

基于改进 YOLO v5 的皮蛋裂纹在线检测方法

汤文权¹ 陈灼廷¹ 王东桥¹ 范维^{1,2} 王巧华^{1,3}

(1. 华中农业大学工学院, 武汉 430070; 2. 农业农村部长江中下游农业装备重点实验室, 武汉 430070;

3. 国家蛋品加工技术研发分中心, 武汉 430070)

摘要: 为了解决裂纹皮蛋分选中存在的效率低、人力成本高等问题,提出了一种基于改进 YOLO v5 的皮蛋裂纹在线检测方法。使用 EfficientViT 网络替换主干特征提取网络,并采用迁移学习对网络进行训练,分别得到 YOLO v5n_EfficientViTb0 和 YOLO v5s_EfficientViTb1 两个模型。YOLO v5n_EfficientViTb0 为轻量化模型,相较于改进前参数量减少 14.8%,浮点数计算量减少 26.8%;YOLO v5s_EfficientViTb1 为高精度检测模型,平均精度均值为 87.8%。采用 GradCAM++ 对模型可视化分析,得出改进模型减少了对背景区域的关注度,证明了改进模型的有效性。设计了视频帧的目标框匹配算法,实现了视频中皮蛋的目标追踪,依据皮蛋的检测序列实现了对皮蛋的定位和裂纹与否的判别。轻量化模型的判别准确率为 92.0%,高精度模型的判别准确率为 94.3%。研究结果表明,改进得到的轻量化模型为运算能力较差的皮蛋裂纹在线检测装备提供了解决方案,改进得到的高精度模型为生产要求更高的皮蛋裂纹在线检测装备提供了技术支持。

关键词: 皮蛋; 裂纹检测; YOLO v5; EfficientViT; 目标跟踪; GradCAM++

中图分类号: TS253.7 **文献标识码:** A **文章编号:** 1000-1298(2024)02-0384-09

OSID:



Crack Detection Method for Preserved Eggs Based on Improved YOLO v5 for Online Inspection

TANG Wenquan¹ CHEN Zhuoting¹ WANG Dongqiao¹ FAN Wei^{1,2} WANG Qiaohua^{1,3}

(1. College of Engineering, Huazhong Agricultural University, Wuhan 430070, China

2. Key Laboratory of Agricultural Equipment in Mid-lower Yangtze River,

Ministry of Agriculture and Rural Affairs, Wuhan 430070, China

3. National Egg Processing Technology Research and Development Branch Center, Wuhan 430070, China)

Abstract: With the aim to address the issues of low efficiency and high labor costs in crack detection and sorting of preserved eggs, a method for online crack detection based on an improved version of YOLO v5 was proposed. The backbone feature extraction network was replaced with the EfficientViT network, and the network was trained by using transfer learning, resulting in two models: YOLO v5n_EfficientViTb0 and YOLO v5s_EfficientViTb1. YOLO v5n_EfficientViTb0 served as a lightweight model, reducing the parameter size by 14.8% and the floating point operations by 26.8% compared with that of the original model. YOLO v5s_EfficientViTb1, on the other hand, was a high-precision detection model with an average precision mean of 87.8%. Through the utilization of GradCAM++ for model visualization and analysis, it was discovered that the improved model demonstrated a decreased focus on the background region. This finding served as evidence supporting the effectiveness of the enhancements implemented in the model. Moreover, a target box matching algorithm was designed for video frames to enable object tracking of preserved eggs in videos. Based on the detection sequence of preserved eggs, the algorithm achieved localization of the eggs and discrimination between cracked and intact ones. The lightweight model achieved a discrimination accuracy of 92.0%, while the high-precision model achieved an accuracy of 94.3%. These research findings indicated that the improved lightweight model provided a

收稿日期: 2023-06-19 修回日期: 2023-09-06

基金项目: 国家自然科学基金面上项目(32072302)和湖北省重点研发计划项目(2023BBB036)

作者简介: 汤文权(1997—),男,博士生,主要从事图像处理和深度学习研究,E-mail: 349346946@qq.com

通信作者: 王巧华(1970—),女,教授,博士生导师,主要从事农畜产品无损检测技术及装备研究,E-mail: wqh@mail.hzau.edu.cn

solution for preserved egg crack detection equipment with lower computational capabilities, while the improved high-precision model offered technical support for preserved egg crack detection equipment with higher production requirements.

Key words: preserved egg; crack detection; YOLO v5; EfficientViT; object tracking; GradCAM + +

0 引言

皮蛋作为我国传统美食,不仅风味独特,且具有抗炎症^[1]、抗肿瘤^[2]等功效,深受广大消费者的喜爱。皮蛋在腌制过程中的相互挤压,腌制完成后清洗、搬运过程中发生的磕碰等,会导致皮蛋表面产生裂纹。裂纹皮蛋不易保存,腐败变质后不仅危害人体健康,更会感染周围皮蛋、散发恶臭,严重影响皮蛋的销售,因此裂纹皮蛋的检测是皮蛋在装箱出厂之前的重要环节。目前工厂并没有一种有效的手段剔除裂纹皮蛋,而人工分选过程繁杂、费时费力、人力成本巨大,严重制约了我国皮蛋产业向工业化、自动化、智能化方向发展。

