



重庆工商大学学报(自然科学版)

Journal of Chongqing Technology and Business University(Natural Science Edition)

ISSN 1672-058X,CN 50-1155/N

《重庆工商大学学报(自然科学版)》网络首发论文

题目：基于改进 YOLOv5s 的果园苹果检测算法
作者：王彦辉，汪军
收稿日期：2024-04-02
网络首发日期：2024-05-07
引用格式：王彦辉，汪军. 基于改进 YOLOv5s 的果园苹果检测算法[J/OL]. 重庆工商大学学报(自然科学版). <https://link.cnki.net/urlid/50.1155.N.20240506.1225.002>



网络首发：在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

出版确认：纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

基于改进 Yolov5s 的果园苹果检测算法

王彦辉, 汪军

安徽工程大学 计算机与信息学院, 安徽 芜湖 241000

摘要: 【目的】针对果园环境下现有的苹果检测识别精度低以及模型大、计算量高的问题, 提出一种基于 Yolov5s 的改进型果园苹果检测算法。【方法】首先在 Yolov5s 算法中使用 Soft_NMS 来替换模型原有的 NMS, 改变原模型对预测框的处理方式, 从而减少果园环境下苹果遮挡重叠带来的错检、漏检情况, 提升检测精度。然后再使用 OTA 来优化原模型的标签分配方式, 将标签分配看作最优传输问题, 并完整利用上下文信息减少模糊框的数量, 更好地处理苹果密集遮挡问题, 进一步提高模型对果园苹果的检测性能。最后使用 FasterNet 块替换主干网络中的卷积模块作为新的特征提取网络, 降低模型的参数量和计算量, 从而实现模型的轻量化。【结果】实验结果显示, 在果园苹果数据集上, 改进的算法模型相对于原模型, mAP 提升了 2.8%, 同时参数量和计算量分别下降了 21%和 29%。与其他同系列主流检测模型相比, 改进模型具有更高的检测精度, 更低的参数量和计算量。【结论】改进的模型更适合果园场景下的苹果检测, 并可作为一种解决方案来参考、应用于相关领域。

关键词: 苹果识别; Yolov5s; 非极大值抑制算法; 最优传输分配; FasterNet

中图分类号: TP302.7

文献标识码: A

Orchard apple detection algorithm based on improved Yolov5s

WANG Yanhui, WANG Jun

Anhui Polytechnic University, School of Computer and Information Science, Wuhu 241000, China

Abstract: 【Objective】 The objective of this study is to propose an improved orchard apple detection algorithm based on Yolov5s in order to solve the problems of low accuracy, large model and high computational cost in orchard environment. 【Method】 Firstly, the Soft_NMS was used in the Yolov5s algorithm to replace the original NMS of the model, and the processing method of the prediction box of the original model was changed, so as to reduce the false detection and missed detection caused by the overlapping of apple occlusion in the orchard environment, and improve the detection accuracy. Then, OTA is used to optimize the label assignment method of the original model, which is regarded as the optimal transmission problem, and the context information is fully used to reduce the number of fuzzy boxes, better deal with the problem of dense occlusion of apples, and further improve the detection performance of the model for orchard apples. Finally, the convolution module in the backbone network is replaced by FasterNet blocks as a new feature extraction network, which reduces the number of parameters and computations of the model, so as to realize the lightweight of the model. 【Result】 The experimental results show that compared with the original model, the mAP of the improved algorithm model is increased by 2.8%, and the number of parameters and computational costs are reduced by 21% and 29%, respectively. Compared with other mainstream detection models of the same series, the improved model has higher detection accuracy, lower number of parameters and lower calculation cost. 【Conclusion】 The improved model is more suitable for apple detection in orchard scenarios, and can be used as a solution for reference and application in related fields.

Key words: apple recognition; Yolov5s; Non-Maximum Suppression algorithm; Optimal Transport Assignment; FasterNet

1 引言

根据 2021-2022 年的数据, 全国苹果产量约为 4597.3 万吨, 占比为 56.4%^[1]。这一数字显示苹果作

为我国重要的经济作物的地位。苹果的生产消费水平在世界上处于领先地位, 并且进出口贸易非常活跃。采摘是保证苹果经济的重要环节, 对后续的加工、运

收稿日期: 2024-04-02。

作者简介: 王彦辉 (2000-), 男, 安徽芜湖人, 硕士研究生, 从事计算机视觉研究。

通讯作者: 汪军 (1975-), 男, 教授, 从事智能感知、图像处理与模式识别研究.Email: wangjun@ahpu.edu.cn

输和销售环节有直接影响。然而，由于我国城镇化进程和人口老龄化的加剧，阶段性劳动力不足。这意味着依靠人工采摘将变得更加困难和费时，并增加人工成本和其他成本。因此，使用机械化代替人工采摘是社会发展的必然要求^[2]。随着传感器技术的发展，苹果采摘机器人得到迅速发展。对于采摘机器人而言，准确的苹果检测和识别是后续采摘的基础和关键^[3]。只有准确检测和识别苹果，才能实现高效的采摘。因此，研究苹果的检测识别技术对于推动苹果采摘工作向自动化发展具有重要意义^[4]。这项技术的应用可以取代手工采摘的费时费力，减少果农的劳动量，提高生产效率，促进农业自动化发展。

