

基于改进 YOLOv5s 的轻量化牛油果成熟度检测方法

孙毅¹ 张双德¹

SUN Yi ZHANG Shuangde

摘要

为实现牛油果成熟度的快速精确检测,提升牛油果物流链的自动化水平,提出了一种轻量化 YOLOv5s 模型检测方法。首先,将 RepGhost 作为轻量化主干网络引入为模型的主干网络,大幅度加快检测速度;其次,添加激活函数 Mish 模块,使得模型更好地实现相关特征传递。在此基础上引入数据增强,提升模型泛化能力和模型检测的性能。结果表明,改进后的 YOLOv5s 模型对未成熟、半成熟、成熟牛油果的检测精度分别提升了 5.1、4.1、9.6 个百分点,平均精度提升了 6.3 个百分点, @map0.5 提升了 7.1 个百分点,检测速度达到 117 帧 /s。对比当前主流检测模型 YOLOv5s,改进后的模型性能更加优异,能够为牛油果物流链条提供一种轻量化、高精度的自动化成熟度检测方案。

关键词

牛油果成熟度检测: RepGhost; Mish; YOLOv5; 轻量化

doi: 10.3969/j.issn.1672-9528.2024.10.024

0 引言

牛油果也称作鳄梨、油梨、酪梨，含有丰富的不饱和脂肪酸、膳食纤维、维生素和矿物质，是一种具有极高营养价值的热带、亚热带水果，并享有“森林奶油”的美誉。随着中国加入世界贸易组织和全球化进程的加快，牛油果开始逐渐在世界各地风靡起来。经济合作与发展组织（OECD）和联合国粮农组织（FAO）发布的《中国农业展望报告（2021—2030）》显示，预计到2030年，牛油果将成为商业化程度最高的热带水果，预计出口额或将达到83亿美元，成为最具经济价值的水果之一。同时，我国海南、广东、广西、云南、四川、福建现存大量种植牛油果园^[1]。

牛油果属于后熟果品，通常会在果实未成熟的情况下采摘、存储、运输、销售。牛油果的高价值性、后熟性使得牛油果物流面临巨大的人力成本、仓储管理成本等一系列复杂问题。因此，实现自动化牛油果成熟度检测有助于大幅度削减人力成本，实现自动化果品成熟度分类，提升果品仓储品质。

目前国内外已产出大量水果相关的自动化成熟度检测研究成果^[2-9],并在涉及光谱领域、声学领域、机器视觉技术和深度学习网络模型方等领域取得一定的研究成果。刘拥民提出一种基于注意力机制的轻量化YOLOv5s蓝莓检测算法,实现了轻量化的蓝莓检测^[10];Rongli Gai提出一种改进后的YOLO-V4模型,在櫻桃目标成熟检测上表现比原版有显著

提升,平均精度 (mAP) 达到 93.6%, 优于其他 V4 版本^[11]。目前牛油果成熟度的图像检测方法数量相对较少, 存在着果品成熟检测效率较低的现象。综上所述, 水果成熟度检测的方法, 实验以轻量化主干网络模型 RepGhost 为主体架构的方法, 大幅度减少模型的参数量, 进而加快模型的成熟度识别速度。同时, 实验引入激活函数 Mish 和数据增强, 提升模型对特征的理解, 改进的 YOLOv5 网络结构图如图 1 所示。

