

基于深度学习的移动端缺陷蛋检测系统研究

范 维¹ 胡建超¹ 王巧华^{1,2} 汤文权¹

(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)

摘要: 针对缺陷鸡蛋差异性大、人工检测主观性强、实时性差, 消费者存在食品安全隐患等问题, 提出一种基于深度学习的移动端缺陷蛋无损检测系统, 实现对裂纹蛋和血斑蛋的实时检测。首先, 建立改进的轻量级卷积神经网络 MobileNetV2_CA 模型, 以 MobileNetV2 原网络为基础, 通过嵌入坐标注意力机制、调整宽度因子、迁移学习等操作对其进行优化, 并进行 PC 端检测对比试验。试验结果表明: 建立的 MobileNetV2_CA 模型验证集准确率达 93.93%, 召回率为 94.73%, 单个鸡蛋平均检测时间为 9.9 ms, 对比改进前 MobileNetV2 模型准确率提升 3.60 个百分点、召回率提升 4.30 个百分点、检测时间缩短 2.62 ms; MobileNetV2_CA 模型的参数数量为 2.36×10^6 , 较原 MobileNetV2 网络模型降低 31.59%。然后, 利用 NCNN 深度学习框架对 MobileNetV2_CA 模型进行训练, 并通过格式转换部署至 Android 移动端, 进行 NCNN 深度学习训练模型的移动端检测验证, 及其与 TensorFlow Lite 深度学习模型的对比分析。试验结果表明: NCNN 深度学习模型对缺陷蛋平均识别准确率达到 92.72%, 单个鸡蛋平均检测时间为 22.1 ms, 库文件大小仅 2.7 MB, 均优于 TensorFlow Lite, 更能满足实际应用要求。

关键词: 缺陷鸡蛋; 无损检测; 移动端; 深度学习; 模型优化

中图分类号: TS253; TP391

文献标识码: A

文章编号: 1000-1298(2023)03-0411-10

OSID:



Detection System Study of Defective Egg on Mobile Devices Based on Deep Learning

FAN Wei¹ HU Jianchao¹ WANG Qiaohua^{1,2} TANG Wenquan¹

(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)

Abstract: Aiming at the problems of large diversity of defective eggs, as well as the strong subjectivity and poor real-time detection of artificial detection, and the potential risk of food safety for end-consumers, a non-destructive testing system based on deep learning for defective eggs on mobile device was proposed to realize real-time detection of cracked eggs and bloody eggs. An improved lightweight convolutional neural network MobileNetV2_CA model was firstly established. MobileNetV2 network was taken as the original framework, it was further optimized by embedding coordinate attention mechanism, adjusting width factor, transfer learning and other parameters. The PC detection was also performed for comparison. Results showed that the MobileNetV2_CA model presented the validation accuracy of 93.93%, the recall rate of 94.73%, and the average detection time of 9.9 ms for a single egg, which was 3.60 percentage points higher, 4.30 percentage points higher, and 2.62 ms shorter than the original MobileNetV2 model, respectively. The parameter score of MobileNetV2_CA model was only 2.36×10^6 , which was 31.59% lower than the original MobileNetV2 network model. In addition, the NCNN deep learning framework was used to train MobileNetV2_CA model, which was further applied to Android mobile terminal through format conversion. The verification of mobile terminal detection of NCNN deep learning training model was investigated and compared with TensorFlow Lite deep learning model. Results showed that the NCNN deep learning model had an average recognition accuracy of 92.72%, an average

收稿日期: 2022-05-26 修回日期: 2022-07-13

基金项目: 国家自然科学基金面上项目(32072302, 31871863)和扬州市科技计划项目(YZ2020047)

作者简介: 范维(1987—), 男, 讲师, 博士, 主要从事农产品无损检测、数字制造与精密数控装备研究, E-mail: fanwei@mail.hzau.edu.cn

通信作者: 王巧华(1970—), 女, 教授, 博士生导师, 主要从事农产品智能化检测与装备研究, E-mail: wqh@mail.hzau.edu.cn

detection time of 22.1 ms for a single egg, and the library file size of 2.7 MB, indicating its better performance than TensorFlow Lite and meeting the requirement of practical applications. The effectiveness of the proposed system based on deep learning was finally demonstrated.

Key words: defective egg; non-destructive detection; mobile devices; deep learning; model optimization

0 引言

随着人们生活水平的提高,消费者对鸡蛋的品质安全要求与日俱增。然而,2018 年国家食品安全监督抽检结果显示,1310 批抽检鸡蛋中有 5.88% 为变质不合格蛋^[1],消费者的食品安全问题仍面临挑战。目前,鸡蛋生产企业广泛采用人工照蛋分拣,科研人员提出了机器视觉、敲击振动、光谱分析等检测方法^[2-4],针对某类特定鸡蛋缺陷均取得了较好的检测效果,但其检测装置往往依托于生产线,体积庞大且价格昂贵,无法切实服务于超市、农贸市场等自由选购的消费者。鸡蛋经过生产企业检测分装之后,在消费者购买前会经过多次运输和储存,由于受到振动、挤压及温度湿度变化等外部条件的影响,会产生不同类型的缺陷蛋^[5-6],主要包括:裂纹、血斑和散黄等。缺陷鸡蛋经过存放或消费者购买之后,极易变质,引发食品安全问题^[7],因此亟需一种直接服务于消费者的移动端缺陷鸡蛋快速无损检测系统。

基于神经网络的图像分类技术广泛应用在鸡蛋检测领域,文献[8]提出利用鸡蛋透射图的颜色信息与哈夫值单位间的关系建立误差反传(Back propagation,BP)神经网络模型,对鸡蛋新鲜度进行分级,整体识别正确率为 97.48%。文献[9]应用机器视觉系统和模糊逻辑根据鸡蛋大小和缺陷对鸡蛋进行分级,例如内部血斑、裂纹和蛋壳破损。卷积神经网络(Convolutional neural networks,CNN)减小了神经网络所需的训练参数,文献[10]提出利用 AlexNet 卷积神经网络,将早期种鸭蛋透射图像输入到卷积神经网络并搭建模型,完成对种鸭蛋早期受精信息的无损判别,准确率为 98.87%。以上研究证明了深度学习对禽蛋的无损检测具有可行性,但传统的鸡蛋检测系统都是固定在流水线上,通过工业相机和计算机端进行建模识别,无法服务于消费者,且传统的计算机端深度学习神经网络参数量大、收敛速度较慢,不利于其在移动端进行检测应用。

本文以裂纹和血斑等缺陷蛋为研究对象,根据模型计算量以及内存大小、识别准确率等因素,搭建裂纹和血斑缺陷蛋的改进 MobileNetV2_CA 识别模型并与传统 CNN 神经网络进行对比分析。通

过深度学习推理框架 NCNN 进行识别模型训练与移植,与 TensorFlow Lite 深度学习框架进行对比,以完成缺陷蛋识别模型在 Android 手机端的部署,实现基于深度学习的 Android 手机端缺陷蛋实时无损检测。

