



DOI:10.12404/j.issn.1671-1815.2403425

引用格式:周晖,杨洁,赵祥飞. YOLOv8 轻量化的果园复杂环境下苹果检测算法[J]. 科学技术与工程, 2025, 25(6): 2274-2283.

Zhou Hui, Yang Jie, Zhao Xiangfei. Lightweight improvement of YOLOv8 for Apple Detection in complex orchard environments[J]. Science Technology and Engineering, 2025, 25(6): 2274-2283.

农业科学

YOLOv8 轻量化的果园复杂环境下苹果检测算法

周晖, 杨洁*, 赵祥飞

(西南林业大学机械与交通学院, 昆明 650224)

摘 要 针对目前复杂果园环境下苹果目标检测算法存在模型参数量大、计算复杂度高,难以在计算资源匮乏的设备上应用的问题,提出了一种改进 YOLOv8 的果园复杂环境下轻量化苹果目标检测算法 YOLOv8n-Apple。引入骨干网络 VanillaNet,减少模型参数量,降低模型复杂度;将原始模型 C2f 模块替换为 C2fGhost 模块,通过较少的卷积运算来获得相似特征图进一步减少模型参数;使用轻量级上采样算子 CARAFE,避免传统上采样算子语义缺失和感受野过小的问题;由于传统损失函数不能完全捕捉到目标之间的相对位置和大小差异,采用 WIoU 边界框作为回归损失函数。收集包含远景顺光、远景背光、近景顺光、近景背光等成熟苹果照片共计 3 120 张,从不同角度和背景进行采集,并改进数据增强,避免数据集单个不确定性;本文提出果园环境下改进后的苹果检测模型平均检测精度分别比 SSD、Faster R-CNN、YOLOv5、YOLOv7、YOLOv8 高 7.5 个百分点、4.8 个百分点、2.2 个百分点、3.8 个百分点和 3.4 个百分点,达到 90%,检测速度达到 286 帧,模型大小 1.8 MB,比原始模型提高了 41 帧,模型大小仅有其 60.0%。

关键词 神经网络; 苹果检测; 轻量化; YOLOv8; VanillaNet

中图法分类号 S225.93;

文献标志码 A

Lightweight Improvement of YOLOv8 for Apple Detection in Complex Orchard Environments

ZHOU Hui, YANG Jie*, ZHAO Xiang-fei

(College of Mechanical and Transportation, Southwest Forestry University, Kunming 650224, China)

[Abstract] Addressing the issues of large model parameters and high computational complexity in apple target detection algorithms for complex orchard environments, which hinder application on devices with limited computational resources, an improved and lightweight apple target detection algorithm named YOLOv8n-Apple based on YOLOv8 was proposed. The backbone network, yaniaNet, was introduced to reduce model parameters and complexity. The original C2f module in the model was replaced with the C2fGhost module, which further decreased model parameters by obtaining similar feature maps through fewer convolutional operations. The lightweight upsampling operator CARAFE was utilized to address the issues of semantic loss and excessively small receptive fields associated with traditional upsampling operators. Given that traditional loss functions cannot fully capture the relative position and size differences between targets, the WIoU bounding box was adopted as the regression loss function. A dataset comprising 3 120 images of mature apples in various scenarios, including distant and close views under front-light and backlight conditions, was collected from diverse angles and backgrounds, to mitigate potential dataset uncertainties. The improved apple detection model for orchard environments demonstrated an average detection accuracy of 90%, which was 7.5, 4.8, 2.2, 3.8, and 3.4 percentage points higher than SSD, Faster R-CNN, YOLOv5, YOLOv7, and YOLOv8, respectively. The detection speed reached 286 frames per second, and the model size was reduced to 1.8 MB, representing an improvement of 41 frames per second compared to the original model, while occupying only 60.0% of size.

[Keywords] neural network; apple detection; lightweight; YOLOv8; VanillaNet

收稿日期: 2024-05-09; 修订日期: 2024-12-17

基金项目: 云南省教育厅科学研究基金(2023J0711/0111723084); 云南省农业联合专项(202301BD070001-001); 云南省专业学位研究生教学案例库建设项目(503210305); 中国学位与研究生教育学会农林学科工作委员会项目(2021-NLZX-YB14/503210401)

第一作者: 周晖(1996—),男,汉族,湖北黄梅人,硕士研究生。研究方向:图像处理、计算机视觉。E-mail:718196955@qq.com。

*通信作者: 杨洁(1973—),女,汉族,河南光山人,博士,副教授。研究方向:机电一体化、机器视觉。E-mail:351725623@qq.com。

投稿网址:www.stae.com.cn

据国家统计局数据显示,2022 年中国苹果总产量 $4\,757.18 \times 10^4$ t,同比增长 3.26%^[1-2]。苹果产量的提升对采摘效率提出了挑战,但目前中国苹果采摘仍以人工为主,效率低、成本高。基于此,采用深度学习目标检测算法的农业机器人得到应用。然而,深度学习算法模型计算量大、模型复杂度高,农业机器人在实际工作环境中部署困难,因此在保证模型的精确度和实时性在满足要求的情况下,如何减少硬件部署成本是目前面临的问题,基于此,开展果园复杂环境下机器视觉的轻量化苹果目标检测对农业机器人或者无人农机开发具有重要意义^[3]。

