

基于改进 MobileNetV3 的笼养蛋鸡声音分类识别方法

衡一帆¹ 盛哲雅² 严煜¹ 谷月¹ 周昊博² 王树才¹

(1. 华中农业大学工学院, 武汉 430070; 2. 华中农业大学动物科学技术学院、动物医学院, 武汉 430070)

摘要: 为实现笼养蛋鸡声音的准确分类,实现蛋鸡健康、情绪、生产状态等信息的智能化、非接触式检测,提出了一种基于改进 MobileNetV3 的笼养蛋鸡声音分类识别方法。以欣华二号蛋鸡为研究对象,采集蛋鸡在笼养条件下发出的热应激声、惊吓声、产蛋声以及鸣唱声,经过声音预处理将一维声音信号转化为三维梅尔频谱图,建立了包括 8 541 幅梅尔频谱图的蛋鸡声音数据集。通过在 MobileNetV3 中引入高效通道注意力 (Efficient channel attention, ECA) 模块,提高了笼养蛋鸡声音分类准确率。试验结果表明,MobileNetV3-ECA 模型准确率、召回率、精确率以及 F1 分数分别达到 95.25%、95.16%、95.02%、95.08%,相比原始模型分别提高 1.99、2.08、2.00、2.04 个百分点。通过与分别引入坐标注意力 (Coordinate attention, CA)、卷积块注意力模块 (Convolutional block attention module, CBAM) 的模型对比,引入 ECA 模块后模型准确率分别提高 2.11、2.03 个百分点,其他指标同样有更明显的提高。与 ShuffleNetV2、DesNet121 和 EfficientNetV2 模型相比,MobileNetV3-ECA 准确率分别提高 1.99、2.03、2.50 个百分点。本文提出的基于 MobileNetV3-ECA 的蛋鸡声音分类识别方法,能够有效且准确地实现对包括热应激声在内的不同种类蛋鸡声音分类识别,为蛋鸡规模化养殖中的自动化、智能化声音检测提供了算法支持,为禽舍巡检机器人功能优化提供了参考,同时为规模化笼养蛋鸡热应激预警开辟了思路。

关键词: 笼养蛋鸡; 声音分类; MobileNetV3; 高效通道注意力; 梅尔频谱图; 卷积神经网络

中图分类号: S24; S831.4

文献标识码: A

文章编号: 1000-1298(2025)04-0427-09

OSID:



Method for Sound Classification and Recognition for Caged Laying Hens Based on Improved MobileNetV3

HENG Yifan¹ SHENG Zheyu² YAN Yu¹ GU Yue¹ ZHOU Haobo² WANG Shucai¹

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

2. College of Animal Science and Technology and College of Veterinary Medicine, Huazhong Agricultural University, Wuhan 430070, China)

Abstract: In order to achieve accurate classification of caged laying hens' sounds and intelligent, non-contact detection of laying hens' health, emotion, production status and other information, a caged laying hens' sound classification and recognition method based on improved MobileNetV3 was proposed. The heat stress sound, fright sound, egg-laying sound and singing sound produced by laying hens under cage conditions were collected from Xinhua No. 2 laying hens as research object, the one-dimensional sound signals were transformed into three-dimensional Mel-spectrograms after sound pre-processing, and the laying hens' sound data set consisting of 8 541 Mel-spectrograms was established. The accuracy of sound classification for caged laying hens was improved by introducing the efficient channel attention (ECA) module in MobileNetV3. The experimental results showed that the MobileNetV3-ECA model achieved 95.25%, 95.16%, 95.02% and 95.08% of accuracy, recall, precision and F1 score, representing an enhancement of 1.99, 2.08, 2.00 and 2.04 percentage points, respectively, in comparison with the original model. Comparing the models with the introduction of coordinate attention (CA) and convolutional block attention module (CBAM) respectively, the accuracy of the model was improved by 2.11 and 2.03 percentage points with the introduction of the ECA module. Significant improvements were also seen in other metrics. The accuracy of MobileNetV3-ECA was improved by 1.99, 2.03 and 2.50

收稿日期: 2024-03-04 修回日期: 2024-04-08

基金项目: 农业生物育种国家科技重大专项 (2023ZD0407106)

作者简介: 衡一帆 (1996—), 女, 博士生, 主要从事深度学习和农业智能检测技术研究, E-mail: a1715622858@163.com

通信作者: 王树才 (1966—), 男, 教授, 博士生导师, 主要从事农产品智能化检测技术与装备研究, E-mail: wsc01@mail.hzau.edu.cn

percentage points compared with that of ShuffleNetV2, DesNet121 and EfficientNetV2. The MobileNetV3 – ECA-based sound classification and recognition method for laying hens proposed provided algorithmic support for automated and intelligent sound detection in the large-scale breeding of laying hens, and also provided a reference for the function optimization of poultry house inspection robots, and opened up a way of thinking for heat stress early warning of large-scale caged laying hens.

Key words: caged laying hens; sound classification; MobileNetV3; efficient channel attention; Mel-spectrogram; convolutional neural network

0 引言

我国是世界上最大的鸡蛋消费国和主要出口国之一^[1],蛋鸡规模化养殖与鸡蛋产量直接相关,我国多以笼养模式为主。随着近些年高新技术的推广应用,一些大型养殖场已引入巡检机器人,如悬轨式机器人 ChickenBoy^[2]、自主导航机器人 PoultryBot^[3]等,但由于各自应用场景不同,功能并不完善,所以仍无法摆脱人工巡检。人工巡检工作强度大、主观性强,极易造成鸡只漏检和误检^[4]。而动物发声可反映其个体健康状况以及需求,能在一定程度上反映其生理生长信息,可对动物机体状况以及动物对环境的应激情况进行分析,作为评价动物舒适度的辅助方法^[5-8]。通过声音监测实现对动物行为监测、疾病预警、繁殖期调控、饲养环境监测等,已成为近年来养殖领域研究热点之一^[9]。

