

果实目标深度学习识别技术研究进展

宋怀波^{1,2} 尚钰莹^{1,2} 何东健^{1,2}

(1. 西北农林科技大学机械与电子工程学院, 陕西杨凌 712100;
2. 农业农村部农业物联网重点实验室, 陕西杨凌 712100)

摘要: 机器视觉技术是果实目标识别与定位研究的关键。传统的目标识别算法准确率较低、检测速度较慢,难以满足实际生产的需求。近年来,深度学习方法在果实目标识别与定位任务中表现出了优良的性能。本文从数据集制备与果实目标识别模型两方面进行综述,总结了数据集制备相关的有监督、半监督和无监督3种方法的特点,按照深度学习算法的发展历程,归纳了基于深度学习的果实目标检测和分割技术的常用方法及其实际应用,轻量化模型的研究进展及其应用情况,基于深度学习的果实目标识别技术面临的问题和挑战。最后指出基于深度学习的果实目标识别方法未来发展趋势为:通过弱监督学习来降低模型对数据标签的依赖性,提高轻量化模型的检测速度以实现果实目标的实时准确检测。

关键词: 果实; 深度学习; 目标识别; 分割

中图分类号: TP391.4 文献标识码: A 文章编号: 1000-1298(2023)01-0001-19 **OSID:** 

Review on Deep Learning Technology for Fruit Target Recognition

SONG Huaibo^{1,2} SHANG Yuying^{1,2} HE Dongjian^{1,2}

(1. College of Mechanical and Electronic Engineering, Northwest A&F University, Yangling, Shaanxi 712100, China
2. Key Laboratory of Agricultural Internet of Things, Ministry of Agriculture and Rural Affairs, Yangling, Shaanxi 712100, China)

Abstract: Machine vision technology is the key of fruit target recognition and positioning. Traditional target recognition algorithm has low accuracy and slow detection speed, which is difficult to meet the needs of actual production. In recent years, deep learning methods have shown excellent performance in fruit target recognition and localization tasks. The fruit target recognition algorithm based on deep learning has the advantages of high detection progress and fast detection speed, so it is widely used in the fruit target recognition task under different scenes and has achieved many good achievements. The data set preparation and fruit target recognition models were reviewed. Firstly, the characteristics of supervised, semi-supervised and unsupervised methods related to dataset preparation were summarized. Secondly, according to the development process of deep learning algorithm, the common methods and practical applications of deep learning-based fruit target detection and segmentation technology were summarized, the previous research on the detection and segmentation of fruit objects such as apple, citrus and tomato under different natural scenes was summarized, and the research progress and application of lightweight model were summarized. Thirdly, the problems and challenges of deep learning-based fruit target recognition technology were summarized. In the end, the future development trend of deep learning-based fruit target recognition methods was pointed out, that was, weakly supervised learning would be used to reduce the dependence of models on data labels, and the detection speed of lightweight models would be improved to achieve real-time and accurate detection of fruit targets.

Key words: fruit; deep learning; target recognition; segmentation

0 引言

随着水果种植产业的迅速发展及劳动力资源的

日益紧缺,开发智能化、自动化农业智能装备的需求在不断增加,果实采摘机器人已成为农业发展的重
要方向^[1]。利用机器人进行采摘作业不仅可以提

高采摘效率且降低了劳动成本,有利于提高果农的经济效益^[2]。实现自然场景下果实的准确识别与定位,可为果实采摘机器人的视觉系统提供关键的技术支持^[3-4]。利用机器视觉技术对果实目标进行检测,对于果实的生长监测、产量预测^[5-6]、果实分拣等任务也具有重要意义,是实施精准农业技术的重要步骤之一^[7]。本文以果实采摘为例,对果实目标识别技术研究成果进行综述。

果实目标识别方法主要包括基于手工设计特征的传统识别方法和基于卷积神经网络(Convolutional neural network, CNN)的深度学习方法两种。传统目标检测算法主要包括区域选择、特征提取和分类3个步骤。传统目标检测算法相对成熟,然而,在复杂的自然场景下,果实目标检测任务仍存在难点:果实种类以及外观形态具有多样性;光照条件变化;复杂天气情况;复杂背景影响等。这些问题使得手工设计特征的过程更加复杂^[8-9],且传统的目标检测算法存在检测精度较低、检测速度较慢、模型实时性较差、普适性不强等缺点,应用传统目标检测算法进行果实目标检测难以满足果实采摘机器人的实际作业要求。基于深度学习的果实目标检测模型是一种端到端的检测模型,可将目标的特征提取、特征选择和特征分类融合在同一模型中^[10]。深度学习模型具有高度的层次结构和强大的学习能力^[11],在复杂视觉信息与目标感知融合方面具有独特优势^[12]。

虽然深度学习技术在果实目标识别方面取得了很好的效果和进展,然而距离实际作业应用还有一定的距离。如图1所示,本文对苹果、番茄、柑橘等28种果实的相关识别研究成果进行检索(图中没有标注数量的均为1),并以此为基础总结归纳国内外果实目标识别的研究进展、关键技术,分析果实目标识别任务存在的问题和面临的挑战,并对未来发展趋势进行展望,以期对果实目标识别任务的后续研究提供参考。

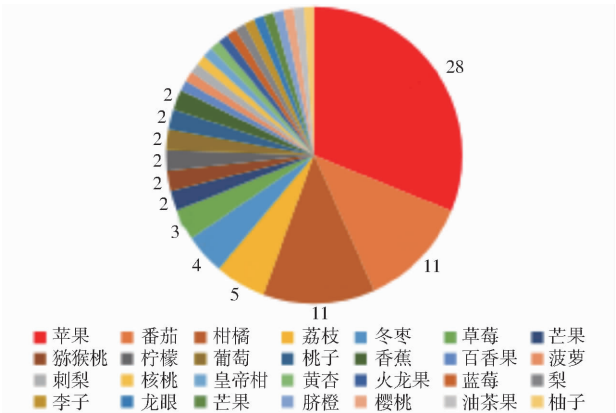


图1 引文中涉及到的果实种类及数量

Fig.1 Species and quantity of fruit involved in citation

1 传统果实目标识别方法

传统的果实目标识别方法是基于颜色、几何形状、纹理等特征对果实目标进行分类、检测和分割^[13]。基于颜色特征的果实目标识别方法主要通过选取合适的颜色模型,利用果实目标与背景区域的像素颜色特征差异,将果实目标与背景分开。利用 YCbCr 颜色模型对荔枝图像进行阈值分割,去除复杂背景,可实现荔枝果实与果梗的识别,综合识别率为 95.50%^[14]。基于归一化红绿色差的苹果分割方法可实现红色苹果与绿色背景的分割,然而当果实目标的颜色与背景颜色相似时,仅利用简单的颜色特征进行果实目标分割难以取得较好的效果^[3]。以归一化的 g 分量和 HSV 颜色空间中 H 、 S 分量为特征参数的支持向量机(Support vector machine, SVM)分类器和以超绿算子($2G-R-B$)为特征的阈值分类器,设计一种用于近色背景中绿色苹果目标识别的混合分类器,该方法平均识别正确率为 89.30%^[15]。

当果实目标与背景的颜色特征较为接近,或者光照条件对果实颜色的影响较大时,相比于颜色特征,利用果实与背景之间的形状和纹理特征的差异可以取得更好的分割效果。利用颜色、形状和纹理特征可识别自然环境中的绿色柑橘,其正确率为 75.30%^[16]。基于边缘曲率分析的重叠番茄识别方法对轻微遮挡的重叠番茄识别正确率为 90.90%,对遮挡率在 25% ~ 50% 之间的番茄目标的识别正确率为 76.90%^[17]。利用柠檬、柑橘等水果近球形的形态特征,结合其深度图像,实现对果实中心点的定位及果实图像的分割,可解决光照和近景色所造成的识别精度较低的问题^[18]。

随着传统目标检测算法的不断发展,手工设计特征的算法性能逐渐趋于饱和^[19],然而,传统目标检测算法仍存在以下局限性:在生成候选区域的过程易产生大量冗余区域;在复杂背景下,基于低级视觉线索设计的特征描述符较难提取具有代表性的语义信息^[20]。所以,对于复杂场景下的果实目标检测,例如:背景较为复杂、目标之间存在严重遮挡、光照不均等情况,传统果实目标检测算法已不能满足需求,且传统目标检测算法在检测速度和模型大小方面均难以满足果实采摘机器人的要求,因此,基于传统方法的果实目标识别技术较难推广到实际应用领域。

2 基于深度学习的果实目标识别方法

深度学习是一种具有多层次表征的学习方

法^[21],深度学习方法通过引入多层感知器结构,利用低级特征形成高级特征,用于最终的目标检测任务^[1]。与传统的目标检测算法相比,深度学习在图像分类、目标检测和识别方面优势明显。由于自然场景下的果实目标具有空间分布随机、存在重叠遮挡、形状多样等非结构化特征,而深度卷积网络可以自动从训练数据中学习特征。因此,深度学习方法可以在复杂的自然场景下表现出更加强大的果实目标识别能力。根据检测组件和目标区域的识别结果,深度学习模型可以分为分类检测模型和分割模型。如图 2a、2b 所示,分类检测模型包括图像分类和目标检测,目标检测的输出是目标类别及其边界框的区域,目标检测任务在完成图像分类任务的同时利用边界框反映目标的位置信息。图像分割是指根据图像属性与目标图像的一致性,对特征相对一致的目标图像进行分割,使同一子区域的特征具有一定的相似性和差异性^[22]。分割模型需要精确的像素级掩码进行目标分割^[12]。如图 2c、2d 所示,分割模型主要包括语义分割和实例分割,语义分割为每个像素分配特定的类别标签,但并不区分同一类别的多个对象,实例分割为每个目标分配单独的分类像素级掩码,可以区分同一类别的不同目标。学者们对于果实目标检测和分割的研究大多是基于在目标识别领域表现较好的网络模型,根据不同果实目标生长的自然环境以及果实目标本身的特点,对网络模型的输入端、骨干网络等结构进行改进,或是引入注意力机制、迁移学习、特征融合、密集连接等操作,以实现提高目标识别的效果,提高模型在复杂场景下的鲁棒性,或是实现模型的轻量化等目标。

2.1 数据集制备方法

图 3 为基于深度学习的果实目标识别及应用的

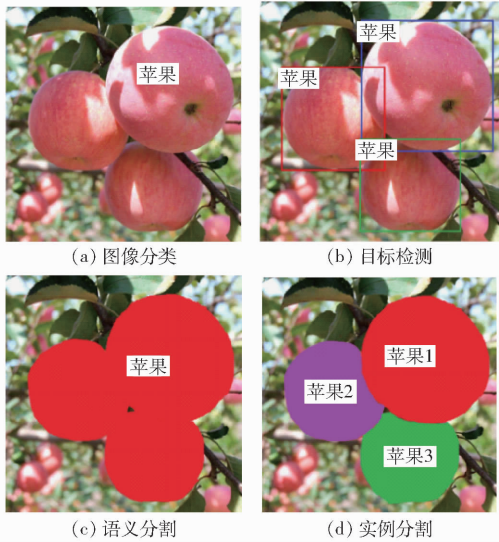


图 2 不同识别任务的比较

Fig. 2 Comparison of different recognition tasks

基本步骤。首先是采集数据,通过对田间采集到的果实目标图像进行处理分析,以实现不同场景下、不同品种、不同生长阶段的果实目标识别任务。接着针对不同目标识别任务的特点进行数据预处理(数据集制备),该过程既包括利用图像增强技术进行图像的颜色、亮度、对比度等的调整,或对图像进行旋转、翻转、裁剪等操作,使输入到网络中的图像更适合于特定的目标识别任务或实现数据集规模的扩大,也包括对数据集进行标注,实现用于深度学习任务的标签文件制作。根据训练数据是否含有标签信息,可将训练过程分为监督学习、无监督学习和半监督学习 3 类。

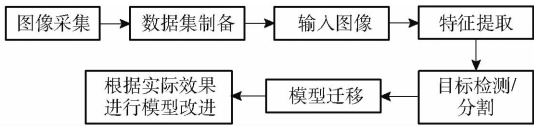


图 3 果实目标识别的基本步骤

Fig. 3 Basic steps of fruit target recognition

基于深度学习的目标识别任务需要一定规模的数据集进行训练,目前基于深度学习的果实目标检测方法大多是基于监督学习,通过向网络中输入一定数量的图像及其对应的标签文件进行模型训练,以提取目标的特征,实现目标识别任务。监督学习是解决分类和回归问题的常用方法^[23]。对于田间果实目标识别任务,训练数据在一定程度上决定了目标检测任务的效果,其规模取决于果实图像的视觉复杂度和深度学习模型的种类。由于果实生长的田间环境较为复杂多变,该环境下采集到的果实目标可能存在遮挡、光照不均、果实目标大小不一、颜色变化等情况,网络的特征提取过程相对困难,因此需要较大规模的数据。对于网络结构较复杂、层数较深的网络,其精确度较高,然而也需要大量的训练数据^[24],且对于不同品种果实的检测任务,都需制定合适的标注策略。