随着机器视觉快速发展^[4],图像处理技术得到应用,常用的图像预处理方法有图像滤波、二值化和边缘提取^[5],但传统方法对数据量庞大的图像流处理速度慢、特征提取时间长,因此近几年越来越多的专家学者开始进行深度学习算法研究,为了对网络模型进行轻量化处理,武星等^[6]对传统 YOLOv3 深度卷积神经网络架构进行改进,设计一种同构残差块串联的特征提取网络结构,简化目标检测的特征图尺度。王卓等^[7]以 YOLOv4 为基础模型,使用轻量化特征提取网络 MobileNet v3,并在特征融合阶段引用深度可分离卷积(DWConv),降低模型计算量;同时为了保证模型精度,引入坐标注意力机制,强化网络特征提取能力。胡广锐等^[8]引入高效通道注意力(efficient channel attention, ECA)和混洗注意力(shuffle attention, SA)两种轻量化视觉注意力模块,构建了混洗注意力与双卷积层(shuffle attention and double convolution layer, SD-CLayer)模块,提高了检测模型对背景与果实特征的提取能力。公徐路等^[9]将 YOLOv5s 的骨干网络更改为 ShuffleNet v2 轻量化网络,引入 CBAM(convolutional block attention module)注意力模块使模型关注苹果叶片小目标病害。Gu 等^[10]在 YOLOv7-tiny 模型的基础上,使用深度可分离卷积来代替 ELAN 中的普通卷积,减少模型的参数量。Chen 等^[11]提出了一种改进的 YOLOv7 网络模型应用在柑橘园,引入小物体检测层、轻量级卷积和 CBAM(convolutional block attention module)注意力机制,实现多尺度特征提取和融合,减少模型的参数数量。Chen 等^[12]在 YOLOv8n-Pose 的基础上提出了一种改进的 YOLOv8-GP(YOLOv8-grape and picking point)模型,解决葡萄串和采摘点同时检测的问题, C2f 中的瓶颈被替换为包含 EMA(高效多尺度注意力)的 FasterNet Block,从而产生了 C2f-Faster-EMA。减少了冗余计算和内存访问,从而更有效地提取空间特征。

Han 等^[13]提出了改进的 VEW-YOLOv8n 方法,主干网络采用了轻量级的重参数化 VanillaC2f 模块,降低了复杂度和参数数量,并采用扩展的激活函数来增强模型的非线性表达能力。

针对农业机器人在自主工作中检测苹果目标,由于深度学习模型的复杂化,其模型参数十分庞大,导致农业机器人在进行实时检测时会消耗大量的运算成本,在计算能力不足的硬件平台上难以进行实时检测,本文研究以 YOLOv8n 算法为基础,提出一种改进后的轻量化苹果检测算法 YOLOv8n-Apple。针对 YOLOv8n 原始模型骨干网络计算量大、复杂度高的问题,将其替换为 VanillaNet 轻量化骨干网络,降低模型复杂度;将原始模型的 C2f 模块替换为 C2fGhost 模块,使用较少的卷积运算来获得相似特征图,进一步减少模型参数量;采用轻量级上采样算子 CARAFE,增强特征融合效果;替换 YOLOv8n 原始模型边界框损失函数,使用更合理的梯度增益分配策略。

1 YOLOv8n 目标检测算法原理

YOLO 算法是 Redmon 等^[14]提出的一种新的一阶段目标检测算法,与其他算法对比,YOLO 算法在检测速度和检测准确度综合方面性能更强。YOLOv8n(图 1)由骨干网络(Backbone)、特征融合网络(Neck)以及检测头(Head)三部分组成。在 YOLOv8n 骨干网络中,采用轻量化的 C2f 模块替换 YOLOv5 中的 C3 模块,改善特征提取效果,提升目标检测性能。并且模型的第一层的卷积由 YOLOv5 中的 6×6 卷积改为 3×3 卷积,有效减轻模型参数量和计算复杂度;在特征融合网络中,YOLOv8n 摒弃了 YOLOv5 中 1×1 卷积的降采样层,并将 C3 模块替换为 C2f 模块,在减少模型参数的同时进一步增强特征融合的效果;在检测头设计中,YOLOv8n 为了提升模型的计算效率,在检测头设计上采用解耦头的设计,将分类任务和回归任务进行分开处理。

2 YOLOv8n-Apple 检测算法

本文研究提出一种基于改进 YOLOv8n 的轻量化苹果检测算法,该模型结构如图 2 所示,其中轻量化骨干网络 VanillaNet^[15]结构如图 3 所示,该结构相较于 YOLOv8n 原始模型骨干网络,层数少、计算简单、模型参数小。将收集到的原始数据集进行数据划分,然后把训练集用于模型训练,得到模型训练权重,最后将模型训练权重用于测试集测试,将测试结果进行验证分析。

