

方国文, 何 超, 王鑫泽. 基于 YOLOv8n 的轻量级巴旦木果实识别方法[J]. 江苏农业学报, 2024, 40(9): 1662-1670.

doi: 10.3969/j.issn.1000-4440.2024.09.10

# 基于 YOLOv8n 的轻量级巴旦木果实识别方法

方国文, 何 超, 王鑫泽

(西南林业大学机械与交通学院, 云南 昆明 650000)

**摘要:** 在果园环境下, 快速精准识别巴旦木果实对提升巴旦木采摘机器人的作业精度和效率至关重要。为减少果园场景中因树叶遮挡或果实重叠导致的巴旦木果实漏检现象, 降低计算量和参数量, 提高果实识别模型的性能和准确度, 本研究在 YOLOv8n 模型的基础上, 利用 ContextGuide 模块替换原模型中主干网络(Backbone)部分基本构成单元 C2f 中的 Bottleneck 模块, 利用 BiFPN 模块替代原模型中颈部网络(Neck)部分中的 PANet 模块, 同时引入 MPDIoU 损失函数替换原模型中的 CIoU 损失函数, 提出了一种改进的轻量级巴旦木果实检测模型(YOLOv8n-BCG)。并利用公开的巴旦木影像数据集对优化后的模型性能进行比较分析。结果表明, 改进后模型参数量仅为 1.528 M, 平均精度值( $mAP_{0.50:0.95}$ )为 69.7%, 相比于原 YOLOv8n 模型提升 0.5 个百分点。与 YOLOv5s、YOLOv5n、YOLOv7-tiny、Faster R-CNN 等模型相比, YOLOv8n-BCG 模型具有更低的浮点计算量和更高的检测精度值。本研究结果可为高效的巴旦木果实采摘机器人自动化作业提供技术支持。

**关键词:** 巴旦木; 果实识别; BiFPN; ContextGuide; MPDIoU 损失函数; YOLOv8n

**中图分类号:** TP391.4      **文献标识码:** A      **文章编号:** 1000-4440(2024)09-1662-09

## Lightweight almond fruit recognition method based on YOLOv8n

FANG Guowen, HE Chao, WANG Xinze

(College of Mechanical and Transportation Engineering, Southwest Forestry University, Kunming 650000, China)

**Abstract:** In the orchard environment, rapid and accurate identification of almond fruits is very important to improve the operation accuracy and efficiency of almond picking robots. In order to reduce the missed detection of almond fruits caused by leaf occlusion or fruit overlap in the orchard scene, reduce the amount of calculation and parameters, and improve the performance and accuracy of the fruit recognition model, based on the YOLOv8n model, this study used the ContextGuide module to replace the Bottleneck module in the basic component unit C2f of the Backbone part of the original model, used the BiFPN module to replace the PANet module in the Neck part of the original model, and introduced the MPDIoU loss function to replace the CIoU loss function in the original model. An improved lightweight almond fruit detection model (YOLOv8n-BCG) was proposed. The performance of the optimized model was compared and analyzed by using the public almond image data set. The results showed that the number of parameters of the improved model was only 1.528 M, and the mean average precision ( $mAP_{0.50:0.95}$ ) was 69.7%, which was 0.5 percentage points higher than that of the original YOLOv8n model. Compared with YOLOv5s, YOLOv5n, YOLOv7-tiny and Faster R-CNN models, YOLOv8n-BCG model had lower floating-point calculation and higher detection accuracy. The results of this study can provide technical support for

efficient automatic operation of almond picking robots.

**Key words:** almond; fruit recognition; BiFPN; ContextGuide; MPDIoU loss function; YOLOv8n

收稿日期: 2023-12-04

基金项目: 云南省高层次人才项目(YNWR-QNBJ-2018-066、YNQR-CYRC-2019-001)

作者简介: 方国文(2001-), 女, 安徽合肥人, 硕士研究生, 主要研究方向为图像识别。(E-mail) fgw09182023@163.com

通讯作者: 何 超, (Tel) 15887130986; (E-mail) hcsmile@163.com

巴旦木, 又名扁桃, 是蔷薇科桃亚属植物。巴旦木果仁富含植物油、蛋白质、淀粉、糖, 并含有少量维

生素 A、维生素 B<sub>1</sub>、维生素 B<sub>2</sub> 和消化酶、杏仁素酶以及钙、镁、钠、钾、铁、钴等营养元素,具有降低胆固醇、有益心脏和肠道、控制体重和血糖水平等功效,越来越受消费者喜爱<sup>[1]</sup>。目前巴旦木果实的采摘仍然依赖传统的人工方式,这种方式成本高、效率低,无法满足巴旦木产业的发展需求。因此,研发巴旦木果实自动化采摘机器人已经成为不可避免的趋势<sup>[2]</sup>。果园环境条件下巴旦木果实的快速准确识别,对巴旦木果实的自动化智能采摘尤为重要<sup>[3]</sup>。

传统的计算机目标识别主要依靠形状、纹理和色泽等特征对图像进行分割与目标识别。陈礼鹏等<sup>[4]</sup>采用图像 RGB 分量<sup>[5]</sup>,实现猕猴桃果实的识别。贾伟宽等<sup>[6]</sup>利用 3 种降噪方法进行苹果果实的夜间识别。李斌等<sup>[7]</sup>基于 RGB<sup>[8]</sup> 和 HSL<sup>[9]</sup> 颜色模型提取枇杷果实的形态特征,从而进行枇杷果实碰伤等级的检测。李昕等<sup>[10]</sup>采用 Hough 变换算法<sup>[11]</sup> 实现有遮挡的油茶果实识别。近年来,随着卷积神经网络等深度学习方法的提出,基于数据驱动的目标检测模型在作物果实识别及分级<sup>[18]</sup>、病虫害检测<sup>[19]</sup> 以及产量预估<sup>[20]</sup> 等研究中得到初步应用<sup>[21]</sup>。在深度学习的方法中,李恒等<sup>[22]</sup> 基于 GhostNet 模型<sup>[23]</sup> 提取主干特征信息,并采用复杂双向多尺度融合算法<sup>[24]</sup> 对提取的主干特征信息进行融合,实现了绿色类圆果实的识别,模型占用内存 11.8 M,且每幅照片检测所需时间仅 37 ms,识别精度达 96.8%。储鑫等<sup>[25]</sup> 利用 MobileNetv1 网络<sup>[26]</sup> 代替 YOLOv4 模型<sup>[27]</sup> 中原有的主干网络 CSPDarknet53<sup>[28]</sup> 进行特征提取,并在特征融合网络 PANet 中引入深度可分离卷积<sup>[29]</sup> 代替原有的 3×3 标准卷积,同时在主干网络中嵌入卷积块注意力模块(CBAM),进而进行番茄叶部病害识别,结果发现改进后的模型比原模型参数减少 80%,检测时间减少 59.1%,识别精度提升 0.6 个百分点。代云等<sup>[30]</sup> 设计 Mob-darknet-52 特征提取网络代替 YOLOv3 模型中的 Darknet-52 模块,并结合多尺度检测模型,对 YOLOv3 模型进行了改进,并利用改进后的模型进行密集环境下槟榔果实的快速识别,识别精度和效率均比原模型有所提升。

