

基于轻量化卷积神经网络的笼养鸡异常发声识别方法

唐瑜嵘^{1,2} 解彬彬^{2,3} 范甲骏^{2,3} 袁超^{2,3} 李尚民⁴ 童海兵⁴ 沈明霞^{2,3}

(1. 南京农业大学工学院, 南京 210031; 2. 农业农村部养殖装备重点实验室, 南京 210031;

3. 南京农业大学人工智能学院, 南京 210031; 4. 江苏省家禽科学研究所, 扬州 225125)

摘要: 针对笼养鸡舍复杂声环境下鸡只异常发声识别易受背景噪声干扰、异常样本稀缺及类别不平衡影响而导致识别稳定性不足的问题,提出一种基于轻量化卷积神经网络(CNN)的鸡只发声三分类识别方法。以鸡舍采集的“叫声”“喷嚏类”和“其他”3类声音为研究对象,构建包含691段音频样本的数据集,按训练集、验证集和测试集进行划分,并将音频转换为128×130的对数梅尔频谱图作为模型输入。在此基础上设计轻量级SimpleCNN模型,并与Inception V3和ResNet-50进行对比试验。结果表明:SimpleCNN在测试集上的总体准确率达到0.95,“叫声”“喷嚏类”和“其他”3类的F1分数分别为0.95、0.91和0.98,对少数类“喷嚏类”和噪声类“其他”均表现出较好的识别能力;相比之下,Inception V3和ResNet-50的总体准确率分别为0.56和0.32,且存在明显类别偏置,其中ResNet-50出现类别塌缩现象。混淆矩阵与可解释性分析表明,SimpleCNN能够更有效地聚焦目标发声事件的关键时频区域,有效降低“喷嚏类”与“其他”之间的混淆。研究结果表明,轻量化CNN结构在小样本、强噪声条件下具有较好的判别稳定性和部署潜力,可为鸡舍异常声音监测与非接触式健康预警提供技术支撑。

关键词: 笼养鸡舍; 鸡只发声识别; 对数梅尔频谱图; 轻量化卷积神经网络

中图分类号: S126

文献标识码: A

文章编号: 1000-1298(2026)15-0024-12

OSID:



Lightweight Convolutional Neural Network-based Method for Abnormal Vocalization Recognition in Cage-housed Chickens

Tang Yurong^{1,2} Xie Binbin^{2,3} Fan Jiajun^{2,3} Yuan Chao^{2,3} Li Shangmin⁴ Tong Haibing⁴ Shen Mingxia^{2,3}

(1. College of Engineering, Nanjing Agricultural University, Nanjing 210031, China

2. Key Laboratory of Livestock Equipment, Ministry of Agriculture and Rural Affairs, Nanjing 210031, China

3. College of Artificial Intelligence, Nanjing Agricultural University, Nanjing 210031, China

4. Jiangsu Institute of Poultry Science, Yangzhou 225125, China)

Abstract: Aiming to address the limited stability of abnormal vocalization recognition in caged layer houses under complex acoustic conditions, where strong background noise, scarce abnormal samples, and class imbalance pose major challenges, a lightweight convolutional neural network (CNN) for three-class chicken sound classification was proposed. A dataset of 691 audio clips, including calls, sneeze-like sounds, and other sounds, was constructed from recordings collected in commercial hen houses and divided into training, validation, and test sets. Each clip was converted into a 128×130 log-Mel spectrogram as model input. A lightweight SimpleCNN was developed and compared with Inception V3 and ResNet-50. Results showed that SimpleCNN achieved an overall test accuracy of 0.95, with F1-scores of 0.95, 0.91, and 0.98 for calls, sneeze-like sounds, and other sounds, respectively, demonstrating strong recognition performance for both the minority abnormal class and the noise class. In contrast, Inception V3 and ResNet-50 achieved overall accuracies of 0.56 and 0.32, respectively, and showed clear class bias, with ResNet-50 exhibiting class collapse. Confusion matrix and interpretability analyses further indicated that SimpleCNN more effectively focused on key time-frequency regions of target vocal events and reduced confusion between sneeze-like sounds and other sounds. These findings suggested that lightweight CNNs offered robust discrimination and practical deployment potential under

收稿日期: 2026-04-10 修回日期: 2026-04-15

基金项目: 江苏省重点及面上项目(BE2022379)

作者简介: 唐瑜嵘(1997—),男,博士生,主要从事智慧养鸡技术研究,E-mail: 2022212003@stu.njau.edu.cn

通信作者: 沈明霞(1964—),女,教授,博士生导师,主要从事智慧畜牧业研究,E-mail: mingxia@njau.edu.cn

small-sample, high-noise conditions, providing technical support for abnormal sound monitoring and non-contact health warning in layer houses.

Key words: cage-house chicken; poultry vocalization recognition; log – Mel spectrogram; lightweight convolutional neural network

0 引言

动物发声是反映其生理状态、行为活动和环境适应能力的重要信息载体。对于规模化家禽养殖而言,鸡只发声中包含丰富的健康、应激与福利状态信息,是实现非接触式连续监测的重要信号来源^[1-5]。随着笼养规模不断扩大,鸡舍内饲养密度高、人工巡检成本高且主观性较强,传统依赖人工观察的健康监测方式已难以满足精准养殖和智能化管理的需求。相比体表测温和人工观察等方式,基于声音的监测方法具有非侵入、连续性强、易部署等优点,因此逐渐成为智能养殖研究的重要方向之一^[1-5]。

在畜禽声音识别研究中,相关方法经历了由传统特征工程向深度学习表征学习的发展过程。早期研究主要提取梅尔频率倒谱系数(MFCC)、短时能量、过零率和谱熵等手工声学特征,并结合支持向量机、k 近邻和随机森林等浅层模型实现发声分类或异常检测^[3,6-7]。该类方法具有实现简单、可解释性较强等特点,在受控场景下能够取得较好的识别效果。但在实际鸡舍环境中,风机、湿帘、饲喂和饮水系统等设备噪声持续存在,声场条件复杂且非平稳,导致传统手工特征对瞬态异常声和局部频带差异的表达能力有限,模型鲁棒性和跨场景泛化能力受到制约^[6,8-9]。

近年来,卷积神经网络(CNN)、长短期记忆网络(LSTM)、卷积循环神经网络(CRNN)及注意力机制等深度学习方法在声学事件识别领域取得了较好效果,并逐步应用于家禽发声分析^[4,10-12]。相关研究表明,深度模型能够通过对时频图像的自动特征学习,提高复杂背景下的识别精度;同时,面向实际应用的轻量化网络和边缘推理方法也受到关注^[13-19]。此外,已有研究还从二氧化碳浓度、空腹压力、疫苗反应、情绪应激和雏鸡性别识别等方面验证了鸡只发声信号在健康与行为评估中的应用潜力^[20-26]。但是不同类型模型在鸡舍发声识别任务中的适用性并不相同。浅层模型通常依赖人工设计特征,结构简单、训练成本较低,在样本有限时具有一定稳定性,但其特征表达能力受限,难以充分刻画异常呼吸声等短时瞬态事件的局部时频模式;深层模型具有更强的非线性表征能力和特征自动学习能力,能够在大规模数据条件下挖掘更复杂的判别信

息,但在小样本、强噪声和类别不平衡场景下,往往面临参数冗余、训练不稳定及多数类偏置等问题;轻量化模型则在参数规模、计算开销和部署适应性方面具有优势,更适合边缘设备或现场实时监测场景,但其在真实鸡舍复杂声环境中的识别能力、少数类检出效果及稳定性仍有待进一步验证。特别是针对“喷嚏类”等少数异常事件,模型不仅需要具备足够的局部特征提取能力,还应在有限样本条件下保持较好的泛化性能与类别平衡能力。

总体来看,现有研究为家禽发声智能解析提供了重要基础,但针对真实笼养鸡舍场景下的异常发声识别,仍存在以下不足:现有研究多关注一般性的声音分类精度,对小样本条件下模型鲁棒性和训练稳定性的讨论相对不足;针对少数异常类别和类别不平衡问题的研究仍不够深入,模型在异常事件检出中易出现漏检或误报;关于浅层模型、深层模型与轻量化模型在真实复杂鸡舍场景中的适用边界尚缺乏系统对比分析^[4,8-9,13]。因此,有必要围绕小样本、强噪声和类别不平衡条件下的鸡只异常发声识别,探索兼顾识别性能、模型复杂度与部署可行性的轻量化建模方法。

在研究对象选择上,本文选取“叫声”“喷嚏类”和“其他”3 类声音事件,主要基于鸡舍现场监测需求、声学可分性及健康关联性的综合考虑。其中,“叫声”作为鸡群日常最常见的发声类型,可反映鸡只基础行为活动与正常声学状态,是识别异常声音的重要参照类别;“喷嚏类”包括喷嚏、咳嗽、甩鼻声等异常呼吸相关短时瞬态声,与呼吸道异常、应激反应及潜在健康风险具有较直接关联,是开展异常检测与早期预警的重点对象;“其他”则包含设备噪声、行为杂音及非目标个体发声等复杂背景声,设置该类有助于提升模型对真实鸡舍复杂声环境的适应能力,避免将环境扰动误判为异常事件。因而,上述 3 类既覆盖了正常发声、潜在异常发声和现场干扰声 3 种典型情形,又能够为后续由异常事件识别进一步拓展到鸡群健康状态评估提供基础。