随着机器学习技术的发展,学者将家禽养殖与机器学习相结合。利用鸡只声音的时域、频域等特征实现了鸡只的禽流感发声检测^[10-11]、新城疫发声检测^[12]等,但传统机器学习方法在针对单一时域或频域特征进行分类时,其识别效果并不稳定,面对不同输入特征时鲁棒性较差。深度学习在图像、语音识别和分类方面取得的效果远超先前相关技术。在家禽养殖领域,学者利用目标检测算法实现了对家禽的行为识别^[13]、疾病诊断^[14]、体温检测^[15]等。针对家禽发声,相关学者结合图像处理技术和 Gabor 滤波器提取蛋鸡声谱图中的声纹信息,并采用 BP 神经网络模型对二维声谱图像进行训练、分类,其平均精度和灵敏度均能达到 92% 以上,并通过开发 SVM 分类器,使用音频技术实现了对栖架散养条件下蛋鸡的热舒适度评估^[16-17]。余礼根等^[18]以蛋鸡的产蛋声、饲喂声、尖叫声等为研究对象,提取声音的融合特征参量,构建了基于遗传算法优化 BP 神经网络的蛋鸡发声分类识别模型,识别精确度、灵敏度均达到 90% 以上。一些学者根据鸡舍内环境噪声、行为噪声以及鸡声的频谱差异,使用卷积神经网络识别患禽流感病鸡的声音;同时提出了使用家禽发声深度网络来识别感染新城疫后的家禽声音,用

于新城疫的早期检测;通过提取雏鸡声音,将音频特征输入到深度学习模型中进行训练和分类,并根据结果预测雏鸡性别^[19-21]。

以往的研究为蛋鸡的声音分类打下了坚实基础,实现了对蛋鸡多种声音的分类识别,但在规模化笼养过程中,往往由于体质量、羽毛密度、健康水平等鸡只个体差异造成鸡只的热中性区有所不同^[22],个别鸡只的耐热能力下降,热应激反应加剧,产蛋率和蛋壳质量也会随之下降,严重时甚至会造成蛋鸡死亡^[23-24]。因此对蛋鸡声音分类识别的准确率仍有待提高,而且国内针对蛋鸡异常发声的研究大多集中在其疾病发声的识别,对层叠式笼养下的蛋鸡热应激发声的研究仍有不足。

轻量化模型 MobileNetV3 由于参数量少、计算量小、推理时间短,更适用于存储空间和功耗受限的场景,更加适配后续禽舍巡检机器人在蛋鸡声音在线检测领域中进行的移动端嵌入式开发。为进一步提高笼养蛋鸡声音分类的准确率,并为层叠式笼养下蛋鸡热应激声音的准确高效识别提供参考,本文提出一种改进 MoblieNetV3 作为分类模型,将蛋鸡在高温下的热应激声音以及在正常环境温度下的声音转换为梅尔频谱图进行分类,以期笼养蛋鸡智能无人化养殖提供技术参考,同时也为后续蛋鸡热应激行为在线检测提供基础。

1 材料与方法

1.1 声音信号采集

声音信号采集于湖北省武汉市华中农业大学试验鸡场,研究对象为 310 d 左右的欣华二号蛋鸡,饲养模式为层叠式笼养,如图 1a 所示。笼舍内配有饮水线、饲料槽、风机等,数据采集时间为 09:00—15:00,采集过程中,蛋鸡鸣唱声、产蛋声、惊吓声均在环境温度为 24 ~ 27℃ 时采集,此温度范围为正常饲养期间的鸡舍温度,通过风机和湿帘调节;热应激声音为环境温度 36 ~ 37℃ 时采集,在该温度下鸡只热应激程度较重,处于无法维持生理功能的临界状态^[25-26],由笼舍内安装的智能温度时间控制器(图 1b)来实时调整笼舍内温度。

声音采集设备选用如图 1c 所示的深圳华声迪科

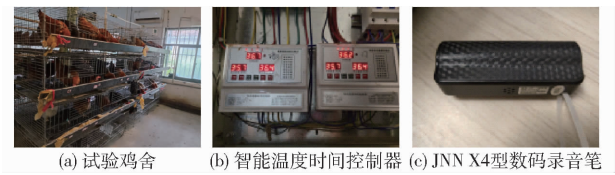


图 1 试验环境及声音采集设备

Fig. 1 Test environment and sound acquisition equipment

技有限公司的 JNN X4 型数码录音笔(16 GB,采样频率 48 kHz,位深度 16 bit),在笼舍内选取 3 个位置固定设备进行采样,每个采样点间距 1.5~2.0 m,声音数据以 WAV 格式储存,通过 USB 连接传输到计算机内进行处理。

