第42卷 第3期
2023年 5月

华 中 农 业 大 学 学 报
Journal of Huazhong Agricultural University

Vol.42 No.3
May 2023,115～122

康俊琪,肖德琴,刘又夫,等.基于梅尔谱图和改进ResNet34模型的鸭蛋裂纹识别算法[J].华中农业大学学报,2023,42(3):115-122.
DOI:10.13300/j.cnki.hnlkxb.2023.03.014

# 基于梅尔谱图和改进ResNet34模型的鸭蛋裂纹识别算法

康俊琪,肖德琴,刘又夫,孔馨月,殷建军

华南农业大学数学与信息学院/农业农村部华南热带智慧农业技术重点实验室,广州 510642

**摘要** 针对鸭蛋裂纹人工检测受主观性影响造成精度波动大等问题,利用ResNet34网络模型,提出1种基于梅尔谱图的鸭蛋裂纹识别算法。首先利用敲蛋装置收集敲蛋声音数据,再将音频转化成梅尔谱图,构建梅尔谱图数据集,然后搭建ResNet34模型,引入迁移学习机制训练模型,再通过Adam优化算法更新梯度,增加注意力机制模块并将卷积结构替换为深度可分离卷积以对网络模型进行改进,并调整参数进行优化,最后利用模型对鸭蛋裂纹进行识别。结果显示:改进的ResNet34DP_CA网络模型检测的平均准确率为92.4%,对比原始ResNet34网络模型,平均准确率提高5.5个百分点,参数量减少32%;对比其他网络模型VGG16、MobileNetv2和EfficientNet,平均准确率分别提高10.9、13.7、16.3个百分点,识别时间为21.5 ms。结果表明,所提出的基于梅尔谱图和改进ResNet34模型的鸭蛋裂纹识别算法,能够有效地对鸭蛋裂纹进行检测识别。

**关键词** 梅尔谱图;无损检测;深度学习;模型优化;卷积神经网络;鸭蛋裂纹识别

**中图分类号** TP18;S879.3 **文献标识码** A **文章编号** 1000-2421(2023)03-0115-08

鸭蛋富含多种营养成分,易于被人体消化吸收,也可加工成皮蛋、咸蛋等蛋制品,风味独特,深受我国消费者喜爱[1]。然而鸭蛋在运输和加工过程中易破损,引起微生物污染,导致鸭蛋腐败变质,对整批鸭蛋及其后续加工造成影响,有损产业的经济效益[2]。因此,在加工前对蛋品进行分级和筛选很有必要。禽蛋分级主要是针对蛋品的裂纹、破损、质量、体积、污斑、血斑、新鲜程度等进行品质划分[3]。目前我国禽蛋裂纹检测工业化程度低,大部分工厂采用人工照蛋法和敲蛋法作为主要检测手段,检测效率低且准确度易受到操作人员水平影响。

近年来,智能技术越来越多地被应用于禽蛋检测行业,常用的有通过模拟人工检测方式的机器视觉技术与声学检测,还包括光学检测、动力学检测[4]等。Bao等[5]采用负拉普拉斯高斯(LoG)算子、滞后阈值法和局部拟合图像指标识别裂纹区域,对裂纹鸡蛋的检测准确率为92.5%。Garcia-alegre等[6]将鸡蛋的RGB图像进行数字图像处理,对鸡蛋的蛋壳外形缺陷进行识别并分类,得到分类时间最快为100 ms,分类准确率提高至82%～92%。上述研究基于禽蛋的图像进行试验,对光照、场景等条件要求严格,限制了其在实际生产中的应用。王树才等[7]通过自研机器人采集禽蛋图像,经过数字图像处理中的阈值变换、平滑、腐蚀等操作将图像进行分割,最后再利用机器臂进行分类,综合分类准确率达到94.8%。Priyadumkol等[8]研发出一种在大气和真空压力下拍摄鸡蛋外壳图像的分类装置,利用大气和真空情况下的图像差异判断裂纹,检测正确率94%。Nasiri等[9]应用VGG神经网络结合裂纹鸡蛋图像特征、选择、分类,实现裂纹鸡蛋的识别,准确率达94.5%。传统卷积神经网络检测模型参数量大,检测速度较慢,所需计算资源较多。

针对上述检测模型参数大难以部署、检测速度慢以及工厂应用化程度较低等问题,本研究提出1种基于梅尔谱图与改进ResNet34的鸭蛋裂纹识别算法(algorithm of duck-egg shell crack identification based on MEL spectrum and ResNet34,简称DCA-MR),模拟实际生产环境搭建试验装置,通过提取装置敲蛋

