

文章编号: 1674-8085 (2022) 06-0064-07

# 基于 PP-YOLO 深度学习模型的赣南脐橙 果实识别方法

章倩丽<sup>1,2</sup>, \*李秋生<sup>1,2</sup>

(1. 赣南师范大学智能控制工程技术研究中心, 江西, 赣州 341000; 2. 赣南师范大学物理与电子信息学院, 江西, 赣州 341000)

**摘 要:** 果实检测在研究脐橙采摘机械化发展中有着重要作用, 然而不良天气条件将对目标果实的检测和识别产生不利影响。针对雾天和雨天情形下脐橙果实图像模糊、噪声复杂, 检测速度较慢和准确率较低的问题, 通过采用单阶段目标检测网络 PP-YOLO 来研究不良天气条件下赣南脐橙果实的识别。通过主干网络 ResNet 提取特征并结合 FPN (特征金字塔网络) 进行特征融合实现多尺度检测, 且基本实现端到端检测。实验结果表明, 所提出的 PP-YOLO 检测模型可实现雾天和雨天情况下赣南脐橙检测任务, mAP 分别为 89.06% 和 91.01%, 识别效率分别可达到 75.30 fps 和 75.44 fps, 可以尝试在脐橙采摘机器人的研制中加以应用。

**关键词:** 目标检测; 果实识别; 机器视觉; PP-YOLO

中图分类号: TP391.41

文献标识码: A

DOI: 10.3969/j.issn.1674-8085.2022.06.010

## FRUIT RECOGNITION METHOD OF GANNAN NAVEL ORANGE BASED ON PP-YOLO DEEP LEARNING MODEL

ZHANG Qian-li<sup>1,2</sup>, \*LI Qiu-sheng<sup>1,2</sup>

(1. Research Center of Intelligent Control Engineering Technology, Ganzhou, Jiangxi 341000, China;

2. School of Physics and Electronic Information, Gannan Normal University, Ganzhou, Jiangxi 341000, China)

**Abstract:** Fruit detection plays an important role in studying the development of navel orange picking mechanization. However, adverse weather conditions will have an adverse impact on the detection and identification of target fruits. Aiming at the problems of blurry images, complex noise, slow detection speed and low accuracy rate of navel orange fruit under foggy and rainy days, this paper uses a single-stage target detection network PP-YOLO to study the identification of Gannan navel orange fruit under bad weather conditions. Feature extraction is achieved by the backbone network ResNet and feature fusion by combining FPN (feature pyramid network), and end-to-end detection is basically realized. The experimental results show that the proposed PP-YOLO detection model can realize the Gannan navel orange detection task under fog and rainy days, the mAP is 89.06% and 91.01%, and the recognition efficiency can reach 75.30 and 75.44fps, respectively, which can be tried to be applied in the development of navel orange picking robot.

**Key words:** machine vision; fruit identification; target detection; PP-YOLO

收稿日期: 2022-02-13; 修改日期: 2022-03-06

基金项目: 国家自然科学基金项目(42061027)

作者简介: 章倩丽(1997-), 女, 江西抚州人, 硕士生, 主要从事智能信息处理方面的研究(E-mail: zhangqianli@gnnu.edu.cn);

\*李秋生(1976-), 男, 江西南康人, 教授, 博士, 硕士生导师, 主要从事智能信息处理、目标识别与跟踪等方面的研究(E-mail: liqiusheng@gnnu.edu.cn).