针对真实笼养鸡舍声环境中背景噪声强、异常发声样本稀缺及类别分布不均衡,导致异常声音识别稳定性不足的问题,本文将重点研究小样本、强噪声条件下鸡只异常发声的稳健识别。以“叫声”“喷

嚏类”和“其他”3类声音事件为研究对象,提出一种基于对数梅尔频谱图和轻量化卷积神经网络的鸡只发声三分类方法。通过构建来源于真实鸡舍环境的音频数据集,以 SimpleCNN 为核心模型,并与 Inception V3 和 ResNet-50 进行对比,将研究重点放在轻量化模型对复杂鸡舍声环境的适应性验证上,系统分析其在小样本条件下的识别稳定性、少数类检出能力及实际部署可行性。

1 材料与方法

1.1 试验场景

本研究数据于 2024 年 12 月 10—30 日采集于河北省石家庄市行唐县峪口禽业规模化笼养蛋种鸡舍,试验对象为 150~170 日龄京粉 6 号蛋种鸡。鸡舍音频采集采用农业农村部养殖装备重点实验室自主研发的声音采集装置(图 1),该装置由拾音模块和信号采集存储模块组成,支持连续实时录音。系统采样频率为 22 050 Hz,量化位数为 16 bit,可较完整保留鸡只发声及鸡舍背景声的时域与频域信息。拾音器布设于鸡舍两侧墙体,距地面约 1.5 m、距笼具约 30 cm,能够满足复杂声环境下长时程、多点位声音信号的稳定采集需求,为后续异常发声识别与声学特征提取提供可靠数据来源。

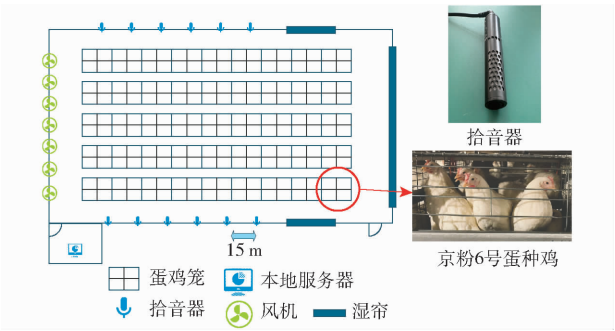


图 1 鸡舍布局与数据采集示意图
Fig. 1 Schematic of chicken coop layout and data collection

采集期间,鸡舍通风系统由饲养员根据每日外界环境情况间歇运行;饲喂与饮水系统每日于 08:30—09:30、12:30—13:30 和 18:30—19:30 运行;拾音设备连续 24 h 工作。上述工况设置旨在覆盖实际生产中常见的背景声源及位置差异,提高模型对现场监测条件的适应性以及在真实复杂声场下的抗噪性与泛化能力^[16]。

在数据整理基础上,围绕鸡舍声学监测需求构建了 3 类声音事件数据集:正常鸡叫声,记为“叫声”;鸡打喷嚏、咳嗽、甩鼻声等异常呼吸相关的短时瞬态声,统称为“喷嚏类”;公鸡声和非目标环境背景声,记为“其他”。其中,“其他”类主要包括公

鸡打鸣、鸡群拍翅、踏笼等行为杂音,以及风机运转、饲喂机械等典型设备噪声。具体分类标准见表 1。表中“叫声”“喷嚏类”和“其他”3 类的声学特征描述,主要依据本研究对原始波形、时频谱图及能量分布的人工复核结果,并结合已有关于鸡只发声监测、呼吸异常检测及鸡舍复杂声景分析的相关研究进行归纳^[17]。

表 1 声音类别特征

Tab. 1 Acoustic characteristics of sound classes			
类别	特征描述	时域特征	能量分布
叫声	单次叫声节奏平稳	波形平滑	较为均匀
喷嚏类	短促、急促	波峰明显	瞬时能量高
其他	随机性强	缺少明显特征	较为分散

1.2 数据集构建与标注规则

为刻画鸡舍空间结构与设备布局引入的声学差异,数据采集采用与 1.1 节一致的固定点位分区布设方案,在鸡舍风机端与湿帘端(进风口侧)分别设置拾音器采集点,拾音器安装于鸡舍两侧墙体,进行 24 h 连续、多日录音采集,以获取不同空间位置条件下的声音分布特征。每个声音样本由人工从连续录音中截取形成。对于持续时间较短的异常呼吸声事件,截取窗口以目标声事件为中心,并适当保留前后少量上下文背景,以完整覆盖其时域与频域特征。本文不采用将同一连续录音机械切分为多个高度重叠相邻样本的方式扩充数据集;若同一录音段中包含多个相邻声事件,则按事件边界分别标注并截取,避免同一声源事件在不同样本中重复出现。数据集划分时以“事件段”为最小独立单元,保证来源于同一连续录音段、同一时间邻域或同一声源事件的样本不会同时出现在训练集、验证集和测试集中,从而尽可能避免信息泄漏并提高评估结果的可靠性。

如表 2 所示,最终共获得 691 段音频样本,其中“叫声”约占 53% (主要包含日常鸣叫与产蛋相关叫声)，“喷嚏类”约占 25%，“其他”约占 22%。数据集首先按比例 8:2 划分为训练集与测试集;随后在训练集中再按 9:1 划分为训练子集与验证集,即训练/验证/测试比例约为 72:8:20,验证集仅用于超参数选择与 early stopping (早停策略)的监控。划分过程采用按类别比例的分层抽样,并以“录音段”

表 2 数据集划分结果

Tab. 2 Dataset partitioning results				
声音类别	训练集	验证集	测试集	总数
	样本数	样本数	样本数	
叫声	264	29	74	367
喷嚏类	124	14	35	173
其他	109	12	30	151

为最小单位进行划分,保证同一录音段,以及同一天、同一采集位置样本不会同时出现在不同集合中,以避免数据泄漏。测试集中各类别保留约 20% 的样本,以确保评估结果具有代表性与统计意义。

1.3 音频特征表示

在特征提取前,首先对原始音频进行去噪与幅值归一化处理(图 2)。具体地,将波形振幅线性归一到 $[-1,1]$ 区间,以减小录制增益、采集距离变化等因素引起的幅度差异,从而提升信号一致性并有利于后续特征提取的稳定性与准确性^[18]。随后,将时域信号转换为对数梅尔频谱图^[19]作为模型输入特征。具体而言,采用 Librosa(用于音频信号分析的开源 Python 库)计算功率谱并映射至 128 维梅尔滤波器组,得到尺寸为 128(频率维) $\times$ 130(时间帧)的梅尔谱矩阵;进一步对谱能量进行对数压缩,以降低动态范围并增强弱能量成分的可分辨性^[16]。图 3 为“喷嚏类”样本的梅尔谱示例,图 4 为鸡只叫声与喷嚏类的谱图对比。与原始波形相比,梅尔谱能够更直观地表征发声事件的时频结构,为卷积神经网络提取局部时频纹理特征(如谐波能量带与瞬态宽带脉冲)提供有效输入表示^[19]。

在比较不同声学表示方法并结合相关研究的基础上^[20],本文选取对数梅尔频谱图作为模型输入特征,而未采用单帧频谱或 WMFCC 作为主要判别依据。其主要原因在于,本研究关注的“喷嚏类”属于持续时间短、瞬态特征突出的声事件,其判别信息主要体现在时频联合结构上,如突发能量块、宽带瞬态分量及其时间演化过程。相比之下,单帧幅度频谱或整段平均频谱难以保留关键时序信息,且在鸡舍非平稳背景噪声干扰下更易受到整体能量分布变化的影响^[21]。对数梅尔频谱图能够在保留声事件时

