

金寿祥,周宏平,姜洪喆,等. 采摘机器人视觉系统研究进展[J].江苏农业学报,2023,39(2):582-595.

doi:10.3969/j.issn.1000-4440.2023.02.033

## 采摘机器人视觉系统研究进展

金寿祥, 周宏平, 姜洪喆, 孙梦梦

(南京林业大学机械电子工程学院,江苏 南京 210037)

**摘要:** 视觉是采摘机器人感知外部环境的重要手段之一。采摘机器人的采摘效率与准确率很大程度上受到其视觉子系统识别与定位性能影响。近年来,研究者围绕采摘机器人的视觉系统开展了大量研究。伴随人工智能深度学习的发展,机器人视觉研究取得了较大进步,采摘机器人进入了实际应用阶段。本文针对采摘机器人视觉系统研究现状,主要从视觉数据获取方法、采摘点识别技术和采摘点定位技术3个方面进行总结分析,指出了当前采摘机器人采摘点识别与定位过程中面临的一些挑战,最后对未来采摘机器人视觉系统的发展进行了展望。

**关键词:** 视觉数据; 采摘点识别; 三维定位; 深度学习

**中图分类号:** TP242.6<sup>+</sup>2      **文献标识码:** A      **文章编号:** 1000-4440(2023)02-0582-14

## Research progress on visual system of picking robot

JIN Shou-xiang, ZHOU Hong-ping, JIANG Hong-zhe, SUN Meng-meng

(School of Mechanical and Electronic Engineering, Nanjing Forestry University, Nanjing 210037, China)

**Abstract:** Vision is one of the important means for picking robots to perceive the external environment. The picking efficiency and accuracy of the picking robots are largely affected by the recognition and positioning performance of its vision subsystem. In recent years, a lot of researches have been carried out regarding the visual system of picking robots. With the development of artificial intelligence deep learning, research of robot vision has made great progress, which opened the stage of practical application of picking robots. In view of the research status of the picking robots visual system, this paper summarized and analyzed from three aspects, such as visual data acquisition method, picking point recognition technology and picking point positioning technology, and pointed out some challenges faced by the current picking robots in the process of picking point identification and positioning. Finally, the development of visual system for picking robots in the future was prospected.

**Key words:** visual data; picking point identification; three dimensions positioning; deep learning

果实采摘需要大量劳动力参与<sup>[1]</sup>,而近年来农业劳动力出现了老龄化与短缺的问题<sup>[2]</sup>。户外露天环境下阶段性、机械性和重复性的劳动使农业面临的劳动力短缺问题更加严峻<sup>[3]</sup>。劳动力的减少

导致农作物生产能力下降与采摘成本上升<sup>[4]</sup>。因此实现采摘过程的自动化和智能化是解决劳动力短缺问题的最佳方案之一<sup>[5]</sup>。采摘机器人通常含有3个子系统:视觉系统、动作执行系统和末端执行器<sup>[6]</sup>。视觉系统负责对采摘环境中的果实进行识别与定位,动作执行系统和末端执行器负责采摘。而视觉系统的运行又可分为3个阶段:数据获取、目标识别、三维定位。视觉系统识别效率与定位精度将直接影响到采摘机器人的工作效率与成功率<sup>[7-8]</sup>。

在农业生产中,由于果实生长环境不同、空间分布杂乱,果实的大小、形状和颜色等特性也会有所不同<sup>[9]</sup>。自然环境中的果实普遍存在着因果实重

收稿日期:2022-05-20

基金项目:国家自然科学基金青年科学基金项目(32102071);江苏省农业科技自主创新基金项目[CX(20)3040];江苏省高等学校自然科学研究项目(21KJB220013)

作者简介:金寿祥(2000-),男,安徽安庆人,硕士研究生,主要研究方向为图像识别、目标检测技术。(E-mail) sx@njfu.edu.cn

通讯作者:周宏平,(E-mail) hpzhou@nifu.edu.cn

叠<sup>[10]</sup>、果实被叶片和枝条遮挡<sup>[11]</sup>、光照条件变化<sup>[12]</sup>而导致的果实图像不清晰的现象。另外,采摘机器人执行采摘动作时对环境产生的扰动,可能干扰之前的定位结果。上述因素给视觉数据的获取和采摘点的识别与定位带来了巨大挑战。

深度学习未出现之前,关于采摘机器人视觉系统中的识别技术已有大量研究。按照技术发展过程可分为以下2类:①传统数字图像处理技术,如基于单一特征的颜色、形状、纹理等特征分析<sup>[13-15]</sup>;②基于传统机器学习的识别方法,如K-means聚类算法、K近邻(K-nearest neighbor, KNN)算法、支持向量机(Support vector machine, SVM)算法等<sup>[16-17]</sup>。近年来,受人工智能技术发展与硬件性能大幅提升的影响,深度学习技术在图像识别、目标检测上取得了非凡的成就<sup>[18]</sup>。相比于传统数字图像处理与机器学习基于颜色、纹理、形状等特征匹配的果实识别方法,深度学习利用深层神经网络强大的特征自学习能力在果实识别与采摘点定位任务中取得了较好的效果<sup>[19-21]</sup>。

本文拟对近年来采摘机器人视觉系统的研究进展进行综述,为后续的研究提供参考。

## 1 视觉数据的获取方法

视觉是一种感知环境的重要方法,是目标识别与定位的基础。计算机视觉数据主要有二维平面数据和三维空间数据。视觉系统由图像传感器和图像处理单元组成,是采摘机器人感知外部环境的重要手段<sup>[22]</sup>。

### 1.1 二维数据的获取

二维视觉数据通常指二维彩色图像,由图像传

感器获得,被广泛用于目标识别。图像传感器按原理可分为电荷耦合器件(Charge coupled device, CCD)传感器和互补金属氧化物半导体(Complementary metal-oxide-semiconductor, CMOS)传感器<sup>[23]</sup>。其中,单目相机是使用最广泛的图像采集设备,其结构简单、价格便宜、使用要求低。RGB (Red, green, blue)图像是最常见的二维图像色彩模式,也有部分学者使用HSV (Hue, saturation, value)、Lab、YIQ (Luminance, in-phase, quadrature-phase)等色彩模式<sup>[24-26]</sup>。彩色图像中包含的颜色、纹理、形状等特征常被用于果实的识别。如,马帅等<sup>[27]</sup>使用HUAWEI mate30设备采集梨园中结果期的梨树RGB图像,构建了自然环境下的梨果实识别数据集,但是不包含夜间情况。针对此问题,Xiong等<sup>[28]</sup>利用CCD相机在人工照明条件下采集夜间绿葡萄的RGB图像,解决了绿葡萄夜间检测和采摘的问题。此外,毕松等<sup>[29]</sup>使用Basler acA2440-20gc相机获取不同光照条件、果实被遮挡等情况下的柑橘RGB图像,实现了复杂环境下的柑橘识别。

由于彩色图像缺少深度信息,难以直接用于采摘点三维定位,通常是将彩色图像与深度图像融合,作为采摘点三维定位的依据<sup>[30]</sup>。即先在彩色图像上对果实进行识别与采摘点定位,再将彩色图像上的坐标映射到相应的深度图上,以该点深度值作为采摘点深度坐标。

### 1.2 三维数据的获取

三维视觉数据是在二维数据的基础上添加了深度方向数据,可为采摘点的三维空间定位提供基础。表1列出了目前常用的三维数据获取方法。

表1 常用三维数据采集设备比较

Table 1 Comparison of common equipments for visual data acquisition

设备	原理	优点	局限性
双目相机	视差	结构简单;分辨率较高	计算复杂度高;对光照变化敏感
RGB-D 相机	SL	短距离检测精度高	容易受到环境光的影响
	TOF	有效检测距离远;受光照影响小	检测精度较低;设备能耗较高
激光雷达	TOF/FMCW	可以在无光环境中使用;检测距离远	检测精度不高;硬件成本高

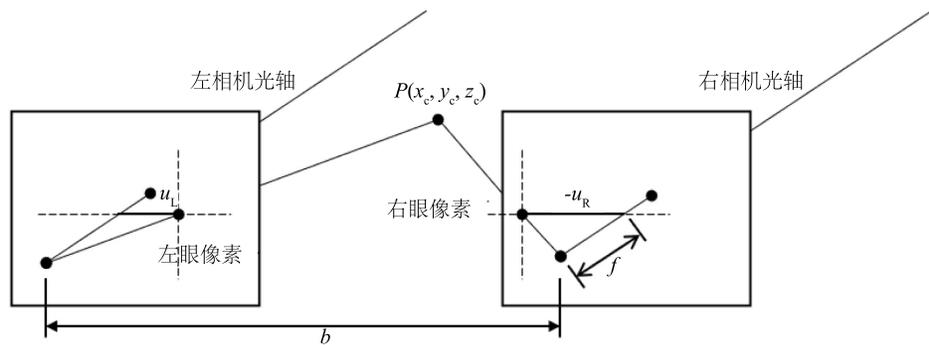
SL:结构光(Structured light);TOF:飞行时间(Time of flight);FMCW:调频连续波雷达(Frequency-modulated continuous wave)。

