ViT通过注意力机制可以关注图像的不同区域,并在整个图像中实现信息集成。从现有研究^[18]看出,在使用足够的数据进行训练时,ViT表现性能优于先进的CNN网络。

图3显示了用于薯形和大小分类的ViT模型。整体上马铃薯分级模型主要由嵌入层、Transformer骨干层和分类层构成,并将其命名为PotatoViT模型。由于标准ViT模型将嵌入的一维序列作为输入,使得二维图像信息无法直接输入网络。为了ViT模型能够处理图像信息,采用图像切分思想将输入图像细分为不重叠的图像块。

Transformer为处理二维的图像信息,对输入ViT模型的任意一张图像I可以表示为:

$$I \in \mathbf{R}^{w \times h \times c}, \tag{1}$$

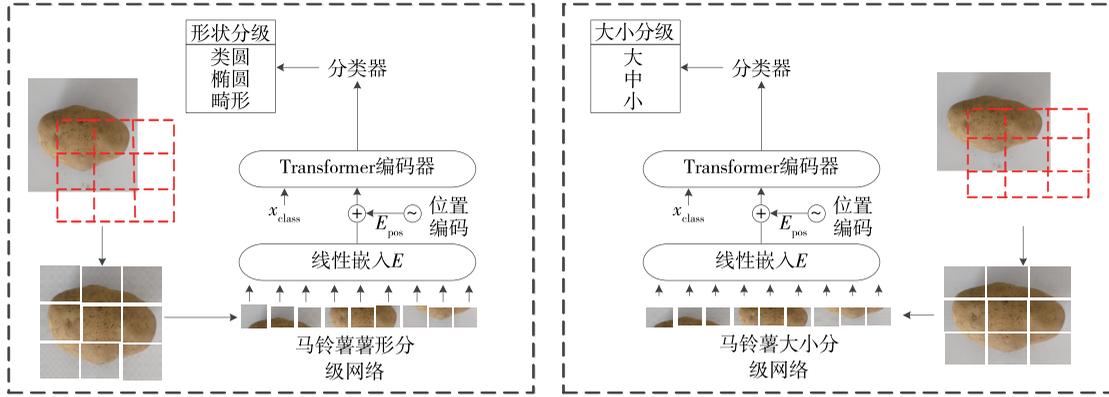


图3 马铃薯薯形与大小分级网络模型

Figure 3 Potato shape and size hierarchical network model

式中:

w ——图像的宽,mm;

h ——图像的高,cm;

I ——输入的任意图像;

R ——输入图像对应的三维矩阵;

c ——图像的通道数。

为了处理二维图像,将输入图像重塑为时序的二维图像块,具体表示如下:

$$I \in \mathbf{R}^{n \times (p^2 \cdot c)}, \quad (2)$$

式中:

p ——割块尺寸,mm;

n ——输入序列的长度($n = \frac{w \times h}{p \times p}$),mm;

c ——通道数。

2.2.2 线性嵌入层 利用可学习的线性投影将扁平化的图像块 x_n 映射到 D 维嵌入空间。ViT 模型将嵌入的图像块视为一组图像块集合,并不考虑它们的顺序关系。为了对图像块空间信息进行编码,利用编码器学习特定位置的嵌入向量,并将其添加到斑块嵌入中,以保留位置信息。将图像分割的图像块、空间位置信息和标签类别重塑为一维向量作为 Transformer 编码器输入,如式(3)所示^[19]。

$$z_0 = [x_{\text{class}}; x_1 E; x_2 E; \dots; x_n E] + E_{\text{pos}}, \quad (3)$$

式中:

z_0 ——编码器的输入序列;

x_{class} ——可学习的分类标记;

x_n ——第 n 个图像块的特征向量;

E ——可学习的嵌入矩阵, $E \in \mathbf{R}^{n \times (p^2 \times c) \times D}$;

E_{pos} ——序列的位置信息, $E \in \mathbf{R}^{(n+1) \times D}$ 。

2.2.3 Transformer 编码器 Transformer 编码器主要由多头注意力模块(MSA)和全连接前馈密集块(MLP)两个核心模块构成,具体结构如图4所示。为了协调图块序

列中不同元素之间的多个复杂关系,在每个层中部署MSA模块。在每个MSA和MLP的头部和尾部添加Layer Norm(LN)处理。因此,第 l 层MSA的输出表示为:

$$z'_l = M_{\text{SA}}(L_N(z_{l-1})) + z_{l-1}, \quad (4)$$

$$z_l = M_{\text{LP}}(L_N(z'_l)) + z'_l, \quad (5)$$

式中:

l ——第 l 层的编码器,其中 $l=1, \dots, L$;