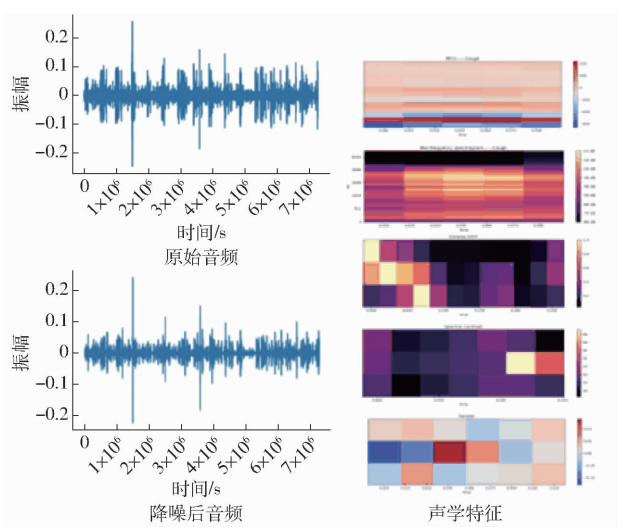


图 2 声音降噪示意图
Fig. 2 Schematic of audio denoising

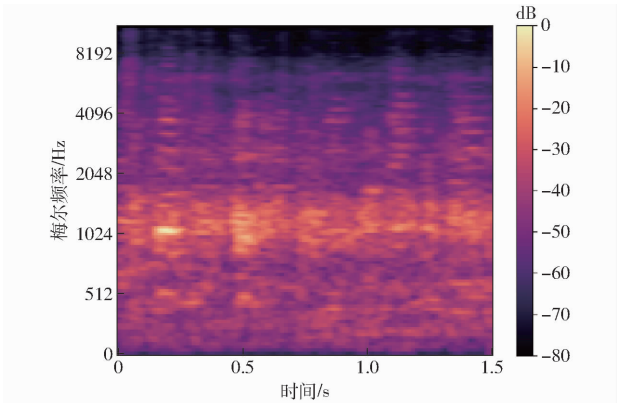


图 3 鸡只喷嚏类声音的梅尔对数功率谱
Fig. 3 Mel logarithmic power spectrum of chicken sneeze sound

频结构的基础上,通过梅尔尺度压缩和对数变换减小动态范围差异,更有利于卷积神经网络提取局部时频纹理特征,并增强对设备噪声和采集位置差异

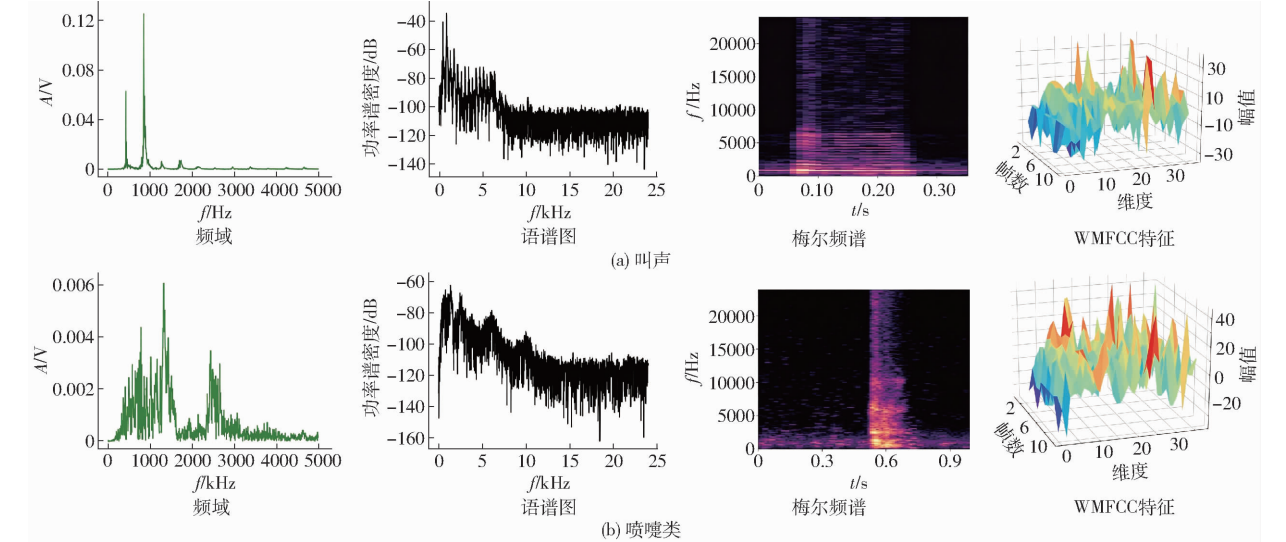


图 4 鸡只叫声与喷嚏类声音的不同图谱对比
Fig. 4 Comparison of spectrograms for chicken calls and sneezes

的适应性^[22]。

相较而言,MFCC 虽具有较好的工程可解释性和计算效率,但其低维压缩表示可能削弱对喷嚏类局部瞬态纹理的刻画能力,在复杂工况下对参数设置和分布漂移也更为敏感^[23]。本文未将 WMFCC 作为主特征,主要是考虑到鸡只品种和疾病类型差异较大,频带加权策略尚缺乏统一标准,且该方法对权重设定较为敏感;在鸡舍背景噪声随设备工况变化的场景下,不恰当的频带加权可能引入额外偏置,进而影响模型稳定性。基于上述考虑,本文最终采用对数梅尔频谱图作为主要输入特征;后续结合特征消融与融合试验,进一步评估 MFCC、WMFCC 及可学习加权策略对少数类异常事件识别性能的提升潜力。

2 试验

2.1 模型结构设计

为系统评估不同复杂度卷积神经网络在笼养鸡舍发声三分类任务中的适用性,本文选取自建轻量级浅层网络 SimpleCNN、经典多尺度卷积网络 Inception V3 以及深层残差网络 ResNet - 5 作为对比模型。3 类模型分别对应轻量级、中深层多尺度和深层残差 3 种典型模型架构,能够较为全面地反映不同网络复杂度条件下模型性能与应用特征的差异。具体选择依据如下:

(1)特征需求

鸡只“叫声”与“喷嚏类”信号在对数梅尔频谱图上的差异,主要体现为局部时频纹理分布、瞬态能量聚集形式及其随时间演化的结构差别^[24]。卷积神经网络在局部感受野建模和层次化特征提取方面具有天然优势,能够有效挖掘谱图中的局部相关信息与判别性模式。因此,将 CNN 作为本研究发声识别任务的核心建模方法,具有明确的理论适配性和方法合理性。

(2)结构代表性

SimpleCNN 结构简洁、参数规模较小,能够表征轻量级浅层网络在有限特征抽取深度下的识别能力;Inception V3 通过并行卷积核实现多尺度特征提取,能够刻画不同时间尺度与频率尺度上的声学模式,代表中深层多尺度网络的典型思路;ResNet - 50 则借助残差连接增强深层特征学习能力,代表高容量深层网络在复杂表征学习中的性能上限。基于上述结构差异,将三者纳入统一试验框架,有助于从网络深度、参数规模与特征建模方式等多个维度开展系统比较。

(3)工程应用价值

笼养鸡舍异常发声识别任务通常面临样本规模有限、环境噪声复杂以及类别分布不均衡等现实约束,模型不仅需要具备较高识别精度,还应兼顾鲁棒性、训练稳定性与部署可行性^[25]。选取 SimpleCNN、Inception V3 和 ResNet - 50 进行对比,有助于揭示轻量模型中深层多尺度模型深层残差模型在性能收益与计算代价之间的差异关系,从而为实际鸡舍环境中的模型选型与系统部署提供依据。3 种模型均以对数梅尔频谱图作为输入,并输出“叫声”“喷嚏类”和“其他”3 类的后验概率分布,以保证比较条件的一致性。

在此基础上,本文构建 SimpleCNN 的目标并不在于单纯提升网络深度或增加结构复杂度,而在于面向鸡舍异常发声识别的小样本、高噪声与资源受限等应用场景,建立一种参数规模可控、训练过程稳定且具备较强样本适应能力的基础模型。基于这一考虑,SimpleCNN 的设计重点放在以较少层数提取关键时频局部特征,并通过控制模型复杂度降低深层网络在有限样本条件下的过拟合风险,从而提升模型在实际应用环境中的可用性与稳健性。