上述研究结果表明,深度学习技术是果园中果实精准识别较好的方法。但在果园小目标果实检测中,上述模型常受到终端检测设备计算能力的限制,如何对计算量庞大的目标检测模型进行轻量化处理,是目前果园中果实小目标检测的一个热点问题。本研究

拟在 YOLOv8n<sup>[31]</sup> 模型的基础上,利用 BiFPN<sup>[32]</sup> 模块代替通用的 PANet 模块, ContextGuide<sup>[33]</sup> 代替 C2f 中的 Bottleneck 模块,构建轻量化模型 YOLOv8n-BCG,为巴旦木果实的自动化采摘提供依据和支撑。

## 1 材料与方法

### 1.1 数据集的获取

本研究选择公开的巴旦木影像数据集 ACFR-Orchard-Fruit-Dataset 为材料进行巴旦木果实的识别和模型比较。该数据集来自澳大利亚悉尼大学和澳大利亚野外机器人技术中心,共包含 620 张巴旦木结果期果树图片,图像分辨率为 300 像素×300 像素,每张图片中平均约有 24 颗巴旦木果实。为避免模型训练过程中出现过拟合现象,提高数据样本的多样性,采用旋转、裁剪、调节图像亮度、加入噪声以及上述 4 种增强方式两两结合,旋转与裁剪结合、旋转与调节图像亮度结合、旋转与加入噪声结合、裁剪与调节图像亮度结合、裁剪与加入噪声结合、调节图像亮度与加入噪声相结合以及 4 种方法相结合的方法来进行原始数据集的增强,增强后共得到巴旦木果树图像 7 440 张。由于部分原始图像本身较模糊,数据增强后除去模糊不清以及过度曝光的 233 张图像,得到 7 207 张图像作为数据集。对数据集图像中的巴旦木果实位置进行标注,并按照 8:1:1 的比例将数据集随机划分为训练集、验证集和测试集进行试验。增强后的部分图像如图 1 所示。

### 1.2 YOLOv8n 模型及改进

YOLOv8 系列模型<sup>[31]</sup> 是 Ultralytics 公司于 2023 年发布的图像分割、目标检测模型,其精度和速度较以前版本均有显著提升。该模型主要包括特征提取的主干网络(Backbone)、特征融合的颈部网络(Neck)以及检测头(Head)等。YOLOv8 系列模型采用 C2f 模块替换 YOLOv5 模型中的 C3 模块,使得模型在保证轻量化的同时获得更加丰富的梯度流信息,并重新融入无锚框思想,即通过多个关键点或者中心点与边界的信息来表示物体,使得模型非常适合密集目标的检测。YOLOv8 模型的识别损失主要包括类别分类损失和边界框回归损失,其中类别分类损失(Cls.Loss)采用交叉熵损失表征,而边界框回归损失(Bbox.Loss)则采用交并比(CIOU)表征,同时还引入了 DFL 损失来处理数据不平衡的问题。总体的损失函数(L)如式(1)所示:

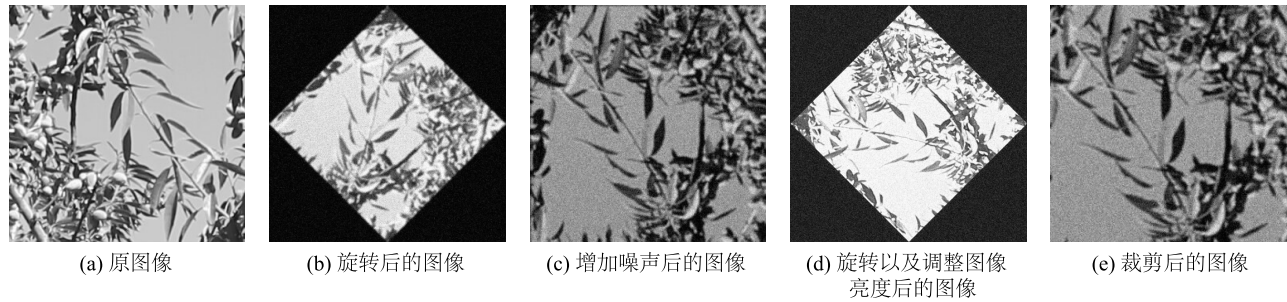


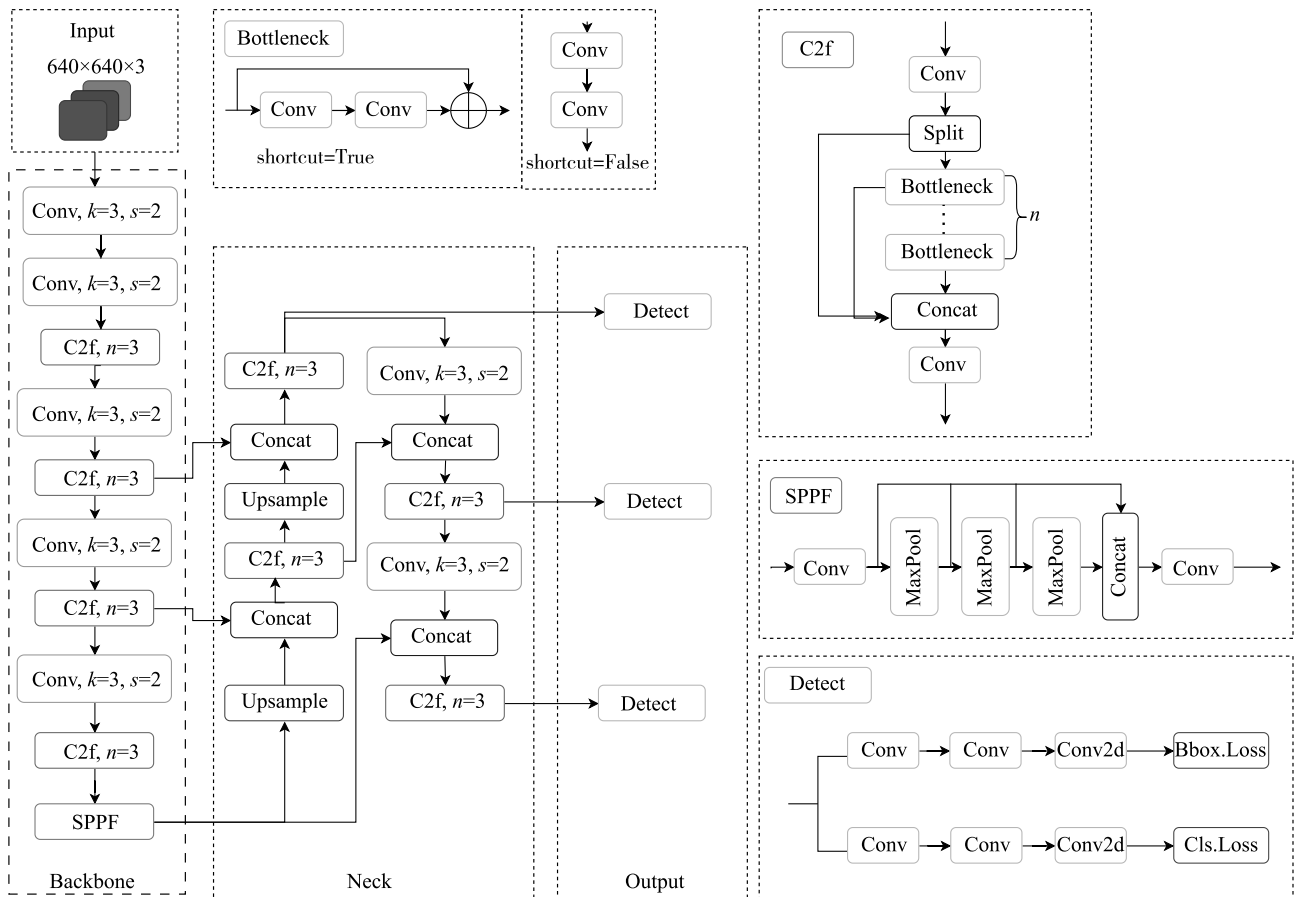
图 1 ACFR-Orchard-Fruit-Dataset 数据集图像示例及其 4 种增强方法增强后的图像

Fig.1 An image from ACFR-Orchard-Fruit-Dataset dataset and the images enhanced by four enhancement methods

$$L = \text{Loss}_{CIUO} + \text{Loss}_{DFL} + \text{Loss}_{CLS} \quad (1)$$

YOLOv8n 是 YOLOv8 系列模型中的一个轻量

化版本,其推理速度最快,占用内存最小,其结构如图 2 所示。



Input: 输入; Backbone: 主干网络; Concat: 融合; Upsample: 上采样; Shortcut: 最短路径; Neck: 颈部网络; Output: 输出; C2f: 特征提取模块; Split: 分割模块; Bottleneck: 一部分通过一系列的卷积、归一化和激活操作进行处理,最后生成的特征图会与另一部分直接传递的特征图在连接模块进行拼接; SPPF: 空间金字塔池化模块; Detect: 检测头; Bbox.Loss: 边界框回归损失函数; Cls.Loss: 分类损失函数; Conv: 卷积处理;  $k$ : 卷积核的大小;  $s$ : 步长;  $n$ : 瓶颈层数。

图 2 YOLOv8n 模型简图

Fig.2 Structure diagram of YOLOv8n model

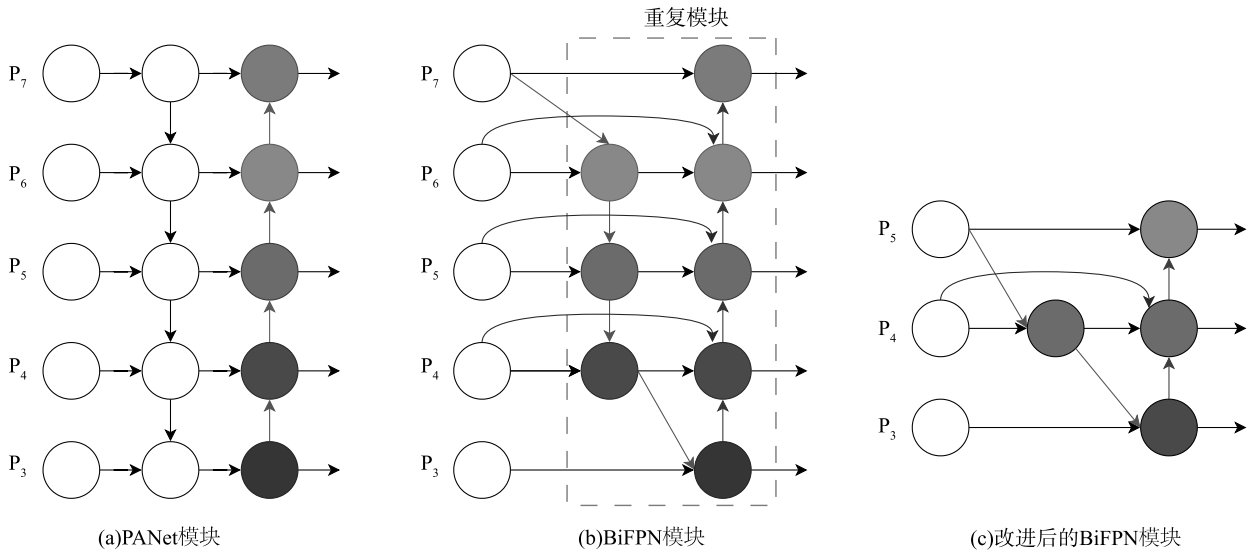
原始 YOLOv8n 模型计算量庞大,结构复杂,导致其在目标识别时对终端设备的计算能力有较高的

