

基于 YOLO 11n – LDA 的苹果成熟度识别与低损采摘方法

牛金星^{1,2} 杨新源^{1,2} 张 涛^{1,2} 师树恒^{1,2} 于青源^{1,2} 毕明博^{1,2}

(1. 华北水利水电大学机械学院, 郑州 450011; 2. 郑州市机器人视觉与智能控制重点实验室, 郑州 450011)

摘要: 针对不同光照下苹果成熟度识别困难,不同成熟度苹果使用同一采摘模式损伤率较高的问题,提出了一种改进 YOLO 11n 不同光照下苹果成熟度识别模型 YOLO 11n – LDA,并与柔性夹爪结合实现不同成熟度苹果低损采摘的方法。首先,在 YOLO 11n 网络中的 SPPF 模块中加入具有大型可分离内核注意力 (Large separable kernel attention, LASK) 模块,增大模型感受野;然后在 C3k2 模块中加入动态蛇形卷积 (Dynamic snake convolution, DySnakeConv),提高模型多尺度特征感知能力;其次将 Neck 网络替换为渐进式特征金字塔 (Asymptotic feature pyramid network, AFPN) 网络,可以减少语义间隙,提高模型特征融合的稳定性。最后使用 ANSYS 仿真分析出不同成熟度苹果的最大安全采摘力,融合到柔性夹爪中实现低损采摘。在 PApple_RGB – D – Size 数据集中苹果分为食用成熟、采摘成熟、未成熟 3 类,改进后模型整体的平均精度均值 (mAP@0.5) 为 87.7%;食用成熟、采摘成熟和未成熟的平均精度均值 (AP@0.5) 分别为 86.6%、89.0% 和 87.5%;整体的 mAP@0.5 和 3 类成熟度的 AP@0.5 相较于 YOLO 11n 分别提高 3.9、8.0、1.6、2.1 个百分点。在 Orchard apple maturity 数据集中改进模型的 mAP@0.5 为 95.5%,相较于 YOLO 11n 提高 0.9 个百分点。采摘实验中检测率、采摘成功率和损伤率分别为 99.0%、84.0% 和 4.8%。

关键词: 苹果;成熟度识别;低损采摘;柔性夹爪;YOLO 11n – LDA

中图分类号: S126

文献标识码: A

文章编号: 1000-1298(2026)11-0343-11

OSID:



Apple Ripeness Detection and Low-damage Harvesting Based on YOLO 11n – LDA

NIU Jinxing^{1,2} YANG Xinyuan^{1,2} ZHANG Tao^{1,2} SHI Shuheng^{1,2} YU Qingyuan^{1,2} BI Mingbo^{1,2}

(1. School of Mechanical Engineering, North China University of Water Resources and Electric Power, Zhengzhou 450011, China

2. Zhengzhou Key Laboratory of Robot Vision and Intelligent Control, Zhengzhou 450011, China)

Abstract: Aiming to address the issue of high damage rates caused by a single harvesting mode for apples of varying maturities, an apple maturity identification model, YOLO 11n – LDA, was proposed based on YOLO 11n, combined with a flexible gripper to achieve low-damage harvesting for apples at different maturity stages. Firstly, the large separable kernel attention (LSKA) module was integrated into the SPPF module of the YOLO 11n network to expand the model's receptive field. Subsequently, dynamic snake convolution (DySnakeConv) was added to the C3k2 module to enhance the model's perception of multi-scale features. Furthermore, the asymptotic feature pyramid network (AFPN) replaced the neck network to bridge the semantic gap and improve feature fusion stability. Finally, the maximum safe picking force for apples of different maturities was determined via ANSYS simulation, and a comprehensive harvesting test platform, comprising a collaborative robotic arm, a flexible fin gripper, a camera, and an industrial control computer was constructed for experimental testing. Experimental results on the PApple_RGB – D – Size dataset showed that, with maturity classified into eating-ripe, harvest-ripe, and unripe, the improved model achieved an overall mean average precision (mAP@0.5) of 87.7%. The AP@0.5 for eating-ripe, harvest-ripe, and unripe apples were 86.6%, 89.0%, and 87.5%, respectively; the overall mAP@0.5 and the AP@0.5 for the three maturity categories was increased by 3.9, 8.0, 1.6, and 2.1 percentage points compared with that of YOLO 11n. In the orchard apple maturity dataset, the improved model achieved an mAP@0.5 of 95.5%, representing an increase

收稿日期: 2025 – 10 – 16 修回日期: 2026 – 01 – 14

基金项目: 河南省重点研发专项(231111112700)

作者简介: 牛金星(1981—),男,教授,博士,主要从事机器人视觉与智能控制研究,E-mail: 63719302@qq.com

of 0.9 percentage points compared with that of YOLO 11n. In harvesting experiments, the detection rate, harvesting success rate, and damage rate were 99.0%, 84.0%, and 4.8%, respectively. The enhancement of the vision model improved detection accuracy, while the visual-tactile fusion method provided a reference scheme for low-damage harvesting.

Key words: apple; ripeness detection; low-damage harvesting; flexible gripper; YOLO 11n – LDA

0 引言

不同成熟度苹果的销售方式和理化性质不同,若不加以区分不但会增加后期分拣成本,而且在机器人采摘过程中也会对果肉造成伤害。不同成熟度苹果理化性质的变化导致了其屈服极限发生了改变,于是根据不同苹果成熟度设置不同的抓取力来实现低损采摘至关重要^[1-2]。低损采摘的前提是成熟度的正确识别。

近年来,基于机器视觉的水果成熟度识别在农业方面得到了广泛应用^[3-5],文献[6]将密集连接卷积网络(Densely connected convolutional networks, DenseNet)加入到 YOLO v3 中,增强其特征传播能力,与原始 YOLO v3 相比,该模型在苹果成熟度识别中的准确率提高了 2.4 个百分点。文献[7]提出一种基于改进 YOLO v5 的苹果分级模型,通过引入 Mish 激活函数、DIoU Loss 损失函数和通道注意力机制(Squeeze-and-excitation, SE),对已采摘苹果成熟度的分级准确率提高 4.2 个百分点。文献[8]针对苹果的成熟度识别提出了一种基于 YOLO v7 的改进模型,实现未成熟、半成熟和成熟 3 类果实的分类,相较于基线模型精度提高 1.9 个百分点。以上研究虽能实现对苹果成熟度的识别,但仍存在模型体积较大、推理速度较慢等问题,难以满足在真实果园环境下的自动化采摘任务^[9-11]。

在低损采摘领域,文献[12]建立苹果的果皮、果肉和果核的粘弹性模型,使用 ANSYS 仿真抓取过程得出其夹持器的最大抓取速度,为低损抓取提供参考。由于近年来低损、无损抓取需求的快速增长,文献[13]提出了一种仿生鱼鳍手,尽管鱼鳍手的材料和设计都是从柔性的角度出发,但抓取物体总会给目标施加力。文献[14]提出了一种鱼鳍手力感知的方法,通过深度学习反过程映射出形变和力的关系,最后通过相机观察鱼鳍手在抓取过程中的变形来得出感知力,此方法一定程度上解决了柔性夹爪力感知这一难题,但同时其复杂的系统和多过程的转换也限制了其使用场景。