随着深度学习技术的不断发展，卷积神经网络不仅在其他领域发挥出重要作用^[5]，也在苹果检测领域发挥出巨大的优势。深度学习神经网络通过大规模数据训练，学习到丰富的苹果特征，如形状、颜色和纹理^[6]。这有助于提高苹果的检测识别精度，并能自动从图像和视频中检测和定位苹果^[7]。目前的目标检测算法可以分为两类：一种是一阶段检测模型，以 Yolo（You Only Look Once）系列为代表^[8-10]；另一种是两阶段检测模型，以 RCNN 为代表（RCNN^[11]、Fast-RCNN^[12]、Faster-RCNN^[13]、Cascade R-CNN^[14]）。本文采用了 Yolo 系列的 Yolov5s 模型，相比两阶段检测模型，虽然精度略有降低，但具有更快的检测速度，适应实时性要求，更适用于实际检测场景。而在苹果的自然生长过程中，会存在果实间相互遮挡重叠和树枝、树叶遮挡的问题^[15]，这些因素都会影响目标检测的精确度。薛月菊等^[16]使用改进的 Yolov2 算法来检测遮未成熟的芒果，虽达到了一定的检测精度和检测速度，但仍会有漏检的问题。LU 等^[17]提出了一种改进的 Yolov4 模型，用于检测果园环境下不同成熟度的苹果，模型虽达到了一定的检测精度，但也增加了模型参数量和计算量。周桂红等^[18]通过改进的 Yolov4 模型可以实现对高分辨率全景图中的苹果进行识别，但在处理遮挡问题方面和模型大小上仍有改进空间。张境峰等^[19]提出的 Des-yolov4 算法解决了复杂环境中苹果的误检问题，但在检测遮挡重叠的苹果仍有不足。高宽祥等^[20]改进的 Yolov5 算法提高了对果园水果的检测能力，但未能处理果实重叠遮挡的问题也未能降低模型大小。

通过上述说明可知，虽然苹果检测技术在不断的进步，检测精度在不断上升，但面对复杂的果园环境中遮挡重叠等问题仍感到力不足，仍有错检漏检情况的发生，同时模型仍然较大，难以在资源受限的环境下部署应用。这也为后续的研究提供了努力的方向。

因此本文继续针对遮挡重叠问题导致检测精度低，模型参数量大、计算量高等问题进行研究，实现模型精度的提升和轻量化。首先针对性构建数据集，然后在 Yolov5s 模型上通过替换 NMS 为 Soft_NMS，改变原模型对预测框的处理方式，从而减少果园环境下苹果遮挡重叠带来的错检、漏检情况，提升检测精度。同时引入 OTA 方法，将标签分配看作最优传输问题，并完整利用上下文信息减少模糊框的数量，进一步提高对密集遮挡目标的检测能力。最后使用 FasterNet 块替换原模型主干网络的卷积模块实现轻量化。

2 网络结构设计 with 算法改进

2.1 Yolov5s 网络结构

Yolov5 是 Yolo 系列中的一种基于深度学习的目标检测算法。相较于之前的 Yolo 系列版本，它采用了新的设计和优化来提高检测性能和检测速度，引入数据增强技术，如随机缩放、裁剪和翻转，以提升模型的鲁棒性和泛化能力。它还采用了多尺度检测策略，通过在多个尺度下进行图像检测，增强对不同大小目标的检测能力。此外，Yolov5 还采用了自适应训练策略，根据目标的难易程度动态调整训练过程，以提高模型在复杂场景中的性能。Yolov5 系列包括 Yolov5m、Yolov5l 和 Yolov5x 等不同的变体，相比其他变体，Yolov5s 具有较小的模型大小和更快的推理速度。此外，Yolov5s 在精度上也表现不错，因此更适合在资源受限的设备上部署，为后续的实际应用奠定了基础。

Yolov5s 主要由主干网络的 CSPDarknet53（Cross Stage Partial）、颈部网络的 PANet（Path Aggregation Network）、头部网络的检测头三部分组成，其具体结构如图 1 所示。