要求。本研究在 YOLOv8n 原始模型的基础上,通过模块优化提出了一种轻量化的目标识别检测模

型——YOLOv8n-BCG。模型的优化主要体现在 3 个方面。一是在主干网络(Backbone)部分基本构成单元 C2f 中利用 ContextGuide 模块替代 Bottleneck 模块,二是在颈部网络(Neck)中利用 BiFPN 模块替代原来的 PANet 模块,三是利用 MPDIoU 损失函数替代原来的边界框损失函数 CIoU。

1.2.1 加权双向特征金字塔网络(BiFPN) 原始 YOLOv8n 模型中 Neck 网络是利用 PANet 模块进行不同特征层的数据融合(图 3a)。该模块通过不同特征层的简单相加实现数据融合,该方案在识别大小均匀的目标时是可靠的。但本研究中识别的巴旦木果实大小不一,特征层的简单相加会导致融合后

输出特征数据不精确。为解决这一问题,本研究借鉴 Tan 等<sup>[32]</sup>提出的 BiFPN 模块代替 PANet 模块进行数据融合。BiFPN 模块能为不同特征层分配不同权重,从而提高模型的目标识别性能。BiFPN 移除 1 个未进行特征融合的中间节点,同时为了加强特征融合,在同层输入节点与输出节点间增加特征分支(图 3b)。本研究为实现轻量化计算,对原始的 BiFPN 模块在保留双向跨尺度连接的基础上,去掉多余支路,仅使用 P3、P4、P5 通道进行输出,这样能在不影响特征融合的情况下进一步精简网络结构,使模型更加轻量化(图 3c)。



$P_3 \sim P_7$  为模型特征层数,第 3 层至第 7 层。

图 3 PANet 模块结构和 BiFPN 模块结构

Fig.3 PANet module structure and BiFPN module structure

特征融合过程中,不同特征层分辨率不同,因而其对输出特征的贡献亦不相同。本研究利用的巴旦木影像数据集大多包含较多的识别目标,BiFPN 模块可快速融合不同特征层的数据信息,挖掘果实深层信息,减少环境复杂而造成漏检、误检。BiFPN 的带权特征融合公式为:

$$O = \sum_i \frac{w_i}{\varepsilon + W} \times I_i \quad (2)$$

式中, $O$ :输出信息, $I_i$ :第  $i$  层输入信息, $\varepsilon$ :用于约束数值振荡的极小值学习率,取值为 0.000 1, $w_i$  为第  $i$  特征层权重, $W$  为所有特征层总权重。

以图 4 中的 P4 特征融合为例,具体计算为:

$$P_4^{td} = \text{Conv} \left[ \frac{w_1 * P_4^{\text{in}} + w_2 * \text{Resize}(P_5^{\text{in}})}{w_1 + w_2 + \varepsilon} \right] \quad (3)$$

$$P_4^{\text{out}} = \text{Conv} \left[ \frac{w'_1 * P_4^{\text{in}} + w'_2 * P_4^{td} + w'_3 * \text{Resize}(P_3^{\text{out}})}{w'_1 + w'_2 + w'_3 + \varepsilon} \right] \quad (4)$$

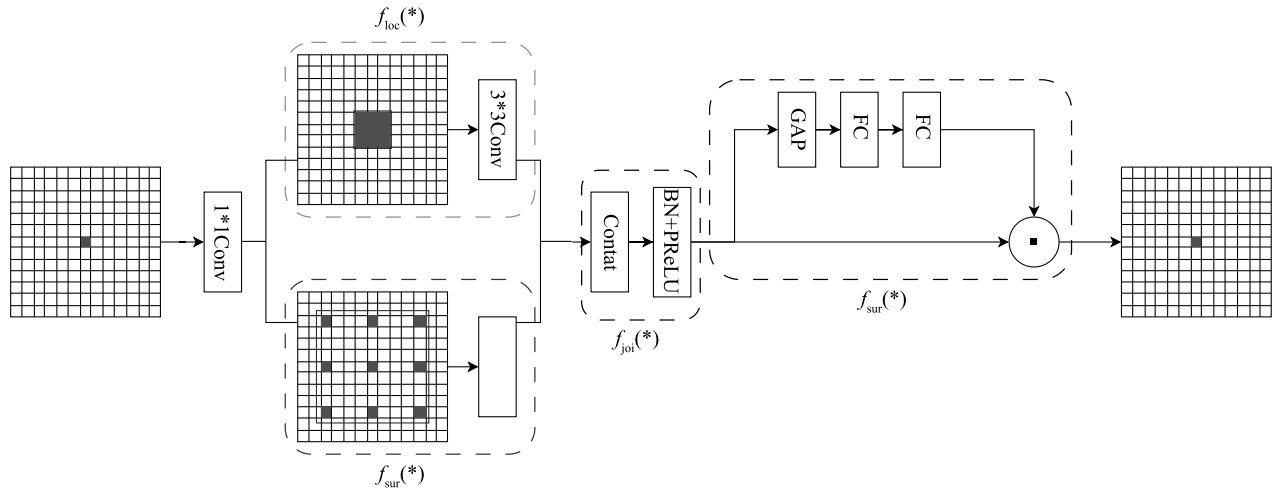
式(3)、(4)中, $P_4^{td}$  为自上而下(top-down)路径中 P4 的中间特征, $P_4^{\text{out}}$  为自下而上(bottle-up)路径中 P4 的输出特征, $P_4^{\text{in}}$  为 bottle-up 路径中的输入特征,Conv 为对应的卷积操作,Resize 为上采样或下采样操作, $w_1$ 、 $w_2$ 、 $w_3$  分别为第 1 层、第 2 层和第 3 层权重, $w'_1$ 、 $w'_2$ 、 $w'_3$  分别为特征融合后的第 1 层、第 2 层和第 3 层权重, $\varepsilon$  为极小非零常数,取值为 0.000 1。



**1.2.2 ContextGuide (CG) 模块** ContextGuided Network (CGNet) 是 Wu 等<sup>[33]</sup>提出的语义分割轻量级模型。CGNet 模型包含 51 层, 3 个阶段, 每个阶段的通道数分别为 32、64、128。假设输入图像大小为  $680 \times 680 \times 3$ , 对图像分别进行  $1/2$ 、 $1/4$ 、 $1/8$  的下采样, 输出图像大小为  $85 \times 85$ , 这样能较好地保留边缘信息。每个阶段的第 1 层输入是来自上 1 个阶段的第 1 层和最后 1 层的输出组合, 有利于特征重用和传播。CGNet 主要是由 CG 模块构建而成, 其可以较好地获取上下层图像的纹理特征, 并通过逐通道重新加权, 引入全局上下层图像的特征进一步改善联合特征。CG 模块中的  $f_{loc}$  提取器用于提取局部特征,  $f_{sur}$  提取器用于提取周围上下层图像的特征,  $f_{joi}$  提取器用于提取联合特征,  $f_{glo}$  提取器用于提取全局