基于监督学习的数据标注过程耗时耗力,效率低下,且对某些特定的识别任务,数据标注过程需要相关专家提供指导。因此如何更有效地标记数据并使用更少的样本进行有效学习是目前该领域的关键问题。利用中小型数据进行半监督学习以获得高精度的结果,为此类研究提供了借鉴,半监督学习的训练数据中仅部分图像有对应的标签,且带标签的数据所占比例较小,通过从带标签的数据中提取到的局部特征进行图像分类^[25]。

无监督学习可以完全脱离数据集标注过程,模型仅利用图像作为训练数据,其通过学习数据的结构,并从数据中提取可区分的信息或特征,将输入映射到特定输出^[26]。聚类是最基本的无监督学习之

一,其目标是将数据分成相似数据点的聚类^[27]。因此,无监督学习在图像分割领域应用较多。尽管如此,监督学习仍然是目前的主流方法,关于半监督学习和无监督学习的研究相对较少。

2.2 基于深度学习的果实目标检测方法

目标检测任务可以分为目标定位和目标分类。随着计算机算力的提升和数据规模的不断扩大,基于深度卷积神经网络(Deep convolutional neural network, DCNN)的目标检测技术逐渐得到发展。自 AlexNet^[28]应用于图像分类任务并赢得 ILSVRC-2012 比赛冠军以来,诸多学者开始致力于 DCNN 的

研究和应用。图4为基于深度学习的目标检测算法的发展历程,图中橙色箭头表示无锚框目标检测算法。基于深度学习的目标检测主要分为两大类:两阶段检测和单阶段检测。两阶段目标检测将目标定位和目标分类任务分离开,首先生成候选区域,再对区域进行分类。其代表算法有 R-CNN^[29]、SPPNet^[30]、Fast R-CNN^[31]、Faster R-CNN^[32]等。单阶段目标检测省去了生成候选区域的过程,直接生成目标的类概率和位置坐标,其过程比两阶段目标检测简单。单阶段目标检测的代表算法有 SSD 系列、YOLO 系列等。

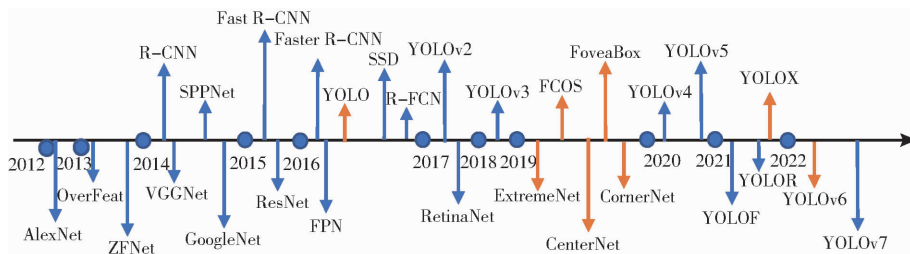


图4 基于深度学习的目标检测算法发展历程

Fig. 4 Development of object detection algorithm based on deep learning

2.2.1 两阶段果实目标检测方法

两阶段检测方法又称为基于候选区域的检测方法。将传统的机器学习方法与 CNN 相结合,提出一种基于 R-CNN 的检测框架,通过选择性搜索获得尽可能多的候选区域,利用 CNN 代替人工提取区域的特征并使用 SVM 进行分类。SPPNet 引入自适应大小的池化,其运行速度比 R-CNN 更快。Fast R-CNN 利用兴趣池化区域(Region of interest Pooling, RoI Pooling)层代替空间金字塔池化(Spatial pyramid pooling, SPP)层,加快了模型的速度。由于 SPPNet 和 Fast R-CNN 生成的候选区域数量过多,导致了大量的计算消耗,因此其应用场景受到了限制。一种区域生成网络(Region proposal network, RPN)方法用于生成候选区域,其输入为骨干网络输出的特征图,输出为一组矩形的候选区域,且每个区域均有一个目标得分^[33]。

Faster R-CNN 模型用 RPN 取代 Fast R-CNN 中的选择性搜索,且通过网络共享的方式生成候选区域,利用 Softmax 分类器完成训练和学习过程,其检测性能有了大幅提高,被广泛地应用于目标检测任务。Faster R-CNN 算法由特征提取器、RPN 和 Fast R-CNN 模块构成。首先对输入网络中的图像进行特征提取,再将提取到的特征输入 RPN 和 Fast R-CNN,生成建议矩形框。如文献[7,33],利用 Faster R-CNN 模型可实现自然场景下的柑橘果实和芒果花穗的检测,然而其 mAP 均相对较低,检测效果不理想。

利用迁移学习的模型对 Faster R-CNN 的模型结构进行改进,可以提高模型的泛化性能和检测精度。如文献[34],利用迁移学习训练基于 Faster R-CNN 的柑橘目标识别模型,可有效降低训练模型的平均损失,且模型平均准确率较高。文献[35]利用迁移学习微调的 AlexNet 网络替换 Faster R-CNN 原始的特征提取层,可解决广域复杂环境中的猕猴桃因枝叶遮挡或部分果实重叠遮挡所导致的识别精度较低的问题。

为了提高检测准确率,降低目标的漏检率,同时提高模型在目标存在遮挡、目标形态和大小存在差异等复杂场景下的鲁棒性,一些学者对 Faster R-CNN 算法进行了改进。基于数据平衡进行数据扩增,可解决 Faster R-CNN 模型检测不同成熟度冬枣的样本数量相差悬殊导致的识别率较低的问题^[36]。利用预训练的 VGG16 网络替换 Faster R-CNN 原始的特征提取层,并改进 RPN 的结构,可提高模型对在体青皮核桃和苹果果实的检测精度的漏检率^[37-38]。将兴趣区域校准引入 Faster R-CNN 框架,可提高 Faster R-CNN 模型对不同形态刺梨的检测精度^[39]。融合 RGB 图像和深度信息,并融合全局和局部信息,可提高 Faster R-CNN 对小目标百香果的检测效果^[40]。

用于目标检测的基于区域的全卷积网络(Region-based fully convolutional networks, R-FCN),由共享的全卷积网络构成,可实现整个图像上的共享计算,有效减少了参数冗余,并引入位置敏感分数图解决了图像分类中的平移不变性与目标检

测中的平移可变性间的矛盾,该模型具有较快的训练和检测速度^[41]。利用 ResNet-44 替换 R-FCN 的原始特征提取网络,可有效识别重叠、枝叶遮挡、模糊及表面有阴影的苹果目标,并简化网络结构^[42]。两阶段果实目标检测算法的相关研究成果如表 1 所示。

表 1 基于两阶段算法的果实目标识别研究成果

Tab.1 Research on fruit target recognition based on two-stage algorithm

果实种类	方法名称	技术原理	技术特点与性能指标	文献序号
绿色柑橘	Faster R-CNN	利用参数调整后的 Faster R-CNN 实现绿色柑橘的识别	对于果径较小和单幅图像中果实数量较多的情况检测效果不理想,其 mAP 为 85.49%	[7]
芒果花穗	Faster R-CNN	利用 Faster R-CNN 实现芒果花穗的定位与计数	模型的 F1 值为 56.00%,精确率为 66.00%,召回率为 48.00%	[33]
柑橘	迁移学习的 Faster R-CNN	利用迁移学习训练 Faster R-CNN 模型实现柑橘目标的识别	模型对于亮度不匀、存在遮挡等情形具有较好的鲁棒性,其 mAP 为 86.60%,最大召回率为 93.00%	[34]
猕猴桃	迁移学习的 Faster R-CNN	利用改进 AlexNet 作为 Faster R-CNN 的特征提取层,实现广域复杂环境下遮挡猕猴桃目标的识别	通过减少全连接层参数压缩网络,减少计算量,模型识别广域复杂环境下存在遮挡情况的猕猴桃果实图像的 mAP 为 96.00%	[35]
冬枣	Faster R-CNN	利用数据平衡进行数据扩增并利用 Faster R-CNN 实现不同成熟度的冬枣识别	数据平衡方法增强了 Faster R-CNN 模型的泛化效果,模型的 mAP 为 98.50%	[36]
青皮核桃	改进 Faster R-CNN	在 Faster R-CNN 的卷积层加入批归一化处理,利用双线性插值法改进 RPN 结构和构建混合损失函数等方式改进模型的适应性	实现了检测速度与精度的相对平衡,模型的精确率为 97.71%,召回率为 94.58%,F1 值为 96.12%,检测速度为 227 ms/幅	[37]
苹果	改进 Faster R-CNN	利用 VGG19 作为 Faster R-CNN 的特征提取层,实现采摘机器人对苹果的识别	有效提高了检测苹果的准确率,其 mAP 为 82.40%	[38]
刺梨	改进 Faster R-CNN	选择基于 Faster R-CNN 的交替优化方式,改进 Faster R-CNN 的区域特征聚集方式提高目标矩形框的精确率	可实现刺梨果实 11 种形态的识别,其召回率最低为 81.40%,最高为 96.93%,精确率最低为 85.63%,最高为 95.53%,F1 值最低为 87.50%,最高为 94.99%	[39]
百香果	改进 Faster R-CNN	融合全局和局部特征增强 Fast R-CNN 的全局和局部信息	提高了检测小目标百香果的精确率和召回率,其精确率、召回率和 F1 值分别为 93.10%、88.50% 和 94.60%	[40]
绿色苹果	改进 R-FCN	利用 ResNet-44 替换 R-FCN 的原始特征提取网络以进行网络简化	可有效识别复杂背景下的绿色苹果,其识别的召回率为 85.70%,准确率为 95.10%	[42]

2.2.2 单阶段果实目标检测方法

虽然两阶段目标检测算法的检测精度较高,但其参数数量和计算量较大,检测速度较慢,难以完成实时检测任务,影响了采摘机器人的工作效率。

表 2 列出了基于单阶段目标检测算法的果实目标识别研究成果。单阶段目标检测算法又称为基于回归的检测方法。LIU 等^[17]提出了一种利用单个深度网络进行目标检测的 SSD 模型,其核心是利用小型卷积滤波器进行多尺度特征映射,生成并预测固定的默认边界框的类别得分和偏移量。由于 SSD 模型实现了端对端的训练,其具有易于训练和集成的优点,与两阶段检测方法相比,SSD 基本实现了检测速度与精度的相对平衡,因此被广泛地应用于果实目标检测任务。利用 ResNet-101 模型替换 SSD 的原始主干网络,可实现苹果、荔枝等 4 种水果的识别,且该模型的检测精度高于原始 SSD 模型^[43]。SSD 模型采用特征金字塔来检测不同尺度的目标,然而由浅层网络生成的小目标特征缺乏足够的语义

信息,导致其对小目标的检测性能较差。FSSD^[78]是一种增加了特征融合的 SSD 模型,该模型在传统 SSD 的基础上增加一个轻量级、高效的特征融合模块,对不同比例的特征图进行融合,以提升对小目标的检测性能。基于改进的轻量化 FSSD 模型可实现灵武长枣的检测,该方法可为灵武长枣的智能化采摘提供一定的技术支持^[44]。基于多重特征增强与特征融合的 MFEFF-SSD 模型可实现无人机拍摄图像中小目标荔枝的检测,然而该方法存在误检和漏检的情况^[45]。基于改进 FSSD 的柚子目标检测模型可有效改善绿叶被误检为膨大期柚子果实的情况^[46]。

YOLO (You only look once) 是 REDMON 等^[79]在 2015 年提出的一种目标检测算法,它是深度学习时期的第一个单阶段目标检测算法。YOLO 将目标检测任务看作单一的回归问题,仅用单个网络便可实现多个边界框的位置和类别预测。YOLO 检测速度快,对背景的误检率较低,且泛化性能较好,然而,