目前国内外对于鸡蛋裂纹检测方法的研究较多^[3-4],其中机器视觉技术和声波信号分析技术所具有的高效、高精度、可应用性强等优势成为鸡蛋裂纹检测的主要检测技术^[5],被国内外学者广泛使用。在利用声学信号分析技术检测鸡蛋裂纹的研究中,研究者们主要采用敲击鸡蛋的方式采集鸡蛋的声学信号,采用递归定量分析^[6]、Hilbert - Huang 变换和多重主成分分析^[7-8]、皮尔逊相关系数和多元方差分析^[9]等方法提取特征参数,建立了有效的模型。

相较于新鲜蛋,经过腌制的皮蛋蛋壳更薄,更容易因敲击而产生新的裂纹,因此本研究采用机器视觉的方法检测皮蛋裂纹。在机器视觉技术检测鸡蛋裂纹的研究中,起初研究者多采用传统图像处理的方式来检测鸡蛋裂纹,文献[10]采用负 LOG 算子和滞后阈值算法来识别鸡蛋壳上的裂纹;文献[11]提出了一种改进粒子群算法检测鸡蛋裂纹;文献[12]提出了一种基于改进蝗虫算法的模糊 C-均值算法和自适应极值逆向学习机制优化 Canny 算子进行鸡蛋裂纹检测。随着卷积神经网络 (Convolutional neural network, CNN) 的迅猛发展以及各种模型在农业领域的成功应用^[13],研究者逐步转向 CNN 来检测鸡蛋裂纹,文献[14]使用 CNN 和深度学习开发了一套自动视觉鸡蛋检测系统检测鸡蛋裂纹,文献[15]所建立 CNN 模型对鸡蛋裂纹识别准确率可达 96.92%。

鸡蛋裂纹图像获取方式多采用透射式,可使其裂纹信息在图像上的表现更加明显,然而皮蛋因其透光性差,多数皮蛋表面存在大量斑点及脏污等干

扰信息,使得皮蛋裂纹检测的难度远高于鸡蛋裂纹检测,导致目前少有皮蛋裂纹检测相关文献。王芳等^[16]通过对皮蛋的偏振光图像进行处理,初步实现了皮蛋裂纹的识别,前期已将 CNN 和深度学习应用于皮蛋裂纹在线检测中^[17],证明了机器视觉技术检测皮蛋裂纹的可行性,深度学习模型相较于传统图像处理技术拥有更强的抗干扰能力和更好的鲁棒性,在皮蛋裂纹检测中有更好的表现,但仅综合考虑皮蛋的 4 个面信息,在实际应用中效果不佳。

YOLO v5 模型的优异性能高度依赖运行设备的运算能力,在运算能力较差的嵌入式平台部署后,模型的检测速度会显著下降,难以满足在线检测对检测速度的要求。为了解决皮蛋裂纹在线检测问题,本文对 YOLO v5 模型进行改进,采用 EfficientViT 网络替换原 YOLO v5 模型的特征提取网络,以实现模型检测精度的提升和模型参数数量的减少,并开发一套适用于生产线的皮蛋目标追踪算法,以实现皮蛋多角度信息的综合判别。

1 数据获取和处理

1.1 数据采集

在如图 1 所示的三通道蛋品输送台上在线采集皮蛋表面图像,其图像采集装置结构示意图如图 2 所示,其中光源由两个条形光源对称分布于相机两侧,可确保光线分布均匀,避免光线集中。数据采集过程中,将裂纹皮蛋和完整皮蛋随机摆放在三通道蛋品输送台上,皮蛋会随着三通道蛋品输送台向前传送的同时保持自转。



图 1 三通道蛋品输送台

Fig. 1 Three-channel egg conveyor platform

为了提高模型的鲁棒性,所采集的数据包含有微小裂纹皮蛋图像、裂纹位于图像边缘的图像、未放置皮蛋的空位图像、完整皮蛋图像等多种情况,具体情况如图 3 所示。

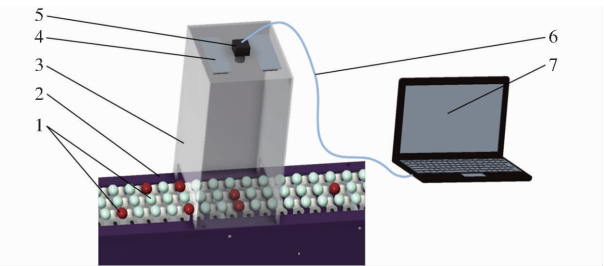


图2 皮蛋图像采集装置示意图

Fig.2 Schematic of preserved egg image acquisition device

1. 皮蛋 2. 三通蛋品输送台 3. 暗箱 4. 光源 5. 工业相机
6. 数据线 7. 计算机

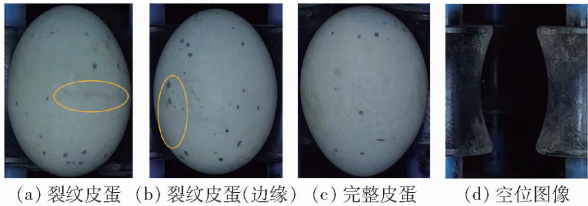


图3 皮蛋数据样例

Fig.3 Samples of preserved egg data

1.2 数据预处理

采集到的原始图像数据为 1 920 像素 × 1 080 像素,如图 4a 所示。为了便于模型训练,输入图像往往长宽相等,原始图像中的有效信息主要集中于图像中部,因此按照如图 4b 的方式将其裁剪为 1 080 像素 × 1 080 像素的图像,裁剪结果如图 4c 所示。