图1 Yolov5s 网络结构

Fig. 1 Yolov5s network structure

主干网络: Yolov5s 使用 CSPDarknet53 作为主干网络。CSPDarknet53 由多个卷积层和残差块组成，能够在不同尺度下提取和组合特征。通过逐步提取图像的特征，从低级特征逐渐转化为高级特征，如边缘、纹理、语义和抽象概念，这为后续的目标检测任务提供特征信息。因此使用 CSP 结构有助于提高网络的效率和准确性。

颈部网络: 采用 PANet 作为连接主干网络和头部网络的中间层，用于跨层级特征融合。PANet 通过上下文感知和信息融合来提高目标检测的精度，它能够处理来自不同层级的特征图。

头部网络: Yolov5s 的头部结构由一系列卷积层和全连接层组成。头部结构负责生成目标检测的边界框和类别预测。Yolov5s 采用了一种简化的头部结构，使用 3 个不同尺度的输出层来预测不同大小的目标。而检测是模型的最后一部分，它利用一些后处理技术（如非极大值抑制，NMS）将网络输出的特征图转换为目标检测结果 包含边界框的位置坐标和类别概率。

2.2 Yolov5s 算法改进

为了应对果园的复杂环境中 Yolov5s 模型的检测精度不高的问题，本文对 Yolov5s 后处理的部分算法进行改进优化，替换传统的非极大值抑制（Non-Maximum Suppression, NMS）算法为 Soft_NMS 以及使用 OTA 优化原有的标签分配策略，提出了一项基于 Yolov5s 的复杂果园环境检测识别算法，同时使

用 FasterNet 替换主干网络达到轻量化的目的。

2.2.1 Soft_NMS

果园环境中的苹果在生长的过程中往往会密集地分布在树上不同的区域，苹果间会相互重叠遮挡，原模型中的 NMS 算法会将与最高置信度框重叠的其他框抑制掉，导致遮挡重叠的苹果被错误地抑制掉，从而出现错检、漏检的情况使得检测精度下降。

在 Yolov5 原模型中的 NMS 步骤为：

首先，对每个类别的预测框按照置信度高低进行排序，将具有较高置信度的框放置在前面。

从置信度最高的预测框开始，计算其与其他预测框之间的重叠度。而交并比（Intersection Over Union, IOU）是用来衡量边界框与边界框间（或区域与区域间）重叠程度的指标。它通过计算两个框的交集面积与并集面积的比值来度量它们的重叠程度。

如果某个预测框与置信度最高的框之间的重叠程度（IOU）超过预先设定的阈值，那么会将该预测框抑制掉，因为它很可能对应着同一个目标物体。然后不断的对余下的预测框进行判断和抑制，直到所有的预测框都被处理完毕。NMS 可以描述为如下的公式：

$$S_i = \begin{cases} S_i, iou(M, b_i) < N_t \\ 0, iou(M, b_i) \geq N_t \end{cases} \quad (1)$$

最后，经过 NMS 处理后，每个类别都只保留了置信度最高的框 并移除与其重叠程度较高的冗余框。

Soft_NMS（Soft Non-Maximum Suppression）在原来的 NMS 的基础上引入一个衰减函数，加入惩罚机制，以概率的形式来降低与最高置信度框重叠的其他框的得分，而不再是直接将得分大于阈值的全置为 0。加入线性衰减函数的 Soft_NMS 公式为

$$S_i = \begin{cases} S_i, iou(M, b_i) < N_t \\ S_i(1 - iou(M, b_i)), iou(M, b_i) \geq N_t \end{cases} \quad (2)$$

其中 S_i 为第 i 个预测框的得分， N_t 为阈值， b_i 表示第 i 个预测框， M 表示得分最高的预测框。

引入高斯衰减函数的公式为：

$$S_i = S_i e^{-\frac{iou(M, b_i)^2}{\sigma}}, \forall b_i \notin D \quad (3)$$

式（3）中，当 σ 取值为 0.5 时衰减幅度最大，预测框筛选速度最快。

引入衰减函数后会降低检测框的得分，而不再直接设置为 0，这就可以保留某些得分高的预测框，从而在后续处理中也有被重新视为正确预测框的可能性。通过这种软性的抑制方式，能够有效缓解模型错