1 材料与方法

1.1 数据集准备

试验样本为江西省三清山的山林散养粉壳鸡蛋,通过模拟鸡蛋在运输中相互碰撞,获取到裂纹蛋 200 枚;自带血斑的鸡蛋样本较少,故通过高温下贮藏新鲜鸡蛋,使其内部产生黑色斑点^[9,11],获取到血斑蛋 100 枚;对比样本为正常新鲜鸡蛋 200 枚,试验鸡蛋样本共计 500 枚。本试验设计一款便携式鸡蛋图像采集装置,易于携带且满足移动端使用要求,结构如图 1 所示。

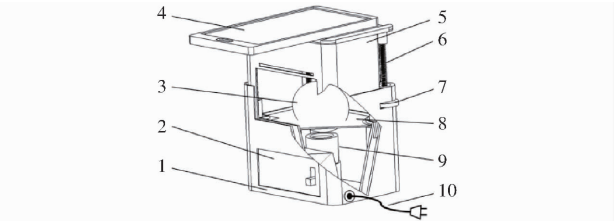


图 1 便携式鸡蛋图像采集装置示意图

Fig.1 Schematic of portable egg image acquisition device

- 1.底座 2.锂电池组 3.试验鸡蛋 4.手机 5.上盖 6.丝杠
7.螺母 8.蛋托 9.光源 10.电源线

通过手机搭配该装置采集到的鸡蛋图像含正常蛋图像 200 幅、裂纹蛋图像 200 幅、血斑蛋图像 100 幅。将分辨率为 2 090 像素×2 787 像素的鸡蛋原始图像统一裁剪为 320 像素×320 像素,以减少图像中的无用信息,结果如图 2 所示,由此加快模型训练速度,减小训练所占内存容量。

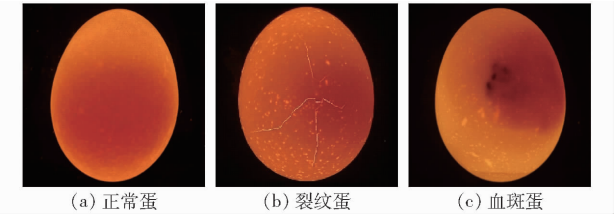


图 2 裁剪后试验鸡蛋图像

Fig.2 Test egg image after cutting

此外,为避免试验鸡蛋样本数量不均而导致训练过程中出现过拟合或泛化能力较差的问题,本试验通过平移旋转、随机裁剪、随机亮度增强、垂直翻

转4种图像增强方法同时对3个数据集的鸡蛋图像进行扩充,得到正常蛋、裂纹蛋、血斑蛋图像各1000幅,并按照7:2:1的比例划分为训练集、测试集、验证集。

1.2 试验方法

如图3所示,基于Android手机的缺陷蛋检测系统的实现主要由3个步骤组成,首先对裂纹蛋、血斑蛋、正常蛋进行图像采集与图像预处理,并针对模

型训练划分数据集;随后选择卷积神经网络进行模型训练,考虑手机受硬件资源的影响,选择轻量级网络 MobileNetV2 并对其进行改进,通过消融试验与热力图可视化 (Gradient-weighted class activation map, Grad - CAM) 对模型性能进行验证;最后将缺陷蛋识别模型进行格式转换并移植到手机端,并对软件操作界面进行开发,实现 Android 手机端对缺陷蛋进行无损检测。

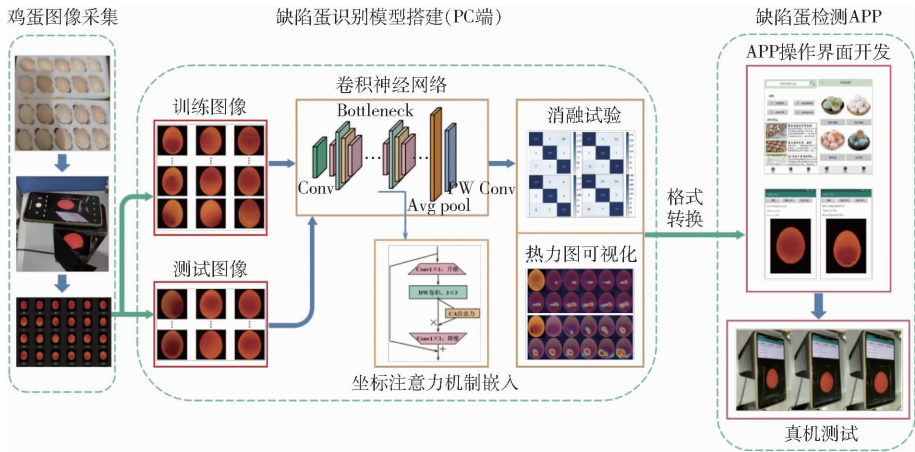


图3 基于 Android 手机的缺陷蛋检测系统实现流程
Fig.3 Implementation process of defective egg detection system based on Android mobile phone

2 缺陷蛋识别模型构建与 Android 端部署

2.1 PC 端模型构建与训练参数设计

2.1.1 改进的 MobileNetV2 模型

移动设备受到硬件资源和算力的限制,难以运行复杂的神经网络。MobileNet^[12] 作为轻量级神经网络,能较好地满足神经网络在移动端、嵌入式设备等硬件资源不足的平台进行实时识别的应用。研究人员对其进行优化,在原网络基础上引入线性瓶颈和反向残差结构并采用深度可分离卷积,提出 MobileNetV2 网络^[13],通过引入 SE 注意力机制及设

置批规范化进一步对其进行改进,提出 MobileNetV3 网络^[14]。

本文针对缺陷蛋与正常蛋的差异,通过分析模型的大小、准确率、识别速度等因素,最终以更稳定的 MobileNetV2 网络为基础网络,并对其进行优化, MobileNetV2 网络模型包含 17 个 Bottleneck 层 (每个 Bottleneck 包含两个逐点卷积层和一个深度卷积层)、1 个标准卷积层 (Conv)、2 个逐点卷积层 (PW Conv),共计 54 层可训练参数层。