表 2 基于单阶段算法的果实目标识别研究成果

Tab.2 Research on fruit target recognition based on one-stage algorithm

果实种类	方法名称	技术原理	技术特点与性能指标	文献序号
苹果、荔枝、脐橙、皇帝柑	改进 SSD	利用 ResNet-101 模型替换 SSD 的原始主干网络,并运用迁移学习方法和随机梯度下降算法优化模型	模型具有较好的泛化性和鲁棒性,其检测 4 种水果的 AP 为 88.40%,在遮挡面积低于 50% 的情况下 F1 值为 96.12%	[43]
灵武长枣	改进 FSSD	采用改进的 DenseNet 模型作为 SSD 的主干网络,利用 Inception 模块替换 SSD 模型中的前 3 个额外层,并引入特征融合结构	在一定程度上简化了模型结构并提高了检测精度,在不加载预训练模型的情况下其 mAP 为 96.60%,模型参数量为 1.99×10^6	[44]
荔枝	改进 FSSD	利用感受野模块提升主干网络的特征提取能力,使用高效空间金字塔增强浅层特征,采用改进的路径聚合网络进行多尺度特征融合,提升荔枝小目标的检测效果	该模型检测荔枝的 AP 为 55.79%,存在误检和漏检	[45]
柚子	改进 SSD	以改进的 Inceptionv3 作为 FSSD 的骨干网络,并采用 Focal Loss 损失函数改善由于正负样本不平衡导致的检测器误检的情况	在背景复杂的情况下有效提高了膨大期柚子的检测准确率,其检测柚子的精确率为 93.70%	[46]
绿色芒果	YOLOv2	利用 YOLOv2 实现树上绿色芒果的识别	该模型检测芒果果实的 AP 为 86.43%	[47]
绿色芒果	改进 YOLOv2	采用带密集连接的 Tiny-yolo-dense 作为 YOLOv2 的主干网络,并利用训练样本目标前景区域标注加强芒果前景区域卷积特征的学习	提高了 YOLOv2 芒果检测模型在重叠遮挡场景下的性能,该方法检测重叠芒果果实的精确率为 90.17%,召回率为 87.63%	[48]
草莓	YOLOv3	基于 YOLOv3 实现草莓果实的识别	该模型在果实遮挡、重叠、密集等复杂环境下具有较好的鲁棒性,其 mAP 为 87.51%,模型对成熟草莓的识别准确率为 97.14%,召回率为 94.46%,对未成熟草莓的识别准确率为 96.51%,召回率为 93.61%	[49]
荔枝	YOLOv3	基于 YOLOv3 实现夜间荔枝果实的检测	该方法在高亮度、正常亮度和低亮度下的 AP 分别为 96.78%、99.57% 和 89.30%	[50]
冬枣	改进 YOLOv3	在 YOLOv3 模型的第 37、52 层后分别嵌入 SE Block 结构,增强了特征图的特征表达能力	与 YOLOv3 模型相比,该模型的召回率提升 4.36 个百分点,mAP 提升 4.78 个百分点,综合评价指标提升 2.38 个百分点,且检测速度无明显降低	[51]
柑橘	改进 YOLOv3	在 YOLOv3 的第 2 个卷积层模块中添加 3 层最大池化层	该模型对枝干遮挡果实、树叶遮挡果实、重叠果实、正常无遮挡果实的综合 F1 值为 91.90%	[52]
柑橘	改进 YOLOv3	在 YOLOv3 的骨干网络中添加 Dense Block 结构,实现了网络多层特征的复用和融合	该模型在夜间柑橘测试集上的 mAP 为 90.75%	[53]
苹果	改进 YOLOv3	将 DarkNet53 网络中的残差模块与 CSPNet 结合,利用 SPP 模块实现果实的全局和局部特征的融合,采用 Soft NMS 算法增强对重叠遮挡果实的识别能力,采用基于 Focal Loss 和 CIoU Loss 的联合损失函数对模型进行优化	该模型在保持检测精度的同时降低了网络的计算量,并提高了极小果实目标的召回率,该方法检测苹果目标的 mAP 为 96.30%,F1 值为 91.80%,在 GPU 下的平均检测速度为 27.8 t/s	[54]
菠萝	改进 YOLOv3	在 YOLOv3 的骨干网络中加入 DenseNet 和 SPP 模块	该模型在轻微遮挡的数据集上的 F1 值为 93.00%,AP 为 97.55%,随着遮挡程度的增加,F1 值和 AP 分别下降为 89.00% 和 91.47%	[55]
番茄花	改进 YOLOv3	将 YOLOv3 骨干网络的输入更改为多输入网络,增加特征信息的冗余,修改 DarkNet53 结构以适应不同尺度输入网络结构	与原始 YOLOv3 模型相比,该方法的检测精度有所提高,其检测番茄花朵的 mAP 为 82.79%	[56]
苹果	轻量化 YOLOv3	该模型由 5 个相同结构的残差块串联组成特征提取网络,在 2 个尺度上进行苹果检测,并采用深度可分离卷积替换普通卷积	该方法有效减少了模型的参数量和计算量,并提高了检测速度和精度,该模型的 F1 值为 94.57%,mAP 为 94.69%	[57]
番茄	改进 YOLOv3-Tiny	采用含有残差模块的 Darknet-20 主干网络实现特征提取和检测速度的提升,并借鉴 YOLOv3-Tiny 检测结构,融合多尺度信息	该方法提高了特征提取过程的速度,其检测番茄果实的 mAP 为 97.13%,精确率为 96.36%,召回率为 96.03%,检测时间为 7.72 ms	[58]
苹果、油桃、黄杏、李子	YOLOv4	基于 YOLOv4 实现苹果、油桃、黄杏和李子的检测	检测苹果、油桃、黄杏和李子的 mAP 分别为 85.72%、86.54%、85.21% 和 86.13%	[59]

续表 2

果实种类	方法名称	技术原理	技术特点与性能指标	文献序号
香蕉	YOLOv4	基于 YOLOv4 实现香蕉串的检测	模型检测香蕉串的 AP 为 99.55% ,单幅图像的平均检测时间为 44.96 ms	[60]
番茄	改进 YOLOv4	在 YOLOv4 的检测框内通过 HSV 方法对番茄的红色区域进行分割,并将分割面积在检测框中达到一定占比的番茄作为目标输出	当分割面积在检测框中的占比为 16% 时该模型的性能最好,其正确率为 94.77%	[61]
番茄	改进 YOLOv4	融合残差神经网络构建了一个新的骨干网络 R-CSPDarkNet53,建立前后层之间的跳转连接,以确保低维小目标特征的不完全丢失,将原 SPP 网络中的深度可分离卷积模型替换为最大池,实现了特征信息的重用和多尺度融合	该模型具有较好的鲁棒性,在自然场景中的番茄检测精确率和召回率分别为 88.00% 和 89.00% ,其 AP 为 94.44%	[62]
柑橘	改进 YOLOv4	以 CSPResNest50 作为骨干网络以适应尺寸较小的春见柑橘果实,采用递归特征金字塔网络进行特征递归融合	该模型提高了具有密集遮挡的小尺寸目标的检测精度,其 mAP 为 94.60% ,平均检测速度为 51 f/s	[63]
苹果	改进 YOLOv4	采用 SE block 和 NL block 两种视觉注意机制从通道和非局部两个方面整合高级特征,提高网络对背景与果实的特征捕捉能力,由不同尺寸的特征图实现不同大小幼果的坐标和类别计算	该模型可在增加少量参数的基础上显著提升网络对苹果幼果的感知能力,其在测试集上的 AP 为 96.90%	[64]
苹果	改进 YOLOv4	将非局部注意力模块和卷积块注意力模型添加到 YOLOv4 模型的基线中以提高模型对低质量图像的检测效果	该模型检测测试集中的高光/阴影、模糊、严重遮挡的图像的 AP 分别为 98.00%、96.20% 和 97.00%	[65]
苹果	改进 YOLOv4	将 CBAM 引入 YOLOv4 基线,引入自适应层和密集连接,并优化 YOLOv4 的回归损失函数	该模型检测早期、中期和收获期的“Envy”品种苹果的精确率分别为 86.20%、87.50% 和 92.60%。然而对于冠层较密的“Scifresh”品种,其检测效果尚需进一步提高	[66]
蓝莓	改进 YOLOv4-Tiny	融合 CBAM 和 YOLOv4-Tiny 实现不同成熟度的蓝莓识别	该模型检测蓝莓果实的 AP 为 96.24% ,平均检测时间为 5.72 ms,网络结构占用内存 24.20 MB	[67]
番茄	改进 YOLOv4-Tiny	将 CBAM 模块集成到 YOLOv4-Tiny 的骨干网络,使用 Mish 激活函数确保深层卷积中提取特征的准确性,并使用密集连接的卷积网络来加强全局特征融合	该模型在复杂的自然环境中具有较好的稳定性,其在 Nvidia GTX 2060 GPU 上的检测速度为 111 f/s,其 mAP 为 97.90%	[68]
草莓	改进 YOLOv4-Tiny	减少骨干网络中 CSPNet 模块的数量和精简 CSPNet 的网络结构	与原始 YOLOv4-Tiny 模型相比,改进后的模型 mAP 仅降低了 0.62% ,但其检测速度提高了 25.93%	[69]
苹果花朵	通道剪枝 YOLOv4	利用通道剪枝精简 YOLOv4 模型	修剪后的 YOLOv4 苹果花朵检测模型的参数数量减少了 96.74% ,其模型占用内存为 12.46 MB,推理时间减少了 39.47% ,其 mAP 为 97.31% ,仅比修剪前的模型低 0.24%	[70]
柑橘	YOLOv5	基于 YOLOv5s 实现自然场景中的柑橘检测	该方法 mAP 为 95.40% ,模型占用内存为 13.70 MB,单幅图像的检测时间为 70 ms	[71]
苹果花朵	YOLOv5	基于 YOLOv5s 实现自然场景中的苹果花朵检测	该方法的精确率为 87.70% ,召回率为 94.00% ,mAP 为 97.20% ,模型占用内存为 14.09 MB,检测速度为 60.17 f/s	[72]
油茶果	YOLOv5	基于 YOLOv5s 实现自然场景中的油茶果识别	该方法的精确率为 90.73% ,召回率为 98.38% ,mAP 为 98.71% ,模型占用内存为 14.08 MB,单幅图像的检测时间为 12.7 ms	[73]
番茄	改进 YOLOv5	采用基于交并比的 CIoU 目标位置损失函数,利用函数计算选择最佳锚框尺寸	该模型对夜间环境下的番茄绿色果实和红色果实的 AP 分别为 96.20% 和 97.60%	[74]
樱桃	改进 YOLOv5	利用离线和在线数据增强的方式提高樱桃的检测精度,骨干网络中增加 Transformer 模块,利用可以进行双向加权融合的 BiFPN 对 Neck 结构进行改进,并在 Head 结构中增加浅层下采样的 P2 模块	该模型检测樱桃果实的精确率为 97.60% ,召回率为 89.90% ,mAP 为 95.20%	[75]
苹果	改进 YOLOv5m	改进原 YOLOv5m 骨干网络中的 BottleneckCSP 模块,将 SE 模块嵌入到模型的骨干网络,改进原 YOLOv5m 架构中输入中等尺寸目标检测层的特征图的跨接融合方式,并改进网络的初始锚框尺寸	该模型的召回率、准确率、mAP 和 F1 值分别为 85.90%、81.00%、80.70% 和 83.40% ,单幅图像的平均识别时间为 25 ms	[76]

YOLO 算法存在以下局限性:YOLO 的每个网格只能有两个预测框并预测一个类别,因此其对小目标的检测性能较差;YOLO 从大量的训练数据中提取目标的特征,若测试数据中目标的长宽比与训练数据有较大不同,则网络检测效果欠佳;损失函数无差别地处理大小边界框的误差,大边界框的误差和小边界框的误差对交并比(Intersection over union, IoU)的影响差异较大。

YOLOv2^[80]在 YOLO 的基础上对损失函数、骨干网络等进行改进,同时引入了锚点框、批量归一化、高分辨率分类器等结构,YOLOv2 在检测速度、准确率等方面均有较大提升。基于 YOLOv2 可实现绿色芒果的检测,在图像中包含的芒果数量较多或者光线较暗的情况下,其检测效果不理想^[47]。对于存在遮挡或重叠的目标,YOLOv2 的检测效果有待提高。采用带密集连接的 Tiny-yolo-dense 作为 YOLOv2 的主干网络,可提高 YOLOv2 芒果检测模型在重叠遮挡场景下的性能,然而该方法需要复杂的前景区域标注过程^[48]。

YOLOv3^[81]引入了特征金字塔网络(Feature pyramid network, FPN)^[82]和 ResNet^[83]结构,同时采用 DarkNet53 作为骨干网络,增加了多尺度预测结构,使网络的检测精度得到了提升。利用 YOLOv3 可实现复杂环境中草莓和荔枝的检测,然而该方法的效果受到光照强度的影响。由于果实生长的自然环境较为复杂,枝叶和果实间的遮挡、复杂的光照情况等为果实目标的准确检测带来困难,针对复杂场景中的果实目标检测以及小目标果实的检测,YOLOv3 的检测性能尚需提高,对 YOLOv3 模型进行改进可以提高模型在复杂环境下的识别效果。利用 YOLOv3-SE 网络模型可实现枝叶遮挡、果实密集重叠等复杂场景下的冬枣识别^[51]。为实现柑橘采摘机器人的识别定位,在 YOLOv3 模型中增加最大池化层可增强模型对采摘场景的理解,提高柑橘目标识别的准确率^[52]。如文献[53-56],利用 DenseNet、CSPNet 和 SPP 模块等对 YOLOv3 的骨干网络进行改进,可提高其在夜间环境、遮挡目标和小目标等复杂场景下的检测性能。由于 DarkNet53 的网络层数过多,导致网络的运算量较大,检测速度较慢,对于一些场景较为简单的果实目标检测任务,可以通过简化网络层数以减小模型复杂度并提高检测速度。如文献[57-58],通过精简 YOLOv3 的骨干网络,可简化目标检测的特征图尺度,实现模型的轻量化,且模型在检测速度和准确率方面均有显著提高。