本文选取的 SimpleCNN、Inception V3、ResNet - 50 模型分别如图 5 ~ 7 所示。

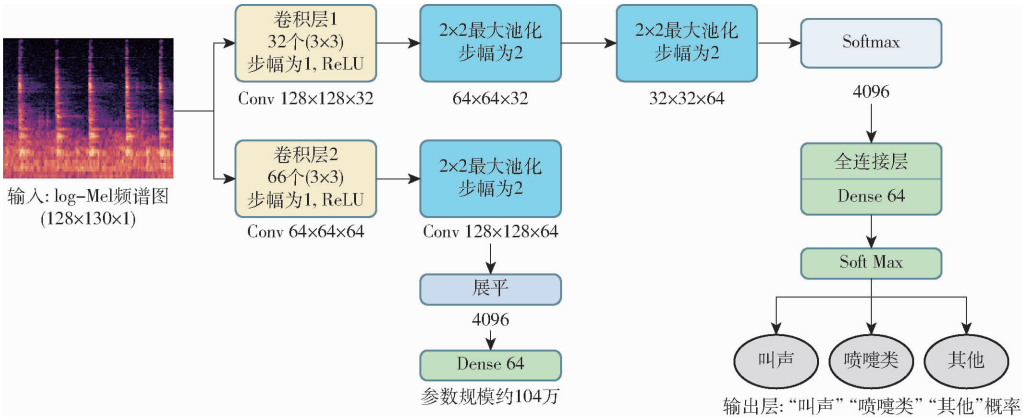


图 5 SimpleCNN 模型结构图
Fig. 5 SimpleCNN model structure

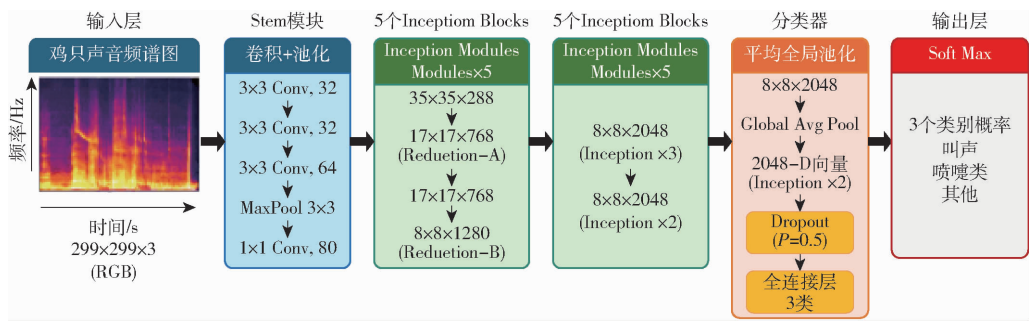


图 6 Inception V3 模型结构图

Fig. 6 Inception V3 model structure

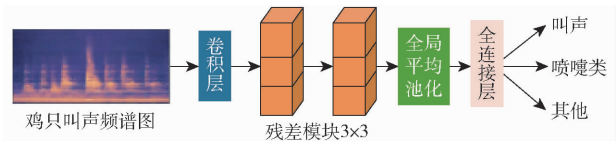


图 7 ResNet - 50 模型结构图

Fig. 7 ResNet - 50 model structure

2.2 超参数设置

3 种模型均采用自适应矩估计优化器进行参数更新,初始学习率设为 0.001,并使用默认动量参数 $\beta_1 = 0.9$ 、 $\beta_2 = 0.999$ 。损失函数选用多类别交叉熵,主要评价指标为分类准确率。训练过程中,批大小设置为 32,最大训练轮数设为 50,验证集来自前文所述训练集内部划分(占全数据集约 8%),用于 early stopping 与学习率衰减的监控。为避免过拟合并提升训练效率,采用 early stopping 策略:在验证集上监控准确率,当验证准确率连续 5 个 epoch 未提升时停止训练,并保留验证性能最优的模型参数。同时,为增强训练稳定性并加速收敛,采用自适应学习率衰减策略:当验证损失连续 3 个 epoch 未下降时,将学习率衰减为原来的 1/2。试验观察表明,Inception V3 与 ResNet - 50 由于网络更深、参数量更大,在训练初期收敛速度相对较慢且指标波动更明显;引入上述学习率衰减策略后,训练过程趋于稳定。最终 3 种模型通常在 20 个 epoch 达到收敛。

为保证比较结果的可解释性与可复现性,3 种模型采用一致的输入特征形式与相同的数据划分策略,并在同一训练策略下优化;对比模型仅更换网络主干与分类头结构,其余训练超参数保持一致。对于 Inception V3 与 ResNet - 50,均不加载预训练权重并移除原始顶层分类层,以避免外部数据迁移带来的干扰,从而更客观地评估不同网络容量在当前鸡舍数据规模下的适配性。

2.3 评价指标

为全面评估三分类模型在测试集上的识别性能,本文采用总体准确率、精确率、召回率和 F1 分数等指标进行综合评价。本文分别计算每一类别的精

确率、召回率和 F1 分数,同时报告宏平均(对各类别指标进行算术平均)以减少样本数量差异对整体评价的偏置,使评价更加关注少数类表现。

本文进一步绘制三分类混淆矩阵以开展误差分析。混淆矩阵为 3×3 的计数矩阵,其中元素 M_{ij} 表示真实类别为 i 的样本被预测为类别 j 的数量。通过混淆矩阵可直观定位主要误差来源,例如判断模型是否倾向于将“喷嚏类”误分为“其他”,或将部分“叫声”混淆为“其他”等,从而为后续结构改进与数据增强策略提供依据。

此外,为辅助理解类别不平衡对评价指标的影响,本文同步统计并报告测试集中各类别样本的支持数(即每类真实样本数量),以便对宏平均指标与类别级指标进行一致性解释。

3 结果与分析

3.1 模型总体性能

图 8 为 SimpleCNN、Inception V3 与 ResNet - 50 在测试集上的三分类性能,评价指标包括总体准确率及各类别的精确率、召回率、F1 分数。总体而言,SimpleCNN 在准确率与类别均衡性两方面均表现最优:其准确率达到 0.95,并在“叫声”“喷嚏类”“其他”3 类上均获得较高且相对一致的精确率、召回率、F1 分数。相比之下,Inception V3 与 ResNet - 50 的性能显著劣化,且呈现出典型的“类别偏置”或“边界失稳”现象,说明在当前数据规模与噪声条件下,高容量网络未能带来预期收益。

从类别级指标看,SimpleCNN 在 3 类发声识别任务中表现最为均衡。“叫声”类别 F1 分数为 0.95,“喷嚏类”为 0.91,“其他”为 0.98,说明该模型既能稳定识别正常鸣叫,也能较好区分异常短时事件与背景噪声,兼顾异常检出与噪声抑制能力。

相比之下,Inception V3 表现出明显的类别偏置。“叫声”类别召回率达 1.00,但精确率仅为 0.55;“喷嚏类”F1 分数为 0,“其他”类别召回率仅为 0.13。说明其预测结果主要集中于“叫声”类别,

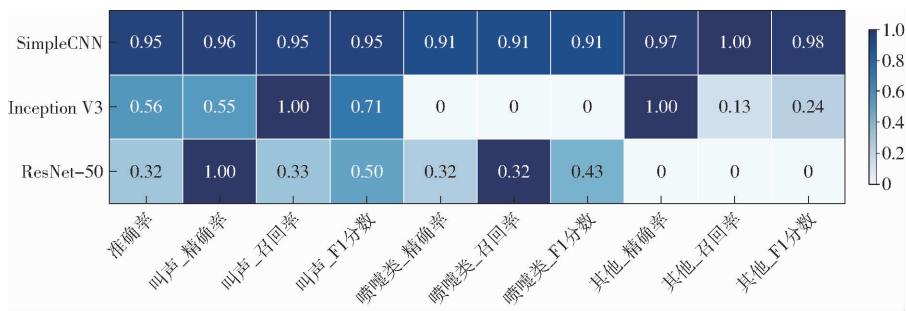


图 8 3 种模型在测试集上的分类性能热力图

Fig. 8 Heatmap of classification performance of three models on test set

未能形成对少数类的有效判别。在小样本、强噪声和类别不平衡条件下,较大参数量且从头训练的深层网络更易形成偏向多数类的决策边界。

ResNet-50 则表现为“喷嚏类”高召回率、低精确率:“喷嚏类”召回率为 1.00,但精确率仅为 0.27;“其他”类别 F1 分数同样为 0。说明该模型虽能覆盖异常样本,但误报较多,且对背景噪声缺乏稳定识别能力。总体来看,在本研究场景下,深层高容量网络并未表现出更优判别性能,反而暴露出其在复杂噪声、小样本和类别不平衡条件下的泛化不足。

从工程应用角度看,ResNet-50 的上述分类模式将明显削弱系统的实际可用性。一方面,“喷嚏类”的大量误报会导致异常报警频繁触发,降低预警结果的可信度;另一方面,“其他”类别无法被可靠识别,意味着设备运行噪声、环境扰动及非目标行为声音难以被有效剔除,从而持续干扰系统决策,影响报警机制的稳定性。其原因可能在于,鸡舍场景

中的短时瞬态噪声在局部谱纹理特征上与“喷嚏类”存在一定相似性,当深层网络在小样本条件下难以学习到稳定、清晰的类别边界时,往往倾向于通过扩大异常类覆盖范围来提高召回率,进而导致误报显著增加。由此可见,较深网络结构在本研究场景下并未有效转化为更优的类别判别能力,反而暴露出其在复杂噪声、小样本及类别不平衡条件下的泛化局限性。

3.2 混淆矩阵分析

为进一步刻画模型的误差来源与类别间混淆关系,本文给出 3 种模型在测试集上的混淆矩阵,如图 9 所示。类别顺序为“其他”“喷嚏类”“叫声”。对角线元素表示正确分类,非对角线元素表示误分类。识别标签 0、1、2 分别对应“其他”“喷嚏类”“叫声”。矩阵中的数字表示样本计数,并可换算为该真实类别内部的比例(%),便于直观比较不同类别的误判模式。