上下层图像的特征(图 4)。CG 模块还采用两种残差连接, 一是连接 CG 模块输入和  $f_{joi}$  输出, 称为局部残差学习 (LRL), 二是连接 CG 模块输入和  $f_{glo}$  输出, 称为全局残差学习 (GRL)。相比较而言, GRL 能促进网络中的信息传递, 有助于模型学习复杂的特征并改善训练期间的梯度反向传播, 具有更好的效果。

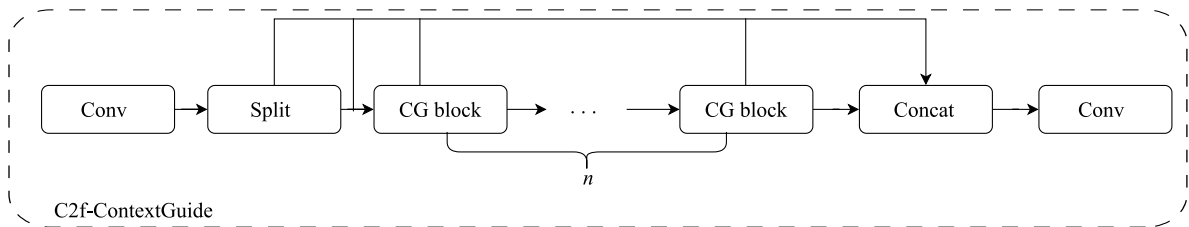
YOLOv8n 主干网络 (Backbone) 部分基本构成单元 C2f 虽可以保证目标识别的准确性和对图像细节的感知能力, 但其浮点计算量和参数量均较高。所以本研究中将 C2f 中的 Bottleneck 模块替换为 CG 模块, 通过在网络的不同阶段利用局部和全局上下文信息来提高识别准确率, 同时保持模型的轻量级特性, 确保边缘信息不丢失, 提高模型对特征信息的获取能力, 增加模型的鲁棒性。改进后的 C2f 如图 5 所示。



1 \* 1 Conv: 1 \* 1 卷积; 3 \* 3 Conv: 3 \* 3 卷积; Contat: 两个卷积结合; BN: 数据归一化层; PReLU: 激活函数; GAP: 全局平均池化; FC: 全连接网络结构; 圆中的  $\cdot$  表示元素乘法;  $f_{loc}(\cdot)$ : 局部特征提取器;  $f_{sur}(\cdot)$ : 周围上下层图像特征提取器;  $f_{joi}(\cdot)$ : 联合特征提取器。

图 4 CG 模块结构图

Fig.4 CG module structure



Conv: 卷积层; Split: 分割; CG block: CG 模块; Concat: 模块融合。

图 5 改进后的 C2f

Fig.5 Improved C2f

**1.2.3 MPDIoU 损失函数** 利用 YOLOv8n 进行巴旦木果实目标检测时, 目标边界框的位置对目标框

宽度计算具有重要的作用。YOLOv8n 中默认的损失函数为 CIoU。CIoU 函数能综合考虑预测框与真

实框重叠面积、中心点距离及宽高比 3 个因素,有利于模型充分学习目标框的特性,但在预测边界框与真实边界框宽高比一致、宽度与高度值不一致时,CIoU 函数的有效性受到影响。为此,本研究中采用具有更高边界框回归效率和精度的 MPDIoU 函数为改进模型的损失函数。MPDIoU 函数通过最小化预测边界框与真实框之间左上和右下点距离,使模型能更好地实现训练、目标检测和图像分割。MPDIoU 损失函数计算公式如下:

$$Loss_{MPDIoU} = 1 - IoU + \frac{d_1^2}{w^2 + h^2} + \frac{d_2^2}{w^2 + h^2} \quad (5)$$

$$d_1^2 = (x_1^{prd} - x_1^{gt})^2 + (y_1^{prd} - y_1^{gt})^2 \quad (6)$$

$$d_2^2 = (x_2^{prd} - x_2^{gt})^2 + (y_2^{prd} - y_2^{gt})^2 \quad (7)$$

式中,  $Loss_{MPDIoU}$  为 MPDIoU 损失函数,  $IoU$  为预测框和真实框的交并比,  $w$ 、 $h$  为输入图像的宽和高,  $d_1$  为预测框的左上点与真实框左上点的直线距离,  $d_2$  为预测框的右下点与真实框右下点的直线距离,  $x_1^{prd}$ 、 $y_1^{prd}$  为预测框左上点的坐标;  $x_2^{prd}$ 、 $y_2^{prd}$  为预测框右下点的坐标;  $x_1^{gt}$ 、 $y_1^{gt}$  为真实框左上点的坐标;  $x_2^{gt}$ 、 $y_2^{gt}$  为真实框右下点的坐标。

MPDIoU 函数不但能充分体现预测框和真实框的重叠程度、中心点距离、宽度和高度偏差等因素,简化了计算过程,还能提高预测框的回归精度和收敛速度,减少预测框的冗余性。

### 1.3 模型系统开发环境及训练参数

本研究中改进 YOLOv8n 模型在 Windows10 操作系统下开发,所用语言为 Python3.8, CPU 型号为 Intel Core i9-13900kf, 运行内存为 64 GB, GPU 型号为 NVIDIA RTX 4090, 显存为 24 GB, 并使用统一计算设备架构(CUDA)和 CUDA 深度学习库(CUDNN)对图形处理器(GPU)进行加速。基于深度学习框架 Pytorch 1.12.1 进行模型训练,具体参数设置如下:学习率 0.01, 动量 0.937, 权重衰减率为 0.000 5, 批量大小 32, 迭代次数 600, 图片尺寸为 300 像素×300 像素。

### 1.4 模型改进措施的消融试验

为验证各个改进模块对于巴旦木果实识别效果的影响,本研究在原始 YOLOv8n 模型的基础上设计 5 组改进措施的消融试验。方案 1 是采用 BiFPN 替换 Neck 中的 PANet 模块;方案 2 是将基本构成单元 C2f 中 Bottleneck 模块替换为 ContextGuide 模块;方案 3 是采用 MPDIoU 损失函数替代 CIoU 函数实现边界框回归;方案 4 是采用 BiFPN、ContextGuide 模块分别替换