YOLOv4 的输入端引入了 Mosaic 数据增强操作,其骨干网络在 DarkNet53 的基础上融合了

CSPNet,采用 SPP 和 FPN+路径聚合网络(Path aggregation network, PAN)作为瓶颈结构,并采用 CIOU_loss 作为预测端的损失函数^[84],与 YOLOv3 相比,其检测速度和准确率都有了较大提升。利用 YOLOv4 可实现自然场景下苹果、油桃、黄杏、李子及香蕉串检测^[59-60]。为提高 YOLOv4 网络在复杂场景下的果实识别效果,一些学者将颜色空间模型、残差神经网络、递归特征金字塔和视觉注意机制等与 YOLOv4 网络相结合,在一定程度上提高了模型的检测性能。基于 YOLOv4+HSV 的成熟番茄识别方法可解决遮挡和光照不均引起的番茄目标误识别的问题^[61]。一种融合残差神经网络和 YOLOv4 的番茄检测方法可解决自然环境中光照变化、背景干扰和叶片遮挡等因素对番茄采摘机器人的检测精度的影响^[62]。一种基于特征递归融合 YOLOv4 网络的 FR-YOLOv4 检测模型可实现自然场景下密集分布的小目标春见柑橘的检测和计数^[63]。如文献[64-66],将 SE 模块、CBAM 视觉注意机制等与 YOLOv4 模型相结合,可实现低质量苹果幼果、不同颜色和品种的苹果果实检测。

YOLOv4-Tiny 对 YOLOv4 的网络结构进行了精简,YOLOv4-Tiny 采用 CSPDarkNet53-Tiny 作为骨干网络,并将 YOLOv4 中的 Mish 激活函数修改为 Leaky_ReLU 激活函数。YOLOv4-Tiny 模型的参数量更少,网络结构更简单,且检测速度更快^[67]。利用 CBAM 视觉注意机制对 YOLOv4-Tiny 模型进行改进,可实现复杂环境下番茄和蓝莓果实的快速识别,并有效解决遮挡、逆光和小目标识别准确率低的问题^[67-68]。通过减少 YOLOv4-Tiny 草莓检测模型的骨干网络中 CSPNet 模块的数量和精简 CSPNet 的网络结构,可进一步简化模型,提高模型的检测速度^[69]。另一种简化 YOLOv4 模型的方法是进行通道剪枝,其本质是通过识别网络的通道来消除不重要的通道及其相关的输入和输出关系^[85],简化后的模型可以减少需要存储的参数数量,并且具有较低的硬件要求,使其更易于部署在嵌入式设备和移动终端等小型计算平台上^[86]。利用通道剪枝的 YOLOv4 可实现自然场景下的苹果花朵实时准确检测,剪枝后模型的检测精度基本不变,但模型的参数量和尺寸得到大幅压缩,且检测速度有较高的提升^[70]。

YOLOv5 的输入端引入了自适应锚框计算,以适应不同尺寸的目标,YOLOv5 的骨干网络中引入了切片操作,并将 CSPNet 同时应用于骨干网络和瓶颈网络,YOLOv5 的输出端采用了 GIoU_loss 损失函数。YOLOv5 包含 5 种体系结构,分别为 YOLOv5-nano、YOLOv5s、YOLOv5m、YOLOv5l 和 YOLOv5x,其

主要区别在于特征提取模块和卷积核在网络特定位置的数量不同。与 YOLOv4 相比, YOLOv5 的检测速度和精度都有较大的性能提升。如文献[71 - 73], 利用 YOLOv5s 目标检测模型可以实现自然场景下的柑橘、苹果花朵、油茶果的准确快速检测, 模型具有较好的鲁棒性, 且模型的尺寸较小, 适用于模型迁移。通过优化 YOLOv5 的损失函数, 可有效提高模型对于遮挡番茄目标的识别准确率^[74]。对 YOLOv5 模型的骨干网络进行简化或引入视觉注意机制模型, 可在一定程度上减小模型的尺寸并提高其目标识别的效果。利用 Transformer 模块对 YOLOv5 的骨干网络进行改进, 并利用 BiFPN 改进其 Neck 结构, 可提高樱桃果实的识别准确率^[75]。利用 Bottleneck 模块对 YOLOv5m 的骨干网络进行改进, 并引入 SE 视觉注意机制模块, 可提高模型对苹果目标的检测速度和精度^[76]。

2.2.3 无锚框目标检测算法

自从 RPN 提出以来, 基于锚框的目标检测算法已经成为目标检测模型的主流, 且取得了较好的检测效果。然而, 基于锚框的检测器存在以下的缺点和局限性: 为实现高召回率, 基于锚框的检测器需要设计各种尺度和形状的锚框, 以覆盖图像中不同尺度和形状的物体, 这些冗余的锚框在训练时大多会被标记为负样本, 导致训练中的正负样本不均衡; 锚框的尺寸、长宽比、数量等参数设置决定了检测器的性能, 因此锚框的设计过程较为复杂, 且训练结果可能受到人为经验的影响; 大量的锚框会增加整个检测过程的计算成本。

为了克服基于锚框的检测器的缺点, 基于无锚框检测器的目标检测算法逐渐兴起, 无锚框检测器不需预先设定锚框, 直接对图像进行目标检测。YOLOv1 是目标检测领域最早的无锚框模型, 它将目标检测视为一个空间分离的边界框和相关概率回归问题, 可直接从图像中预测边界框的位置和分类概率。该方法运算速度较快, 但其召回率较低, 且检测精度不理想。如图 5 所示, 无锚框检测可大致分为两类: 基于锚点检测和基于关键点检测^[87]。锚点检测器通过将真实框编码为锚点, 锚点是特征图上的像素点, 其位置与特征相关联, 其代表算法有 CenterNet^[88]、FCOS^[89]、FoveaBox^[90] 等。关键点检测器通过预测包围盒中的几个关键点的位置, 如角点、中心或极值点, 将关键点解码到预测框中, 其代表算法有 CornerNet^[91]、ExtremeNet^[92] 等。

表 3 列出了无锚框目标检测算法的果实目标识别研究成果。对 FCOS 的骨干网络和损失函数进行改进实现自然场景下的苹果果实检测, 可在提高检

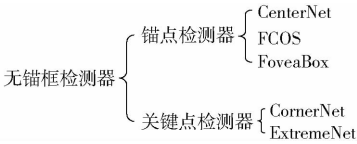


图 5 无锚框检测器分类

Fig. 5 Classification of anchor-free detector

测性能的同时降低正负样本比例失衡带来的误差^[93]。利用残差特征金字塔网络对 FCOS 的骨干网络进行改进, 可实现光照变化和阴影重叠条件下的绿色苹果识别^[94]。改进 FoveaBox 模型的骨干网络和特征提取网络可以提高模型对不同规模的绿色苹果的召回率^[95]。利用 Tiny Hourglass - 24 网络对 CenterNet 的骨干网络进行改进可实现密集场景下的多苹果目标的快速识别, 然而该方法对于重度遮挡和果实表面亮度较高的情景存在个别目标漏检的情况^[96]。利用改进的 MobileNetv3 作为 CenterNet 模型的骨干网络实现苹果目标的识别, 可以在保证检测精度的前提下减小模型的尺寸并提高其检测速度^[97]。

YOLOX^[101] 是 YOLO 的无锚框版本, 其沿用了 YOLOv4 的特征提取网络 CSPDarknet 和 YOLOv5 的 Mosaic 增强技术, 并创新了解耦检测头、无锚框和 SimOTA 标签分配策略^[102], YOLOX 模型的设计简单, 且性能更具优势。利用 YOLOX 可实现冬枣的高精度检测和计数^[98]。通过改进 YOLOX - S 的激活函数和损失函数, 并转移浅层特征, 可实现小目标猕猴桃的检测, 该方法在减小了模型参数数量的同时提高了模型的检测精度^[99]。利用加入 CBAM 的轻量级网络 ShuffleNetv2 对 YOLOX - Tiny 的骨干网络进行改进, 可实现对苹果果实的高精度和实时性检测^[100]。

2.3 基于深度学习的果实目标分割方法

2.3.1 基于深度学习的语义分割模型

语义分割是将图像的每个像素分配给预定义类别集合中的相应类别标签的任务, 其目的是对图像中的每一个像素点进行分类。用于目标识别的 CNN 网络通常由卷积层、池化层和全连接层组成。然而, 全连接层会损失目标的位置信息, 因此 CNN 无法完成目标的分割任务。全卷积网络 (Full convolutional networks, FCN)^[103] 是基于深度学习的语义分割模型的重要成果, 通过将 CNN 网络中的全连接层替换为卷积层, 得到全部由卷积层构成的 FCN, 与 CNN 相比, FCN 可同时保留目标的位置信息和语义信息, 可在像素层次上进行分类, 完成目标分割任务。

RONNEBERGER 等^[104] 提出的 U - Net 网络被

表 3 基于无锚框算法的果实目标识别研究成果

Tab.3 Research on fruit target recognition based on anchor-free algorithm

果实种类	方法	技术原理	技术特点与性能指标	文献序号
苹果	改进 FCOS	利用 DarkNet19 作为 FCOS 的骨干网络,采用联合交并比和焦点损失对损失函数进行改进	该方法提高了模型的检测性能,同时降低了正负样本比例失衡带来的误差,其精确率为 96.00%, mAP 为 96.30%	[93]
绿色苹果	改进 FCOS	利用残差特征金字塔网络替换 FCOS 的原始骨干网络,并添加双层卷积注意网络以恢复不完整绿色水果的边缘	该模型检测苹果果实的 AP 为 81.20%	[94]
绿色苹果	改进 FoveaBox	采用 EfficientNetV2-S 和加权双向特征金字塔网络作为 FoveaBox 的骨干网络和特征提取网络,采用自适应训练样本选择方法直接选择正样本和负样本	该模型可以提高不同规模绿色果实的召回率,其检测绿色苹果的 mAP 为 62.30%	[95]
苹果	改进 CenterNet	采用 Tiny Hourglass-24 网络作为 CenterNet 的骨干网络	该方法在密集场景下的 AP 为 93.63%, F1 值为 93.00%, 单幅图像的平均识别时间为 69 ms	[96]
苹果	改进 CenterNet	采用改进的 MobileNetv3 作为 CenterNet 的骨干网络	该模型的 AP 为 88.90%, 模型占用内存为 14.2 MB,检测速度为 8.1 f/s	[97]
冬枣	YOLOX	利用 YOLOX 实现了冬枣的高精度检测和计数		[98]
猕猴桃	改进 YOLOX-S	改进 YOLOX-S 的激活函数和损失函数,设计一种新的多尺度特征集成结构,消除 YOLOX 模型中用于检测大目标的特征映射	与原始的 YOLOX-S 相比,该模型检测猕猴桃图像的 AP 提高 6.52%, 模型参数量减少 44.80%, 检测速度提高 63.90%	[99]
苹果	改进 YOLOX-Tiny	采用加入 CBAM 的轻量级网络 ShuffleNetv2 作为 YOLOX 的骨干网络,在 PANet 网络中加入自适应空间特征融合模块	该模型在验证集上的 AP 为 96.76%, 精确率为 95.62%, 召回率为 93.75%, F1 值为 95.00%, 检测速度为 65 f/s。该模型在采用 TensorRT 加速的嵌入式设备 Jetson Nano 上的半精度浮点 (FP16) 精度为 26.3 f/s	[100]

广泛地应用于语义分割任务,U-Net 包括编码器和解码器结构,编码器网络利用池化层进行下采样操作,其作用是进行特征提取,解码器网络利用反卷积进行上采样操作,编码器部分与解码器部分近似对称,整体网络呈“U”形结构。在上采样的过程中,利用跳跃连接操作可以将该级的特征图与编码器结构中其对应位置的特征图进行融合,使得解码器能够获取更多高分辨率特征,有利于提高分割精度。利用 U-Net 可实现苹果目标的分割,结果表明,当果实可以通过颜色区分时,传统分割算法的效果优于 U-Net,当测试集数据与训练集相似时,U-Net 的分割效果较好^[105]。

CHEN 等^[106]结合深度卷积网络和概率图模型 (DenseCRFs),提出了 DeepLab 语义分割模型,DeepLab 以 VGG-16 为基础网络进行改进,首先将 VGG-16 的全连接层替换为卷积层,并移除原网络的最后两个池化层,使用空洞卷积进行上采样,在扩大了感受野的同时减小了参数量。DeepLabv2^[107]对 DeepLab 进行了改进,DeepLabv2 以 ResNet101 模型作为基础网络,并引入带有空洞卷积的空间金字塔结构 (Atrous spatial pyramid pooling, ASPP),在多个尺度上进行图像分割,可以处理不同尺度的目标,与 DeepLab 相比,DeepLabv2 的分割精度有所提升。

SUN 等^[108]利用 DeepLab-ResNet 实现了苹果花、梨花和桃花的目标分割,该模型的平均 F1 值为 80.90%。DeepLabv3^[109]分别利用 ResNet101 和 Xception 作为骨干网络,并将深度可分离卷积应用于 ASPP 结构,在保持性能的同时有效降低了模型的计算复杂度。利用 ResNet 和 DenseNet 结构对 DeepLabv3 的骨干网络进行改进,可实现荔枝花朵的分割^[110]。KANG 等^[111]提出了一种用于苹果语义分割的 DasNet 网络,该模型尝试分别利用 ResNet-50、ResNet-101 和轻量化网络 LW-net0 作为骨干网络,利用门控特征金字塔网络进行多级特征的融合,并采用 ASPP 增强目标的多尺度特征提取。结果表明,以 ResNet-101 为骨干网络的 DasNet 模型在语义分割和目标检测任务中表现最好,其检测苹果目标的 F1 值为 83.20%,其分割苹果目标的 F1 值为 87.60%。