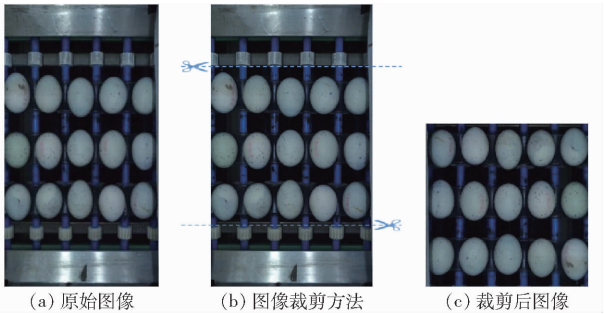


图4 数据预处理

Fig.4 Data preprocessing

本文共采集到图像 237 幅,采用 Labellmg 对图像中完全出现在视野中的皮蛋进行标注,其中完整皮蛋图像标签 1 213 个,裂纹皮蛋标签 1 139 个。按照 8:2 的比例将数据集划分训练集和验证集,数据集样本分布情况如表 1 所示。

表 1 数据集样本分布情况

Tab.1 Distribution of dataset samples

类别	图像数量	裂纹皮蛋标签数量	完整皮蛋标签数量
训练集	189	926	987
验证集	48	213	226
合计	237	1 139	1 213

2 研究方法

2.1 YOLO v5 模型

YOLO v5 包括 YOLO v5n、YOLO v5s、YOLO v5m、YOLO v5l、YOLO v5x 这 5 个不同大小和性能的版本,可以满足不同的场景需求,其模型结构如图 5 所示。其中 YOLO v5n 和 YOLO v5s 具有模型小、运算速度快的特点,常被应用于设备运算性能较差或需要实时检测的场景中,文献[18]采用改进的 YOLO v5n 模型建立了可部署在移动设备或嵌入式设备中的轻量化垃圾检测模型,文献[19-20]利用 YOLO v5s 模型实现了自然场景下的苹果花朵和油茶果的快速高精度识别和定位,模型具有较好的鲁棒性,且模型尺寸较小,适用于模型迁移和部署。YOLO v5m、YOLO v5l 和 YOLO v5x 这 3 个模型具有较强的识别能力,往往被应用于检测任务复杂、难度大和数据量大的场景下,由于本研究任务中图像背景较为单一,仅为二分类任务,分类任务比较简单,且受皮蛋检测生产线中设备运算能力的限制,故本文选用 YOLO v5n 和 YOLO v5s 作为基础网络进行研究。

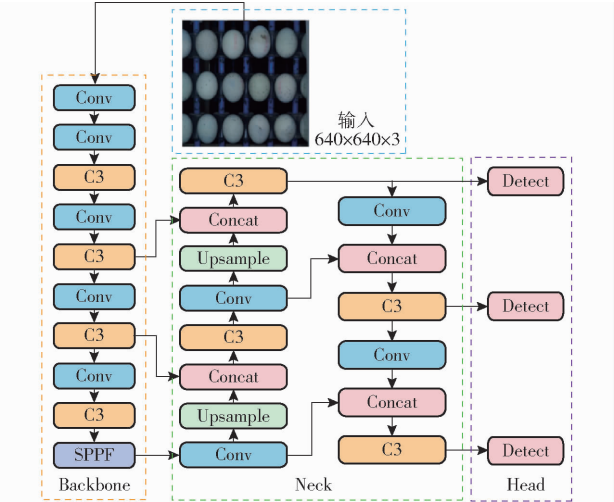


图5 YOLO v5 模型结构

Fig.5 YOLO v5 model architecture

2.2 EfficientViT 网络

EfficientViT(Efficient vision transformer) 是一种针对在运算能力有限的边缘设备上部署深度学习模型的情况下所设计的轻量级多尺度注意力模型。相较于以前的深度学习模型依赖于大量的自注意、硬件低效的大内核卷积或复杂的拓扑结构来获得良好的性能,EfficientViT 通过模型轻量化、硬件高效化的操作实现了全局感受野和多尺度学习,在保持良好性能的同时减少了模型计算量和存储需求^[21]。

Transformer 模型是自然语言处理领域的一种成功模型^[22], ViT (Vision transformer) 模型将 Transformer 模型的自注意力机制应用于计算机视觉

任务中,取得了较好的图像检测效果^[23]。然而传统 ViT 模型在运算能力有限的边缘设备上应用时受到了计算速度和存储效率的限制,为了克服这些限制, EfficientViT 采用了以下优化措施:

(1)采用轻量级基于 ReLU 的全局注意力替代自注意力机制来获得全局感受野,利用矩阵乘法的关联性和基于 ReLU 的全局注意力,将计算复杂度从二次降低到线性,同时保持功能完整。