1.2.1 双目相机 双目相机一般由2个型号相同的单目相机按相应的位置组合而成,通常是水平放置。双目相机通过空间中任意一点的位置在2个相机生成图像中的视差计算该点的深度值<sup>[31]</sup>,原理如

图1所示。

根据图1中的几何关系可得:

$$\frac{Z_c - f}{b - (u_L - u_R)} = \frac{Z_c}{b} \quad (1)$$



$f$ : 相机焦距;  $b$ : 基线距离;  $P(x_c, y_c, z_c)$ : 空间中任意一点;  $u_L$ :  $P$  点在左侧相机  $X$  轴上的坐标;  $u_R$ :  $P$  点在右侧相机  $X$  轴上的坐标。

图1 双目相机测距模型

Fig.1 Ranging model of binocular camera

即可得到  $P$  点的深度坐标  $Z_c$ :

$$Z_c = \frac{fb}{u_L - u_R} \quad (2)$$

其中  $(u_L - u_R)$  即为视差。同理,按照几何关系也可算出  $P$  点横坐标和纵坐标与左右相机之间的关系。但是在相机成像过程中还存在一定程度的畸变,在实际使用中需要通过矫正来抵消畸变影响<sup>[32]</sup>。

双目相机硬件成本低、深度图分辨率高,是较早出现的三维视觉数据采集设备<sup>[33]</sup>。例如, Ling 等<sup>[34]</sup>为了实现番茄自动采摘,使用2个相同的水平放置的 CCD 相机组成的双目相机采集番茄果实的彩色图,建立了一个番茄识别数据集,用于番茄识别,成功率达到了96%。再根据双目测距原理得到番茄的深度坐标,机械手末端定位误差不超过10 mm。同样的,庞超凡<sup>[35]</sup>使用双目相机采集了不同遮挡情况下的小金橘的彩色图像,用来对小金橘进行识别定位,但是对多个被遮挡过多的小金橘定位效果并不佳。为提高多目标定位精度,曹春卿等<sup>[36]</sup>提出一种基于双目视觉技术的多目标苹果识别定位模型,可以准确识别多种场景下的目标苹果。

根据双目相机成像原理,每个像素点的深度信息都需要通过计算获得,计算量较大。另外,双目相机十分依赖图像的特征匹配寻找左右图像中相同的像素点,使得算法复杂度高<sup>[37]</sup>。而且双目相机是一种被动成像相机,所以对照明条件非常敏感,在光照不足、背光和强光等环境下的使用效果较差。

1.2.2 RGB-D 相机 RGB-D (Red, green, blue-depth) 相机一般由一个 RGB 相机和一个深度相机组合而成,可输出视场内的彩色图像和深度点云数

据<sup>[38]</sup>。根据成像原理,深度相机又可分为结构光 (Structured light, SL) 型、飞行时间 (Time-of-flight, TOF) 型和主动立体红外 (Active infrared stereo, AIRS) 型<sup>[39]</sup>,如图2、图3所示。

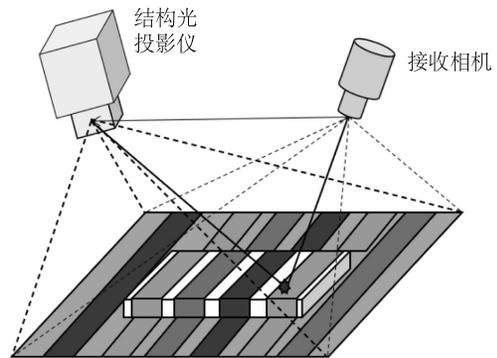


图2 结构光型深度相机测距原理

Fig.2 Ranging principle of structured light depth camera

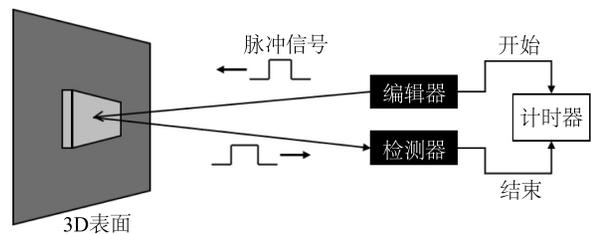


图3 飞行时间型深度相机测距原理

Fig.3 Ranging principle of time-of-flight depth camera

结构光型、飞行时间型和主动立体红外型都使用了红外光作为深度信息的获取方式,差别在于红外光的形式不同。结构光型相机通过发射特定编码后的红外光到被测物体表面,利用红外光从发射到接受过程中的失真程度提取深度信息,典型的设备

有 Kinect v1、Xtion PRO Live 等;飞行时间型相机通过测量红外线从发射到经物体表面反射回接收器的时间计算深度信息,常见的设备有 Kinect v2、Creative Sens3D 等;AIRS 是结构光型相机与双目相机的结合,使用 2 个红外线接收器,弥补了双目相机在室外场景使用的缺点,代表产品有 RealSense D400 系列<sup>[39-42]</sup>。例如,刘景娜<sup>[43]</sup>设计了一种基于 Kinect v1 传感器的番茄信息采集系统,实现了番茄植株外形的三维重建与果实的识别与计数。孙宝霞等<sup>[44]</sup>基于 Kinect v2 与笔记本电脑组成的系统采集了不同光照度下的 920 幅龙眼图像作为龙眼识别数据集。张勤等<sup>[45]</sup>使用 RealSense D435i 相机拍摄不同照明条件下的葡萄串,利用采集到的葡萄串 RGB-D 融合图像实现了复杂环境下的葡萄串识别与采摘点的定位,采摘点定位成功率达到 93.83%、深度误差仅为 3 mm。

结构光型相机近距离检测精度高,不受物体纹理影响,但是其检测精度会随着检测距离变远而极度下降,且易受强光干扰,不适合在室外工作环境中使用;飞行时间型相机检测距离较远,受光照变化和物体纹理的影响较小,但其深度图分辨率相对较低,功耗较高,硬件成本也高<sup>[46]</sup>。

**1.2.3 激光雷达** 激光雷达(Light detection and ranging, LiDAR)按原理可分为飞行时间法和相干探测法两类,都是主动发射激光到被测物体表面生成点云数据,如图 4 所示。激光雷达在自动驾驶目标检测领域已被广泛使用<sup>[47-48]</sup>。近年来,有部分学者将其应用在果实识别与定位上,并取得不错的效果。如,Gené-Mola 等<sup>[49]</sup>为了实现夜间无光条件下的苹果识别,在夜间使用 Velodyne VLP-16 激光雷达扫描仪生成了苹果树的 3D 点云数据,利用苹果树的树叶、枝干和果实对激光的反射率不同,实现了果园中苹果的识别与定位。为了减少遮挡对扫描结果的影响,Mendez 等<sup>[50]</sup>使用 ScanStation P20 型激光扫描仪从 4 个方位扫描橙树,得到橙树四周的三维点云数据。同样的,Tang 等<sup>[51]</sup>使用 Faro Focus 3D X330 地面激光扫描仪获取油茶树点云数据,得到的扫描结果不仅包含点云的强度信息,还包含对应物体表面的 RGB 颜色信息。

激光雷达是一种主动测量的方法<sup>[52]</sup>,测量效果不受光照条件的影响,可以直接用于夜间果实检测<sup>[53]</sup>。但是,激光雷达的价格相比普通相机较贵,

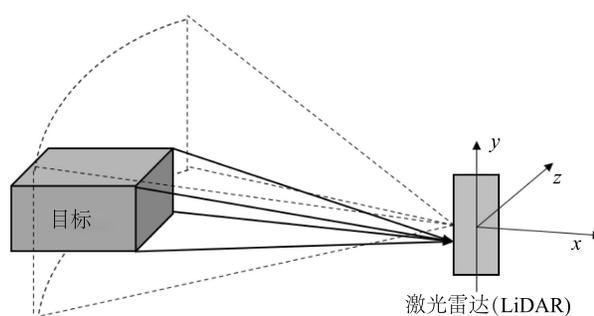


图 4 激光雷达原理

Fig.4 Principle of light detection and ranging

且分辨率不高,另外,对 3D 点云数据的识别技术正处于发展阶段,使基于激光雷达方法的果实识别效果通常差于使用 RGB-D 相机的效果<sup>[54]</sup>。但随着深度学习技术在点云数据处理上的发展,未来的情况可能会有所好转。

## 2 果实识别技术的研究进展

果实的识别是智能化采摘过程中的第一个环节,也是最关键的一步<sup>[55]</sup>,识别的准确率直接影响到后续采摘点的三维定位精度。随着科技的进步,果实识别技术从基于数字图像处理的方法<sup>[56]</sup>发展到基于机器学习的识别方法<sup>[57]</sup>,再发展到基于深度学习的识别方法<sup>[58]</sup>,识别模型的准确性、实时性和鲁棒性不断提高,这也使得采摘机器人能被实际应用。表 2 介绍了一些果实的识别效果。

### 2.1 基于传统数字图像处理的识别方法

人们在观察物体时,通常先注意到物体的形状、颜色和纹理等特征,并基于此判断物体的类别。传统数字图像处理方法就使用了类似原理,先由研究人员通过个人经验筛选出待采摘果实的特征,再利用计算机对所获彩色图像进行特征匹配,当某个区域的匹配率高于设定的阈值时,就认为该区域存在目标果实。