## 0 引言

国内外农业生产的发展趋势都在沿着规模化、机械化、现代化方向发展,例如嫁接机器人、施肥和喷雾机器人、采摘机器人等设备的陆续研制及优化,未来在农业领域对智能机器的需求将不断增加<sup>[1]</sup>。我国是脐橙种植大国,根据江西省赣州市果业局统计,赣南地区是脐橙种植面积世界第一<sup>[2]</sup>,年产量世界第三、全国最大的脐橙主产区<sup>[3]</sup>。成熟赣南脐橙的精准高效采收是脐橙整个生产过程的关键,影响着脐橙的贮藏、运输、包装等环节。目前脐橙主要依靠人力收获,使得劳动力和生产成本增加。这种落后的方式严重制约了脐橙产业的健康快速发展,实现脐橙采收机械化,可以大幅度降低成本的同时也促进了脐橙产业的发展<sup>[4]</sup>。

机器视觉通过计算机模拟人类视觉采集并学习图像中的特征<sup>[5]</sup>,对果实进行识别、检测、定位,为采收机器人的研究提供重要参考<sup>[6]</sup>。自1968年视觉系统第一次被引入用于柑橘果实识别以来<sup>[7]</sup>,国内外众多学者针对自然环境下的果实识别方法开展了多方面研究工作,包括光照变化、阴影覆盖、着色度变化、枝叶遮挡、果实重叠等不同条件下果实识别任务所需的特征组合,以及从彩色、深度图像中提取上述特征和识别果实的方法<sup>[8]</sup>。

采摘机器人的研究主要有苹果<sup>[9]</sup>、樱桃<sup>[10]</sup>、番茄<sup>[11]</sup>等果蔬采摘机器人<sup>[12]</sup>,针对脐橙图像主要是品质检测<sup>[13]</sup>和分级<sup>[14]</sup>、缺陷检测<sup>[15]</sup>和病虫害检测<sup>[16]</sup>、脐橙表面残留物<sup>[17]</sup>的研究,真正将脐橙图像用于目标识别和定位中的研究较少。徐越等<sup>[18]</sup>利用颜色特征对重叠苹果果实进行识别,该方法在光照等自然条件好、成熟果实和未成熟果实差异明显的情况下可以达到较好的效果,但其适用性有一定的局限。高新浩等<sup>[19]</sup>利用卷积神经网络视觉处理算法对水果果实与复杂背景进行分割分类及二维定位,并结合景深传感数据进行果实3D定位,定位精度及果蔬采摘成功率高,但该算法计算复杂度较

大。Li等<sup>[20]</sup>研究了对多个荔枝串检测定位,先用语义分割分割出背景、果实和树枝,结合聚类算法聚类出同一荔枝串的果实分枝,通过主成分分析拟合出线性信息对应分枝位置,单个荔枝支路的检测精度达到83.33%,检测时间为0.464 s。宋怀波等<sup>[21]</sup>设计了针对复杂环境下的重度粘连小麦籽粒的检测方法,基于YOLOv5、混合深度可分离卷积以及压缩激励,模型检测精确率达到93.15%。但是该方法仍存在重度粘连难以识别,且检测速度较为缓慢。

传统的视觉定位系统对待检测物体所处的环境有很强的依赖关系,在恶劣天气条件下脐橙果实图像识别更加困难,有效解决雨天和雾天条件下的脐橙果实识别问题,可为实现雨天和雾天脐橙采摘机器人的作业提供技术支撑。目前,真正将脐橙图像用于目标识别和定位的研究少有报道,针对雾天和雨天条件下的研究更是鲜见。本研究利用赣南脐橙果实图像在不同特殊天气下的特征差异,基于PP-YOLO单阶段目标检测算法,对恶劣天气下的脐橙果实进行检测研究。

## 1 模型结构

基于锚框的单阶段目标检测网络主要组成部分有:主干网络、检测颈和检测头。PP-YOLO是PaddleDetection基于YOLOv3优化和改进的目标检测模型,采用加入可变形卷积的残差网络ResNet替换DarkNet作为主干网络提取特征,输出feature map,检测颈使用特征金字塔网络FPN进行特征融合,检测头沿用YOLOv3检测头分类及定位<sup>[22]</sup>。网络结构如图1所示。