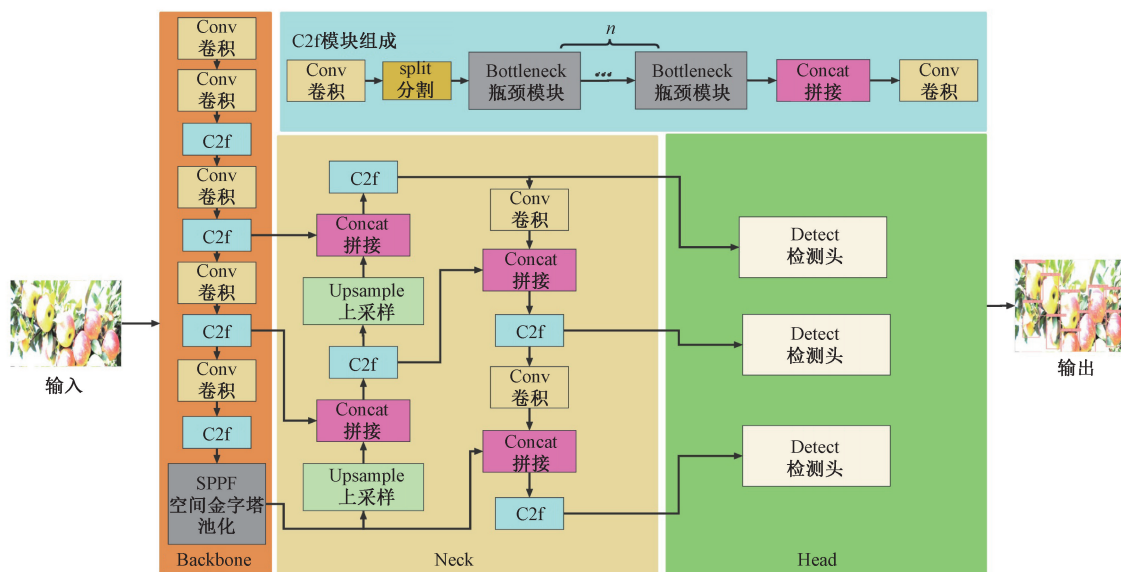


图1 YOLOv8n 的网络结构

Fig. 1 Network architecture of YOLOv8n

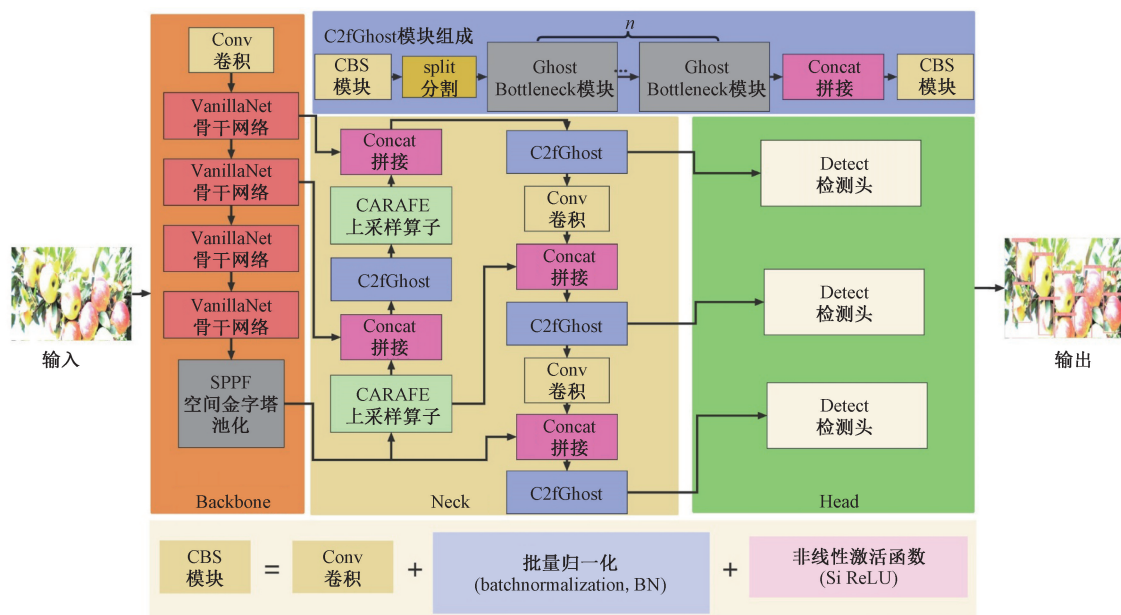


图2 YOLOv8n-Apple 的网络结构

Fig. 2 Network architecture of YOLOv8n-Apple

2.1 轻量化骨干网络 VanillaNet

在视觉任务处理中,骨干网络主要是对图像进行特征提取。YOLOv8n 的骨干网络采用的是 DarkNet-53,该结构设计复杂、模型参数大,为此本次实验采用轻量化 VanillaNet 来替换 DarkNet-53, VanillaNet 包含简单的卷积层和池化层,没有复杂的分支结构^[16],因此有效降低了模型大小,并提升模型计算率。它的结构图如图3所示(以6层的网络结构为例),该网络结构包含3个部分。在第Ⅰ部分 Stem 结构中,将原始3通道图像经过下采样处理转换成C通道的特征图。在第Ⅱ部分中,不断使用步

长为2的池化层来对特征图的尺寸进行调整。在第Ⅲ部分中,使用全连接层对分类结果进行输出。在网络结构上 VanillaNet 设计简单,同时在处理计算机视觉任务时性能卓越,有效避免网络结构的深度、额外的分支以及复杂的注意力机制,产生结构简化的网络,解决了目前深度学习框架过于复杂的问题,适应于资源非常有限的硬件部署环境。

2.2 C2fGhost 模块

在普通卷积中,由于提取特征时很多特征通道上的特征映射非常相似,增加网络模型参数数量及浮点数运算次数(FLOPs)。为了使用较少的卷积来

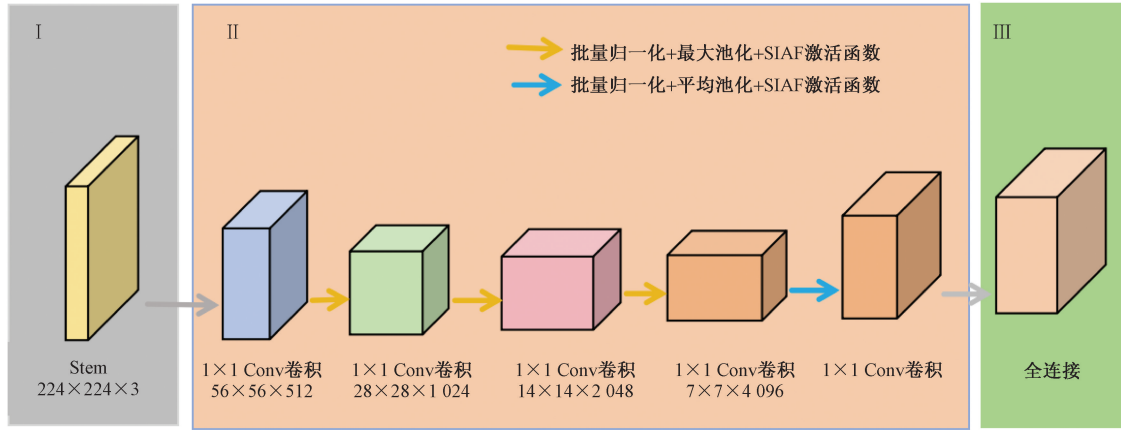


图3 VanillaNet 网络结构

Fig. 3 VanillaNet network architecture

获得相似特征图,Ghost 模块(图4)先通过卷积来提取物体部分特征,再运用线性变换来获得物体更多的表面特征,最终将特征映射相似的两组特征信息拼接起来投射到输出通道的维度上,实现使用较少的参数和运算量获得相同特征映射的效果。