**2.1.1 基于单一特征识别** 基于单一特征的果实识别方法只采用一种果实特征进行识别,需要果实与背景环境有明显的特征区别时才具有可行性。由于该类模型只使用单一特征,所以其结构一般较为简洁,便于操作。

根据果实的生长特性,当可采摘果实与背景颜色差异明显时可使用颜色特征来对果实进行识别。如 Zhuang 等<sup>[59]</sup>通过自适应增强柑橘 RGB 图像中 RG 通道色度,使柑橘果实与背景差异增大,再利用

大津阈值法、形态学操作、凸包操作方法从图像中定位潜在柑橘区域,但该方法受果实颜色与背景差异程度影响,为此,柳长源等<sup>[60]</sup>使用深度图像提出了一种基于果实形态特征的果实识别方法,利用各像素点的梯度向量筛选果实中心点,再以果实中心点搜索果实边界点,平均识别准确率达到84.8%。

表2 采摘机器人视觉系统识别性能

Table 2 Recognition performance of picking robots visual system

识别作物	所用相机	工作环境	检测方法	识别效果			
				准确率(%)	召回率(%)	FI值	检测时间(ms)
苹果 <sup>[24]</sup>	RGB相机	果园	K-means加形态学操作	82.33			806.0
苹果 <sup>[62]</sup>	RGB相机	果园	改进的YOLOv3	94.75	93.09	0.93	54.0
苹果 <sup>[63]</sup>	RGB相机	果园	改进的Mask-RCNN	93.75	89.99		111.0
番茄 <sup>[64]</sup>	RGB相机	温室	颜色特征加形态学处理	80.80	87.50		30.0
番茄 <sup>[65]</sup>	RGB相机	温室	改进的YOLO	96.36	96.03		7.7
荔枝 <sup>[66]</sup>	RGB加TOF	果园	YOLACT	89.70		0.83	154.0
草莓 <sup>[67]</sup>	RGB相机	温室	改进的Mask-RCNN	95.78	95.41		
柑橘 <sup>[68]</sup>	RGB加TOF	果园	改进的YOLOv3			0.91	
柑橘 <sup>[69]</sup>	RGB相机	果园	改进的YOLOv5	93.10	84.30		23.6

TOF:飞行时间型相机。FI:准确率和召回率的调和平均数。

但是,基于单一特征的果实识别方法忽略了果实的其他特征,损失了许多有用的特征信息,导致其对环境敏感程度较高、鲁棒性差。对于一些成熟期果实颜色与背景差异不大,且形状、纹理变化不明显的果实,只选择一种特征很难达到果实识别精度要求。

2.1.2 基于多特征融合识别 彩色图像中包含着形状、颜色、纹理等多种特征信息,使用单一特征进行果实识别会导致大量有效信息被浪费。因此,为了充分利用图像中的信息、提高模型的准确性和鲁棒性,部分学者使用多个特征结合来识别目标果实<sup>[70-72]</sup>。不同特征的优缺点存在明显差异:颜色特征对光照度差异较为敏感;形状与纹理特征对果实重叠和遮挡较为敏感。当这些特征被组合使用时,它们优点互补,对环境变化的敏感程度就会下降,从而提高模型的识别准确率<sup>[73]</sup>。例如,汪杰等<sup>[74]</sup>提出了一种HSV和形状特征融合的花椒识别方法,该方法采用同态滤波进行光照补偿,再利用圆度特征排除背景干扰,提高了复杂背景下花椒的识别准确率。Wu等<sup>[75]</sup>提出了一种融合颜色特征和三维几何特征的水蜜桃识别方法。该方法首先使用颜色特征

Lin等<sup>[61]</sup>提出了一种基于轮廓信息的果实识别方法,使用其推导出的形状描述符表示任意部分的几何特性,再通过特征匹配得到感兴趣的子片段,然后使用改进霍夫变换聚合这些子片段获得候选果实,该方法在柑橘、番茄、南瓜、芒果的识别应用中都取得了不错的效果。

与方向特征对图像进行初步分割,再根据分割结果采用欧式距离算法聚类三维点云数据,最后利用深度直方图分布情况来判断该区域是否为水蜜桃果实,与传统颜色分割方法相比,该方法的分割准确率达到了98.99%。

但是,目前数字图像处理技术中的果实识别方法几乎都需要设置颜色、形状或纹理等阈值,然而每幅图像的最佳阈值都存在差异,导致识别结果的好坏取决于训练数据的质量。基于人工特征提取的方法又存在一定的局限性,造成特征遗漏和利用不充分等现象,影响模型识别准确率。

## 2.2 基于传统机器学习的识别方法

机器学习是人工智能领域的一个重要分支。简单来说,机器学习就是通过算法不断迭代,从大量数据中寻找相应特征。机器学习根据输入数据是否需要被标注,可分为监督学习和无监督学习两类<sup>[76]</sup>。

2.2.1 监督学习 监督学习从人工标注的数据中学习规律,学习效果依赖于数据预处理的质量,需要大量、准确的人工标注数据集。监督学习通过建立数据值与标签值之间的映射关系,经过不断迭代数据、调整模型参数值,最终使模型接近真实映射关

系<sup>[77]</sup>。常见的监督学习算法有 K 最近邻算法 (K-nearest-neighbor, KNN)、AdaBoost 分类器算法、支持向量机 (Support vector machine, SVM)<sup>[78]</sup>。例如, Ling 等<sup>[34]</sup>利用 AdaBoost 分类器检测番茄图像滑动子窗口中包含的是番茄果实还是背景,再使用平均像素值 (APV) 分类器对 AdaBoost 分类器选择的图像进行分割,识别成功率达到 98%。雷欢等<sup>[79]</sup>首先利用颜色直方图提取了苹果的颜色特征,再考虑苹果表皮的局部与全局纹理特征,基于多特征融合与 SVM 构建苹果品种的识别模型,在推理耗时仅 2 ms 的情况下,平均识别准确率达到 94%。

监督学习通过循环迭代学习数据特征与标注信息之间的关系,建立分类、回归模型。因此,模型对于数据标注的要求较高,不允许出现太多的异常值,否则将会导致模型识别效果变差。

**2.2.2 无监督学习** 无监督学习与监督学习相反,一般不要求对数据进行特别分类标注,省去了标注数据的时间。相反,无监督学习算法基于统计学方法,学习总结数据中未体现的规律,将相似特征的数据进行聚类<sup>[80]</sup>。无监督学习算法中,K-means ( $k$  均值聚类) 算法最具有代表性。K-means 算法通过给定的  $k$  值在数据集中找到  $k$  个聚类群落,再反复迭代数据,最后将数据集中的数据按相似程度聚类到  $k$  个群落中,从而达到分类效果。例如,Luo 等<sup>[81]</sup>基于有效颜色分量,利用 K-means 算法获取葡萄园图像中的葡萄簇的像素区域,实现了葡萄簇与背景的分割。Jiao 等<sup>[24]</sup>利用 K-means 算法对 Lab 图像中的苹果与背景进行初步分割,再利用形态学中的侵蚀与膨胀操作提取苹果准确轮廓,能够快速识别重叠苹果。

尽管无监督学习算法省去了数据标注环节,节约了训练周期,但训练结果存在一些不确定性。训练时所用的数据集对聚类模型影响过大,导致无监督聚类算法的鲁棒性较差。另外,K-means 聚类算法需要人为设定  $k$  值,而不同的  $k$  值也会使聚类结果产生较大的差异。传统机器学习方法在简单的果实识别任务中被大量使用,但是在识别背景环境复杂、形态差异不明显的果实时,识别效果还达不到要求。

### 2.3 基于深度学习的识别方法

深度学习扩充了机器学习的范围,属于机器学习的延伸领域<sup>[82]</sup>。深度学习利用深度神经网络和

随机梯度下降 (Stochastic gradient descent, SGD) 算法自动学习、提取目标特征。深层网络能从大量样本中学习到复杂特征,从而使模型获得更好的识别效果。近些年来果实识别研究的相关文献表明,基于深度学习的目标检测方法 (图 5) 已经成为目前果实识别的主要方法。