L ——Transformer有 L 层编码器;

$L_N(\cdot)$ ——归一化算子;

$M_{\text{SA}}(\cdot)$ ——多头注意力模块层的编码;

$M_{\text{LP}}(\cdot)$ ——全连接前馈密集块的编码;

z_{l-1} ——第 $l-1$ 个编码器模块输出值;

z'_l ——第 l 个多头注意力模块的输出;

z_l ——编码后的图像表示。

3 试验结果与分析

3.1 试验环境参数设置与训练策略

模型的训练和测试试验在 Ubuntu 18.04 系统下进行,硬件系统规格:GPU为NVIDIA GeForce RTX 3090,CPU为Intel XEON,所有代码都使用Pytorch实现,该框架是利用python编写的开源深度神经网络库。模型训练

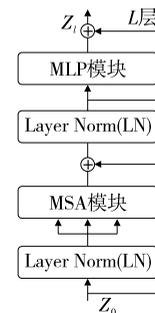


图4 Transformer 编码器模块结构

Figure 4 Transformer encoder module structure

参数一般设置:初始学习率为 0.001,减少系数为 0.1, batch size 为 8,训练轮数 200。为了提高 PotatoViT 模型的泛化能力和鲁棒性,模型训练分 3 个阶段。

(1) 首先利用 ImageNet 数据集训练 PotatoViT 模型,获得模型初始权重。

(2) 将薯形样本集和大小样本集组合为 1 个数据集,通过迁移学习利用融合数据集对初始模型进行进一步训练,获得初步分级模型。

(3) 利用薯形样本集和大小样本集分别二次调整初步分级模型,最终得到薯形分级模型和大小分级模型。

3.2 性能评价指标

利用精确度、召回率、模型 F_1 得分和准确率指标衡量 PotatoViT 模型的分级性能。根据分类器的真类和预测类的组合,可以将样本分为真阳性(T_p)、假阳性(F_p)、真阴性(T_n)和假阴性(F_n)4 种类型。具体计算式如式(6)~式(9)所示^[20-21]。

$$A_{cc} = \frac{T_p + T_n}{T_p + T_n + F_p + F_n}, \quad (6)$$

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

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

$$F_1 = \frac{2P \cdot R}{P + R}, \quad (9)$$

式中:

A_{cc} ——准确率;

P ——精确度;

R ——召回率;

T_p ——正确识别的正样本数量;

F_p ——错误识别的正样本数量;

T_n ——正确识别的负样本数量;

F_n ——错误识别的负样本数量;

F_1 ——模型 F_1 得分。

为了评估分类模型在所有类中的预测精度,对模型 F_1 得分平均值 μ_{F_1} 进行计算,如式(10)所示^[22]。

$$\mu_{F_1} = \frac{\sum_{i=1}^M F_1}{M}, \quad (10)$$

式中:

M ——马铃薯标签的数量。

3.3 分级精度评价

3.3.1 分块策略对分级精度的影响 马铃薯图像在输入 PotatoViT 模型时需要将图像分割为相同尺寸的图像块,并将分割的图像块依据空间位置有序排列并输入网络模型。图块尺寸(p)直接决定输入时序的长度, p 越小,时序越长,反之亦然。为了筛选 PotatoViT 模型最优图块尺寸参数,将 p 设置为 16×16 和 32×32 两种情形开展试

验,并将对应参数的模型命名为 PotatoViT-16 和 PotatoViT-32。表 1 统计了两种分块策略的马铃薯薯形和大小分级精度。 p 设置为 16×16 时,薯形分级的准确率和模型 F_1 得分平均值(μ_{F_1})分别为 96.36%,94.75%,相较于 p 设置为 32×32 时的提高了 2.82,4.03 个百分点。在大小分级上, p 设置为 16×16 时的准确率和模型 F_1 得分平均值 μ_{F_1} 分别为 89.66%,85.16%,较于 p 设置为 32×32 时的提高了 3.28,4.26 个百分点。试验结果表明, p 设置为 16×16 时,PotatoViT 模型能够有效实现马铃薯薯形和大小分级任务,模型判决性能和鲁棒性较好。因此,后续试验中选取 PotatoViT-16 作为分级模型。

表 1 不同图块尺寸下马铃薯形状分级精度评估

Table 1 Potato shape grading accuracy evaluation under different block size %

形状	16×16		32×32	
	准确率	μ_{F_1}	准确率	μ_{F_1}
薯形分级	96.36	94.75	93.54	90.72
大小分级	89.66	85.16	86.38	80.90