针对缺陷鸡蛋的裂纹、血斑特征识别,设计改进的 MobileNetV2 网络结构如图 4 所示,网络的具体

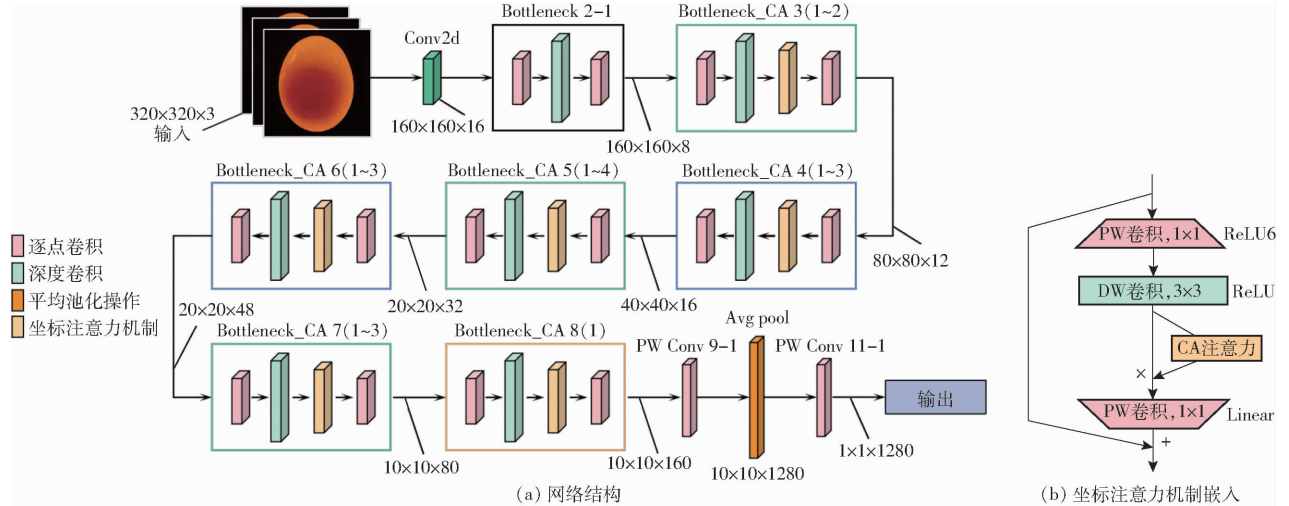


图4 改进 MobileNetV2 网络结构
Fig.4 Improved network structure of MobileNetV2

改进内容包括：

(1)将坐标注意力机制嵌入在 MobileNetV2 网络的反向残差结构中(Bottleneck 层),如图 4b 所示,在 Bottleneck 层的深度卷积之后嵌入 CA(坐标注意力)模块后再进行逐点卷积操作。网络的第 3~18 层 Bottleneck 中均嵌入 CA 模块,命名为 Bottleneck_CA,以此增强网络对重要特征的关注度,抑制噪声来提高模型精度。

(2)在精度下降可接受范围内设置宽度因子 α 为 0.75,使网络轻量化,减少模型计算量。

(3)将网络输入由 $224 \times 224 \times 3$ 改为 $320 \times 320 \times 3$,增加网络预测精度。改进后的模型以 MobileNetV2_CA 命名。

改进后的 MobileNetV2_CA 网络模型参数如表 1 所示,表中 t 、 c 和 s 分别表示扩展因子、输出特征矩阵深度和步距。

表 1 MobileNetV2_CA 网络参数

Tab.1 MobilenetV2_CA network parameters

输入尺寸	操作	t	c	s
$320^2 \times 3$	Conv1-1 3×3		32	2
$160^2 \times 32$	Bottleneck2-1	1	16	1
$160^2 \times 16$	Bottleneck_CA3-1	6	16	2
$80^2 \times 16$	Bottleneck_CA3-2	6	24	1
$80^2 \times 24$	Bottleneck_CA4-1	6	24	2
$40^2 \times 24$	Bottleneck_CA4-2	6	24	1
$40^2 \times 24$	Bottleneck_CA4-3	6	32	1
$40^2 \times 32$	Bottleneck_CA5-1	6	32	2
$20^2 \times 32$	Bottleneck_CA5-2	6	32	1
$20^2 \times 32$	Bottleneck_CA5-3	6	32	1
$20^2 \times 32$	Bottleneck_CA5-4	6	64	1
$20^2 \times 64$	Bottleneck_CA6-1	6	64	1
$20^2 \times 64$	Bottleneck_CA6-2	6	64	1
$20^2 \times 64$	Bottleneck_CA6-3	6	96	1
$20^2 \times 96$	Bottleneck_CA7-1	6	96	2
$10^2 \times 96$	Bottleneck_CA7-2	6	96	1
$10^2 \times 96$	Bottleneck_CA7-3	6	160	1
$10^2 \times 160$	Bottleneck_CA8-1	6	320	1
$10^2 \times 320$	PW Conv9-1 1×1		1 280	1
$10^2 \times 1 280$	Avg pool 10×10			
$1^2 \times 1 280$	PW Conv11-1 1×1		3	1

2.1.2 坐标注意力机制

坐标注意力机制^[15]分别从水平和垂直方向对输入特征图进行全局池化操作并得到两个独立的方向感知特征图,每个方向感知特征图都能与输入特征沿某一个方向保持依赖关系,从而得到其位置信息。图 5 为 CA 模块与传统注意力机制(Squeeze-and-excitation,SE)模块^[16]、卷积模块注意力机制(Convolutional block attention module,CBAM)模

块^[17]的结构对比,三者都是通过池化操作获取通道注意力特征,SE 模块使用全局池化操作获取通道特征信息,如图 5a 所示;CBAM 模块在 SE 模块的基础上使用全局最大池化和全局平均池化对通道特征信息进行更深入的提取,如图 5b 所示;而 CA 模块则是对输入特征的 X 和 Y 方向上进行全局平均池化,分别提取水平和垂直方向的特征信息,如图 5c 所示。CA 机制不仅解决了其他注意力机制因全局池化而造成的位置信息丢失等问题,也避免计算量的大量增加,更加适用于移动端的神经网络。

CA 机制为了能在空间上精确捕捉位置信息,分别在水平方向和垂直方向采用尺寸为 $H \times 1$ 和 $1 \times W$ 的池化核进行全局池化操作,从而获得输入特征图 X 、 Y 轴的位置信息,计算过程为

$$z_h = \frac{1}{W} \sum_{0 \leq j \leq W} F(H, j)$$

(1)

$$z_w = \frac{1}{H} \sum_{0 \leq i \leq H} F(i, W)$$

(2)

式中 W 、 H ——输入特征宽、高
 $F(i, j)$ ——特征图在 (i, j) 位置的值得
 z_w 、 z_h ——水平方向、垂直方向编码所得到的
一维特征

将这两种变换分别沿两个方向聚合,使得注意力块在一个空间方向上获取依赖关系,并在另一个空间方向上保存精确的位置信息,以此来完成坐标信息的嵌入。为了利用由此产生的表征,提出坐标注意力生成,旨在编译显著区域的权重。坐标注意力生成首先需要将水平和垂直的池化结果拼接在一起,过程为

$$f = \delta(F_1(z_h, z_w))$$

(3)

式中 F_1 ——将水平和垂直池化结果的拼接操作
 δ ——采样比

方向感知和位置敏感图计算公式为

$$g_h = \sigma(F_h(f_h))$$

(4)

$$g_w = \sigma(F_w(f_w))$$

(5)

式中 f_h 、 f_w —— f 分解的 2 个单独张量
 g_h 、 g_w —— f_h 、 f_w 通过两个 1×1 卷积最终生成的一对方向感知、位置敏感图
 σ ——拓展系数
 F_h 、 F_w ——对 f_h 、 f_w 进行 1×1 的卷积操作
函数

然后对 g_h 和 g_w 进行拓展,并通过像素点相乘互补来应用于特征图 F ,由此得到注意力的权重数据,再将各权重进行相乘获得坐标注意力输出特征 C ,计算公式为

$$C = F(i, j) g_h(i) g_w(j)$$

(6)