2.3.2 基于深度学习的实例分割模型

实例分割是将语义标签和实例标签分配给所有像素,以分割对象实例,实例分割可以提供比语义分割更详细的图像信息,例如检测对象的位置和数量。其中 Mask-RCNN 为果实目标实例分割中最具代表性的算法。表 4 列出了利用基于深度学习的实例分割算法进行果实目标识别的研究成果。

表 4 基于深度学习的实例分割算法的果实目标识别研究成果

Tab.4 Research on fruit target recognition based on deep learning instance segmentation algorithm				
果实种类	方法	技术原理	技术特点与性能指标	文献序号
草莓	Mask R-CNN	基于 Mask R-CNN 实现自然场景中草莓的分割	其 AP 为 95.78%,召回率为 95.41%,平均联合交叉率为 89.85%。与传统方法相比,该方法对于重叠以及不同光照情况下的果实目标都具有更好的通用性和鲁棒性	[112]
苹果	Mask R-CNN	基于 Mask R-CNN 实现自然场景中苹果果实的分割	该方法的 AP 为 85.99%,F1 值为 86.00%,然而对于高度遮挡的苹果目标及小目标,该方法存在误分割的情况	[113]
葡萄	Mask R-CNN	基于 Mask R-CNN 实现自然场景中葡萄串的分割	该方法的精确率为 92.00%,召回率为 86.00%,F1 值为 88.90%	[114]
柑橘	Mask R-CNN	基于 Mask R-CNN 实现自然场景中柑橘类果实的分割	该方法对不同类型柑橘类水果的 AP 为 88.15%,召回率为 79.85%	[115]
番茄	Mask R-CNN	利用迁移学习预训练的 Mask R-CNN 实现自然场景下番茄果实的分割	当 IoU 设置为 0.5 时,该模型分割番茄果实的 AP 为 73.50%	[116]
苹果	改进 Mask R-CNN	采用 ResNet + DenseNet 作为模型的骨干网络	该方法的精确率为 97.31%,召回率为 95.70%	[117]
苹果	改进 Mask R-CNN	采用融合注意机制和特征金字塔网络的 ResNet 代替 Mask R-CNN 的原始骨干网络	该方法可实现不同颜色、相互重叠、被枝叶遮挡等情况下的苹果精确分割。其召回率、精确率和 F1 值分别为 97.10%、95.80% 和 96.40%,平均检测时间为 0.25 s/幅	[118]
番茄	改进 Mask R-CNN	修改网络的输入层实现 RGB 和深度图像的双模数据融合,并采用多类预测子网络来解耦水果和茎的像素级类别预测	该模型的精确率为 93.76%,召回率为 94.47%,与改进前相比,其准确率和召回率分别提高 11.53% 和 11.53%	[119]
番茄	改进 Mask R-CNN	修改输入层实现 RGB 图像、深度图像和红外图像的多源信息融合	该模型的准确率为 98.30%	[120]
苹果	DasNet-v2	DasNet-v2 在 FPN 网络中添加实例分割的分支,用于预测目标的边界框和掩码,对 FPN 结构进行简化,并采用“带空洞卷积的编码器-解码器”结构,提高分割的准确性	以 ResNet-101 为骨干网络的 DasNet-v2 检测苹果果实的召回率和精确率分别为 86.80% 和 88.00%,其分割精确率为 87.30%,以 ResNet-18 为骨干网络的 DasNet-v2 在模型占用内存和检测速度方面优势明显。其检测苹果果实的召回率和精确率分别为 85.00% 和 87.00%,分割苹果果实的精确率为 86.60%	[121]

HE 等^[122]提出了一种用于实例分割的 Mask R-CNN 网络,该网络通过在 Faster R-CNN 中添加一个用于预测目标掩码的并行分支实现。在非结构化环境中,Mask R-CNN 不仅能准确识别目标类别并用边界框标出目标区域,还能在像素级别上从背景中提取目标区域。Mask R-CNN 由 3 个阶段组成,首先利用 ResNet 骨干网络从输入图像中提取特征图;其次,特征图被输入到 RPN 用以生成候选区域;最后,这些候选区域被映射到原始图像中像素点对应的位置,以在共享特征映射中提取相应的目标特征,然后分别输出到全连接层和 FCN,用于目标分类和实例分割。

如文献[112-115],基于 Mask R-CNN 可实现草莓果实、苹果果实、葡萄串和柑橘类果实的目标分割,且该方法在复杂场景下具有较好的鲁棒性。对于部分农业场景中数据量不足的问题,如文献[116],利用迁移学习预训练的 Mask R-CNN 模型可提高对自然场景下番茄果实的分割精度。如文献[117-118],利用 ResNet、DenseNet 等网络对 Mask R-CNN 的原始骨干网络进行改进,可实现重

叠、遮挡等情况下苹果目标的精确分割。通过融合 RGB 图像、深度图像、红外图像等多源信息,对图像的颜色、形状空间位置等特征进行深度挖掘,可提高 Mask R-CNN 分割目标的准确率和鲁棒性。如文献[119-120],融合深度图像或红外图像等多源信息作为 Mask R-CNN 的输入,可提高模型对番茄果实的定位精度并提高模型的鲁棒性。

基于 DasNet 语义分割模型,在其 FPN 结构中添加实例分割的分支,开发一种用于进行苹果果实实例分割的 DasNet-v2 网络^[121],该模型在田间实地测试中表现出优良的性能。

2.4 基于轻量化模型的果实目标识别方法

自然场景下的果实目标识别任务存在遮挡、光照不均等一系列挑战。为提高果实目标识别任务的精度,基于深度学习的果实目标识别网络在不断加深,以适应越来越复杂的目标检测任务。然而,随着网络深度的加深,模型的参数量和计算复杂度也在不断上升,模型大小和计算成本的爆炸性增长为模型在嵌入式设备上的部署带来了新的挑战^[123-124]。目前,研究人员致力于研究轻量化的目标检测网络,

以兼顾模型的移植部署、检测速度和检测精度,以期
为果园采摘机器人的发展提供技术支持。

依据网络的结构层次,可以将网络的轻量化划分
为模型的轻量化设计和模型压缩两大类,图 6 列
出了常用轻量化模型和模型压缩的方法。

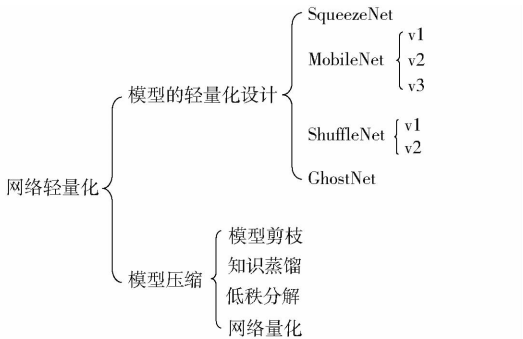


图 6 网络轻量化的常用方法

Fig. 6 Common method of network lightweight

模型压缩是在原有模型的基础上进行修改,而
轻量化模型则是在设计模型时就遵从轻量化的思
想,例如采用深度可分离卷积、分组卷积等卷积方
式,减少卷积的数量、增加网络的并行度、减少网
络的碎片化程度等^[125]。相比于模型压缩,轻量化
模型的设计能够更大程度地减小模型的参数量和计
算量,并提高模型的检测速度,因此轻量化模型是未
来目标检测算法用于嵌入式设备移植和移动端的主
要发展方向。如图 6 所示,近年来表现优秀的轻量
化网络主要有 SqueezeNet^[126]、MobileNet 系列、
ShuffleNet 系列、GhostNet^[125]等。表 5 列出了基
于轻量化模型的果实目标识别研究成果。

LANDOLA 等^[126]提出的 SqueezeNet 是最早的
轻量化模型设计,SqueezeNet 使用了新的网络架构
“Fire 模块”,整个 SqueezeNet 网络由若干“Fire
模块”的堆叠组成。“Fire 模块”由压缩层和扩展
层组成,其中压缩层仅由 1×1 的卷积核组成,扩
展层由 1×1 和 3×3 的卷积核组成,Fire 模块的
设计大大减少了模型的参数量和计算量,Squeeze
Net 的模型占用内存为 0.5 MB。

HOWARD 等^[134]基于深度可分离卷积提出了
MobileNet 轻量化模型。该模型由深度可分离卷
积与普通卷积模块堆叠组成,深度可分离卷积是
将普通卷积拆分为深度卷积和逐点卷积的操作,利
用深度可分离卷积可大大降低网络的计算量。以
MobileNet 模型分别替换 YOLOv4 和 SSD 的骨干
网络,可实现龙眼果实的检测,其中 MobileNet -
YOLOv4 模型具有更好的泛化性能^[127]。Mobile
Netv2^[135]网络中引入了线性瓶颈结构和倒残差
结构,进一步压缩了模型占用内存。利用 Mobile
Netv2 对 YOLOv3 检测模型的骨干网络进行

改进,可实现酿酒葡萄和柑橘的检测,与原始
YOLOv3 模型相比,该方法可在保证检测精度的情
况下大幅提升检测速度并压缩模型的尺寸^[128 - 129]。
MobileNetv3^[130]在 MobileNet2 的倒残差结构的基
础上加入了 SE 注意力模块和 h-swish 激活函数,并
精简了卷积层的结构,同时引入了 NAS 模块搜索和
NetAdapt 层搜索结构进行网络结构的优化,与
MobileNet2 相比,其准确率更高且减少了延迟。利
用 MobileNetv3 模型可实现香蕉、柠檬、柑橘等水
果的检测,其检测精确率和检测速度均优于 Xception
和 DenseNet 模型^[130]。利用 MobileNetv3 对 YOLOv4
检测模型的骨干网络进行改进,开发一种 YOLOv4 -
MobileNetv3 轻量化模型,可实现火龙果、密集圣
女果和苹果果实的准确快速识别,该方法在检测速
度和模型尺寸方面具有显著优势^[131 - 133]。

ZHANG 等^[137]提出了 ShuffleNet 轻量化模型,
该模型提出了逐点组卷积和通道混洗操作,利用逐
点组卷积降低模型的计算复杂度,并利用通道混洗
操作解决组卷积造成的信息交互问题。对于一定的
计算复杂度,ShuffleNet 可以保证更多的特征映射
信道,提高模型的性能。在此基础上,ShuffleNetv2^[138]
模型引入了通道分裂(Channel split)操作,在保
证模型准确率的同时进一步降低了模型复杂度,提
升了模型运行速度。

HAN 等^[125]提出了 GhostNet 轻量化模型,其基
础模块为“Ghost Module”,Ghost 模块可以通过生
成“影子”特征图来减少卷积操作,进而减小模型
的计算量,GhostNet 的准确率高 MobileNetv3,且
其模型的计算复杂度小于 MobileNetv3。且 Ghost
Module 可用于替换许多经典目标检测网络中的普
通卷积模块,以减小模型的参数量和计算量,目前
Ghost Module 应用于简化果实目标检测模型的研
究较少,该方法可为果实目标检测模型的轻量化
提供新的研究思路。

3 存在的挑战和未来趋势展望

3.1 存在的挑战

基于深度学习的果实目标识别研究取得了一
定的研究成果,然而距离模型的实际应用仍存在以
下难点:

(1)大规模数据集的获取存在一定困难。为提
高模型的准确率,果实目标识别任务中要求采集的
数据样本数量足够大,为提高并验证模型的鲁棒性,
数据集集中的图像还应该包含有无遮挡、枝干叶片
遮挡和不同果实目标间的遮挡、不同的光照情况、
图像中包含单目标和多目标果实等情况。由于果实生长