3.3.2 PotatoViT 模型分级精度评估 为探讨 PotatoViT 模型对每一类马铃薯的识别性能和差异性,将薯形测试集应用到 PotatoViT-16 模型建立混淆矩阵,如表 2 所示。混淆矩阵能够直观反映预测正确和错误的样本数。混淆矩阵的对角线为预测正确的马铃薯图像数量,其他位置为误判的马铃薯图像数量。同时,利用 P 、 R 和 F_1 指标评价 PotatoViT-16 模型的马铃薯分级精度。整体来看,畸形类别的识别精度最高, P 、 R 和 F_1 分别达到 97.49%,98.98%,98.23%,3 个评价指标均高于类圆和椭圆。类圆和椭圆类别之间误判的概率大,而畸形样本识别的稳定性较好。主要原因是测试样本中类圆和椭圆的类间差异小,特征区分度低,导致两类样本容易误判断。试验结果表明,PotatoViT-16 模型对薯形分级是可行的,能够满足马铃薯在线分级精度要求。

同样,利用 PotatoViT-16 模型统计了大小分级混淆矩阵,如表 3 所示。从表 3 可以看出,模型对 3 类样本的识别精度整体较高, P 、 R 和 F_1 均达到 80% 以上。但是

表 2 PotatoViT-16 模型的薯形分级精度

Table 2 Potato shape classification accuracy of PotatoViT-16 model

薯形	预测的样本数			$P/\%$	$R/\%$	$F_1/\%$
	类圆	椭圆	畸形			
类圆	252	10	0	96.18	91.97	94.02
椭圆	20	201	2	90.13	93.93	91.99
畸形	2	3	194	97.49	98.98	98.23

表 3 PotatoViT-16 模型的大小分级精度

Table 3 Size classification accuracy of PotatoViT-16 model %

大小	预测的样本数			P/%	R/%	F ₁ /%
	大	中	小			
大	224	25	11	86.15	89.24	87.67
中	18	208	21	84.21	82.21	83.20
小	9	20	168	85.28	84.00	84.63

3 类样本之间存在错误识别样本,中等级样本错误识别为大类别样本和小类别样本的数据相对较多。由于马铃薯外观形态多样化,受到训练数据集的限制,模型未能完全捕捉和学习不同形态样本,使得马铃薯大小分级精度整体低于薯形分级。

3.3.3 不同网络模型的分级精度评估 基于深度卷积网络(DCNN)方法已被应用于许多目标分类任务中。为了验证研究提出的 PotatoViT-16 模型的优越性,将其与目前典型的 CNN 网络模型 VGG16、ResNet50 和 MobileNetV3 进行了比较。在比较的 3 种模型中,VGG16 是浅层网络,ResNet50 是深层网络,而 MobileNetV3 是轻量化的网络。利用马铃薯训练集分别对上述 4 种模型进行训练,然后利用测试集对不同模型的性能进行评估,表 4 显示了不同模型 μ_{F_1} 和准确率的比较。通过对测试结果的分析可以看出,MobileNetV3 底层信息语义值相对较低,分级效果较差。由于 ResNet50 残差单元的添加和网络层数的增加,使其表达的特征也更丰富,相应的分类性能与 PotatoViT-16 模型接近。PotatoViT-16 模型对薯形分级和大小分级的整体结果均优于其他对比网络,能够实现高准确率和 μ_{F_1} 性能,同时降低权重大小和计算复杂度。

表 4 不同网络模型对比

Table 4 Comparison of different network models %

模型	薯形分级		大小分级	
	准确率	μ_{F_1}	准确率	μ_{F_1}
VGG16	94.13	91.58	85.14	78.97
ResNet50	95.86	94.03	89.01	84.14
MobileNetV3	91.13	87.24	80.12	72.59
PotatoViT-16	96.36	94.75	89.66	85.16

4 结论

(1) 研究测试了两种分块策略对 ViT 模型分级精度的影响。使用 16×16 分块策略的模型准确率和 F_1 得分平均值高于 32×32 分块策略,表明图块尺寸设置为 16×16 更适合马铃薯薯形及大小分级任务。

(2) PotatoViT-16 模型是一种提取全局特征的非局

部模型,能够增强全局特征提取,对薯形分级和大小分级的准确率达到 96.36%,89.66%。试验结果表明,研究提出的 PotatoViT-16 模型与传统 DCNN 模型相比具有更好的识别性能。但 PotatoViT-16 模型对马铃薯大小分级在标注不准确或者不一致的情况下精度低,且对拍摄位置、图像分辨率有严格的要求,无法有效适应实际生产高效分级要求。在未来的研究工作中将进一步改进网络模型结构和挖掘马铃薯外形与大小检测的敏感特征,利用多任务学习实现多个指标的同步预测。