本文提出一种低损采摘方案,首先构建一个基于 YOLO 11n 的改进模型,提高对不同成熟度苹果的识别准确率,随后通过 ANSYS Workbench 仿真出

不同成熟度苹果的最大安全采摘力,并将薄膜压力传感器贴于鱼鳍型柔性夹爪表面实时反馈接触力,最终通过闭环交互实现对不同成熟度苹果的低损采摘,以为低损采摘提供技术支撑。

1 材料与方法

1.1 数据集

苹果图像来自莱里达大学公开的富士苹果数据集 PApple_RGB – D – Size^[15],该数据集共有 3 925 幅图像,图像分辨率为 1 024 像素 × 1 024 像素,格式为 PNG,包含不同成熟度类型和不同光照情况,拍摄距离为 0.3 ~ 0.8 m。但数据集内部比例不均衡^[16],为了保证 3 种成熟度内部比例均衡且图像特征清晰,经挑选后使用 1 464 幅图像。

1.2 数据集构建

首先根据苹果成熟度对数据集进行分类:分为食用成熟、采摘成熟和未成熟 3 种类型。本文借用苹果表面颜色信息来判断成熟度,当表面红色面积占 90% 时认定为食用成熟,红色面积在 70% ~ 90% 之间认定为采摘成熟,红色面积在 70% 以下认定为未成熟。随后按照比例 8:2 划分为训练集和验证集,并保证训练集和验证集中 3 种成熟度的图像数量都为 1:1:1。

划分完成后使用图像标注软件 LabelImg 进行数据标注,得到 YOLO 格式的 txt 文件,标注的标签类别为 A、B、C,分别代表食用成熟、采摘成熟、未成熟。部分图像样本如图 1 所示。

1.3 YOLO 11n – LDA 苹果成熟度识别模型建立

1.3.1 YOLO 11n – LDA 模型

YOLO 11 (You only look once 11) 是 Ultralytics 公司在 2024 年 9 月 30 日发布的神经网络模型^[17],相较于之前版本,YOLO 11 引入了 C3k2 (C3 cross stage partial with kernel size 2) 和 C2PSA (channel-to-pixel spatial attention) 两个全新模块,提升了其特征提取能力,从而能够更精确地实现复杂任务。同样 YOLO 11 也分为 n、s、m、l 和 x 版本,网络深度依次加深,检测精度和计算量依次增大,由于该模型最终要部署到移动设备上在户外实现苹果采摘,综合考虑硬件资源、响应速度和能耗 3 方面因素,本文选择使用 YOLO 11n 作为基线模型。为了提升在不同光



图 1 复杂果园环境下图像样本

Fig. 1 Image samples in complex orchard environments

照下对不同成熟度苹果的特征提取能力,首先在 SPPF 中添加 LSKA 模块构建 SPPF-LSKA 模块,

然后在 C3k2 模块中残差块中增加 DySnakeConv 构建 C3k2-DySnake 模块,提升模型对苹果轮廓特征的理解能力,最后将 Neck 网络替换为特征交互能力更强的 AFPN 网络,构建了 YOLO 11n-LDA 识别模型,整体模型网络结构如图 2 所示。

1.3.2 SPPF-LSKA 模块

在自然果园环境中,受光照变化等因素的影响,苹果表面的颜色和纹理特征易被破坏,导致判别成熟度的关键信息颜色分布和果皮纹理的表征能力下降。如何准确获取苹果图像的特征,对于提高识别准确率至关重要。为此本研究在 SPPF (Spatial pyramid pooling fast) 模块的特征输出部分添加 LSKA 注意力机制,从而构建 SPPF-LSKA 模块。SPPF 模块通过多尺度空间金字塔池化融合不同感受野的特征,初步提升对光照下苹果目标的尺度适应性,但在复杂光照条件下效果并不理想。

LSKA 注意力机制基于大核注意力 (Large kernel attention, LKA) 的设计思想,如图 2 右侧所示。通过将传统的二维卷积核分解为两个级联的一

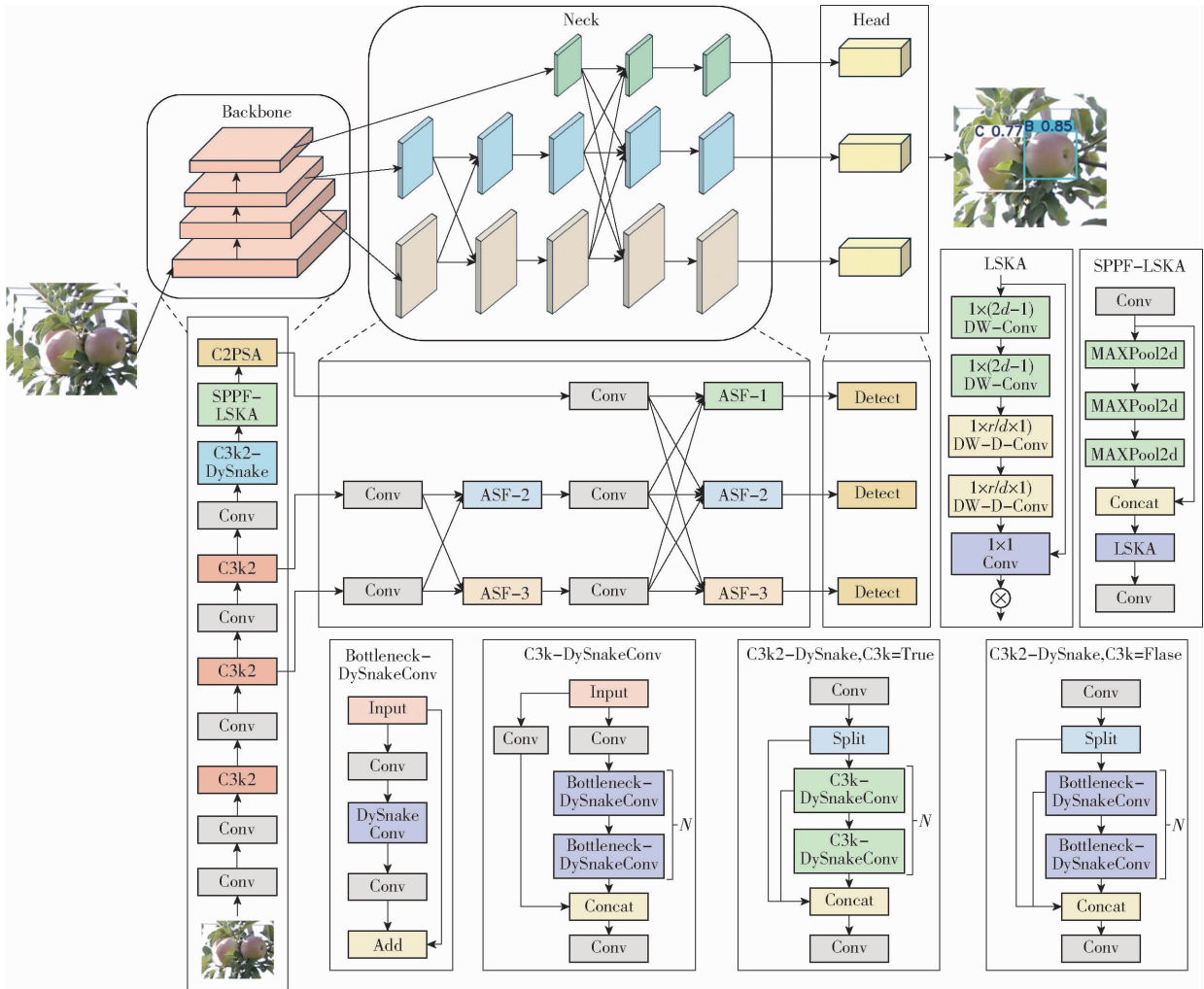


图 2 YOLO 11n-LDA 网络结构图

Fig. 2 Architecture of YOLO 11n-LDA

维可分离卷积核,显著降低了计算复杂度和内存占用,其中 DW-Conv 为深度卷积,DW-D-Conv 为扩张深度卷积, $\otimes$ 为逐元素乘积, r 和 d 表示不同的卷积核大小^[18]。