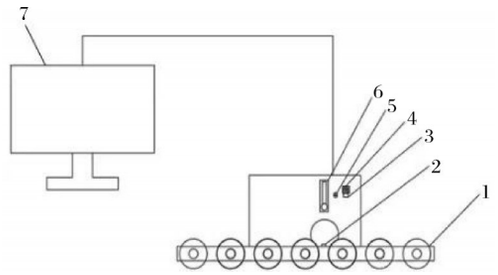
收稿日期:2022-09-29
基金项目:国家现代农业产业技术体系建设专项(CARS-42-13)
康俊琪,E-mail:kjq@stu.scau.edu.cn
通信作者:殷建军,E-mail:jianjunyin@scau.edu.cn

声音的梅尔特征谱图作为数据集，选择ResNet34网络进行结构调整优化。在敲蛋声音数据集上完成模型训练与测试，并与卷积神经网络VGG16[10]、ResNet34[11]、MobileNetv2[12]、EfficientNet[13]进行对比，以期实现裂纹鸭蛋的快速检测识别，提高生产经济效益，并为智能化、现代化家禽工厂的建设提供技术指导。

## 1 材料与方法

### 1.1 试验装置

本试验数据采自于农业农村部华南热带智慧农业技术重点实验室模拟生产线。模拟生产线由传送带、敲击装置、声音收集装置、计算机组成。传送带的尺寸为50 cm×150 cm，传送速度为5～8 m/min，匀速可调。敲蛋锤以3次/s的固定速率对传送带上鸭蛋进行敲击，通过传送带上方声音收集装置采集声音。声音收集装置采样率设置为48 kHz，采样位数为32 bit，通道数为双通道。装置示意图如图1所示。




1. 传送轴 Transmission shaft；2. 对射式光电传感器 Counter-launched photoelectric sensor；3. A/D转换器 A/D conversion；4. 微处理器 Micro processor；5. 声音收集器 Sounds collection equipment；6. 敲击锤 Knocking head；7. 计算机 Computer.

图1 鸭蛋敲击实验装置示意图

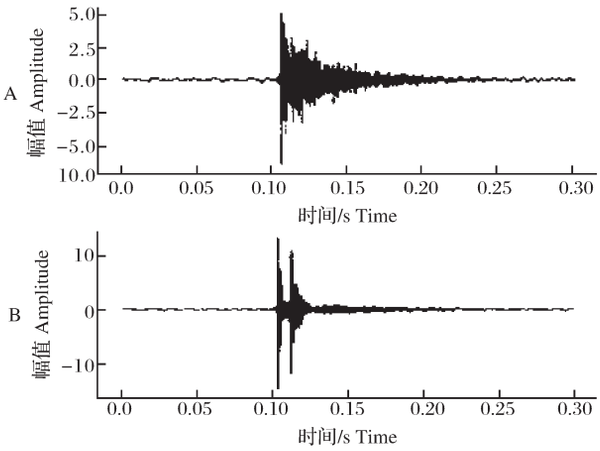
Fig.1 Schematic diagram of knocking duck egg device

### 1.2 数据采集

试验样本为市场购入的青壳鸭蛋。模拟鸭蛋在生产运输中相互碰撞，共得到110枚裂纹蛋，330枚完好鸭蛋。每批次选取10枚裂纹鸭蛋和30枚完好鸭蛋，分批置入传送带。鸭蛋通过传送带匀速送至敲击装置下，由声音收集装置采集保存敲击的音频。得到敲击裂纹蛋声样本330个，敲击完好蛋样本990个。样本的时域波形图如图2所示。

### 1.3 基于梅尔谱图的鸭蛋裂纹识别算法(DCA-MR)的建立

采用ResNet34网络构建鸭蛋裂纹识别模型，ResNet34网络具有识别速度快、准确性高等优点。敲蛋音频信号先通过端点检测、预加重、分帧加窗等




A：裂纹鸭蛋 Crack egg；B：完好鸭蛋 Intact egg.

图2 敲蛋音频时域波形图

Fig.2 Time-domain curves of knocking duck egg

操作将声音信号转换为可视化的梅尔谱图，然后借助增加了注意力机制与替换残差结构后改进的ResNet34网络模型进行特征提取，再基于预训练模型训练并对超参数进行调整，最后利用模型对数据进行分类。该算法整体结构如图3所示。

### 1.4 基于梅尔频率的音频转换

Mel频率倒谱系数(Mel frequency cepstrum coefficient，MFCC)是基于人耳听觉特性所提出的特征提取方法，利用人耳听觉与频率大小具有非线性关系这一特点，着重提取声音的低频部分的特征而忽略高频部分的影响，其提取步骤[14]如下：

(1)端点检测。所采集到的敲蛋音频信号中，由于敲击之间存在时间间隔，音频存在无效空白片段，因此需要对信号进行端点检测，确定敲击信号的起始点与终止点，减少空白段对模型检测的影响。本研究选用基于改进型双门限语音端点检测算法[15]对音频信号进行检测。