### 1.1 残差结构

简单地将一系列卷积层和池化层堆叠搭建更深层的网络结构,随着网络层数不断加深,计算成本较高,不利于模型训练,且会导致梯度消失或梯度爆炸现象和退化问题。若每层误差梯度小于1,在反向传播过程中,每传播一层都乘这个系数,网络越深连乘次数越多,梯度将越来越小,最终导致

梯度消失;反之,若每层误差梯度大于1,则梯度将越来越大,最终导致梯度爆炸。网络层数加深导致训练误差不降反升,即为退化。针对这一类问题,何凯明等<sup>[23]</sup>提出残差网络结构,该模型局部结构如图2。

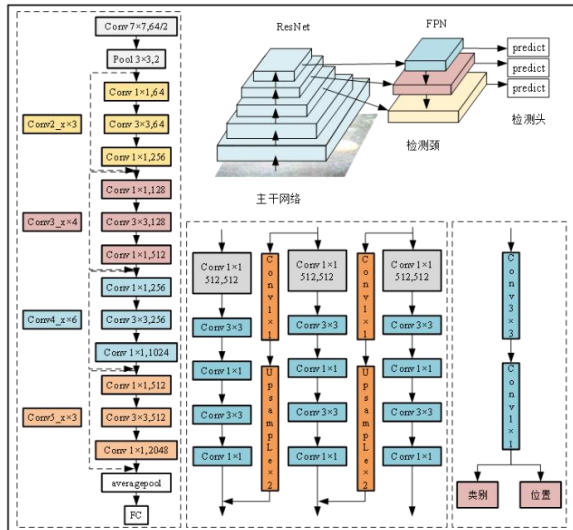


图1 PP-YOLO 结构图

Fig. 1 PP-YOLO structure diagram

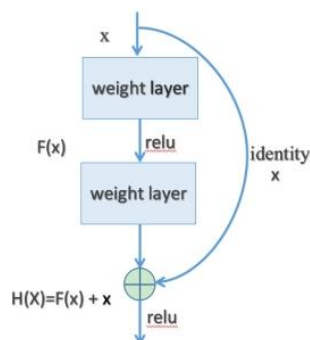


图2 残差结构示意图

Fig. 2 Schematic diagram of the residual structure

主分支将输入特征矩阵通过卷积层和激活函数,侧分支由输入直接连接到输出作恒等映射,主分支和侧分支所得到的特征矩阵在同一维度,相加后经过激活函数。通过残差结构加深网络,可有效降低训练误差。

## 1.2 特征金字塔网络

特征金字塔结构,把强位置信息的低层特征和高语义信息的高层特征进行融合,达到多尺度预测,其结构如图3所示。

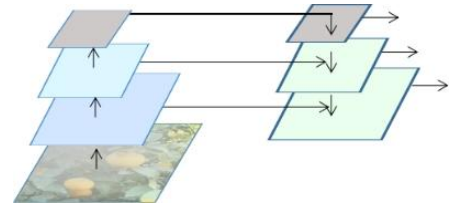


图3 特征金字塔结构图

Fig. 3 Characteristic pyramid structure diagram

从底向上的路径,通过一些卷积进行特征提取,得到不同尺寸的feature map,同时从顶到下的路径,feature map进行卷积和上采样,将强位置信息和高语义信息横向连接,得到多尺度特征,送到检测头再进行预测分类和精准定位<sup>[24]</sup>。

## 1.3 迁移学习

迁移学习<sup>[25]</sup>在深度学习过程中被普遍使用,迁移即将已学习网络的浅层网络的参数迁移到新网络中,新网络不仅能够识别底层通用特征,还能快速学习新数据集的高维特征。训练模型时,模型的损失值随时间推移而慢慢收敛,最终趋于稳定,其权值初始化需耗费大量时间,通过迁移学习可加快训练的进程,并且在数据集较小时也能得到较好的效果。