PANet 模块和 Bottleneck 模块;方案 5 在采用 BiFPN、ContextGuide 模块分别替换 PANet 模块和 Bottleneck 模块的同时,采用 MPDIoU 损失函数实现边界框回归。

### 1.5 不同模型检测精度比较及识别案例分析

为了进一步说明本研究提出的 YOLOv8n 改进模型的性能,利用其与当前主流卷积神经网络模型(YOLOv8n、YOLOv5n、YOLOv5s、YOLOv7、Faster R-CNN)进行巴旦木果实识别和模型性能的比较。为直观说明不同模型的识别准确率,本研究对不同光线环境背景和不同遮挡情况的 2 个场景巴旦木果实的识别情况进行个例分析。

### 1.6 评价指标

由于本研究中只有巴旦木果实 1 个识别目标,因此本研究的识别属于二分类问题,故选择平均精度值<sup>[34]</sup>(交并比阈值 0.50 的  $mAP_{0.50}$  和交并比阈值 0.50 至 0.95 的  $mAP_{0.50:0.95}$ )、模型参数量、浮点计算量(GFLOPs)4 个指标来评价模型精度和计算效率。

## 2 结果与分析

### 2.1 消融试验结果

消融试验结果如表 1 所示。从表中可以看出,采用原始 YOLOv8n 模型分类的平均精度值  $mAP_{0.50}$  和  $mAP_{0.50:0.95}$  分别为 0.984 和 0.692,参数量和浮点计算量分别为 3.011M 和  $1.8.2 \times 10^9$  次。与原始 YOLOv8n 模型相比,采用 BiFPN 替换 Neck 中的 PANet 模块后(方案 1),由于受到骨干网络池化操作和卷积影响,参数量和浮点计算量分别减少 33.7% 和 12.2%,导致模型的  $mAP_{0.50}$  和  $mAP_{0.50:0.95}$  亦有所下降,分别下降了 0.3 和 2.0 个百分点。将 C2f 中 Bottleneck 模块替换为 ContextGuide 模块后(方案 2),在模型参数和浮点计算量减少的同时,模型识别精度均有所提升,  $mAP_{0.50}$  和  $mAP_{0.50:0.95}$  分别增加 0.2 个百分点和 2.6 个百分点。采用 MPDIoU 损失函数实现边界框回归的方案 3,参数量和浮点计算量略有减少,平均精度值  $mAP_{0.50}$  下降 0.1 个百分点,平均精度值  $mAP_{0.50:0.95}$  提高 0.5 个百分点。同时采用 BiFPN 与 ContextGuide 模块分别替换 PANet 模块和 Bottleneck 模块的方案 4,参数量和浮点计算量分别下降 49.1% 和 28.0%,  $mAP_{0.50}$  和  $mAP_{0.50:0.95}$  均下降 1.0 个百分点。在方案 4 的基础上,进一步利用 MPDIoU 损失函数实现边界框回归后的方案 5,参数量和浮点计算量分别下降 49.3%

和 29.3%, 平均精度值  $mAP_{0.50}$  下降 0.3 个百分点, 而平均精度值  $mAP_{0.50:0.95}$  增加 0.5 个百分点, 即方案 5

在平均精度变化不大的同时, 参数量和浮点计算量得到大幅下降, 达到了模型轻量化的要求。

表 1 消融试验结果

Table 1 Results of ablation test

试验方案	改进方法			参数量 (M)	浮点计算量 ( $\times 10^9$ 次, 1 s)	$mAP_{0.50}$ (%)	$mAP_{0.50:0.95}$ (%)
	BiFPN	ContextGuide	MPDIoU				
原始模型 YOLOv8n	×	×	×	3.011	8.2	98.4	69.2
方案 1	✓	×	×	1.996	7.2	98.1	67.2
方案 2	×	✓	×	2.543	6.7	98.6	71.8
方案 3	×	×	✓	3.006	8.1	98.3	69.7
方案 4	✓	✓	×	1.532	5.9	97.4	68.2
方案 5	✓	✓	✓	1.528	5.8	98.1	69.7

✓ 表示原始模型 YOLOv8n 中相应模块被替换; × 表示原始模型 YOLOv8n 中相应模块未被替换。  $mAP_{0.50}$ : 交并比阈值为 0.50 的平均精度值;  $mAP_{0.50:0.95}$ : 交并比阈值为 0.50 至 0.95 的平均精度值。

## 2.2 不同模型对巴旦木果实的识别性能

不同模型对巴旦木样本集的识别性能如表 2 所示。从表中可以看出, 与本研究提出的 YOLOv8n 改进模型相比, Faster R-CNN 模型和 YOLOv7-tiny 模型参数量多、浮点计算量高, 且模型精度低; YOLOv5s 模型有更高的精度, 但参数量和浮点计算量高, 这将导致模型的识别效率低, 难以满足生产需求; YOLOv5n 模型的参数量和浮点计算量分别增加 63.8% 和 22.4%, 且  $mAP_{0.50}$  和  $mAP_{0.50:0.95}$  分别下降 2.5 个百分点和 11.6 个百分点; 原始 YOLOv8n 模型虽然精度相差不大, 但参数量和浮点计算量高, 即识别效率低。因此, 本研究提出的 YOLOv8n 改进模型既有较高的识别精度, 又有更高的识别效率, 更适合在嵌入式设备中应用。

表 2 不同模型识别性能对比结果

Table 2 Comparison results of recognition performance of different models

模型	参数量 (M)	浮点计算量 ( $\times 10^9$ 次, 1 s)	$mAP_{0.50}$ (%)	$mAP_{0.50:0.95}$ (%)
YOLOv8n	3.011	8.2	98.4	69.2
YOLOv5n	2.503	7.1	95.6	58.1
YOLOv5s	9.111	23.8	99.0	73.3
YOLOv7-tiny	6.014	13.2	94.4	54.2
Faster R-CNN	137.000	273.4	65.9	36.5
YOLOv8n-BCG	1.528	5.8	98.1	69.7

YOLOv8n: 原始模型; YOLOv8n-BCG: 本研究改进的轻量级模型; YOLOv5n、YOLOv5s、YOLOv7-tiny、Faster R-CNN 分别为当前主流卷积神经网络模型。  $mAP_{0.50}$ : 交并比阈值为 0.50 的平均精度值;  $mAP_{0.50:0.95}$ : 交并比阈值是 0.50 至 0.95 的平均精度值。

