

# 基于 YOLO-FFD 的水果品种和新鲜度识别方法

## Fruit variety and freshness recognition method based on YOLO-FFD

鄢 紫 陈良艳 刘卫华 赖华清 叶 胜

YAN Zi CHEN Liangyan LIU Weihua LAI Huaqing YE Sheng

(武汉轻工大学电气与电子工程学院, 湖北 武汉 430023)

(School of Electrical and Electronic Engineering, Wuhan Polytechnic University, Wuhan, Hubei 430023, China)

**摘要:**目的:改善现有水果识别与分级方法依赖于人工操作和复杂设备的情况。方法:提出了一种轻量化模型 YOLO-FFD(YOLO with fruit and freshen detection),该模型以 YOLOv5 框架为基础,基于深度可分离卷积和 GELU 激活函数设计轻量化模块 LightweightC3 作为主干特征提取网络的基本单元,减少模型参数数量和计算量,加快模型的收敛速度;使用大内核深度可分离卷积模块 EnhancedC3 改进原模型的颈部,抑制信息丢失并增强模型的特征融合能力,提高模型的检测精度;采用 GSConv 代替特征融合网络中的普通卷积,使模型进一步轻量化。结果:提出模型的平均精度均值达到了 96.12%,在 RTX 3090 上速度为 172 帧/s,在嵌入式设备 Jetson TX2 上速度为 20 帧/s。相比于原始 YOLOv5 模型,平均精度均值提高了 2.21%,计算量减少了 26%,在 RTX 3090 和 Jetson TX2 上的速度分别提高了 2 倍和 1 倍。结论:YOLO-FFD 模型能够满足识别水果品种和新鲜度的需求,且在复杂场景下错检、漏检情况均有改善。

**关键词:**水果;新鲜度;品种识别;轻量化;深度学习;目标检测

**Abstract: Objective:** In order to improve the situation that existing fruit recognition and classification methods rely on manual operation and complex equipment. **Methods:** A lightweight model YOLO-FFD (YOLO with fruit and freshen detection) was proposed, which based on the YOLOv5 framework. Firstly, LightweightC3 was designed as the basic unit of the backbone feature extraction network based on the depth separable convolution and GELU activation function, which reduced the number of model parameters and computation, and speeds up the convergence of the model. Secondly,

**基金项目:**湖北省高校优秀中青年科技创新团队项目(编号:T2021009)

**作者简介:**鄢紫,女,武汉轻工大学在读硕士研究生。

**通信作者:**陈良艳(1982—),女,武汉轻工大学副教授,博士。

E-mail: chenliangyan@whpu.edu.cn

**收稿日期:**2023-05-15 **改回日期:**2023-09-13

EnhancedC3, a large kernel depth separable convolution module, was used to improve the neck of the original model, suppressed information loss and enhance the feature fusion ability of the model, so as to improve the detection accuracy of the model. Finally, GSConv was used to replace the common convolution in the feature fusion network to further lighten the model. **Results:** The experimental results showed that the average accuracy of the proposed model reached 96.12%, the FPS on RTX 3090 was 172, and the speed on the embedded Jetson TX2 was 20 frames per second. Compared with the original YOLOv5 model, the mAP was improved by 2.21%, the calculation amount was reduced by 26%, and the speed was increased by two times.

**Conclusion:** YOLO-FFD can meet the requirement of identifying fruit varieties and freshness, and improve the falsely detection and missing detection in complex scenes.

**Keywords:** fruit; freshness; variety identification; lightweight; deep learning; object detection

水果品质管理问题涉及到人们身体健康和经济活动,因此对水果的检测和新鲜度分拣变得非常重要。目前,中国水果分拣主要依赖于人工操作,采用计算机算法进行自动化分级是发展趋势。对水果的检测和新鲜度分级主要有传统机器学习方法以及深度学习方法。

目前大多数研究都围绕水果识别或多品种分类展开<sup>[1-3]</sup>,传统的计算机视觉方法解决水果新鲜度分级问题大多通过 K-近邻(KNN)、支持向量机(SVM)、人工神经网络(ANN)等机器学习方法建立模型。刘禾等<sup>[4]</sup>利用人工神经网络来分析检测苹果体形并进行分类,达到了 80% 的识别率。Semary 等<sup>[5]</sup>使用支持向量机和主成分分析(PCA)从 HSV 和 RGB 颜色空间的纹理和颜色特征中检测番茄样本表面缺陷,达到了 92% 的准确率。Moallem 等<sup>[6]</sup>使用支持向量机、人工神经网络和 K-近邻 3 种机器学习技术分别检测苹果的腐烂或感染部分,结果显示使用 SVM 分类器的评分结果最高,为 92.5%。以上传统机器学习方法在检测果蔬等产品的新鲜度方面是有效的,