本研究采用迁移学习的思想,选择在公共数据集ImageNet上预训练好的分类模型进行权重初始化,减少训练所需的时间以及资源。

## 2 数据集制作

### 2.1 图像采集

图像于2020年11月14日17:00~18:30之间在江西省赣州市赣南师范大学的脐橙园所采集,从各个角度对脐橙进行拍摄阴天时的脐橙图片,拍摄设备为1200万以上像素的手机。总共采集脐橙图像500张,筛选删除模糊不清以及不适用的图片后保留了450张制作数据集。包含完整无遮挡、树枝树叶树干遮挡、脐橙果实相互遮挡的脐橙图像,且2/3以上的图像包含10个以上脐橙果实,图像保存格式为jpg格式。

## 2.2 制作数据集

雾天图片共 444 张,雨天图片共 445 张。将阴天拍摄图片作为输入,通过加雾算法和加雨算法生成带雾和雨的图片。由于数据集为仿真生成,与实际应用环境中获取的图像存在一些差异,加载此数据集训练的模型真实应用效果有待进一步探究。图片示例如图 4 所示。



图 4 部分雾天和雨天的成熟脐橙图片

Fig. 4 Partial picture of ripe navel oranges on foggy and rainy days

首先将图像数据按照 7:3 的比例划分,即雾天 310 张作为训练集,剩余 134 张作为测试集;雨天 310 张作为训练集,剩余 135 张作为测试集。训练集和测试集互斥。利用图像标注工具labelImg进行标注,格式选择VOC格式,得到带有图片内目标类别和相应位置信息的xml文件。编辑并运行python脚本,遍历划分好的图像文件和标注信息文件,生成最终模型训练所需训练集和测试集文件列表。图片文件、标注信息文件、训练集和测试集文件列表、标签文件,即构成了模型的数据集,具体数据如表 1 所示。

表 1 恶劣天气下脐橙数据集

Table 1 Navel orange dataset in severe weather

|       | 雾天数据               | 雨天数据            |
|-------|--------------------|-----------------|
| 训练集   | 310                | 310             |
| 测试集   | 134                | 135             |
| 总计    | 444                | 445             |
| 图像分辨率 | 400×300<br>300×300 | 400×300 300×300 |

## 2.3 数据增强

由于自制的数据集图片数量较少,通过采用数据增强操作扩充脐橙数据集,使得模型训练学习到较完整的特征,减少数据集过于简单造成的过拟合现象。数据集输入模型之后,引入随机因子在线读取每个batch图像数据,并进行随机裁剪、平移、旋转、重叠。因为每次训练batch图像不一致,以及增强的随机因子极大地扩大了数据集,确保模型能够有效地进行训练。

## 3 模型训练

### 3.1 训练平台

模型训练基于百度大脑AI studio,支持在线云计算,采用开源的深度学习框架PaddlePaddle,物体检测统一框架PaddleDetection。本试验运行所需的软硬件信息和环境配置如下:4核CPU,RAM 32 GB,GPU为Tesla V100,显存 16 GB,磁盘 100 GB,Python 3.7,PaddlePaddle 1.8.4,CUDA功能 70,驱动API版本 10.1,运行时API版本 9.0,cuDNN7.6。

### 3.2 参数设置

初始化采用预训练好的分类模型权重,由于使用单卡GPU,且训练图像大约 300 张,所以batch\_size设置为 2,初始学习率设置为 0.001,最大迭代次数设置为 15000。数据集制作使用VOC数据格式,mAP评估方法采用VOC的 11point方法。11point,即将所有预测结果按得分从大到小排序,分别计算召回率大于等于从 0 到 1 每隔 0.1 一个点(共 11 个点)的最大精度值,计算 11 次的结果相加求和计算平均。

### 3.3 评估指标及结果分析

目标检测任务常用网络模型性能评价指标为精确率(precision)、召回率(recall)、平均精度(AP)和mAP等。计算precision-recall曲线下的面积即平均精度(AP),mAP即计算每类AP的平均。mAP位于区间[0, 1],是目标检测算法最重要的评价指标之一。