检漏检的问题。

2.2.2 最优传输分配

最优传输分配 (Optimal Transport Assignment, OTA) 是基于最优传输理论的一种方法, 它将标签分配视为最优传输问题。它通过计算网络生成的预测框与真实的人工标注框之间的距离矩阵, 并利用线性规划方法找到最优分配方案, 从而有效地匹配预测框与真实标注框之间的对应关系, 使总距离最小化。原 YOLOv5s 网络主要关注于单个对象的最优分配策略, 而很少从全局的角度关注理解上下文信息。对于本实验中密集苹果目标来说, 遮挡重叠的模糊框就难以完整利用上下文信息, 从而出现对目标错检漏检的问题。在目标检测中, 遮挡的本质原因在于一些预测框无法与真实标注框完全匹配的情况, 以及一个真实标注框对应多个预测框的情况。因此通过使用 OTA 来优化标签分配策略, 可以很好地解决这种不完整匹配和多对一匹配的情况, 并结合全局信息更准确地度量预测框与真实标注框之间的差异从而有效解决苹果的密集遮挡带来检测精度低的问题。

OTA 的最优传输理论方法应用在苹果的目标检测中, 将 GT (Ground Truth) 视为供应商, 预测框视为需求者, 然后在 GT 和所有预测框之间计算运输成本, 并通过寻找最低运输成本的合适映射关系来进行标签分配。假定 GT 的数量为 m , 预测框的数量为 n , 将标签从 GT 运输到预测框的成本 c^{fg} 为:

$$c_{ij}^{fg} = L_{cls}(P_j^{cls}(\theta), G_i^{cls}) + \alpha L_{reg}(P_j^{box}(\theta), G_i^{box}) \quad (4)$$

其中 θ 是模型的参数, L_{cls} 表示分类损失 Focal loss, L_{reg} 为回归损失 IOU loss, G_i 为真实的标签, α 是可调节的超参数, P_j^{cls} 与 P_j^{box} 表示的是 a_j 的预测类别和检测框。

在训练过程中, 除了正标签分配之外, 许多锚框被视为负样本。而一个负样本的运输成本为:

$$c_{ij}^{fg} = L_{cls}(P_j^{cls}(\theta), \emptyset) \quad (5)$$

其中, \emptyset 表示的背景类。

通过上述方法计算每个 GT 和每个预测框的成本, 以及计算每个 GT 需要的预测框数量 k , 就可以通过使用 Sinkhorn-Knopp 算法求得最小成本的传输 π^* 。具体的算法公式为:

$$\begin{aligned} \pi^* &= \pi_{ij} : \\ \min_{\pi} & \sum_{i=1}^m \sum_{j=1}^n c_{ij} \pi_{ij} \\ \sum_{i=1}^m \pi_{ij} &= d_j, \sum_{j=1}^n \pi_{ij} = s_i, \sum_{i=1}^m s_i = \sum_{j=1}^n d_j \\ \pi_{ij} &\geq 0, i=1, 2, \dots, m, j=1, 2, \dots, n \end{aligned} \quad (6)$$

其中 s_i 表示 GT 拥有的正标签数, d_j 表示预测框对应的标签。

最终计算得到最小传输成本 π^* 后, 根据每个预测框分配最多标签的 GT 来实现标签的最优分配。

2.2.3 FasterNet 轻量化

轻量化的模型不仅能够在保持相对较高的检测性能的同时, 还能够减少模型的计算复杂度, 使其能够在资源有限的环境中高效运行, 因此模型的轻量化对后面的实际应用十分重要。

FasterNet 是一种轻量化的特征提取网络, 它基于部分卷积 (PConv) 和逐点卷积 (PWConv)。FasterNet 由四个阶段组成, 每个阶段都包含一个嵌入层或合并层, 以及一系列 FasterNet 块。嵌入层用于空间下采样和通道数量扩展, 而合并层则将不同阶段的特征融合起来。FasterNet 块是 FasterNet 的核心部件, 它由一个 PConv 层和两个 PWConv 层组成, 并以倒置残差结构的形式实现。PConv 层可以有效地提取空间特征, 而 PWConv 层可以进一步提高网络的效率。

FasterNet 块的结构如图 2 所示, 它的主要实现为: 输入特征图后进行 PConv 操作, 然后再接两个 PWConv, 第一的目的是减少特征图的通道数, 从而减低计算成本, 第二个是用于恢复特征图的通道数, 最后结合输入与输出特征图得到最终的输出结果。

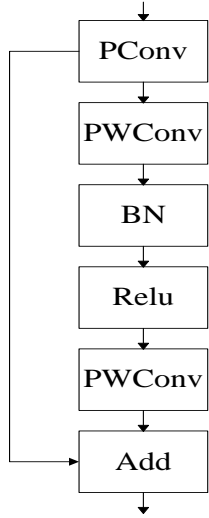


图2 FasterNet 块
Fig. 2 FasterNet block

在本文中通过使用 FasterNet 块来替换 Yolov5s 主干网络中原有的卷积模块来降低原网络的参数和计算量，从而达到轻量化的目的，改进的主干网络如图 3 所示。