但是这些模型主要基于手工特征提取方法,耗时而且依赖于所研究的对象类型,具有局限性。

基于深度学习的方法近年来在农业领域也得到了更广泛的应用。基于深度学习的模型可以自动提取相关特征,无需人工干预。Suharjito 等<sup>[7]</sup>在 MobileNetV1 的基础上基于相应的标准将油棕分成 6 个等级,达到了 81.1% 的精度。Gai 等<sup>[8]</sup>基于 YOLOv4 网络引入 DenseNet 结构检测不同成熟度樱桃果实的位置,准确率达到了 94.7%。Chen 等<sup>[9]</sup>利用 YOLOv5s 先识别柑橘再结合该目标的显著性图输送到 ResNet34 网络中确定果实成熟度等级,最终成熟度检测准确率达到了 95.07%。Ananthanarayana 等<sup>[10]</sup>提出了一种利用 SSD 网络和 MobileNetV2 网络进行水果新鲜度分类和检测的方法,所提出的方法在图像分类和目标检测任务中的准确率分别达到 97% 和 62%。尽管上述研究采用深度学习的方法在一定程度上有效解决了水果新鲜度分类的问题,但这些模型大多数都是仅针对单一品种,泛化能力不高,背景单一忽略真实场景下因遮挡重叠因素造成的识别困难,抗噪能力弱。

针对日常复杂场景下,识别多种水果不同新鲜度级

别的需求,克服密集场景、遮挡障碍等困难,同时为了便于模型部署在移动端设备和满足实时检测任务的速度和精度要求,网络结构应轻量化。研究拟提出一种水果新鲜度检测算法 YOLO-FFD,设计一种基于深度可分离卷积的轻量化模块 LightweightC3 作为主干网络的基本单元,并引入 GSConv 和大核卷积块 EnhancedC3 改进特征融合网络,以期为实现日常场景下水果品种及新鲜度识别的精度和速度要求。

## 1 水果品种和新鲜度识别算法

### 1.1 YOLO-FFD 模型结构

YOLO-FFD 模型结构如图 1 所示,其以 YOLOv5n 为基本框架做了三方面的改动。

- (1) 在特征提取网络中提出 LightweightC3 模块作为主干网络的基本构建模块,对模型进行轻量化,减少参数量。
- (2) 设计 EnhancedC3 模块在模型的特征融合部分取代原有的 C3 模块,充分利用特征信息,提高模型的特征表达能力,提高检测精度。
- (3) 替换颈部网络中的普通卷积为 GSConv,进一步压缩模型参数量。

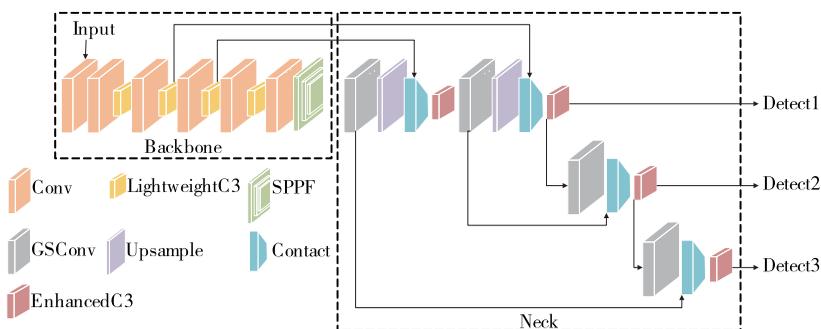


图 1 YOLO-FFD 模型结构图

Figure 1 The structure of YOLO-FFD model

### 1.2 特征提取网络的改进

1.2.1 基本单元 一阶段目标检测算法 YOLOv5 具有运行速度快、检测效果好的优点。该网络由输入端 (Input)、主干网络 (Backbone)、颈部网络 (Neck) 和检测端 (Head) 四部分组成。目前包含有 5 个不同版本,即 YOLOv5n、YOLOv5s、YOLOv5m、YOLOv5l 和 YOLOv5x。其中 YOLOv5n 网络的深度和宽度最小,而其他网络则在其基础上加深和加宽。因此,选用 YOLOv5n 为基础框架。