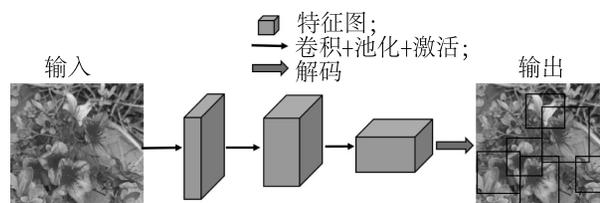
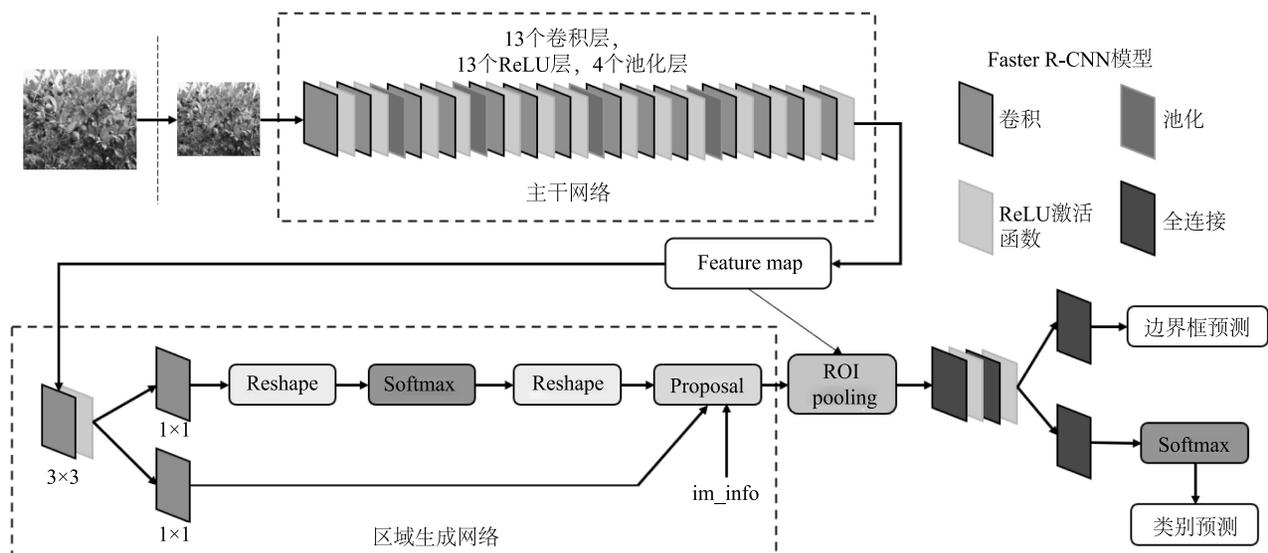


图 5 深度学习目标识别模型示意

Fig.5 Schematic diagram of deep learning target recognition model

**2.3.1 两阶段目标识别算法** 两阶段目标识别、分割算法由生成候选框和候选框分类 2 个阶段组成。第一阶段由区域生成网络 (Region proposal network, RPN) 结构生成一系列目标候选框,第二阶段利用深层神经网络对特征图上每个候选框所对应的特征样本进行分类,再经过非极大值抑制 (Non-maximum suppression, NMS) 处理,剔除冗余的样本<sup>[83]</sup>,如图 6 所示。由于该类算法需要在检测之前生成预选框,导致模型整体结构复杂、训练时间较长、推理速度较慢。例如,李旭等<sup>[84]</sup>基于 Vgg16 为主干的 Faster-RCNN 网络检测复杂背景下的茶叶嫩芽,当采用一芽一叶/二叶标注时,识别准确率与召回率分别为 98% 和 76%,单张图像检测用时 0.20 s,相比数字图像处理的方法快了 0.52 s,且效果更好。针对采摘点难以直接预测的问题,Zheng 等<sup>[85]</sup>在 Mask-RCNN 模型的 ROIAlign 层后添加关键点检测分支用来预测芒果图像中采摘点位置,最终采摘点检测的最佳精度和召回率分别达到了 98.4% 和 90.8%。Wang 等<sup>[86]</sup>在 Faster-RCNN 基础上,使用添加注意力机制的 ResNet50 作为主干网络提高特征提取能力,采用软非极大值抑制方法降低重叠果实漏检率,平均精度达到了 98.46%。

虽然两阶段识别算法在果实识别上获得了不错的成绩,但是由于其复杂的模型结构与庞大的计算量,导致该类算法运行速度难以达到实时检测 (运行速度  $\geq 30$  FPS) 的需求,这也是两阶段算



Feature map:特征层;Reshape:重塑;Softmax:指数归一化函数操作;Proposal:生成感兴趣区域;ROI pooling:感兴趣区域池化;im\_info:图像信息。

图 6 Faster R-CNN 模型网络结构

Fig.6 Network structure of faster R-CNN model

法的通病。而目前移动端平台的性能很难达到此类模型的最佳运行条件,因此,该类算法实际应用的并不多。

**2.3.2 一阶段目标识别算法** 与两阶段目标识别算法不同,一阶段识别算法省去了提取候选框的步骤,直接预测边界框的大小、置信度和类别概率信息。因此,一阶段算法与二阶段算法相比,模型结构更为简单、训练时间更短、运行速度更快<sup>[87]</sup>。由于目前移动端设备计算能力有限,所以在综合考虑检测精度与模型大小的情况下,轻量级的一阶段算法,如YOLO与YOLACT算法被大量使用在果实识别中。例如Liu等<sup>[62]</sup>将YOLOv3算法中用来检测物体位置的方形预测框设置成圆形,用来识别图像中的番茄,减少了模型参数,提高了对圆形果实的拟合程度,相比于未改进之前,该模型的召回率和精度分别提高了2个和3个百分点,达到了93.09%和94.75%。Xu等<sup>[88]</sup>基于YOLOv5s算法对自然环境下的花椒果实进行检测,平均识别精度(MAP)达到了0.945,在NVIDIA Jetson TX2设备上运行速度达到了每秒33帧,实现了实时检测。何斌等<sup>[89]</sup>为实现夜间情况下番茄果实的快速识别,提出了一种基于改进YOLOv5的识别方法(图7),通过在损失函数中增加中心距离度量,提高了被遮挡与暗光下的果实识别精度。杨福增等<sup>[90]</sup>使用基于无锚框的

CenterNet目标检测模型,实现了密集分布场景下的苹果果实识别,该方法省去了锚框和非极大值抑制过程,平均识别精度为93.63%。

与二阶段算法相比,一阶段算法通常能以较少的模型参数量与计算量取得相对较高的果实识别准确率。因此,在果实识别的研究中,越来越多的学者选择采用或者基于一阶段目标检测算法进行改进,构建不同的果实识别模型。

### 3 采摘点定位技术

图像中的果实被识别之后,还需要确定采摘点相对于机器人的空间坐标,才能指导机器人利用动作执行系统与末端执行器完成采摘动作<sup>[91]</sup>。对于不同种类的果实,采摘点有的在果实上,有的在果梗上。如图8所示,苹果<sup>[92]</sup>、柑橘<sup>[93]</sup>等较大的果实,一般选择逐个采摘,通常以果实中心作为采摘点;对于番茄<sup>[94]</sup>、荔枝<sup>[95]</sup>、葡萄<sup>[96]</sup>等小型簇状果实,基本选择整串采摘,一般将果实上部的果梗上的一个点作为采摘点。

采摘点定位准确性将直接决定采摘机器人的采摘效果。通过果实在图像中的位置可以得到采摘点在像素坐标系下的二维坐标,还需要将其转化为机械臂坐标系下的三维坐标,才能正确指引采摘装置完成采摘,流程如图9所示。

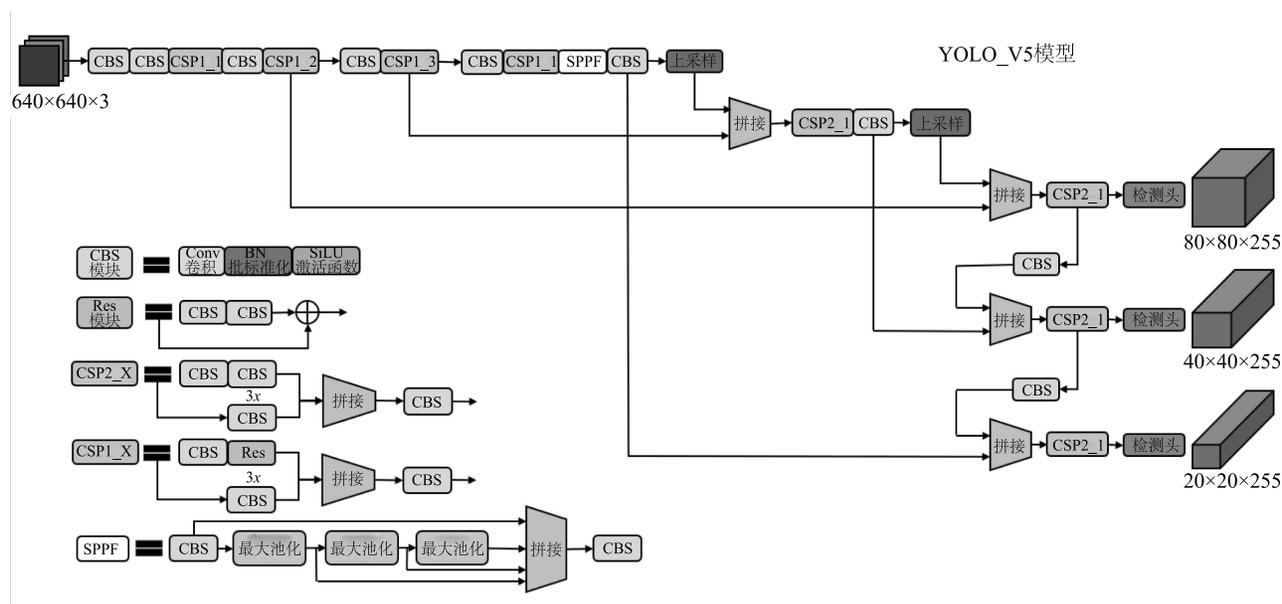


图7 YOLO v5 模型网络结构

Fig.7 Network structure of YOLO v5 model



A: 苹果; b: 柑橘; c: 番茄; d: 葡萄; ☆: 采摘点。

图8 果实采摘点示意

Fig.8 Schematic diagram of fruit picking points

### 3.1 采摘点的二维定位

针对采摘点在果实上位置不同的情况, 一般利用果实与采摘点的位置关系, 将识别出的果实边界框或掩膜的中心作为采摘点, 得到采摘点的像素坐标。例如, 任亚婧等<sup>[97]</sup>使用形态学操作获取图像中的苹果掩膜的质心作为采摘点。但是, 果梗通常是与背景颜色相近的细长条状体, 对于采摘点在果梗或枝干上的情况, 很难直接确定采摘点在果梗上的位置, 所以一般采用间接方法定