(2)预加重。预加重能够放大声音信号的高频部分。经过预加重滤波器后，噪声会受到抑制。

(3)分帧加窗。将敲蛋音频信号划分为短时间帧，并在短帧内执行特征提取。为了实现帧之间的平滑过渡，连续帧之间采用50%的重叠。然后使用汉明窗加以平滑，相比于普通矩形窗函数，通过汉明窗可以降低傅里叶变换后旁瓣大小以及减少频谱泄露[16-17]。

(4)快速傅里叶变换(fast Fourie-transform)。通过快速傅里叶变换(FFT)将时域音频信号变换为频域信号。

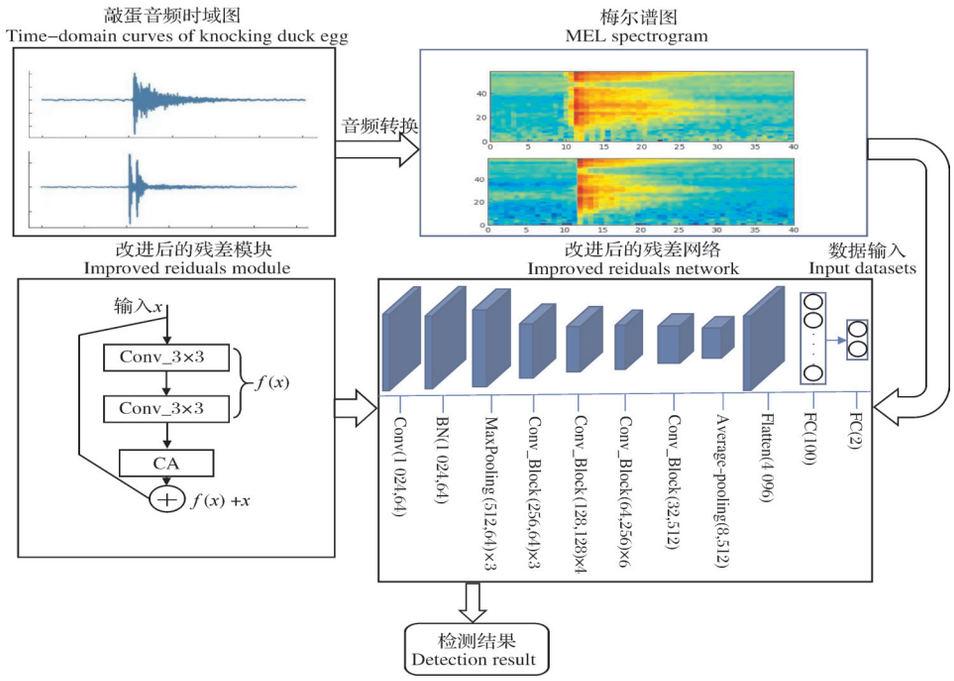



图 3　鸭蛋裂纹识别算法结构图

Fig.3　Processing procedure of the method of duck egg crack identification

(5)功率谱计算。将时域信号转化为频域进行后续的频率分析。

(6)梅尔频率滤波器。使用滤波器组对频域的幅值进行精简。功率谱 $P(k)$ 通过一组 Mel 尺度三角形滤波器获得 Mel 谱。在每个频率下，计算 $P(k)$ 和滤波器 $H_m(k)$ 的乘积，定义具有 40 个滤波器的三角形滤波器组。将滤波器应用于信号的功率谱后，最终得到梅尔谱图，如图 4 所示。

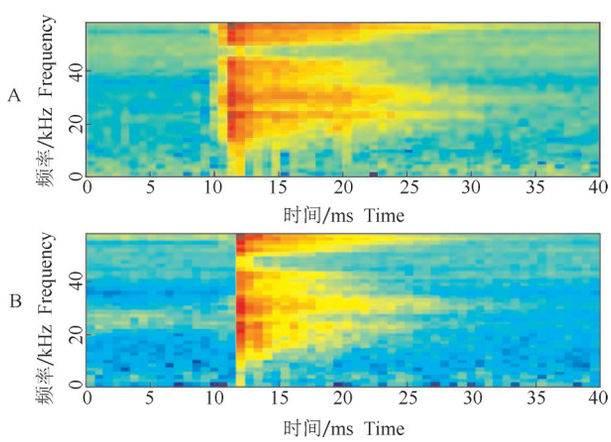



图 4　敲击完好蛋(A)和裂纹蛋(B)的声音梅尔谱图

Fig.4　Mel spectrum of sound of knocking intact egg(A) and crack egg (B)

### 1.5　ResNet 图像特征提取模型的构建与改进

1)ResNet34 模型的构建。在 ResNet 网络提出之前，一般认为随着网络层数的加深，模型效果会越好。然而简单的堆叠网络会不可避免地带来梯度爆炸和网络退化等问题，最终导致模型准确率下降，甚至结果不如浅层模型。针对此问题，He 等[10]提出 ResNet 模型，使用残差模块的结构如图 5 所示，通过恒等映射的方法，将深层梯度直接传回浅层，有效解决了网络退化问题，随着网络深度增加，模型性能可以稳步提升。