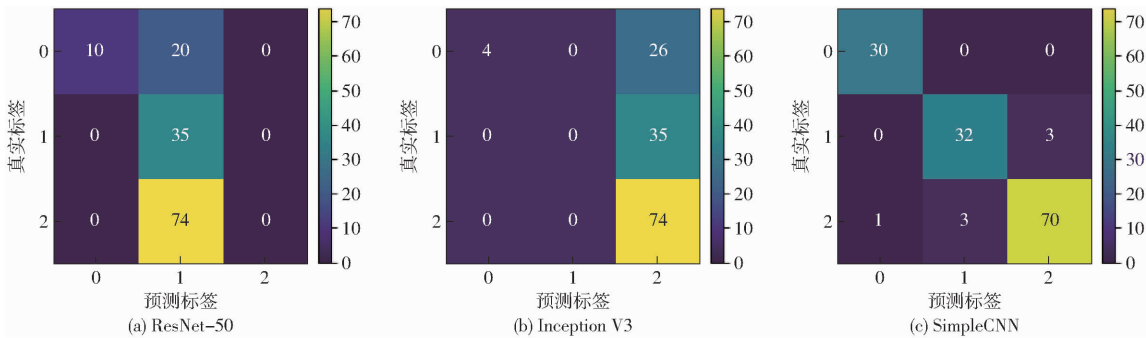


图 9 不同模型在测试集上的混淆矩阵

Fig. 9 Confusion matrices of different models on test set

从类别级指标来看,SimpleCNN 在 3 类发声识别任务中表现出较好的均衡性与稳定性。“叫声”类别的精确率、召回率和 F1 分数分别为 0.96、0.95 和 0.95;“喷嚏类”3 项指标均为 0.91;“其他”类别分别为 0.97、1.00 和 0.98。结果表明,SimpleCNN 不仅能够有效学习正常鸣叫的稳定时频特征,而且对异常短时事件和复杂背景噪声均具有较强区分能力,兼顾了异常检出与噪声抑制效果。

相比之下,Inception V3 表现出明显的类别失衡

特征。其“叫声”类别召回率达到 1.00,但精确率仅为 0.55,F1 分数为 0.71,说明模型存在显著的多数类偏置;“喷嚏类”的精确率、召回率和 F1 分数均为 0,表明该类别在测试阶段被完全漏检;“其他”类别虽有较高精确率,但召回率仅为 0.13,F1 分数为 0.24。总体来看,Inception V3 的预测结果主要集中于“叫声”类别,未能形成对少数类的有效判别,说明在小样本、强噪声和类别不平衡条件下,深层高容量网络更易形成偏向多数类的决策边界。

ResNet-50 则呈现出“喷嚏类”高召回率、低精确率的误差特征。“喷嚏类”召回率为 1.00,但精确率仅为 0.27,F1 分数为 0.43,说明模型虽能覆盖大多数真实异常样本,但误报较多;“叫声”类别精确率为 1.00,而召回率仅为 0.33,F1 分数为 0.50,表明其对该类判断较为保守;“其他”类别的精确率、召回率和 F1 分数均为 0,说明其对环境噪声缺乏稳定识别能力。由此可见,在本研究场景下,较深网络结构并未转化为更优的判别性能,反而暴露出其在复杂噪声、小样本和类别不平衡条件下的泛化局限性。

3.3 Grad-CAM 可视化解释

为解释不同模型在鸡舍场景下的判别依据,本文采用梯度加权类激活映射(Grad-CAM)对梅尔频谱输入的关键响应区域进行可视化分析(图 10)。从左到右依次为 SimpleCNN、Inception V3、ResNet-50;从上到下依次为“其他”“喷嚏类”“叫声”样本。颜色由冷到暖表示对当前类别判定的贡献由低到高(红/黄色为高响应区域)。结果表明,SimpleCNN 的高响应区域与目标事件的主要时频能量结构具有较高一致性:对“喷嚏类”样本,响应主要集中于中高频瞬态能量带及其邻近区域;对“叫声”样本,响

应则主要分布于稳定能量带和类谐波结构所在频段。这说明 SimpleCNN 更倾向于聚焦于发声事件本身的关键时频证据,而非对环境背景或整体能量变化产生过度响应,更符合鸡舍连续监测与异常报警的实际需求。

相比之下,Inception V3 与 ResNet-50 的响应区域相对分散,部分样本对大范围频带或背景区域表现出较强响应,表明其判别过程可能更多依赖全局谱或背景能量模式。在小样本和强噪声条件下,这种响应方式更易引发类别偏置、漏检或误报,与混淆矩阵中观察到的误判现象相一致。综上,Grad-CAM 结果从可解释性角度进一步表明,轻量化的 SimpleCNN 更能稳定关注与鸡只发声事件相关的关键时频区域,因此具有更高的判别可靠性。

3.4 叫声与喷嚏类异常呼吸声的频谱差异分析

为从声学机理角度解释模型对“叫声”“喷嚏类”两类事件的判别依据,本文进一步对两类样本进行频域分析,如图 11 所示。整体来看,两类声音在频谱形态与关键频率指标上呈现出稳定差异,可为后续模型设计与误差归因提供支持。

首先,从单类幅度谱的统计分布(图 11a、11b)