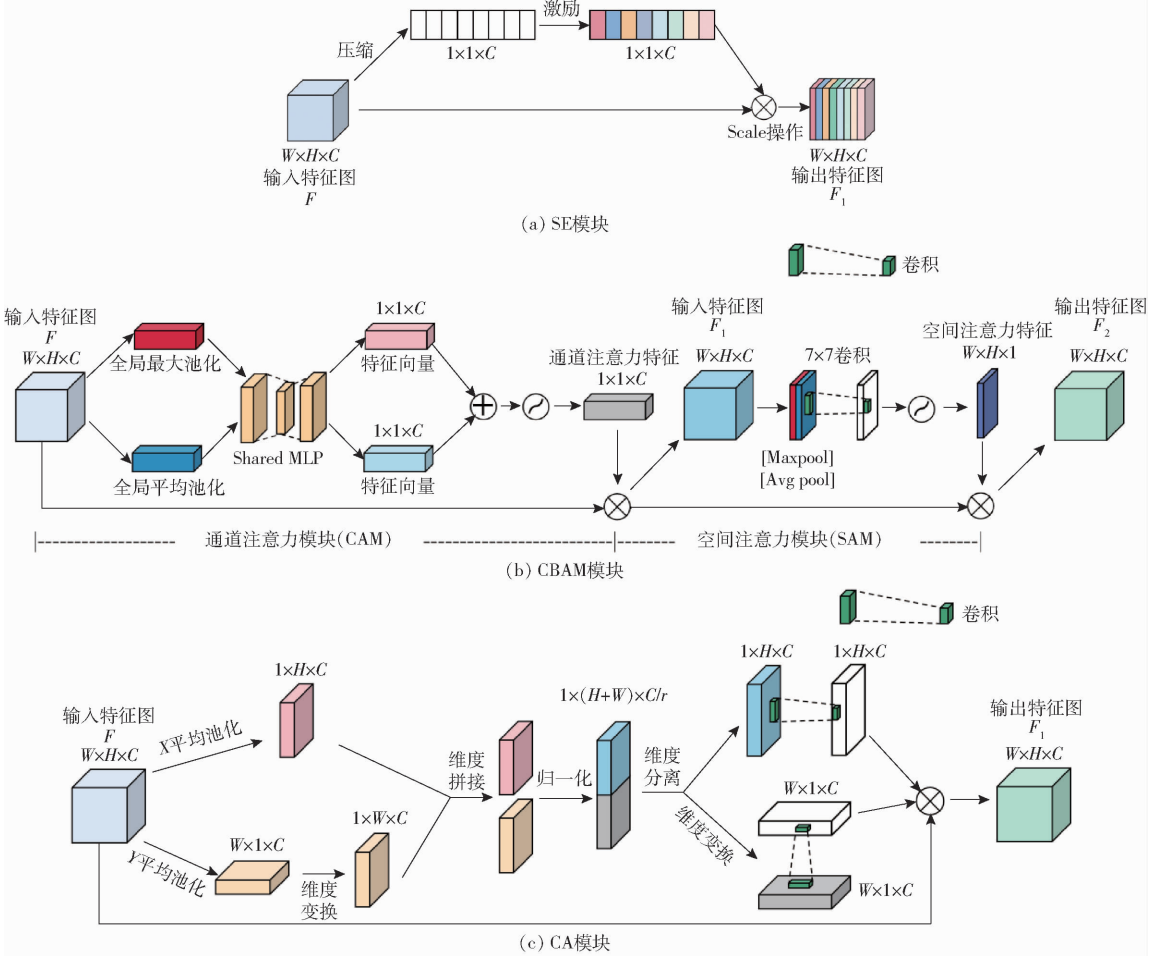


图 5 注意力机制结构对比图

Fig.5 Structure comparison of attention mechanism

2.1.3 超参数

超参数包括宽度因子和分辨率因子,宽度因子 α 可以改变神经网络输入特征的通道数,分辨率因子 β 可以改变输入特征的尺寸。本试验输入特征图已通过裁剪降低了像素值,故分辨率因子 β 设置不变。以 MobilenetV2 原型网络为基础,在 0.25 ~ 1.0 范围内设置宽度因子 α 并进行对比。宽度因子与模型的计算量、参数量、准确率呈正相关,而 α 小于 0.6 时,模型准确率低于 85%,基本不满足检测要求,故对比分析宽度因子 0.6 ~ 1.0 范围内对模型参数影响,结果如表 2 所示。综合考虑,选择宽度因子设置为 0.75,在对缺陷蛋识别准确率只降低 2.03 个百分点的情况下模型的计算量减小了 37.10%,使性能得到大幅提升,能更好地应用于移动端模型部署。

2.1.4 模型训练

模型训练硬件平台参数如下:处理器为 AMD 锐龙 7 4800H,内存为 16 GB,搭配显卡为 NVIDIA GeForce RTX 2060。软件平台:Pytorch 1.9.0 + Python 3.7。

受硬件条件约束,批次大小设置为10,试验过

表 2 不同宽度因子下的模型参数

Tab.2 Model parameters under different width factors			
宽度因子	最大准确率/%	模型计算量	模型参数量
1.00	90.33	4.67×10^8	3.45×10^6
0.95	90.14	4.46×10^8	3.22×10^6
0.90	89.84	4.11×10^8	3.07×10^6
0.85	89.45	3.70×10^8	2.81×10^6
0.80	88.95	3.27×10^8	2.52×10^6
0.75	88.36	2.64×10^8	2.17×10^6
0.70	86.43	2.42×10^8	2.09×10^6
0.65	84.11	2.17×10^8	1.94×10^6
0.60	81.38	1.84×10^8	1.73×10^6

程发现训练迭代次数超过 200 次时缺陷蛋的识别率增长开始减缓,故本模型迭代轮数设置为 300。初始学习率设置为 0.001,并采用指数衰减法不断更新,其中衰减速度设置为 100,衰减率为 0.96。利用 Adam 优化器寻找最优梯度下降方向,参数设置为默认。此外,为了避免模型从零开始学习,本研究利用原 MobileNetV2 网络在 ImageNet 数据集上的部分权重信息对 MobileNetV2_CA 模型进行迁移学习,以优化模型初始权重参数,加快模型的收敛速度。

2.2 模型移植与 Android 平台设计

2.2.1 模型移动端部署

PC 端所训练出的深度学习参数往往保存在指定的模型中,并不能应用在所有硬件平台,模型移植到手机端首先需要经过参数提取与格式转换。图 6 为缺陷蛋识别模型部署至 Android 端的原理图,其中 NCNN 是一个针对移动平台优化的高性能神经网络推理框架,相对于其他移动端深度学习框架,NCNN 没有任何第三方库依赖,支持跨平台操作,精