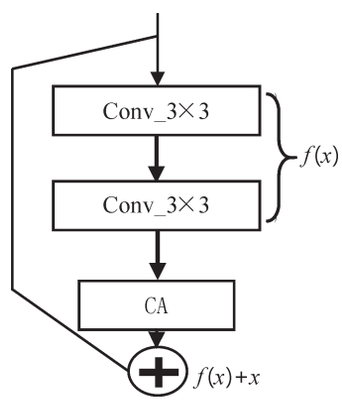



图 5　残差模块

Fig.5　Reiduals module

本研究对 ResNet34 的改进包括两方面。第一，引入注意力机制(coordinate attentions，CA)[18]。以敲蛋音频的梅尔谱图为数据集，在特征提取各位置的局部信息的过程中，梅尔谱图的冗余信息会对特征提取造成干扰，从而对模型性能造成影响，CA 能

减少冗余信息的权重，提高模型检测的准确率。第二，用深度可分离卷积替换ResNet网络残差结构中的传统卷积。先利用深层卷积对输入层的每个通道分别进行卷积操作，然后通过逐点卷积提取不同通道在相同空间位置上的特征信息。相比传统卷积，深度可分离卷积可在不影响分类精度的前提下降低网络参数量，通过迁移学习的方法提高训练效率[19]。即先通过在数据集ImageNet上训练得到初始权重，再将该权重迁移到改进ResNet34网络模型中，最后对预训练后的网络进行训练微调，从而实现网络模型参数和权重在梅尔谱图数据集上的复用[20]。

2）坐标注意力机制（CA）。首先，CA分别沿着水平方向和垂直方向对输入特征图进行全局池化操作，得到2个不同尺寸的方向感知特征图（$C\times H\times 1$，$C\times 1\times W$），每个方向感知特征图都能与输入特征沿某个方向保持依赖关系，从而得到其位置信息。然后，通过垂直方向池化操作进行特征图维度变换，将水平方向与垂直方向变换后的矩阵进行拼接，经过卷积、标准化处理后输入ReLU激活函数。最后，将矩阵在垂直方向上分离，分别进行卷积后进行升维操作，将垂直方向与水平方向的2个矩阵与恒等映射的原矩阵进行乘法操作得到输出矩阵，可以解决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。CA模块结构如图6所示。

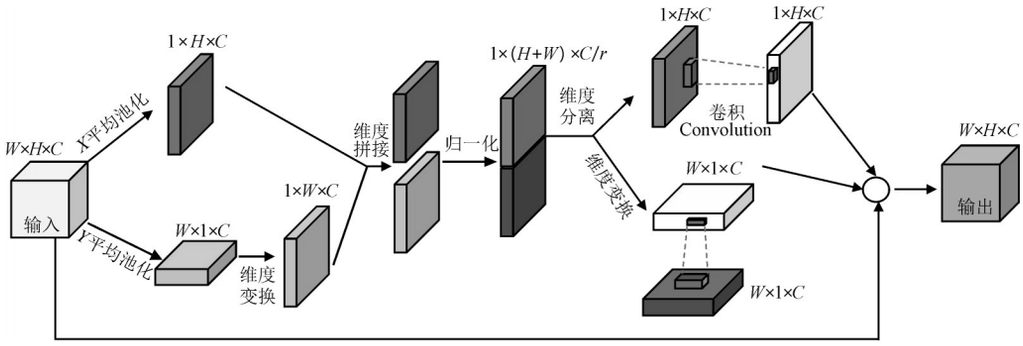



图6 CA注意力机制模块结构示意图

Fig.6 CA diagram of attention mechanism

3）深度可分离卷积。网络层数的加深带来模型参数量的显著增加，由于终端设备的计算资源有限，过多的参数将对模型的部署带来困难。通过将ResNet34模型中的传统卷积替换成深度可分离卷积，在保持准确率与检测速度平衡的时候，尽可能地减小参数量能够为后续模型的部署提供可能。深度可分离卷积由逐通道深度卷积（depthwise convolution）和逐点卷积（pointwise convolution）两阶段组成。其具体步骤为：首先对输入特征图不同的通道采用不同的卷积核进行卷积操作，得到深度卷积特征图；再将深度卷积特征图通过逐点卷积使特征在通道上叠加。因为深度卷积单独对输入层每个通道进行的卷积运算割裂了输入特征图在不同通道上同一位置的联系，因此需要通过逐点卷积叠加操作生成新的特征矩阵，从而实现特征图在深度方向的加权操作。当输入特征图尺寸为$H\times W\times C_1$（C为通道数），卷积核大小为$k\times k$，输出特征图尺寸为$H\times W\times C_2$时，传统卷积产生的参数量为$k\times k\times C_1\times C_2$，而使用图5所示的深度可分离卷积的参数量为$k\times k\times C_1+C_1\times C_2$，参数量变为原来的$1/C_2+1/k^2$。本研究将ResNet中BasicBlock中的卷积替换为深度可分离卷积得到ResNet34DP模型。