SPPF-LSKA 为最终的改进模块,特征图输入进 SPPF-LSKA 模块后,首先通过池化层对不同尺度特征进行压缩与降维,以减少噪声和冗余信息的干扰,同时增大模型感受野。随后借助连接层将局部区域内的特征值进行融合,并输入至 LSKA 模块。在 LSKA 中采用上述核分解的处理方式,并结合串联卷积策略,从而有效保留实例中的关键特征,提取出更为鲁棒且具有代表性的图像特征,提高模型对不同成熟度苹果图像的理解能力。

1.3.3 C3k2-DySnake 模块

苹果表面红绿相接的区域为苹果成熟度的过渡区域,该过渡区域本质上是一条蜿蜒的曲线,此外苹果轮廓与背景的交界处本质上也是蜿蜒的曲线,成熟度检测的关键点也在于要准确提取这些曲线的特征。但曲线细长的特征会导致它们易受光线的影响,无论是强光的高亮还是暗光的阴影都会导致这些曲线难以辨认。为此本研究引入了动态蛇形卷积(DySnakeConv)。动态蛇形卷积能动态调整卷积核形状以适应目标结构,从而实现自适应地关注细长

和蜿蜒的局部特征,同时它通过 x 轴和 y 轴两个方向的卷积从不同视角观察结构特征,通过多视角特征融合,使模型能够更全面地理解苹果表面纹理和颜色的变化^[19]。

如图 3 所示,在动态蛇形卷积中,标准卷积沿 x 轴和 y 轴方向扩展,以中心网格为起点,按前一网格位置叠加 $[-1,1]$ 的偏移量确定后续网格,保证符合线性形态。然后对分数偏移位置用双线性插值处理,最终形成 9×9 感受野,能自适应聚焦细长弯曲局部特征,提升检测准确性。表达式为

$$K_{i \pm c} = \begin{cases} (x_{i+c}, y_{i+c}) = \left(x_i + c, y_i + \sum_i^{i+c} \Delta y\right) \\ (x_{i-c}, y_{i-c}) = \left(x_i - c, y_i + \sum_{i-c}^i \Delta y\right) \end{cases} \quad (1)$$

$$K_{j \pm c} = \begin{cases} (x_{j+c}, y_{j+c}) = \left(x_j + \sum_j^{j+c} \Delta x, y_j + c\right) \\ (x_{j-c}, y_{j-c}) = \left(x_j + \sum_{j-c}^j \Delta x, y_j - c\right) \end{cases} \quad (2)$$

式中 $K_{i \pm c}$ 、 $K_{j \pm c}$ ——卷积核 K 在网格的具体位置
 (x, y) ——特征图上的坐标
 (i, j) ——中心点坐标索引
 c ——距离中心点的偏移距离,取 0、1、2、3、4
 Δx 、 Δy ——学习偏移量

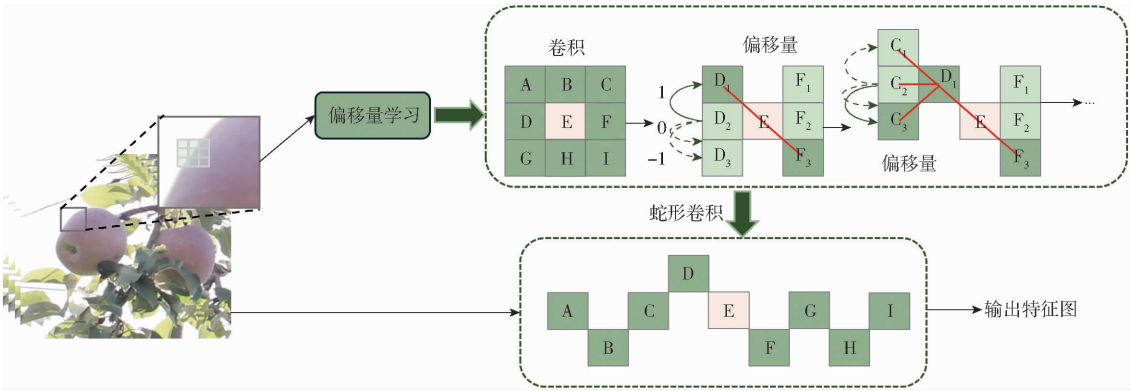


图 3 蛇形卷积结构
Fig.3 Serpentine convolution structure

C3k2 中残差块的作用是通过跳跃连接避免梯度消失,使网络能学习输入与输出的残差映射,有效提取高级特征。本研究在残差块中增加 DySnakeConv,它能够通过水平、垂直和标准卷积的组合,自适应关注苹果表面的细长纹理等关键特征,提升对成熟度的感知能力。其结构如图 2 下方所示。其工作流程为:首先通过卷积层对输入特征进行初步提取,然后进行分支处理,一组利用动态蛇形卷积动态调整卷积核,增强模型的灵活性和适应性,处理完成后与不做处理的另一组相加,实现深度特征的融合。这种结构有效提升了模型对曲线特征的

提取能力。
1.3.4 AFPN 模块
YOLO 11 的 Neck 网络采用 PAFPN 的架构思路,在特征融合过程中其非相邻层级特征需经多中间层传递才能融合,导致苹果成熟度相关的关键特征易丢失;并且 PAFPN 采用固定的特征融合权重,在特征融合过程中无法动态调整特征权重,这就导致在特征融合过程中如果模型对特征的权重造成误判,就很可能造成最后识别结果的错误。
针对以上两点问题,本文提出使用渐进式特

征融合金字塔 (AFPN) 来解决此类问题,AFPN 有以下两点优势:一是采用渐进式融合,将低层级特征逐步纳入高层级特征,缩小非相邻层语义差距,完整保留不同光照下苹果的细节特征与成熟度语义信息,避免传递损耗;在本文中,提取了 3 种不同层次的特征,分别为小层、中层和大层,表示为 C3、C4 和 C5。在第 1 阶段首先融合 C3 和 C4,然后最后加入 C5,这有效规避了 C3 和 C5 的不良融合,并减少了它们之间的语义差距^[20],其结构如图 2 上方的 Neck 部分所示。二是通过自适应空间融合动态加权多层级特征,可抑制光照引发的特征矛盾,重点突出成熟度相关有效特征,从而更精准捕捉不同光照下苹果成熟度差异,提升检测准确度,ASFF 用于为不同层级的特征分配不同的权重,其结构如图 4 所示。在多层级引入不同的空间权重时,特征融合增强了关键层次的重要性,同时抑制了来自不同对象冲突信息的影响,本文中融合了 3 个层次的特征,表达式为

$$y_{ij}^l = \alpha_{ij}^l x_{ij}^{1 \rightarrow l} + \beta_{ij}^l x_{ij}^{2 \rightarrow l} + \gamma_{ij}^l x_{ij}^{3 \rightarrow l} \tag{3}$$

$$\alpha_{ij}^l + \beta_{ij}^l + \gamma_{ij}^l = 1 \tag{4}$$

式中 $\alpha_{ij}^l, \beta_{ij}^l, \gamma_{ij}^l$ ——不同层次的特征对于第 l 层的空间权重
 y_{ij}^l ——多层特征自适应空间融合的输出
 $x_{ij}^{1 \rightarrow l}, x_{ij}^{2 \rightarrow l}, x_{ij}^{3 \rightarrow l}$ ——来自第 1、2、3 层经调整后用于匹配第 l 层的特征向量