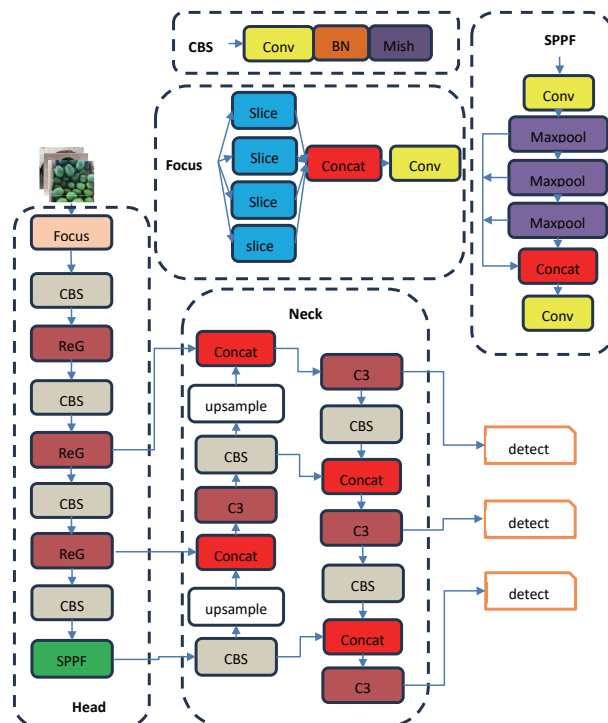


图 1 改进 YOLOv5 流程图

1. 武汉轻工大学电气与电子工程学院 湖北武汉 430023

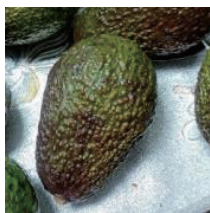
1 材料与方法

1.1 数据采集

实验研究选来自不同渠道的不同成熟度牛油果图片数据,数据集的来源渠道分为两大类:一部分来自互联网图像采集,数据集共计 1502 张;其他来自人工拍摄,使用手机型号为 Redmi K70E,共计数数据集共计 443 张。所构成数据集细分为不同成熟度,图 2 为采集图像样例。



(a) 未成熟果实: 未成熟牛油果表现为绿色或浅青色果皮,果皮表现出浅色光泽,果皮无棕色或黑色斑点。未成熟果实的果柄呈现浅色状态。



(b) 半熟果实: 半熟牛油果表现出深绿色、黑绿色,一些果实附带有有一定数量的棕色或黑色斑点。半熟果实的果柄呈绿色转向棕色状态。



(c) 成熟果实: 成熟牛油果果皮显暗绿色、暗紫色,颜色均匀,无斑点。牛油果果柄显淡棕色,果柄表现为直线状。

图 2 不同成熟度牛油果图片

1.2 数据集制作

根据不同成熟情况,将牛油果果实图片数据分为三大类别,分别是未成熟果实、半成熟果实、全熟果实。通过数据标注软件 Labelimg 对果实数据集进行相关标注,标注文件格式保存为 .xml。标注的过程中,对于果实严重遮挡和图像模糊的地方不予标注,防止模型学习到不正确的特征属性,以避免出现模型精度下降等问题。

由于数据集样本存在一定的非均衡性,未成熟果品和数据图片和标签数量远高于半熟果品和全熟果品。为提升模型训练的效果和提高模型的鲁棒性,防止产生因为局部类型的图片数量稀缺而导致的模型训练过拟合的问题。实验采取融合数据增强的方法,使用旋转、镜像、椒盐噪声、对比度、抠图等随机组合方法,对不同类型的数据集进行不同比例的数据增强(AUG),尽可能使得不同类型的数据标签数量保持均衡,增强后数据集图片共计 3153 张,数据增强后如图 3 所示,其中(a)为原图,(b)(c)为上述方法随机组合增强。



(a) 原始图像



(b) 增强后图像: 随机调整对比度、随机去除图像的一部分



(c) 增强后图像: 随机旋转图像、随机镜像、随机调整对比度

图 3 数据增强

将增强后的数据集进行分割,使得训练集、验证集、测试集以 8:1:1 的方式进行随机抽取,得到 2521:316:316 的图像数量集合。最后,将图像增强后的数据集转化为 YOLO 能够使用的 .txt 格式供模型训练。增强前后图像标签数量分布如表 1 所示。