用 Ghost Bottleneck 模块替换基础网络 YOLOv8n 中的 C2f 模块,形成 C2fGhost 模块,不仅能减少模型大小,还能提升对冗余信息的提取速度,实现模型轻量化的同时保证模型检测的准确性。

2.3 轻量级上采样算子 CARAFE

在自然环境下收集果实数据集,由于背景环境复杂多变,导致其基础模型在上采样过程中语义信息缺失并且感受野过小,不能实现很好的特征融合效果。为了模型能输出较强的特征信息,在 YOLOv8n 的基础模型上用轻量级上采样算子 CARAFE^[17]来替换掉传统上采样算子,避免语义信息缺失和感受野过小的问题。

轻量级上采样算子 CARAFE 由上采样核预测模块和内容感知重组模块组合而来,其中上采样核预测模块通过产生上采样核,根据下采样特征图上的信息映射对上采样核重新进行权重分配,实现更好的特征融合。内容感知模块在上采样过程中,能保留更多的物体边界信息。其中采样核预测模块计算方法,公式为

$$W_{l'} = \Psi[N(X_l, X_{\text{encoder}})] \quad (1)$$

式(1)中: $W_{l'}$ 为预测位置核的坐标; Ψ 为采样核预测模块; $N(X_l, X_{\text{encoder}})$ 为原始特征图 $X_{\text{encoder}} \times X_{\text{encoder}}$ 子区域; X_l 为输入特征图位置坐标; X_{encoder} 为原始特征图子区域大小参数。

经过内容感知重组模块处理,将原始特征图的领域和预测位置核的坐标进行特征融合,公式为

$$X_{l'} = \Phi[N(X_l, K_{\text{up}}), W_{l'}] \quad (2)$$

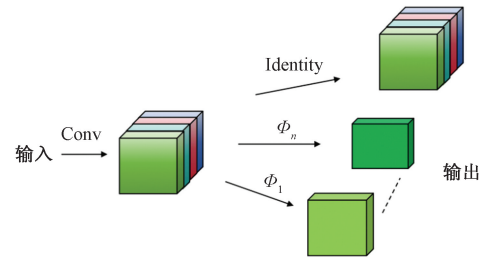
Conv 为卷积;Identity 为恒等映射; Φ_1 为线性变换 1; Φ_n 为线性变换 n

图4 Ghost 模块结构

Fig. 4 Ghost module structure

式(2)中: $X_{l'}$ 为融合后的新特征图; Φ 为内容感知重组模块; $N(X_l, K_{\text{up}})$ 为输入特征图 $K_{\text{up}} \times K_{\text{up}}$ 子区域; $W_{l'}$ 为预测位置核的坐标。

2.4 边界框回归损失函数优化

交集-并集比例(intersection over union, IOU)是一种描述框之间的重合度的方式。在回归任务中,可通过“目标框”与“预测框”的比值来衡量框的回归程度,但传统的目标检测损失函数主要关注预测框和真实框之间的 IOU,并不能完全捕捉到目标之间的相对位置和大小差异,无法完全展现出目标检测模型的性能。Tong 等^[18]提出了动态非单调的聚焦机制,设计了 WIoU(Wise-IoU)。采用“离群度”概念来替代 IoU 对锚框进行预测,并使用更合理的梯度增益分配策略,使得损失函数聚焦于普通质量的锚框,提升检测器检测性能。实验采用 WIoU V3,计算公式为

$$\beta = \frac{L_{\text{IoU}}^*}{\bar{L}_{\text{IoU}}} \in [0, +\infty) \quad (3)$$

$$r = \frac{\beta}{\delta \alpha^{\beta-\delta}} \quad (4)$$

$$L_{\text{WIoUv3}} = r L_{\text{WIoUv1}} \quad (5)$$

式中: β 为离群度; L_{IoU}^* 为具有动量 m 的指数移动

值; $\bar{L}_{\text{IoU}}$ 为具有动量 m 的指数移动平均值; α, δ 分别为超参数, 不同的超参数适用于不同的模型和数据集; r 为梯度增益; L_{WIoUv1} 为基于注意力的边界框损失; L_{WIoUv3} 为处于动态变化中在最高梯度增益的边界框的条件下产生的损失函数。

2.5 数据集介绍

实验使用的数据集来源于昆明西山区团结印象苹果园和网络图片 (<https://universe.roboflow.com/peterjbloch-gmail-com/apple-vision>), 共计 3 120 张。数据集详细信息如图 5 所示, 部分数据集图片如图 6 所示。