1.2 声音降噪及切分

由于采集到的声音仍含有大量噪声,为了降低

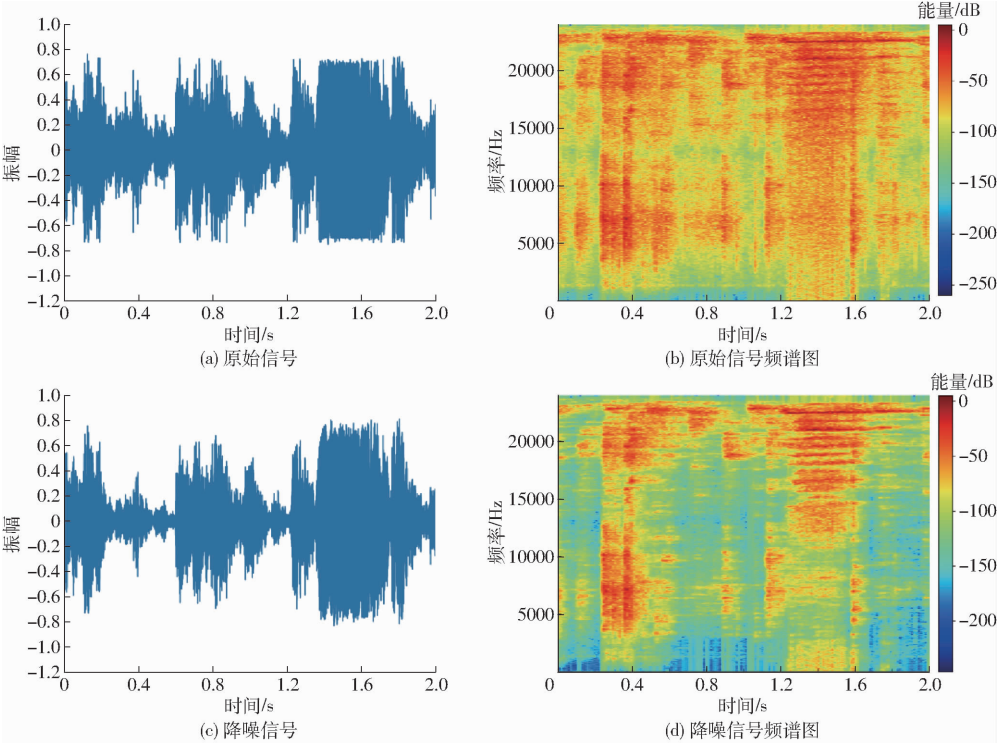


图 2 蛋鸡声音降噪前后效果对比

Fig. 2 Effects of noise reduction on laying hens before and after sound reduction

通过人工方式将降噪后的声音数据在 Adobe Audition 2022 中进行切分,每个声音样本持续时间为 2 s,最终有效声音数据共 8 541 个片段,具体声音类型、样本数量及声音内容如表 2 所示。

表 2 样本声音类型及描述

Tab. 2 Sample sound types and definitions		
声音类型	样本数量/个	声音描述
热应激声	2 151	蛋鸡在热应激试验中发出的急促鸣叫声
惊吓声	1 733	蛋鸡受人类惊吓时发出的尖叫声
产蛋声	2 334	蛋鸡产蛋后发出的咯咯及咯哒声
鸣唱声	2 323	蛋鸡在无外界刺激并自由进食、饮水下发出的声音

噪声对后续试验结果的影响,通过 Adobe Audition 2022 对其进行自适应降噪,降噪参数调整见表 1,降噪后效果如图 2 所示。

表 1 降噪参数

Tab. 1 Noise reduction parameters	
参数	数值
降噪幅度/dB	20
噪声量/%	70
微调噪声基准/dB	2
信号阈值/dB	3.2
频谱衰减率/(ms·60 dB ⁻¹)	200
宽频保留/Hz	300
FFT 大小	1 024

1.3 声音信号特征提取

声音信号作为一维信号,只能反映信号时频特性。在进行声音识别分类时,为了使模型实现更好的分类效果,需要提供更高质量、包含更多信息的输入。梅尔频谱图(Mel-spectrogram)作为一种常用的声音特征提取方式,展示了信号的能量在时间和梅尔尺度频率上的分布,而梅尔尺度是一种基于人类听觉感知的非线性频率尺度,提供了一种比较直观的频率和时间的表示。因此本文将采集到的一维声音信号转换为含有时间、频率和能量三维表达的梅尔频谱图作为声音识别模型的输入,可以对声音实现更精确的分类结果。为了进一步降低音频噪声在声音提取过程中的影响,同时降低采集设备所带来

的音频混叠、失真等,需要对采集到的音频进行预处理,主要包括预加重、分帧、加窗等。特征提取过程如图 3 所示。

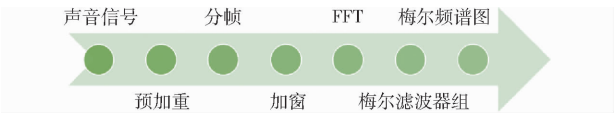


图 3 MFCC 特征提取过程
Fig.3 MFCC feature extraction process

为了加强音频信号中的高频信息,采用数字滤波器实现声音信号预加重,预加重输入和输出关系为

$$y(n) = x(n) - \alpha x(n - 1) \tag{1}$$

式中 $y(n)$ ——输出信号
 α ——预加重系数,取 0.97
 $x(n)$ ——声音信号第 n 个采样点
 $x(n - 1)$ ——声音信号第 $n - 1$ 个采样点

图 4 为预加重前后音频信号频谱对比。

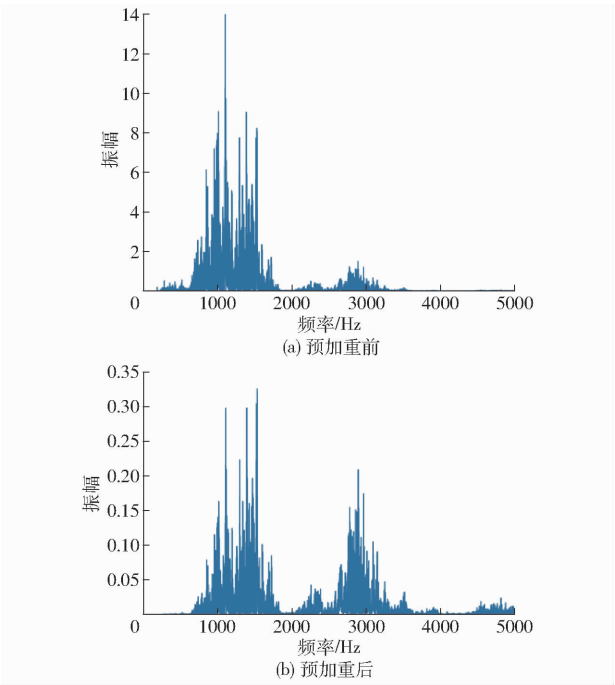


图 4 预加重前后的音频信号频谱对比
Fig.4 Comparison of audio signal spectrum before and after pre-emphasis