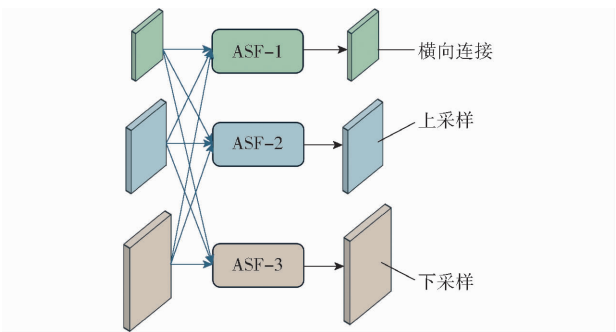


图 4 自适应空间融合操作
Fig. 4 Adaptive spatial fusion operation

综上,AFPN 可以逐步融合不同尺度的特征,以

便更好地理解不同光照环境下苹果表面的成熟度特征。同时,自适应空间融合机制可以根据不同情况调整注意力,更好地处理复杂光照下的苹果表面特征与背景特征之间的权重。

1.4 低损采摘方法

上文通过改进视觉模型对苹果成熟度实现有效分类,但要实现低损采摘则需将视觉信息转化为对应成熟度的苹果力学特性。不同成熟度苹果的果肉在弹性模量、泊松比等本构参数上存在显著差异,直接决定了其最大安全采摘力的大小。所以本文通过有限元仿真测出不同成熟度苹果的力学特性,并将其与视觉结合。借助视觉识别出苹果成熟度后匹配有限元仿真得出的最大安全采摘力,此力值作为柔性夹爪的力控设定点,确保抓取过程在最大安全采摘力内进行。

1.4.1 鱼鳍手柔性夹爪

在抓取苹果的过程中,夹爪与苹果之间不适当的接触力会使苹果产生淤伤。由于苹果自身为椭圆形,而刚性夹爪形状固定难以适应苹果的形状,所以在抓取过程中往往接触面积较小,导致接触应力较大,从而导致苹果淤伤^[21]。柔性夹爪材质柔软,具有一定的自适应性,在抓取物体时能对被抓物形成包络,这样其能增大与苹果的接触面积,降低接触应力^[22-23]。所以本文选取鱼鳍手柔性夹爪进行苹果采摘。

1.4.2 有限元仿真最大安全采摘力

测量苹果最大安全采摘力的常见方法有两种:一是随机采摘不同成熟度的苹果然后使用夹爪进行抓取,抓取后保持静止使用光谱测量;二是借助有限元仿真软件进行仿真测试。鉴于第 1 种方案需要消耗大量人力物力,且目前开源的苹果数据较多,所以选用有限元仿真进行测试。首先根据文献[24]所提供的富士苹果的特征数据,使用 Solidworks 2020 构建了果核-果肉-果皮的复合苹果模型。然后使用 ANSYS Workbench 2021 进行仿真试验。苹果数据如表 1 所示。

表 1 各成熟度苹果理化性质

Tab. 1 Physicochemical properties of apples at different maturity stages

成熟度	果皮			果肉			果核		
	弹性模量/MPa	泊松比	密度/(kg·m ⁻³)	弹性模量/MPa	泊松比	密度/(kg·m ⁻³)	弹性模量/MPa	泊松比	密度/(kg·m ⁻³)
未成熟	10.78	0.35	995	5.04	0.25	995	20	0.35	995
采摘成熟	7.18	0.35	1 037	5.12	0.21	1 037	20	0.35	1 037
食用成熟	5.23	0.35	1 047	3.85	0.35	1 047	20	0.35	1 047

首先将模型导入 ANSYS Workbench 中,材料方面因为只需要检测出苹果的最大安全采摘力,并不

需要模拟损伤随时间的演变,所以定义为弹性材料。苹果核和果肉以及果肉和果皮均采用绑定连接。单

元选择为 Solid186, 网格类型为四面体, 网格节点数为 349 840, 单元数为 194 677。设置苹果上下果萼为固定面, 施加力位置为柔性夹爪在抓取苹果时其梁单元对苹果的施加力位置^[25], 结合实际抓取情况共设置 5 处受力位置, 受力面积为薄膜压力传感器检测面积, 为 78.5 mm², 其仿真结果如图 5 所示。虽然果皮位于苹果最外侧且较薄, 但得益于其弹性模量较大, 在采摘过程中仍是果肉最易受到损伤^[26], 所以仿真分析中只展示了果肉部分。当果肉

所受应力大于其屈服极限时代表损失不可逆, 即造成淤伤, 所参考不同成熟度果肉的屈服极限为文献[24]的公开数据, 未成熟、采摘成熟、食用成熟果肉的屈服极限为 0.291、0.271、0.211 MPa。综合仿真结果与上述数据得出食用成熟、采摘成熟、未成熟的**最大安全采摘力**分别为 5.3、6.1、6.8 N, 在实际应用时为了减小不可控因素的影响, 本文取 10% 的安全裕度, 所以本文中实际采用的最大安全采摘力为 4.7、5.4、6.1 N。