## 2.3 不同模型的识别案例分析

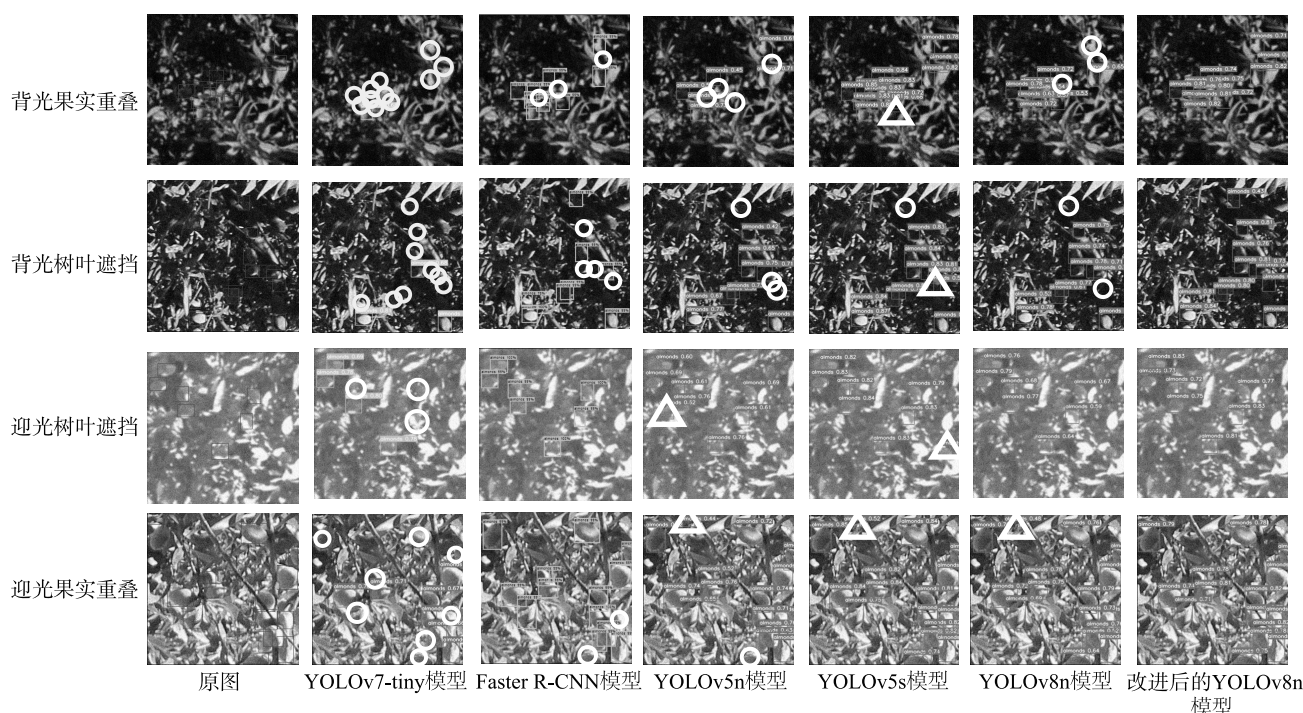
不同模型对不同光线背景(迎光和背光)、不同遮挡情况(树叶遮挡和果实重叠)下的巴旦木果实的识别情况如图 6 所示。从图中可以看出, YOLOv7-tiny 的识别效果最差, 漏检数量不论在哪种情况下都最高。Faster R-CNN 在迎光树叶遮挡的情况下, 果实识别率达到最佳, 但其他情况下均有较严重的漏检。YOLOv5n 能够较好识别巴旦木果实, 但在光线不好的情况下易出现错检、漏检, 且重叠果实的识别性能较差。YOLOv5s 能够较好地识别果实, 且背光条件下识别准确率高, 但仍存在错检和漏检情况。YOLOv8n 在迎光树叶遮挡的情况下未出现漏检、错检现象, 但光线较差时, 对重叠果实的识别准确率较低, 漏检数量多, 在果实重叠时存在误检现象。本研究建立的 YOLOv8n 改进模型对于枝叶遮挡、果实重叠、光线不佳的不良情况拥有很高的识别能力, 无漏检、错检的果实, 在实际的果园环境下拥有更优异的检测效果。

综上所述, 经过算法优化后得到的 YOLOv8n-BCG 模型在巴旦木实例检测中能够达到更好的检测效果, 在保证识别准确度的前提下, 能有效地防止误检和漏检情况发生。

## 3 结 论

(1) 本研究在 YOLOv8n 模型的基础上, 利用 BiFPN 模块代替 Neck 中的 PANet 模块、Context-Guide 模块代替骨干部分 C2f 结构单元中的 Bottleneck 模块、MPDIoU 损失函数代替 CIoU 损失函数,





圆圈为漏检果实,三角形为误检果实,正方形均为检测正确的果实。

图6 不同模型的识别效果

Fig.6 The recognition effect of different models

构建了 YOLOv8n 改进模型——YOLOv8n-BCG 模型。结合公开的巴旦木影像数据集 ACFR-Orchard-Fruit-Dataset 对模型进行了比较,发现改进后的模型能有效解决识别目标的重叠问题,突出重要层次的信息,减少冗余信息、模型参数及浮点计算量,提高训练速度和识别精度,能较好地实现果园环境下的巴旦木果实轻量化识别。

(2)消融试验结果表明不同的改进模块对模型的参数量、浮点计算量和精度有不同的影响,综合使用3个模块代替方案后模型参数和浮点计算量比原模型分别下降49.3%和29.3%,平均精度值( $mAP_{0.50:0.95}$ )提高0.5个百分点,效果最好。不同模型比较结果亦显示改进后的 YOLOv8n-BCG 模型不同光线背景、不同遮挡情况下漏检率、误检率最低。上述结果说明本研究构建的 YOLOv8n-BCG 模型非常适合于巴旦木果实智能采摘装备的研发。