由于在极短时间内声音信号的频率信息不变,对声音信号进行分帧处理,帧长选择为 25 ms,帧移为 10 ms,分帧结束后,对每一帧添加窗函数,选择汉明窗函数实现信号加窗。汉明窗定义公式为

$$H(n) = \begin{cases} 1 - \beta - \beta \cos(2\pi n/N) & (0 \leq n \leq N) \\ 0 & (n > N) \end{cases} \tag{2}$$

式中 N ——窗口长度
 β ——窗口形状调节参数
根据 β 取值获得不同的汉明窗,此处 β 取

0.46。

对加窗后的每一帧信号进行傅里叶变换,得到信号的功率图谱,最后构建梅尔滤波器组,将功率图谱转化为梅尔频谱图,其映射关系为

$$m = 2\,595 \lg \left(1 + \frac{f}{700} \right) \tag{3}$$

梅尔滤波器 $H_m(k)$ 计算公式为

$$H_m(k) = \begin{cases} 0 & (k < f(m - 1)) \\ \frac{k - f(m - 1)}{f(m) - f(m - 1)} & (f(m - 1) \leq k \leq f(m)) \\ \frac{f(m + 1) - k}{f(m + 1) - f(m)} & (f(m) < k \leq f(m + 1)) \\ 0 & (k > f(m + 1)) \end{cases} \tag{4}$$

式中 m ——滤波器序号
 f ——频率 k ——频点索引
 $f(m - 1)$ ——第 m 个滤波器起始点
 $f(m)$ ——第 m 个滤波器中间点
 $f(m + 1)$ ——第 m 个滤波器结束点

1.4 蛋鸡声音数据集构建

随着人工智能的快速发展,深度学习也逐渐应用到声音识别与分类中。在蛋鸡声音分类识别中,将含有更多发声信息的梅尔频谱图作为分类网络输入,可以实现更好的分类效果。经过预处理获得的 8 541 幅蛋鸡声音梅尔频谱图,按照比例 7:3 分为训练集和验证集,其中训练集共 5 974 幅梅尔图像,4 种声音在数据集中的具体分布情况见表 3,图 5 为 4 种声音信号与对应的梅尔频谱图。

表 3 蛋鸡发声梅尔频谱图数据集样本数量
Tab.3 Distribution of Mel-spectrogram dataset for laying hens

标签	声音类型	训练集	验证集	总计
A	热应激声	1 505	646	2 151
B	惊吓声	1 213	520	1 733
C	产蛋声	1 630	704	2 334
D	鸣唱声	1 626	697	2 323

2 蛋鸡声音识别分类模型构建

2.1 分类网络模型构建

MobileNetV3^[27] 在 MobileNetV2 的基础上进行改进,在继承 MobileNetV1 的深度可分离卷积和 MobileNetV2 的倒残差结构的同时引入了压缩和激励网络(Squeeze-and-excitation networks, SE),使用 Hard-swish 激活函数代替了原有的 ReLU6 激活函数,并重新设计了网络耗时层结构,在模型预测准确率和计算延时方面均有所改进。经过声音预处理后

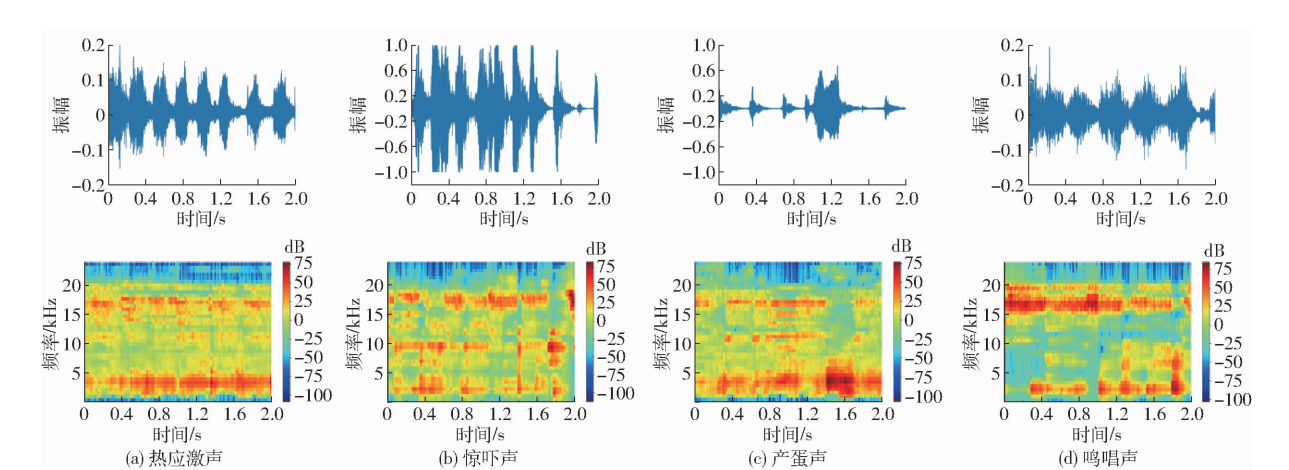


图5 4种蛋鸡声音信号及对应梅尔频谱图

Fig. 5 Sound signals of four types of laying hens and corresponding Mel-spectrograms

得到的梅尔频谱图,仍存在一定程度的噪声干扰,采用 MobileNetV3 提取准确有效的特征难度依然较大。为提高模型识别精度,对 MobileNetV3 网络进行改进,在倒残差结构中引入高效通道注意力 (Efficient channel attention, ECA) [28]。