原始 YOLOv5n 的主干网络采用普通卷积构成的 C3 模块进行特征提取,受 MobileNet 的启发,使用深度可分离卷积取代普通卷积,把对输入进行的滤波组合的连续步骤分解为两步,降低计算量和模型大小,减少图像处理所需的内存,并加快目标的检测速度。因此,采用轻量级

的 LightweightC3(LC3)模块作为原始 YOLOv5n 的主干网络的基本单元。其结构如图 2 所示,将输入的特征图分为两个分支,其中一个分支进行普通卷积后再经过一

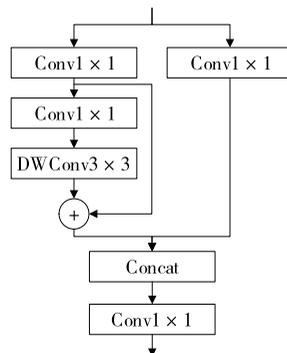


图 2 LightweightC3 模块

Figure 2 LightweightC3 module

个由普通卷积和深度可分离卷积组成的残差结构,与另一个经过普通卷积的分支的结果进行 Concat 操作,最后经过一个  $1 \times 1$  普通卷积后得到输出的特征图。

深度可分离卷积是由分组卷积和点卷积两个卷积组成,如图 3 所示,首先由分组卷积对输入特征图的每个通道分别使用一个卷积核进行卷积,然后将卷积核的输出进行拼接得到输出。接着由逐点卷积对分组卷积输出的

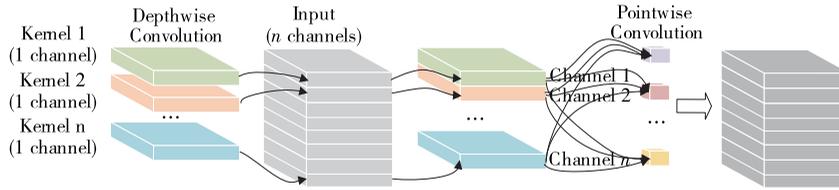


图 3 深度可分离卷积

Figure 3 Depth-separable convolution

$$Q_{SC} = k^2 \times H_{out} \times W_{out} \times M \times N, \quad (1)$$

$$Q_{DWC} = k^2 \times H_{out} \times W_{out} \times M + M \times N \times H_{out} \times W_{out}, \quad (2)$$

$$\frac{Q_{DWC}}{Q_{SC}} = \frac{1}{N} + \frac{1}{K^2}, \quad (3)$$

式中:

$Q_{SC}$ ——普通卷积的计算量;

$Q_{DWC}$ ——深度可分离卷积的计算量;

$H_{in}, W_{in}, M$ ——输入特征图的长、宽以及通道数;

$H_{out}, W_{out}, N$ ——输出特征图的长、宽以及通道数;

$k$ ——卷积核的大小。

深度可分离卷积的参数量是普通卷积的  $\frac{1}{N} + \frac{1}{K^2}$  倍,

因此引入深度可分离卷积可以减少模型的参数量,节约计算成本。而且它基本不影响卷积层的特征提取能力,因此在改进的网络中引入深度可分离卷积是很有必要的。

**1.2.2 激活函数** 由于各品种水果的种内相似度和种间相似性都相对较高,因此在 LightweightC3 模块中采用 GELU 激活函数(gaussian error linear unit,高斯误差线性单元),以提高模型的泛化能力。GELU 激活函数结合了 dropout 策略和 ReLU 函数丢弃不重要的激活信息的共性。dropout 按照指定的概率值对输入神经元随机乘以 0,ReLU 则按照输入值的正负对输入神经元乘以 0 或 1。而 GELU 激活函数就是将神经元输入乘以 0 或 1 的概率  $\Phi(x)$  与神经元输入  $x$  本身结合起来,当  $x$  变大,  $\Phi(x)$  随之变大,  $x$  越容易被保留,当  $x$  变小,  $\Phi(x)$  也随之变小,  $x$  越容易被丢弃。这样操作既保留了概率性也提高了对于神经元输入值本身的依赖性。GELU 激活函数的数学表达式如式(4)所示。

$$GELU(x) = x\Phi(x) \approx 0.5x(1 + \tanh$$

$$\left[ \sqrt{\frac{2}{\pi}}(x + 0.044715x^3) \right]), \quad (4)$$

特征图进行通道融合得到最后的输出特征图。假设分别使用  $k \times k$  大小的普通卷积和  $k \times k$  大小的深度可分离卷积对尺寸大小为  $H_{in} \times W_{in} \times M$  的输入特征图进行运算,最终得到大小为  $H_{out} \times W_{out} \times N$  的输出特征图,那么普通卷积的计算量  $Q_{SC}$  如式(1)所示,深度可分离卷积的计算量  $Q_{DWC}$  如式(2)所示,深度可分离卷积的计算量与普通卷积计算量的比值如式(3)所示。