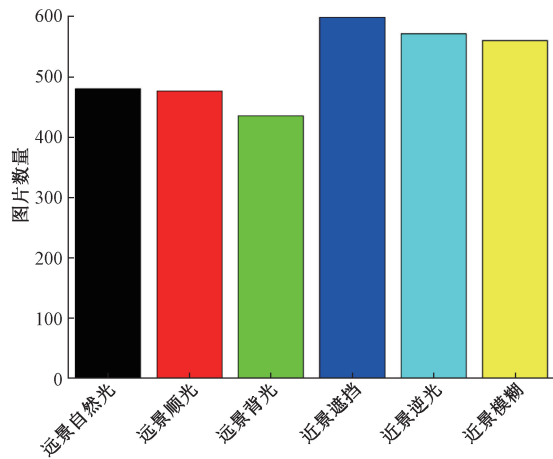


图 5 数据集详细信息

Fig. 5 Dataset information

2.6 改进数据增强

为了模拟现实场景中复杂的环境背景, 减少样

本信息单调对算法鲁棒性的影响, 对原始数据集进行前期处理。本次实验采用数据增强的方式对原始图片进行一系列的变换和操作, 产生更多的训练样本, 增强模型对目标特征的学习能力, 提高模型的泛化能力和复杂自然环境下的鲁棒性, 降低模型过拟合性能。在对原始的 3 120 张照片进行数据增强后, 获得了 7 292 张图像。通过 python 脚本文件随机将数据集划分成训练集、验证集、测试集, 其比例为 8:1:1, 可以得到训练集图片 5 833 张, 验证集和测试集图片分别为 729 张。实验通过使用 opencv-python 编写脚本来对数据随机进行加噪声、改变亮度、平移、旋转、镜像、cutout 处理, 对单张图片采用一种或多种处理手段来模拟苹果采摘机器人在工作时面临的强光、暗光、画面抖动、画面不全、遮挡等情况, 继续增强算法在果园复杂环境下的鲁棒性, 数据增强效果如图 7 所示。

3 结果与分析

3.1 实验环境

实验在 AutoDL 平台上实现, 实验运行环境如下: 采用 Ubuntu18.04 操作系统, CPU: Intel(R) Xeon(R) Platinum 8255C, GPU: NVIDIA GeForce RTX 2080ti, Python 版本: Python3.8。

3.2 性能评价指标

目标检测模型的性能需要进行定量分析, 本文研究从精确率 (precision)、平均准确率均值 (mean average precision, mAP)、每秒帧率 (frames per second, FPS)、模型的参数量 (params) 和计算量



图 6 部分数据集图片

Fig. 6 Some images from the dataset



图7 数据增强效果

Fig. 7 Data augmentation effects

(GFLOPs)5个指标来评估模型。精确率表示指模型预测为正的样本中实际也为正的样本占被预测为正的样本的比例。平均准确率均值在检测任务中表示检测到所有目标的平均精度值,值越高表示预测框与真实框之间拟合的更准确。FPS表示单位时间内处理图像的数量,值越高说明检测任务的实时性越好。模型的参数量(params)用来表示模型大小,值越大表示模型部署要求的硬件资源越多,部署越困难。计算量(GFLOPs)用来评价模型的计算量大小,GFLOPs值越低表示模型计算量小、计算效率高。

3.3 模型的训练

实验将数据集划分成训练集、验证集、测试集3个部分,其比例8:1:1。模型训练过程中部分参数的设置如下:轮次(Epoch)为150,批量大小(batch-size)为16,学习率为0.01,动量(momentum)为0.937。在模型训练过程中,网络模型的损失函数越小说明模型训练后的权重和真实值拟合的越好。原始YOLOv8n算法和本文算法损失函数变化情况如图8所示。从图8中可知,训练前期损失值下降较快,随着轮次增加,损失值慢慢趋于稳定,当轮次为100左右时模型基本完成收敛,并且在整个训练过程中模型没有出现拟合现象。相较于原始YOLOv8n算法,本文所提出来的YOLOv8n-Apple算法收敛更快,过程更平滑,模型训练权重更好。

3.4 消融实验

将训练好的权重文件导入相应模型,用测试集来进行实验,为了能直观地观察到各个改动部分对实验结果的影响,对整个测试过程采取消融实验的

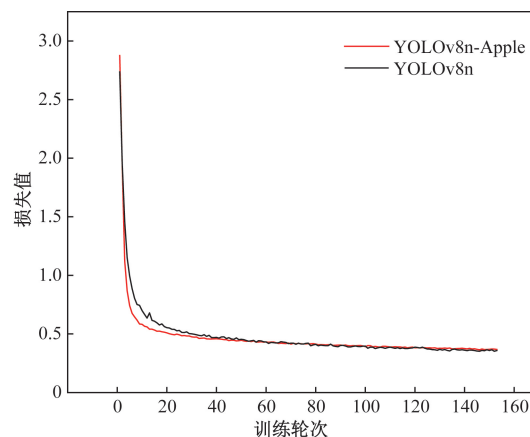


图8 训练损失函数曲线

Fig. 8 Training loss function curve

方式进行,消融实验结果如表1所示,不同改进模型对应的精确率(precision)曲线如图9所示。

从图9来看,随着模型结构的改进,模型的精度值稳步提升,并最终趋于稳定。其中,本文提出的YOLOv8n-Apple算法精确度曲线稳定在93.9%,比消融实验其他改进阶段的模型都高,比原始模型高1.8%,实验表明YOLOv8n-Apple算法检测苹果目标有更好的精确性。

从表1中可见,将骨干网络替换为VanillaNet模块后,模型参数量由3.0M降为1.9M,计算量由原始模型的8.1GFLOPs降为5.5GFLOPs, FPS由245帧提升为301帧,同时提升了mAP@0.5:0.95值,提升了大约0.7个百分点。因为采用VanillaNet作为骨干网络,没有多余的分支连接,减少了模型运算量并且还能保持良好的精度和实时性。引入C2fGhost模块后,能有效降低模型大小,模型大小由

表 1 消融实验结果
Table 1 Ablation experiment results

VanillaNet	C2fGhost	CARAFE	wise-IOU	Precision/%	mAP@0.5;0.95/%	Params/M	GFLOPs	FPS
×	×	×	×	92.1	86.6	3.0	8.1	245
✓	×	×	×	92.4	87.3	1.9	5.5	301
✓	✓	×	×	92.5	88.2	1.5	4.9	312
✓	✓	✓	×	93.2	89.1	1.8	5.1	280
✓	✓	✓	✓	93.9	90.0	1.8	5.1	286