ECA 注意力机制利用自适应的全局平均池化和一维卷积操作,对每个通道的特征进行自适应的加权,通过学习通道之间的相关性,提取更具有代表性的特征,在避免降维的同时高效实现了局部跨通道交互。通过在原模型的基础上引入 ECA 注意力机制,弥补了由于 SE 注意力机制对输入特征图进行降维时造成的通道之间依赖关系的损失。改进后的整体网络结构如图 6 所示。

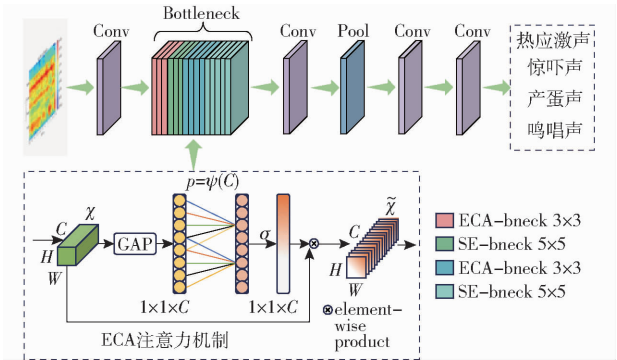


图6 MoblieNetV3 - ECA 整体网络结构图

Fig. 6 Overall network structure of MoblieNetV3 - ECA

ECA 注意力机制结构中,在对输入特征图进行全局平均池化 (Global average pool, GAP) 后直接进行一维卷积, p 为自适应一维卷积内核大小,即跨通道信息交互作用的覆盖范围,通过内核大小为 p 的一维卷积来实现通道之间的信息交互, C 为通道维数,二者之间的非线性关系为

$$p = \psi(C) = \left\lfloor \frac{\ln C}{\gamma} + \frac{b}{\gamma} \right\rfloor_{\text{odd}} \tag{5}$$

其中 $\gamma = 2, b = 1$, 卷积后的权重 ω 满足

$$\omega = \sigma(C_{1D}(y)) \tag{6}$$

式中 σ ——Sigmoid 激活函数

C_{1D} ——一维卷积 y ——通道

2.2 模型试验环境及参数设置

为了测试改进 MobileNetV3 蛋鸡声音识别模型效果,用得到的 8 541 幅梅尔频谱图输入模型进行训练和验证。软件环境设置为 Python 3.8,深度学习框架为 Pytorch 1.7.1,程序运行硬件配置为 Intel (R) Core (TM) i5 - 9300HF CPU @ 2.40 GHz, NVIDIA GeForce GTX 1650 显卡,运行内存 16 GB, 64 位 Windows 11 操作系统。

训练过程中, Batch size 统一设置为 8, 迭代次数设置为 250, 每 10 次保存一次权值, 采用随机梯度下降法训练模型, 学习率、动量、权重衰减分别为 0.01、0.9、0.000 5。

2.3 评价指标

为了全面评价 MobileNetV3 - ECA 模型性能及有效性,采用准确率 (f_{Acc})、精确率 (f_{Pre})、召回率 (f_{Rec})、F1 分数 (F_{1score}) 作为综合评估指标。梅尔频谱图分类识别时,当前识别图像的实际类别数为正,其余类别总数均为负。

3 结果与分析

3.1 改进模型性能检测

为验证改进 MobileNetV3 - ECA 模型性能,分别对 4 种蛋鸡声音的梅尔频谱图进行了分类识别。模型在训练集和验证集上的损失值变化如图 7a 所示,训练达到 30 次后,训练集和验证集的损失值下降速度逐渐减慢;训练达到 180 次后,训练集和验证集的损失值趋于平缓,最终收敛值及变化趋势在训练集和验证集上均保持了一致,说明 MobileNetV3 - ECA 模型整体收敛趋势较好,且在迭代次数内未发生过拟合现象。模型在训练集上各指标变化如图 7b 所

示,从整体上看,模型在训练集上达到 180 次后,其精确率、准确率、召回率均达到 90% 以上,F1 分数也超过 90%,4 项指标变化趋势一致且逐渐趋于稳定,初步验证了 MobileNetV3 - ECA 模型在蛋鸡声音分类中的有效性。

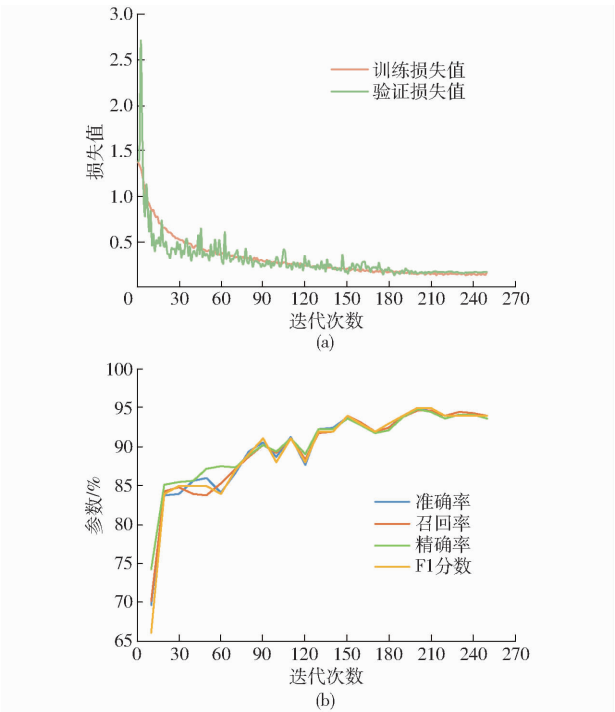


图 7 MobileNetV3 - ECA 在训练集和验证集上的性能曲线

Fig. 7 Performance of MobileNetV3 - ECA on training and validation sets

为了进一步评价改进模型性能,对模型在验证集上各个类别精确率和召回率以及 F1 分数进行了统计。分类结果如表 4 所示,在单个类别识别中,MobileNetV3 - ECA 模型对蛋鸡热应激声识别效果最好,模型精确率、召回率、F1 分数分别为 100%、99.69%、99.84%,与之相比,蛋鸡惊吓声识别精确率和 F1 分数相对较低,但也均不低于 89.74%,其他两种声音识别精确率、召回率以及 F1 分数也不低于 92.76%。总体来看,引入 ECA 注意力机制后,它