表 1 数据集标签数量

	增强前	增强后
未成熟	3117	3554
成熟	1561	2921
半熟	712	2669

2 RepGhost 轻量化模块

传统 YOLOv5 的网络模型主要分为 YOLOv5x、YOLOv5l、YOLOv5m 等,他们的模型参数量为 205.7、109.1、49。相对轻量化 v5s 版本,它的参数也高达 16 Gflops。庞大的参数量给边缘设备的模型部署带来了巨大的挑战。传统的 v5 版本主干网络中使用了需要消耗大量计算资源的 C3 模块,而许多边缘设备是很难提供其所需的庞大计算力。因此研究使用了改进后的主干网络 RepGhost 取代了传统 C3 模块,使得模型轻量化得到巨大提升。

RepGhost 基于特征复用技术,对以往的 Ghost 进行相关改进。传统 Ghost 通过一些低成本操作来生成更多的特征图,并将他们与原始特征图 concat 在一起,进而产生更多的特征图,特征复用技术的出现扩大了通道的数量和网络容量,同时参数量也会处在一个较低的水平。虽然 concat 是一种 0paramas 的操作,但它不代表在硬件设备处理的计算成本

也是如此,这是由于参数量 FLOPs 不完全等于模型在硬件设备上运行的直接指标。鉴于复杂的内存复制, add 表现出优于 concat 的效率。此外, add 操作还具备特征融合作用,而特征融合过程是在权值空间完成的,不会引入额外的推理时间。两种方法可以用公式表示为:

$$y = \text{concat}([x, \Phi_1(x), \Phi_2(x) \dots \Phi_{s-1}(x)]) \quad (1)$$

$$y = \text{Add}([x, \Phi_1(x), \Phi_2(x) \dots \Phi_{s-1}(x)]) \quad (2)$$

式中: x 为输入, 记 $x \in \mathbf{R}^{N \times C \times H \times W}$; y 为输出, 记 $y \in \mathbf{R}^{N \times C \times H \times W}$; $\Phi_i(x), \forall i=1,2,\dots,s-1$, 记应用于 x 上的其他神经网络层。相关 RepGhost 模块重参数化过程演进如图 4 所示^[12]。

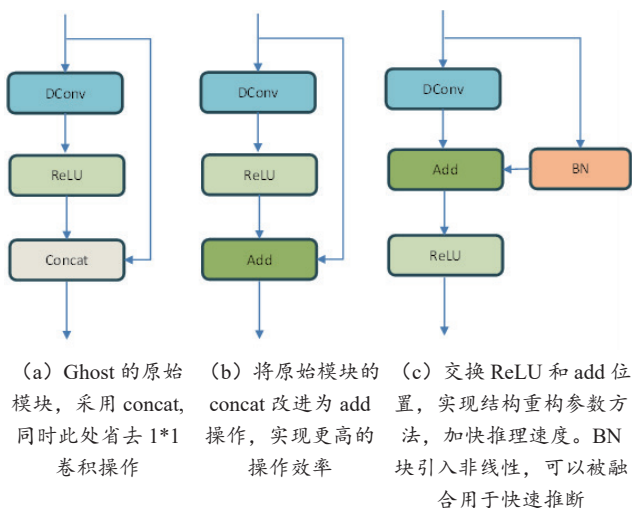


图 4 重参数化演进过程

RepGhost 颈部网络结构如图 5 所示, 其中 (a) 是 RepGhost 网络训练的 block, (b) 是 RepGhost 网络推理的 block。