度损失和内存占用低,且计算速度快,适合进行移动端的模型移植^[18]。缺陷蛋模型成功部署到 Android 手机,首先需要将 Pytorch 所训练的.pth 模型文件转化为开放神经网络交换模型(Open neural network exchange,ONNX)^[19],并检查模型是否错误;随后通过 onnx2ncnn 工具将 ONNX 模型转化为 NCNN 模型,并对模型进行量化操作;最后针对 APP 的功能需求,搭建 Android 项目将 NCNN 模型部署到手机端,并进行检测精度测试。

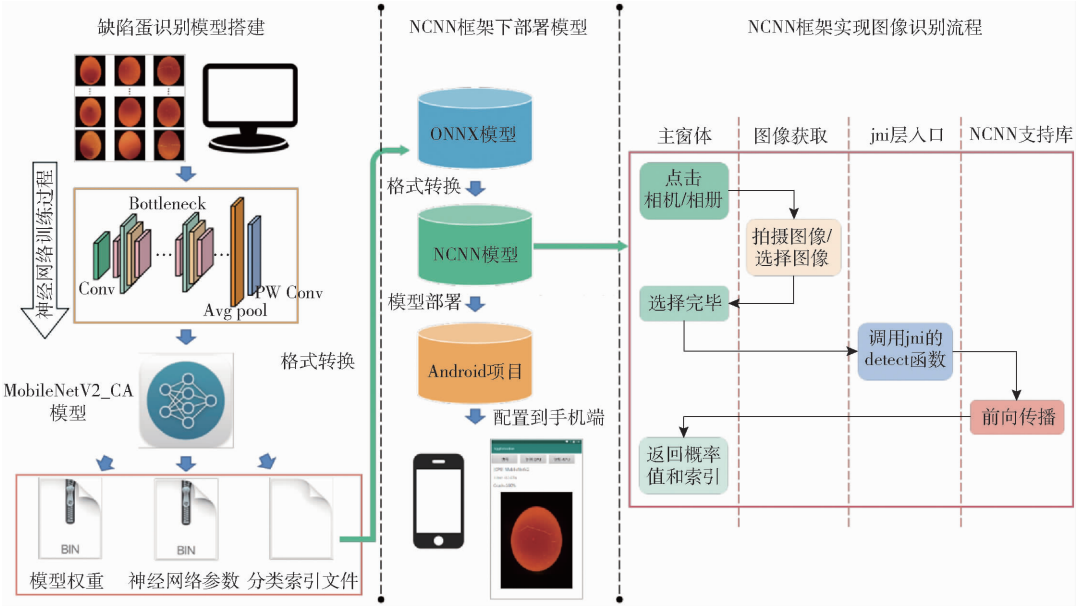


图 6 缺陷蛋识别模型部署至 Android 端流程图

Fig. 6 Flowchart of deployment of defective egg identification model to Android

2.2.2 APP 开发

所设计缺陷蛋检测 APP 主要包括:图像获取、图像预处理、缺陷蛋检测和检测结果显示 4 个模块,软件主界面如图 7a 所示。用户可通过图像获取模块用手机进行鸡蛋图像拍摄获取。图像预处理模块可将输入鸡蛋图像转变为 320 像素×320 像素,从而提高检测速度、降低背景干扰。缺陷蛋检测模块可调用 NCNN 或 TensorFlow Lite 模型进行检测分类^[20]。结果显示模块将识别预测得到的裂纹蛋、血斑蛋和正常蛋进行分类显示。

另外,缺陷蛋检测 APP 识别过程中,软件会调用手机相机并对鸡蛋进行识别,结果如图 7b、7c 所示,NCNN 框架精细的调度控制策略,使得模型在手机 CPU 和 GPU 运行下均有优化,试验发现,通过 GPU 加速下运行软件识别速率更快。

3 模型识别试验结果与对比分析

3.1 MobileNetV2_CA 模型试验

3.1.1 消融试验

为验证 PC 端 MobileNetV2_CA 网络的各项改进对训练模型的提升效果,分别针对坐标注意力机制、超参数、迁移学习进行消融试验。首先以 MobileNetV2 网络为基础,分别在相同的网络层之间添加 CA 模块、SE 模块、CBAM 模块和不嵌入注意力机制,对比测试集中裂纹蛋、血斑蛋、正常蛋的综合准确率,结果为如图 8 所示的混淆矩阵,综合准确率分别为 94.83%、93.16%、93.83%、90.33%,表明模型在添加注意力机制后准确率有明显提高,且本模型添加的坐标注意力机制在对缺陷蛋的检测上表现最好。

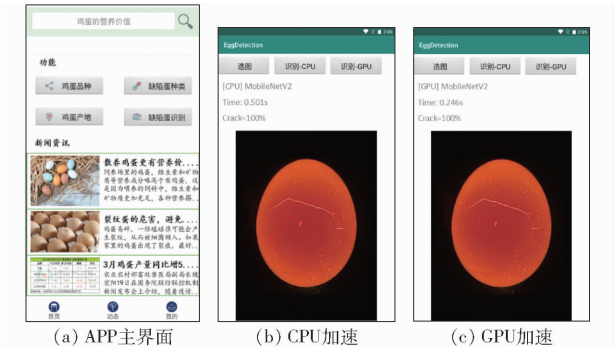


图 7 缺陷蛋识别 APP 界面

Fig. 7 Defective egg identification APP interface

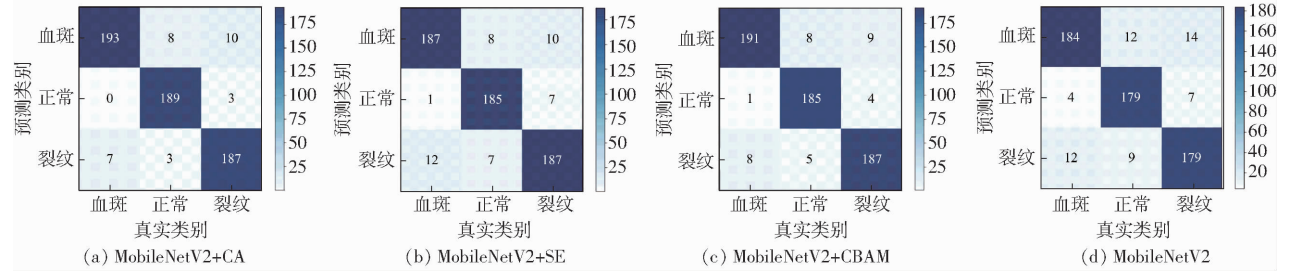


图8 MobileNetV2 嵌入不同注意力机制得到的混淆矩阵

Fig. 8 Confusion matrices obtained by embedding different attention mechanisms in MobileNetV2

随后设置宽度因子并进行迁移学习,缺陷蛋测试集的最高准确率、参数量对比结果如表3所示,分析可知,嵌入坐标注意力机制后模型最高准确率提高4.50个百分点,而模型参数量也有所增加;通过设置宽度因子为0.75后,模型准确率降低1.97个百分点,但参数量减小 1.28×10^6 ,极大地优化了模型内存大小;最后通过迁移学习,使模型收敛更好,准确率提升1.07个百分点。综合对比,MobileNetV2_CA较改进前模型准确率提升3.60个百分点,模型体积减小31.59%,最终对裂纹蛋、血斑蛋、正常蛋的识别率分别达到96.43%、91.52%、93.84%,综合识别率为93.93%,为移植入手机端提供技术支持。