式中:

$\Phi(x)$ —— $x$  的高斯正态分布的累积分布。

### 1.3 特征融合网络的改进

VGG 的出现使得大核卷积逐渐被淘汰,虽然大核卷积带来的大的感受野获得了丰富的特征信息,但随之暴涨的计算量使得其成本是昂贵的,然而最近的研究<sup>[11-12]</sup>发现通过应用深度可分离卷积可以改善这一情况,在不增加太多参数的情况下有效扩大模型的感受野。因此为了增强网络的特征提取能力,在模型的特征融合部分针对原有的 C3 模块进行改进,设计了一种由大核深度分离卷积构成 EnhancedC3(EC3)模块,如图 4 所示,它是由一个  $3 \times 3$  的普通卷积和一个  $5 \times 5$  的深度可分离卷积取代原始的 C3 模块中瓶颈层中的一个  $1 \times 1$  的普通卷积和一个  $3 \times 3$  的普通卷积,以增加有效感受野,利于对于密集目标的检测任务场景。

当特征映射到达模型的颈部时,特征的通道尺度达到最大值,宽度和高度达到最小值,因此在此处使用 GSConv<sup>[13]</sup>可以减少冗余信息,处理连接的特征映射,减少推理时间和降低检测的计算复杂度,进一步对模型进

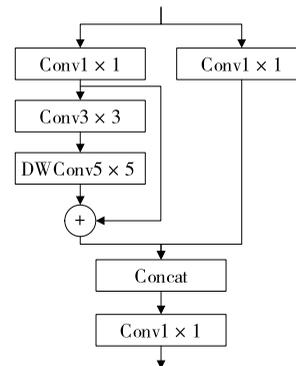


图 4 EnhancedC3 模块

Figure 4 EnhancedC3 module

行精简。GSConv (Ghost-shuffle convolution)是由普通卷积和深度可分离卷积拼接组成的卷积。如图 5 所示,对输入的特征图进行卷积后通过深度可分离卷积得到另一个特征张量。根据通道数将两个特征张量连接起来。最后,利用通道混洗操作(shuffle)将生成的信息渗透到特征

张量的各个部分。通道混洗操作是一种均匀混合策略,这种方法可以在不增加计算量的情况下完成通道间的信息融合。经过混洗操作使来自普通卷积的信息完全混合到深度可分离卷积的输出中,在不同的通道上均匀地交换本地特征信息,实现通道之间信息的流通。

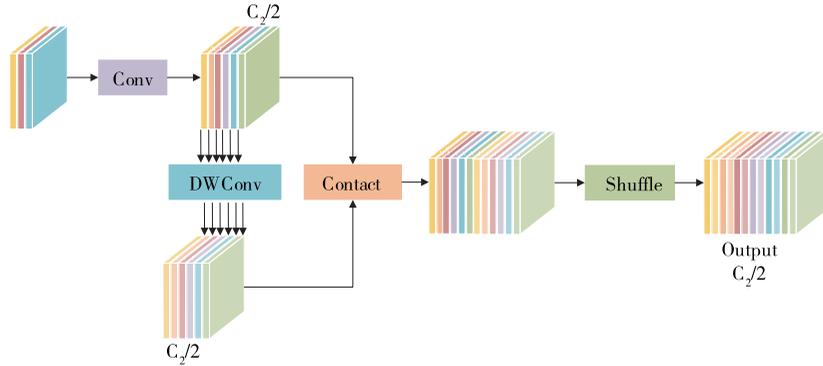


图 5 GSConv 结构图

Figure 5 The structure of GSConv

## 2 试验与结果分析

### 2.1 试验环境及数据集

试验环境的处理器型号为 Intel(R) Core(TM) i9-12900K CPU@ 3.20 GHz,内存为 128 GB,显卡型号为 NVIDIA GeForce RTX 3090(24 G 显存),开发语言使用 Python 3.8.13,深度学习框架为 Pytorch 1.11.0,CUDA 11.3,CUDNN 8.2.0。