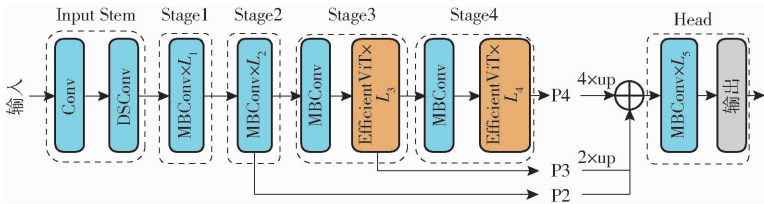


图 6 EfficientViT 模型结构

Fig. 6 EfficientViT model architecture

EfficientViT 根据各模块的复用次数组合分为性能不同的 4 个版本: EfficientViTb0、EfficientViTb1、EfficientViTb2 和 EfficientViTb3,网络各模块的复用次数如表 2 所示, EfficientViTb0 和 EfficientViTb1 因其更简单的网络结构和更少的参数量,成为了本文的研究首选。

表 2 不同版本的 EfficientViT 网络参数

Tab.2 Network parameters of different versions of EfficientViT

模型	L_1	L_2	L_3	L_4	L_5
EfficientViTb0	2	2	2	2	1
EfficientViTb1	2	3	3	4	3
EfficientViTb2	3	4	4	6	3
EfficientViTb3	4	6	6	9	3

2.2.1 DSConv 模块

DSConv (Dynamic separable convolution) 深度可分离卷积将标准卷积分解为两个步骤:深度卷积 (Depthwise convolution) 和逐点卷积 (Pointwise convolution)。深度卷积是在每个输入通道上应用一个独立的卷积核,相当于对输入的每个通道进行单独的卷积操作,可以显著减少计算量。逐点卷积是在所有输入通道上使用 1×1 的卷积核,将深度卷积的输出进行通道间线性组合。DSConv 采用量化方式显著提升了网络的速度和降低内存的使用率,同时保持较高的性能,可以很好地替换标准卷积,常用于轻量级模型的设计,使得模型能够在资源受限的环境下高效运行^[24]。

2.2.2 MBConv 模块

MBConv 由一个 1×1 的普通卷积、一个 $k \times k$ 的深度卷积、一个逐点卷积、一个 Dropout 层构成。MBConv 模块通过轻量级设计、提升模型性能、捕获

(2)采用可学习的线性层生成特征图代替图像编码过程,多尺度小卷积核和基于 ReLU 的全局注意力对特征图进行聚合,实现自注意力和全局感受野与多尺度学习的结合。

EfficientViT 的设计思想为在边缘设备上进行实时计算和推理提供了新的可能性,使得深度学习模型能够更广泛地应用于边缘计算场景,其模型结构如图 6 所示。

局部信息,可以在保持模型轻量级的同时提供较高的性能和表示能力^[25]。

2.2.3 EfficientViT 模块

EfficientViT 模块由一个轻量级多尺度全局自注意力 (Multi-scale attention, MSA) 模块和一个 MBConv 组成。轻量级 MSA 模块负责捕获全局信息,而 MBConv 负责捕获局部信息。轻量级 MSA 模块平衡了性能和效率,在 MSA 模块加入了全局感受野和多尺度学习来提供强大的性能,并用轻微的容量损失换取计算效率的显著提高。

2.3 迁移学习

迁移学习 (Transfer learning) 是一种机器学习方法,利用已经在一个任务上训练好的模型作为起点,来解决另一个类似但不完全相同的任务。在深度学习领域,特别是计算机视觉任务中,迁移学习已经成为一种非常有效的技术。卷积神经网络往往需要大量的标注数据和计算资源,迁移学习可利用预训练模型学到的特征,减少对大量标注数据的需求,加快训练速度并提高模型性能,图 7 为本文数据集在原始 YOLO v5 各版本模型分别采用预训练权重和不采用预训练权重的训练效果对比,图 7 中采用平均精度均值 (Mean average precision, mAP) 作为评价指标。

由图 7 可知,在 COCO 2017 数据集上获取的预训练权重进行迁移学习,可以显著提升模型的检测效果,因此本文在训练模型时,采用 COCO 2017 数据集上训练得到的权重文件作为预训练权重进行迁移学习训练。

2.4 YOLO v5 模型改进

为了实现 YOLO v5 模型的轻量化,研究者将 MobileNetv3 作为 YOLO v5s 的主干特征提取网络,

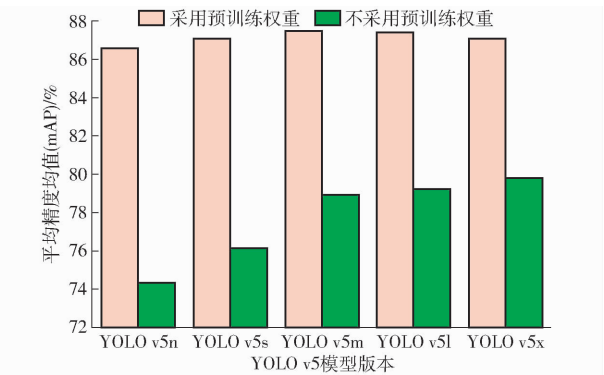


图7 采用预训练权重与不采用预训练权重模型精度对比

Fig.7 Comparison of model accuracy with pretrained weights and without pretrained weights