表3 MobileNetV2_CA 消融试验结果

Tab.3 MobileNetV2_ CA ablation test results

网络模型	因素			参数 量	最高准 确率/ %	召回 率/%	特异 度/%
	坐标注 意力	宽度 因子	迁移 学习				
MobileNetV2	√	0.75	√	3.45×10^6	90.33	90.43	95.17
				2.17×10^6	88.36	87.13	92.77
				3.64×10^6	94.83	94.93	97.40
				3.45×10^6	91.40	92.97	95.93
				2.36×10^6	93.93	94.73	97.33

3.1.2 Grad-CAM 可视化结果

随机挑选测试集中的裂纹蛋和血斑蛋图像各一幅,通过 Grad-CAM^[21]的方法生成 MobileNetV2_CA 模型主要层的可视化热力图,结果如图9所示。由图9可知,随着网络的不断传播,模型不断地调整所关注的区域,每一次卷积操作之后,模型对裂纹特征和血斑特征的关注度不断增加,热力图表达也更有区分度。对于裂纹蛋,模型首先关注到裂纹聚集处的特征信息;对于血斑蛋,模型则是首先关注到鸡蛋中心区域的特征信息。

3.1.3 MobileNetV2_CA 与轻量级 CNN 网络对比

为进一步验证分析所提出优化模型移植到 Android 手机端的识别性能,随机选取60个验证集样本,选取5种轻量级CNN网络 MobileNetV3 -

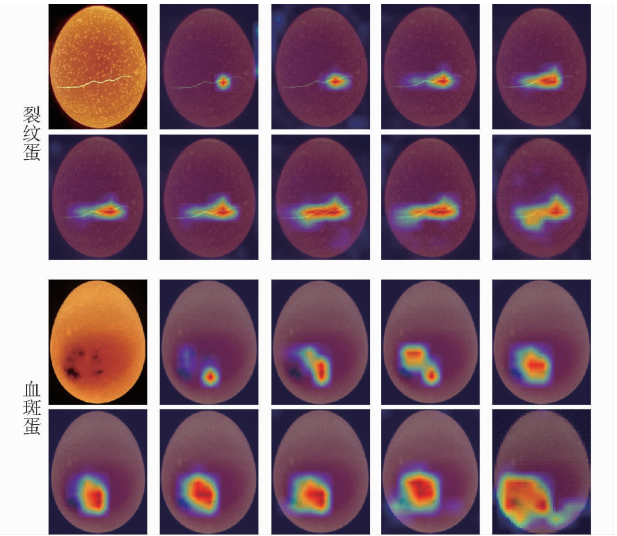


图9 模型各层特征可视化热力图

Fig. 9 Visual thermodynamic diagram of characteristics of each layer of model

Large、MobileNetV3 - Small、SqueezeNet^[22-23]、MobileNetV2、EfficientNet^[24]模型与 MobileNetV2_CA 进行试验对比,同时记录模型每个训练周期的训练损失率和测试集准确率,便于掌握模型的训练情况。

图10为该6个网络的训练过程,图10a和图10b分别表示准确率和损失值随迭代次数增加的变化曲线。结果显示,迭代次数超过150次之后,各个网络训练的准确率和损失值均趋于平稳,对比迭代300次后各个模型对缺陷蛋测试的准确率、召回率、特异度、参数量及平均检测时间,结果如表4所示。

从表4中可知,EfficientNet 模型能较好地平衡网络宽度、深度与分辨率之间的关系,在识别准确率方面达到94.87%,6个网络中表现最好;但参数量方面,EfficientNet 模型表现不佳,本文改进的 MobileNetV2_CA 模型参数量为 2.36×10^6 ,占用内存最小;此外,各模型的检测速度对比如图11所示, MobileNetV2_CA 模型平均检测时间为9.9 ms,表现最佳。对比改进前 MobileNetV2 模型,准确率提高3.60个百分点,召回率提高4.30个百分点,检测时

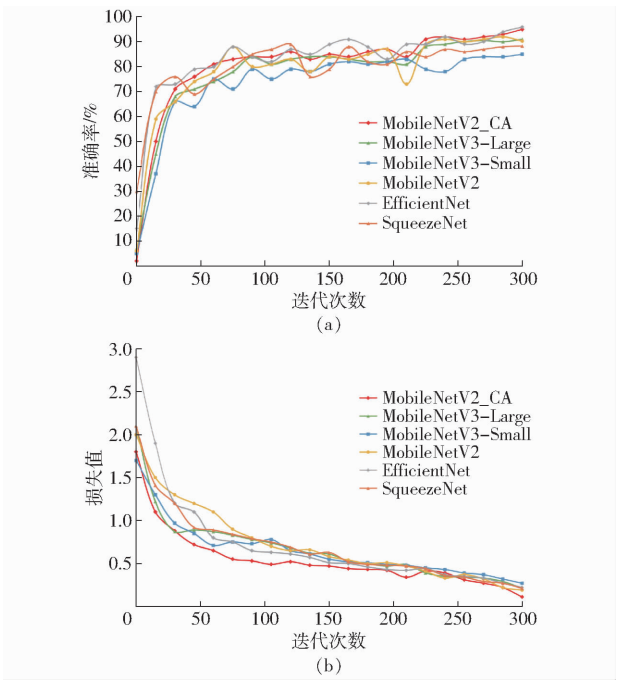


图 10 6 种卷积神经网络分类准确率及损失值对比
Fig. 10 Comparison of classification accuracy and loss rate of six convolutional neural networks

表 4 模型对比结果
Tab.4 Model comparison results

模型	参数量	准确率/%	召回率/%	特异度/%	平均检测时间/ms
MobileNetV2_CA	2.36×10^6	93.93	94.73	97.33	9.90
MobileNetV3-Large	5.34×10^6	90.97	91.26	92.14	15.18
MobileNetV3-Small	2.27×10^6	85.11	86.35	87.99	11.53
MobileNetV2	3.45×10^6	90.33	90.43	91.43	12.52
EfficientNet	1.09×10^7	94.87	95.61	97.85	32.22
SqueezeNet	4.85×10^6	88.32	88.76	90.07	55.18

间缩短 2.62 ms。综合考虑, MobileNetV2_CA 网络在识别准确率比 EfficientNet 模型低 0.94 个百分点的情况下, 参数量只有其 21.83%, 具有低消耗、高精度的特点, 故将其作为部署到手机移动端的缺陷

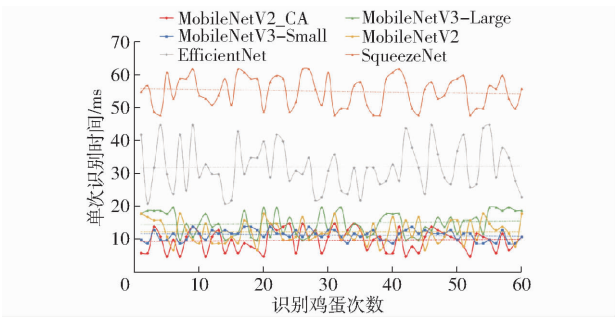


图 11 6 种卷积神经网络识别速率对比
Fig. 11 Comparison of recognition rates of six convolutional neural networks