由于水果目标检测领域缺乏可供研究的完善的开源数据集,数据集采用 kaggle 网站开源图片以及百度图库网络爬虫所得图片构建。考虑到水果种类繁多且根据地域网分布,数据集中选取具有代表性的 3 种常见水果苹果、香蕉和柑橘,图片标签分为六类,分别是新鲜苹果(fresh\_apple)、不新鲜苹果(stale\_apple)、新鲜香蕉(fresh\_banana)、不新鲜香蕉(stale\_banana)、新鲜柑橘(fresh\_orange)以及不新鲜柑橘(stale\_orange)。为了避免数据集中图片有效信息不充足对于模型训练的影响,也为了更贴合实际生活场景中水果检测的环境,该数据集中既包含实验室中单调背景下拍摄的单一水果图片,也包括日常生活场景下的同品种多数量水果和多品种的图片。数据集共包括 15 372 张图片,其中含苹果 5 496 张、香蕉 5 170 张、柑橘 4 261 张以及多品种水果图片 445 张,按照 7 : 3 的比例将其划分为训练集和测试集。

### 2.2 网络训练与评价指标

选用平均精度均值(mAP)、参数量、计算量和每秒处理帧数(FPS)作为模型的评价指标。mAP 是衡量模型对各个类别整体检测效果的常用指标,是所有类别平均精度(AP)的平均值,而 AP 是以召回率(R)为横坐标,精度

(P)为纵坐标的 PR 曲线下的面积,模型的召回率是预测正确结果在所有正样本中的比例,即“查全率”,精度是模型预测的所有结果中,预测正确的比例,即“查准率”。参数量是模型的所有参数的数量,计算量是模型的总浮点运算次数,这两者都是用来反映模型的轻量化程度。FPS 是模型每秒处理的帧数,FPS 越大则说明模型处理图片的速度越快,实时性越好。具体的计算公式:

$$P = \frac{T_P}{T_P + F_P} \times 100\%, \quad (5)$$

$$R = \frac{T_P}{T_P + F_N} \times 100\%, \quad (6)$$

$$A_P = \int_0^1 P(R) dR, \quad (7)$$

$$m_{AP} = \frac{1}{c} \sum_{i=0}^c A_{P,i}, \quad (8)$$

式中:

- P——模型所有结果中预测正确的比例,%;
- $T_P$ ——检测为正样本的正样本(真正样本);
- $F_P$ ——检测为正样本的负样本(虚假正样本);
- $F_N$ ——检测为负样本的正样本(虚假负样本);
- R——模型正样本中预测正确的比例,%;
- $A_P$ ——平均精度均值,%;
- $m_{AP}$ ——所有类别的平均精度的均值,%;
- c——类别数;
- $A_{P,i}$ ——第 i 个类别的 AP 值。

图 6 展示了训练过程中损失值下降过程。随着训练次数的递增,模型在前 25 轮次急速下降,迭代至 75 轮时趋于平缓,最终稳定在 0.015 附近。两组数据分别是采用 SiLU 和 GELU 作为主干网络激活函数时的损失变化曲

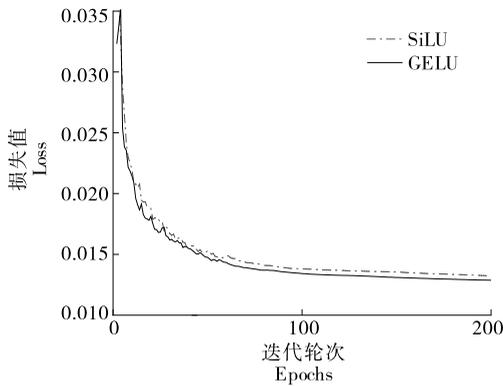


图 6 损失下降曲线图

Figure 6 Loss decline curve

线。由此可得,在相同轮次下,采用 GELU 时的损失值更小,它提高了模型损失值下降的速度,使模型更快收敛。

### 2.3 试验结果及分析

2.3.1 轻量化试验 YOLOv5 虽然检测精度高,运行速度快,但它复杂的网络结构和较大的参数量使其不易在移动设备上部署。为了降低模型参数量和运算量,现将 YOLOv5 模型的主干特征提取网络替换为轻量级网络。为了进行正确有效的模型轻量化,将研究提出的 YOLO-FFD 模型与目前已有的经典轻量级网络 MobileNetV3<sup>[14]</sup>、ShuffleNetV2<sup>[15]</sup>、GhostNet<sup>[16]</sup>、PP-LCNet,替换原 YOLOv5 中的主干网络 CSPDarknet 进行对比试验。