精确率(precision)和召回率(recall)计算公式如



式(1)、(2)所示:

$$\text{精确率(precision)} = \frac{TP}{TP + FP} \quad (1)$$

$$\text{召回率(recall)} = \frac{TP}{TP + FN} \quad (2)$$

$TP$  真阳 (True Positives)、 $TN$  真阴 (True Negatives)、 $FP$  伪阳(False Positives)、 $FN$  伪阴(False Negatives)具体含义见表 2。

在本次赣南脐橙目标检测任务中,可作以下理解:

a.  $TP$ : 脐橙被正确地识别成了脐橙。

b.  $TN$ : 树叶树枝天空等背景被正确地识别为树叶、树枝、天空等背景。

c.  $FP$ : 树叶树枝天空等背景被错误地识别为脐橙。

d.  $FN$ : 脐橙被错误地识别为树叶、树枝、天空等背景。

表 2 相关表述

Table 2 Related expressions

|                  | P  | N              |
|------------------|----|----------------|
|                  | TP | TN             |
| T (正样本被正确识别为正样本) |    | (负样本被正确识别为负样本) |
|                  | FP | FN             |
| F (负样本被错误识别为正样本) |    | (正样本被错误识别为负样本) |

雾天和雨天PP-YOLO模型训练过程的损失变化曲线如图 5 所示,由于采用迁移学习在底层卷积层中共享的权重,训练损失值在短期内迅速下降并收敛。

图 6 所示为在雾天和雨天条件下使用PP-YOLO模型训练脐橙检测网络的平均精度值随迭代次数变化曲线。受益于预训练权重,平均精度值在较短的迭代轮次内即得到了较好的效果,后期随迭代次数增加而上升,最后趋于饱和。雾天mAP达到了89.06%,雨天达到了91.01%。

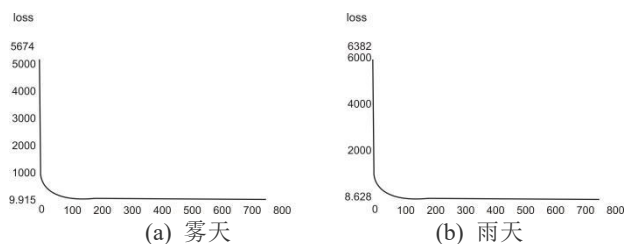
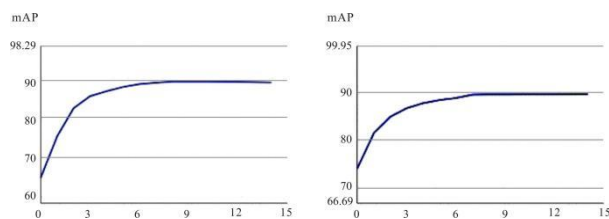


图 5 雾天及雨天训练损失值变化曲线

Fig. 5 Variation curves of training loss values on foggy and rainy days.



(a) 雾天

(b) 雨天

图 6 雾天及雨天 mAP 变化曲线

Fig. 6 MAP change curves on foggy and rainy days

使用相同的雾天和雨天数据集设计对比试验,比较YOLOv3 模型和本文的PP-YOLO模型对赣南脐橙的检测识别性能,实验所得mAP和检测效率结果如表 3 所示。

表 3 恶劣天气下脐橙检测结果

Table 3 Navel orange test results in bad weather

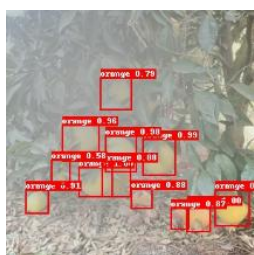
|     | 雾天        |           | 雨天        |           |
|-----|-----------|-----------|-----------|-----------|
|     | YOLOv3    | PP-YOLO   | YOLOv3    | PP-YOLO   |
| mAP | 85.65%    | 89.06%    | 88.27%    | 91.01%    |
| 速度  | 67.35 fps | 75.30 fps | 68.92 fps | 75.44f ps |