表 5 基于轻量化模型的果实目标识别研究成果

Tab.5 Research on fruit target recognition based on lightweight model

果实种类	方法	技术原理	技术特点与性能指标	文献序号
龙眼	改进 MobileNet	采用 MobileNet 分别替代 SSD 和 YOLOv4 模型的原始骨干网络	其检测无人机拍摄的龙眼图像的 mAP 分别为 66.53% 和 89.73%,其平均检测时间分别为 128 ms 和 68 ms,模型占用内存分别为 27.3 MB 和 46.5 MB,且 MobileNet – YOLOv4 模型具有更好的泛化性能	[127]
葡萄	改进 MobileNetv2	采用 MobileNetv2 代替 YOLOv3 的原始骨干网络	在减小模型计算量的同时增大了目标检测速度。该模型在测试集上的平均 AP 为 81.20%,F1 值为 85.63%,平均每幅葡萄图像的检测时间为 6.29 ms,浮点运算量为 10.14×10^9 ,网络尺寸 44 MB,但该模型对于重叠面积超过一半的葡萄串的定位效果不理想	[128]
柑橘	改进 MobileNetv2	采用 MobileNetv2 代替 YOLOv3 的原始骨干网络,使用混合训练与迁移学习结合的预训练方式提高模型的泛化能力	该模型在测试集上的 F1 值和 AP 分别为 93.69% 和 91.13%,在 GPU 上的检测速度为 246 f/s,对单幅图像 (416 像素 $\times$ 416 像素)的推理时间为 16.9 ms,模型占用内存为 28 MB	[129]
香蕉、柠檬、柑橘等水果	MobileNetv3	利用 MobileNetv3 实现香蕉、柠檬和柑橘等水果的检测	该模型准确率为 93.64%,单幅图像检测时间为 8.4 ms	[130]
火龙果	改进 MobileNetv3	采用 MobileNetv3 代替 YOLOv4 的原始骨干网络	该模型的 AP 为 96.48%,召回率为 95.00%,模型所占内存为 2.7 MB	[131]
圣女果	改进 MobileNetv3	采用 MobileNetv3 代替 YOLOv4 的原始骨干网络,在特征金字塔结构中引入有利于小目标检测的特征层,实现细粒度检测,利用深度可分离卷积降低模型的运算量	当圣女果的重叠度为 0.50 时,该模型在测试集上的调和均值、AP 和精确率分别为 0.99、99.74% 和 99.15%	[132]
苹果	改进 MobileNetv3	采用 MobileNetv3 代替 YOLOv4 的原始骨干网络,将深度可分离卷积引入特征融合网络,在模型中引入坐标注意力机制,并将跨域迁移与域内迁移相结合	模型在保证低误检率的同时提高了检测速度,其检测苹果果实的 AP 为 92.23%,模型所占存储空间为 54.1 MB,浮点数计算量为 6.92×10^9 ,在台式计算机及嵌入式平台 Jetson AGX Xavier 上的检测速度分别为 15.34 f/s 和 15.11 f/s	[133]

的自然环境复杂多变,存在许多人为不可控因素,且果实生长具有一定的周期性,必须在特定的时间段完成图像采集任务,因此大规模数据的采集是目前果实目标识别任务的难点之一。

(2)提高模型在复杂场景下的稳定性、泛化性和鲁棒性。由于采摘机器人作业的自然环境中存在各种复杂多变的情景,用于果实目标检测的模型需要兼具较高的稳定性、泛化性和鲁棒性,才能保证其稳定的作业效果和较高的作业效率。因此,在保证识别准确率的同时提高模型在复杂场景下的表现性能是目前目标识别领域的难题。

(3)提高模型的通用性。由于田间生长的果实目标在不同的生长阶段具有不同的颜色和大小,因此研发通用性较高的果实目标识别模型有利于进行果实的生长监测和提高采摘机器人决策的准确性。

3.2 未来研究趋势展望

(1)由于数据采集和标注任务需耗费大量的时间和人力,且大规模的数据采集任务存在一定的困难,利用较少的样本数据进行模型训练并减小标签数据的数量,对于降低人工成本和提高检测的灵活性非常重要^[28]。而目前对于小规模数据集的模型、半监督和无监督模型用于果实目标检测的研究相对较少。因此,小规模数据模型和弱监督模型将是未来果实目标识别模型的发展方向。

(2)针对轻量化模型设计,部分学者已经进行相关研究并取得了一些进展,在保证识别精度的前提下,用于嵌入式设备的模型不仅要求较小的模型尺寸,还应该保证较快的检测速度以实现实时检测。因此,后续的研究应着重于提高轻量化模型在算力有限的嵌入式设备上的检测速度,开发可用于边缘设备进行实时准确检测果实目标的模型。

参 考 文 献

[1] ZHOU H, WANG X, AU W, et al. Intelligent robots for fruit harvesting: recent developments and future challenges[J]. Precision Agriculture, 2022, 23(5): 1856 – 1907.

- [2] 卢军, 桑农. 变化光照下树上柑橘目标检测与遮挡轮廓恢复技术[J]. 农业机械学报, 2014, 45(4): 76–81.
LU Jun, SANG Nong. Detection of citrus fruits within tree canopy and recovery of occlusion contour in variable illumination[J]. Transactions of the Chinese Society for Agricultural Machinery, 2014, 45(4): 76–81. (in Chinese)
- [3] 司永胜, 乔军, 刘刚, 等. 苹果采摘机器人果实识别与定位方法[J]. 农业机械学报, 2010, 41(9): 148–153.
SI Yongsheng, QIAO Jun, LIU Gang, et al. Location of fruits for apple harvesting robot[J]. Transactions of the Chinese Society for Agricultural Machinery, 2010, 41(9): 148–153. (in Chinese)
- [4] GENE-MOLA J, VILAPLANA V, ROSELL-POLO J R, et al. Multi-modal deep learning for Fuji apple detection using RGB–D cameras and their radiometric capabilities[J]. Computers and Electronics in Agriculture, 2019, 162: 689–698.
- [5] GAN H, LEE W S, ALCHANATIS V, et al. Immature green citrus fruit detection using color and thermal images[J]. Computers and Electronics in Agriculture, 2018, 152: 117–125.
- [6] BAZAME H C, MOLIN J P, ALTHOFF D, et al. Detection, classification, and mapping of coffee fruits during harvest with computer vision[J]. Computers and Electronics in Agriculture, 2021, 183: 106066.
- [7] 熊俊涛, 刘振, 汤林越, 等. 自然环境下绿色柑橘视觉检测技术研究[J]. 农业机械学报, 2018, 49(4): 45–52.
XIONG Juntao, LIU Zhen, TANG Linyue, et al. Visual detection technology of green citrus under natural environment[J]. Transactions of the Chinese Society for Agricultural Machinery, 2018, 49(4): 45–52. (in Chinese)
- [8] XIAO Y, TIAN Z, YU J, et al. A review of object detection based on deep learning[J]. Multimedia Tools and Applications, 2020, 79(33–34): 23729–23791.
- [9] TANG Y, CHEN M, WANG C, et al. Recognition and localization methods for vision-based fruit picking robots: a review[J]. Frontiers in Plant Science, 2020, 11: 510–526.
- [10] 赵永强, 饶元, 董世鹏, 等. 深度学习目标检测方法综述[J]. 中国图像图形学报, 2020, 25(4): 629–654.
ZHAO Yongqiang, RAO Yuan, DONG Shipeng, et al. Survey on deep learning object detection[J]. Journal of Image and Graphics, 2020, 25(4): 629–654. (in Chinese)
- [11] KAMILARIS A, PRENAFETA-BOLDÚ F X. Deep learning in agriculture: a survey[J]. Computers and Electronics in Agriculture, 2018, 147: 70–90.
- [12] LI Y, FENG Q, LI T, et al. Advance of target visual information acquisition technology for fresh fruit robotic harvesting: a review[J]. Agronomy, 2022, 12(6): 1336.
- [13] TIAN Y, YANG G, WANG Z, et al. Instance segmentation of apple flowers using the improved Mask R–CNN model[J]. Biosystems Engineering, 2020, 193: 264–278.
- [14] 熊俊涛, 邹湘军, 陈丽娟, 等. 基于机器视觉的自然环境中成熟荔枝识别[J]. 农业机械学报, 2011, 42(9): 162–166.
XIONG Juntao, ZOU Xiangjun, CHEN Lijuan, et al. Recognition of mature litchi in natural environment based on machine vision[J]. Transactions of the Chinese Society for Agricultural Machinery, 2011, 42(9): 162–166. (in Chinese)
- [15] 张春龙, 张焜, 张俊雄, 等. 近色背景中树上绿色苹果识别方法[J]. 农业机械学报, 2014, 45(10): 277–281.
ZHANG Chunlong, ZHANG Ji, ZHANG Junxiong, et al. Recognition of green apple in similar background[J]. Transactions of the Chinese Society for Agricultural Machinery, 2014, 45(10): 277–281. (in Chinese)
- [16] KURTULMUS F, LEE W S, VARDAR A. Green citrus detection using ‘eigenfruit’, color and circular gabor texture features under natural outdoor conditions[J]. Computers and Electronics in Agriculture, 2011, 78(2): 140–149.
- [17] 项荣, 应义斌, 蒋焕煜, 等. 基于边缘曲率分析的重叠番茄识别[J]. 农业机械学报, 2012, 43(3): 157–162.
XIANG Rong, YING Yibin, JIANG Huanyu, et al. Recognition of overlapping tomatoes based on edge curvature analysis[J]. Transactions of the Chinese Society for Agricultural Machinery, 2012, 43(3): 157–162. (in Chinese)
- [18] 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.
LIU Changyuan, LAI Nanxu, BI Xiaojun. Spherical fruit recognition and location algorithm based on depth image[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(10): 228–235. (in Chinese)
- [19] ZOU Z, SHI Z, GUO Y, et al. Object detection in 20 years: a survey[DB/OL]. (2019–05–16)[2022–08–10]. arXiv: 1905.05055v2. <https://arxiv.org/abs/1905.05055>.
- [20] WU X, SAHOO D, HOI S C H. Recent advances in deep learning for object detection[J]. Neurocomputing, 2020, 396: 39–64.
- [21] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015, 521(7553): 436–444.
- [22] ZHANG Y, XIE Q, ZHANG C. Key algorithms for segmentation of copperplate printing image based on deep learning[J]. Mobile Information Systems, 2021, 2021(6): 9940801.
- [23] CARUANA R, NICULESCU MIZIL A. An empirical comparison of supervised learning algorithms[C]// Proceedings of the 23rd International Conference on Machine Learning, Pennsylvania, 2006: 161–168.
- [24] KOIRALA A, WALSH K B, WANG Z, et al. Deep learning-method overview and review of use for fruit detection and yield estimation[J]. Computers and Electronics in Agriculture, 2019, 162: 219–234.
- [25] XIAO Y, TIAN Z, YU J, et al. A review of object detection based on deep learning[J]. Multimedia Tools and Applications, 2020, 79(33): 23729–23791.
- [26] HASAN A S M M, SOHEL F, DIEPEVEEN D, et al. A survey of deep learning techniques for weed detection from images[J]. Computers and Electronics in Agriculture, 2021, 184: 106067.
- [27] DOS SANTOS FERREIRA A, FREITAS D M, DA SILVA G G, et al. Unsupervised deep learning and semi-automatic data labeling in weed discrimination[J]. Computers and Electronics in Agriculture, 2019, 165: 104963.
- [28] KRIZHEVSKY A, SUTSKEVER I, HINTON G. ImageNet classification with deep convolutional neural networks[J]. Advances in Neural Information Processing Systems, 2017, 60(6): 84–90.