如表 1 所示,采用轻量级主干网络后,其模型计算量基本都得到了大幅度减少。5 组试验中,ShuffleNetV2 作为主干网络时虽然得到了较好的轻量化表现,但精度损失较大;MobileNetV3 和 PP-LCNet 精度稍微高一些但比原始网络仍有损失;GhostNet 作为主干网络损失的精度最少但轻量化程度也是最低的。以上模型都因为轻量化处理损失了检测精度,而使用 LightweightC3 模块替换主干网络中的 C3 模块时,提高检测速度的同时并未带来性能损失,不仅模型计算量降低了 26%,精度还提高了 0.79%,FPS 也上涨到 120。因此选用 LightweightC3 模块搭建模型的主干特征提取网络。

表 1 不同骨干网络的试验结果

Table 1 Experimental results of different backbone networks

主干网络	平均精度 均值/%	每秒处理帧数/ (帧·s <sup>-1</sup> )	参数量	计算量
ShuffleNetV2	74.43	73	0.22×10 <sup>6</sup>	0.5×10 <sup>9</sup>
PP-LCNet	89.12	75	0.76×10 <sup>6</sup>	1.6×10 <sup>9</sup>
MobileNetV3	90.37	51	0.80×10 <sup>6</sup>	1.2×10 <sup>9</sup>
GhostNet	93.64	52	2.44×10 <sup>6</sup>	3.7×10 <sup>9</sup>
LightweightC3	94.70	120	1.42×10 <sup>6</sup>	3.4×10 <sup>9</sup>

2.3.2 消融试验 为了验证各种改进方法的有效性,进行了一系列消融试验。以原始 YOLOv5n 模型为基准,在显卡为 RTX3090 的 PC 平台和嵌入式设备 Jetson TX2 平台上使用相同的数据集和相同的超参数分别进行 5 组消融试验,结果如表 2 所示。将原始的 C3 模块改为 LightweightC3 模块之后,计算量减少了 26%,模型得到进一步轻量化的同时精度增加了 1.39%,mAP 增加了 0.79%。在原始网络中的颈部使用 EnhancedC3 模块时,精度和 mAP 分别上涨了 2.14% 和 1.3%。EnhancedC3 的引入使模型中不同通道的特征信息得到了充分的流通与融合,提高了模型的检测精度。同时加入 LightweightC3 和 EnhancedC3 模块时,模型减少了 24% 的计算量,mAP 上涨 1.77%。采用 LightweightC3 模块、在颈部引入 EnhancedC3 模块和 GSConv 时,改进后的模型比较改进前模型,模型的精度增加了 1.81%,mAP 增加了 2.21%,模型计算量减少了 26% 的同时在 RTX 3090 上速度为原来的 3 倍,在嵌入式设备 Jetson TX2 上速度达到了 20 帧/s,整体性能得到提升。

2.3.3 与其他算法对比试验 为了验证所提 YOLO-FFD 模型的性能与优势,与目前经典的轻量化目标检测算法分别在 RTX3090 平台和嵌入式设备 Jetson TX2 平台上进行对比试验,选择的目标检测算法包括 SSD<sup>[17]</sup>、YOLOv3-Tiny<sup>[18]</sup>、YOLOv4-Tiny<sup>[19]</sup>、YOLOv5-Lite。如表 3 所示,相较于 YOLOv3-Tiny、YOLOv4-Tiny、

表 2 消融试验结果<sup>†</sup>

Table 2 Ablation results

YOLOv5n	+ EC3	+ LC3	+ GSConv	平均精度 均值/%	每秒处理帧数/(帧·s <sup>-1</sup> )		计算量
					RTX 3090 平台	Jetson TX2 平台	
√	—	—	—	93.91	55.2	11.1	4.6×10 <sup>9</sup>
√	—	√	—	94.70	119.9	18.5	3.4×10 <sup>9</sup>
√	√	—	—	95.21	82.7	16.7	4.2×10 <sup>9</sup>
√	√	√	—	95.68	164.2	17.4	3.5×10 <sup>9</sup>
√	√	√	√	96.12	172.1	20.2	3.4×10 <sup>9</sup>

† “√”表示使用该模块。

表 3 不同模型在研究数据集的对比试验

Table 3 Comparative experiments of different models in the dataset of this paper

网络	平均精度 均值/%	每秒处理帧数/(帧·s <sup>-1</sup> )		计算量
		RTX 3090 平台	Jetson TX2 平台	
SSD	88.54	101.7	1.9	274.8×10 <sup>9</sup>
Faster R-CNN	93.67	16.3	0.8	402.0×10 <sup>9</sup>
YOLOv3-Tiny	77.06	128.4	12.1	12.9×10 <sup>9</sup>
YOLOv4-Tiny	77.38	108.0	6.4	17.6×10 <sup>9</sup>
YOLOv5-Lite	92.04	62.1	15.3	2.6×10 <sup>9</sup>
YOLO-FFD	96.12	172.1	20.2	3.4×10 <sup>9</sup>