结果表明,采用PP-YOLO对雾天和雨天条件下的赣南脐橙有较优的识别效果,检测效率也明显更高: 1) 使用同一模型检测不同天气环境下的脐橙时,雨天的识别精度和速度要略高于雾天,因为雨天图像更易于识别; 2) 在同一天气情况下,PP-YOLO 与 YOLOv3 相比较, mAP 分别提高了 3.41%和 2.74%,检测效率分别加快了 7.95 fps和 6.52 fps。相较于YOLOv3, PP-YOLO模型主干网络采用了残差网络,同时引入可变形卷积在原卷积采样位置增加偏移,使得卷积核在每一个元素上额外增加一个方向参数,扩大训练过程中卷积核的范围,捕捉图像中的细节信息,能更好地提取图像特征。因此,PP-YOLO模型检测性能较高于YOLOv3。此外,雾天图像噪声较复杂于雨天图像检测难度也较高,因此雨天检测性能较好。

### 3.4 模型检测效果

分别选择雾天和雨天没有标注训练的脐橙图片采用训练好的YOLOv3 以及PP-YOLO模型进行预测,其结果如图 7 和图 8。结果表明,本次采用的PP-YOLO模型泛化性较好,对于未经训练的雾天

和雨天情况下脐橙图像数据漏检、误检果实较少。但是,仍存在部分果实误检情况,且雾天较雨天检测效果要更差。综上所述,在恶劣天气下,本实验提出模型对赣南脐橙仍拥有较好的检测能力,且对部分遮挡情况下的脐橙也有较好的识别效果。



(a) YOLOv3



(b) PP-YOLO

图7 雾天检测结果示例图

Fig. 7 Example diagram of foggy weather detection results

## 4 结论

针对在雾天和雨天等恶劣天气下脐橙果实检测速度和准确率较低这一问题,本研究运用基于YOLOv3改进的PP-YOLO目标检测模型对恶劣天气下赣南脐橙的果实识别问题进行了研究。结合迁移ImageNet数据集训练的权重模型到脐橙检测任务中,节省了训练时间。在雾天和雨天检测准确率

分别可达89.06%和91.01%,检测效率分别可达75.30 fps和75.44 fps。在提高检测精度的同时,也均衡了检测速度。实验结果表明,相比YOLOv3模型,该PP-YOLO模型具有更高的识别准确率和识别效率,可以有效地检测恶劣天气下的赣南脐橙果实,满足实际应用中采摘机器人的实时检测要求。当然,由于数据集数量相对较少,模型可能未学习到全部特征;且雾天和雨天的数据为基于阴天脐橙图像用算法生成,可能与真实情况存在距离,所以在实际应用中的结果还有待进一步研究。



(a) YOLOv3



(b) PP-YOLO

图8 雨天检测结果示例图

Fig. 8 Example of rainy day test result

## 参考文献:

- [1] 王家硕,陈炳舟,杨广召,等.农业机器人在农业中应用[J].安徽农学通报,2020,26(20):132-133.
- [2] 徐丽明,刘旭东,张凯良,等.脐橙采摘机器人末端执行器