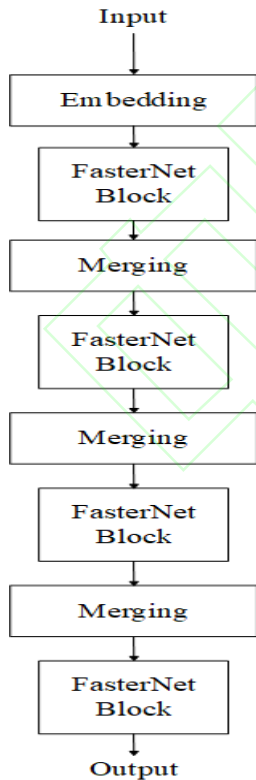


图3 改进的主干网络
Fig. 3 Improved backbone networking

2.3 改进 Yolov5s 苹果检测

在经过改进后的苹果检测流程如图 4 所示，输入图像后先对图像进行调整大小、归一化等预处理操作，然后进行特征提取，并根据提取的特征图生成锚框。

然后根据 OTA 算法分配标签，再将分配的真实边界框与相应的锚框匹配并预测每个物体的类别概率，最后再经过 Soft_NMS 后输出结果。

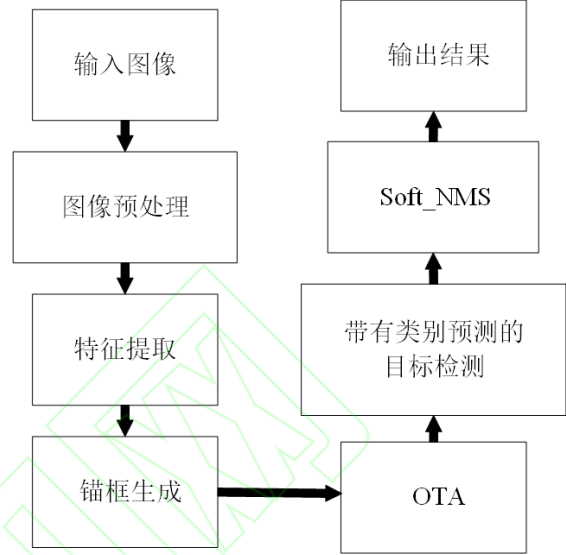


图4 苹果检测算法流程图

Fig. 4 Flow chart of Apple's detection algorithm

3 仿真实验与结果分析

3.1 数据集构建与处理

由于果园内的情况是复杂多样的，而目前公开的数据集大多数较为单一，单一的数据集难以包含各种重叠遮挡的情况，使得模型无法学习到更多的有效特征，不仅模型精度难以提升，还会导致错检漏检情况频繁出现，因此针对遮挡情况构建果园苹果数据集是非常有必要的。本文使用的数据集图像均从真实果园收集并包含枝、叶遮挡、果实重叠，混合遮挡等多种遮挡情况，共得到分辨率为 960×1280 图像 1357 张并进行人工标注。标注使用标注工具 Labelimg 进行标注，按实际情况分为无遮挡，枝叶遮挡，果实遮挡，果实枝叶混合遮挡等类别。收集到图像后会先对图像去近似、去模糊，以提高图像的质量和可分析性，然后对图像进行批量旋转，扩充训练数据集，增加样本的多样性和数量，通过旋转可模拟不同角度和视角下的图像变化，使模型能够更好地适应不同的旋转变换。在实验开始前还会通过调整对比度，曝光等手段来丰富数据集以达到模拟实际的果园场景的效果。实验中所使用的数据集将划分出训练集、测试集和验证集，划分的比例为 8: 1: 1。

3.2 实验环境

本文实验使用了一台计算机进行网络训练，其中 CPU 为 i7-5820K，显卡为 GeForce GTX TITAN X，内存为 32G，硬盘容量为 1.2TB，操作系统为 Ubuntu 20.04。在模型训练过程中，使用了动量值 0.937，衰

减值 0.0005，学习率 0.01，并设置了预热轮数为 3，批次大小为 8，总共进行了 100 轮的训练。

3.3 训练结果

Yolov5 使用三种损失：分类损失、定位损失和置信度损失来衡量模型的性能。Yolov5 中损失函数的具体组成形式如式（7）所示。

$$loss = \lambda_c * loss_{box} + \lambda_{cls} * loss_{cls} + \lambda_{conf} * loss_{conf} \quad (7)$$

$$loss_{box} = 1 - IOU \quad (8)$$

$$loss_{cls} = -\sum_i (p_i * \log(C_i) + (1 - p_i) * \log(1 - C_i)) \quad (9)$$

$$loss_{conf} = -\sum_i (C_i * \log(p_i) + (1 - C_i) * \log(1 - p_i)) \quad (10)$$