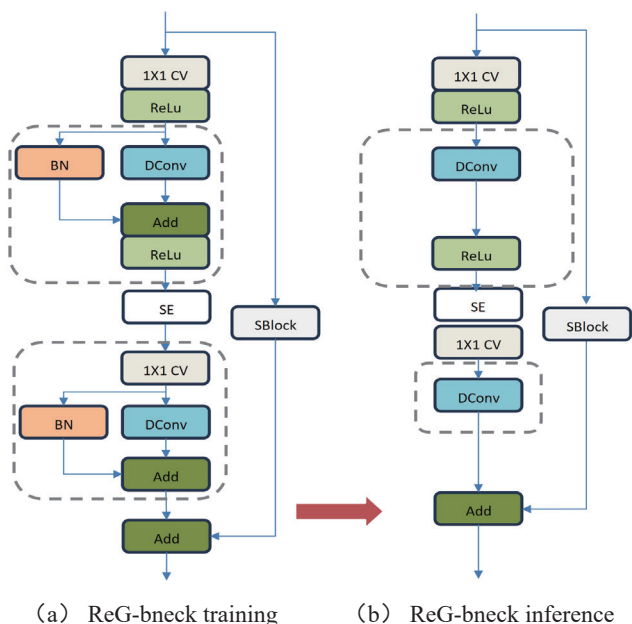


图 5 ReG 网络结构

其中 (a)、(b) 包含的 1*1 卷积和激活函数 ReLU 产生了输入通道数量减半的效果; 引入 SE 层和 1*1 卷积提升模型对特征的识别性, 同时使得通道能够与后续匹配; RepGhost 模块中输出维度是输入的 1 倍。整体 RepGhost 颈部模型在训练和推理上节约大量资源, 结构重参数化的技术使得推理过程中仅仅含有 2 个分支, 进而提升推理性能和节约内存成本。

3 激活函数 Mish

YOLOv5 原始默认使用 SiLU 作为网络的激活函数, 具体公式为:

$$\text{SiLu}(x) = \frac{x}{1 + e^{-x}} \quad (3)$$

但是, 这种激活函数在面对过大、过小的数据时, 可能导致梯度爆炸或消失的局面。引入 Mish 为改进的激活函数, 这种函数没有尖锐拐点和饱和区域, 具有更强的平滑性, 整体优于 SiLU 函数, 函数对比如图 6 所示。

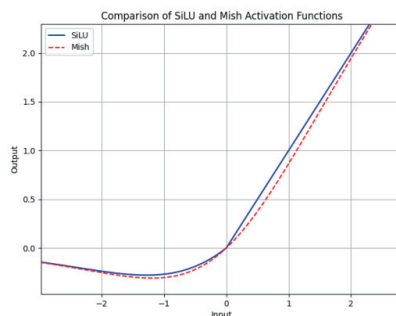


图 6 激活函数

新的激活函数具备更好的梯度传播性质, 使得模型训练更加平稳, 为增强为了加强模型的检测能力, 实验引入了新的激活函数 Mish, 具体公式^[13]为。

$$\text{Mish}(x) = x * \tanh(\ln(1 + e^x)) \quad (4)$$

实验对比了不同激活函数 SiLU、harswish、Mish, 对比结果如图 7 所示。可以明显发现激活函数 SiLU 和 hardswish 在 200 轮次训练后, 收敛性弱于激活函数 Mish。因此, 函数 Mish 具备更强的特征传递、更高的稳定性以及提高复杂环境下的检测能力。

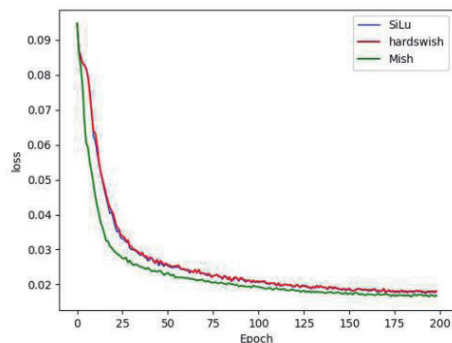


图 7 不同激活函数 loss

4 实验环境配置与相关网络训练参数

研究使用的操作系统为 Windows 11、CPU 为英特尔酷睿 i5-13400F 处理器、GPU 型号为 GTX1660-Super 6 GB、运行内存为 16 GB，编程语言为 Python3.8，使用的深度学习框架为 PyTorch，GPU 加速库 CUDA 11.7。为提高网络训练效率，设置网络初始学习率为 0.001，批处理量设置为 32，迭代次数为 200 轮。

5 模型评价指标

牛油果的成熟度检测要考虑到各种指标经行综合评价。研究选用精确率（precision， P ）、召回率（recall， R ）、平均精度值 @mAP50-90、检测速度（FPS）等指标对模型进行相关评价，相关计算公式为：

$$P = \frac{TP}{TP + FP} \times 100\%$$
 (5)

$$R = \frac{TP}{TP + FN} \times 100\%$$
 (6)

$$mAP = \frac{\sum_{n=1}^n AP(K)}{K}$$
 (7)