注:VanillaNet 为轻量化骨干网络;C2fGhost 为改进后的 C2f 模块;CARAFE 为轻量级上采样算子;wise-IOU 为损失函数;Precision 为精确度;mAP@0.5;0.95 为 mAP@0.5;0.95 表示 IoU 阈值从 0.5 开始,以 0.05 为步长,直到 0.95,计算得到的平均准确率均值;Params 为模型的参数量;GFLOPs 为计算量;FPS 为每秒帧率。

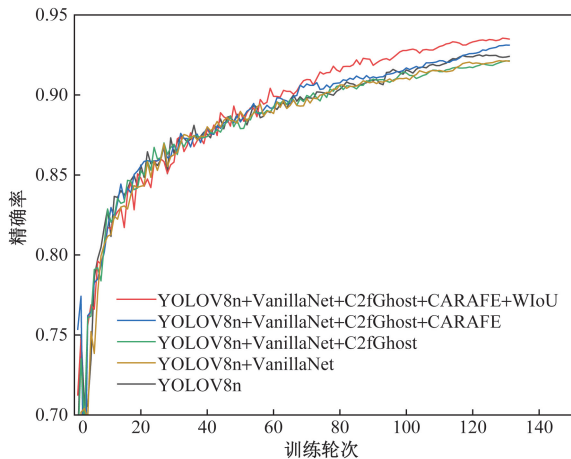


图 9 不同模型精确率曲线
Fig. 9 Precision curves of different models

1.9 M 下降为 1.5 M,计算量由 5.5GFLOPs 下降为 4.9GFLOPs, FPS 由于计算量的减少,从 301 帧提升到 312 帧,提升了 11 帧。在特征融合网络中引入 CARAFE 模块后,通过联系上下文信息进行特征融合,虽然增加模型部分运算量,却显著提升模型的检测精度,mAP@0.5;0.95 值由 88.2 到 89.1,涨了 0.9 个百分点,参数量和计算量分别增加了 0.3 M 和 0.2GFLOPs,导致 FPS 值下降到 280 帧,但仍能满足日常工作中的实时检测任务。将损失函数替换为 Wise-IOU 后,模型的检测精度提升了 0.9 个百分点,FPS 从 280 帧上升为 286 帧。最终模型保持了较低的参数量和计算量,仅为 YOLOv8n 的 60% 和 63%,因此本文提出的改进后的网络模型较原始模型相比,能够降低模型大小,同时保持良好的检测精度和检测实时性。

3.5 与当前先进算法对比

与当前主流目标检测模型性能对比,实验结果如表 2 所示,可知改进后的 YOLOv8n 苹果目标检测算法在 mAP@0.5;0.95、模型参数量、模型计算量等方面表现最好。与原始模型相比,mAP@0.5;0.95 值提升了 3.4 个百分点,模型的参数量减少 1.2M,模型的计算量减少 3GFLOPs,FPS 提升 41 帧。

表 2 与主流目标检测模型性能对比
Table 2 Performance comparison with mainstream object detection models

模型	mAP@0.5;0.95	Params/M	GFLOPs	FPS
SSD	82.5	26.3	62.7	78
Faster R-CNN	85.2	137.1	370.2	14
YOLOV5	87.8	7.03	15.8	170
YOLOV7	86.2	6.02	13.1	143
YOLOV8	86.6	3.0	8.1	245
本文模型	90.0	1.8	5.1	286

注:mAP@0.5;0.95 为 mAP@0.5;0.95 表示 IoU 阈值从 0.5 开始,以 0.05 为步长,直到 0.95,计算得到的平均准确率均值;Params 为模型的参数量;GFLOPs 为计算量;FPS 为每秒帧率。

综上所述,本文提出的改进 YOLOv8n 算法在复杂的果园环境下,不仅降低了硬件环境安装的成本,同时保持了良好的检测精度,满足任务检测实时性要求。

为了能直观对比原始 YOLOv8n 算法和改进后算法的区别,设置 2 组照片测试算法分别在远景、近景背景下检测果园环境下苹果目标准确性。对比结果如图 10 所示,从图 10 可以看出,改进前的苹果检测算法表现较差,容易出现漏检现象,并且对小目标检测效果较差,不能很好地适应果园环境下苹果目标检测,而实验算法可以较为准确地检测出在设定条件下的苹果目标。并且本文研究利用 Grad-CAM 图对 YOLOv8n 算法和改进后的算法进行可视化效果比较,如图 11 所示。红色部分表示模型对目标区域的关注程度,通过热力图可以看出,改进后的算法较原始算法相比,对苹果目标给予了更强的关注,有助于网络模型对关键特征的学习,提升目标检测精度。

4 结论

在果园复杂环境下,基于 YOLOv8n 基准模型提出一种轻量化苹果检测算法,采用轻量化的骨干网络 VanillaNet,在有效降低模型参数同时保持良好的检测精度;引入 C2fGhost 模块,使用较少的卷积来获得相似特征图,进一步减少模型参数,提高嵌入