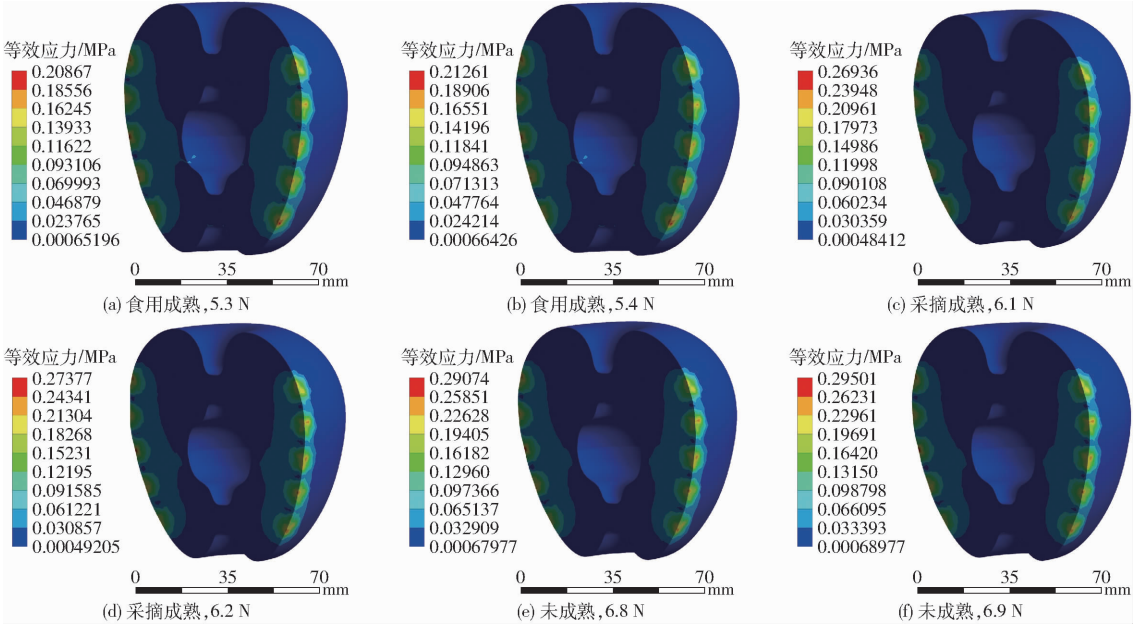


图 5 不同成熟度苹果仿真结果

Fig. 5 Simulation results of apples at different maturity stages

1.4.3 薄膜压力传感器标定

本文选用薄膜压力传感器来反馈夹爪对苹果的抓取力, 传感器型号为 RFP602, 由深圳市洛迦科技有限公司生产。量程为 0 ~ 10 N, 误差 0.05 N。该传感器具有较高的柔性和精度, 依靠压敏电阻来反馈力值, 压力越大, 电阻越小, 反馈力越大, 在使用前选择使用砝码对其进行标定, 然后使用 Origin 软件对标定结果进行拟合。

在 Origin 中使用指数增长模型对其进行拟合, 基本公式为

$$P = A_1 e^{\frac{d_i}{t_1}} + y_0 \tag{5}$$

- 式中 P ——压力, N
- d_i ——模数转换后的数字量
- A_1 ——响应幅值系数
- t_1 ——衰减常数
- y_0 ——基准偏移量

拟合最终结果为 $P = 0.09e^{\frac{x}{11191.65}} - 0.15$, 其 R^2 为 0.964, 表示其拟合性能较好。

2 结果与分析

2.1 视觉模型性能分析

2.1.1 实验平台与评价指标

视觉模型的训练和测试所使用计算机硬件为英特尔至强® 银 4210R 处理器和 NVIDIA GTX3090 (24 GB) 显卡。网络采用 PyTorch 1.12.1 框架构建, 通过 Anaconda3 创建深度学习环境, 配置包含 Python 3.10 和 CUDA 11.3。训练参数为: 输入图像尺寸为 640 × 640, 训练轮数设定为 300, 批量大小设定为 32。本文以准确率、召回率、平均精度均值 (Mean average precision, mAP@0.5)、模型内存占用量和总时间等指标作为模型性能评价标准。其中 mAP@0.5 是指 IoU 阈值设为 0.5 时的平均精度均值。

2.1.2 消融实验

为了说明本文所采用的 LSKA、DySnakeConv、AFPN 模块的有效性, 以 YOLO 11n 为基线模型, 用上述所规定的评价标准进行了 4 组消融实验, 结果如表 2 所示。为评估改进模型对不同成熟度苹果的

识别性能,其 AP@0.5 指标如表 3 所示。

表 2 消融实验结果
Tab.2 Ablation experiment results

LSKA	DySnake Conv	AFPN	准确 率/%	召回 率/%	mAP@ 0.5/ %	模型内 存占用 量/MB	总时 间/ ms
×	×	×	75.2	75.2	83.8	5.5	2.0
✓	×	×	82.3	77.4	85.1	6.3	3.3
✓	✓	×	83.3	75.1	87.4	6.4	5.2
✓	✓	✓	81.8	81.3	87.7	6.6	3.0

注：“✓”表示使用该模块，“×”表示不使用该模块，下同。

表 3 不同成熟度 AP@0.5 实验结果

Tab.3 Experimental results for different maturity
levels of AP@0.5 %

LSKA	DySnake Conv	AFPN	食用成熟	采摘成熟	未成熟
×	×	×	78.6	87.4	85.4
✓	×	×	82.8	88.7	87.1
✓	✓	×	85.1	89.3	87.6
✓	✓	✓	86.6	89.0	87.5

由表 2 可知,对比 YOLO 11n,在 SPPF 中加入 LSKA 模块后模型的 mAP@0.5 和召回率分别提升 1.3、2.2 个百分点,从表 3 来看食用成熟识别的 AP@0.5 提升 4.2 个百分点。为了进一步提升对食用成熟的检测性能,引入了 DySnakeConv 模块,相较于基线模型整体 mAP@0.5 和食用成熟的 AP@0.5 分别提升 3.6、6.5 个百分点,但召回率有所下降。本模型是在 P5 的高特征层的 C3k2 中添加 DySnakeConv,可以使模型提取到更为丰富的特征,减小光照的影响,但同时也会导致此特征层所提取的特征与 P4 特征层和 P3 特征层不同,会存在语义差距,且由于 YOLO 11 的 Neck 网络是 PAFPN 架构,在特征融合时还存在特征衰减,会加剧不同特征层之间的语义差距,所以召回率会降低,但 DySnakeConv 使模型更准确地提取特征使检测准确率提高,且准确率的提升幅度大于召回率的下降幅度,所以最终 mAP@0.5 也会提升。为适配果园机械臂采摘场景对近距离大目标的检测需求,本模型仅在 P5 高层特征图中引入 DySnakeConv 模块。因为模型的 P3 与 P4 层主要负责中小目标特征提取,与当前任务目标关联较弱;若在全尺度特征层中添加该模块,则会引入冗余计算,导致模型参数量与推理时间显著增加。

为了改善语义差距造成的召回率下降,本文使用 AFPN 替换了 Neck 网络对特征进行渐进式融合,最终召回率提升 6.1 个百分点,整体 mAP@0.5 提升 3.9 个百分点,食用成熟、采摘成熟和未成熟的

AP@0.5 分别提升 8.0、1.6、2.1 个百分点,每幅图像处理总时间仅需 3.0 ms,模型检测对比如图 6 所示。实验结果验证了改进措施的有效性,为后续研究提供参考依据。

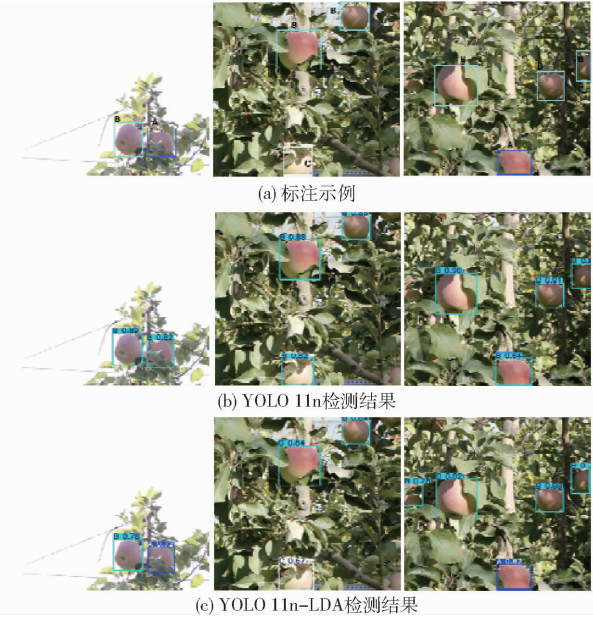


图 6 模型检测对比图

Fig. 6 Model detection comparison charts

2.1.3 与其他 YOLO 模型对比

为了验证模型性能,选择与主流的轻量级网络 YOLO v8n、YOLO 11n、YOLO 12n、YOLO 13n 以及高性能模型 YOLO v8m、YOLO 11m 在相同环境下进行实验,结果如表 4 所示。