蛋识别模型。

同时, 提取 6 种网络模型的最后一层卷积层所提取特征的可视化热力图并进行对比, 结果如图 12 所示。对于裂纹蛋, SqueezeNet 网络的特征关注度区域最小, 特征提取效果最差, EfficientNet 和 MobileNetV2_CA 的特征提取效果表现均较佳, 但是 EfficientNet 对裂纹特征的提取易受到鸡蛋边缘的影响, 稳定性不如 MobileNetV2_CA 模型。对于血斑蛋, MobileNetV3-Small 的特征提取效果表现最差, MobileNetV2_CA 关注的区域最丰富, 效果最佳。

3.2 移动端的深度学习识别检测试验

进行移动端的深度学习识别检测试验, 在 PC 端对数据集在 Pytorch 1.9.0 框架下进行训练, 得到可供移动端移植的 pb 模型, 通过格式转换并封装, 随后部署到测试 Android 手机中, 参与测试的 Android 手机型号为 Xiaomi 8, 操作系统为 Android 8.0, 处理器为高通骁龙 845, 手机镜头焦距为 20 mm。

缺陷蛋识别 APP 的应用如图 13 所示, 手机通过暗箱顶部开口获取到鸡蛋图像并进行检测, 手机屏幕上会显示所检测鸡蛋的相关信息。图 13 分别为裂纹蛋、血斑蛋、正常蛋在 NCNN 深度学习框架训练模型的检测结果, 手机屏幕上显示着所

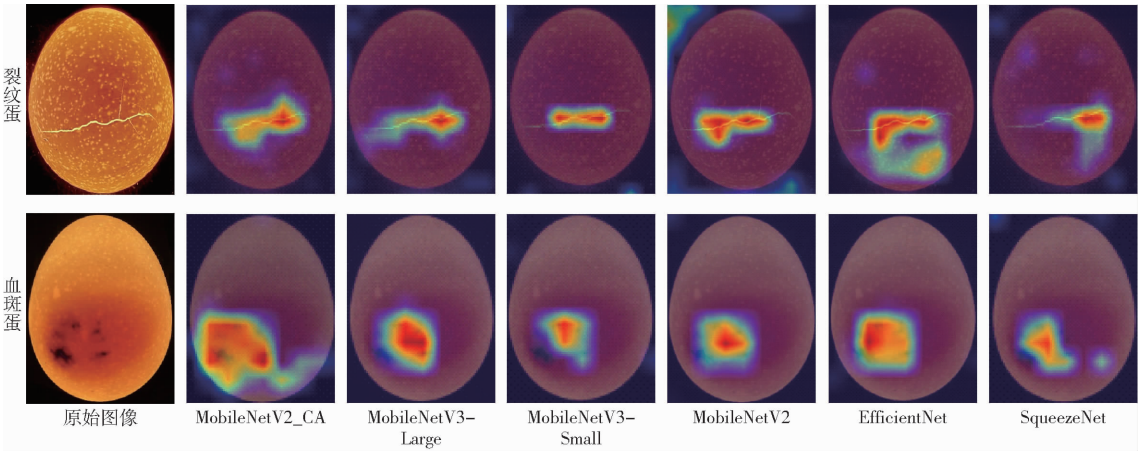


图 12 各网络提取特征可视化对比结果
Fig. 12 Visual comparison results of extracted features of each network

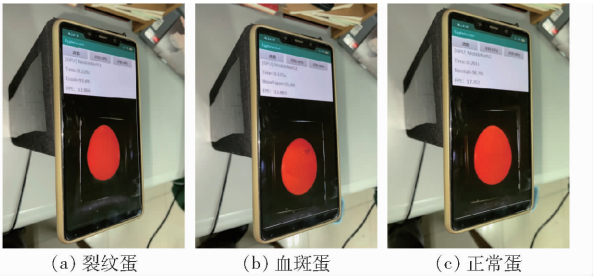


图 13 手机端缺陷蛋识别结果

Fig. 13 Identification result diagrams of defective egg on mobile terminal

用模型名称、加速方式、检测速率、识别结果等参数,初步验证了所设计深度学习缺陷蛋识别 APP 的有效可行。

为验证所设计 NCNN 深度学习框架训练模型内存和性能的均衡性,选取另一移动端深度学习框架 TensorFlow Lite^[25]进行对比,首先在 PC 端对同一数据集在 TensorFlow 2.2.0 框架下进行训练,得到.pth 模型,通过格式转换并封装,随后部署到同一款测试真机中。随机挑选 60 枚待检测鸡蛋,包括裂纹蛋、血斑蛋、正常蛋各 20 枚进行识别,对比两者的识别速率、模型精度损失、库文件大小、CPU 占用率以及模型占用内存大小等要素。

对比结果显示:①模型识别速度。GPU 加速下的 NCNN 框架对单个鸡蛋的平均检测时间为 22.1 ms。无论通过 GPU 加速还是 CPU 加速,NCNN 框架下的缺陷蛋识别速率都比 TensorFlow Lite 更快,而移动端受手机硬件限制,识别速度均低于 PC 端。②模型识别准确率,可以验证模型在移植过程中的精度损失情况。NCNN 和 TensorFlow Lite 分别为 92.72% 和 91.15%,相比于在 PC 端的识别准确率 93.93%,NCNN 损失程度更小。③库文件大小。NCNN 在不依赖第三方库的情况下,库的大小为 2.7 MB,仅为 TensorFlow Lite 的 24%,更能满足 Android 手机端不占用较大内存时使用。④软件使

用时的 CPU 占用率。NCNN 使用过程中 CPU 占用率为 TensorFlow Lite 的 3 倍左右,对手机性能要求较高,可能导致电量的消耗加快。综上,选择所设计 NCNN 深度学习框架进行缺陷蛋 MobileNetV2_CA 识别模型的训练,并部署至 Android 移动端,综合性能更好。