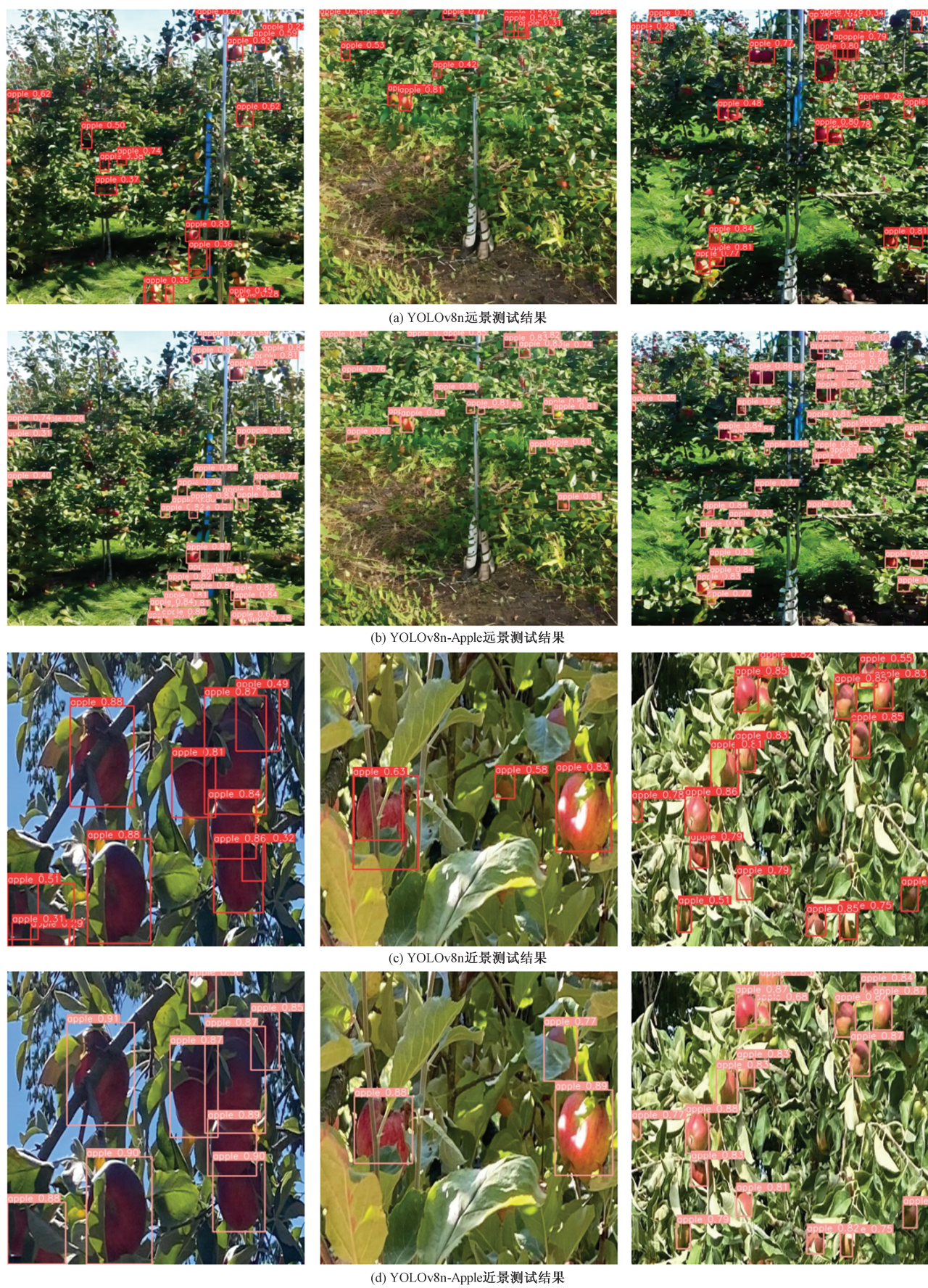


图 10 改进前后效果对比

Fig. 10 Comparison of effects before and after improvement

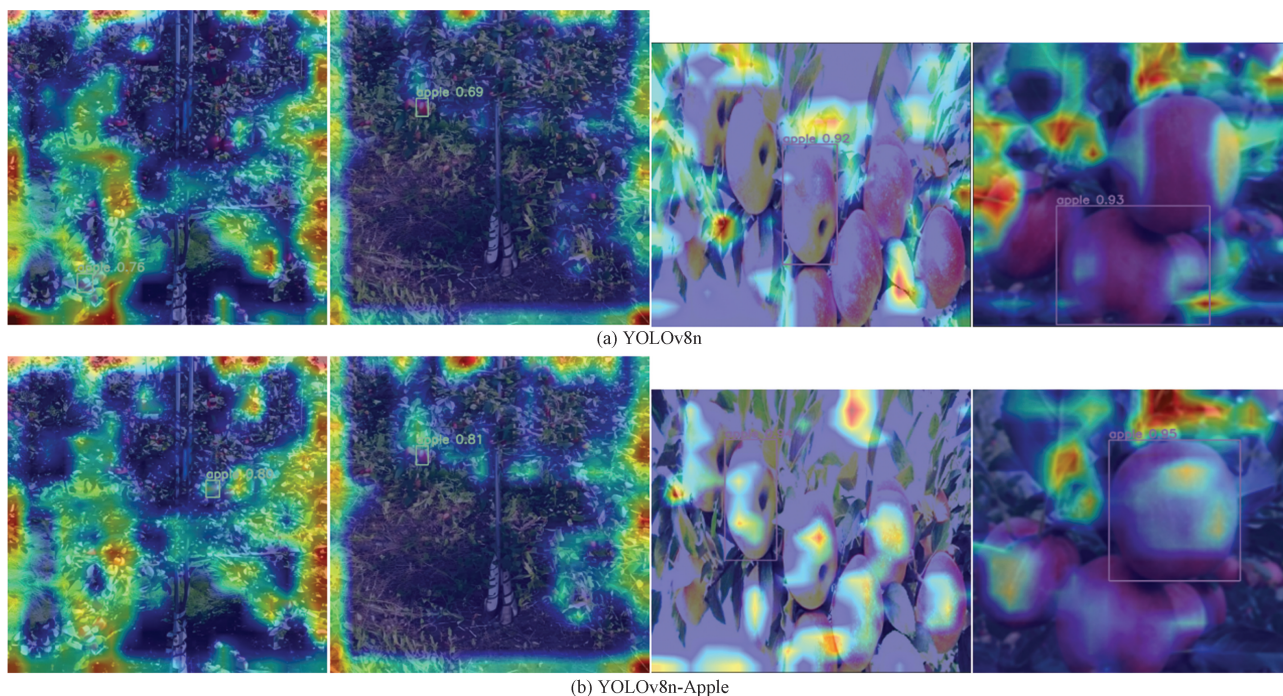


图 11 不同模型的热力图效应

Fig. 11 Heatmap effects of different models

式设备部署适应性;使用上采样算子 CARAFE 后,实现更好的特征融合,提升目标检测精度;将损失函数替换为 WIoU,采用的梯度分配增益比基础模型的损失函数更合理,进一步提高目标检测效果。