### 1.6 模型训练

鸭蛋裂纹检测模型训练的试验环境参数为：处理器Intel 11th Gen Intel（R）Core（TM）i5-11400 @ 2.60 GHz，内存16 GB，显卡NVIDIA GeForce RTX 3070，操作系统Window 10（64位），编程语言Python3.8.6。使用PyTorch深度学习框架并利用计算机统一设备架构（compute unified device architecture，CUDA）对训练进行加速。

### 1.7 网络训练

将1 320个梅尔谱图数据按8∶1∶1比例划分为训练集、验证集、测试集，通过训练集训练模型，在验证集上评估模型的性能，利用测试集检验模型稳定性。通过PyTorch深度学习框架加载ResNet34-pre.pth预训练模型。梅尔特征谱图尺寸均设置为224像素×224像素，批大小（batch size）设置为4，模型迭代次数为500，利用Adam优化算法来调整学习率。学习率决定目标函数能否收敛至局部最小值与何时收敛至最小值。设置过小，模型收敛速度较慢；设置过大则出现不收敛，无法达到最优解。

通过验证集准确率、测试集准确率、损失、模型

大小和测试时长5个评价指标判定鸭蛋裂纹识别模型的性能。

准确率表示检测正确的样本占全部样本的比值，计算公式如下：

$$\text{Accuracy}=\frac{M}{N}\times 100\% \tag{1}$$

式(1)中，$M$代表分类正确的样本个数，$N$代表全部样本个数。

损失即损失函数的收敛值，选用交叉熵损失函数作为本研究所有网络模型的Loss函数，训练集损失和验证集损失定义为：

$$\text{Loss}=-\sum_{i=1}^{n} y_i \lg(p_i) \tag{2}$$

式(2)中，Loss代表交叉熵损失的值；$n$代表分类类别的个数；$p_i$是表示预测类别为$i$的概率；$y_i$为标签，当预测类别与真实标签相同，则$y_i$=1，否则$y_i$=0。

模式大小为训练完成后的网络模型大小。测试时长为同一试验平台下测试1张图片所耗费的时间。

## 2　结果与分析

### 2.1　不同注意力机制模块性能对比

注意力机制能够增加模型特征提取中图像关键信息所占权重。为证明CA注意力机制的有效性，选取ResNet34原始模型与分别添加了SE[21]通道注意力、CBAM[22]注意力、CA注意力机制的模块进行对比试验。SE通道注意力机制模块通过激励、挤压操作计算出每个通道的权重。CBAM注意力机制模块分为通道注意力与空间注意力。注意力机制模块添加位置固定，其他参数设置不变。试验效果如表1所示。

**表1　注意力机制试验结果对比**

**Table 1　Comparison of the results of different kinds of attention**

| 模型 Model | 参数量 Parameter | 准确率/% Accuracy | 平均检测时间/ms Average detection time |
|---|---|---|---|
| ResNet34 | $2.13\times10^7$ | 87.8 | 45.1 |
| ResNet34_CA | $2.14\times10^7$ | 92.2 | 43.5 |
| ResNet34_SE | $2.20\times10^7$ | 89.8 | 40.4 |
| ResNet34_CBAM | $2.25\times10^7$ | 90.6 | 46.2 |
| ResNet34DP_CA | $1.45\times10^7$ | 92.4 | 21.5 |

由表1可知，3种增加了注意力机制模块的模型较原始ResNet34模型检测准确率分别增加了4.4、2.0、2.8个百分点。相比于原始ResNet34模型，注意力机制增加了模型对于不同信息的关注度，使模型参数量增加的同时也提高了模型性能。改进的ResNet34DP_CA模型的准确率为92.4%，比原始ResNet模型准确率增加了4.6个百分点，参数量较ResNet34_CA减少了32%。结果表明，注意力机制可以提高模型的性能和准确性。同时，参数量的减少有助于在计算资源较少的物联网设备上实现高效、稳定的部署。

### 2.2　学习率的选择

选取合适的学习率，对模型的训练具有重要作用。在网络结构与超参数相同的情况下，选取不同数量级的参数值0.001、0.000 1、0.000 01进行试验，评估模型的训练效果，得到最优学习率。如图7所示，在学习率为0.000 1时，模型能取得较好的收敛效果以及较高的准确率。