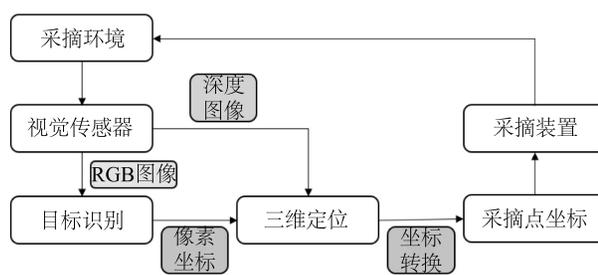


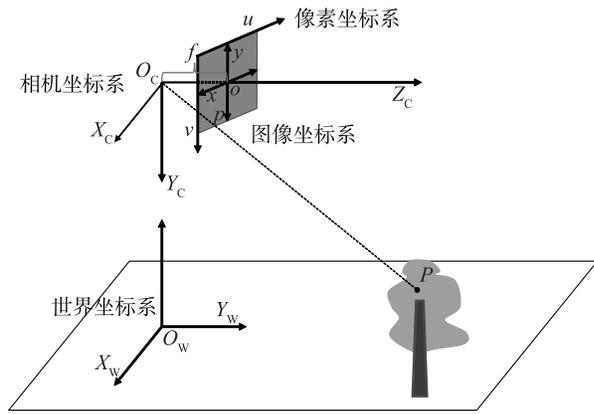
图9 采摘点定位流程

Fig.9 Picking point positioning process

位。根据果实的生长特性, 先确定果实在彩色图像中的位置, 再利用果实、果梗与采摘点之间的自然位置关系, 在图像中推断出果梗上适合采摘的位置作为采摘点。例如, Luo 等<sup>[81]</sup>为了实现葡萄的自动采摘, 利用  $k$  均值聚类与有效颜色分割算法获取重叠葡萄簇的轮廓, 根据每个像素区域的几何信息确定葡萄果梗感兴趣区域, 再利用重心等几何约束关系计算出每个葡萄果梗上的采摘点坐标。同时, 为实现整串番茄的采收, Rong 等<sup>[98]</sup>先使用 YOLCAT++ 网络分割出番茄图像中的番茄簇和果梗, 再判断番茄簇与果梗的位置关系, 剔除无效的果梗掩膜, 然后利用最小二乘法拟合果梗的掩膜, 得到对应的果梗曲线, 通过几何计算, 将果梗曲线的中点作为番茄串采摘点。

经过上述采摘点的定位流程得到的是采摘点在

像素坐标系下的二维坐标,如图 10 中的  $P$  点,该坐标无法直接使用,还需要将其转化到机械臂坐标系(世界坐标系)下的三维坐标,才能正确指引采摘装置执行采摘动作。



$X_w Y_w Z_w$ :世界坐标系,单位为 m,原点为  $O_w$ ;  $X_c Y_c Z_c$ :相机坐标系,原点为  $O_c$ ,单位为 mm,光心  $O_c$  为原点;  $xy$ :图像坐标系,单位为 mm,原点为  $O$ ;  $uv$ :像素坐标系,单位为 pixel;  $P$ :世界坐标系中的一个点;  $p$ :  $P$  在图像坐标系上的投影;  $f$ :相机焦距。

图 10 坐标系位置关系

Fig.10 Position relation of coordinate system

### 3.2 采摘点的三维定位

采摘点的三维定位是由 2D 向 3D 转换的过程。在确定了采摘点的像素坐标后,通过图像与像素坐标系之间的转换关系,得到采摘点在图像坐标系下的二维坐标,利用相机与图像坐标系之间的转换关系,计算得到采摘点在相机坐标系下的三维坐标<sup>[64,99]</sup>。再通过手眼标定方法,获得相机坐标系与采摘机器人坐标系之间的转换关系,最后,通过上述的转换关系得到采摘点在机器人坐标系下的三维坐标<sup>[100]</sup>。

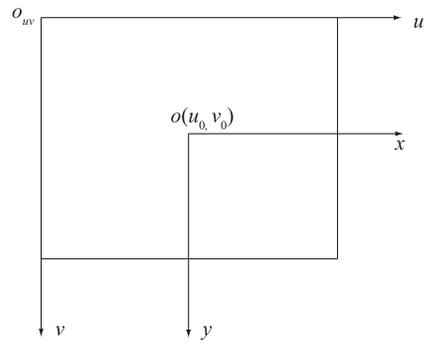
3.2.1 图像坐标系与像素坐标系的关系 如图 11 所示,像素坐标系原点为图像左上角  $O_w$  点,反映像素排列情况<sup>[101]</sup>。图像坐标系的原点为相机光轴与成像平面的交点,通常情况下是成像平面的中点<sup>[102]</sup>。2 个坐标系的转换关系如下:

$$\begin{cases} u = \frac{x}{d_x} + u_0 \\ v = \frac{y}{d_y} + v_0 \end{cases} \quad (3)$$

将公式(3)转换为齐次矩阵形式:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{d_x} & 0 & u_0 \\ 0 & \frac{1}{d_y} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (4)$$

其中  $d_x$  与  $d_y$  分别表示一列和一行像素所对应的图像坐标系中的距离,单位为 mm/pixel。  $(u, v)$  为采摘点在像素坐标系中的坐标,  $(x, y)$  表示采摘点在图像坐标系中的坐标,  $(u_0, v_0)$  为图像坐标系原点  $O$  在像素坐标系上的坐标。



$O, xy, uv$  见图 10 注。  $O_w$  为像素坐标系  $uv$  的原点,  $(u_0, v_0)$  为图像坐标系原点  $O$  在像素坐标系上的坐标。

图 11 图像坐标系与像素坐标系的关系

Fig.11 Relationship between image coordinate system and pixel coordinate system

3.2.2 相机坐标系与图像坐标系的关系 如图 12 所示,相机坐标系是三维坐标系,反映空间中目标物与相机的位置关系。从相机坐标系到图像坐标系是 3D 到 2D 的投影,投影关系如下:

$$\begin{cases} x = f \frac{X_c}{Z_c} \\ y = f \frac{Y_c}{Z_c} \end{cases} \quad (5)$$

其中,  $(X_c, Y_c, Z_c)$  为采摘点  $P$  在相机坐标系下的坐标。

将公式(5)转换成矩阵形式:

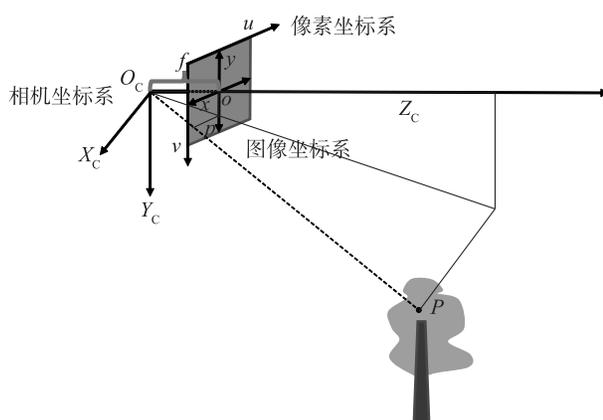
$$Z_c \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} \quad (6)$$

公式(6)展现了相机坐标系到图像坐标系的转换关系,其中  $f$  为相机焦距。  $Z_c$  为采摘点到相机坐标系

原点的深度距离,即采摘点在深度图像与 RGB 图像对应位置上的值。通常是取深度图上与 RGB 图像中识别出的果实掩膜或边界框对应区域的值,经过一些过滤、取平均值等操作,得到采摘点的深度信息<sup>[95]</sup>。

3.2.3 机械臂坐标系与相机坐标系的关系 机械臂与相机存在 2 种位置关系,相机固定称为“眼固定”、相机在机械臂上称为“眼在手上”,如图 13 所示,其中 base 代表基座, end 代表机械臂末端执行器<sup>[103-104]</sup>。