表 4 各 YOLO 模型检测结果

Tab.4 Detection results for each YOLO model

模型	准确 率/%	召回率/ %	mAP@0.5/ %	模型内存占 用量/MB	总时 间/ms
YOLO v8n	78.9	77.1	84.3	6.3	6.7
YOLO v8m	80.1	75.4	86.2	52.0	5.1
YOLO 11n	77.5	75.2	83.8	5.5	2.0
YOLO 11m	77.3	76.7	85.0	40.5	4.5
YOLO 12n	71.3	74.0	81.7	5.5	3.2
YOLO 13n	79.8	70.7	83.7	5.4	4.1

本文所提出的 YOLO 11n-LDA 在 mAP@0.5 上相较于轻量级网络 YOLO v8n、YOLO 11n、YOLO 12n 和 YOLO 13n 分别提升 3.4、3.9、6.0、4.0 个百分点,模型内存占用量相较于 YOLO 11n 仅增加 1.1 MB。并且相较于主流的高性能检测网络 YOLO v8m 和 YOLO 11m 在 mAP@0.5 上也有 1.5、2.7 个百分点的提升,同时在模型内存占用量方面分别降低了 45.4 MB 和 33.9 MB。实验表明 YOLO 11n-LDA 在检测精度、检测速度和模型内存占用量方面均有优势。

2.1.4 模型普适性验证

前述实验验证了模型在背景干扰较大的环境下的鲁棒性。然而实际作业环境也涵盖背景干扰较小的情况。为了全面评估模型在不同环境下的普适性,本文又引入了由文献[27]公开的 Orchardapple maturity 数据集进行补充实验,该数据集包含 1 989 幅图像,分为“过熟、成熟、早熟、未成熟”4 个类别,按照比例 8:2 划分为训练集和验证集,部分图像样本如图 7 所示,检测结果如表 5 所示。

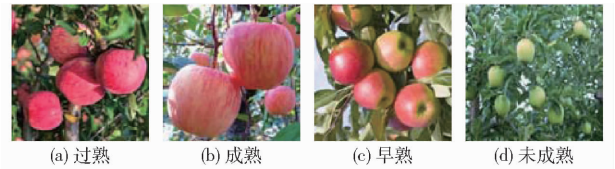


图 7 简单环境下图像样本

Fig. 7 Image samples in simple environment

表 5 数据集模型检测结果

Tab. 5 Dataset model detection results

模型	准确 率/%	召回 率/%	mAP@0.5/ %	模型内存 占用量/MB	总时 间/ms
YOLO 11n	89.8	89.8	94.6	5.5	27.0
YOLO 11n-LDA	90.7	91.1	95.5	6.7	32.8

由表 5 可知在背景干扰较小的数据集上,尽管 YOLO 11n 已经达到了较高的精度,但改进模型在准确率、召回率和 mAP@0.5 上仍分别提升了 0.9、1.3、0.9 个百分点。

2.2 低损采摘实验分析

2.2.1 整体实验平台搭建

实验平台主要包含 Intel D455 型相机、法奥意威 FR3 型协作机械臂、易爪机器人鱼鳍手型柔性夹爪、工控机以及洛迦科技 RFP602 型薄膜压力传感器。同时从郑州市中牟县采集了 100 枚带有果枝的苹果,其中采摘成熟和食用成熟各 50 枚,如图 8 所示。利用铁丝将采集的苹果固定于与苹果树相似的树上,模拟真实的采摘环境。其整体实验环境如图 9 所示,从采集到测试完成控制在 24 h 内,确保数据真实性。



图 8 采摘苹果样例

Fig. 8 Apple picking example



图 9 实验平台

Fig. 9 Experimental platform

1. 工控机 2. 易爪机器人鱼鳍手型柔性夹爪 3. Intel D455 型相机 4. 法奥意威 FR3 型协作机械臂

其中苹果平均直径和高度分别为 75.3 mm 和 63 mm;深度相机 RGB 分辨率为 1 280 像素×800 像素,帧率为 30 f/s,有效测量范围为 0.6~6 m;协作机械臂最大工作半径为 0.62 m,最大负载质量为 3 kg;柔性夹爪由伺服电机驱动,其最大张开角为 135°;薄膜压力传感器厚度为 0.4 mm,响应时间 10 ms 以内;工控机操作系统为 Ubuntu 20.04,搭配 RTX3060 显卡和 Intel i5-8500 处理器。

本研究采用眼在手外配置进行手眼标定。将标定板固定于机械臂末端,控制其多位姿移动,同步采集 RGB-D 图像与末端位姿数据。通过提取棋盘格特征并结合机械臂运动学参数,利用最小二乘法求解相机与基坐标系的变换矩阵。该矩阵将视觉三维坐标精确映射至机械臂基坐标系,为后续运动规划与抓取任务提供准确的定位基准,确保作业精度。

实验平台整体由工控机控制,旨在实现苹果成熟度判别、安全力匹配、坐标解算及机械臂与夹爪的闭环控制。整体通讯结构如图 10 所示。

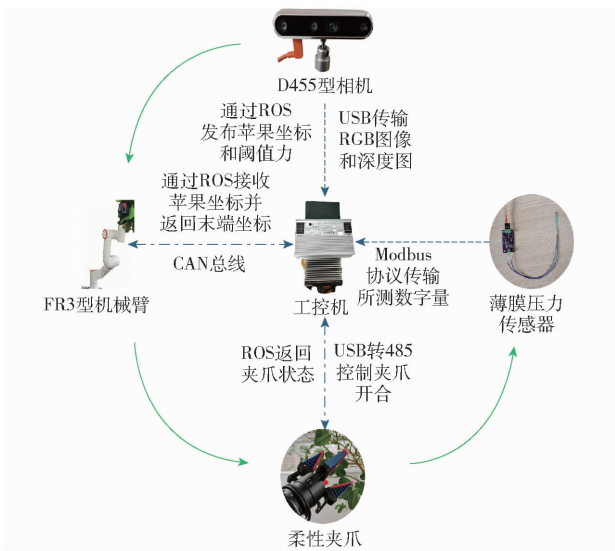


图 10 实验平台各部件通讯关系

Fig. 10 Communication relationships among components of experimental platform

具体作业流程为: D455 型相机通过 USB 接口传输 RGB-D 图像, 视觉算法识别苹果成熟度并匹配最大安全采摘力, 同时结合手眼标定矩阵将目标从相机坐标系转换至世界坐标系。工控机通过 ROS 话题发布坐标, 并经由 CAN 总线驱动机械臂抵达目标位置。随后, 通过 RS485 指令控制夹爪闭合, 期间薄膜压力传感器利用 Modbus 协议实时回传模数转换后的数字量。系统依据 1.4.3 节所标定的公式实时计算抓取力, 一旦达到设定的最大安全采摘力, 夹爪立刻停止闭合。最后机械臂通过订阅夹爪的状态信号, 当夹爪完成闭合后, 机械臂执行采摘动作将苹果运送至存放点释放, 完成自动化作业。