实地采集和网上收集成熟季节树上苹果照片建立数据集,验证改进后算法优越性。结果表明:改进后的算法在复杂自然环境下平均检测精度达到 90.0%,较 YOLOv8n 原始模型提高 3.4 个百分点,并且有效降低模型大小,检测速度提升 41 帧,为农业机器人选择果园复杂环境下苹果目标检测算法提供了参考。

参 考 文 献

- [1] Chen M Y, Shao M A, Wei X R, et al. Response of the vertical distribution of soil water and nitrogen in the 5 m soil layer to the conversion of cropland to apple orchards in the Loess Plateau, China [J]. Agriculture, Ecosystems & Environment, 2022, 333. DOI: 10.1016/j.agee.2022.107960.
- [2] Zhao L H, Gao X D, He N N, et al. Ecohydrological advantage of young apple tree-based agroforestry and its response to extreme droughts on the semiarid Loess Plateau [J]. Agricultural and Forest Meteorology, 2022, 321. DOI: 10.1016/j.agrformet.2022.108969.
- [3] 王成军, 韦志文, 严晨. 基于机器视觉技术的分拣机器人研究综述 [J]. 科学技术与工程, 2022, 22(3): 893-902.
Wang Chengjun, Wei Zhiwen, Yan Chen, et al. Review on sorting robot based on machine vision technology [J]. Science Technology and Engineering, 2022, 22(3): 893-902.
- [4] 孔令云, 吴楚枫, 尹果果. 基于图像处理技术的混凝土表面构造仿真 [J]. 科学技术与工程, 2020, 20(6): 2361-2368.
Kong Lingyun, Wu Chufeng, Yin Guoguo. Concrete surface structure simulation based on image processing technology [J]. Science Technology and Engineering, 2020, 20(6): 2361-2368.
- [5] 马文礼, 李治平, 卢婷, 等. 机器视觉在油气勘探开发中的应用现状 [J]. 科学技术与工程, 2018, 18(17): 112-119.
Ma Wenli, Li Zhiping, Lu Ting, et al. Application status of machine vision in oil and gas exploration and development [J]. Science Technology and Engineering, 2018, 18(17): 112-119.
- [6] 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.
- [7] 王卓, 王健, 王泉雄, 等. 基于改进 YOLO v4 的自然环境苹果轻量化检测方法 [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 YOLO v4 [J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(8): 294-302.
- [8] 胡广锐, 周建国, 陈超, 等. 融合轻量化网络与注意力机制的果园环境下苹果检测方法 [J]. 农业工程学报, 2022, 38(19): 131-142.
Hu Guangrui, Zhou Jianguo, Chen Chao, et al. Fusion of the lightweight network and visual attention mechanism to detect apples in orchard environment [J]. Transactions of the Chinese Society of Agricultural Engineering, 2022, 38(19): 131-142.
- [9] 公徐路, 张淑娟. 基于改进 YOLOv5s 的苹果叶片小目标病害轻量化检测方法 [J]. 农业工程学报, 2023, 39(19): 175-184.
Gong Xulu, Zhang Shujuan. Lightweight detection of small target

- diseases in apple leaf using improved YOLOv5s[J]. Transactions of the Chinese Society of Agricultural Engineering, 2023, 39(19): 175-184.
- [10] Gu B, Wen C, Liu X, et al. Improved YOLOv7-Tiny complex environment citrus detection based on lightweighting[J]. Agronomy, 2023, 13(11). DOI: 10.3390/agronomy13112667.
- [11] Chen J, Liu H, Zhang Y, et al. A multiscale lightweight and efficient model based on YOLOv7: applied to citrus orchard[J]. Plants, 2022, 11(23). DOI: 10.3390/plants11233260.
- [12] Chen J, Ma A, Huang L, et al. Efficient and lightweight grape and picking point synchronous detection model based on key point detection[J]. Computers and Electronics in Agriculture, 2024, 217. DOI: 10.1016/j.compag.2024.108612.
- [13] Han B, Lu Z, Dong L, et al. Lightweight non-destructive detection of diseased apples based on structural reparameterization technique [J]. Applied Sciences, 2024, 14(5). DOI: 10.3390/app14051907.
- [14] Redmon J, Divvala S, Girshick R, et al. Youonly look once: unified, real-time object detection [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 779-788.
- [15] Chen H, Wang Y, Guo J, et al. Vanillanet: the power of minimalism in deep learning[J]. Advances in Neural Information Processing Systems, 2023: arxiv-2305.12972v2.
- [16] 王春梅, 刘欢. YOLOv8-VSC: 一种轻量级的带钢表面缺陷检测算法[J]. 计算机科学与探索, 2024, 18(1): 151-160. Wang Chunmei, Liu Huan. YOLOv8-VSC: lightweight algorithm for strip surface defect detection[J]. Journal of Frontiers of Computer Science and Technology, 2024, 18(1): 151-160.
- [17] Wang J, Chen K, Xu R, et al. Carafe: content-aware reassembly of features[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Long Beach: IEEE, 2019: 3007-3016.
- [18] Tong Z, Chen Y, Xu Z, et al. Wise-IoU: bounding box regression loss with dynamic focusing mechanism[J]. arXiv preprint arXiv: 2301.10051, 2023.