YOLOv5-Lite,研究提出的模型的运行速度和检测精度最高,浮点运算量仅次于 YOLOv5-Lite。说明研究所提出 YOLO-FFD 模型明显改善了轻量化模型所带来的精度损失,减少了检测时的漏检问题,在运行速度和检测精度上都有一定的提升,提高了在移动端设备上部署模型的可行性。

2.3.4 检测效果展示 为了直观对比改进模块对于原始模型的有效性,现对不同环境背景条件下水果图片作为测试对象进行新鲜度检测。图 7(a)为 YOLOv5n 模型误将背景物检测为目标,而 YOLO-FFD 模型检测结果正常;图 7(b)为密集场景下测试结果,原始模型不仅漏检了许多目标,而且检测到的目标也出现错检情况,而 YOLO-FFD 模型准确地分辨出了目标的新鲜程度,漏检情况也得到了极大改善;图 7(c)为出现遮挡物时检测结果,YOLOv5n 模型无法识别目标,而 YOLO-FFD 模型识别出目标,虽然存在漏检情况但相比原始模型结果得到改善;图 7(d)为多种类水果混合的场景,YOLOv5n 模型误将背景中带花纹的黄色木板错检为香蕉,多个被遮挡住的目标也未能识别,而 YOLO-FFD 模型未出现错检,漏检情况也得到改善;图 7(e)为对研究选取的 3 种水果按新鲜与否分成的 6 类对象的检测情况,YOLOv5n 模型未能完全识别出所有种类的对象,存在一处漏检和一处错检情况,此外,被遮挡的不新鲜香蕉被重复识别,出现了两个检测框,而 YOLO-FFD 模型对各个类别的目标均能准确识别,置信度也明显提高。

改进后模型的主干网络中由于 LightweightC3 模块的添加,增强了目标的位置信息,过滤掉复杂背景带来的冗余信息或冲突信息,只保留特征的关键信息,提高模型对于重要目标位置的关注度,因此模型对于目标的错检减少。同时,在模型的特征融合网络中引入了 EnhancedC3 模块和 GSConv,丰富了模型的感受野,保留了更多融合后的目标特征信息,其对于密集目标的检测能力大大提高,改进后模型在漏检方面表现优于改进前模型。试验结果表明,提出的 YOLO-FFD 模型对于水果新鲜度检测精度有明显提升,对不同种类及不同新鲜程



左图为 YOLOv5n 模型检测结果,右图为 YOLO-FFD 模型检测结果  
图 7 模型改进前后对比效果图

Figure 7 Comparison before and after model improvement

度水果在不同场景下的错检、漏检情况有明显改善。

### 3 结论

为了满足日常密集场景下水果品种及新鲜度识别的需求,研究提出了一种轻量化的水果目标检测与新鲜度

分类模型YOLO-FFD。网络基于YOLOv5n网络框架,将特征提取网络基本单元替换为由深度可分离卷积构成的轻量化模块LightweightC3可以使模型参数数量和运算量大幅减少,采用基于大核深度可分离卷积的EnhancedC3替换特征融合网络中的C3模块可以提高模型的特征融合能力,弥补因模型参数量减少而造成的精度损失;引入的GSConv替换特征融合网络中原来的普通卷积可以对模型进行进一步轻量化处理,提高模型的检测精度。YOLO-FFD模型平均精度均值达到了96.12%,在RTX 3090上的检测速度达到了172帧/s,在Jetson TX2上的检测速度达到了20帧/s,相较于原始模型平均精度均值提高了2.21%,模型计算量减少了26.1%,在RTX 3090上的检测速度增加了2倍,在嵌入式设备Jetson TX2上的速度增加了1倍,提升了部署在嵌入式设备端实时检测的可行性。