2.2.2 低损采摘实验

首先将苹果分为 10 组, 每组包含食用成熟和采摘成熟各 5 枚, 挂在树上模拟真实的果园环境, 机械臂控制路径规划部分采用 RRT-Connect (Rapidly-exploring random tree connect) 算法。采摘流程如图 11 所示。

采摘时, 首先通过 ROS 中的 Rviz 连接机械臂, 初始化整个采摘系统 (图 11a), 然后相机捕获 RGB 图像和深度图并传输给视觉模型, 模型根据 RGB 图像和深度图可识别出苹果成熟度、对应的最大安全采摘力以及坐标, 并进行坐标变换 (图 11b), 机械臂获取坐标后从初始位置 (图 11c) 移动至采摘位置 (图 11d), 随后夹爪开始闭合 (图 11e) 并实时检测夹爪对苹果的抓取力 (图 11f), 当抓取力达到所设定的最大安全采摘力后夹爪停止闭合, 并进行采摘 (图 11g), 最后将苹果运输至存放点完成采摘 (图 11h)。

为了评估所提出视觉模型 YOLO 11n-LDA 以及“视-触”融合方案的有效性, 将采摘后的苹果置于相同的室温环境中, 在 3 d 和 7 d 时统计其褐变情况, 并对实验结果进行了整理, 如表 6 所示。

采摘实验中, 检测率被定义为模型成功定位苹果目标的能力, 该指标与成熟度的正确分类不同, 漏检的原因为严重遮挡, 误检的主要原因为苹果表面有大面积光照, 若是少面积光照情况, 如图 11b 所示, 所改进模型可以正确识别出成熟度。采摘失败的主要原因为苹果受树枝遮挡夹爪无法形成有效包裹、部分位置 RRT-Connect 算法无法规划以及机械臂路径规划中未加入对苹果位姿的考虑, 会出现夹爪从果梗的树枝上方抓取苹果无法将其摘下等情况。损伤方面, 食用成熟的苹果被误检为采摘成熟时, 会匹配错误的最大安全采摘力, 导致其采摘力大于最大安全采摘力。采摘实验总计检测率 99.0%、成熟度正确识别率 93.0%、成功采摘率 84.0%、

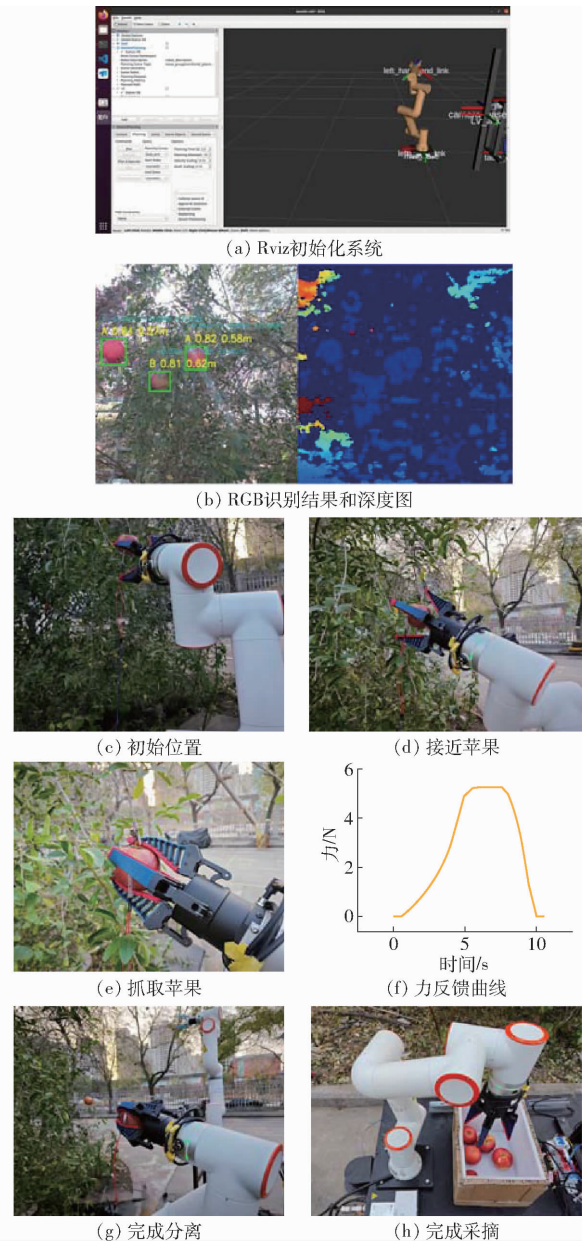


图 11 采摘流程图

Fig. 11 Harvesting flowcharts

表 6 实验结果统计

Tab. 6 Statistics of experimental results

组别	检出目标数	成熟度正确分类个数	成功采摘个数	损伤个数	单果采摘时间/s
1	10	9	8	0	15.8
2	10	10	9	1	16.2
3	10	10	8	0	14.7
4	9	8	8	0	14.8
5	10	9	8	1	15.2
6	10	10	9	1	14.5
7	10	9	9	0	15.3
8	10	9	8	0	14.8
9	10	10	9	1	14.9
10	10	9	8	0	15.2
总计	99	93	84	4	

损伤率 4.8% 以及平均单果采摘时间为 15.1 s。

3 结 论

(1)提出了改进的苹果成熟度检测模型 YOLO 11n-LDA,在背景干扰较大的数据集中其 mAP@0.5 为 87.7%,相 较 于 YOLO v8n、YOLO 11n、YOLO 12n、YOLO 13n、YOLO v8m 和 YOLO 11m 分别提高 3.4、3.9、6.0、4.0、1.5、2.7 个百分点。食用成熟、采摘成熟和未成熟的 AP@0.5 相较于 YOLO 11n 分别提高 8.0、1.6、2.1 个百分点。模

型内存占用量为 6.6 MB,每幅图像处理时间为 3.0 ms。

(2)在背景干扰较小的数据集中 YOLO 11n-LDA 的 mAP@0.5 为 95.5%,相较于 YOLO 11n 提高 0.9 个百分点,验证了该模型的普适性。

(3)通过采摘实验得出该方案可以较好地实现低损采摘的目的,其检测率、成熟度正确识别率、成功采摘率和损伤率分别为 99.0%、93.0%、84.0% 和 4.8%,可为后续低损采摘提供参考。

参 考 文 献

[1] 史帅旗,杨福增,刘志杰,等.复合气动式苹果采摘手设计与试验[J].农业机械学报,2024,55(10):93-105.
SHI Shuaiqi, YANG Fuzeng, LIU Zhijie, et al. Design and experiment of composite pneumatic apple picking manipulator[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(10): 93-105. (in Chinese)

[2] SUN J, SUN L, ZHAO G, et al. Triboelectric force feedback-based fully actuated adaptive apple-picking gripper for optimized stability and non-destructive harvesting[J]. Computers and Electronics in Agriculture, 2025, 237: 110725.

[3] GU Z, MA X, GUAN H, et al. Tomato fruit detection and phenotype calculation method based on the improved RTDETR model[J]. Computers and Electronics in Agriculture, 2024, 227: 109524.

[4] YUAN J, FAN J, SUN Z, et al. Deployment of CES-YOLO: an optimized YOLO-based model for blueberry ripeness detection on edge devices[J]. Agronomy, 2025, 15(8): 1948.