4 结论

- (1)建立识别准确率、体积等综合性能更适用于移动端移植的 MobileNetV2_CA 神经网络模型。通过对 MobileNetV2 原始网络嵌入坐标注意力机制、设置宽度因子、迁移学习等优化操作,使得模型在体积减小 31.59% 的情况下准确率提升 3.60 个百分点,且收敛速度更快,最终对裂纹蛋、血斑蛋、正常蛋的识别率分别达到 96.43%、91.52%、93.84%,综合识别率为 93.93%。
- (2)对比 MobileNetV2_CA、MobileNetV3 - Large、MobileNetV3 - Small、MobileNetV2、EfficientNet、SqueezeNet 网络,所建立 MobileNetV2_CA 模型验证集准确率达到 93.93%,召回率为 94.73%,单个鸡蛋平均检测时间为 9.9 ms,具有低消耗、高精度的特点。利用 Grad - CAM 可视化方法对比 6 种模型最后一层卷积网络的热力图,表明所建立 MobileNetV2_CA 稳定性好,不易受其他特征影响。
- (3)进行了移动端深度学习识别模型部署与试验分析。随机挑选裂纹蛋、血斑蛋、正常蛋各 20 枚作为检测对象,以 Android 8.0 操作系统,对比 NCNN 和 TensorFlow Lite 两种移动端深度学习框架下的模型识别性能,得到 NCNN 深度学习框架 CPU 占用率相对更高,对缺陷蛋平均识别准确率达到 92.72%,单个鸡蛋平均检测时间为 22.1 ms,库文件大小仅 2.7 MB,均能更好地满足实际检测需求,表明所提出基于深度学习的移动端缺陷蛋无损快速检测系统有效可行。

参 考 文 献

[1] 吕冰峰,刘敏,裴新荣. 2018 年鸡蛋的国家食品安全监督抽检结果分析[J]. 食品安全质量检测学报,2020,11(1):319-323.
LÜ Bingfeng, LIU Min, PEI Xinrong. Results analysis of the national food safety supervision and sampling inspection of eggs in 2018[J]. Journal of Food Safety and Quality, 2020,11(1):319-323. (in Chinese)

[2] 李庆旭,王巧华. 基于原型网络的小样本禽蛋图像特征检测方法[J]. 农业机械学报,2021,52(11):376-383.
LI Qingxu, WANG Qiaohua. Feature detection method of small sample poultry egg image based on prototypical network[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021,52(11):376-383. (in Chinese)

[3] 王明,于金莹,胡雁翔,等. 无损检测技术在鸡蛋品质检测中的应用研究进展[J]. 食品科技,2021,46(4):268-272.
WANG Ming, YU Jinying, HU Yanxiang, et al. Application of nondestructive testing technology in detection of egg quality[J]. Food Science and Technology, 2021, 46(4):268-272. (in Chinese)

[4] GENG L, YAN T, XIAO Z, et al. Hatching eggs classification based on deep learning[J]. Multimedia Tools Applications, 2018,77: 22071-22082.

[5] GUANJUN B, MIMI J, YI X, et al. Cracked egg recognition based on machine vision[J]. Computers and Electronics in

- Agriculture, 2019, 158(3):159–166.
- [6] TURKOGLU M. Defective egg detection based on deep features and bidirectional long-short-term-memory[J]. Computers and Electronics in Agriculture, 2021, 185: 106152.
- [7] ABDULLAH M H, NASHAT S, ANWAR S A, et al. A framework for crack detection of fresh poultry eggs at visible radiation[J]. Computers and Electronics in Agriculture, 2017, 141: 81–95.
- [8] 梁丹,李平,梁冬泰,等. 基于机器视觉的鸡蛋内外品质一体化检测与分级系统[J]. 中国食品学报,2020,20(11):247–254.
LIANG Dan, LI Ping, LIANG Dongtai, et al. Integrated inspection and grading system of the egg quality based on machine vision[J]. Journal of Chinese Institute of Food Science and Technology, 2020,20(11):247–254. (in Chinese)
- [9] NASIRI A, OMID M, TAHERI-GARAVAND A. An automatic sorting system for unwashed eggs using deep learning[J]. Journal of Food Engineering, 2020, 283(1):110036.
- [10] 李庆旭,王巧华,顾伟,等. 基于深度学习的种鸭蛋孵化早期受精信息无损检测[J]. 农业机械学报,2020,51(1):188–194.
LI Qingxu, WANG Qiaohua, GU Wei, et al. Non-destructive testing of early fertilization information in duck egg laying based on deep learning[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020,51(1): 188–194. (in Chinese)
- [11] 王巧华,王彩云,马美湖. 基于机器视觉的鸭蛋新鲜度检测[J]. 中国食品学报, 2017,17(8): 268–274.
WANG Qiaohua, WANG Caiyun, MA Meihu. Duck eggs' freshness detection based on machine vision technology[J]. Journal of Chinese Institute of Food Science and Technology, 2017,17(8): 268–274. (in Chinese)
- [12] HOWARD A G, ZHU M, CHEN B, et al. MobileNets: efficient convolutional neural networks for mobile vision applications[J]. arXiv preprint arXiv:1704.04861, 2017.
- [13] SANDLER M, HOWARD A, ZHU M, et al. MobileNetV2: inverted residuals and linear bottlenecks[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018:4510–4520.
- [14] SUHARJITO B, GNE A, JSP A. Oil palm fresh fruit bunch ripeness classification on mobile devices using deep learning approaches[J]. Computers and Electronics in Agriculture, 2021, 188: 106359.
- [15] HOU Q, ZHOU D, FENG J. Coordinate attention for efficient mobile network design[J]. arXiv preprint arXiv:2103.02907, 2021.
- [16] HU Jie, SHEN Li, SUN Gang, et al. Squeeze-and-excitation networks[J/OL]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019. DOI: 10.1109/TPAMI.2019.2913372.
- [17] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[C] // European Conference on Computer Vision. Springer, Cham, 2018.
- [18] ZHOU Zhongxian, SONG Zhenzhen, FU Longsheng, et al. Real-time kiwifruit detection in orchard using deep learning on Android smartphones for yield estimation[J]. Computers and Electronics in Agriculture, 2020, 179:105856.
- [19] LE T D, BERCEA G T, CHEN T, et al. Compiling ONNX neural network models using MLIR[J]. arXiv preprint. arXiv: 2008.08272, 2020.
- [20] 谷典典,石屹宁,刘讚哲,等. 基于元算子的深度学习框架缺陷检测方法[J]. 计算机学报,2022,45(2):240–255.
GU Diandian, SHI Yining, LIU Huanzhe, et al. Defect detection for deep learning frameworks based on meta operators[J]. Chinese Journal of Computers, 2022,45(2): 240–255. (in Chinese)
- [21] SELVARAJU R R, COGSWELL M, DAS A, et al. Grad-CAM: visual explanations from deep networks via gradient-based localization[J]. International Journal of Computer Vision, 2020, 128(2):336–359.
- [22] 刘阳,高国琴. 采用改进的 SqueezeNet 模型识别多类叶片病害[J]. 农业工程学报,2021,37(2):187–195.
LIU Yang, GAO Guoqin. Identification of multiple leaf diseases using improved SqueezeNet model[J]. Transactions of the CSAE, 2021, 37(2):187–195. (in Chinese)
- [23] IANDOLA F N, HAN S, MOSKEWICZ M W, et al. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5 MB model size[J]. arXiv preprint. arXiv.1602.07360, 2016.
- [24] TAN Mingxing, LE Q V. EfficientNet: rethinking model scaling for convolutional neural networks[J]. arXiv preprint. arXiv: 1905.11946, 2019.
- [25] ABADI M, AGARWAL A, BARHAM P, et al. TensorFlow: large-scale machine learning on heterogeneous distributed systems[J]. arXiv preprint. arXiv:1603.04467, 2016.