式中：TP 为模型正确预测为正类样本数量，FP 为模型错误预测为正类样本数量，FN 为模型错误预测为负类样本数量， K 为检测样品的类别， P 为检测正确的样本占有检测样本数量的比例， R 为检测正确的样本占有检测样本中正确样本数量的比例，mAP 为所有类别 AP 的平均值。YOLOv5s 模型和相关改后的模型指标如表 2 所示。

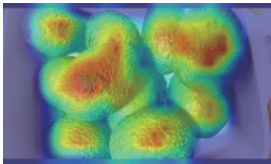
表 2 各模型性能表

	P	R	@mAP0.5	GFLOPS	FPS
YOLOv5s	0.825	0.844	0.882	15.8	64
YOLOv5s+ReG	0.84	0.823	0.894	4.3	117
YOLOv5s+ReG+Mish	0.878	0.803	0.899	4.3	117
YOLOv5s+ReG+Mish+Aug (ours)	0.888	0.89	0.953	4.3	117

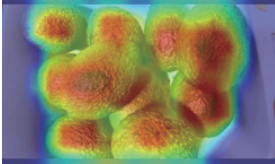
基于改进的 YOLOv5s 加速了模型识别速度，同时模型具备了更强的识别能力，实验使用热力图作为一种可视化工具来验证改进模型，以便直观地展示模型对被检测图像的关注程度。相关果实成熟的识别热力图如图 8 所示，（a）为原始被测对象，（b）为 YOLOv5s 模型的热力图，（c）为最终改进的模型热力图。从热力图上看，最终改进的模型在图像特征处理和推断上相比于（b）表现更加优异。最终检测结果效果如图 9 所示。



(a) 原图



(b) YOLOv5s

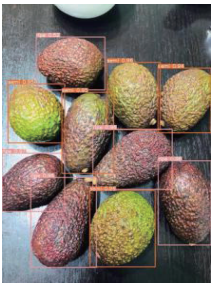


(b) ours

图 8 模型热力图



(a) 改进前



(b) 改进后

图 9 模型检测效果图

6 结论

实验提出的基于 YOLOv5 的改进轻量化网络模型，检测速度提升至 117 帧/s，准确率提升了 6.3%，召回率提升了 4.6%，@mAP0.5 提升了 7.1%。相比于原模型 YOLOv5s，改进后模型兼顾轻量化和高准确率等特点，使得模型更方便部署在边缘设备，同时能基本满足牛油果果实的成熟度识别。

参考文献：

[1] 奎国秀, 祁春节, 李冬梅, 等. 全球牛油果产业发展态势及对我国牛油果产业发展的思考 [J]. 中国果树, 2023(1): 122-127.

[2] 巨志勇, 易成, 周重臣, 等. YOLO-Rice: 一种基于 YOLOv5 的水稻虫害检测 [J/OL]. 控制工程, 1-10[2024-05-10]. <https://doi.org/10.14107/j.cnki.kzgc.20231000>.

[3] 黄华, 张昊, 胡晓林, 等. 基于改进 YOLO v5 的复杂环境下花椒簇识别与定位方法 [J]. 农业机械学报, 2024, 55(3): 243-251.

[4] 胡根生, 谢一帆, 鲍文霞, 等. 面向无人机遥感图像中茶叶枯病检测的轻量化网络 [J/OL]. 农业机械学报, 1-18[2024-05-10]. <http://kns.cnki.net/kcms/detail/11.1964.S.20240228.1722.042.html>.

[5] 司永胜, 孔德浩, 王克俭, 等. 基于 CRV-YOLO 的苹果中心花和边花识别方法 [J]. 农业机械学报, 2024, 55(2): 278-286.