[5] QIU Z, OU W, MO D, et al. BGWL-YOLO: a lightweight and efficient object detection model for apple maturity classification based on the YOLOv11n improvement[J]. Horticulturae, 2025, 11(9): 1068.

[6] TIAN Y, YANG G, WANG Z, et al. Apple detection during different growth stages in orchards using the improved YOLO-V3 model[J]. Computers and Electronics in Agriculture, 2019, 157: 417-426.

[7] 崔祥.基于视觉检测的苹果分级技术研究[D]. 镇江:江苏大学, 2023.
CUI Xiang. Research and implementation of apple grading technology based on visual inspection[D]. Zhenjiang: Jiangsu University, 2023. (in Chinese)

[8] 苗荣慧,李港澳,黄宗宝,等.基于 YOLO v7-ST-ASFF 的复杂果园环境下苹果成熟度检测方法[J].农业机械学报, 2024,55(6):219-228.
MIAO Ronghui, LI Gang'ao, HUANG Zongbao, et al. Maturity detection of apple in complex orchard environment based on YOLO v7-ST-ASFF[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(6): 219-228. (in Chinese)

[9] TANG Z, XU L, LI H, et al. YOLOC-tiny: a generalized lightweight real-time detection model for multiripeness fruits of large non-green-ripe citrus in unstructured environments[J]. Frontiers in Plant Science, 2024, 15: 1415006.

[10] YU H, QIAN C, CHEN Z, et al. Ripe-detection: a lightweight method for strawberry ripeness detection[J]. Agronomy, 2025, 15(7): 1645.

[11] YANG X, ZHAO W, WANG Y, et al. Lightweight and efficient deep learning models for fruit detection in orchards[J]. Scientific Reports, 2024, 14(1): 26086.

[12] JI W, QIAN Z, XU B, et al. Apple viscoelastic complex model for bruise damage analysis in constant velocity grasping by gripper[J]. Computers and Electronics in Agriculture, 2019, 162: 907-920.

[13] WILSON M. Festo drives automation forwards[J]. Assembly Automation, 2011, 31(1): 12-16.

[14] XU W, ZHANG H, YUAN H, et al. A compliant adaptive gripper and its intrinsic force sensing method[J]. IEEE Transactions on Robotics, 2021, 37(5): 1584-1603.

[15] FERRER M, RUIZ-HIDALGO J, GREGORIO E, et al. Simultaneous fruit detection and size estimation using multitask deep neural networks[J]. Biosystems Engineering, 2023, 233: 63-75.

[16] 李会宾.智能苹果采摘机器人的视觉精准感知算法研究[D]. 北京:中国农业科学院, 2023.
LI Huibin. Research on visual precision perception algorithm for intelligent apple picking robot[D]. Beijing: Chinese Academy of Agricultural Sciences, 2023. (in Chinese)

[17] JOCHER G, QIU J. Ultralytics YOLO11[EB/OL]. (2024-09-30)[2025-10-10]. <http://github.com/ultralytics/ultralytics>.

[18] LAU K W, PO L M, REHMAN Y A U. Large separable kernel attention: rethinking the large kernel attention design in CNN[J]. Expert Systems with Applications, 2024, 236: 121352.

[19] QI Y, HE Y, QI X, et al. Dynamic snake convolution based on topological geometric constraints for tubular structure

segmentation[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023: 6070 – 6079.

[20] YANG G, LEI J, ZHU Z, et al. AFPN: asymptotic feature pyramid network for object detection[C]//2023 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2023: 2184 – 2189.

[21] JI W, HE G, XU B, et al. A new picking pattern of a flexible three-fingered end-effector for apple harvesting robot[J]. Agriculture, 2024, 14(1): 102.

[22] AN B, CHOI T, KIM U. Linkage integrated fin ray gripper capable of safe adaptive grasping for tomato harvesting[J]. Computers and Electronics in Agriculture, 2025, 232: 110118.

[23] LEE L Y, TERRILE S, NAM S, et al. Fin – A – Rays: expanding soft gripper compliance via discrete arrays of flexible structures[J]. Soft Robotics, 2026, 1(13): 121 – 135.

[24] SLUPSKA M, KUPRIANIUK S, STOPA R, et al. Finite element modeling of apple tissue mechanics: a comparative study of elastic and elastoplastic behavior across ripening stages[J]. Journal of Food Engineering, 2026, 402: 112696.

[25] ZHANG Z, ZHOU J, YI B, et al. A flexible swallowing gripper for harvesting apples and its grasping force sensing model[J]. Computers and Electronics in Agriculture, 2023, 204: 107489.

[26] 陈思宇. 苹果力学损伤特性研究及采收试验平台设计[D]. 沈阳: 沈阳农业大学, 2022.

CHEN Siyu. Study on mechanical damage characteristics of apple and design of harvesting test platform [D]. Shenyang: Shenyang Agricultural University, 2022. (in Chinese)

[27] TOM T. Orchard apple maturity dataset [EB/OL]. (2025 – 07 – 10) [2026 – 01 – 03]. <https://doi.org/10.6084/m9.figshare.29533841.v1>.

(上接第 281 页)

[21] 雷霆, 陈刚, 何颖波, 等. 冲击载荷作用下 PZT-5 压电陶瓷的力电特性[J]. 高压物理学报, 2019, 33(5): 129 – 135.

LEI Ting, CHEN Gang, HE Yingbo, et al. Dynamic behavior of PZT – 5 piezoelectric ceramics under impact loading[J]. Chinese Journal of High Pressure Physics, 2019, 33(5): 129 – 135. (in Chinese)

[22] 王礼立. 应力波基础[M]. 北京: 国防工业出版社, 2005.

[23] 刘延柱, 陈立群, 陈文良. 振动力学[M]. 北京: 高等教育出版社, 2019.

[24] 孔祥安, 江晓禹, 金学松. 固体接触力学[M]. 北京: 中国铁道出版社, 1999.

[25] 万丽荣, 陈博, 杨扬, 等. 单颗粒煤岩冲击放顶煤液压支架尾梁动态响应分析[J]. 煤炭学报, 2019, 44(9): 2905 – 2913.

WAN Lirong, CHEN Bo, YANG Yang, et al. Dynamic response of single coal-rock impacting tail beam of top coal caving hydraulic support[J]. Journal of China Coal Society, 2019, 44(9): 2905 – 2913. (in Chinese)

[26] 程光明, 庞建志, 唐可洪, 等. 压电陶瓷发电能力测试系统的研制[J]. 吉林大学学报(工学版), 2007, 37(2): 367 – 371.

CHENG Guangming, PANG Jianzhi, TANG Kehong, et al. Development of measuring system for electricity generating capacity of piezoelectric ceramics[J]. Journal of Jilin University (Engineering and Technology Edition), 2007, 37(2): 367 – 371. (in Chinese)

[27] XU L, HANSEN A C, LI Y, et al. Numerical and experimental analysis of airflow in a multi-duct cleaning system for a rice combine harvester[J]. Transactions of the ASABE, 2016, 59(5): 1101 – 1110.

[28] CHAI X, ZHOU Y, XU L, et al. Effect of guide strips on the distribution of threshed outputs and cleaning losses for a tangential-longitudinal flow rice combine harvester[J]. Biosystems Engineering, 2020, 198: 223 – 234.