- [29] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, 2014: 580–587.
- [30] HE K, ZHANG X, REN S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37: 1904–1916.
- [31] GIRSHICK R B. Fast R–CNN[C]//Proceedings of the IEEE International Conference on Computer Vision, Santiago, 2015: 1440–1448.
- [32] REN S, HE K, GIRSHICK R, et al. Faster R–CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137–1149.
- [33] WANG Z, UNDERWOOD J, WALSH K B. Machine vision assessment of mango orchard flowering[J]. Computers and Electronics in Agriculture, 2018, 151: 501–511.
- [34] 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.
BI Song, GAO Feng, CHEN Junwen, et al. Detection method of citrus based on deep convolution neural network[J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(5): 181–186. (in Chinese)
- [35] 穆龙涛, 高宗斌, 崔永杰, 等. 基于改进 AlexNet 的广域复杂环境下遮挡猕猴桃目标识别[J]. 农业机械学报, 2019, 50(10): 24–34.
MU Longtao, GAO Zongbin, CUI Yongjie, et al. Kiwifruit detection of far-view and occluded fruit based on improved AlexNet[J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(10): 24–34. (in Chinese)
- [36] 王铁伟, 赵瑶, 孙宇馨, 等. 基于数据平衡深度学习的不同成熟度冬枣识别[J]. 农业机械学报, 2020, 51(增刊1): 457–463, 492.
WANG Tiwei, ZHAO Yao, SUN Yuxin, et al. Recognition approach based on data-balanced Faster R–CNN for winter jujube with different levels of maturity[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(Supp.1): 457–463, 492. (in Chinese)
- [37] 樊湘鹏, 许燕, 周建平, 等. 基于改进卷积神经网络的在体青皮核桃检测方法[J]. 农业机械学报, 2021, 52(9): 149–155, 114.
FAN Xiangpeng, XU Yan, ZHOU Jianping, et al. Green walnut detection method based on improved convolutional neural network[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(9): 149–155, 114. (in Chinese)
- [38] ZHANG J, KARKEE M, ZHANG Q, et al. Multi-class object detection using Faster R–CNN and estimation of shaking locations for automated shake-and-catch apple harvesting[J]. Computers and Electronics in Agriculture, 2020, 173: 105384.
- [39] 闫建伟, 赵源, 张乐伟, 等. 改进 Faster R–CNN 自然环境下识别刺梨果实[J]. 农业工程学报, 2019, 35(18): 143–150.
YAN Jianwei, ZHAO Yuan, ZHANG Lewei, et al. Recognition of *Rosa roxbunghii* in natural environment based on improved Faster R–CNN[J]. Transactions of the CSAE, 2019, 35(18): 143–150. (in Chinese)
- [40] TU S, PANG J, LIU H, et al. Passion fruit detection and counting based on multiple scale Faster R–CNN using RGB–D images[J]. Precision Agriculture, 2020, 21(5): 1072–1091.
- [41] DAI J, LI Y, HE K, et al. R–FCN: object detection via region-based fully convolutional networks[J]. Advances in Neural Information Processing Systems, 2016, 29: 379–387.
- [42] 王丹丹, 何东健. 基于 R–FCN 深度卷积神经网络的机器人疏果前苹果目标的识别[J]. 农业工程学报, 2019, 35(3): 156–163.
WANG Dandan, HE Dongjian. Recognition of apple targets before fruits thinning by robot based on R–FCN deep convolution neural network[J]. Transactions of the CSAE, 2019, 35(3): 156–163. (in Chinese)
- [43] 彭红星, 黄博, 邵园园, 等. 自然环境下多类水果采摘目标识别的通用改进 SSD 模型[J]. 农业工程学报, 2018, 34(16): 155–162.
PENG Hongxing, HUANG Bo, SHAO Yuanyuan, et al. General improved SSD model for picking object recognition of multiple fruits in natural environment[J]. Transactions of the CSAE, 2018, 34(16): 155–162. (in Chinese)
- [44] 王昱潭, 薛君蕊. 改进 SSD 的灵武长枣图像轻量化目标检测方法[J]. 农业工程学报, 2021, 37(19): 173–182.
WANG Yutan, XUE Junrui. Lightweight object detection method for Lingwu long jujube images based on improved SSD[J]. Transactions of the CSAE, 2021, 37(19): 173–182. (in Chinese)
- [45] 彭红星, 李荆, 徐慧明, 等. 基于多重特征增强与特征融合 SSD 的荔枝检测[J]. 农业工程学报, 2022, 38(4): 169–177.
PENG Hongxing, LI Jing, XU Huiming, et al. Litchi detection based on multiple feature enhancement and feature fusion SSD[J]. Transactions of the CSAE, 2022, 38(4): 169–177. (in Chinese)
- [46] 肖德琴, 蔡家豪, 林思聪, 等. 基于 IFSSD 卷积神经网络的柚子采摘目标检测模型[J]. 农业机械学报, 2020, 51(5): 28–35, 97.
XIAO Deqin, CAI Jiahao, LIN Sicong, et al. Grapefruit detection model based on IFSSD convolution network[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(5): 28–35, 97. (in Chinese)
- [47] 熊俊涛, 刘振, 林睿, 等. 自然环境下树上绿色芒果的无人机视觉检测技术[J]. 农业机械学报, 2018, 49(11): 23–29.
XIONG Juntao, LIU Zhen, LIN Rui, et al. Unmanned aerial vehicle vision detection technology of green mango on tree in natural environment[J]. Transactions of the Chinese Society for Agricultural Machinery, 2018, 49(11): 23–29. (in Chinese)
- [48] 薛月菊, 黄宁, 涂淑琴, 等. 未成熟芒果的改进 YOLOv2 识别方法[J]. 农业工程学报, 2018, 34(7): 173–179.
XUE Yueju, HUANG Ning, TU Shuqin, et al. Immature mango detection based on improved YOLOv2[J]. Transactions of the CSAE, 2018, 34(7): 173–179. (in Chinese)
- [49] 刘小刚, 范诚, 李加念, 等. 基于卷积神经网络的草莓识别方法[J]. 农业机械学报, 2020, 51(2): 237–244.

- LIU Xiaogang, FAN Cheng, LI Jianian, et al. Identification method of strawberry based on convolutional neural network[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(2): 237–244. (in Chinese)
- [50] LIANG C, XIONG J, ZHENG Z, et al. A visual detection method for nighttime litchi fruits and fruiting stems[J]. Computers and Electronics in Agriculture, 2020, 169: 105192.
- [51] 刘天真, 滕桂法, 苑迎春, 等. 基于改进 YOLOv3 的自然场景下冬枣果实识别方法[J]. 农业机械学报, 2021, 52(5): 17–25.
- LIU Tianzhen, TENG Guifa, YUAN Yingchun, et al. Winter jujube fruit recognition method based on improved YOLOv3 under natural scene[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(5): 17–25. (in Chinese)
- [52] 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, 72.
- YANG Changhui, LIU Yanping, WANG Yi, et al. Research and experiment on recognition and location system for citrus picking robot in natural environment[J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(12): 14–22, 72. (in Chinese)
- [53] 熊俊涛, 郑镇辉, 梁嘉恩, 等. 基于改进 YOLOv3 网络的夜间环境柑橘识别方法[J]. 农业机械学报, 2020, 51(4): 199–206.
- XIONG Juntao, ZHENG Zhenhui, LIANG Jiaen, et al. Citrus detection method in night environment based on improved YOLOv3 network[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(4): 199–206. (in Chinese)
- [54] 赵辉, 乔艳军, 王红君, 等. 基于改进 YOLOv3 的果园复杂环境下苹果果实识别[J]. 农业工程学报, 2021, 37(16): 127–135.
- ZHAO Hui, QIAO Yanjun, WANG Hongjun, et al. Apple fruit recognition in complex orchard environment based on improved YOLOv3[J]. Transactions of the CSAE, 2021, 37(16): 127–135. (in Chinese)
- [55] LIU T, NIE X, WU J, et al. Pineapple (*Ananas comosus*) fruit detection and localization in natural environment based on binocular stereo vision and improved YOLOv3 model[J/OL]. Precision Agriculture, 2022. <https://doi.org/10.1007/s11119-022-09935-x>.
- [56] 赵春江, 文朝武, 林森, 等. 基于级联卷积神经网络的番茄花期识别检测方法[J]. 农业工程学报, 2020, 36(24): 143–152.
- ZHAO Chunjiang, WEN Chaowu, LIN Sen, et al. Tomato florescence recognition and detection method based on cascaded neural network[J]. Transactions of the CSAE, 2020, 36(24): 143–152. (in Chinese)
- [57] 武星, 齐泽宇, 王龙军, 等. 基于轻量化 YOLOv3 卷积神经网络的苹果检测方法[J]. 农业机械学报, 2020, 51(8): 17–25.
- WU Xing, QI Zeyu, WANG Longjun, et al. Apple detection method based on light-YOLOv3 convolutional neural network[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(8): 17–25. (in Chinese)
- [58] 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.
- LIU Fang, LIU Yukun, LIN Sen, et al. Fast recognition method for tomatoes under complex environments based on improved YOLO[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(6): 229–237. (in Chinese)
- [59] 张晴晖, 孔德肖, 李俊菽, 等. 基于逆运动学降维求解与 YOLOv4 的果实采摘系统研究[J]. 农业机械学报, 2021, 52(9): 15–23.
- ZHANG Qinghui, KONG Dexiao, LI Junqiu, et al. Design of fruit picking system based on inverse kinematics dimension reduction and YOLOv4[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(9): 15–23. (in Chinese)
- [60] FU L, WU F, ZOU X, et al. Fast detection of banana bunches and stalks in the natural environment based on deep learning[J]. Computers and Electronics in Agriculture, 2022, 194: 106800.
- [61] 李天华, 孙萌, 丁小明, 等. 基于 YOLOv4 + HSV 的成熟期番茄识别方法[J]. 农业工程学报, 2021, 37(21): 183–190.
- LI Tianhua, SUN Meng, DING Xiaoming, et al. Tomato recognition method at the ripening stage based on YOLOv4 and HSV[J]. Transactions of the CSAE, 2021, 37(21): 183–190. (in Chinese)
- [62] ZHENG T, JIANG M, LI Y, et al. Research on tomato detection in natural environment based on RC-YOLOv4[J]. Computers and Electronics in Agriculture, 2022, 198: 07029.
- [63] 易诗, 李俊杰, 张鹏, 等. 基于特征递归融合 YOLOv4 网络模型的春见柑橘检测与计数[J]. 农业工程学报, 2021, 37(18): 161–169.
- YI Shi, LI Junjie, ZHANG Peng, et al. Detecting and counting of spring-see citrus using YOLOv4 network model and recursive fusion of features[J]. Transactions of the CSAE, 2021, 37(18): 161–169. (in Chinese)
- [64] 宋怀波, 江梅, 王云飞, 等. 融合卷积神经网络与视觉注意机制的苹果幼果高效检测方法[J]. 农业工程学报, 2021, 37(9): 297–303.
- SONG Huaibo, JIANG Mei, WANG Yunfei, et al. Efficient detection method for young apples based on the fusion of convolutional neural network and visual attention mechanism[J]. Transactions of the CSAE, 2021, 37(9): 297–303. (in Chinese)
- [65] JIANG M, SONG L, WANG Y, et al. Fusion of the YOLOv4 network model and visual attention mechanism to detect low-quality young apples in a complex environment[J]. Precision Agriculture, 2022, 23(2): 559–577.

- [66] LU S, CHEN W, ZHANG X, et al. Canopy-attention – YOLOv4-based immature/mature apple fruit detection on dense-foliage tree architectures for early crop load estimation[J]. Computers and Electronics in Agriculture, 2022, 193: 106696.
- [67] 王立舒, 秦铭霞, 雷洁雅, 等. 基于改进 YOLOv4 – Tiny 的蓝莓成熟度识别方法[J]. 农业工程学报, 2021, 37(18): 170 – 178.
WANG Lishu, QIN Mingxia, LEI Jieya, et al. Blueberry maturity recognition method based on improved YOLOv4 – Tiny[J]. Transactions of the CSAE, 2021, 37(18): 170 – 178. (in Chinese)
- [68] 杨坚, 钱振, 张燕军, 等. 采用改进 YOLOv4 – tiny 的复杂环境下番茄实时识别[J]. 农业工程学报, 2022, 38(9): 215 – 221.
YANG Jian, QIAN Zhen, ZHANG Yanjun, et al. Real-time recognition of tomatoes in complex environments based on improved YOLOv4 – tiny[J]. Transactions of the CSAE, 2022, 38(9): 215 – 221. (in Chinese)
- [69] ZHANG Y, YU J, CHEN Y, et al. Real-time strawberry detection using deep neural networks on embedded system (rtsd-net): an edge AI application[J]. Computers and Electronics in Agriculture, 2022, 192: 106586.
- [70] WU D, LV S, JIANG M, et al. Using channel pruning-based YOLOv4 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.
- [71] CHEN S, XIONG J, JIAO J, et al. Citrus fruits maturity detection in natural environments based on convolutional neural networks and visual saliency map[J]. Precision Agriculture, 2022, 23(5): 1515 – 1531.
- [72] 尚钰莹, 张倩如, 宋怀波. 基于 YOLOv5s 的深度学习在自然场景苹果花朵检测中的应用[J]. 农业工程学报, 2022, 38(9): 222 – 229.
SHANG Yuying, ZHANG Qianru, SONG Huaibo. Application of deep learning using YOLOv5s to apple flower detection in natural scenes[J]. Transactions of the CSAE, 2022, 38(9): 222 – 229. (in Chinese)
- [73] 宋怀波, 王亚男, 王云飞, 等. 基于 YOLO v5s 的自然场景油茶果识别方法[J]. 农业机械学报, 2022, 53(7): 234 – 242.
SONG Huaibo, WANG Ya'nan, WANG Yunfei, et al. *Camellia oleifera* fruit detection in natural scene based on YOLO v5s [J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(7): 234 – 242. (in Chinese)
- [74] 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.
HE Bin, ZHANG Yibo, GONG Jianlin, et al. Fast recognition of tomato fruit in greenhouse at night based on improved YOLO v5[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(5): 201 – 208. (in Chinese)
- [75] 张志远, 罗铭毅, 郭树欣, 等. 基于改进 YOLOv5 的自然环境下樱桃果实识别方法[J]. 农业机械学报, 2022, 53(增刊 1): 232 – 240.
ZHANG Zhiyuan, LUO Mingyi, GUO Shuxin, et al. Cherry fruit detection method in natural scene based on improved YOLOv5[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(Sup. 1): 232 – 240. (in Chinese)
- [76] 闫彬, 樊攀, 王美茸, 等. 基于改进 YOLOv5m 的采摘机器人苹果采摘方式实时识别[J]. 农业机械学报, 2022, 53(增刊 1): 28 – 38, 59.
YAN Bin, FAN Pan, WANG Meirong, et al. Real-time apple picking pattern recognition for picking robot based on improved YOLOv5m[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(Sup. 1): 28 – 38, 59. (in Chinese)
- [77] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[C] // Proceedings of the European Conference on Computer Vision, Amsterdam, 2016: 21 – 37.
- [78] LI Z, ZHOU F. FSSD: feature fusion single shot multibox detector[DB/OL]. (2018 – 05 – 17) [2022 – 08 – 10]. arXiv: 1712.00960v1. <https://doi.org/10.48550/arXiv.1712.00960>.
- [79] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, 2016: 779 – 788.
- [80] REDMON J, FARHADI A. YOLO9000: better, faster, stronger[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, 2017: 6517 – 6525.
- [81] REDMON J, FARHADI A. YOLOv3: an incremental improvement[DB/OL]. (2018 – 04 – 08) [2022 – 08 – 10]. arXiv: 1804.02767v1. <https://doi.org/10.48550/arXiv.1804.02767>.
- [82] LIN T Y, DOLLÁR P, GIRSHICK R B, et al. Feature pyramid networks for object detection[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, 2017: 936 – 944.
- [83] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, 2016: 770 – 778.
- [84] BOCHKOVSKIY A, WANG C, LIAO H. YOLOv4: optimal speed and accuracy of object detection[DB/OL]. (2020 – 04 – 23) [2022 – 08 – 10]. arXiv:2004.10934v1. <https://doi.org/10.48550/arXiv.2004.10934>.
- [85] MOLCHANOV P, MALLYA A, TYREE S, et al. Importance estimation for neural network pruning[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Long Beach, 2019: 11256 – 11264.
- [86] LIU Z, LI J, SHEN Z, et al. Learning efficient convolutional networks through network slimming[C] // Proceedings of the IEEE International Conference on Computer Vision, Venice, 2017: 2755 – 2763.
- [87] ZHANG T, JIN B, JIA W. An anchor-free object detector based on soften optimized bi-directional FPN[J]. Computer Vision and Image Understanding, 2022, 218: 103410.
- [88] ZHOU X, WANG D, KRÄHENBÜHL P. Objects as points[DB/OL]. (2019 – 04 – 25) [2022 – 08 – 10]. arXiv:1904.07850v2. <https://arxiv.org/abs/1904.07850>.
- [89] TIAN Z, SHEN C, CHEN H, et al. FCOS: fully convolutional one-stage object detection[C] // Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, 2019: 9627 – 9636.
- [90] KONG T, SUN F, LIU H, et al. FoveaBox: beyond anchor-based object detector[J]. IEEE Transactions on Image Processing, 2020, 29: 7389 – 7398.