### 参考文献

- [1] ROCHA A, HAUAGGE D C, WAINER J, et al. Automatic fruit and vegetable classification from images[J]. *Computers and Electronics in Agriculture*, 2010, 70(1): 96-104.
- [2] DUBEY S R, JALAL A S. Fruit and vegetable recognition by fusing colour and texture features of the image using machine learning[J]. *International Journal of Applied Pattern Recognition*, 2015, 2(2): 160-181.
- [3] ZHANG Y D, DONG Z, CHEN X, et al. Image based fruit category classification by 13-layer deep convolutional neural network and data augmentation[J]. *Multimedia Tools and Applications*, 2019, 78: 3 613-3 632.
- [4] 刘禾, 汪懋华. 水果果形判别人工神经网络专家系统的研究[J]. *农业工程学报*, 1996(1): 171-176.  
LIU H, WANG M H. Study on neural network expert system for fruit shape judgement[J]. *Transactions of the Chinese Society of Agricultural Engineering*, 1996(1): 171-176.
- [5] SEMARY N, THARWAT A, EL-HARIRI E, et al. Fruit-based tomato grading system using features fusion and support vector machine[C]// *Intelligent Systems' 2014*. Warsaw: Springer, 2014: 401-410.
- [6] MOALLEM P, SERAJODDIN A, POURGHASSEM H. Computer vision-based apple grading for golden delicious apples based on surface features[J]. *Information Processing in Agriculture*, 2017, 4(1): 33-40.
- [7] SUHARJITO, ELWIREHARDJA G N, PRAYOGA J S. Oil palm fresh fruit bunch ripeness classification on mobile devices using deep learning approaches [J]. *Computers and Electronics in Agriculture*, 2021, 188: 106359.
- [8] GAI R L, CHEN N, YUAN H. A detection algorithm for cherry fruits based on the improved YOLO-v4 model[J]. *Neural Computing and Applications*, 2023, 35(19): 13 895-13 906.
- [9] CHEN S M, XIONG J T, JIAO J M, et al. Citrus fruits maturity detection in natural environments based on convolutional neural networks and visual saliency map[J]. *Precision Agriculture*, 2022, 23(5): 1 515-1 531.
- [10] ANANTHANARAYANA T, PTUCHA R, KELLY S C. Deep learning based fruit freshness classification and detection with CMOS image sensors and edge processors[J]. *Electronic Imaging*, 2020, 32(12): 172-177.
- [11] 陈实. 轻量化改进的YOLOv5与道路病害实时检测[D]. 武汉: 武汉大学, 2023: 23.  
CHEN S. Lightweight improved YOLOv5 and real-time detection of road damage[D]. Wuhan: Wuhan University, 2023: 23.
- [12] 赵九霄, 张馨, 史凯丽, 等. 基于改进Tiny-YOLOv5l算法的串型番茄定位与计数[J]. *食品与机械*, 2022, 38(12): 79-86.  
ZHAO J X, ZHANG X, SHI K L, et al. Localization and counting of string tomatoes based on improved Tiny-YOLOv5l algorithm[J]. *Food & Machinery*, 2022, 38(12): 79-86.
- [13] 梁秀满, 肖寒. 基于SDD-YOLO的轻量级带钢缺陷实时检测算法[J/OL]. *中国测试*. (2023-01-10) [2023-05-13]. <http://kns.cnki.net/kcms/detail/51.1714.TB.20230109.1648.002.html>.  
LIANG X M, XIAO H. Lightweight strip defect real-time detection algorithm based on SDD-YOLO[J]. *China Measurement & Test: 1-8*. (2023-01-10) [2023-05-13]. <http://kns.cnki.net/kcms/detail/51.1714.TB.20230109.1648.002.html>.
- [14] HOWARD A, SANDLER M, CHU G, et al. Searching for mobilenetv3 [C]// *Proceedings of the IEEE/CVF international conference on computer vision*. Seoul: IEEE, 2019: 1 314-1 324.
- [15] MA N, ZHANG X, ZHENG H T, et al. Shufflenet v2: Practical guidelines for efficient cnn architecture design[C]// *Proceedings of the European Conference on Computer Vision (ECCV)*. Munich: Springer, 2018: 116-131.
- [16] 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: IEEE, 2020: 1 580-1 589.
- [17] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot MultiBox detector[C]// *Computer Vision-ECCV 2016*. Amsterdam: Springer, 2016: 21-37.
- [18] REDMON J, FARHADI A. YOLO9000: Better, faster, stronger[C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Hawaii: IEEE, 2017: 7 263-7 271.
- [19] 王淑青, 黄剑锋, 张鹏飞, 等. 基于YOLOv4神经网络的小龙虾质量检测方法[J]. *食品与机械*, 2021, 37(3): 120-124.  
WANG S Q, HUANG J F, ZHANG P F, et al. Crayfish quality detection method based on YOLOv4[J]. *Food & Machinery*, 2021, 37(3): 120-124.