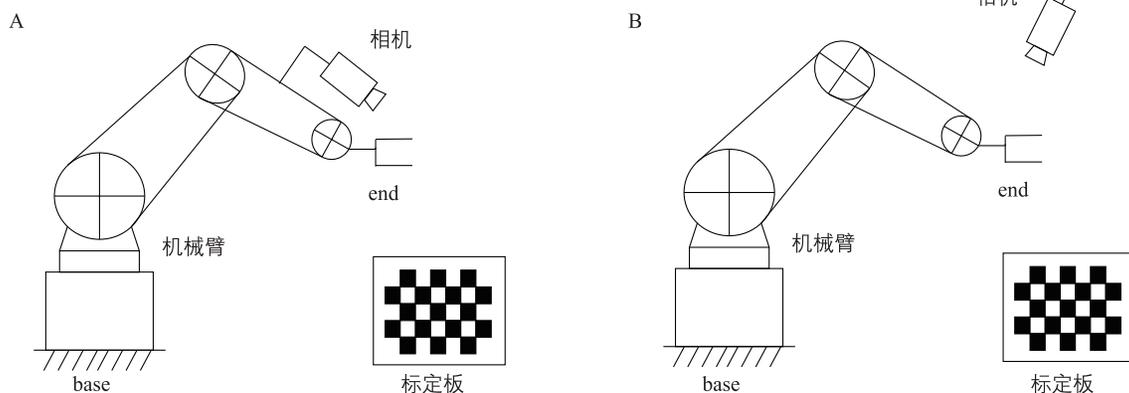
当“眼在手上”时,相机标定的目的是确定机械臂末端与相机坐标系的转换关系;当“眼固定”时,相机标定的目的则是机械臂与相机坐标系的转换关系。2 种转换本质上都是 2 个空间坐标系之间的转换,存在如下的转换关系:



$O_c, O, xy, P, f$  见图 10 注。

图 12 相机坐标系与图像坐标系的关系

Fig.12 Relationship between camera coordinate system and image coordinate system



A:眼在手上;B:眼固定。base:基座;end:机械臂末端执行器。

图 13 相机与机械臂的位置关系

Fig.13 Positional relationship between camera and mechanical arm

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R} & \mathbf{T} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad (7)$$

其中,  $\mathbf{R}$  为旋转矩阵,  $\mathbf{T}$  为平移矩阵,  $(X_w, Y_w, Z_w)$  为采摘点在世界坐标系下的坐标。手眼标定方法通过相机拍摄不同位置的标定板,建立约束方程,得到包含转换矩阵  $\mathbf{X}$  的方程  $\mathbf{AX}=\mathbf{XB}$ , 其中  $\mathbf{A}、\mathbf{B}$  已

知,可求得  $\mathbf{X}$ <sup>[105]</sup>:

$$\mathbf{X} = \begin{bmatrix} \mathbf{R} & \mathbf{T} \\ 0 & 1 \end{bmatrix} \quad (8)$$

式中,  $\mathbf{X}$  为世界坐标系向相机坐标系的转换矩阵。

通过上述推导的坐标系转换关系,由公式(4)、公式(6)、和公式(7)可得到世界坐标系到像素坐标系的转换关系:

$$Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{d_x} & 0 & u_0 \\ 0 & \frac{1}{d_y} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{R} & \mathbf{T} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} = \mathbf{MX} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad (9)$$

式中,  $d_x$ 、 $d_y$  与  $f$  为相机内参, 用矩阵  $M$  表示, 可通过相机标定获得<sup>[106]</sup>。根据采摘点像素坐标 ( $u$ ,  $v$ ) 和深度坐标  $Z_c$ , 利用公式 (9) 就可计算出采摘点的三维坐标。

例如, Li 等<sup>[107]</sup> 基于无人机搭建了一个龙眼的自动化采摘平台, 利用以 MobileNet 为主干的 YOLOv4 网络识别 RGB 图像中的龙眼, 以识别到的龙眼边界框与果梗的交点作为采摘点, 再将图像上的采摘点映射到深度图上获得采摘点深度坐标, 计算得到采摘点的空间坐标。采摘点的识别准确率达到 88.89%, 采摘成功率达到 66.67%。同时, Li 等<sup>[99]</sup> 设计了一种茶叶采摘机器人, 该机器人通过 RGB-D 相机采集茶叶茎 RGB-D 融合图像, 利用 YOLOv3 模型识别茶叶茎, 选择茶叶茎与边界框下边的交点作为采摘点, 再通过坐标转换关系得到采摘点在机械臂坐标系下的三维坐标, 茶叶采摘成功率为 83.18%。同时, 为了简化采摘点从 2D 向 3D 的转换过程, Gené-Mola 等<sup>[49]</sup> 利用移动地面激光扫描仪 (MTLS) 生成苹果树的 3D 点云图像, 根据苹果果实、叶片和枝干对激光雷达反射率的不同对其进行了分割, 定位和识别的成功率分别达到了 87.5% 和 82.4%, 取得了与基于 RGB-D 方法相近的成功率, 且该方法不受光照变化的影响, 还可以直接从 3D 点云图中获取目标物的三维坐标。因此, 采摘点三维定位流程不一定要遵循先 2D 再 3D 的定位流程, 如何缩短采摘点三维坐标定位流程, 也是未来研究的重要方向。获得采摘点在采摘空间中的三维坐标后, 将该坐标转化并输入到采摘执行器, 通过预先设计的轨迹规划策略, 即可控制采摘装置完成果实采摘的任务。

## 4 发展趋势

经过上文的分析可以发现, 深度学习目标检测技术和高性能硬件的使用, 使得采摘机器人的果实识别能力有了较大提升。但是, 目前的机器人视觉在复杂环境下对果实的识别能力和人类相比仍有差距。未来的研究过程中, 如何继续在提高模型的识别准确率的同时简化模型结构仍是重点。因此, 深度学习技术将来仍然是果实识别的主要方法。从目前的采摘机器人视觉系统技术的发展情况看, 本文总结了以下几个发展趋势:

(1) 数据共享。训练深度神经网络模型需要大

量数据, 目前基于一种果实数据训练出的模型只能识别一种果实, 但神经网络的能力远不止于此。未来可以通过数据共享, 训练出丰富的、通用的果实识别模型, 减少重复性的研究。

(2) 规范化种植。当前, 果实检测过程面临着果实受树叶和枝条遮挡、光照变化等不利因素影响, 极大地增加了检测难度。采用规范化种植, 合理安排种植间隙、定期修剪枝干, 可以减少上述不利情况的发生, 降低果实识别难度、提高采摘点定位精度和效率。

(3) 夜间识别。机器人相比人工的优势在于其不间断、重复性的机械劳动。但是目前的采摘点定位研究主要是针对白天的情况展开, 很少考虑到夜间照明情况。这使得采摘机器人丧失了这一重要优势。因此, 实现夜间环境下采摘点的精准定位将大幅提高采摘机器人的使用效率和应用前景。

(4) 高性能硬件。受硬件性能的影响, 图像传感器获得的图像分辨率远小于人眼, 且计算能力也不足以支撑高像素图像计算。因此, 高性能的图像传感器和图像处理器也是机器视觉的研究重点, 它将直接提高视觉系统的检测性能。

(5) 5G 传输技术。面对越来越庞大的检测模型, 移动端设备难以有效承载。利用 5G 数据传输技术, 将视觉传感器获取的数据传输到云端进行计算, 再将检测结果反馈给机器人主控制器, 控制机械手完成果实采摘。利用 5G 传输的快速性和云端的强大运算性能, 可以大大节约计算时间, 提高采摘效率。

### 参考文献:

- [1] 张文翔, 张兵园, 贡宇, 等. 果蔬采摘机器人机械臂研究现状与展望[J]. 中国农机化学报, 2022, 43(9): 232-237.
- [2] WANG Z H, XUN Y, WANG Y K, et al. Review of smart robots for fruit and vegetable picking in agriculture [J]. International Journal of Agricultural and Biological Engineering, 2021, 14(6): 33-54.
- [3] VERBIEST R, RUYSEN K, VANWALLEGHEM T, et al. Automation and robotics in the cultivation of pome fruit: where do we stand today? [J]. Journal of Field Robotics, 2021, 38(4): 513-531.
- [4] BAI Y H, GUO Y X, ZHANG Q, et al. Multi-network fusion algorithm with transfer learning for green cucumber segmentation and recognition under complex natural environment [J]. Computers and Electronics in Agriculture, 2022, 194: 1-27.