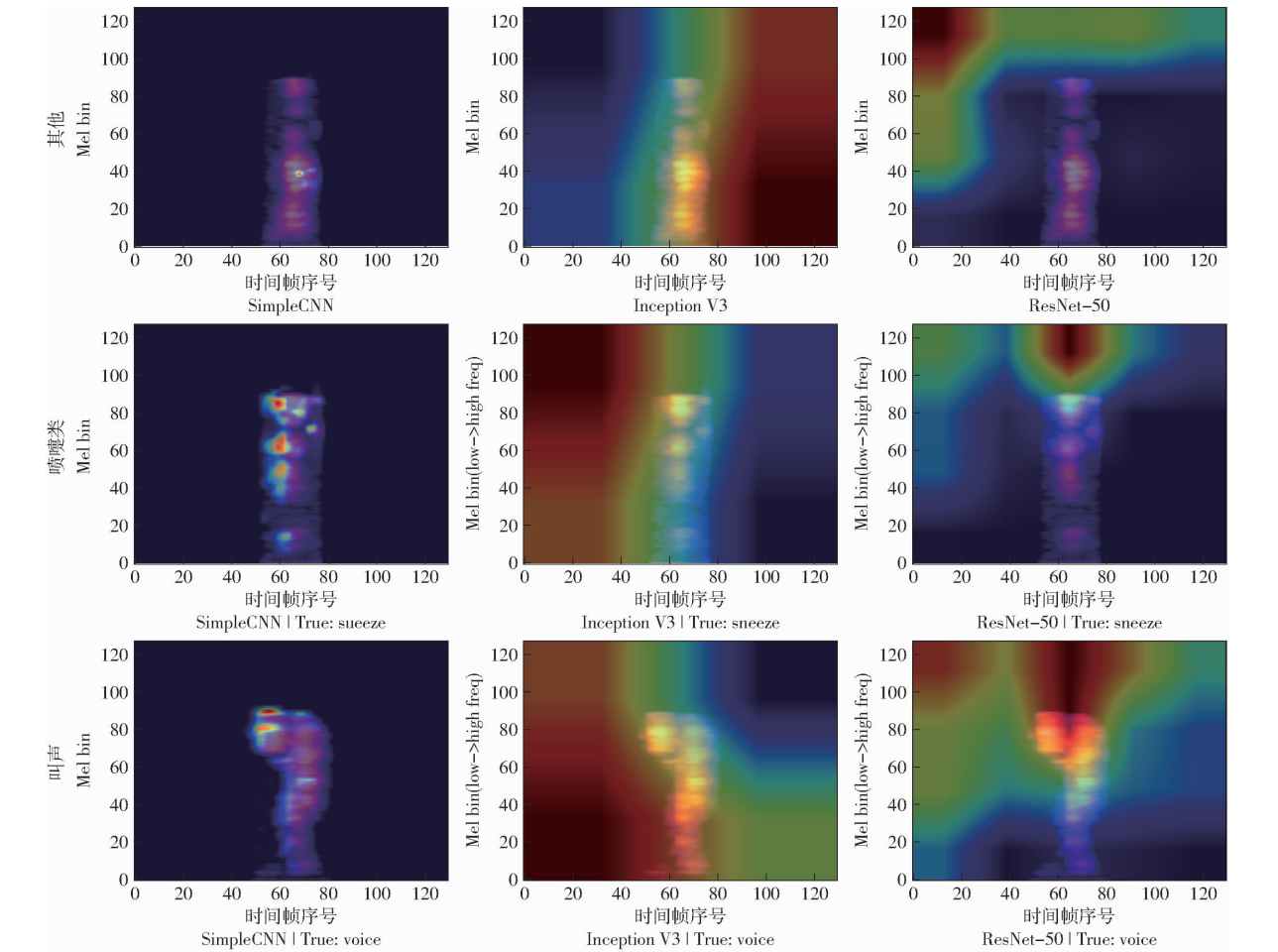


图 10 不同模型在梅尔频谱图上的 Grad-CAM 可视化结果

Fig. 10 Grad-CAM visualization results of different models on Mel spectrogram

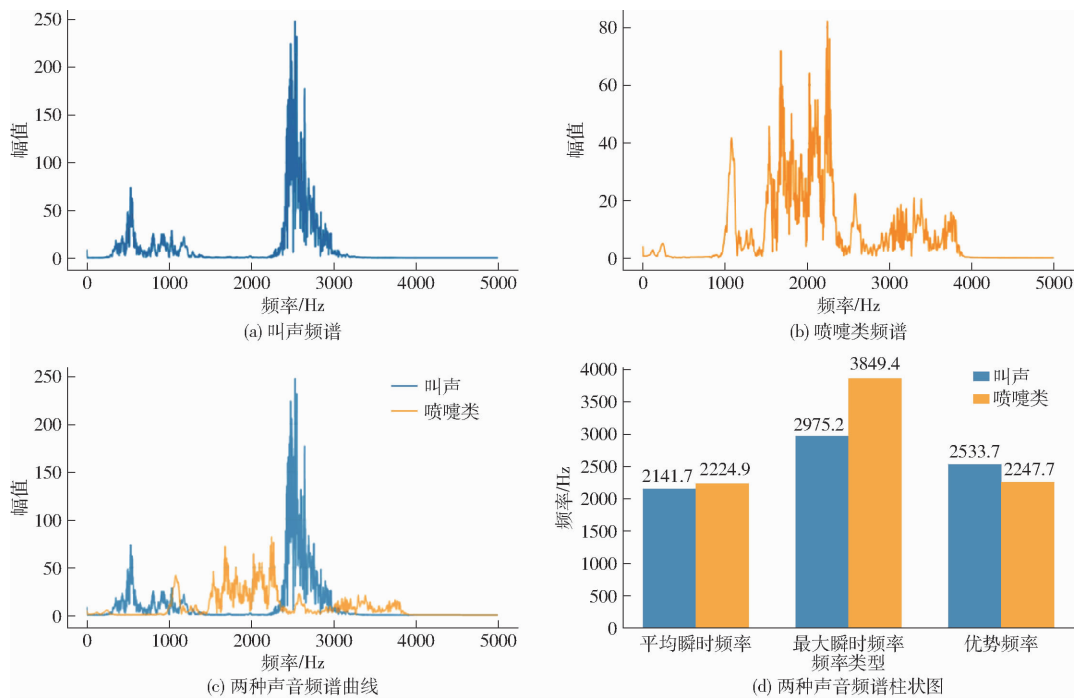


图 11 叫声与喷嚏类的频谱对比及关键频率指标

Fig. 11 Spectral comparison of voice and sneeze sounds and statistics of key frequency indicators

可见,“叫声”在 2 ~ 3 kHz 区间呈现明显的能量聚集,并形成相对突出的主峰结构,表明其频谱分布具有较强的集中性与稳定性,符合较为持续、周期性发声所对应的共振/谐波特征。相比之下,“喷嚏类”的谱能量在更宽的频带范围内分布,中频段存在多个局部峰值,同时高频段仍维持一定幅度响应,整体表现出更典型的宽带瞬态特征。上述差异与两类声源的产生机制一致:叫声对应较稳定的声带振动过程,而喷嚏类属于短时呼吸冲击事件,往往包含更丰富的瞬态宽带成分,因此其频谱能量分布更为离散。

其次,叠加谱对比结果(图 11c)进一步揭示了“叫声”与“喷嚏类”在频域上的共性与差异。两类声音在中频段均存在较明显的能量重叠,说明该频段可能是类别混淆的主要来源;而在高频段,喷嚏类更常表现出能量延伸或局部尖峰增强特征,因此可作为区分两类事件的重要频域线索。这一现象也可在一定程度上解释混淆矩阵中少量“叫声-喷嚏类”误判,即当部分叫声样本叠加背景噪声或出现突发高频成分时,其局部谱形可能接近喷嚏类的瞬态结构,从而增加边界样本的判别难度。

关键频率统计指标(图 11d)为上述差异提供了定量支撑,据结果表明,喷嚏类的最大频率为 3 849.4 Hz,高于叫声的 2 975.2 Hz,说明其更易产生高频瞬态成分;两类声音的平均频率整体接近,但喷嚏类略高于叫声,表明其频谱重心整体偏向高频。相比之下,叫声的主导频率高于喷嚏类,说明其主要能量峰位置更集中、分布更稳定。总体来看,喷嚏类

具有最大频率更高、频带分布更宽的特征,而叫声则表现为主峰更集中、主导频率更稳定,这为模型区分两类声音提供了较为明确的声学依据。

3.5 特征表示选择与声学可解释性分析

特征表示的选择对鸡舍环境下的细粒度发声识别具有重要影响。本文采用对数梅尔频谱图(log-Mel spectrogram)作为卷积神经网络的输入,而未直接使用原始波形或传统 MFCC 特征。其主要原因在于,log-Mel 具有二维时频结构,能够更完整地保留局部谱形与时间变化信息,便于 CNN 以类似纹理学习的方式提取判别特征,如叫声较稳定的谐波/共振结构,以及喷嚏类更突出的瞬态宽带能量分布。相比之下,MFCC 虽能在一定程度上表征谱形信息,但其低维压缩过程可能削弱短时异常事件所对应的局部细粒度差异。

为验证 log-Mel 作为输入特征的合理性,本文在相同数据划分、网络结构和训练策略下,对 log-Mel、MFCC 和原始波形输入进行了对比试验(表 3)。结果表明,log-Mel 在准确率、宏 F1 分数、宏召回率等指标上均取得最优结果,尤其在“喷嚏类”这一少数类异常事件上的精确率、召回率和 F1 分数表现最佳,说明其更有利于表征短时瞬态异常呼吸声的局部时频纹理与宽带能量特征。相比之下,MFCC 受限于低维压缩,对细粒度谱形差异的保留不足;原始波形输入整体性能最弱,表明在小样本、强噪声条件下,端到端时域建模对训练稳定性和特征学习能力要求更高,难以充分发挥优势。

表 3 不同输入特征表示性能比较

Tab.3 Comparison of different input representations						
输入特征表示	准确	宏 F1	宏召	喷嚏类	喷嚏类	喷嚏类
	率	分数	回率	精确率	召回率	F1 分数
log-Mel	0.94	0.93	0.94	0.90	0.89	0.89
MFCC	0.90	0.88	0.88	0.84	0.79	0.81
Raw waveform (1-D CNN)	0.87	0.84	0.85	0.74	0.74	0.75

为揭示时域与频域信息在声学表征中的互补作用,图 12 展示了叫声样本 (5h_chicken_call_27.wav) 的时域波形与幅度频谱。时域上,该样本有效发声时长约 0.19 s,呈现两个较明显的能量集中阶段,具有突出的瞬态特征;频域上,其主要频率成分集中于 1~2.5 kHz,并伴随多个显著谱峰,同时在更高频段仍存在较弱但清晰的频率分量,表现出典型的类谐波或多共振峰结构。上述结果表明,时域波形有助于刻画能量包络及其动态变化,频域谱则更适于揭示频率组成特征,而梅尔频谱图能够在统一框架下同时表征时间维与频率维信息,使模型综合利用谐波纹理、能量带分布和时间包络等线索,从而更稳定地区分叫声、喷嚏类与环境噪声。

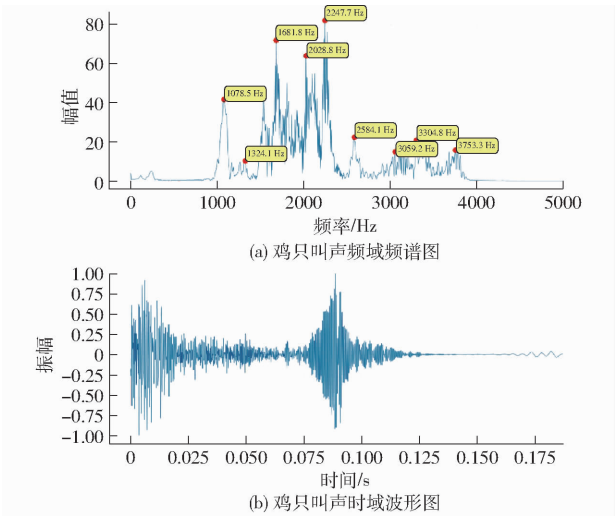


图 12 叫声样本的时域波形与幅度频谱
Fig. 12 Time-domain waveform and amplitude spectrum of chicken call sample

此外,本文还尝试了以原始波形作为深度模型输入的建模方式,但其整体性能低于基于频域特征的方案。其原因在于,在采样频率为 22 050 Hz、输入长度统一为 1 s 的条件下,每条样本包含约 2.2×10^4 个采样点,端到端波形建模对模型容量、参数优化和训练稳定性要求更高;同时,原始波形幅值易受录制距离、采集朝向及背景噪声变化影响,导致跨场景泛化难度增加。相比之下,梅尔频谱图在一定程度上模拟了人耳对频率的非线性感知特性,因而具有更好的鲁棒性与可解释性,也更便于将模型关注