- 设计与试验[J]. 农业工程学报, 2018, 34(12): 53-61.
- [3] 李追风, 时文涛, 姬志发, 等. 国内外农业机器人研究进展分析[J]. 南方农机, 2022, 53(5): 156-158.
- [4] 彭威. 脐橙收获机的研究现状及前景[C]. 以创新驱动为引领, 加快“中国制造 2025”战略实施研讨会暨 2016 年第六届全国地方机械工程学会学术年会论文集[A]. 2016.
- [5] 胡潇. 可伸缩采摘机械手目标识别与定位研究[D]. 南昌: 江西农业大学, 2020.
- [6] 仇裕淇, 黄振楠, 阮昭, 等. 机器视觉技术在农业生产智能化中的应用综述[J]. 机械研究与应用, 2019, 32(2): 202-206.
- [7] Schertz C E. Human energy expenditure for picking oranges [J]. *Journal of Agricultural Engineering Research*, 1967, 12(4): 281-284.
- [8] Gongal A. Sensors and systems for fruit detection and localization: A review [J]. *Computers & Electronics in Agriculture*, 2015, 116(C): 8-19.
- [9] Kuznetsova A, Maleva T, Solovie V. Using YOLOv3 algorithm with pre- and post-processing for apple detection in fruit-harvesting robot [J]. *Agronomy*, 2020, 10(7):1016.
- [10] Zhang Q, Zuo J, Yu T, et al. Visual recognition system of cherry picking robot based on Lab color model [C]. 2017 IOP Conference Series: Earth and Environmental Science, 2017, 100: 012099.
- [11] Tong S, Qian D, Fang J, et al. Residual vibration suppression of a tomatoes' string picking robot: A cascade design method [C]. 2016 International Conference on Advanced Mechatronic Systems(ICAMechS).IEEE, 2017.
- [12] 王丹丹, 宋怀波, 何东健. 苹果采摘机器人视觉系统研究进展[J]. 农业工程学报, 2017, 33(10):59-69.
- [13] Liu C, Yang S, Deng L. Determination of internal qualities of newhall navel oranges based on NIR spectroscopy using machine learning [J]. *Journal of Food Engineering*, 2015, 161(sep.): 16-23.
- [14] 胡发焕, 董增文, 匡以顺. 基于机器视觉的脐橙品质在线分级检测系统[J]. 中国农业大学学报, 2016, 21(3):112- 118.
- [15] Xie X, Ge S, Xie M, et al. Image matching algorithm of defects on navel orange surface based on compressed sensing [J]. *Journal of Ambient Intelligence and Humanized Computing*, 2018, 18:833.
- [16] Zhang L, Liu M, Li J, et al. Detecting navel orange canker with hyperspectral imaging [C]. 2011 Fourth International Conference on Intelligent Computation Technology and Automation, 2011: 106-109.
- [17] Hu H, Huang L, Liu M, et al. Nondestructive determination of cu residue in orange peel by laser induced breakdown spectroscopy [J]. *Plasma science & Technology*, 2015, 17(8):711-715.
- [18] 徐越, 李盈慧, 宋怀波, 等. 基于Snake模型与角点检测的双果重叠苹果目标分割方法[J]. 农业工程学报, 2015, 31(1):196-203.
- [19] 高新浩, 王占山, 温贻芳, 等. 基于景深传感技术的水果采摘视觉定位系统设计[J]. 湖北农机化, 2019(10): 45-46.
- [20] Li J, Tang Y, Zou X, et al. Detection of fruit-bearing branches and localization of litchi clusters for vision-based harvesting robots [J]. *IEEE Access*, 2020, 8: 117746-117758.
- [21] 宋怀波, 王云飞, 段援朝, 等. 基于YOLOv5- MD 的重度粘连小麦籽粒检测方法[J/OL]. 农业机械学报, 2022. <http://kns.cnki.net/kcms/detail/11.1964.S.20220225.1020.004.html>.
- [22] Long X, Deng K, Wang G, et al. PP-YOLO: An effective and efficient implementation of object detector [J]. *arXiv:2007.12099*, 2020.
- [23] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[J]. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 770-778.
- [24] Lin T Y, Dollar P, Girshick R, et al. Feature pyramid networks for object detection [J]. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017: 936-944.
- [25] 张雪松, 庄严, 闫飞, 等. 基于迁移学习的类别级物体识别与检测研究与进展[J]. 自动化学报, 2019, 45 (7):1224-1243.