- [91] LAW H, DENG J. CornerNet: detecting objects as paired keypoints[J]. International Journal of Computer Vision, 2019, 128: 642 – 656.
- [92] ZHOU X, ZHUO J, KRÄHENBÜHL P. Bottom-up object detection by grouping extreme and center points[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, 2019: 850 – 859.
- [93] 龙燕, 李南南, 高研, 等. 基于改进 FCOS 网络的自然环境下苹果检测[J]. 农业工程学报, 2021, 37(12): 307 – 313. LONG Yan, LI Nannan, GAO Yan, et al. Apple fruit detection under natural condition using improved FCOS network[J]. Transactions of the CSAE, 2021, 37(12): 307 – 313. (in Chinese)
- [94] LIU M, JIA W, WANG Z, et al. An accurate detection and segmentation model of obscured green fruits[J]. Computers and Electronics in Agriculture, 2022, 197: 106984.
- [95] JIA W, WANG Z, ZHANG Z, et al. A fast and efficient green apple object detection model based on Foveabox[J]. Journal of King Saud University-Computer and Information Sciences, 2022, 34(8, Part A): 5156 – 5169.
- [96] 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. YANG Fuzeng, LEI Xiaoyan, LIU Zhijie, et al. Fast recognition method for multiple apple targets in dense scenes based on CenterNet[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(2): 265 – 273. (in Chinese)
- [97] XIA X, SUN Q, SHI X, et al. Apple detection model based on lightweight anchor-free deep convolutional neural network[J]. Smart Agriculture, 2020, 2(1): 99 – 110.
- [98] ZHANG Y, ZHANG W, YU J, et al. Complete and accurate holly fruits counting using YOLOX object detection[J]. Computers and Electronics in Agriculture, 2022, 198: 107062.
- [99] ZHOU J, HU W, ZOU A, et al. Lightweight detection algorithm of kiwifruit based on improved YOLOX-S[J]. Agriculture, 2022, 12(7): 993.
- [100] JI W, PAN Y, XU B, et al. A real-time apple targets detection method for picking robot based on ShufflenetV2-YOLOX[J]. Agriculture, 2022, 12(6): 856.
- [101] GE Z, LIU S, WANG F, et al. YOLOX: exceeding YOLO series in 2021[DB/OL]. (2021-08-06)[2022-08-10]. arXiv:2107.08430v2. <https://doi.org/10.48550/arXiv.2107.08430>.
- [102] ZHANG J, KE S. Improved YOLOX fire scenario detection method[J]. Wireless Communications and Mobile Computing, 2022, 2022(21): 9666265.
- [103] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, 2015: 3431 – 3440.
- [104] RONNEBERGER O, FISCHER P, BROX T. U-net: convolutional networks for biomedical image segmentation[M]. Springer, 2015: 234 – 241.
- [105] HÄNNI N, ROY P, ISLER V. A comparative study of fruit detection and counting methods for yield mapping in apple orchards[J]. Journal of Field Robotics, 2020, 37(2): 263 – 282.
- [106] CHEN L C, PAPANDREOU G, KOKKINOS I, et al. Semantic image segmentation with deep convolutional nets and fully connected crfs[DB/OL]. (2016-06-07)[2022-08-10]. arXiv:1412.7062v4. <https://doi.org/10.48550/arXiv.1412.7062>.
- [107] CHEN L, PAPANDREOU G, KOKKINOS I, et al. Deeplab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 40(4): 834 – 848.
- [108] SUN K, WANG X, LIU S, et al. Apple, peach, and pear flower detection using semantic segmentation network and shape constraint level set[J]. Computers and Electronics in Agriculture, 2021, 185: 106150.
- [109] CHEN L, ZHU Y, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[C]//Proceedings of the European Conference on Computer Vision, Munich, 2018: 801 – 818.
- [110] 熊俊涛, 刘柏林, 钟灼, 等. 基于深度语义分割网络的荔枝花叶分割与识别[J]. 农业机械学报, 2021, 52(6): 252 – 258. XIONG Juntao, LIU Bolin, ZHONG Zhuo, et al. Litchi flower and leaf segmentation and recognition based on deep semantic segmentation[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(6): 252 – 258. (in Chinese)
- [111] KANG H, CHEN C. Fruit detection and segmentation for apple harvesting using visual sensor in orchards[J]. Sensors, 2019, 19(20): 4599.
- [112] YU Y, ZHANG K, YANG L, et al. Fruit detection for strawberry harvesting robot in non-structural environment based on Mask R-CNN[J]. Computers and Electronics in Agriculture, 2019, 163: 104846.
- [113] GENE-MOLA J, SANZ-CORTIELLA R, ROSELL-POLO J R, et al. Fruit detection and 3D location using instance segmentation neural networks and structure-from-motion photogrammetry[J]. Computers and Electronics in Agriculture, 2020, 169: 105165.
- [114] SANTOS T T, DE SOUZA L L, DOS SANTOS A A, et al. Grape detection, segmentation, and tracking using deep neural networks and three-dimensional association[J]. Computers and Electronics in Agriculture, 2020, 170: 105247.
- [115] YANG C, XIONG L, WANG Z, et al. Integrated detection of citrus fruits and branches using a convolutional neural network[J]. Computers and Electronics in Agriculture, 2020, 174: 105469.
- [116] 张美娜, 王潇, 刘亚楠, 等. 温室远程监控系统人机交互与番茄识别研究[J]. 农业机械学报, 2022, 53(10): 363 – 370. ZHANG Meina, WANG Xiao, LIU Ya'nian, et al. Human computer interaction and tomato recognition in a greenhouse remote monitoring system[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(10): 363 – 370. (in Chinese)

- [117] JIA W, TIAN Y, LUO R, et al. Detection and segmentation of overlapped fruits based on optimized Mask R – CNN application in apple harvesting robot[J]. Computers and Electronics in Agriculture, 2020, 172: 105380.
- [118] WANG D, HE D. Fusion of Mask R – CNN and attention mechanism for instance segmentation of apples under complex background[J]. Computers and Electronics in Agriculture, 2022, 196: 106864.
- [119] XU P, FANG N, LIU N, et al. Visual recognition of cherry tomatoes in plant factory based on improved deep instance segmentation[J]. Computers and Electronics in Agriculture, 2022, 197: 106991.
- [120] 王文杰, 贡亮, 汪韬, 等. 基于多源图像融合的自然环境下番茄果实识别[J]. 农业机械学报, 2021, 52(9): 156 – 164.
WANG Wenjie, GONG Liang, WANG Tao, et al. Tomato fruit recognition based on multi-source fusion image segmentation algorithm in open environment[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(9): 156 – 164. (in Chinese)
- [121] KANG H, CHEN C. Fruit detection, segmentation and 3D visualisation of environments in apple orchards[J]. Computers and Electronics in Agriculture, 2020, 171: 105302.
- [122] HE K, GKIOXARI G, DOLLÁR P, et al. Mask R – CNN [C] // Proceedings of the IEEE International Conference on Computer Vision, Venice, 2017: 2961 – 2969.
- [123] ZHU L, XIE Z, LUO J, et al. Dynamic object detection algorithm based on lightweight shared feature pyramid[J]. Remote Sensing, 2021, 13(22): 4610.
- [124] WANG C, WANG Z, LI K, et al. Lightweight object detection model fused with feature pyramid[J/OL]. Multimedia Tools and Applications, 2022. <https://doi.org/10.1007/s11042-022-12127-4>.
- [125] HAN K, WANG Y, TIAN Q, et al. Ghostnet: more features from cheap operations [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, 2020: 1580 – 1589.
- [126] IANDOLA F N, HAN S, MOSKEWICZ M W, et al. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size [DB/OL]. (2016 – 02 – 24) [2022 – 08 – 10]. arXiv:1602.07360v4. <https://doi.org/10.48550/arXiv.1602.07360>.
- [127] LI D, SUN X, ELKHOCHLAA H, et al. Fast detection and location of longan fruits using UAV images[J]. Computers and Electronics in Agriculture, 2021, 190: 106465.
- [128] 李国进, 黄晓洁, 李修华, 等. 采用轻量级网络 MobileNetV2 的酿酒葡萄检测模型[J]. 农业工程学报, 2021, 37(17): 168 – 176.
LI Guojin, HUANG Xiaojie, LI Xiuhua, et al. Detection model for wine grapes using MobileNetV2 lightweight network[J]. Transactions of the CSAE, 2021, 37(17): 168 – 176. (in Chinese)
- [129] 吕石磊, 卢思华, 李震, 等. 基于改进 YOLOv3 – LITE 轻量级神经网络的柑橘识别方法[J]. 农业工程学报, 2019, 35(17): 205 – 214.
LÜ Shilei, LU Sihua, LI Zhen, et al. Orange recognition method using improved YOLOv3 – LITE lightweight neural network [J]. Transactions of the CSAE, 2019, 35(17): 205 – 214. (in Chinese)
- [170] CAO B, ZHANG B, ZHENG W, et al. Real-time, highly accurate robotic grasp detection utilizing transfer learning for robots manipulating fragile fruits with widely variable sizes and shapes[J]. Computers and Electronics in Agriculture, 2022, 200: 107254.
- [131] 王金鹏, 高凯, 姜洪喆, 等. 基于改进的轻量化卷积神经网络火龙果检测方法[J]. 农业工程学报, 2020, 36(20): 218 – 225.
WANG Jinpeng, GAO Kai, JIANG Hongzhe, et al. Method for detecting dragon fruit based on improved lightweight convolutional neural network[J]. Transactions of the CSAE, 2020, 36(20): 218 – 225. (in Chinese)
- [132] 张伏, 陈自均, 鲍若飞, 等. 基于改进型 YOLOv4 – LITE 轻量级神经网络的密集圣女果识别[J]. 农业工程学报, 2021, 37(16): 270 – 278.
ZHANG Fu, CHEN Zijun, BAO Ruofei, et al. Recognition of dense cherry tomatoes based on improved YOLOv4 – LITE lightweight neural network[J]. Transactions of the CSAE, 2021, 37(16): 270 – 278. (in Chinese)
- [133] 王卓, 王健, 王泉雄, 等. 基于改进 YOLOv4 的自然环境苹果轻量级检测方法[J]. 农业机械学报, 2022, 53(8): 294 – 302.
WANG Zhuo, WANG Jian, WANG Xiaoxiong, et al. Lightweight real-time apple detection method based on improved YOLOv4[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(8): 294 – 302. (in Chinese)
- [134] HOWARD A G, ZHU M, CHEN B, et al. MobileNets: efficient convolutional neural networks for mobile vision applications [DB/OL]. (2017 – 04 – 17) [2022 – 08 – 10]. arXiv:1704.04861v1. <https://doi.org/10.48550/arXiv.1704.04861>.
- [135] SANDLER M, HOWARD A, ZHU M, et al. MobileNetV2: inverted residuals and linear bottlenecks [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, 2018: 4510 – 4520.
- [136] HOWARD A G, SANDLER M, CHU G, et al. Searching for MobileNetV3 [C] // Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, 2019: 1314 – 1324.
- [137] ZHANG X, ZHOU X, LIN M, et al. ShuffleNet: an extremely efficient convolutional neural network for mobile devices [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, 2018: 6848 – 6856.
- [138] MA N, ZHANG X, ZHENG H, et al. ShuffleNetV2: practical guidelines for efficient CNN architecture design [C] // Proceedings of the European Conference on Computer Vision, Munich, 2018: 122 – 138.