- [5] 姚成胜,肖雅雯,杨一单. 农业劳动力转移与农业机械化对中国粮食生产的关联影响分析[J]. 农业现代化研究, 2022,43(2): 1-15.
- [6] MONTOYA-CAVERO L, DÍAZ DE LEÓN TORRES R, GÓMEZ-ESPINOSA A, et al. Vision systems for harvesting robots: produce detection and localization[J]. Computers and Electronics in Agriculture, 2022,192: 1-27.
- [7] 李会宾,史云. 果园采摘机器人研究综述[J]. 中国农业信息, 2019,31(6): 1-9.
- [8] 刘好,刘洒,杨长辉,等. 基于双目立体视觉的重叠柑橘空间定位[J]. 中国农业科技导报, 2020,22(9): 104-112.
- [9] JUN J, KIM J, SEOL J, et al. Towards an efficient tomato harvesting robot: 3D perception, manipulation, and end-effector[J]. IEEE Access, 2021,9: 17631-17640.
- [10] LV J D, WANG Y J, NI H M, et al. Method for discriminating of the shape of overlapped apple fruit images[J]. Biosystems Engineering, 2019,186: 118-129.
- [11] 刘振宇,丁宇祺. 自然环境中被遮挡果实的识别方法研究[J]. 计算机应用研究, 2020,37(S2): 333-335.
- [12] 何斌,张亦博,龚健林,等. 基于改进 YOLOv5 的夜间温室番茄果实快速识别[J]. 农业机械学报, 2022,53(5): 1-10.
- [13] 雷旺雄,卢军. 葡萄采摘机器人采摘点的视觉定位[J]. 江苏农业学报, 2020,36(4): 1015-1021.
- [14] 陈志健,伍德林,刘路,等. 复杂背景下油茶果采收机重叠果实定位方法研究[J]. 安徽农业大学学报, 2021,48(5): 842-848.
- [15] 汪杰,陈曼龙,李奎,等. 基于 HSV 与形状特征融合的花椒图像识别[J]. 中国农机化学报, 2021,42(10): 180-185.
- [16] 杨帆,李鹏飞,刘庚,等. 橘子采摘机器人目标识别定位方法与实验研究[J]. 西安理工大学学报, 2018,34(4): 460-467.
- [17] 王瑾,王瑞荣,李晓红. 番茄采摘机器人目标识别方法研究[J]. 江苏农业科学, 2021,49(20): 217-222.
- [18] 赵立新,邢润哲,白银光,等. 深度学习在目标检测的研究综述[J]. 科学技术与工程, 2021,21(30): 12787-12795.
- [19] 包晓敏,王思琪. 基于深度学习的目标检测算法综述[J]. 传感器与微系统, 2022,41(4): 5-9.
- [20] 杨长辉,刘艳平,王毅,等. 自然环境下柑橘采摘机器人识别定位系统研究[J]. 农业机械学报, 2019,50(12): 14-22.
- [21] 傅隆生,冯亚利,ELKAMIL T,等. 基于卷积神经网络的田间多簇猕猴桃图像识别方法[J]. 农业工程学报, 2018,34(2): 205-211.
- [22] TANG Y C, CHEN M Y, WANG C L, et al. Recognition and localization methods for vision-based fruit picking robots: a review[J]. Frontiers in Plant Science, 2020,11: 1-17.
- [23] 魏宏飞,张季萌. 基于 CMOS 传感器的采摘机器人计算机视觉系统研究[J]. 农机化研究, 2022,44(12): 221-224.
- [24] JIAO Y H, LUO R, LI Q W, et al. Detection and localization of overlapped fruits application in an apple harvesting robot[J]. Electronics, 2020,9(6): 1-14.
- [25] MALIK M H, ZHANG T, LI H, et al. Mature tomato fruit detection algorithm based on improved HSV and watershed algorithm[J]. IFAC-PapersOnLine, 2018,51(17): 431-436.
- [26] 牛晗,伍希志. 基于大津算法连通域的松果多目标识别定位[J]. 江苏农业科学, 2021,49(15): 193-198.
- [27] 马帅,张艳,周桂红,等. 基于改进 YOLOv4 模型的自然环境下梨果实识别[J]. 河北农业大学学报, 2022,45(3): 105-111.
- [28] XIONG J T, LIU Z, LIN R, et al. Green grape detection and picking-point calculation in a night-time natural environment using a charge-coupled device (CCD) vision sensor with artificial illumination[J]. Sensors, 2018,18(4): 1-17.
- [29] 毕松,高峰,陈俊文,等. 基于深度卷积神经网络的柑橘目标识别方法[J]. 农业机械学报, 2019,50(5): 181-186.
- [30] ZHOU H Y, WANG X, AU W, et al. Intelligent robots for fruit harvesting: recent developments and future challenges[J]. Precision Agriculture, 2022,23(5): 1856-1907.
- [31] 刘好,刘洒,杨长辉,等. 基于双目立体视觉的重叠柑橘空间定位[J]. 中国农业科技导报, 2020,22(9): 104-112.
- [32] 邹朋朋,张滋黎,王平,等. 基于共线向量与平面单应性的双目相机标定方法[J]. 光学学报, 2017,37(11): 244-252.
- [33] 杨皓天,万腾. 葡萄采摘机械臂的双目定位与抓取精度研究[J]. 农机化研究, 2022,44(12): 49-54.
- [34] LING X, ZHAO Y S, GONG L, et al. Dual-arm cooperation and implementing for robotic harvesting tomato using binocular vision[J]. Robotics and Autonomous Systems, 2019,114: 134-143.
- [35] 庞超凡. 基于双目视觉小金果果实的识别及定位采摘研究[D]. 郑州:河南农业大学, 2021.
- [36] 曹春卿,张吴平,李富忠,等. 自然场景下多目标苹果识别定位融合算法研究[J]. 湖北农业科学, 2022,61(7): 145-151.
- [37] 陈炎,杨丽丽,王振鹏. 双目视觉的匹配算法综述[J]. 图论学报, 2020,41(5): 702-708.
- [38] LI Y J, FENG Q C, LI T, et al. Advance of target visual information acquisition technology for fresh fruit robotic harvesting: a review[J]. Agronomy, 2022,12(6): 1-19.
- [39] FU L S, GAO F F, WU J Z, et al. Application of consumer RGB-D cameras for fruit detection and localization in field: a critical review[J]. Computers and Electronics in Agriculture, 2020,177: 1-12.
- [40] GONGAL A, AMATYA S, KARKEE M, et al. Sensors and systems for fruit detection and localization: a review[J]. Computers and Electronics in Agriculture, 2015,116: 8-19.
- [41] SUN Q X, CHAI X J, ZENG Z K, et al. Noise-tolerant RGB-D feature fusion network for outdoor fruit detect[J]. Computers and Electronics in Agriculture, 2022,198: 1-13.
- [42] 赵辉,李浩,岳有军,等. 基于 RGB-D 相机的矮砧苹果识别与定位[J]. 计算机工程与设计, 2020,41(8): 2278-2283.
- [43] 刘景娜. 基于 Kinect 的移动式番茄生长信息采集系统的研制[D]. 南京:南京农业大学, 2020.
- [44] 孙宝霞,郑镇辉,胡文馨,等. 基于 RGB-D 的龙眼实时检测与