其中式（8）为定位损失计算公式，式（9）为分类损失计算公式，式（10）为置信度损失计算公式。 p_i 为模型预测第 i 个类别的概率， C_i 表示第 i 个类别是否为真实框的指示变量， λ_c 、 λ_{cls} 、 λ_{conf} 为平衡因子。

由图 5 中可以看到这三种损失函数均在前 20 轮内快速下降，在后 80 轮内缓慢下降并趋于稳定。因此本文的研究将训练 100 轮结束后的模型作为最后的输出模型。

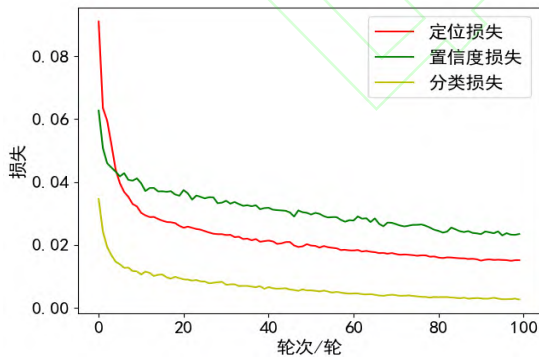


图 5 损失函数变化图

Fig. 5 Diagram of the change in the loss function

3.4 评价指标

在苹果的检测中通常使用均值平均精度（mean Average Precision）、精确率（Precision）及召回率（Recall）作为性能指标。均值平均精度使用 IOU（Intersection over Union）阈值为 0.5(步长为 0.05)的均值进行计算，即用 mAP@0.5 来衡量目标模型的准

确性。下面详细介绍一下各类指标：

（1）均值平均精度。在目标检测任务中，每个类别都有一个平均精度（Average Precision）值，它衡量了模型在单个类别上的准确性。AP 值是通过计算精确率-召回率曲线下的面积得到的。而 mAP（Mean Average Precision）则是所有类别 AP 的均值，它提供了对模型在多个类别上的整体性能的评估。mAP 的取值范围是 0 到 1，数值越高表示模型在多个类别上的性能越好。

（2）参数量。参数量是指模型中需要进行学习和优化的可调整参数的数量。参数量的大小对模型的容量和复杂度产生重要影响。参数量越大时，模型能够更好地拟合细节和复杂性，但也容易过拟合训练数据。相反，参数量较小的模型可能更简单、更具泛化能力，但可能无法捕捉到数据中的复杂模式。通常情况下，参数量和数据及训练所需的计算资源是成正相关的，因此大参数量不适合资源受限的环境。

（3）计算量。计算量即浮点运算的次数，用来衡量算法的时间复杂度。通常用 FLOPs（Floating Point Operations per Second）作为单位进行衡量。模型的计算量与其参数量和数据量是息息相关的。通常情况下，模型的计算量随着参数量和数据量的增加而增加。计算量的多少可以反映模型的计算复杂度、计算资源需求以及模型的训练和推理速度。因此在实际的应用中需要综合考虑计算量和计算资源之间的平衡，以满足任务需求。

3.5 实验结果与分析

3.5.1 消融实验

为更好的说明改良算法的有效性，本文采用消融实验来进行比较说明。基础模型为 Yolov5s 模型，替换原模型中的 NMS 算法为 Soft_NMS，使用最优传输分配 OTA 优化标签分配方法，以及用 FasterNet 替换主干网络。以下所有的模型都使用同样的数据集来进行训练和验证，并在验证集上得到实验结果。使用 mAP@0.5，参数量以及计算量作为指标评价。表 1 为模型改进的数据对比情况。

表 1 消融实验

Table 1 Ablation experiments

模型	mAP@0.5	参数量	FLOPs	fps
Yolov5s	0.785	7.02	16	96
Yolov5s+Soft_NM	0.802	7.02	16	72
S				
Yolov5s+Soft_NM	0.817	7.02	16	78
S+OTA				

Yolov5s+Soft_NM	0.813	5.55	11.3	93
S+OTA+FasterNet				

由表 1 分析可以得到，使用改进方案后，模型的精度明显提高。改进后的网络比原模型 mAP@0.5 上涨了 2.8%，在加入 FasterNet 后虽然平均精度略微下降，但参数量和计算量分别下降了 21% 和 29%，且保证了检测速度。这说明了改进模型的精度和轻量化的优势，降低了模型部署难度，这为后续的实际应用奠定了基础。

热力图能够可视化场景中物体的热能分布，通过不同颜色来表示不同温度区域。其中，红色代表高温区域，蓝色代表低温区域。通过这种的颜色变化可清楚展现物体的分布情况。如图 6 所示，热力图展示了苹果检测结果。可以看出改进后的模型在检测遮挡苹果方面的准确度更高。