通过在通道维度上对特征进行自适应加权,使网络能够更好地聚焦于输入特征图中更重要的特征,抑制梅尔频谱图中噪声对网络的干扰。自适应的全局平均池化层和卷积层使模型能够自动地选择和放大重要的频谱特征,使得模型在验证集上的识别精确率、召回率和 F1 分数平均值均不低于 95.02%,识别效果较优,同时进一步证明了改进 MobileNetV3 模型能够有效且准确地实现对包括热应激声在内的不同种类蛋鸡声音分类识别。

表 4 测试集蛋鸡声音分类结果

hens in test set				%
声音类别	精确率	召回率	F1 分数	
热应激声	100	99.69	99.84	
惊吓声	89.74	92.50	91.10	
产蛋声	94.78	92.76	93.76	
鸣唱声	95.56	95.70	95.63	
平均值	95.02	95.16	95.08	

3.2 不同注意力机制对模型的影响

为了测试 ECA 注意力机制在蛋鸡声音分类中的有效性,选取坐标注意力^[29] (Coordinate attention, CA)、卷积块注意力^[30] (Convolutional block attention module, CBAM) 两种常见模块进行对比试验。将 MobileNetV3 - ECA 中的 ECA 模块分别替换为 CA 模块和 CBAM 模块,在相同数据集、模型结构、试验平台等条件下训练 250 次,训练结束后绘制 3 种注意力机制分类混淆矩阵,如图 8 所示。试验结果如表 5 所示。

由表 5 可知,ECA 模块的自适应特性使模型在进行复杂声音分类时达到了更高精度,在精确率、召回率、准确率和 F1 分数中均达到 3 个模块中的最高。虽然 CA 模块和 CBAM 模块可以提取到输入特征中的重要位置信息,但面对差异较小且较为复杂的声音特征图,更大的位置信息权重反而会影响模型的分类效果。ECA 模块通过全局平均池化和一

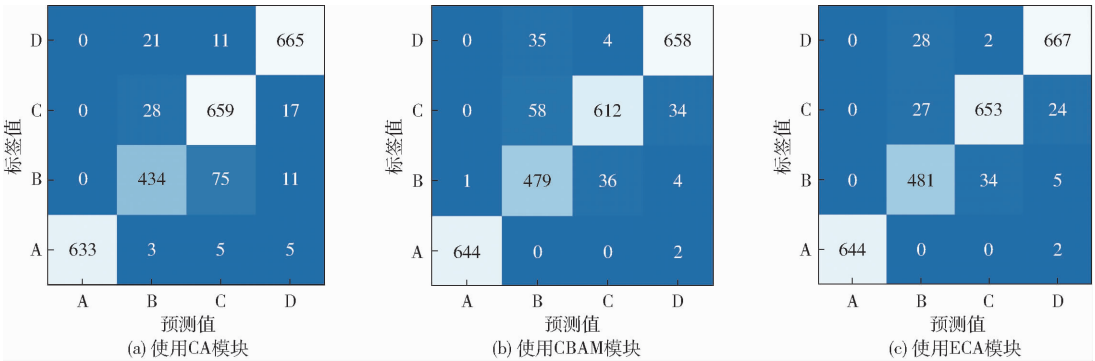


图 8 MobileNetV3 插入不同注意力机制后的分类混淆矩阵

Fig. 8 Classification confusion matrix for MobileNetV3 after inserting different attention mechanisms

表 5 不同注意力机制对比试验结果

Tab.5 Comparison experiments of different attention mechanisms %

模块	准确率	召回率	精确率	F1 分数
CA	93.14	92.62	93.11	92.81
CBAM	93.22	93.29	92.93	93.02
ECA	95.25	95.16	95.02	95.08

维卷积,在初始 SE 模型基础上加强了梅尔频谱图每个通道之间的相关性,提升了通道之间的依赖关系,可以获取更具有代表性、更详细的通道信息特征。因此面对蛋鸡声音的梅尔频谱图,ECA 模块分类效果更好。

图 8 混淆矩阵中坐标轴英文字母分别对应表 3 中不同声音的样本类别,对角线方向为训练集中分类正确的样本数量。由图 8 可知,使用 ECA 注意力机制在惊吓声和鸣唱声中分类效果最好;热应激声识别结果表明,使用 ECA 注意力机制和 CBAM 注意力机制时相同,CA 注意力机制分类效果最差;在产蛋声的识别中,使用 CA 注意力机制分类效果略优于使用 ECA 注意力机制时的效果,使用 CBAM 注意力机制则出现大量误检,其误检数量分别超过使用 CA 注意力机制和 ECA 注意力机制产生误检数的 51%、45%。

分析原因,CA 注意力机制增加了模型对输入的位置信息依赖,过度地关注位置信息,进而忽略了其他重要特征,同时由于 CA 注意力机制中额外增加的卷积层会造成训练数据中的位置信息过拟合,从而降低了模型的泛化能力,导致引入 CA 注意力机制后模型在热应激声识别中效果不佳;CBAM 注意力机制则主要关注输入在空间维度上的特征,进而极容易忽略通道维度上的特征,同时额外增加的卷积层和池化层在增加了参数量的同时,也降低了模型在数据集中的泛化能力,因此引入 CBAM 注意力机制后模型识别产蛋声时反而产生了大量误检;ECA 注意力机制通过引入一个可学习的一维卷积操作,相比于 CA 注意力机制和 CBAM 注意力机制,计算效率更高且参数量更少,降低了数据过拟合的风险,这使得 ECA 模块在保持较低计算复杂度的同时,能够有效地利用通道之间的关系进行注意力加权。通过自适应地选择和加权通道的重要性,从而提升了模型泛化能力,更好地适应不同种类蛋鸡声音分类。综合来看,引入 ECA 注意力机制后的 MobileNetV3 模型在蛋鸡声音分类识别中效果最好,模型泛化能力和准确性相较于其他两种方法也有所提高。