实现了复杂环境下的大田甘蓝实时识别^[26]和复杂果园环境中的百香果检测^[27],研究表明采用轻量化网络替换原 YOLO v5 网络的主干特征提取网络可以在保证模型精度的情况下,实现对模型的轻量化,使其满足嵌入式设备的运算需求。为了进一步减少 YOLO v5 模型的参数量,实现模型的轻量化,降低模型对硬件资源的占用,提升模型的检测性能,本文采用 EfficientViT 的特征提取网络,替换原 YOLO v5 模型的特征提取模块 Backbone 网络,并将原模型的倒金字塔与 YOLO v5 的检测头相连接,改进模型结构如图 8 所示。

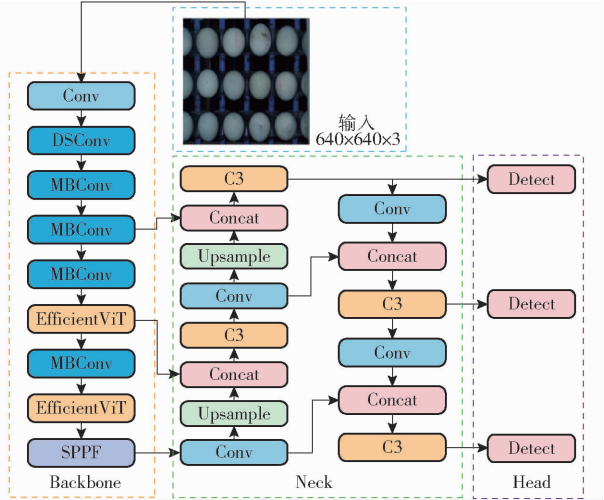


图8 改进 YOLO v5 模型结构

Fig.8 Improved YOLO v5 model architecture

3 模型训练

3.1 训练参数配置

本文针对皮蛋裂纹在线检测问题,使用 Pytorch 框架来改进 YOLO v5 网络结构,为了对比不同模型的检测效果和改进算法的有效性,模型训练环境保持一致。模型的训练平台为:128 GB 内存,搭载 AMD Rezen Threadripper 2920X CPU, GPU 为 NVIDIA RTX2080ti;运行环境为:Windows 10 (64

位)操作系统,Python 3.8 编程语言版本,Pytorch 1.9 深度学习框架,CUDA 版本为 11.3,运行软件为 Pycharm 2020.1。

3.2 评价指标

本文采用 mAP 作为目标检测模型精度的评价指标,参数量和浮点运算数 (Floating point operations, FLOPs) 分别评价模型的空间复杂度和时间复杂度。在目标检测任务中真实区域与预测区域的交叉面积与整体面积的比值被称为交并比 (Intersection over union, IoU),当 IoU 大于阈值 0.5,且检测框预测类别与真实框类别相同,标记为真阳性 (True positive, TP),否则为假阳性 (False positive, FP);真阴性 (True negative, TN)和假阴性 (False negative, FN)同理。精确率 (Precision, P) 为正类且预测正确占有所有正类的比例,召回率 (Recall, R) 为预测出正类占有所有正类的比例。平均精度 (Average precision, AP) 为每一类 $P-R$ 曲线面积的平均。本文采用准确率来评价目标追踪的检测分选效果。

3.3 训练结果对比

本文选取的基础网络分别为 YOLO v5n 和 YOLO v5s,并采用 EfficientViTb0 和 EfficientViTb1 分别对 YOLO v5n 和 YOLO v5s 进行改进,得到了 YOLO v5n_EfficientViTb0 和 YOLO v5s_EfficientViTb1 这两个改进网络。MobileNetv3 和 EfficientNet 常用于替换 YOLO v5 的特征提取网络,本文分别采用 MobileNetv3small 和 EfficientNetb0 替换 YOLO v5n 的特征提取网络,MobileNetv3large 和 EfficientNetb1 替换 YOLO v5s 的特征提取网络,作为本文改进模型的对比网络。在训练 YOLO v5n 和 YOLO v5s 网络时,采用官方给出的预训练权重进行迁移学习训练;在训练改进模型时,本文先将改进网络在 COCO 2017 数据集上进行训练,再将收敛后得到模型权重在本文数据集上进行迁移学习训练,训练得到的各模型性能如表 3 所示。

由表 3 可知, YOLO v5n_EfficientViTb0 相较于 YOLO v5n 网络参数量减少 14.8%,浮点数计算量减少 26.8%,精度略有提升,适合在计算能力较差的硬件设备上部署,相比之下采用 MobileNetv3small 对 YOLO v5n 的改进将会使模型精度大幅下降, EfficientNetb0 虽然可以提升模型精度,但同时也会大幅提升模型的参数量和浮点数计算量; YOLO v5s_EfficientViTb1 相较于 YOLO v5s 虽然网络参数量有所增加,但浮点数计算量减少 12.7%,且 YOLO v5s_EfficientViTb1 网络拥有更高的检测精度,适合

表 3 不同模型性能对比

Tab. 3 Performance comparison of different models

	模型	mAP/%	参数量	浮点计算量
轻量化模型	YOLO v5n	86. 6	1.76×10^6	4.1×10^9
	YOLO v5n_MobileNetv3small	82. 0	1.68×10^6	2.7×10^9
	YOLO v5n_EfficientNetb0	86. 8	4.86×10^6	8.7×10^9
	YOLO v5n_EfficientViTb0	86. 8	1.50×10^6	3.0×10^9
高精度模型	YOLO v5s	87. 1	7.03×10^6	1.58×10^{10}
	YOLO v5s_MobileNetv3large	87. 2	5.89×10^6	1.02×10^{10}
	YOLO v5s_EfficientNetb1	87. 6	9.51×10^6	1.61×10^{10}
	YOLO v5s_EfficientViTb1	87. 8	7.88×10^6	1.38×10^{10}