参考文献

- [1] 肖伟中,高迎春,陈海需,等.小型马铃薯捡拾分级机的设计与试验[J].农机化研究,2019,41(12):130-134.
XIAO W Z, GAO Y C, CHEN H X, et al. Design and test of a small potato pickup and grading machine [J]. Journal of Agricultural Mechanization Research, 2019, 41(12): 130-134.
- [2] 向静,何志良,汤林越,等.结合计算机视觉的马铃薯外部品质检测技术[J].计算机工程与应用,2018,54(5):165-169.
XIANG J, HE Z L, TANG L Y, et al. Research of potato quality detection technology based on computer vision [J]. Comp Engin Appl, 2018, 54(5): 165-169.
- [3] 孔彦龙,高晓阳,李红玲,等.基于机器视觉的马铃薯质量和形状分选方法[J].农业工程学报,2012,28(17):143-148.
KONG Y L, GAO X Y, LI H L, et al. Potato grading method of mass and shapes based on machine vision [J]. Transactions of the Chinese Society of Agricultural Engineering, 2012, 28(17): 143-148.
- [4] SU Q H, KONDO N S, LI M Z, et al. Potato quality grading based on machine vision and 3D shape analysis [J]. Computers and Electronics in Agriculture, 2018, 152: 261-268.
- [5] 王红军,熊俊涛,黎邹邹,等.基于机器视觉图像特征参数的马铃薯质量和形状分级方法[J].农业工程学报,2016,32(8):272-277.
WANG H J, XIONG J T, LI Z Z, et al. Potato grading method of weight and shape based on imaging characteristics parameters in machine vision system [J]. Transactions of the Chinese Society of Agricultural Engineering, 2016, 32(8): 272-277.
- [6] 邓立苗,杜宏伟,徐艳,等.基于机器视觉的马铃薯智能分选方法与实现[J].中国农机化学报,2015,36(5):145-150.
DENG L M, DU H W, XU Y, et al. Implementation of intelligent potato grading method based on vision [J]. Journal of Chinese Agricultural Mechanization, 2015, 36(5): 145-150.
- [7] SHI R, LI T X, YAMAGUCHI Y S. An attribution-based pruning method for real-time mango detection with YOLO network [J]. Computers and Electronics in Agriculture, 2020, 169: 105214.
- [8] WU L, MA J, ZHAO Y H, et al. Apple detection in complex scene using the improved YOLOv4 model [J]. Agronomy, 2021, 11(3): 476.