3.3 不同模型对比试验

为了验证本文模型有效性,选取 ShuffleNetV2^[31]、DesNet121^[32]、EfficientNetV2^[33]、MobileNetV3 4 个

在目标分类识别领域比较出色和具有代表性的网络,在蛋鸡声音识别分类方法选择中具有一定的对比性与参考价值。在相同训练参数、数据集、参数设置、试验平台等条件下,与改进 MobileNetV3 – ECA 模型进行对比试验,每个模型训练 250 次。

5 个模型准确率、召回率、精确率和 F1 分数如表 6 所示。结果表明,MobileNetV3 – ECA 模型在验证集中的表现均优于其他网络。与同为轻量级网络的 ShuffleNetV2 相比,MobileNetV3 – ECA 模型准确率、召回率、精确率和 F1 分数分别提高 1.99、2.08、2.00、2.04 个百分点,与 DesNet121、EfficientNetV2 和 MobileNetV3 模型相比,MobileNetV3 – ECA 模型的准确率提高 2.03、2.50、2.03 个百分点;召回率提高 2.34、2.84、1.92 个百分点;精确率提高 1.89、2.41、2.08 个百分点;F1 分数提高 2.07、2.63、2.05 个百分点。综上所述,本文模型不仅在同类型中有着较优的识别效果,与包括原模型在内的其他分类模型相比仍具有较好表现。

表 6 模型对比试验结果

Tab.6 Model comparison test results %

模型	准确率	召回率	精确率	F1 分数
ShuffleNetV2	93.26	93.08	92.02	93.04
DesNet121	93.22	92.82	93.13	92.95
EfficientNetV2	92.75	92.32	92.61	92.45
MobileNetV3	93.22	93.24	92.94	93.03
MobileNetV3 – ECA	95.25	95.16	95.02	95.08

4 结论

(1)提出了一种基于改进 MobileNetV3 的笼养蛋鸡声音识别分类方法,即 MobileNetV3 – ECA。建立了包括蛋鸡热应激声音在内的蛋鸡声音数据库,将蛋鸡声音转化为梅尔频谱图作为模型输入,实现了笼养蛋鸡声音的准确分类识别。

(2)引入 ECA 模块后的 MobileNetV3 模型在对蛋鸡声音种类进行识别时,各项模型评估指标均有所提高,准确率、召回率、精确率以及 F1 分数分别达到 95.25%、95.16%、95.02% 和 95.08%。

(3)不同注意力机制对比试验表明,ECA 注意力机制在蛋鸡声音识别分类中效果更好,提高了原始模型泛化能力和准确性。与选取的同类轻量级卷积神经网络和卷积神经网络对比,本文模型在蛋鸡声音识别分类中各项指标表现更好。

(4)MobileNetV3 – ECA 识别精度高、准确度高、泛化能力强,在对包括热应激声在内的不同种类的蛋鸡声音分类识别中具有较好的识别效果,为蛋鸡规模化养殖中的自动化、智能化声音检测提供了参考,同时为规模化笼养蛋鸡热应激预警提供了思路。

参 考 文 献

[1] 成祖斌,刘正青,于传军,等. 浅析笼养蛋鸡健康养殖技术研究现状及发展趋势[J]. 中国动物保健, 2020, 22(11): 44-45.
CHENG Zubin, LIU Zhengqing, YU Chuanjun, et al. Analysis of the current status and development trends of healthy breeding techniques for caged egg hens[J]. China Animal Health, 2020, 22(11): 44-45. (in Chinese)

[2] ChickenBoy. Higher productivity and increased welfare with an autonomous robot and artificial intelligence[R/OL]. (2021-01-21) [2024-01-20]. <https://cdn.bigdutchman.com/>.

[3] RICHARDSON R, DALCIMAR C, MARCELO T, et al. Generating action plans for poultry management using artificial neural networks[J]. Computers and Electronics in Agriculture, 2019, 161: 131-140.

[4] 王珑翰,张良,高建波,等. 蛋鸡健康监测和行为识别及死鸡移除系统研究进展[J]. 中国家禽, 2022, 44(10): 83-88.
WANG Longhan, ZHANG Liang, GAO Jianbo, et al. Research progress on health monitoring, behavior recognition and dead chicken removal system for layer production[J]. China Poultry, 2022, 44(10): 83-88. (in Chinese)

[5] 杜晓冬,滕光辉. 蛋鸡声音信号去噪方法对比分析[J]. 农业机械学报, 2017, 48(12): 327-333.
DU Xiaodong, TENG Guanghui. Improved de-noising method of laying hens' vocalization[J]. Transactions of the Chinese Society for Agricultural Machinery, 2017, 48(12): 327-333. (in Chinese)

[6] VANDERMEULEN J, BAHR C, TULLO E, et al. Discerning pig screams in production environments[J]. PLoS One, 2015, 10(4): e0123111.

[7] CHELOTTI J O, VANRELL S R, MILONE D H, et al. A real-time algorithm for acoustic monitoring of ingestive behavior of grazing cattle[J]. Computers and Electronics in Agriculture, 2016, 127: 64-75.

[8] LI N, REN Z, LI D, et al. Automated techniques for monitoring the behaviour and welfare of broilers and laying hens: towards the goal of precision livestock farming[J]. Animal, 2020, 14(3): 617-625.

[9] 赵晓洋. 基于动物发声分析的畜禽舍环境评估[D]. 杭州: 浙江大学, 2019.
ZHAO Xiaoyang. Environment assessment of livestock house based on vocal analysis animals[D]. Hangzhou: Zhejiang University, 2019. (in Chinese)

[10] HUANG J, WANG W, ZHANG T. Method for detecting avian influenza disease of chickens based on sound analysis[J]. Biosystems Engineering, 2019, 180: 16-24.