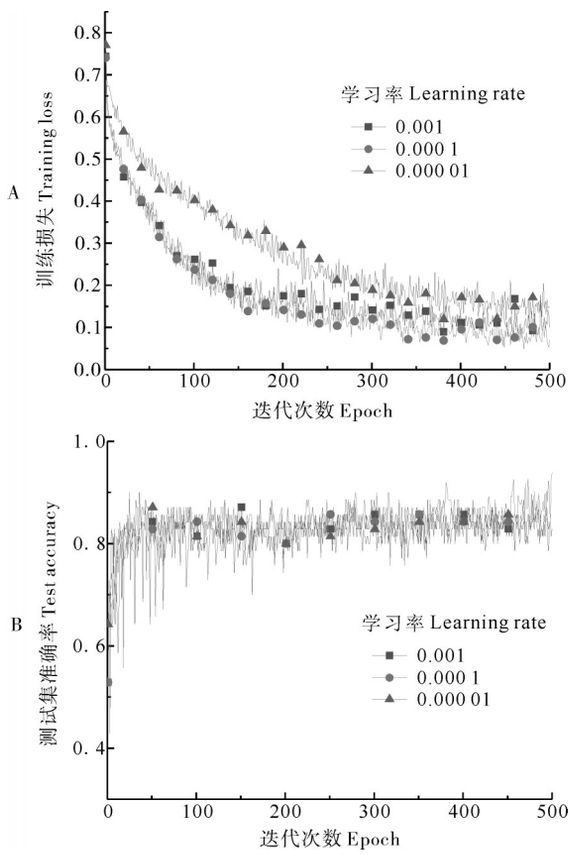



**图7　不同学习率对损失值(A)和准确率(B)的影响**

**Fig.7　Influence of different learning rate on loss value(A) and accuracy(B)**

### 2.3　不同神经网络的对比

选取ResNet34DP_CA模型，分别与VGG16、ResNet34、轻量化模型MobileNetv2、EfficientNet进行对比试验，为保证试验的公平性，网络参数均采用

相同数值，试验结果如表2所示。ResNet34DP_CA较VGG16以及原始ResNet34模型验证集准确率分别提高了10.1、4.7个百分点，测试集上准确率分别提高了11.7、6.2个百分点。相较于轻量化模型MobileNetv2、EfficientNet验证集准确率分别提高12.7、16.3个百分点，测试集上准确率分别提高14.8、16.4个百分点。在验证集与测试集的准确率对比上，ResNet34DP_CA模型准确率波动较其他模型变化小，模型稳定性好。ResNet34DP_CA模型大小相较于VGG-16、ResNet34小。MobileNetV2、EfficientNet在牺牲模型性能的同时，大幅度减小了模型的大小，加快了模型的运算速度，MobileNetV2、EfficientNet模型大小与平均检测时间均小于ResNet34DP_CA，但准确率较低，未能达到生产需求。ResNet34DP_CA模型在减小模型大小和降低平均检测时间的同时保证了检测准确率。

**表2 网络模型准确率对比**

**Table 2 Comparison of the results of different kinds of models**

| 模型<br>Model | 验证集准确率/%<br>Accuracy of validate set | 测试集准确率/%<br>Accuracy of test set | 模型大小/MB<br>Model size | 平均检测时间/ms<br>Average detection time |
|---|---|---|---|---|
| ResNet34DP_CA | 92.5 | 92.3 | 83.4 | 21.5 |
| VGG16 | 82.4 | 80.6 | 441.6 | 78.3 |
| ResNet34 | 87.8 | 86.1 | 98.7 | 44.1 |
| MobileNetv2 | 79.8 | 77.5 | 14.5 | 8.4 |
| EfficientNet | 76.2 | 75.9 | 19.6 | 9.7 |

### 2.4 不同FFT点数和窗移参数下试验对比

不同类型的谱图包含特征信息的能力不同，会对模型性能产生影响。因此，本研究针对不同FFT点数、窗移下的梅尔谱图进行了多次试验，将不同类型梅尔谱图输入模型后得到的识别效果如表3所示。

试验结果表明，在不同的FFT点数与窗移下产生的梅尔谱图用于模型取得的识别效果不同。ResNet34DP_CA与ResNet34模型在输入FFT点数1 024、窗移尺寸512时生成的梅尔谱图取得的识别效果最佳，分别为92.4%和86.9%。VGG-16与EfficientNet模型在输入FFT点数为1 024、窗移尺寸256时生成的梅尔谱图取得的识别效果最佳分别为83.5%和78.3%。MobileNetv2在输入FFT点数512、窗移尺寸256时生成的梅尔谱图取得的识别效果最佳为81.1%。

综上，本研究改进的ResNet34DP_CA模型在FFT点数1 024、窗移尺寸512时效果最佳，准确率为92.4%。

**表3 不同FFT点数和窗移下识别准确率**

**Table 3 Statistics of test results with different FFT point and window shift** %

| FFT点数<br>FFT points | 窗移尺寸<br>Window shift size | ResNet34DP_CA | VGG16 | ResNet34 | MobileNetv2 | EfficientNet |
|---|---|---|---|---|---|---|
| 2 048 | 512 | 90.8 | 82.2 | 82.1 | 80.5 | 76.8 |
| 1 024 | 512 | 92.4 | 81.5 | 86.9 | 78.7 | 76.1 |
| 1 024 | 256 | 91.6 | 83.5 | 83.6 | 80.9 | 78.3 |
| 512 | 256 | 91.1 | 82.6 | 82.7 | 81.1 | 76.2 |