在计算能力较强且对精度要求较高的硬件设备上部署,相比之下采用 MobileNetv3large 对 YOLO v5s 的改进可大幅减少模型的参数量,但对模型精度的提升不明显, EfficientNetb1 对模型精度提升效果不如 EfficientViTb1,而模型参数量和浮点数计算量却远高于 EfficientViTb1。

4 结果分析

4.1 基于 GradCAM ++ 的热力图可视化分析

深度学习模型在进行分类预测时,通常会对输入图像的不同区域产生不同的响应。GradCAM

(Gradient-weighted class activation mapping) 通过计算目标类别相对于模型最后一个卷积层的梯度,来确定每个空间位置对于目标类别的重要程度。通过将这些梯度与卷积层的特征图相乘并求和,生成热力图来可视化模型的注意力区域^[28]。GradCAM ++ 补齐了 GradCAM 中一个图像出现多个相同类别和目标定位差的短板^[29],故本文采用 GradCAM ++ 生成模型的热力图效果来解释 EfficientViT 网络在模型中的作用,模型可视化结果如图 9 所示,图中颜色由蓝到红表示模型对图像区域的关注程度,蓝色表示不关注,红色表示非常关注。

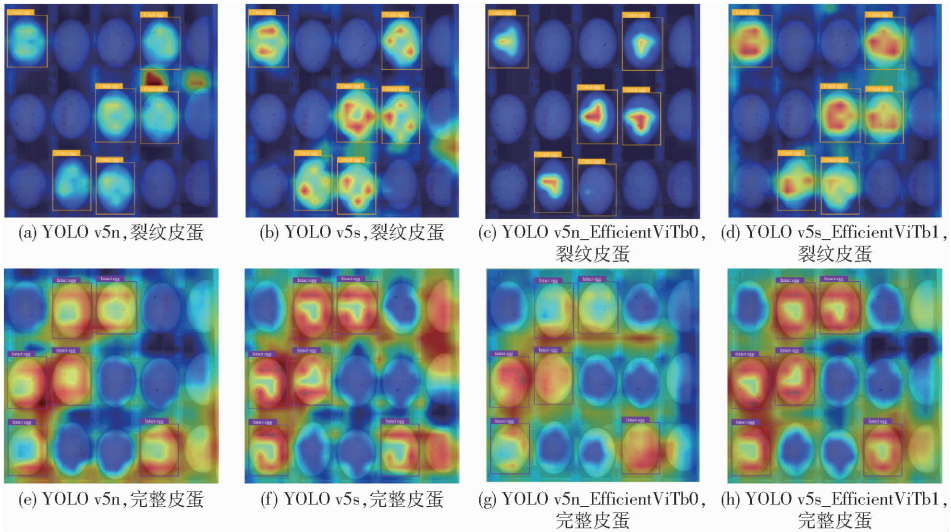


图 9 基于 GradCAM ++ 的可视化结果

Fig. 9 Visualization results based on GradCAM ++

由图 9 可知,基础模型在检测过程中关注了过多的背景区域,这导致了检测精度的下降,而采用 EfficientViT 网络改进之后,由于网络本身的自注意力机制使得模型关注区域集中于皮蛋的表面。相较于小模型,大模型在进行检测时对皮蛋表面信息的关注度更高,因此大模型的精度也更高,模型对皮蛋表面细节的关注度如图 10 所示。

4.2 基于生产线的皮蛋目标追踪

4.2.1 目标追踪算法

由于皮蛋在生产线上运动的过程中,不会出

现目标被遮挡的情况,且目标的运动轨迹和移动速度相对固定,为了降低目标追踪算法的计算量,本文未采用传统目标追踪算法 DeepSORT 算法,仅采用目标框的重合度来匹配同一目标,对同一目标进行追踪,并将同一目标的所有检测结果进行统计,根据统计结果进一步判断模型对裂纹蛋的判别准确率,其程序工作流程图如图 11 所示,图中目标框为调用模型得到皮蛋的目标框,候选框为已被分配编号的上一帧目标框,目标追踪效果如图 12 所示。