- 定位方法[J]. 林业工程学报, 2022,7(3): 150-157.
- [45] 张勤,陈建敏,李彬,等. 基于RGB-D信息融合和目标检测的番茄串采摘点识别定位方法[J]. 农业工程学报, 2021, 37(18): 143-152.
- [46] 彭孝东,时磊,何静,等. 消费级RGB-D相机在农业领域应用现状与发展趋势[J]. 中国农机化学报, 2022,43(4): 206-215.
- [47] 彭育辉,江铭,马中原,等. 汽车自动驾驶关键技术研究进展[J]. 福州大学学报(自然科学版), 2021,49(5): 691-701.
- [48] 罗玉涛,秦瀚. 基于稀疏彩色点云的自动驾驶汽车3D目标检测方法[J]. 汽车工程, 2021,43(4): 492-500.
- [49] GENÉ-MOLA J, GREGORIO E, GUEVARA J, et al. Fruit detection in an apple orchard using a mobile terrestrial laser scanner[J]. *Biosystems Engineering*, 2019,187: 171-184.
- [50] MÉNDEZ V, PÉREZ-ROMERO A, SOLA-GUIRADO R, et al. In-field estimation of orange number and size by 3D laser scanning[J]. *Agronomy*, 2019,9(12): 1-18.
- [51] TANG J, JIANG F G, LONG Y, et al. Identification of the yield of camellia oleifera based on color space by the optimized mean shift clustering algorithm using terrestrial laser scanning[J]. *Remote Sensing*, 2022,14(3): 1-18.
- [52] 温玉维,邓长勇,曾德培,等. 三维激光扫描仪在电力工程实测量中的应用[J]. 测绘通报, 2021(10): 163-167.
- [53] TSOULIAS N, PARAFOROS D S, XANTHOPOULOS G, et al. Apple shape detection based on geometric and radiometric features using a LiDAR laser scanner[J]. *Remote Sensing*, 2020, 12(15): 1-18.
- [54] WU Y T, WANG Y Y, ZHANG S W, et al. Deep 3D object detection networks using LiDAR data: a review[J]. *IEEE Sensors Journal*, 2021,21(2): 1152-1171.
- [55] JIA W K, ZHANG Y, LIAN J, et al. Apple harvesting robot under information technology: a review[J]. *International Journal of Advanced Robotic Systems*, 2020,17(3): 1-16.
- [56] ZHUANG J J, HOU C J, TANG Y, et al. Computer vision-based localisation of picking points for automatic litchi harvesting applications towards natural scenarios[J]. *Biosystems Engineering*, 2019,187: 1-20.
- [57] JIAO Y H, LUO R, LI Q W, et al. Detection and localization of overlapped fruits application in an apple harvesting robot[J]. *Electronics*, 2020,9(6): 1-14.
- [58] 宁政通,罗陆锋,廖嘉欣,等. 基于深度学习的葡萄果梗识别与最优采摘定位[J]. 农业工程学报, 2021,37(9): 222-229.
- [59] ZHUANG J J, LUO S M, HOU C J, et al. Detection of orchard citrus fruits using a monocular machine vision-based method for automatic fruit picking applications[J]. *Computers and Electronics in Agriculture*, 2018,152: 64-73.
- [60] 柳长源,赖楠旭,毕晓君. 基于深度图像的球形果实识别定位算法[J]. 农业机械学报, 2022,53(10): 228-235.
- [61] LIN G C, TANG Y C, ZOU X J, et al. Fruit detection in natural environment using partial shape matching and probabilistic Hough transform[J]. *Precision Agriculture*, 2020,21(1): 160-177.
- [62] LIU G X, NOUAZE J C, TOUKO P L, et al. YOLO-Tomato: a robust algorithm for tomato detection based on YOLOv3[J]. *Sensors*, 2020,20(7): 1-20.
- [63] 高梦圆,马双宝,董玉婕,等. 基于实例分割苹果采摘机器人视觉定位与检测[J]. 江苏农业科学, 2022,50(3): 201-208.
- [64] BENAVIDES M, CANTÓN-GARBÍN M, SÁNCHEZ-MOLINA J A, et al. Automatic tomato and peduncle location system based on computer vision for use in robotized harvesting[J]. *Applied Sciences*, 2020,10(17): 1-21.
- [65] 刘芳,刘玉坤,林森,等. 基于改进型YOLO的复杂环境下番茄果实快速识别方法[J]. 农业机械学报, 2020,51(6): 229-237.
- [66] ZHONG Z, XIONG J T, ZHENG Z H, et al. A method for litchi picking points calculation in natural environment based on main fruit bearing branch detection[J]. *Computers and Electronics in Agriculture*, 2021,189: 1-11.
- [67] YU Y, ZHANG K I, YANG L, et al. Fruit detection for strawberry harvesting robot in non-structural environment based on Mask-RCNN[J]. *Computers and Electronics in Agriculture*, 2019,163: 1-9.
- [68] 杨长辉,刘艳平,王毅,等. 自然环境下柑橘采摘机器人识别定位系统研究[J]. 农业机械学报, 2019,50(12): 14-22.
- [69] 黄彤镔,黄河清,李震,等. 基于YOLOv5改进模型的柑橘果实识别方法[J]. 华中农业大学学报, 2022,41(4): 170-177.
- [70] TANG Y F, ZHANG Y W, ZHU Y. A research on the fruit recognition algorithm based on the multi-feature fusion[C]. Harbin: ICMCCE, 2020.
- [71] ZHANG C L, ZHANG K F, GE L Z, et al. A method for organs classification and fruit counting on pomegranate trees based on multi-features fusion and support vector machine by 3D point cloud[J]. *Scientia Horticulturae*, 2021,278: 1-9.
- [72] 雷欢,焦泽昱,马敬奇,等. 基于多特征融合与SVM的苹果品种快速识别算法[J]. 自动化与信息工程, 2020,41(4): 13-17.
- [73] WU J G, ZHANG B H, ZHOU J, et al. Automatic recognition of ripening tomatoes by combining multi-feature fusion with a bi-layer classification strategy for harvesting robots[J]. *Sensors*, 2019, 19(3): 1-22.
- [74] 汪杰,陈曼龙,李奎,等. 基于HSV与形状特征融合的花椒图像识别[J]. 中国农机化学报, 2021,42(10): 180-185.
- [75] WU G, LI B, ZHU Q B, et al. Using color and 3D geometry features to segment fruit point cloud and improve fruit recognition accuracy[J]. *Computers and Electronics in Agriculture*, 2020,174: 1-8.
- [76] 杨晓静,张福东,胡长斌. 机器学习综述[J]. 科技经济市场, 2021(10): 40-42.
- [77] JIANG T, GRADUS J L, ROSELLINI A J. Supervised machine learning: a brief primer[J]. *Behavior Therapy*, 2020,51(5): 675-687.

- [78] 郑太雄, 江明哲, 冯明驰. 基于视觉的采摘机器人目标识别与定位方法研究综述[J]. 仪器仪表学报, 2021, 42(9): 28-51.
- [79] 雷欢, 焦泽昱, 马敬奇, 等. 基于多特征融合与SVM的苹果品种快速识别算法[J]. 自动化与信息工程, 2020, 41(4): 13-17.
- [80] ALZUBI J, NAYYAR A, KUMAR A. Machine learning from theory to algorithms: an overview [C]. Bangalore, INDIA: NCCI, 2018.
- [81] LUO L F, TANG Y C, LU Q H, et al. A vision methodology for harvesting robot to detect cutting points on peduncles of double overlapping grape clusters in a vineyard[J]. Computers in Industry, 2018, 99: 130-139.
- [82] 余凯, 贾磊, 陈雨强, 等. 深度学习的昨天、今天和明天[J]. 计算机研究与发展, 2013, 50(9): 1799-1804.
- [83] GUO Y M, LIU Y, OERLEMANS A, et al. Deep learning for visual understanding: a review[J]. Neurocomputing, 2016, 187: 27-48.
- [84] 李旭, 李振海, 杨海滨, 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): 217-214.
- [85] ZHENG C, CHEN P F, PANG J, et al. A mango picking vision algorithm on instance segmentation and key point detection from RGB images in an open orchard [J]. Biosystems Engineering, 2021, 206: 32-54.
- [86] WANG P, NIU T, HE D J. Tomato young fruits detection method under near color background based on improved Faster R-CNN with attention mechanism[J]. Agriculture, 2021, 11(11): 1-13.
- [87] 李章维, 胡安顺, 王晓飞. 基于视觉的目标检测方法综述[J]. 计算机工程与应用, 2020, 56(8): 1-9.
- [88] XU Z B, HUANG X P, HUANG Y, et al. A real-time *Zanthoxylum* target detection method for an intelligent picking robot under a complex background, based on an improved YOLOv5s architecture [J]. Sensors, 2022, 22(2): 1-15.
- [89] 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识别[J]. 农业机械学报, 2022, 53(5): 201-208.
- [90] 杨福增, 雷小燕, 刘志杰, 等. 基于CenterNet的密集场景下多苹果目标快速识别方法[J]. 农业机械学报, 2022, 53(2): 265-273.
- [91] MONTOYA-CAVERO L, TORRES R D D L, GÓMEZ-ESPINOSA A, et al. Vision systems for harvesting robots: produce detection and localization [J]. Computers and Electronics in Agriculture, 2022, 192: 1-27.
- [92] LI T, FENG Q C, QIU Q, et al. Occluded apple fruit detection and localization with a frustum-based point-cloud-processing approach for robotic harvesting[J]. Remote Sensing, 2022, 14(3): 1-18.
- [93] 熊棣文, 孔文斌, 冯洋. 在树柑桔果实识别与定位技术发展现状及展望[J]. 中国南方果树, 2021, 50(2): 185-190.
- [94] BENAVIDES M, CANTÓN-GARBÍN M, SÁNCHEZ-MOLINA J A, et al. Automatic tomato and peduncle location system based on computer vision for use in robotized harvesting[J]. Applied Sciences, 2020, 10(17): 1-21.
- [95] LI J H, TANG Y C, ZOU X J, et al. Detection of fruit-bearing branches and localization of litchi clusters for vision-based harvesting robots[J]. IEEE Access, 2020, 8: 117746-117758.
- [96] 罗陆锋, 邹湘军, 熊俊涛, 等. 自然环境下葡萄采摘机器人采摘点的自动定位[J]. 农业工程学报, 2015, 31(2): 14-21.
- [97] 任亚娟, 张宁宁, 徐媛媛, 等. 基于视觉识别的成熟苹果识别及采摘定位系统[J]. 现代电子技术, 2021, 44(11): 73-77.
- [98] RONG J C, DAI G L, WANG P B. A peduncle detection method of tomato for autonomous harvesting [J]. Complex & Intelligent Systems, 2021, 8(4): 2955-2969.
- [99] LI Y T, HE L Y, JIA J M, et al. In-field tea shoot detection and 3D localization using an RGB-D camera[J]. Computers and Electronics in Agriculture, 2021, 185: 1-12.
- [100] 王芳, 崔丹丹, 李林. 基于深度学习的采摘机器人目标识别定位算法[J]. 电子测量技术, 2021, 44(20): 162-167.
- [101] 李瑞龙. 自动驾驶场景下的三维目标检测技术研究[D]. 长春: 中国科学院大学(中国科学院长春光学精密机械与物理研究所), 2022.
- [102] 齐锐丽. 基于机器视觉花椒目标识别与定位技术研究[D]. 汉中: 陕西理工大学, 2020.
- [103] 周俊, 刘锐, 张高阳. 基于立体视觉的水果采摘机器人系统设计[J]. 农业机械学报, 2010, 41(6): 158-162.
- [104] 胡小平, 左富勇, 谢珂. 微装配机器人手眼标定方法研究[J]. 仪器仪表学报, 2012, 33(7): 1521-1526.
- [105] KLAUS H S, GERD H. Optimal hand-eye calibration [C]. Beijing: IEEE, 2006.
- [106] 卜令昕. 结构化果园苹果收获机器人关键技术研究[D]. 杨凌: 西北农林科技大学, 2021.
- [107] LI D H, SUN X X, ELKHOULHAA H, et al. Fast detection and location of longan fruits using UAV images[J]. Computers and Electronics in Agriculture, 2021, 190: 1-15.

(责任编辑: 陈海霞)