## 3 讨 论

本研究提出1种基于梅尔特征谱图与ResNet34网络的鸭蛋裂纹识别算法。通过收集敲击鸭蛋蛋壳的声音音频，将音频经过端点检测、预加重等操作转化为梅尔特征谱图，利用卷积神经网络对谱图进行特征提取并分类。本算法改进的ResNet34DP_CA模型将ResNet网络中的卷积替换为深度可分离卷积，使得网络的模型大小降低16%，参数量降低32%。注意力机制能够识别不同局部信息的重要性，通过增加注意力机制，提升了模型的性能，试验结果表明，ResNet34DP_CA模型的验证集准确率达92.5%，测试集准确率92.3%，平均准确率92.4%。与VGG16、ResNet34、MobileNetv2和EfficientNet 4种卷积神经网络进行对比，发现ResNet34DP_CA模型的性能最佳，平均准确率分别提高了10.9、5.5、

13.7、16.3个百分点，在不降低检测准确率的前提下，将识别时间缩短到21.5 ms，可满足实际生产的要求。

采用敲击的方法能规避计算机视觉方法中检测不全面的问题，利用梅尔特征谱图提取方法为禽蛋声学检测提供了新的思路。该算法能够解决生产过程中人工对鸭蛋裂纹检测的准确率不稳定、效率低等问题，在保持较高准确率的同时与参数量、检测速度之间保持了平衡，为检测装备的改进与智能化工厂的建设提供了指导。

虽然本研究提出的鸭蛋裂纹识别算法有较高的准确率和检测速度，能对鸭蛋是否存在裂纹进行检测，但还不能对裂纹大小进行划分。在试验过程中，装置敲击时激励点与裂纹部位的距离会对试验结果造成影响，下一步研究中应针对上述问题对模型进行优化，减少数据采集过程中敲击点位置对检测结果的影响，并增加以裂纹大小为判断依据的精准分类功能，进一步提高模型精度与速度。

## 参考文献 References


[1] 朱玲娇，张宇，许春芳．科技创新推动我国禽蛋产业健康发展[J]．养殖与饲料，2019(1)：23-27.ZHU L J，ZHANG Y，XU C F. Scientific and technological innovation promotes the healthy development of poultry and egg industry in China[J]. Animals breeding and feed，2019(1)：23-27(in Chinese).

[2] DONG X G，ZHANG B B，DONG J，et al.Egg freshness prediction using a comprehensive analysis based on visible near infrared spectroscopy[J].Spectroscopy letters，2020，53(7)：512-522.

[3] LORENC Z，PAŚKO S，KURSA O，et al.Spectral technique for detection of changes in eggshells caused by *Mycoplasma synoviae* [J].Poultry science，2019，98(9)：3481-3487.

[4] 祝志慧，叶子凡，杨凯．种蛋孵化信息的无损检测研究进展[J]．华中农业大学学报，2021，40(4)：94-101. ZHU Z H，YE Z F，YANG K. Progress of nondestructive detection of hatching information of breeding eggs[J].Journal of Huazhong Agricultural University，2021，40(4)：94-101(in Chinese with English abstract).

[5] BAO G J，JIA M M，XUN Y，et al. Cracked egg recognition based on machine vision[J].Computers and electronics in agriculture，2019，158：159-166.

[6] GARCIA-ALEGRE M C，RIBEIRO A，GUINEA D，et al. Eggshell defects detection based on color processing[C]// Proceedings of SPIE：machine vision applications in industrial inspection Ⅷ.Columbia：Society of Photo Optical，2000：280-287.

[7] 王树才，文友先，苏工兵．禽蛋检测与分级智能机器人系统的设计[J]．机械工程学报，2008，44(2)：182-188.WANG S C，WEN Y X，SU G B. Design for system of intelligent robot detecting and grading eggs[J].Journal of mechanical engineering，2008，44(2)：182-188(in Chinese with English abstract).

[8] PRIYADUMKOL J，KITTICHAIKARN C，THAINIMIT S. Crack detection on unwashed eggs using image processing[J]. Journal of food engineering，2017，209：76-82.

[9] NASIRI A，OMID M，TAHERI-GARAVAND A.An automatic sorting system for unwashed eggs using deep learning[J/OL]. Journal of food engineering，2020，283：110036[2022-09-30]. https://doi.org/10.1016/j.jfoodeng.2020.110036.

[10] HE K M，ZHANG X Y，REN S Q，et al.Deep residual learning for image recognition[C]//2016 IEEE conference on computer vision and pattern recognition. June 27-30，2016，Las Vegas，NV，USA.New York：IEEE，2016：770-778.