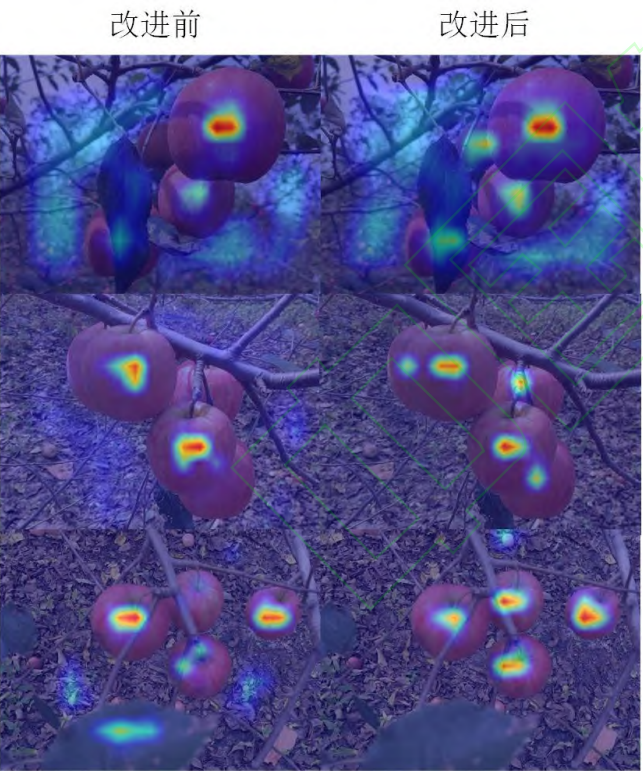


图 6 苹果检测热力图

Fig. 6 Apple detection heatmap

3. 5. 2 对比实验

为了进一步比较分析所提出改进的果园苹果检测模型的性能，将改进的模型与原模型，Yolov3，Yolov6，Yolov7，Yolov8 等 Yolo 系列的主流单目标检测模型进行比较。同样以 mAP@0.5，精确率以及召回率等作为评价指标进行分析评价。表 2 展示了不同模型间的比较结果。

表 2 对比实验

Table 2 Comparative experiments				
模型	mAP@0.5	参数量	FLOPs	fps
Yolov3-tiny	0.749	12.2	19.1	48
Yolov5s	0.785	7.02	16	96
Yolov5s+Soft_NM	0.813	5.55	11.3	93
S+OTA+FasterNet				
Yolov6s	0.756	16.31	45.2	88
Yolov7-tiny	0.794	6.02	13.2	98
Yolov8s	0.778	11.14	28.7	88

通过表 2 的实验数据可知，本文所提出的改进算法在众多算法中表现最好，相较于其他 Yolo 系列的算法模型精确度有着较为明显的提升。mAP@0.5 分别较 Yolov3、Yolov6、Yolov7、Yolov8 高出 9.1%、5.7%、1.9%、3.5%，同时参数量和计算量也低于他们，并保证了检测速度。因此本文所提出的算法最能够适应果园内的复杂场景，以及适合后续的模型部署，这为实际应用提供了参考方案。

4 结论

针对复杂果园场景下现有模型识别精度低以及模型大、计算量高的局限性，本文提出了一种改进的基于 Yolov5s 的果园苹果检测算法，旨在解决遮挡重叠问题提高检测性能以及实现模型的轻量化。本文通过自主收集相关数据集来训练改进的模型并取得了一定的效果。改进的模型通过使用 Soft_NMS 算法来替换原模型中的 NMS 算法，改变了原有的对预测框的处理方式，提升对果园内重叠遮挡的苹果的检测效果。使用 OTA 方法优化标签分配策略，将标签分配看作最优传输问题，并完整利用上下文信息减少模糊框的数量，更好地处理苹果密集遮挡问题，进一步提升了模型的检测性能。以及通过使用 FasterNet 块替换主干网络的卷积模块作为新的特征提取网络，降低模型的参数量和计算量，实现轻量化。本文的算法与基础模型相比均值平均精度提升了 2.8%，参数量和计算量分别下降了 21% 和 29%。与其他同系列主流检测模型相比，改进模型具有更高的检测精度，更低的参数量和计算量。因此本算法可作为一种解决方案来参考、应用于相关领域。

参考文献:

[1] 王壹. 中国苹果产量稳居世界第一 [N]. 农民日报, 2023-11-18(7).
WANG Yi. China's apple output ranks first in the world[N].Farmer's Daily,2023-11-18(7)