#### 参考文献:

- [1] 石鑫,杨豫新,牛长河,等.滚筒刷式巴旦木脱青皮机的设计与试验[J].西北农林科技大学学报(自然科学版),2023,51(12):143-154.
- [2] 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.
- [3] 刘冬.数字化与智能化在农业机械工程设计中的应用[J].中国农业资源与区划,2023,44(7):48-65.
- [4] 陈礼鹏,穆龙涛,刘浩洲,等.基于猕猴桃果萼图像的多目标果实识别方法[J].计算机工程与设计,2018,39(6):1738-1744.
- [5] ABDULLAH M S, HASAN R M A, RAHMAN A N, et al. Versatile recognition of graphene layers from optical images under controlled illumination through green channel correlation method[J]. Nanotechnology, 2023, 34(44):1899-1904.
- [6] 贾伟宽,赵德安,阮承治,等.苹果采摘机器人夜间图像降噪算法[J].农业工程学报,2015,31(10):219-226.
- [7] 李斌,韩昭洋,王秋,等.基于高光谱成像技术的枇杷碰伤等级检测研究[J].光谱学与光谱分析,2023,43(6):1792-1799.
- [8] MAR A, HILMY B, SERGIO V, et al. Object detection and tracking on UAV RGB videos for early extraction of grape phenotypic traits[J]. Computers and Electronics in Agriculture, 2023, 211(20):347-353.
- [9] CHEN B, SHI S, CHEN B, et al. True color 3D imaging optimization with missing spectral bands based on hyperspectral LiDAR[J]. Optics Express, 2021, 29(13):20406-20422.
- [10] 李昕,李立君,高自成,等.改进类圆随机 Hough 变换及其在油茶果实遮挡识别中的应用[J].农业工程学报,2013,29(1):164-170.
- [11] 张彦斐,刘茗洋,宫金良,等.基于两级分割与区域标记梯度



- Hough 圆变换的苹果识别[J]. 农业工程学报, 2022, 38(19): 110-121.
- [12] 刘安旭, 黎向锋, 刘晋川, 等. 改进卷积空间传播网络的单目图像深度估计[J]. 电子测量技术, 2021, 44(23): 78-85.
- [13] 吕伟, 宋轩, 杨欢. 基于深度学习和多源遥感数据的玉米种植面积提取[J]. 江苏农业科学, 2023, 51(23): 171-178.
- [14] 鲍彤, 罗瑞, 郭婷, 等. 基于BERT字向量和TextCNN的农业问句分类模型分析[J]. 南方农业学报, 2022, 53(7): 2068-2076.
- [15] 翟先一, 魏鸿磊, 韩美奇, 等. 基于改进YOLO卷积神经网络的水下海参检测[J]. 江苏农业学报, 2023, 39(7): 1543-1553.
- [16] 向俊, 严恩萍, 姜稼伟, 等. 基于全卷积神经网络和低分辨率标签的森林变化检测研究[J]. 南京林业大学学报(自然科学版), 2024, 48(1): 187-195.
- [17] 阮子行, 黄勇, 王梦, 等. 基于改进卷积神经网络的番茄品质分级方法[J]. 江苏农业学报, 2023, 39(4): 1005-1014.
- [18] DEBANIRANJAN M, NIVA D, KUMAR K M. Deep neural network based fruit identification and grading system for precision agriculture[J]. Proceedings of the Indian National Science Academy, 2022, 88(2): 228-239.
- [19] TANG Z, LIU J J, CHEN Z Y, et al. Improved Pest-YOLO: real-time pest detection based on efficient channel attention mechanism and transformer encoder[J]. Ecological Informatics, 2023, 78(14): 891-898.
- [20] 郭明月, 刘雅晨, 李伟夫, 等. 基于视频跟踪算法的果园猕猴桃产量实时预估[J]. 农业机械学报, 2023, 54(6): 178-185.
- [21] 王大方, 刘磊, 曹江, 等. 基于空洞空间池化金字塔的自动驾驶图像语义分割方法[J]. 汽车工程, 2022, 44(12): 1818-1824.
- [22] 李恒, 南新元, 高丙朋, 等. 一种基于GhostNet的绿色类圆果实识别方法[J]. 江苏农业学报, 2023, 39(3): 724-731.
- [23] SUE R. GhostNets: go global, and local[J]. Arlink, 2013, 33(2): 70-71.
- [24] 赵升, 赵黎. 基于双向特征金字塔和深度学习的图像识别方法[J]. 哈尔滨理工大学学报, 2021, 26(2): 44-50.
- [25] 储鑫, 李祥, 罗斌, 等. 基于改进YOLOv4算法的番茄叶部病害识别方法[J]. 江苏农业学报, 2023, 39(5): 1199-1208.
- [26] 李运寰, 闻继伟, 彭力. 高帧率的轻量级孪生网络目标跟踪[J]. 计算机科学与探索, 2022, 16(6): 1405-1416.
- [27] 王琛, 林威, 胡良鹏, 等. 分体式飞行汽车全自主对接引导系统设计与验证[J]. 浙江大学学报(工学版), 2023, 57(12): 2345-2355.
- [28] 孔维刚, 李文婧, 王秋艳, 等. 基于改进YOLOv4算法的轻量化网络设计与实现[J]. 计算机工程, 2022, 48(3): 181-188.
- [29] LIN C L, WANG T Y, DONG S Y, et al. Hybrid convolutional network combining 3D depthwise separable convolution and receptive field control for hyperspectral image classification[J]. Electronics, 2022, 11(23): 3992-3997.
- [30] 代云, 卢明, 何婷, 等. 基于改进型YOLO的密集环境下槟榔果实的快速识别方法[J]. 食品与机械, 2023, 39(4): 83-88.
- [31] YANG S Z, WANG W, GAO S, et al. Strawberry ripeness detection based on YOLOv8 algorithm fused with LW-Swin Transformer[J]. Computers and Electronics in Agriculture, 2023, 215(5): 1234-1240.
- [32] TAN M, PANG R, LE Q. EfficientDet: scalable and efficient object detection[C]//IEEE. 2020 IEEE/CVF conference on computer vision and pattern recognition (CVPR). Washington: IEEE Computer Society, 2020: 10778-10787.
- [33] WU T Y, TANG S, ZHANG R, et al. CGNet: a light-weight context guided network for semantic segmentation[J]. IEEE Transactions on Image Processing, 2021, 30: 1169-1179.
- [34] HENDERSON P, FERRARI V. End-to-end training of object class detectors for mean average precision[C]//LAI S H, LEPETIT V, NISHIMO K, et al. Computer Vision-ACCV2016. Cham, SW: Springer, 2016: 198-213.

(责任编辑: 石春林)