频带与实际声学现象相对应。因此,在本研究场景下,基于梅尔频谱图的特征表达更适合作为鸡只异常发声识别的输入形式。

3.6 预训练策略与轻量级基线对比试验

由于大模型在小样本条件下未使用预训练权重,进而影响模型比较公平性,本文进一步进行预训练迁移学习及轻量级主流模型对比试验。在保持数据划分方式、输入特征形式、训练轮次以及优化器设置一致的前提下,分别对 Inception V3 和 ResNet-50 设置“随机初始化”与“ImageNet 预训练初始化”两种训练策略,并引入 MobileNetV2、EfficientNet-B0 和 ShuffleNetV2 作为轻量级基线模型。上述模型均面向资源受限场景设计,具有代表性的轻量化结构特点。所有模型均在相同数据划分、输入特征及训练策略下进行训练与测试,以保证比较公平性。所有模型均以对数梅尔频谱图作为输入,并统一将分类头替换为三分类输出层,以尽可能控制无关变量,保证试验比较的可比性。评价指标包括准确率、宏 F1 分数、宏召回率,以及“喷嚏类”的精确率、召回率和 F1 分数。

由表 4 结果可知,在小样本鸡舍声音识别任务中,预训练初始化对中深层模型的性能提升作用较为明显。与随机初始化相比,采用预训练权重后的 Inception V3 和 ResNet-50 在准确率、宏 F1 分数以及“喷嚏类”召回率等指标上均表现出不同程度的改善,说明迁移学习能够在一定程度上缓解深层网络在有限样本条件下面临的训练不足、参数优化困难及过拟合风险。上述结果表明,在原始试验设置下未引入预训练权重,确实可能限制中深层模型性能的充分发挥;但与此同时,这一补充试验也进一步验证了预训练策略对于提升大容量模型在小样本场景下适应性的必要性。

表 4 预训练时 Inception V3 和 ResNet-50 性能

Tab.4 Impact of pre-training on Inception V3 and ResNet-50					
模型	初始化	准确率	ResNet-50		
			宏 F1 分数	宏召回率	喷嚏类召回率
Inception V3	随机	0.88	0.81	0.80	0.74
	预训练	0.90	0.85	0.84	0.78
ResNet-50	随机	0.87	0.80	0.79	0.72
	预训练	0.90	0.84	0.83	0.77

从表 5 的轻量级基线对比结果可以看出, MobileNetV2、ShuffleNetV2 和 EfficientNet-B0 均具有较好的分类性能,说明当前主流轻量化网络能够较好适应鸡舍发声识别任务。然而,SimpleCNN 在参数量更小的情况下,仍取得了与这些主流轻量级

模型相当甚至略优的结果,尤其在“喷嚏类”异常呼吸声识别指标上表现更为稳定。结果表明,在小样本、强噪声和类别不平衡条件下,针对局部时频模式进行建模的浅层轻量化卷积结构更符合该任务的数据特征,因而能够在保证模型简洁性的同时实现较好的综合识别性能。

表 5 轻量级模型与 SimpleCNN 对比
Tab.5 Comparison of lightweight models and SimpleCNN

模型	参数量	准确率	宏 F1 分数	宏召回率	喷嚏类 F1 分数
SimpleCNN	1.2×10^6	0.91	0.86	0.85	0.91
MobileNetV2	3.5×10^6	0.90	0.85	0.84	0.80
EfficientNet - B0	5.3×10^6	0.91	0.85	0.84	0.81
ShuffleNetV2	2.3×10^6	0.90	0.85	0.84	0.80
Inception V3 - pretrained	2.38×10^7	0.90	0.85	0.84	0.80
ResNet50 - pretrained	2.56×10^7	0.90	0.84	0.83	0.79

3.7 多次随机重复试验与模型稳定性分析

鉴于本研究数据规模相对有限,单次训练一测试划分下的试验结果可能受到随机初始化等因素影响,从而引起模型性能波动。为验证 SimpleCNN 在小样本条件下的训练稳定性,本文在保持训练集、验证集和测试集划分不变的前提下,设置 5 个不同随机种子对模型进行重复训练,并统计各项评价指标的均值与标准差。为考察模型对数据划分扰动的敏感性,本文补充开展了多次分层随机划分试验,并报告相应的平均性能。通过上述两类试验,从参数初始化随机性和数据划分扰动两个层面,对 SimpleCNN 在小样本鸡舍声音识别任务中的稳定性进行了系统验证。

表 6 结果显示,在不同随机种子条件下,SimpleCNN 的准确率、宏 F1 分数以及“喷嚏类”F1 分数等指标波动幅度均较小,各项指标的标准差保持在较低水平。这表明该模型对参数初始化及训练过程中的随机性不敏感,具有较好的训练稳定性与结果一致性。同时,多次分层随机划分试验所得平

表 6 多次随机种子重复训练结果

随机种子数	准确率	宏 F1 分数	宏召回率	喷嚏类精确率	喷嚏类召回率	喷嚏类 F1 分数
42	0.91	0.86	0.85	0.82	0.80	0.81
123	0.90	0.85	0.84	0.82	0.80	0.81
2 024	0.91	0.86	0.85	0.82	0.81	0.82
3 407	0.91	0.85	0.84	0.82	0.80	0.81
8 888	0.90	0.85	0.84	0.82	0.80	0.81
平均值 ± 标准差	0.906 ± 0.003	0.854 ± 0.003	0.844 ± 0.003	0.819 ± 0.003	0.802 ± 0.004	0.810 ± 0.004

均性能亦保持相对稳定,进一步说明 SimpleCNN 对数据划分扰动具有较强鲁棒性。总体而言,SimpleCNN 在小样本条件下不仅能够取得较优识别性能,而且表现出较好的重复训练稳定性和泛化可靠性。

3.8 模型优势分析

SimpleCNN 在“喷嚏类”和“其他”两类声音上的性能优势,主要得益于其模型容量、特征建模尺度与当前数据规模之间具有较好的匹配性。作为浅层轻量化网络,SimpleCNN 参数量较小,在小样本条件下更易稳定收敛,并能够在一定程度上降低类别不平衡导致的多数类偏置,因此在学习主类稳定模式的同时,仍能较好保留对少数类瞬态特征的辨识能力。相比之下,Inception V3 和 ResNet - 50 等高容量网络在未使用预训练权重且样本有限的条件下,更易受到类别不平衡和噪声干扰的影响,导致决策边界向多数类偏移,甚至出现类别塌缩现象。

从声学特征角度看,“喷嚏类”通常具有持续时间短、能量突发、频带较宽等瞬态特征,并可能伴随局部高频增强。SimpleCNN 采用 3×3 小卷积核和逐级下采样结构,能够较好地捕捉局部时频纹理及其中尺度模式,因此更适合提取异常发声的关键判别特征。而更深层网络在当前数据条件下更容易学习到与背景噪声或整体能量分布相关的非稳健特征,从而削弱对少数类异常事件的判别稳定性。

综上,在本研究的数据规模和鸡舍噪声条件下,轻量浅层结构表现出更好的泛化稳定性,尤其在少数类异常事件识别方面更具优势。但这一结论主要基于当前样本规模及从头训练设定;若进一步扩大数据规模,或引入预训练与迁移学习策略,高容量模型的性能潜力仍有待进一步挖掘。

4 结论

(1)针对笼养鸡舍复杂声环境下鸡只异常发声识别中存在的噪声干扰强、异常样本少和类别不平衡等问题,构建了基于对数梅尔频谱图的鸡只发声三分类识别方法,并对轻量化 SimpleCNN 与 Inception V3、ResNet - 50 两种深层网络进行了对比研究。

(2)构建了适用于笼养鸡舍真实声环境的鸡只异常发声识别数据与特征表达方法。面向噪声干扰强、异常样本少和类别不平衡等问题,建立了包含“叫声”“喷嚏类”和“其他”3 类声音事件的鸡只发声数据集,共 691 段样本;同时采用对数梅尔频谱图表征声音信号的时频结构,为模型提取局部判别特

征提供了有效输入。

(3)验证了轻量化 SimpleCNN 在小样本、强噪声条件下的识别优势。在相同数据划分和训练条件下,SimpleCNN 在测试集上取得了 0.95 的总体准确率,且“叫声”“喷嚏类”和“其他”3 类的 F1 分数分别达到 0.95、0.91 和 0.98,整体性能优于 Inception V3 和 ResNet-50。结果表明,在当前数据规模与噪声条件下,轻量化网络较深层高容量网络具有更好的任务适配性与泛化稳定性。

(4)揭示了不同模型识别差异的表现特征及其应用意义。混淆矩阵分析表明,SimpleCNN 能够较

好平衡异常事件检出与背景噪声抑制,显著降低“喷嚏类”与“其他”之间的混淆;而 Inception V3 和 ResNet-50 则表现出明显的类别偏置,ResNet-50 甚至出现类别塌缩现象。进一步结合 Grad-CAM 可视化与频谱分析发现,SimpleCNN 对目标事件关键时频区域的响应更为集中,说明轻量化卷积网络在鸡舍非接触式声学监测与异常报警中具有较好的应用潜力。后续仍需结合多位置、多季节和多品种数据开展跨场景验证,并通过引入注意力机制、多尺度特征提取及边缘端实时推理策略,进一步提升模型鲁棒性与现场应用能力。

参 考 文 献

[1] 袁超. 基于深度学习的肉鸡全生命周期咳嗽声识别预警系统设计与实现[D]. 南京:南京农业大学,2023.
Yuan Chao. Design and implementation of a deep learning-based cough sound recognition and early warning system for the full life cycle of broilers[D]. Nanjing: Nanjing Agricultural University, 2023. (in Chinese)