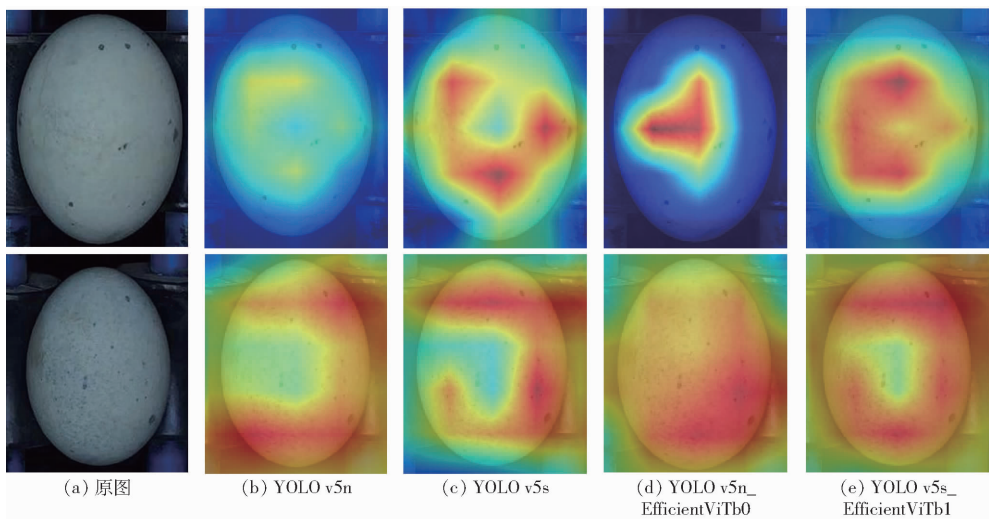


图 10 可视化结果局部放大图

Fig. 10 Visualization of locally magnified images

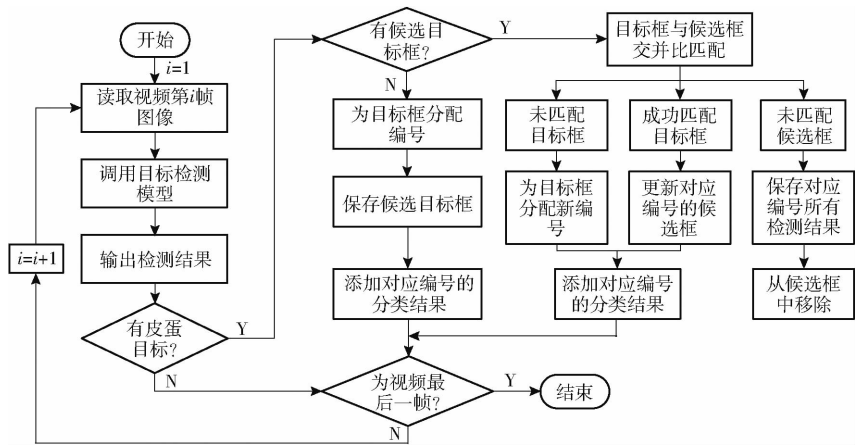


图 11 目标追踪流程图

Fig. 11 Object tracking process flowchart

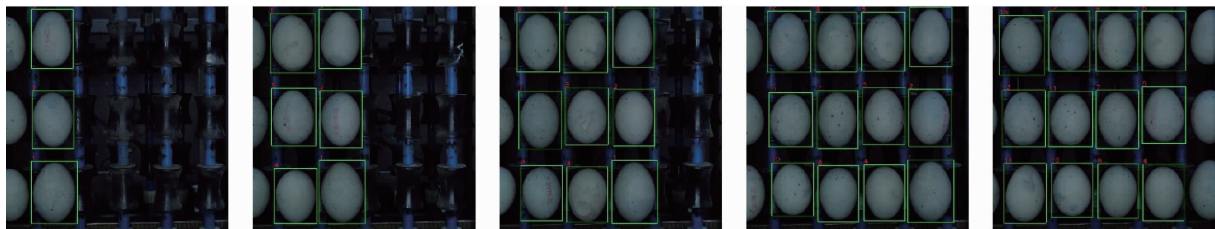


图 12 目标追踪效果

Fig. 12 Object tracking results

算法步骤如下：

(1)从视频的第 1 帧开始读取视频帧图像,将图像输入到训练好的目标检测模型,根据检测结果判断是否有皮蛋目标,没有则直接检测下一帧图像。

(2)若检测到皮蛋目标,则判断上一帧中是否已经获取了候选框,没有候选框则将当前帧所检测到的所有目标框从 1 开始顺序添加编号,并将检测框的分类结果存放在对应编号下。