- [9] 曲福恒, 李金状, 杨勇, 等. 基于改进 DeepLabv3+ 的轻量化作物杂草识别方法[J]. 石河子大学学报(自然科学版), 2024, 42(1): 117-125.
- QU F H, LI J Q, YANG Y, et al. Lightweight crop and weed recognition method based on improved DeepLabv3+ [J]. Journal of Shihezi University (Natural Science), 2024, 42(1): 117-125.
- [10] WU D H, LU S C, JIANG M, et al. Using channel pruning-based YOLO v4 deep learning algorithm for the real-time and accurate detection of apple flowers in natural environments[J]. Computers and Electronics in Agriculture, 2020, 178: 105742.
- [11] XIONG J T, LIU Z, CHEN S M, et al. Visual detection of green mangoes by an unmanned aerial vehicle in orchards based on a deep learning method [J]. Biosystems Engineering, 2020, 194: 261-272.
- [12] 杨森, 冯全, 张建华, 等. 基于轻量卷积网络的马铃薯外部缺陷无损分级[J]. 食品科学, 2021, 42(10): 284-289.
- YANG S, FENG Q, ZHANG J H, et al. Non-destructive classification of potato external defects based on lightweight convolutional network[J]. Food Science, 2021, 42(10): 284-289.
- [13] 汪成龙, 黄余凤, 庄学敏, 等. 基于深度学习的马铃薯畸形检测方法研究[J]. 惠州学院学报, 2018, 38(3): 61-66.
- WANG C L, HUANG Y F, ZHUANG X M, et al. Research on method method potato on deep potato [J]. Journal of Huizhou University, 2018, 38(3): 61-66.
- [14] 张兆国, 张振东, 李加念, 等. 采用改进 YoloV4 模型检测复杂环境下马铃薯[J]. 农业工程学报, 2021, 37(22): 170-178.
- ZHANG Z G, ZHANG Z D, LI J N, et al. Potato detection in complex environment based on improved YoloV4 model [J]. Transactions of the Chinese Society of Agricultural Engineering, 2021, 37(22): 170-178.
- [15] 史方青, 王虎林, 黄华. 基于卷积神经网络的马铃薯芽眼检测识别研究[J]. 中国农机化学报, 2022, 43(6): 159-165.
- SHI F Q, WANG H L, HUANG H. Research on potato buds detection and recognition based on convolutional neural network [J]. Journal of Chinese Agricultural Mechanization, 2022, 43(6): 159-165.
- [16] 贾伟宽, 孟虎, 马晓慧, 等. 基于优化 Transformer 网络的绿色目标果实高效检测模型[J]. 农业工程学报, 2021, 37(14): 163-170.
- JIA W K, MENG H, MA X H, et al. Efficient detection model of green target fruit based on optimized Transformer network [J]. Transactions of the Chinese Society of Agricultural Engineering, 2021, 37(14): 163-170.
- [17] 徐艳蕾, 孔朔琳, 陈清源, 等. 基于 Transformer 的强泛化苹果叶片病害识别模型[J]. 农业工程学报, 2022, 38(16): 198-206.
- XU Y L, KONG S L, CHEN Q Y, et al. Model for identifying strong generalization apple leaf disease using Transformer [J]. Transactions of The Chinese Society of Agricultural Engineering, 2022, 38(16): 198-206.
- [18] 周丽娟, 毛嘉宁. 视觉 Transformer 识别任务研究综述[J]. 中国图象图形学报, 2023, 28(10): 2 969-3 003.
- ZHOU L J, MAO J N. Vision Transformer-based recognition tasks: a critical review[J]. Journal of Image and Graphics, 2023, 28(10): 2 969-3 003.
- [19] BAZI Y, BASHMAL L, RAHHAL MM A, et al. Vision transformers for remote sensing image classification [J]. Remote Sensing, 2021, 13(3): 516.
- [20] CAI G, ZHU Y, WU Y, et al. A multimodal transformer to fuse images and metadata for skin disease classification[J]. The Visual Computer, 2023, 39(7): 2 781-2 793.
- [21] AFONSO M, FONTEIJN H, POLDER G, et al. Tomato fruit detection and counting in greenhouses using deep learning [J]. Frontiers in Plant Science, 2020, 11: 571299.
- [22] REEDHA R, DERICQUEBOURG E, CANALS R, et al. Transformer neural network for weed and crop classification of high resolution UAV images[J]. Remote Sensing, 2022, 14(3): 592.

(上接第 73 页)

- [16] 雷蕾, 石晨. 基于启发式搜索算法的自动采摘机器人路径规划研究[J]. 农机化研究, 2021, 43(7): 240-244.
- LEI L, SHI C. Path planning of automatic picking robot based on heuristic search algorithm[J]. Agricultural Mechanization Research, 2021, 43(7): 240-244.
- [17] 徐岩. 基于改进引力搜索算法的高速并联机器人轨迹优化[J]. 食品与机械, 2022, 38(5): 82-86.
- XU Y. Trajectory optimization of high-speed parallel robots based on improved gravity search algorithm[J]. Food & Machinery, 2022, 38(5): 82-86.
- [18] 张皓宇, 刘晓伟, 任川, 等. 并联机器人正运动学与 NURBS 轨迹规划[J]. 机械设计与制造, 2021, 12(4): 282-292.
- ZHANG H Y, LIU X W, REN C, et al. Parallel robot forward kinematics and NURBS trajectory planning[J]. Mechanical Design & Manufacturing, 2021, 12(4): 282-292.
- [19] 张灵枝, 黄艳, 于英杰, 等. 基于近红外光谱技术的六大茶类快速识别[J]. 食品与生物技术学报, 2024, 43(1): 48-59.
- ZHANG L Z, HUANG Y, YU Y J, et al. Rapid identification of six major tea categories based on near-infrared spectroscopy technology [J]. Journal of Food Science and Biotechnology, 2024, 43(1): 48-59.
- [20] 陈兴彬, 肖舜仁, 闵新和, 等. 基于 Workbench 和 nCode 工具的齿轮疲劳建模与寿命分析[J]. 机床与液压, 2022, 50(13): 149-154.
- CHEN X B, XIAO S R, MIN X H, et al. Gear fatigue modeling and life analysis based on Workbench and nCode tools[J]. Machine Tool and Hydraulic, 2022, 50(13): 149-154.