- [2] 陈青, 殷程凯, 郭自良, 等. 苹果采摘机器人关键技术研究现状与发展趋势[J]. 农业工程学报, 2023, 39(4): 1-15.
CHEN Qing, YIN Chengkai, GUO Ziliang, et al. Current status and future development of the key technologies for apple picking robots[J]. Transactions of the Chinese Society of Agricultural Engineering, 2023, 39(4): 1-15
- [3] 颜清华, 杜玉珠, 李亚芹, 等. 苹果采摘机器人的机械臂路径规划发展现状[J]. 中国科技信息, 2023, (16): 139-141.
YAN Qinghua, DU Yuzhu, LI Yaqin, et al. Development status of robotic arm path planning of apple picking robot[J]. China Science and Technology Information, 2023, (16): 139-141.
- [3] 颜清华, 杜玉珠, 李亚芹, 等. 苹果采摘机器人的机械臂路径规划发展现状[J]. 中国科技信息, 2023(16): 139-141.
YAN Qinghua, DU Yuzhu, LI Yaqin, et al. Development status of robot arm path planning for apple picking robot[J]. China Science and Technology Information, 2023(16): 139-141.
- [4] 丁士宁. 一种基于 YOLOX-s 模型的苹果检测系统[J]. 中国科技信息, 2024(4): 85-88.
DING Shining. An apple detection system based on YOLOX-s model[J]. China Science and Technology Information, 2024(4): 85-88.
- [5] 聂壮壮, 汪军, 黄翔翔. 基于改进 YOLOv5 的护目镜佩戴检测算法[J/OL]. 重庆工商大学学报(自然科学版). <https://link.cnki.net/urlid/50.1155.N.20240112.1019.00>.
NIE Zhuangzhuang, WANG Jun, HUANG Xiangxiang. Based on the improved YOLOv5 goggle wearing detection algorithm[J/OL]. Journal of Chongqing Technology and Business University(Natural Science Edition). <https://link.cnki.net/urlid/50.1155.N.20240112.1019.00>.
- [6] 宋秋霜, 杨金鹏. 基于图像处理的苹果外观缺陷检测[J]. 南方农机, 2024, 55(4): 147-150.
SONG Qiushuang, YANG Jinpeng. Apple appearance defect detection based on image processing[J]. China Southern Agricultural Machinery, 2024, 55(4): 147-150.
- [7] 黄明辉, 程忠. 苹果采摘机器人目标识别与定位方法研究[J]. 南方农机, 2023, 54(16): 135-138.
HUANG Minghui, CHENG Zhong. Research on target recognition and positioning method of apple picking robot[J]. China Southern Agricultural Machinery, 2023, 54(16): 135-138.
- [8] 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. Piscataway: IEEE Press, 2016: 779-788.
- [9] REDMON J, FARHADI A. YOLO9000: better, faster, stronger[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2017: 7263-7271.
- [10] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2023: 7464-7475.
- [11] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition. New York: ACM, 2014: 580-587.
- [12] GIRSHICK R. Fast R-CNN[C]//Proceedings of the IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2015: 1440-1448.
- [13] 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.
- [14] CAI Z, VASCONCELOS N. Cascade R-CNN: delving into high quality object detection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2018: 6154-6162.
- [15] 郝鹏飞, 刘立群, 顾任远. YOLO-RD-Apple 果园异源图像遮挡果实检测模型[J]. 图学学报, 2023, 44(3): 456-464.
HAO Pengfei, LIU Liqun, GU Renyuan. YOLO-RD-Apple orchard heterogenous image obscured fruit detection model[J]. Journal of Graphics, 2023, 44(3): 456-464.
- [16] 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 Chinese Society of Agricultural Engineering, 2018, 34(7): 173-179.
- [17] 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.
- [18] 周桂红, 马帅, 梁芳芳. 基于改进 YOLOv4 模型的全景图像苹果识别[J]. 农业工程学报, 2022, 38(21): 159-168.

ZHOU Guihong, MA Shuai, LIANG Fangfang. Recognition of the apple in panoramic images based on improved YOLOv4 model[J]. Transactions of the Chinese Society of Agricultural Engineering, 2022, 38(21): 159-168.

[19] 张境锋, 陈伟, 魏庆宇, 等. 基于Des-YOLO v4的复杂环境下苹果检测方法[J]. 农机化研究, 2023, 45(5): 20-25.

ZHANG Jingfeng, CHEN Wei, WEI Qingyu, et al. Apple detection method in complex environments using des-YOLO

v4[J]. Journal of Agricultural Mechanization Research, 2023, 45(5): 20-25.

[20] 高宽祥, 段金英, 汪婕. 基于改进 YOLOv5 算法的 果园水果检测[J]. 工业控制计算机, 2023, 36(11): 107-109.

GAO Kuanxiang, DUAN Jinying, WANG Jie. Orchard fruit detection based on improved YOLOv5 algorithm[J]. Industrial Control Computer, 2023, 36(11): 107-109.