[11] SIMONYAN K，ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[DB/OL]. arXiv，2014：1409.1556 [2022-09-30]. https://doi. org/10.48550/arXiv.1409.1556.

[12] SANDLER M，HOWARD A，ZHU M L，et al. MobileNetv2：inverted residuals and linear bottlenecks [C]//2018 IEEE/CVF conference on computer vision and pattern recognition. June 18-23，2018，Salt Lake City，UT，USA.New York：IEEE，2018：4510-4520.

[13] TAN M X，LE Q V.EfficientNet：rethinking model scaling for convolutional neural networks [DB/OL]. arXiv，2019：1905.11946 [2022-09-30].https://doi.org/10.48550/arXiv.1905.11946.

[14] AI O C，HARIHARAN M，YAACOB S，et al. Classification of speech dysfluencies with MFCC and LPCC features[J]. Expert systems with applications，2012，39(2)：2157-2165.

[15] 薛胜尧．基于改进型双门限语音端点检测算法的研究[J]．电子设计工程，2015，23(4)：78-81. XUE S Y. Research on speech endpoint detection based on the improved dual-threshold algorithm[J]. Electronic design engineering，2015，23(4)：78-81(in Chinese with English abstract).

[16] ASTUTI W，SEDIONO W，AIBINU A M，et al. Adaptive short time Fourier transform (STFT) analysis of seismic electric signal (SES)：a comparison of Hamming and rectangular window[C]//2012 IEEE symposium on industrial electronics and applications. September 23-26，2012，Bandung，Indonesia. New York：IEEE，2012：372-377.

[17] TRANG H，LOC T H，NAM H B H.Proposed combination of PCA and MFCC feature extraction in speech recognition system[C]//2014 International conference on advanced technologies for communications (ATC 2014). October 15-17，2014，Hanoi，Vietnam.New York：IEEE，2014：697-702.

[18] HOU Q B，ZHOU D Q，FENG J S.Coordinate attention for efficient mobile network design[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). June 20-25，2021，Nashville，TN，USA. New York：IEEE，2021：13708-13717.

[19] SHORTEN C，KHOSHGOFTAAR T M. A survey on image data augmentation for deep learning[J].Journal of big data，2019，

6(1):1-48.

[20] YOSINSKI J, CLUNE J, BENGIO Y, et al. How transferable are features in deep neural networks?[J]. Advances in neural information processing systems, 2014, 27: 3320-3328.

[21] HU J, SHEN L, SUN G. Squeeze-and-excitation networks [C]//2018 IEEE/CVF conference on computer vision and pattern recognition. June 18 - 23, 2018, Salt Lake City, UT, USA. New York: IEEE, 2018: 7132-7141.

[22] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module [C]//Computer vision: ECCV 2018. September 8-14, 2018, Munich, Germany. Cham: Springer International Publishing, 2018: 3-19.

# Identification algorithm of duck-egg shell crack based on MEL spectrum and improved ResNet34 model

KANG Junqi, XIAO Deqin, LIU Youfu, KONG Xinyue, YIN Jianjun

*College of Mathematics and Informatics, South China Agricultural University/Key Laboratory of Smart Agricultural Technology in Tropical South China, Ministry of Agriculture and Rural Affairs, Guangzhou* 510642, *China*

**Abstract** During the production, operation and processing of duck egg, the egg shells are easily broken and microorganisms including bacteria tend to invade the egg from the shell cracks, which in turn affect the quality of the eggs and damage economic benefits of production. An identification algorithm of duck-egg shell crack based on MEL spectrum was established by using ResNet34 network model to solve the problem of the subjectivity and large fluctuation of accuracy in manual identification of duck egg shell cracks. First, the egg knocker was used to collect the sound data, and the audio was transformed into the MEL spectrum graph to construct the dataset of MEL spectrum graph. Then the ResNet34 model was built and the transfer learning mechanism was introduced to train the model. The gradient was updated by Adam optimization algorithm, the attention mechanism module was added, and the convolution structure was replaced by a deeply separable convolution to improve the network model. The parameters were adjusted for optimization and the duck egg shell cracks were identified with the model. The results showed that the average detection accuracy of the ResNet34DP_CA enhanced network model was 92.4%, which was 5.5 percentage points higher than that of the original ResNet34 network model. The quantity of parameter was reduced by 32%. Compared with other network models including VGG16, MobileNetv2 and EfficientNet, the average accuracy was improved by 10.9, 13.7 and 16.3 percentage points, respectively. The recognition time was 21.5 ms. It is indicated that the established identification algorithm of duck-egg shell crack based on Mel spectrogram and the improved ResNet34 model can efficiently identify the duck-egg shell cracks. It will be of great significance to improve the economic benefits of production and to build an intelligent and modern poultry factory.

**Keywords** MEL spectrum graph; nondestructive testing; deep learning; model optimization; convolutional neural network; identification of duck-egg shell crack

(责任编辑:赵琳琳)