[2] 余礼根,庄晏榕,邱枫,等. 农业领域音频信息技术研究进展[J]. 农业机械学报,2025,56(10):223-248.
Yu Ligen, Zhuang Yanrong, Qiu Feng, et al. Research progress of audio information technology in agriculture[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(10): 223-248. (in Chinese)

[3] 唐瑜嵘,沈明霞,薛鸿翔,等. 人工智能技术在畜禽养殖业的发展现状与展望[J]. 智能化农业装备学报,2023,4(1):1-16.
Tang Yurong, Shen Mingxia, Xue Hongxiang, et al. Current status and prospects of artificial intelligence technology in livestock and poultry breeding[J]. Journal of Intelligent Agricultural Equipment, 2023, 4(1): 1-16. (in Chinese)

[4] 唐瑜嵘,解彬彬,刘龙申,等. 多模态技术在智能养鸡工厂中的研究现状与展望[J]. 南京农业大学学报,2025,48(3):515-527.
Tang Yurong, Xie Binbin, Liu Longshen, et al. Research status and prospects of multimodal technology in intelligent poultry farms[J]. Journal of Nanjing Agricultural University, 2025, 48(3): 515-527. (in Chinese)

[5] 邓为,孙卫东,刘龙申,等. 产蛋高峰期蛋鸡亚健康评价指标体系的构建[J]. 南京农业大学学报,2024,47(4):741-749.
Deng Wei, Sun Weidong, Liu Longshen, et al. Construction of an evaluation index system for sub-health status in laying hens during the peak laying period[J]. Journal of Nanjing Agricultural University, 2024, 47(4): 741-749. (in Chinese)

[6] Lin M, Zhong S, Lin L. Chicken sound recognition using anti-noise mel frequency cepstral coefficients[C]//2015 Third International Conference on Robot, Vision and Signal Processing (RVSP). IEEE,2025.

[7] Whitaker B M, Carroll B T, Daley W, et al. Sparse decomposition of audio spectrograms for automated disease detection in chickens[C]//2014 IEEE Global Conference on Signal and Information Processing (GlobalSIP). IEEE, 2014: 1122-1126.

[8] Mao A, Giraudet C S E, Liu K, et al. Automated identification of chicken distress vocalizations using deep learning models[J]. Journal of the Royal Society Interface, 2022, 19 (191): 20210921.

[9] Imoto K, Mishima S, Arai Y, et al. Impact of data imbalance caused by inactive frames and difference in sound duration on sound event detection performance[J]. Applied Acoustics, 2022, 196: 108882.

[10] Hassan E, Elbedwehy S, Shams M Y, et al. Optimizing poultry audio signal classification with deep learning and burn layer fusion[J/OL]. Journal of Big Data, 2024. <https://doi.org/10.1186/s40537-024-00985-8>.

[11] Srinivasagan R, El S M S, Al-Rasheed M I, et al. Edge intelligence for poultry welfare: utilizing tiny machine learning neural network processors for vocalization analysis[J/OL]. PLOS ONE, 2025, 20(1): e0316920. <https://doi.org/10.1371/journal.pone.0316920>.

[12] Carvalho S P, Grzywalski T, Hou Y, et al. Automated detection of broiler vocalizations a machine learning approach for broiler chicken vocalization monitoring[J]. Poultry Science, 2025, 104(5): 104962.

[13] 袁超,沈明霞,姚文,等. 基于发声特征和深度学习的白羽肉鸡全生命周期咳嗽检测方法[J]. 南京农业大学学报,2023,46(5):975-985.
Yuan Chao, Shen Mingxia, Yao Wen, et al. Cough detection method for white-feathered broilers throughout their entire life cycle based on vocalization features and deep learning[J]. Journal of Nanjing Agricultural University, 2023, 46(5): 975-985. (in Chinese)

[14] Szegedy C, Vanhoucke V, Ioffe S. Rethinking the inception architecture for computer vision[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2016: 2818-2826.

[12] 朱仁军,李骁骛. 鸡舍空气中有害物质的危害与控制[J]. 养禽与禽病防治, 2011(5):29 – 31.

[13] Han H, Zhou Y, Liu Q, et al. Effects of ammonia on gut microbiota and growth performance of broiler chickens[J]. Animals, 2021,11(6):1716.

[14] NY/T 1566—2007 标准化肉鸡养殖场建设规范[S].

[15] 孙俏,张彤,王新阳,等. 基于渐进 DS 证据理论的空气质量评价方法[J]. 北京林业大学学报, 2022,44(3):119 – 128.

Sun Qiao, Zhang Tong, Wang Xinyang, et al. Air quality evaluation method based on progressive DS evidence theory[J]. Journal of Beijing Forestry University, 2022, 44(3): 119 – 128. (in Chinese)

[16] 安维峥,孙钦,马婉蓉,等. 基于改进 D – S 证据理论的水下电子模块可靠性评估方法[J]. 化工自动化及仪表,2026, 53(1):54 – 60.

An Weizheng, Sun Qin, Ma Wanrong, et al. A reliability assessment method for the subsea electronic modules based on improved D – S evidence theory[J]. Chemical Automation and Instrumentation, 2026, 53(1): 54 – 60. (in Chinese)

[17] Suraiya A, Bin C, Derek W, et al. Impacts of air velocity treatments under summer condition; part I—heavy broiler’s surface temperature response[J]. Animals, 2022,12(3):328.

[18] Sadaf S, Ahmad H, Mahdih Z. Effects of cold stress on growth performance, carcass traits and tibia attributes in broiler chickens with thiram-induced dyschondroplasia[J]. Journal of Animal Physiology and Animal Nutrition, 2022,107(2):659 – 670.

[19] Purswell J J, Davis J D, Luck B D, et al. Effects of elevated carbon dioxide concentrations on broiler chicken performance from 28 to 49 days[J]. International Journal of Poultry Science, 2011,10(8):597 – 602.

[20] Jawad K O, Jawad R M, Ghazi Y A A. Design and implementation of smart relay based remote monitoring and controlling of ammonia in poultry houses[J]. International Journal of Computer Applications, 2014,103(8):13 – 18.

[21] 廖旺涛,张向利. 基于优先级和遗传算法的工业任务调度算法[J]. 桂林电子科技大学学报,2026,46(2):150 – 156.

Liao Wangtao, Zhang Xiangli. Industrial task scheduling algorithm based on genetic algorithm and task priority[J]. Journal of Guilin University of Electronic Technology, 2026, 46(2): 150 – 156. (in Chinese)

(上接第 35 页)

[15] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2016: 770 – 778.

[16] Zhou H, Zhu Q, Norton T. Cough sound recognition in poultry using portable microphones for precision medication guidance [J]. Computers and Electronics in Agriculture, 2025,237:110541.

[17] 李文伟,郑永军,杨圣慧,等. 音频信息技术在禽畜养殖与果蔬种植中的应用研究进展[J]. 农业工程学报,2024, 40(7):34 – 49.

Li Wenwei,Zheng Yongjun,Yang Shenghui,et al. Research progress on the application of audio technology in livestock breeding and fruit and vegetable planting[J]. Transactions of the CSAE,2024,40(7):34 – 49. (in Chinese)

[18] Manikandan V, Neethirajan S. AI-powered vocalization analysis in poultry: systematic review of health, behavior, and welfare monitoring[J]. Sensors, 2025, 25(13): 40 – 58.

[19] Manikandan V, Neethirajan S. Decoding poultry welfare from sound: a machine learning framework for non-invasive acoustic monitoring[J]. Sensors,2025, 25(9):2912.

[20] Qiao H, Chen Z, Ji C, et al. SmartEars: a practical framework for poultry respiratory monitoring via spectrogram-based audio classification and AI-assisted labeling[J]. Computers and Electronics in Agriculture, 2026, 241:111241.

[21] Ginovart G J, Iriondo I, Panisello M T, et al. Acoustic detection of vaccine reactions in hens for assessing anti-inflammatory product efficacy[J]. Appl. Sci, 2024, 14:2156.

[22] Ginovart G J, Iriondo I, Panisello M T, et al. Trend and representativeness of acoustic features of broiler chicken vocalisations related to CO₂[J]. Appl. Sci. , 2022,12:10480.

[23] Ginovart G J, Alsina R M, Panisello M T. Acoustic description of bird broiler vocalisations in a real-life intensive farm and its impact on animal welfare: a comparative analysis of recordings[J]. Eng. Proc. , 2020,2:53.

[24] Ginovart G J, Lsina R M, Iriondo S I, et al. Acoustic description of the soundscape of a real-life intensive farm and its impact on animal welfare: a preliminary analysis of farm sounds and bird vocalisations[J]. Sensors, 2020, 20:4732.

[25] Ginovart G J, Iriondo I, Panisello M T, et al. Acoustic detection of the effects of prolonged fasting on newly hatched broiler chickens[J]. Computer Engineering and Applications, 2024, 219:108763.