(3)若已经有了候选框,则将当前帧的检测框与候选框计算交并比,重合度大于阈值 0.9 视为

匹配成功并分配相同的编号,将当前帧的目标框位置更新为同一编号的候选框,并作为新的候选框,将当前帧的分类结果添加到同一编号的检测结果下。

(4)未匹配成功的目标框意味着该皮蛋为新进入相机视野的皮蛋,因此为该目标框分配新的编号并添加到下一帧的候选框中。

(5)未被成功匹配的候选框意味着该候选框对应编号的皮蛋已经离开相机视野,该枚皮蛋检测结束,将该候选框移除,并将该候选框对应编号的所有检测结果保存为文本文件。

- [8] 王芳, 谭佐军, 谢静, 等. 基于声波信号的 HHT 和 Multi-PCA 无损检测鸡蛋蛋壳裂纹[J]. 华中农业大学学报, 2017, 36(4): 102-109.
WANG Fang, TAN Zuojun, XIE Jing, et al. Non-destructive detection of eggshells based on acoustic response coupling with HHT and Multi-PCA[J]. Journal of Huazhong Agricultural University, 2017, 36(4): 102-109. (in Chinese)
- [9] SUN L, FENG S, CHEN C, et al. Identification of eggshell crack for hen egg and duck egg using correlation analysis based on acoustic resonance method[J]. Journal of Food Process Engineering, 2020, 43(8): e13430.
- [10] GUANJUN B, MIMI J, YI X, et al. Cracked egg recognition based on machine vision[J]. Computers and Electronics in Agriculture, 2019, 158: 159-166.
- [11] 张健, 崔英杰. 基于改进粒子群算法的鸡蛋裂纹检测方法[J]. 食品与机械, 2020, 36(7): 136-139, 226.
ZHANG Jian, CUI Yingjie. Egg crack detection based on improved particle swarm optimization[J]. Food and Machinery, 2020, 36(7): 136-139, 226. (in Chinese)
- [12] 涂伟沪, 蔡玲霞, 李学军. 基于改进蝗虫算法优化 Canny 算子的鸡蛋裂纹图像检测[J]. 食品与机械, 2022, 38(2): 167-172, 202.
TU Weihu, CAI Lingxia, LI Xuejun. Egg crack image detection method based on improved grasshopper optimization algorithm and Canny operator[J]. Food and Machinery, 2022, 38(2): 167-172, 202. (in Chinese)
- [13] 宋怀波, 尚钰莹, 何东健. 果实目标深度学习识别技术研究进展[J]. 农业机械学报, 2023, 54(1): 1-19.
SONG Huaibo, SHANG Yuying, HE Dongjian. Review on deep learning technology for fruit target recognition[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(1): 1-19. (in Chinese)
- [14] MIZUNO N, NAKANO Y. Development of visual egg inspection system for poultry farmer using CNN with deep learning[C]//2020 IEEE/SICE International Symposium on System Integration (SII), 2020.
- [15] BOTTA B, GATTAM S S R, DATTA A K. Eggshell crack detection using deep convolutional neural networks[J]. Journal of Food Engineering, 2022, 315: 110798.
- [16] 王芳, 文友先, 谭佐军, 等. 基于偏振光和聚类分析的皮蛋壳裂纹无损检测[J]. 农业工程学报, 2014, 30(9): 249-255.
WANG Fang, WEN Youxian, TAN Zuojun, et al. Nondestructive detecting cracks of preserved eggshell using polarization technology and cluster analysis[J]. Transactions of the CSAE, 2014, 30(9): 249-255. (in Chinese)
- [17] TANG W, HU J, WANG Q. High-throughput online visual detection method of cracked preserved eggs based on deep learning[J]. Applied Sciences, 2022, 12(3): 952.
- [18] 涂成凤, 易安林, 姚涛, 等. 轻量化 YOLO v5n 的高精度垃圾检测算法[J]. 计算机工程与应用, 2023, 59(10): 187-195.
TU Chengfeng, YI Anlin, YAO Tao, et al. High-precision garbage detection algorithm of lightweight YOLO v5n[J]. Computer Engineering and Applications, 2023, 59(10): 187-195. (in Chinese)
- [19] CHEN S, XIONG J, JIAO J, et al. Citrus fruits maturity detection in natural environments based on convolutional neural networks and visual saliency map[J]. Precision Agriculture, 2022, 23(5): 1515-1531.
- [20] 宋怀波, 王亚男, 王云飞, 等. 基于 YOLO v5s 的自然场景油茶果识别方法[J]. 农业机械学报, 2022, 53(7): 234-242.
SONG Huaibo, WANG Ya'nan, WANG Yunfei, et al. *Camellia oleifera* fruit detection in natural scene based on YOLO v5s[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(7): 234-242. (in Chinese)
- [21] CAI H, LI J, HU M, et al. EfficientViT: lightweight multi-scale attention for on-device semantic segmentation[J]. arXiv preprint arXiv:220514756v3, 2023.
- [22] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[J]. Advances in Neural Information Processing Systems, 2017, 30: 5998-6008.
- [23] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16 × 16 words: transformers for image recognition at scale[J]. arXiv preprint arXiv:201011929, 2020.
- [24] NASCIMENTO M G D, FAWCETT R, PRISACARIU V A. DSConv: efficient convolution operator[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 5148-5157.
- [25] TAN M, LE Q. Efficientnet: rethinking model scaling for convolutional neural networks[C]//International Conference on Machine Learning, PMLR, 2019: 6105-6114.
- [26] 翟长远, 付豪, 郑康, 等. 基于深度学习的大田甘蓝在线识别模型建立与试验[J]. 农业机械学报, 2022, 53(4): 293-303.
ZHAI Changyuan, FU Hao, ZHENG Kang, et al. Establishment and experimental verification of deep learning model for on-line recognition of field cabbage[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(4): 293-303. (in Chinese)
- [27] 罗志聪, 李鹏博, 宋飞宇, 等. 嵌入式设备的轻量化百香果检测模型[J]. 农业机械学报, 2022, 53(11): 262-269, 322.
LUO Zhicong, LI Pengbo, SONG Feiyu, et al. Lightweight passion fruit detection model based on embedded device[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(11): 262-269, 322. (in Chinese)
- [28] SELVARAJU R R, COGSWELL M, DAS A, et al. Grad-CAM: visual explanations from deep networks via gradient-based localization[C]//Proceedings of the IEEE International Conference on Computer Vision, 2017: 618-626.
- [29] CHATTOPADHAY A, SARKAR A, HOWLADER P, et al. Grad-CAM++: generalized gradient-based visual explanations for deep convolutional networks[C]//2018 IEEE Winter Conference on Applications of Computer Vision (WACV), IEEE, 2018: 839-847.