Back UNet：基于自注意力机制与 UNet 的婴儿脑 MRI 分割

夏宗辉¹ 庞宇²
XIA Zonghui PANG Yu

摘要

婴儿脑影像分割在神经发育研究和早期疾病诊断中扮演着关键角色。然而，此领域的研究面临婴儿脑组织间对比度低、婴儿脑影像尺寸较大等挑战，这些问题增加了自动化分割方法的研发难度。针对上述挑战，提出一种结合自注意力机制与 UNet 的新型深度学习方法。通过引入自注意力机制，所提出的模型能够更好地捕捉图像中不同组织间的特征信息，强化对全局信息的理解，从而更准确地区分不同脑组织。在公开的婴儿脑 MRI 数据集上进行的实验结果表明，所提出的方法在常用评价指标上显示出较现有技术更优的表现。

关键词

婴儿脑分割；自注意力机制；UNet；MRI

doi: 10.3969/j.issn.1672-9528.2024.10.025

0 引言

随着医学影像技术的持续进步，磁共振成像（MRI）已成为探索和评估婴儿脑部结构和功能的重要工具。婴儿脑影像分割，即从 MRI 中准确分离出不同的脑组织，对于早期诊

断神经发育异常、监测脑部发展以及进行相关神经科学研究至关重要^[1]。准确的脑部图像分割能够为临床医生提供关键信息，帮助他们更好地理解婴儿大脑的发展状况，从而制定出更为有效的治疗策略。

虽然医学影像技术发展迅速，但是分割婴儿脑影像目前还存在多个技术挑战，这些挑战主要包括两点。其一是婴儿脑组织的对比度低：婴儿脑中，如灰质与白质之间的对比度较成人脑显著较低，这一特点使得自动分割算法难以准确区分不同的脑组织。其二是婴儿脑影像尺寸过大：婴儿脑 MRI

1. 长春金融高等专科学校计算机应用与数字媒体学院
吉林长春 130124
 2. 吉林化工学院理学院 吉林吉林 132022
- [基金项目] 2024 年长春金融高等专科学校科研规划项目
(2024JZ018, 2024JZ012)

- [6] 鄢紫, 陈良艳, 刘卫华, 等. 基于 YOLO-FFD 的水果品种和新鲜度识别方法 [J]. 食品与机械, 2024, 40(1): 115-121.
- [7] 施杰, 林双双, 罗建刚, 等. 基于 YOLOv5s 改进模型的玉米作物病虫害检测方法 [J]. 江苏农业科学, 2023, 51(24): 175-183.
- [8] 汤泽政, 伍奕桦, 徐新明, 等. 基于改进 YOLOv7-Tiny 的成熟草莓识别模型研究 [J]. 江西农业大学学报, 2023, 45(6): 1528-1542.
- [9] 朱琦, 周德强, 盛卫锋, 等. 基于 DSCS-YOLO 的苹果表面缺陷检测方法 [J/OL]. 南京农业大学学报, 1-10[2024-05-10]. <http://kns.cnki.net/kcms/detail/32.1148.S. 20231107.1546.002.html>.
- [10] 刘拥民, 张炜, 麻海志, 等. 基于注意力机制的轻量化 YOLOv5s 蓝莓检测算法 [J]. 河南农业科学, 2024, 53(3): 151-157.
- [11] GAI R, CHEN N, YUAN H. A detection algorithm for cherry fruits based on the improved YOLO-v4 model [J]. Neural com-

puting and applications, 2021, 35(19): 1-12.

- [12] ZHOU J L, ZHANG Y Y, WANG J P. RDE-YOLOv7: an improved model based on YOLOv7 for better performance in detecting dragon fruits [J]. Agronomy, 2023, 13 (4): 1042-1059.
- [13] MISRA D. Mish: a self regularized non-monotonic activation function [J/OL]. CoRR. (2019-08-23)[2024-05-13]. <https://arxiv.org/abs/1908.08681>.

【作者简介】

孙毅 (1998—), 男, 湖北武汉人, 硕士研究生, 研究方向: 计算机视觉。

张双德 (1968—), 男, 湖北应城人, 硕士, 教授, 研究方向: 传输与智能信息系统。

(收稿日期: 2024-07-21)