[11] 张铁民,黄俊端. 基于音频特征和模糊神经网络的禽流感病鸡检测[J]. 农业工程学报, 2019, 35(2): 168-174.
ZHANG Tiemin, HUANG Junduan. Detection of chicken infected with avian influenza based on audio features and fuzzy neural network[J]. Transactions of the CSAE, 2019, 35(2): 168-174. (in Chinese)

[12] BANAKAR A, SADEGHI M, SHUSHTARI A. An intelligent device for diagnosing avian diseases: newcastle, infectious bronchitis, avian influenza[J]. Computers and Electronics in Agriculture, 2016, 127: 744-753.

[13] 夏元天,寇旭鹏,薛洪成,等. 基于改进 YOLO-MAO 检测框架的笼养白羽肉鸡行为检测方法[J]. 农业机械学报, 2024, 55(11): 103-111.
XIA Yuantian, KOU Xupeng, XUE Hongcheng, et al. Behavior detection algorithm for caged white-feather broilers based on improved YOLO detection framework[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(11): 103-111. (in Chinese)

[14] 陈佳,丁奇安,刘龙申,等. 基于 YOLO v5 与短时跟踪的鸡只呼吸道疾病早期检测[J]. 农业机械学报, 2023, 54(1): 271-279.
CHEN Jia, DING Qi'an, LIU Longshen, et al. Early detection of broilers respiratory diseases based on YOLO v5 and short time tracking[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(1): 271-279. (in Chinese)

[15] 严煜,盛哲雅,谷月,等. 非接触式笼养蛋鸡核心体温检测方法[J]. 农业机械学报, 2024, 55(8): 312-321.
YAN Yu, SHENG Zheya, GU Yue, et al. Non-contact core body temperature detection method for caged laying hens[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(8): 312-321. (in Chinese)

[16] 杜晓冬,滕光辉, TOMAS Norton, 等. 基于声谱图纹理特征的蛋鸡发声分类识别[J]. 农业机械学报, 2019, 50(9): 215-220.
DU Xiaodong, TENG Guanghui, TOMAS Norton, et al. Classification and recognition of laying hens' vocalization based on texture features of spectrogram[J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(9): 215-220. (in Chinese)

[17] DU X, CARPENTIER L, TENG G, et al. Assessment of laying hens' thermal comfort using sound technology[J]. Sensors, 2020, 20(2): 473.

[18] 余礼根,杜天天,于沁杨,等. 基于多特征融合的蛋鸡发声识别方法研究[J]. 农业机械学报, 2022, 53(3): 259-265.
YU Ligen, DU Tiantian, YU Qinyang, et al. Recognition method of laying hens' vocalizations based on multi-feature fusion[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(3): 259-265. (in Chinese)

[19] CUAN K, ZHANG T, HUANG J, et al. Detection of avian influenza-infected chickens based on a chicken sound convolutional neural network[J]. Computers and Electronics in Agriculture, 2020, 178: 105688.

[20] CUAN K, LI Z, ZHANG T, et al. Gender determination of domestic chicks based on vocalization signals[J]. Computers and Electronics in Agriculture, 2022, 199: 107172.

[21] CUAN K, ZHANG T, LI Z, et al. Automatic newcastle disease detection using sound technology and deep learning method [J]. Computers and Electronics in Agriculture, 2022, 194: 106740.

[22] CHARLES D, WALKER A. Poultry environment problems: a guide to solutions[J]. Poultry Science, 2001, 81(8): 1257 – 1258.

[23] 包婉婉,王炬,李静. 热应激对蛋鸡的影响及其防控措施[J]. 畜牧兽医科技信息,2022(10): 214 – 216.
BAO Wanwan, WANG Ju, LI Jing. The impact of heat stress on laying hens and its prevention and control measures[J]. Animal Husbandry and Veterinary Technology Information, 2022(10): 214 – 216. (in Chinese)

[24] LARA, L J, ROSTAGNO M H. Impact of heat stress on poultry production[J]. Animals, 2013(3): 356 – 369.

[25] 吴晓辉. 鸡热应激状态和耐热性的“非致死”性判断方法研究[D]. 武汉:华中农业大学, 2022.
WU Xiaohui. The methods to identify cicken heat stress states and heat tolerance in non-lethal condition[D]. Wuhan: Huazhong Agricultural University, 2022. (in Chinese)

[26] WU X, ZHENG B, MEI Z, et al. Key parameters of physiological responses to acute heat stress in two commercial layers determined by fisher discriminant analyses[J]. Journal of Thermal Biology, 2023, 117: 103694.

[27] HOWARD A, SANDLER M, CHU G, et al. Searching for MobileNetV3[C] // Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 1314 – 1324.

[28] WANG Q, WU B, ZHU P, et al. ECA – Net: efficient channel attention for deep convolutional neural networks[C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 11534 – 11542.

[29] HOU Q, ZHOU D, FENG J. Coordinate attention for efficient mobile network design[C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 13713 – 13722.

[30] WOO S, PARK J, LEE J Y, et al. Cbam: convolutional block attention module[C] //Proceedings of the European Conference on Computer Vision (ECCV), 2018: 3 – 19.

[31] MA N, ZHANG X, ZHENG H T, et al. ShuffleNet V2: practical guidelines for efficient CNN architecture design[C] //15th European Conference on Computer Vision (ECCV), 2018: 122 – 138.

[32] HUANG G, LIU Z, VAN DER MAATEN L, et al. Densely connected convolutional networks[C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 4700 – 4708.

[33] TAN M, LE Q. Efficientnetv2: smaller models and faster training[C] //International Conference on Machine Learning, 2021: 10096 – 10106.